THE DEVELOPMENT AND DYNAMICS OF THE INFANT SKIN MICROBIOME

BY

ILIANA ROSA SERGHIOU

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

University of East Anglia School of Biological Sciences

Quadram Institute Bioscience

Norwich Research Park

Unilever

September 2024

Word Count: 59,857

©This copy of thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with the current UK Copyright Law. In addition, any quotations or extract must include full attribution.

ACKNOWLEDGEMENTS

As my PhD journey comes to an end, I have had the delight of reflecting on how far I have developed as a researcher over the last four years, and those that have invaluably guided me towards my PhD completion and attainment. With that, I would like to extend my sincere thanks to my three supervisors, Professor Mark Webber, Professor Lindsay Hall and Dr. Barry Murphy, for their assistance and insightful mentoring. I would specifically like to thank Mark for his unwavering support, kindness and patience, Lindsay for her wisdom, contagious enthusiasm and belief in this project, and Barry for his bright and cheery nature, and his refreshing industry perspective and passion for this research. It has been an honour to have been supervised by you all, and you have imparted invaluable research values, that I will forever be grateful for. It was a joy and privilege to work alongside you.

I further wish to thank Dr. Raymond Kiu and Dr. Matthew Dalby for their indispensable bioinformatics teachings that developed a newfound passion and career path for me. I would also like to thank the Quadram Institute's PEARL, sequencing and core Bioinformatics teams (specifically Sarah Phillips, Dave Baker and Dr. Alise Ponsero), and Dr. Charlotte Neumann (from the Med Uni Graz in Austria), for their services, and funding bodies (Unilever, BBSRC and UKRI) that made this project possible. I also wish to thank Dr. Leanne Sims for her contribution and time in proofreading my thesis, and other members of the Webber and Hall research groups, specifically Maria, Agata, Heather and Emma, for their amazing friendship and heartfelt support, especially during the final year of my PhD journey.

A special thank you to my darling fiancé Dr. David for being my biggest supporter and fan and always being proud of my endeavours. I lovingly thank my amazing father Serge, my sisters Alexia and Crystalla, my Aunty Georgia and Yiayia for their immense support and love during this last year, and ongoing wholeheartedness. You are my pillar of strength and guiding light.

Finally, I would like to proudly thank my wonderful, beautiful mother Patrizia for always believing in me and providing everlasting support and unconditional love, especially during this last year with your heroic endless battle with ailment. You are an inspiration.



ABSTRACT

The skin microbiome is organised into microbial communities across the body, with early life a critical point in ecosystem development. Many factors cause temporal community shifts, sometimes causing perturbations and skin disease; highlighting the key role of skin microbes in health. Sequencing methods have allowed us to explore the skin microbiome however many previous studies have used 16S rRNA gene amplicon sequencing, which has its limitations. Currently, there are open questions on how the whole perinatal skin microbiome interacts and functions, before and after birth.

In this study, I used shotgun metagenomics and a controlled healthy pregnancy and infant cohort (Pregnancy and Early Life (PEARL)) to analyse skin microbiome composition, temporal dynamics (bacteriome focused) and microbiome influencing factors. I compared multi body sites (skin, gut and vaginal) to investigate skin microbiome establishment and microbial sharing over time between mothers and their infants.

To achieve these aims, I first optimised a skin specific bacterial DNA extraction method; addressing the limitations of existing commercial kit protocols for low biomass skin samples. The optimised method was validated on a healthy adult and PEARL study subset and sequenced using Shotgun Metagenomics Sequencing (SMS). The resulting data allowed genus, species and strain level profiling, and generation of Metagenome Assembled Genomes (MAGs) to explore the diversity of genomes to aid future determination of functional capacities.

An optimised bacterial DNA recovery method from skin swabs, yielding high quality DNA suitable for shotgun metagenomics, was presented. Its application to the PEARL subset established a SMS generated baseline composition for a healthy infant skin microbiome, highlighting compositional and diversity changes over time. The infant archaeome was also indicated. Sharing of core genera, and their abundance changes over time, was identified between maternal and infant skin, stool and vaginal sites. Strain sharing of these core genera was also presented between mothers and infants. The maternal skin was the main source of microbial transfer to infant skin. A PEARL MAGs catalogue was generated. This research may help develop novel evidence-based probiotics to treat and prevent the increasing occurrence of early life skin disease.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate ‘take down’ action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

TABLE OF CONTENTS

LIST OF PUBLICATIONS AND PRESENTATIONS I

LIST OF TABLES III

LIST OF FIGURES V

LIST OF APPENDICES VIII

LIST OF ABBREVIATIONS IX

CHAPTER 1 1

INTRODUCTION 1

CHAPTER 1 OVERVIEW 2

1.1 Human Skin Microbiome and Health 3

1.1.1 Structure and Function of the Human Skin Microbiome 3

1.1.2 The Human Skin Microbiome and the Immune System 6

1.1.3 Skin and Gut Microbiome Interactions and Links to Immune Responses 7

1.1.4 Early Life Skin Microbiome Perturbation Related Skin Diseases 9

1.2 Factors Influencing the Infant Skin Microbiome 11

1.2.1 Delivery Mode 13

1.2.3 Labour 14

1.2.4 Age 2

1.2.5 Antibiotic Exposure 15

1.2.6 Vaginal Seeding 16

1.2.7 Probiotics and Live Biotherapeutics 17

1.2.8 Gestational Age 18

1.2.9 The Neonatal Skin Structure and Function 18

1.2.10 Environment 19

1.2.11 Skincare 20

1.2.12 Breastfeeding 21

1.2.13 Diet 22

1.2.14 Maternal Obesity 22

1.2.15 Genetics 23

1.2.16 Gender 23

1.3 Current Methods Available to Investigate the Skin Microbiome 24
1.4 Technical Challenges and Unmet Research Questions 28
1.5 Hypothesis 31
1.6 Aims 31
1.7 Objectives 31
CHAPTER 2 33
METHODOLOGY 33
CHAPTER 2 OVERVIEW 34
2.1 Adult Volunteer and PEARL Study Cohorts 35
2.1.1 Ethical Statement and Regulatory Considerations for Data Protection and Donor Confidentiality 35
2.1.2 Study Design 36
2.1.3 Sample Collection and Handling 38
2.2 Sample Processing, Culturing and DNA Extraction Method for Shotgun Metagenomic Sequencing 44
2.3 Qubit Assay 44
2.4 Chapter 3 45
2.4.1 Bacterial Strains 45
2.4.2 Overnight Culture Methods 45
2.4.3 Tapestation Assay 45
2.4.4 Illumina Library Preparation for Shotgun Metagenomic Sequencing 45
2.4.5 Oxford Nanopore Library Preparation for Shotgun Metagenomic Sequencing 47
2.4.6 Data Clean Up 48
2.4.7 Taxonomic Profiling 48
2.4.8 Metagenome Assembled Genomes Assembly 48
2.4.9 Data Visualisation and Statistical Testing 49
2.5 Chapter 4 49
2.5.1 Illumina Library Preparation for Shotgun Metagenomic Sequencing 49
2.5.2 Data Clean Up 51

2.5.3 Taxonomic Profiling 51
2.5.4 Metagenome Assembled Genomes Assembly 54
2.5.5 Data Visualisation and Statistical Testing 55
CHAPTER 3 56
METHOD DEVELOPMENT FOR HIGH MOLECULAR WEIGHT BACTERIAL DNA EXTRACTION SUITABLE FOR SHOTGUN METAGENOMICS FROM SKIN SWABS 56
CHAPTER 3 OVERVIEW 57
3.1 Introduction 58
3.2 Aims 58
3.3 Objectives 59
3.4 Experimental Methods 59
3.4.1 Initial Diluent Testing Experiments 61
3.4.2 Lysis Method Experiments 63
3.4.3 DNA Extraction Validation with Adult Volunteer and PEARL Study Skin Swabs 67
3.5 Data Analysis 69
3.5.1 Initial Diluent Testing 69
3.5.2 Lysis Method Experiments 69
3.5.3 DNA Extraction Validation with Adult Volunteer and PEARL Study Skin Swabs 69
3.6 Results 70
3.6.1 Initial Diluent Testing Experiments 70
3.6.2 Lysis Method Experiments 77
3.6.3 DNA Extraction Validation with Adult Volunteer and PEARL Study Skin Swabs 81
3.7 Discussion 98
3.7.1 Initial Diluent Testing Experiments 98
3.7.2 Lysis Method Experiments 99
3.7.3 DNA Extraction Validation with Adult Volunteer and PEARL Study Skin Swabs 99
3.8 Conclusion 102
CHAPTER 4 103

THE COMPOSITION OF THE DEVELOPING INFANT SKIN MICROBIOME 103
CHAPTER 4 OVERVIEW 104
4.1 Introduction 105
4.2 Aims 105
4.3 Objectives 106
4.4 Experimental Methods 106
4.5 Data Analysis 109
4.5.1 Genus and Species Level Analysis 109
4.5.2 Strain Level Analysis 111
4.5.3 MAGs Analysis 112
4.6 Results 112
4.6.1 The Composition and Baseline of a Healthy Infant Skin Microbiome Over Time 120
4.6.2 Maternal and Infant Skin Archaeal Composition Over Time 137
4.6.3 Maternal and Infant Skin, Stool and Vaginal Composition Diversity Changes Over Time 148
4.6.4 Shared Microbiota Over Time between Maternal and Infant Skin, Stool and Vaginal Body Sites 160
4.6.5 Strain Level Sharing of the Common Microbiota between Maternal and Infant Skin, Stool and Vaginal Body Sites Over Time 174
4.6.6 Generation of a PEARL Study MAGS Genome Catalogue 187
4.7 Discussion 196
4.8 Conclusion 202
CHAPTER 5 204
CONCLUSIONS AND FUTURE WORK 204
CHAPTER 5 OVERVIEW 205
REFERENCES 211
APPENDICES 228
PUBLICATIONS 346

LIST OF PUBLICATIONS AND PRESENTATIONS

PUBLICATIONS

Serghiou IR, Webber MA, Hall LJ. An update on the current understanding of the infant skin microbiome and research challenges. Current Opinion Microbiology. 2023 Oct; 75:102364. Epub 2023 Aug 14. PMID: 37586254 346

Serghiou IR, Baker D, Evans R, Dalby MJ, Kiu R, Trampari E, Phillips S, Watt R, Atkinson T, Murphy B, Hall LJ, and Webber MA. An efficient method for high molecular weight bacterial DNA extraction suitable for shotgun metagenomics from skin swabs. Microbial Genomics. 2023 July; 9(7) 353

Dalby MJ, Kiu R, Serghiou IR, Miyazaki A, Acford-Palmer H, Tung R, Cain S, Phillips S, Kujawska M, Matsui M, Iwamoto A, Taking B, Cox SE, Hall LJ. Gut microbiota and faecal cytokine profiles of rural Cambodian infants and links to early life nutrition and diarrhoeal episodes. npj Biofilms Microbiomes 10, 85 (2024) 367

Jain V, Dalby MJ, Alexander EC, Burford C, Gerasimidis K, Acford-Palmer H, Serghiou IR, Teng, MYN, Kiu R, Zafeiropoulou K, Logan M, Anita Verma A, Davenport M, Hall LJ, Dhawan A. Association of gut microbiota and gut metabolites and adverse outcomes in Biliary Atresia- a longitudinal prospective cohort-controlled study. Hepatology Communications; 8(11): e0550. 383

PRESENTATIONS

International

Invited speaker at the Microbirth International Conference – Oral presentation title: An update on the current understanding of the infant skin microbiome

Invited speaker at the Microbiome Virtual International Forum – Oral presentation title: An efficient method for high molecular weight bacterial DNA extraction suitable for shotgun metagenomics from skin swabs

Invited panel speaker at the Society for Applied Microbiology Early Career Symposium

Poster presentation Microbiology Society Annual Conference – Poster title: A recipe for success: Shotgun metagenomics for the skin microbiome

National

Invited speaker at the Norfolk and Norwich University Hospital (NNUH) Microbiology Department – Oral presentation title: The development and evolution of the infant skin microbiome

Poster presentation at the Microbes in Norwich Conference – Poster title: A recipe for success: Shotgun metagenomics for the skin microbiome

LIST OF TABLES

CHAPTER 2

Table 2.1. Overview of the PEARL Study Design. Extracted from Phillips et al. [329]. 37
Table 2.2. PEARL Infant Participant Metadata 42
Table 2.3. PEARL Participant Sample Collection Timepoints 43

CHAPTER 3

Table 3.1. Pilot Study DNA Yield 61
Table 3.2. Comparison of Extraction Methods 64
Table 3.3. PEARL Study Infant Metadata 68
Table 3.4. DNA Yield from Spiked Swabs 71
Table 3.5. CFU/ml for Each Culture Condition Suspension Replicate 73
Table 3.6. Bacterial DNA Yield (ng/µl) from the Six DNA Extraction Methods 77
Table 3.7. Average Molecular Weight (bp) of DNA Extracted 79
Table 3.8. DNA Yield (ng/µl) of Spiked Swabs Extracted with and without Agarase 79
Table 3.9. Skin Swab DNA Yield/Weight and Sequencing QC Summary 82
Table 3.10. Skin Swab Culture DNA Yield/Weight and Sequencing QC Summary 83
Table 3.11. MAGs Recovered from Ten Infant Skin Swabs at 1 Gb and 5 Gb Sequencing Depth. 98

CHAPTER 4

Table 4.1. Alpha Diversity Index ANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time 153
Table 4.2. Alpha Diversity Index ANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Pregnancy Term 153
Table 4.3. Alpha Diversity Index ANOVA P Value for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Delivery Mode 154
Table 4.4. Alpha Diversity Index ANOVA P Value for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Feeding Mode 154
Table 4.5. Beta Diversity Index PERMANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time 158
Table 4.6. Beta Diversity Index PERMANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Pregnancy Term ... 158
Table 4.7. Beta Diversity Index PERMANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Delivery Mode 159

Table 4.8. Beta Diversity Index PERMANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Feeding Mode 159
Table 4.9. Threshold and Clade Used for the Strain Detection and Identification Analysis 175
Table 4.10. List of Species of Strains Identified in One Maternal Culture Sample and At Least One Infant Sample 186

LIST OF FIGURES

CHAPTER 1

Figure 1.1. Diagram of Skin Layers and Microbiome 3
Figure 1.2. An Overview of the Complex Interlinking Factors Influencing the Infant Skin Microbiome 12
Figure 1.3. Current Amplicon and Shotgun Metagenomic Sequencing Methods . 25

CHAPTER 3

Figure 3.1. RSC Blood DNA Kit Protocol (Yellow Box) and Modifications to Test Initial Diluents and Lysis Procedures 60
Figure 3.2. Workflow Diagram for Measuring Cell Viability 63
Figure 3.3. Total DNA Yield (ng/µl) of Spiked Swabs Processed in 1xPBS and UMQW. 72
Figure 3.4. Bacterial Growth for *E. coli*, *S. aureus* and *P. aeruginosa* that were Processed and Grown in 1x PBS and UMQW 75
Figure 3.5. Example Photos of *E. coli*, *S. aureus* and *P. aeruginosa* Colony Growth 76
Figure 3.6. Total DNA Yield (ng/µl) Observed for each DNA Extraction Method for *E. coli*, *S. aureus* and *P. aeruginosa* 78
Figure 3.7. DNA Yield (ng/µl) from Spiked Swabs (with *E. coli*, *S. aureus* and *P. aeruginosa*) Extracted with and without Agarase 80
Figure 3.8. Skin Swab Aerobic and Anaerobic Cultures Obtained from PEARL Infants. 84
Figure 3.9. Skin Swab Aerobic and Anaerobic Cultures Obtained from Adult Volunteers 85
Figure 3.10. Comparison of DNA Yield (ng/µl) from Swabs Incubated for Different Periods 86
Figure 3.11. Taxonomic Profiles of Skin Swabs from Adult Volunteers Generated using Illumina and ONT Data 88
Figure 3.12. Taxonomic Profiles of Skin Swab Cultures from Adult Volunteers Generated using Illumina and ONT Data 89
Figure 3.13. Taxonomic Profile Reads of Skin Swabs from Adult Volunteers Generated using Illumina and ONT Data 90
Figure 3.14. Taxonomic Profile Reads of Skin Swab Cultures from Adult Volunteers Generated using Illumina and ONT Data 91
Figure 3.15. Taxonomic Profiles of the Skin Swabs and Cultures from PEARL Study Infants Generated using Illumina and ONT Data 92
Figure 3.16. Taxonomic Profile Reads of the Skin Swabs and Cultures from PEARL Study Infants Generated using Illumina and ONT Data 93

Figure 3.17. Alluvial Plot of the Percentage Total Abundant Species Counts between Illumina and ONT Shotgun Metagenomics data from the PEARL Study Infant Skin Swab and Culture Datasets. 95
Figure 3.18. GC-AT Content (%) of the Illumina and Nanopore Swab Datasets from the PEARL Study Infants 96
Figure 3.19. Species and MAGs Recovery from PEARL Infant Skin Swabs. Composition data was generated from PEARL study infant skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform 97

CHAPTER 4

Figure 4.1. Summary of the PEARL Study Total Sample Numbers Collected and Sequenced 108
Figure 4.2. Rarefaction Curves of the Maternal and Infant Datasets at Genus Level 114
Figure 4.3. Rarefaction Curves of the Maternal and Infant Datasets at Species Level 116
Figure 4.4. Taxonomic Relative Abundance Illumina Profiles of Controls to Genus and Species) Level 118
Figure 4.5. Taxonomic Read Illumina Profiles of Controls to Genus and Species Level 119
Figure 4.6. Common Relative Abundant Skin Genera Across all Infants Over Time 125
Figure 4.7. Common Relative Abundant Skin Species Across all Infants Over Time 127
Figure 4.8. Significant Genus and Species Level Relative Abundance Changes Over Time on Infant Skin. 129
Figure 4.9. Individual Infant and Maternal Skin Relative Abundant Genus Level Taxonomic Profiles Over Time 131
Figure 4.10. Individual Infant and Maternal Skin Relative Abundant Species Level Taxonomic Profiles Over Time 133
Figure 4.11. Top 50 Differential Abundant Genera and Species Significant Associations Over Time on Infant Skin 136
Figure 4.12. Archaeal Species Relative Abundance Profiles in Maternal and Infant Skin Swab Datasets 139
Figure 4.13. Archaeal Species Relative Abundance Changes Over Time in Maternal and Infant Skin Swab Datasets 141
Figure 4.14. Archaeal Species Relative Abundance Changes Over Time in Maternal and Infant Skin Swab Datasets Accounting for Pregnancy Term. 143
Figure 4.15. Archaeal Species Relative Abundance Changes Over Time in Maternal and Infant Skin Swab Datasets Accounting for Delivery Mode 145

Figure 4.16. Archaeal Species Relative Abundance Changes Over Time in Maternal and Infant Skin Swab Datasets Accounting for Feeding Mode 147
Figure 4.17. Mother and Infant Genus Level Richness Alpha Diversity Changes Over Time at Multiple Body Sites 149
Figure 4.18. Mother and Infant Species Level Richness Alpha Diversity Changes Over Time at Multiple Body Sites 150
Figure 4.19. Mother and Infant Genus Level Shannon Alpha Diversity Changes Over Time at Multiple Body Sites 151
Figure 4.20. Mother and Infant Species Level Shannon Alpha Diversity Changes Over Time at Multiple Body Sites 152
Figure 4.21. Mother and Infant Genus Level Beta Diversity Changes Over Time at Multiple Body Sites.. 156
Figure 4.22. Mother and Infant Species Level Beta Diversity Changes Over Time at Multiple Body Sites 157
Figure 4.23. Common Shared Relative Abundant Genera Over Time Across all Maternal and Infant Datasets 161
Figure 4.24. Common Shared Relative Abundant Species Over Time Across all Maternal and Infant Datasets 162
Figure 4.25. Individual Variation of the Relative Abundance Changes Over Time of the Common Shared Genera in Maternal and Infant Datasets 165
Figure 4.26. Individual Variation of the Relative Abundance Changes Over Time of the Common Shared Species in Maternal and Infant Datasets..... 168
Figure 4.27. Top 50 Differential Abundant Genera and Species Significant Associations Over Time Within and Across Mother and Infant Body Sites 173
Figure 4.28. Sample Availability Overview 174
Figure 4.29. Strains Shared between Mother and Infant Pairs. 176
Figure 4.30. Strain Transmissibility of the Shared Common Microbiota Between Maternal and Infant Pairs 185
Figure 4.31. Genus Level Tree of Life for Maternal and Infant MAGs Generated 190
Figure 4.32. Pruned Tree of Life for Maternal and Infant MAGs Generated for the Core Genera 192
Figure 4.33. Treemap of the Maternal and Infant Species MAGs Counts 195

LIST OF APPENDICES

Appendix 1. FMH Approval Letter..... 223

Appendix 2. FMH Application 224

Appendix 3. Human Tissue study Protocol 232

Appendix 4. HRGC Human Tissue Registration Form 245

Appendix 5. PIS and CF 249

Appendix 6. Stool Positive Controls 255

Appendix 7. Processing, Culturing and Extracting Bacterial DNA from PEARL Study Skin Swab Samples 256

Appendix 8. 241617-SC-024 MOTION study Stool Sample DNA Extraction Using the Promega Maxwell V3 273

Appendix 9. QIB SOP for General Bacterial Growth 282

Appendix 10. QIB SOP Preparation and Storage of Bacterial Master and Working Stocks 288

Appendix 11. PEARL Sample Sequencing MultiQC Summary 301

Appendix 12. Promega Maxwell RSC 48 Blood DNA Kit Protocol 306

Appendix 13. FastPrep MPBio Kit QIB Protocol ……..314

Appendix 14. Maternal and Infant Genus Level Taxonomic Profiles Over Time 322

Appendix 15. Maternal and Infant Species Level Taxonomic Profiles Over Time 324

Appendix 16. Maternal and Infant Sample Details for the Strain Detection and Identification Analysis ……326

LIST OF ABBREVIATIONS

µl	microliter
µM	micromolar
16s rRNA	16s ribosomal Ribonucleic Acid
AD	Atopic Dermatitis
AT	Adenine-Thymine
ATCC	American Type Culture Collection
BLT	Bead Linked Transposomes
bp	base pairs
BR	Broad Range
BWA	Burrows-Wheeler Aligner
CBA	Colombia Blood Agar
CF	Consent Forms
CFU	Colony Forming Unit
CONCOCT	Clustering cONtigs with COverage and ComposiTion
CoNS	Coagulase-negative staphylococci
CRF	Clinical Research Facility
CS	Caesarean Section
DD	Diaper Dermatitis
DNA	Deoxyribonucleic Acid
dNTPs	Deoxynucleoside Triphosphates
dsDNA	double stranded Deoxyribonucleic Acid
ELSG	Early-Life Skin Genome
ETN	Erythema Toxicum Neonatorum
Fastp	Fast Pre-processor
FastQ	Fast Quality
FastQC	Fast Quality Control
FFPE	Formalin-Fixed Paraffin-Embedded
FMH	Faculty of Medicine and Health Sciences
FMT	Faecal Microbiota Transplantation
Gb	Gigabase
GBS	group-B *Streptococcus*
GC	Guanine-Cytosine
GDM	Gestational Diabetes Mellitus
gDNA	Genomic DNA
GDPR	General Data Protection Regulation
GP	General Practice
GTDB-Tk	Genome Database Taxonomy Toolkit
HELLP	Haemolysis Elevated Liver Enzymes and Low Platelets
HG	Human Genome
HIC	High Income Countries
HIV	Human Immunodeficiency Virus
HMW	High Molecular Weight
HPC	Performance Computing

HRA	Health Research Authority
HRGC	Human Research Governance Committee
HS	High Sensitivity
HSV	Herpes Simplex Virus
ICH GCP	International Conference on Harmonisation Good Clinical Practice
IRIDA	Integrated Rapid Infectious Disease Analysis
ITS	Internal Transcribed Spacer
KMC	kangaroo mother care
LB	Luria Bertani
LBP	Live Biotherapeutics
LIC	Low Income Countries
LMW	Low Molecular Weight
MaAsLin2	Microbiome Multivariable Association with Linear Models 2
MAGs	Metagenome-Assembled Genomes
Mbp	Megabase pairs
MetaBAT	Metagenome Binning with Abundance and Tetra-nucleotide frequencies
Metaphlan4	Metagenomic Phylogenetic Analysis
MetaWRAP	Metagenomic Wrapper
ml	millilitre
MultiQC	Multi-Quality Control
N	Nucleotide
NBI	Norwich Biosciences Institute
ng	nanogram
NHS	National Health Service
NICU	Neonatal Intensive Care Unit
NNUH	Norfolk and Norwich University Hospital
ONT	Oxford Nanopore Technologies
PacBio SMRT	Pacific Biosciences' Single-Molecule Real-Time
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PEARL	Pregnancy and Early Life
PF	Passing Filter
PIS	Participant Information Sheets
pM	picomolar
QC	Quality Control
QIB	Quadram Institute of Biosciences
REC	Research Ethics Committee
RefSeq	Reference Sequence
RNA	Ribonucleic Acid
SBCC	Skin Bacteria Culture Collection
SD	Seborrheic Dermatitis
SEM	Standard Error Mean
seqfu	Sequence Fastx Utilities
SMGC	Skin Microbial Genome Collection
SMS	Shotgun Metagenomic Sequencing
SOP	Standard Operating Procedure

SPRI	Solid Phase Reversible Immobilization
SPRI	Solid Phase Reversible Immobilisation
SSSS	Staphylococcal-Scalded Skin Syndrome
TB1	Tagmentation Buffer 1
TC	Tinea Capitis
UEA	University of East Anglia
UMQW	Ultrapure Milli-Q Water
VD	Vaginal Delivery
VS	Vaginal Seeding
WHO	World Health Organisation

CHAPTER 1

INTRODUCTION

CHAPTER 1 OVERVIEW

This chapter explores what is known about the skin microbiome, particularly infant skin. The structure and function of the skin microbiome is investigated, as is its impact on the immune system and health, specifically skin health. Perturbations to the skin microbiome and resulting effects on the immune system and skin health are discussed; as are early life and longer-term factors that contribute to these perturbations. The current methods available to investigate the skin microbiome are presented, leading to a discussion about the technical challenges and unmet research questions within the skin microbiome field. Finally, the project aims and objectives are outlined.

Chapter 1 forms the basis of a first author review publication: Serghiou, I.R, Webber, M.A., & Hall, L.J. (2023). An update on the current understanding of the infant skin microbiome and research challenges. Current opinion in microbiology, 75:102364.

1.2.4 Age

https://doi.org/10.1016/j.mib.2023.102364

1.1 Human Skin Microbiome and Health

1.1.1 Structure and Function of the Human Skin Microbiome

The skin is the human's largest organ and main interface with the external world [1, 2]. As reviewed in Smythe et al. [3], the skin is made up of multiple layers including the epidermis, dermis and subcutaneous adipose tissue, with hair follicles and glands embedded within these layers; together these are involved in preserving the cutaneous (skin) barrier and homeostatic functions, such as thermoregulation and wound repair [3] (**Figure 1.1**). This first line organ also retains a microbiome described as a complex ecosystem organised into communities at different body sites [4-9].

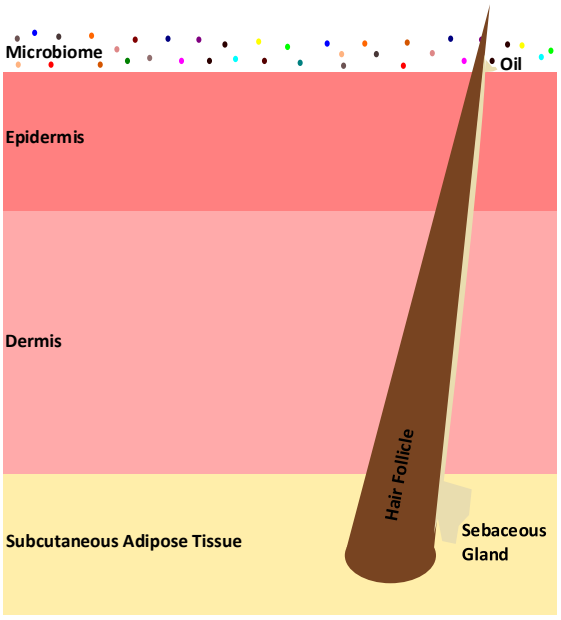


Figure 1.1. Diagram of Skin Layers and Microbiome

The microbial communities start to become site specific with age [4-7]. These communities include bacteria (bacteriome), archaea (archaeome), viruses (virome) and fungi (mycobiome), which collectively have been shown to contain certain metagenomic elements that help them adapt and participate in the host's physiological functions and immunity, and overall skin health [4, 9-12].

What a healthy microbiome should look like is a frequently asked question. A lack of knowledge over the last 14 years, high inter-individual variation in both microbiome composition and factors affecting it prevents a clear answer [13]. What is known is that the adult skin bacteriome is typically colonised with phyla Pseudomonadota (formerly Proteobacteria), Actinomycetota (formerly Actinobacteria), and Bacillota (formerly Firmicutes), whilst the infant skin bacteriome exhibits phyla Bacillota, Actinomycetota, Pseudomonadota, and Bacteroidota

(formerly Bacteroidetes) [1, 7, 14, 15]. Directly after birth the infant skin microbiome is dominated by coagulase-negative staphylococci species (CoNS) (*Staphylococcus epidermidis* and *Staphylococcus capitis*) and *Lactobacillus iners* [16, 17]. Shortly after, these microbes decrease in abundance, allowing gradual colonisation (or transient appearance) by *Acinetobacter*, *Streptococcus*, *Veillonella*, *Bifidobacterium* and *Enhydrobacter* species [14]. Preterm neonates appear to have a more varied skin microbiome and exhibit more *Staphylococcus*, *Klebsiella*, *Serratia*, *Streptococcus*, *Enterococcus* and *Enterobacter* [17]. *Streptococcus* species particularly appear to be more abundant in infancy and decrease with age [1, 18, 19]. As the best studied microbial community on skin, bacteriome collections have been shared, such as the Skin Microbial Genome Collection (SMGC) and Skin Bacteria Culture Collection (SBCC) [20]. These collections recently identified a novel skin-abundant bacterial genus *Candidatus* Pellibacterium in adults and infants.

More recently over the last two years, research has provided more insight into the infant skin fungal communities [21, 22]. The mycobiome typically contains *Malassezia*, *Candida*, *Cladosporium*, *Alternaria*, *Rhodotorula*, *Fusarium*, *Cryptococcus*, *Exophiala*, *Aspergillus* and *Nigrospora* [22-25]. *Candida albicans* has been shown more abundant in areas with skin-to-skin contact, whereas *Cladosporium* and *Alternaria* were more abundant in less-occluded areas [22]. Saheb Kashaf et al. [20] also recently identified three novel *Malassezia* species (*Malassezia auris*, *Malassezia palmae* and *Malassezia rara*) in adults and infants.

There is a lack of infant skin virome research, and current work is focused on disease outcomes in neonates, such as the herpes simplex virus 2 [21]. Many viruses identified in the skin microbiomes of adults and infants have been classified to the class *Caudoviricetes* [20]. The toe web appears to be a site where viruses are abundant, and novel jumbo phages (bacteriophages characterised by having a very large genome > 200 kb) [26] have been identified here [20]. Archaeome research and its role in the skin microbiome is lacking across the lifespan; Nitrososphaerota (previously known as Thaumarchaeota) is currently the main archaeal phylum identified on human skin, with no specific phyla relating to infant skin [12, 27, 28]. The genus *Nitrosocosmicus* have been found to form an integral component of the skin microbiome [29]. The abundance of archaea is particularly increased in children and the elderly, which may be attributed to physical and chemical changes of skin physiology [27, 30].

Overall, research on early-life skin microbial community assembly is currently limited, especially using Metagenome-Assembled Genomes (MAGs); however, an Early-Life Skin Genome (ELSG) catalogue was recently published for infants aged 2–3 months and 12 months [31]. This catalogue comprises 9194 bacterial genomes from 1029 species, 206 fungal genomes from 13 species and 39 viral sequences [31]. The top bacterial (*Streptococcus*, *Corynebacterium*, *Neisseria*, *Bifidobacterium*, *Prevotella*), fungal (*Malassezia*) and viral (*Torque teno virus* and *Gammapapillomavirus*) genera were identified [31].

These mixed microbial communities can be found at different skin body sites, which differ in temperature, moisture, pH, and oxygen availability; these provide distinct microenvironments that supply required nutrients to support specific microbial growth [32, 33], demonstrating a mutualistic interaction between the host and its microbiome [34, 35]. Microenvironments include the face and forehead, which are high in sebaceous glands resulting in an environment suitable for lipophilic microbes, such as *Cutibacterium* (formerly *Propionibacterium*) *and Corynebacterium species* (Actinomycetota) [15, 36, 37]; The gluteal region, axillary vault, and toe web are moist and warm environments allowing Gram-negative bacilli (Pseudomonadota) and Gram-positive *Staphylococcus aureus* (Bacillota) to thrive [2, 36]. The antecubital fossa is a dry desiccated area, which fluctuates in temperature, and harbours diverse microbes [36], such as Pseudomonadota and Bacteroidota (Flavobacteriales) [15].

These distinct microbial communities at different body sites have different functions [4-7]. Research into skin microbiome functionality is still limited, but it is known to contribute to epidermal development and differentiation [38]. More integrative approaches are now being used to study the infant skin microbiome and its function, using combinations of metabolomics, culturomics and metagenomics [20, 22, 39-43]. This has enabled development of the SMGC and SBCC collections from infant and adult skin [20], and identification of three distinct metabolite–microbe clusters in the skin. The first was dominated by *Cutibacterium* linked to hydrophobic elements of the skin barrier, the second associated *Staphylococcus* with amino acids relevant to water holding capacity and pH regulation of the skin surface, and the third was characterised by *Streptococcus* associated with amino acid and sugar abundance [42]. Archaea have been associated with ammonia oxidation and nitrification processes [28]. These processes are known to lower the environmental pH, resulting in a more acidic environment, which has been associated with a more

stable skin barrier [27, 44]. Integrative studies will have implications for treatment of disease and maintenance of skin health through identification of potential strains for probiotic interventions [45].

It is also important to understand the microbiome functions of other body sites, such as the gut and vagina, as these are implicated in health outcomes in combination with the skin [38, 46-50]. The gut microbiome plays a vital role in nutrient uptake and metabolism [50], where enteric microbes produce and modify many biologically active compounds, such as macromolecules, bile acids and vitamins, that support host metabolism [51, 52]. For example, *Akkermansia muciniphila* produce short chain fatty acids (acetate and propionate), which activate certain fatty acid receptors (GPR41 and GPR43) that induce peptide production (GLP-1) [53]. GLP-1 is a principal incretin hormone that promotes glucose-dependent insulinotropic activity, inhibits appetite and food intake, delays gastric emptying, and restores the impaired incretin effect in type 2 diabetes mellitus patients [53]. Another example is *Bifidobacteirum longum*, which produce adenosine derivatives (methylthioadenosine) by converting dietary amino acids [54]. Methylthioadenosine has been shown to ameliorate obesity, thus has an anti-obesogenic effect [54]. In the vagina, *Lactobacillus* species support a low vaginal pH [51, 52] through microbial production of lactic acid, which prevents the development of infectious disease, such as Bacterial Vaginosis [55]. These varying roles of the skin, gut and vaginal microbiomes demonstrate the importance of understanding the mechanisms which underpin host microbial symbiosis, and how microbiota may contribute to immunity and disease [56].

1.1.2 The Human Skin Microbiome and the Immune System

The human skin microbiome influences host immunity [57] and evolution of the infant skin microbiome has been linked to immune system development [18]. The skin microbiome helps prevent pathogen colonisation of the skin [18, 37] and may be important for immune homeostasis and long-term health [1, 57-61].

For example, the skin commensal genus *Staphylococcus* produces lipoteichoic acid, which suppresses skin inflammation during wound repair by acting on a keratinocyte receptor (Toll-like receptor 2), indicating a protective role of the skin microbiome against aberrant inflammatory response [62]. *Staphylococcus epidermidis* has also been shown to modulate the host innate immune response through production of phenol-soluble modulins or lipoteichoic acid, which can inhibit

skin pathogens and inflammation respectively [2]. Mouse models have further shown that the neonatal skin microbiota cause a high influx of regulatory T cells, promoting immune tolerance and preventing inflammation [61, 63]. When this influx was blocked before microbial colonisation, tolerance did not develop [63]. Regulatory T cells were also found to decrease if commensal microbiota colonisation and hair follicle morphogenesis were disrupted [64]. Hair follicle morphogenesis facilitates the accumulation of T cells within the hair follicles in neonatal skin [64]. These hair follicles are also home to commensal microbes that further increase T cell accumulation and promote chemokine production [64]. Therefore, hair follicle morphogenesis and commensal microbes work in concert to produce T cells, which establish and maintain immune homeostasis in the skin, thereby having critical roles in skin immunity development [64].

Certain commensal microbes, such as *Staphylococcus* and *Cutibacterium* species, can act as both beneficial microbes and opportunistic pathogens [2, 65-67]. Continuous exposure of the skin microbiome to external pressures may destabilise the microbiome equilibrium, causing a microbial imbalance and allowing opportunistic pathogens to colonise [65, 68]. A microbial imbalance has been linked with several dermal conditions, including acne, vitiligo, Erythema Toxicum Neonatorum (ETN), Seborrheic Dermatitis (SD), Atopic Dermatitis (AD), Staphylococcal-Scalded Skin Syndrome (SSSS), Diaper Dermatitis (DD) and Tinea Capitis (TC) [9, 14, 22, 24, 59, 69-76]. The mechanisms of perturbations are complex [77] but contributing factors may include resource competition [78], nutritional interdependency resulting in trophic cascades [79], and behavioural changes resulting from internal or external stimuli [80]. It is therefore important to understand the skin microbiome in diseased and healthy individuals to elucidate how the skin microbiome can affect human health through host interactions [4, 81-83].

1.1.3 Skin and Gut Microbiome Interactions and Links to Immune Responses

As well as the skin microbiome, gut microbiomes are known to influence host immunity, with gut microbial imbalances linked to a variety of diseases including obesity, diabetes, inflammatory bowel disease, cancer, necrotizing enterocolitis, neurodevelopment, and allergic diseases such as asthma and AD [31, 46-49, 84-87]. Gut microbial imbalances have elicited cytokine production in infants and mouse models, stimulating T-helper 2 cell inflammation, central to the pathogenesis of many allergic diseases [88, 89]. The skin microbiome has also been suggested

to influence allergic disease relating to skin health, but this is very understudied in comparison to the gut microbiome [46-49].

Understanding the dynamic infant skin microbiome and its two-way interaction with the gut microbiome could allow for interventional opportunities to prevent disease [1, 58]. This link has been supported by literature suggesting that probiotic supplementation in the perinatal period may help prevent paediatric AD [90]. In another example, exposure of the foetus *in utero* to meconium-stained amniotic fluid was suggested to protect the neonate from developing dermatitis and hospital-acquired skin infections [91]. This exposure may increase diversity of the gut microbiome, potentially stimulating the immune system and preventing skin inflammatory conditions [92]. This link between the gut and skin microbiota working in parallel to prevent AD, emphasises the importance of host-microbe interactions during the perinatal period [93]. This demonstrates a gut-skin axis relationship, where dietary catabolic products are utilised by gut microbes, which produce byproducts (i.e. short chain fatty acids) that influence gut physiology, resulting in circulating secretory products (i.e. hormones) around the body that enter the skin [94]. Therefore, ingested compounds and chemicals can impact the skin activity and appearance [94]. Additional research is required to understand the immune interaction between the gut and skin microbiome [95].

Further research could identify paediatric skin disease treatments, and early preventative strategies such as oral and topical probiotics, or specific microbial metabolites [93]. Currently, probiotic interventions, mostly using *Lactobacillus* and *Bifidobacterium* spp., and *Lacticaseibacillus rhamnosus GG*, have shown promising results in treating AD in infants [96-98]. After probiotic treatment with *L. rhamnosus GG*, higher levels of *Prevotella*, *Veillonella* and *Ralstonia* and lower levels of *Stenotrophomonas* and *Microbacterium* were identified, which may have skin health implications. Certain skin and gut microbes (*Acinetobacter* spp., *Haemophilus haemolyticus*, *Bacteroides ovatus*, *Schaalia odontolytica* and *Actinomyces graevenitzi*) have also been identified as potential candidates for probiotic treatment of AD and food allergies [99]. I have recently highlighted multiple clinical trials for skin probiotic interventions in early life which are currently being conducted, or were recently completed [45]; these studies assess the potential of probiotics for treatment of skin disease.

1.1.4 Early Life Skin Microbiome Perturbation Related Skin Diseases

Early life skin microbiome perturbation related skin disease include ETN, SD, AD, SSSS, DD and TC [9, 14, 22, 24, 59, 65, 69-76, 100, 101].

AD is a chronic skin condition common in early childhood and usually progresses to the development of other atopic diseases later in life such as asthma, allergic rhinitis, food allergies and hay fever (known as the atopic march) [100, 102]. Symptoms include dry, itchy skin, skin lesions, and a defective skin barrier, increasing vulnerability to superinfection [102]. Age (children), gender (females), geographic location (high income countries), genetics (filaggrin (FLG) gene) and parental atopic or allergic disease appear to be risk factors for incidence and prevalence of AD [103, 104]. AD-afflicted infants, whether in a lesional or non-lesional state, demonstrated a slower temporal development of the skin microbiome than healthy infants [74], and also have an increased risk of developing allergic rhinitis and asthma [105]. The development and pathogenesis of AD may be determined by barrier dysfunction, inflammation, microbiota, and antibiotics, although there have been conflicting results between studies [106-108]. Skin microbiome profiling of infant AD has shown altered microbiome community composition, with *S. aureus* identified as the main bacterial driver; however, this was not always the case [9, 73, 74]. Other research has suggested a combined effort between *S. aureus* and *S. epidermidis* [71, 108] and some have identified no specific bacterial taxa associated with AD [14]. Mouse models have further demonstrated that *S. aureus* drives AD development and *Corynebacterium bovis* contributed to the pathogenic immune response [109]. There is a clear lack of consensus on the main causative agent of AD, with some believing that early commensal staphylococcal exposure may be critical in prevention [101]; thus, further investigation is warranted.

ETN is another common neonatal skin condition which develops within the first week after birth [59, 76]. Symptoms present as papules and pustules on the body's trunk, thighs, and buttocks – all areas with hair follicles [59, 76]. This condition is rare in premature neonates and common in neonates that were vaginally delivered with greater gestational age (full term) and birthweight [76, 110]. Other risk factors for prevalence of ETN include ethnicity (Caucasian), gender (male), mothers with no gestational risk factors, seasonal (warmer seasons) [111]. Full maturation of the foetal immune system appears to be related to the development of ETN [111]. ETN could be due to an immune response to the establishment of skin commensal

microbiota, through the hair canal; affected areas, around hair follicles, were found to contain primarily *Staphylococcus* [59]. Hair follicles are important in immune system development, and surrounding areas contain pro-inflammatory markers such as Interleukin-1, Interleukin-8, E-selectin, and psoriasin [59, 112, 113]. This suggests that ETN represents an immune response to microbiota in hair follicles [59, 112, 113], though the adaptive benefits of this response are not yet understood.

DD, SD and TC are skin diseases that have been studied in relation to the bacteriome and mycobiome [22, 24, 75]. *Candida* and *S. aureus* have been directly correlated with DD disease severity [22, 75]. Peak incidence of DD appears to occur in infants aged 9-12 months, with risk factors including age, with a link to diet changes (affecting stool consistency and PH), and frequency of nappy changes (less frequent changes result in prolonged contact with urine and faeces, increasing risk of skin inflammation) [75, 114]. Symptoms of DD include erythema (sometimes more severe with maceration, erosions, nodules and punched out ulcerations), erythematous papules and itching [114].

SD is a common chronic skin condition in infants (known as “cradle cap”), being most prevalent at three months of age [69, 70, 115]. Risk factors include age and a family history of atopy [115]. Clinical symptoms include erythema, greasy, and flaky skin, typically affecting the scalp, but it can also affect the face, ears, neck, and diaper area [69, 70]. In adults it typically affects sebum-rich areas such as the scalp, ears, forehead, nasolabial folds, chest, and upper back [70, 116, 117]. Lipophilic yeast *Malassezia* has been a suggested cause [69, 70, 116-118]. *Malassezia* are found in healthy skin microbiota, but can become pathogenic [116, 117]. Both *Malassezia restricta* and *Malassezia globosa* are commonly found on healthy and diseased skin [117-119]. An increased abundance of *Malassezia* yeasts has been recorded on the scalp of patients with SD [116]. *Staphylococcus*, *Streptococcus*, and *Acinetobacter* were found to be more abundant at lesional sites, possibly supplying nutrients promoting *Malassezia* species growth [70]. It is as yet unknown why *Malassezia* species cause SD, as they are also present in healthy skin.

TC is caused by a pathogenic fungus *Microsporum canis*, which has been correlated with an increase in abundance of *Cutibacterium* and *Corynebacterium*, and reduced *Streptococcus* [24]. TC commonly presents with diffuse scaling, erythema and alopecia. TC predominantly occurs in infants and children aged 2-8 years, although is rare compared to the other skin diseases previously described [24, 120]. Age is a

clear risk factor, likely due to adults producing fungistatic long-chain fatty acids by their matured sebaceous glands, which leads to increased resistance of hair to dermatophyte colonisation [24]. Presence of household pets (cats and dogs) are another risk factor as they are reservoirs for TC, therefore a source of transmission to infants/children [120].

1.2 Factors Influencing the Infant Skin Microbiome

There are many complex interlinking factors that could influence microbiome community changes in early life and longer term, resulting in perturbation related skin disease. Recent studies have focused on how some of these factors influence skin microbiome establishment, function and disease [14, 16, 17, 22, 23, 39, 40, 121-128]. Microbiome influential factors have been categorised as intrinsic (genetics, gender and age), which are inherent, and extrinsic (delivery mode, labour, antibiotic exposure, Vaginal Seeding (VS), probiotics, Live Biotherapeutics (LBP), neonatal skin structure and function, environment, skincare, breastfeeding, diet and maternal obesity) [1, 11, 12, 129-131] (**Figure 1.2**). It should be recognised that some research has presented the idea of an *in utero* microbiome [41, 43, 56, 132-142]. However, the detected microbial signals from these studies likely resulted from contamination during the clinical procedures to obtain foetal samples or during DNA extraction and sequencing [43]. This highlights the high-risk nature of the *in utero* environment for contamination once exposed to the outside world [43]. Furthermore, this idea of a microbiome within healthy foetal tissues challenges the fundamental concepts of immunology, clinical microbiology and the obtainment of germ-free mammals [43]. Therefore, a transdisciplinary approach is warranted, incorporating biological, ecological and mechanistic concepts, in addition to contamination controls [43].

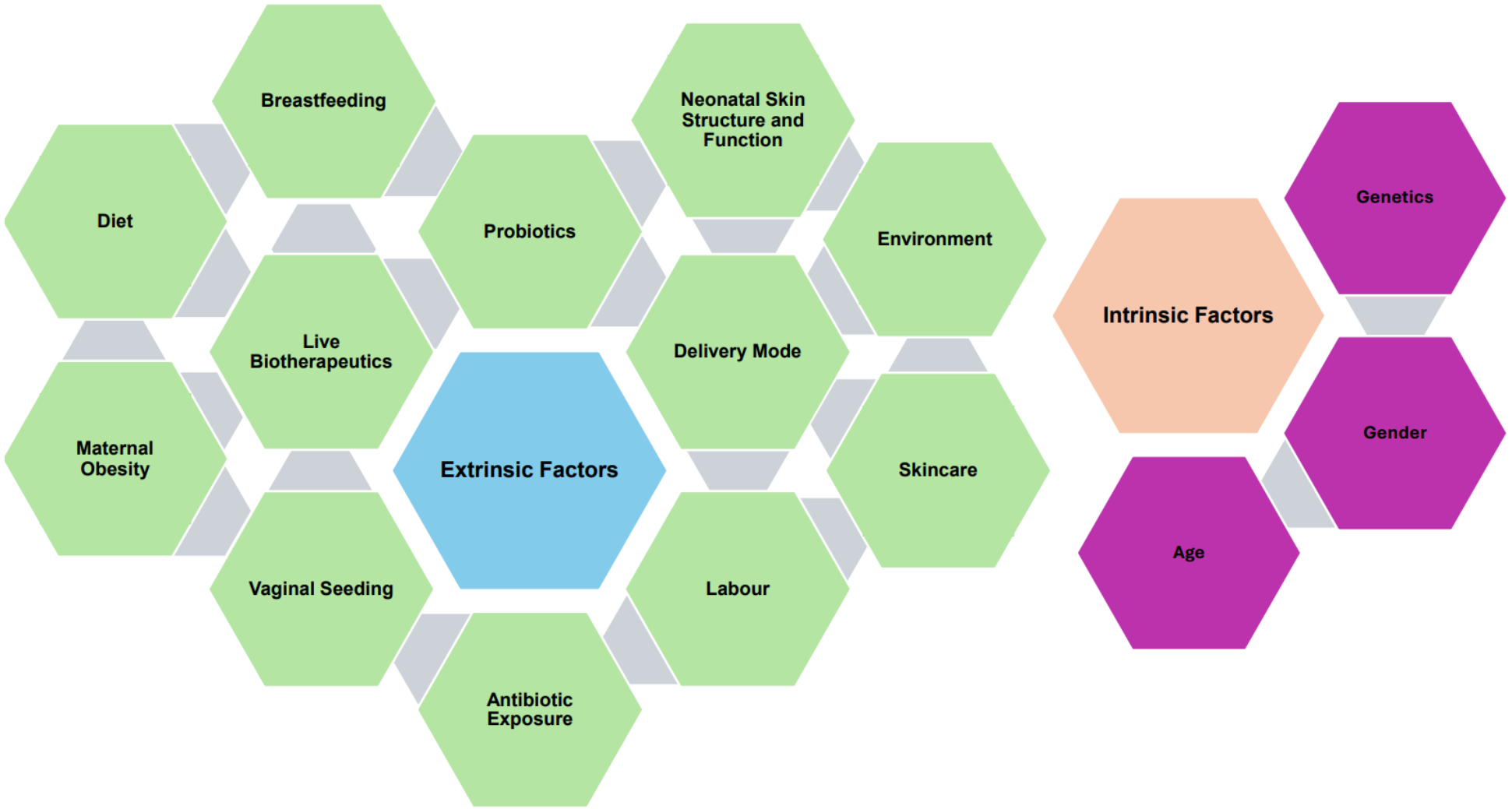


Figure 1.2. An Overview of the Complex Interlinking Factors Influencing the Infant Skin Microbiome

1.2.1 Delivery Mode

In utero, the foetal skin is exposed to amniotic fluid, and undergoes changes after delivery including reduced transepidermal water loss, skin pH and sebaceous activity changes, and increased water content [143]. During early life, the neonatal skin barrier, water handling functions, and the epidermal cell proliferation and keratinization, mature rapidly affecting skin microbial colonisation [143]. Delivery mode, i.e. Caesarean Section (CS) and Vaginal Delivery (VD), influences the skin microbial colonisation, in turn affecting skin barrier function [14, 22].

The International Federation of Gynaecology and Obstetrics have labelled CS use as an “epidemic” due to its global overuse [144]. There is a global concern for the health of women and their children exposed to CS deliveries [145]. CS delivery has been associated with the development of immune related disorders such as asthma, allergies [146-152], AD, type 1 diabetes mellitus, impaired cognitive abilities [153], juvenile arthritis, systemic connective tissue disorders, inflammatory bowel disease, immune deficiencies, leukaemia [147], and obesity [154-159]. This observed link between CS delivery and health may be due to the absence of vaginal microbiota transmission from mothers to neonates at birth and may contribute to the immune system development [56, 129, 132, 154, 159-161].

Research suggests the skin microbiota of vaginally delivered neonates resemble the maternal vaginal microbiota, particularly *Lactobacillus* [56, 129]. Whereas, the gut of VD neonates is dominated by vaginal and gut microbes (e.g. *Bacteroides*, *Bifidobacterium*, and *Lactobacillus*) [159]. These microbes are often absent, or have a delayed development in CS delivered neonates, possibly delaying the development of the neonatal immune system and limiting metabolism of the early life diet, i.e. human breast milk oligosaccharides [159]. Other body sites, such as the nares and oral cavity, in VD neonates, have been shown to harbour microbiota similar to the maternal skin and vaginal microbiota [56]. These studies collectively identify the internal and external transfer of microbes from mothers to neonates.

In contrast, the skin microbiota of neonates delivered by CS has been shown to resemble the maternal skin microbiota, dominated by *Staphylococcus*, *Streptococcus*, *Corynebacterium*, and *Cutibacterium* [56, 129]. Mycobiome-focused studies have also shown *Aspergillus* in higher abundance in CS infants, whilst *C. albicans* and *Rhodotorula* species were found in higher abundance on the skin of VD infants [22, 23]. More recent research has also included comparisons between

elective and emergency CS, as well as investigating modern birthing methods such as VD in water [14]; Rapin et al. identified a higher abundance of *Streptococcus* on the skin of infants born by elective CS and VD in water, compared with normal VD [14]. Dry skin was also observed in all infants born through elective CS, which was associated with higher *Streptococcus* abundance [14].

1.2.3 Labour

Labour onset involves the foetal descent into the vaginal canal and the rupture of the foetal membranes, exposing the foetus to the maternal vaginal microbiota [153, 162]. This concept has been supported through research revealing neonatal microbiota differences between elective CS (unlaboured), emergency CS (more likely laboured) and VD methods [56, 162]. In one study, no significant difference was observed in bacterial skin diversity between elective and emergency CS, though higher *Lactobacillus* abundance was observed in emergency compared with elective CS. This was proposed to be due to initial labour-associated exposure to the vaginal microbiota [14]. Another study identified the presence of labour to be more influential on the neonatal microbiota than the delivery procedure itself [56]. There is a lack of research investigating microbiome differences between delivery methods that include comparisons between elective and emergency CS, therefore require further investigation [162].

An increase in circulating blood cytokine concentrations has been observed in mothers and their neonates delivered vaginally, lasting up to 4 days post-partum. This was not observed in mothers and their neonates delivered by elective CS, likely due to the absence of labour [163]. These delivery mode differences in inflammatory and immune-mediating cytokine levels may contribute to microbiome differences and associated neonatal immune and health outcomes [153]. Asthma, psychosis and celiac disease have all been associated with CS delivery [164-166].

Alternatively, microbiome differences associated with CS delivery could result from an underlying medical indication [162]. By way of illustration, most CS deliveries are performed due to medical indications such as foetal malpresentation, severe preeclampsia and Haemolysis Elevated Liver Enzymes and Low Platelets (HELLP) syndrome, absence of labour, foetal macrosomia, and history of previous CS [167, 168]. Little is known about how these medical indications affect the neonatal microbiota and further research is warranted [167]. While medical indication is the primary reason for a CS, patient autonomy enables women to choose their birth

experience (for reasons such as a preference for a planned delivery date or fear of pain) causing a surge in elective CS without any medical indication, contributing to this CS “epidemic” [169, 170].

It is evident that there is multi body site (i.e. vagina, gut and skin) sharing of microbes between mothers and neonates, and clear neonatal microbiome differences between delivery modes. Despite these differences, from 1 month of age, infants start to develop colonisation patterns driven by body site [1, 56]. The skin microbiome continues to evolve into a stable community [1, 171]. During infancy, the microbiome starts to resemble an adult microbiome [18, 172-174]. Between 3 weeks and 3 months of age, infants start developing the site-specific bacterial profiles resembling adults [1, 4, 56, 175] (as discussed in section **1.1.1**); though these still differ from the adult skin microbiome until late adolescence [68]. Whilst body site specific functional pathways appear in early life, the microbial community and function may start to change when the infant becomes more exposed to environmental factors [56]. This is consistent with other work which suggested that the skin microbiome community is largely defined by age, and only transiently influenced by factors such as delivery mode, drug use, cleaning and moisturising practices [14]. Many studies have shown that age influences the neonatal and infant skin microbiome development, continuing to influence into childhood, adolescence, adulthood and old age [176-181]. Overall, the skin microbiome literature is lacking research on age groups at different stages of skin aging, and age-independent extrinsic factors such as medication and lifestyle should also be considered [177]. More recently, community assembly was found to be driven primarily by ecological drift, suggesting occurrence of stochastic processes (random events evolving over time) [123]. It may be necessary to understand how long delivery mode differences persist in infants before stochastic processes occur, and whether a critical window of microbiome development is embedded within this complex microbial establishment that may link to long term health [182].

1.2.5 Antibiotic Exposure

Irrespective of delivery mode, the pregnancy and birthing process can encounter complications; these may lead to the use of antibiotics, reducing skin microbial diversity [16, 126]. This appears to be transient, followed by recovery and an increase in diversity over time [16, 126]. Antibiotic use in pregnancy leads to colonisation of the infant skin with genitourinary tract-associated pathogens (*Ureaplasma* and *Gardnerella*) [39]. Furthermore, CS delivery is routinely

accompanied with antibiotics (intrapartum antibiotic prophylaxis) and anti-inflammatory pain analgesics [153, 167]. Antibiotics and duration of exposure can disrupt the neonatal gut and skin microbiome, reducing diversity, abundance and richness [36, 183-189]. These perturbations to the neonatal gut and skin were associated with delayed immune development, increasing the risk of developing a variety of skin and health conditions into adulthood including asthma, allergies, obesity, delayed wound healing, vitiligo. Reduced abundance of protective skin bacteria, such as *Lactobacillus*, is also observed. *Lactobacillus* is known to suppress *S. aureus* infections, AD, and acne [190-202], by inhibiting adherence of *S. aureus* to keratinocytes by competitive exclusion [192]. Lactobacilli lysates (postbiotics) have also shown the same binding inhibition of *S. aureus* to keratinocytes, and stimulated keratinocyte production of antimicrobial peptides as a mechanism for clearing intracellular *S. aureus* [203]. These postbiotics have also demonstrated activation of tight junction proteins that appear to restore the skin barrier, thus promote wound healing, and may consequently contribute to the inhibition of *S. aureus* spreading between epidermal cells [203].

These antibiotic research findings could suggest that the routine practices associated with CS delivery are an important contributor to the neonatal gut and skin microbiome disturbance and health outcomes. Some countries try to combat this through administration of intrapartum antibiotic prophylaxis after the cord is clamped to reduce antibiotic exposure to the neonate [153]. A recent randomized controlled trial (NCT06030713) suggested that the effect of timing of prophylactic antibiotic administration (before CS skin incision and after cord clamping) on the infant gut microbiome was minor, and feeding mode was more important in shaping the gut microbiome [204]. However, research on this topic has been mostly associative, leaving the mechanism and cause undetermined [153]. It is further important to consider the limitation of “dead” microbial DNA present in sequence-based analysis; limiting our knowledge of the effects of antibiotic exposure to the neonate within the first days of life [153].

1.2.6 Vaginal Seeding

To prevent CS associated conditions, interventions have been trialled, such as VS [122, 125]. This exposes CS delivered infants to the maternal vaginal microbiome that is ‘bypassed’ during CS delivery, potentially important for neonatal immune system development, although this does not account for important vertical transfer of maternal gut microbes [122, 125, 159]. VS is currently only performed inside of a

research protocol, approved by a human subject's review board, where a gauze soaked in the vaginal fluids is rubbed over the neonate's entire body [159, 205, 206]. This experimental design may lead to the normalisation of neonatal microbial exposure through VS, which could normalise a healthy microbial development [159].

Currently, there are five vaginal seeding randomised placebo-controlled trials in infants investigating how this process correlates with immunity and health outcomes in infants [45]. To date, one trial (ID: NCT03567707) does not appear to have published any data, however, the other four (ID: NCT03298334, ID: NCT03809390, ID: ChiCTR2000031326 and ID: ACTRN12618000339257) have conducted associated studies, some showing that VS can partially 'restore' microbiota profiles in CS babies and normalise neurodevelopment [122, 125, 207-209]. Conversely, two associated studies demonstrated that VS had no significant impacts on gut microbiota composition, growth, anthropometry or allergy risks during the first two years of life [210, 211].

Concerns have been raised regarding VS infection transmission from the mother's vaginal microbiome to the neonate (such as transfer of GBS, Herpes Simplex Virus (HSV) and Human Immunodeficiency Virus (HIV)), however screening steps have been put in place [122, 125, 208, 212]. It is also important to note that this screening is not standard procedure for VD; and suggesting exposure to vaginal secretions could put all neonates at risk of infection, controversially infers that VD is a neonatal health risk [213]. The short- and long-term risks and benefits are still not fully understood, and this could influence future regulations regarding the unique VS procedure as a drug or biologic [159]. There are suggestions of neonatal probiotic administration being a more controllable and safer option to promote a healthy neonatal microbiome and correct perturbations [153, 189, 214].

1.2.7 Probiotics and Live Biotherapeutics

Probiotics are food supplements used for health that contain live organisms and are administered orally [215]. LBPs are biological drugs that contain live organisms, and can be in the form of topical probiotics or skin bacterial transplantation, to treat a specific condition or disease [215]. Oral and topical probiotics have been proven to reduce infection and antibiotic administration [216] and act by inhibiting skin pathogen growth and promoting a microbial balance [217]. Several probiotic studies have suggested beneficial outcomes in paediatric patients with skin disease [218-220]. Multiple clinical trials for skin probiotic interventions in early life are currently

being/have been conducted for treatment of skin disease, which we have recently reviewed [45]. Furthermore, there is an emerging trend, with promising results, in skin bacterial transplantation to treat skin disease [221-223]. With ongoing clinical trials in this research field, we are a step closer for probiotics and LBPs to be implemented in clinical practice [224]. However research is still underway to establish formulation techniques for safe and effective delivery of these therapeutics for a wide spectrum of dermatological conditions, before being implemented in clinical practice [224].

1.2.8 Gestational Age

After much deliberation on the intertwining effects of delivery mode, labour, antibiotic use and VS, there has been a lack of consideration for differences in skin microbiomes of pre-term and full-term neonatal groups, determined by their gestational age, which has been linked with significant microbiome differences [139, 225-227]. These differences may be due to the influence of gestational age on skin maturity at birth, delivery mode and exposure to the Neonatal Intensive Care Unit (NICU); in turn affecting nutritional differences and antibiotic administration [228, 229]. The multitude of factors involved in gestational age microbiome differences could govern neonatal immune development and health [139, 225-227, 230-233].

1.2.9 The Neonatal Skin Structure and Function

The foetal skin starts to mature and function as a barrier defence at 20-24 weeks of gestation [234]. The outermost layer of the skin (stratum corneum) is formed by the third trimester of pregnancy and provides immune and structural protection from colonising microbes and the environment at birth [235, 236]. Premature neonates have an immature stratum corneum, which results in increased temperature fluctuations, permeability, water loss, electrolyte imbalances, and risk of infection [235]. Term neonates have a stratum corneum maturity resembling that of an adult, whereas pre-term infants reach this maturation level 2-9 weeks after birth [235, 237].

Term neonates also present with a vernix caseosa (white, greasy, lipid-rich matrix) on the skin, produced during the last trimester of pregnancy and facilitating the formation of the acid mantle (slightly acidic film on the skin surface acting as a skin barrier) [238-240]. This is helpful to the neonate at birth as they experience environmental changes. The neonate begins with a uniformly neutral skin pH [241], but within 16 hours of birth skin acidity alterations occur, leading to body site specific differences in skin pH [241]. Skin acidity at this stage prevents pathogenic bacteria

colonising, and encourages commensal bacteria to populate [241]. The vernix caseosa is particularly important at birth as the neonate is suddenly exposed to microbes, toxins, oxidative stress, variable temperatures, and humidity; all of which are managed by preventing water loss, regulating temperature, and modulating the innate immunity [238, 239, 242].

The vernix caseosa contains high levels of antimicrobial substances (such as, LL-37 and lysozyme) and antimicrobial peptides (such as lactoferrin and alpha-defensins) that are suggested to work synergistically to inhibit certain pathogens, such as *Klebsiella*, *Bacillus megaterium*, *Listeria monocytogenes*, *Group B Streptococcus*, *and Candida albicans* [239, 240, 243-246]. These antimicrobial properties are found in much higher levels on neonatal skin than adult skin [236]. However, they were unsuccessful in inhibiting *Pseudomonas aeruginosa*, coagulase-negative *Staphylococcus*, and *Serratia marcescens* [244, 245]. The vernix caseosa has demonstrated an imperative function to the neonate at birth but is still relatively understudied and deserves special attention; particularly, with regard to pre-term neonates that have an immature skin structure and function.

1.2.10 Environment

Due to the immature skin structure and function of pre-term neonates, they are predisposed to pathogen colonisation (with CoNS and *Candida* species) resulting in an unstable skin microbiome, sepsis and immune deficiency [247-250]. Environmental factors, such as longer-term hospital stays, including in NICU (a common phenomenon with pre-terms), encourage carriage of potential pathogens to neonatal skin, and there are differences between countries in the carriage of hospital and NICU pathogens; for example, *Mycobacterium tuberculosis* and *Mycobacteroides abscessus* were found on the skin of neonates in a hospital in India [127]; *Streptococcus* was identified on neonates in a hospital in Sweden [14, 127]; and bacterial colonisation with Gram negatives *Klebsiella pneumoniae* and *Serratia marcescens* appeared to be common on neonates born in South African hospitals [128]. Similarities have been identified between the pre-term skin microbiome and the environment of the infant's incubator/ isolator and air swabs; although the skin microbiome swabs clustered separately from incubator/isolator and control swabs [17]. Neonates, particularly pre-term, have a higher risk of common skin commensal microbes becoming pathogenic [60]. Overall, term and pre-term neonates display differences in their skin microbiome, with pre-terms exhibiting more *Staphylococcus*, *Cornebacterium* and *Prevotella*, and term infants

presenting with more *Brevundimonas*, *Flavobacterium*, and *Sphingobacterium*; pre-terms also have reduced diversity and community evenness [36].

Longer-term environmental factors may also affect the developing infant skin microbiome and health into adulthood [251]; these environmental factors include, family members, pets, climate and urbanisation [14, 77, 251-256]. For example, the presence of dogs during pregnancy has been associated with higher representations of *Betaproteobacteriales* and *Ralstonia* on the infant skin [14]. Exposure to an alpine climate resulted in reduced *S. aureus* abundance on the skin of children with AD [255], potentially due to the higher UV radiation, and lower pollution and allergen levels in an alpine climate [255]. Conversely, urban environments have been associated with a higher risk of developing immune and metabolic diseases [257-259]. Urbanisation is associated with High Income Countries (HIC), which coincide with certain factors (human density, housing, urban planning, home architecture, technology, ventilation, exercise, diet, food/water access, clothing, antibiotic exposure, medicine, personal care and cleaning products) that may lead to microbiome changes [256]. However, there is limited research into skin health and disease in relation to environmental factors, especially in Low Income Countries (LIC) as we have previously highlighted [45].

1.2.11 Skincare

Infantile skincare is important in clinical practice. Skincare regimes in hospitals include the use of neonatal skin cleansing practices with chlorhexidine gluconate, shown to be effective for skin antisepsis and reducing bacterial colonisation of neonates [124, 128]. These products can be harsh on the neonatal skin, therefore emollient application may be necessary to improve skin condition, which has previously been shown to increase *S. aureus* colonisation [128]. Capone et al. was the first study to evaluate the impact of commercialised baby skincare products on the infant skin microbiome and found that routine emollient use in infants increased their skin microbiome richness, ceramide and free fatty acid levels, which play an important role in skin barrier integrity [40]. In contrast, subsequent studies have shown no associations between emollient use and skin bacterial composition [14]. Some studies have shown zinc pyrithione, a commonly added ingredient in skincare cosmetics, to inhibit microbial growth [12, 83].

Industry has seen the influence of skincare products on the skin microbiome as a market opportunity to develop microbiome friendly skincare products [13]. Industry

has focused on producing products containing prebiotics or probiotics to restore beneficial microbiota and healthy skin properties [13, 260]. There should be caution when using some commensal strains in skincare as probiotics, as they can become pathogenic; to mitigate this, low risk probiotics such as lactic acid bacteria (e.g. *Lactobacillus*) are commonly used in skincare cosmetics [13]. *Lactobacillus* has shown beneficial effects on skin health, however the mechanisms by which it improves skin health are unknown [261-263]. Probiotic use in skincare cosmetics appears to be a promising method for improving skin health, and in the future could treat skin disease [264, 265]. However, the effect of probiotic use on delicate infant skin remains understudied, especially in pre-term neonates that have an immature skin structure and microbial diversity.

1.2.12 Breastfeeding

Fragile neonatal and long-term skin health could be improved through early life breastfeeding, as breastmilk associated microbiota are known to influence immune development [266-268]. Breastfeeding has been strongly recommended for pre-term and low birth weight neonates by the World Health Organisation (WHO) [269]. However, pre-terms are typically born via CS, which is associated with a delay in breastmilk production and breastfeeding, sub-optimal breastfeeding behaviour, a reduced duration in breastfeeding, and a reduced volume of breastmilk consumption in the first five days of life [270-272]. The use of donor breastmilk could minimise these issues, and is conditionally recommended by the WHO when the mother's breastmilk is unavailable [269]. Breastfeeding has been associated with the direct transfer of transient microbes (family *Prevotellaceae* and genus *Bifidobacterium*) to the infant skin, which was limited to the typical period of breastfeeding [14, 19]. These transient microbes could have important implications in skin community establishment and health; however there are differing opinions as to whether breastmilk has an effect on the infant skin microbiome composition, allergic disease and AD in infants [121, 123]. Another substitute for when the mother's breastmilk is unavailable is formula milk, which has been compared with breastfeeding in relation to infant skin communities with AD and food allergies [121]. This study revealed that breastfeeding altered allergic infant's skin communities to resemble healthy infants more than formula feeding [121]. This emphasises the importance of breastfeeding in shifting the microbiome toward a 'healthy' state [121].

1.2.13 Diet

After the breastfeeding period, the infant's diet may continue to influence the developing skin microbiome [273], however, research is limited. Diet has been shown to affect both skin physiology and microbiome, and can cause adult Acne vulgaris [274, 275]. Cultural differences in diet have also been shown to impact the skin microbiome [276]. For example, cultural dietary differences resulted in a low microbial skin diversity and a dominance of *Corynebacterium* in westerners compared to Amerindians [276]. Significant weight gain (obesity) and a dominance of *Corynebacterium* have been found on the skin of mice that were fed a high fat diet compared to those that were fed a caloric neutral diet, suggesting that diet and associated obesity alters the skin microbiome [277]. Murine models can be useful for studying host-microbial interactions in complex systems as mice are easier to manipulate compared to humans [3]. However, it is standard practice to feed mice (including controls) a chow-based diet, which is composed mainly of plant materials and differs considerably to the composition and variation in a human diet [278]. Another consideration is that mouse skin is loosely connected to the subcutaneous connective tissue, whereas human skin is tightly connected [279], which could influence results. This highlights the importance of using real human skin (skin explants), although, they are often difficult to access and have a specific window of viability; therefore, other models, such as pig skin, may be a better option due to their easier access and similarity in skin structure to humans [279].

1.2.14 Maternal Obesity

Obesity is also an important maternal factor that could impact the infant skin microbiome and health [280]. Obese mothers have a higher risk of longer or failed labour, resulting in a CS delivery [281-283]; leading onto the aforementioned CS associated health outcomes (section **1.2.1)**. Furthermore, obesity is associated with maternal Gestational Diabetes Mellitus (GDM) or type II diabetes, and a high fat diet, shown to alter the mothers' milk microbiome; all may affect neonatal and infant immune development [56, 284]. Another maternal factor possibly affecting the infant skin microbiome are food allergies [14]. However, these maternal factors are understudied in relation to the neonatal and infant skin microbiome, with previous research particularly focused on neonatal and infant gut microbiome and health outcomes [280].

1.2.15 Genetics

Less-studied factors impacting the infant skin microbiome include genetics and gender [14, 23, 126]. Ethnicity (a demographic variable) is linked with heritable genetic variation, which interacts with the skin microbiome [177]. Host genetics and the environment together have been shown to shape the skin microbiome of twins; research in this area may be useful for estimating the scale of genetic and environmental effects on human health [285, 286]. The hosts genetic variations in particular genes relating to skin barrier function, such as the FLG gene, appear to be associated with the abundance of *Corynebacterium jeikeium* [285]. The FLG gene encodes filaggrin, which is a structural protein in the cornified envelop of the stratum corneum critical for skin barrier function [285]. This association of genetic traits with the skin microbiome could potentially identify genomic and microbial targets for skin healthcare. On the contrary, more recent research on genetics, particularly of FLG gene functional deficiency (responsible for impaired skin barrier function and a major risk factor for AD) has shown no association with the skin microbiome [14].

1.2.16 Gender

Studies comparing the impact of gender on the infant skin microbiome are limited and report contrasting results. Some identified no significant effect [18, 23], whereas others have shown a significant difference in the first week of birth, specifically in the groin area, but the reasons were unclear [126]. Longer term into adulthood, there appear to be gender differences, with women having higher skin microbial diversity and richness than men, possibly related to gender physiological and anatomical differences, such as sebum, sweat and hormone production [2, 287]. For example, *Cutibacterium* are more abundant in men than women on the glabellae (forehead) and back of the hands; due to more terminal hair follicles and sebaceous glands in these areas in men than women, which is a desirable environment for *Cutibacterium* [287]. The role of the father in the neonatal microbiome development is very understudied and may also be an important factor to consider [176].

Overall, there are a multitude of factors contributing to the establishment, development and changes to the infant skin microbiome. This insight into the complex nature of the skin microbiome has only been possible through traditional and emerging sequencing methods [13, 288].

1.3 Current Methods Available to Investigate the Skin Microbiome

Traditionally, microbiome research was performed using microbial cultivation techniques [289]. However, these techniques do not capture a true representation of the microbiota within a sample due to culturing bias [290]. Today, DNA-based methods, such as 16S rRNA gene amplicon sequencing and more recent Shotgun Metagenomic Sequencing (SMS), are frequently used for skin microbiome research as they sequence detectable genetic material within a sample [13, 288] (**Figure 1.3**).

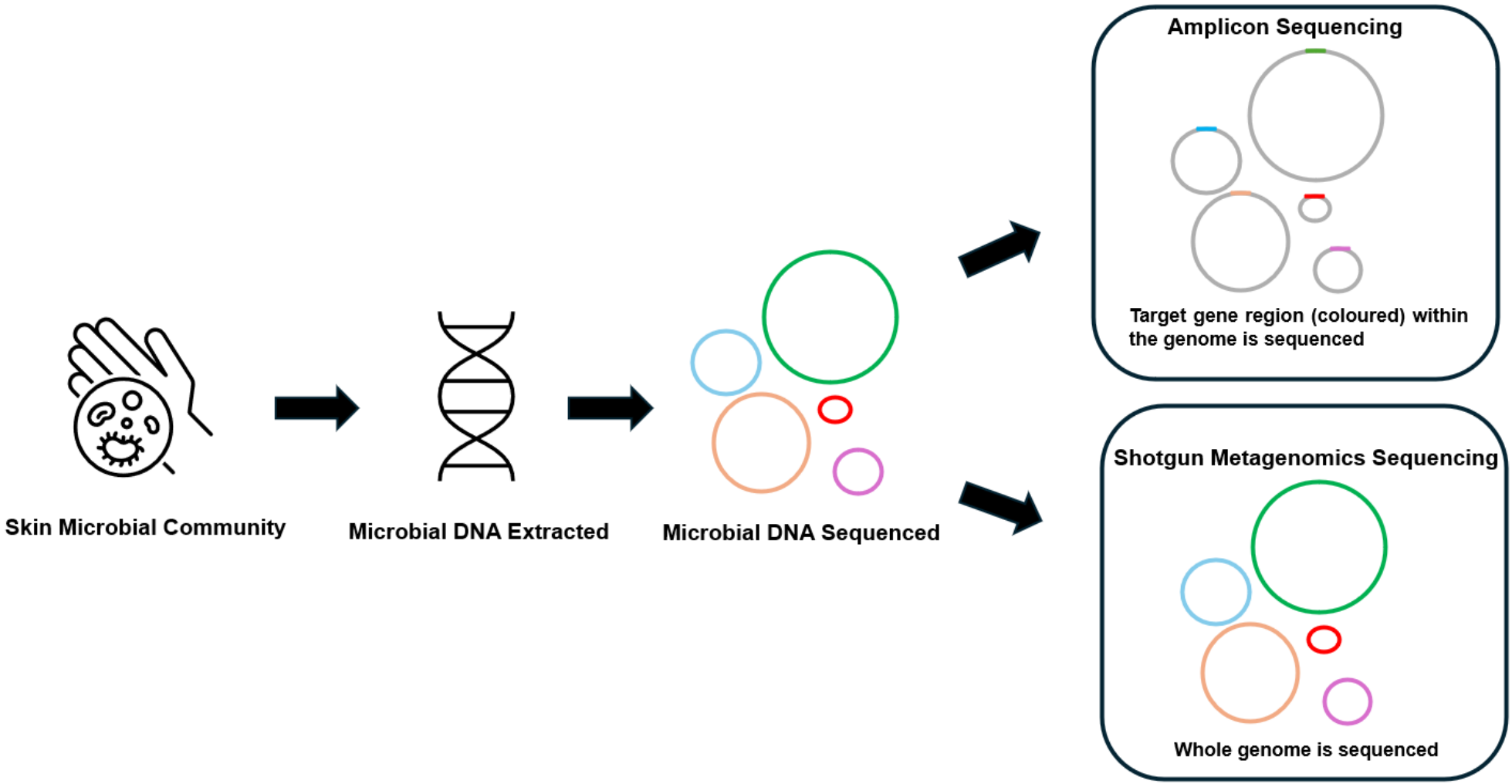


Figure 1.3. Current Amplicon and Shotgun Metagenomic Sequencing Methods

For interpretation of amplicon and SMS results, computational and bioinformatics tools are essential [291, 292]. When using sequencing methods, DNA contamination arising from the DNA extraction kit reagents (kitome), skin to sample or sample to sample, and host DNA should be considered [293, 294]. Furthermore, due to the low microbial biomass of skin samples, these contaminants could result in false positives, which is of particular concern [291, 293, 295-298]. There are bioinformatic tools available to identify and deplete ktiome or other sources of microbial contaminants (SourceTracker, Meta-SourceTracker, decontam, microDecon, Recentrifuge and Squeegee), however there is a risk of removing microbes that aren't contaminants within samples [294]. Host DNA contamination can be depleted before sequencing using mechanical and physiochemical depletion methods, and adaptive sampling (to selectively enrich microbial DNA or selectively deplete host DNA) [299, 300]. Although, adaptive sampling approaches do not reliably target all species and can skew the resulting genomic profiles, and mechanical host DNA depletion can result in some loss of bacteria [300, 301]. Alternatively, host DNA can also be removed after sequencing using bioinformatic depletion methods, which maps host DNA reads to a host reference database [19]. There should also be caution with read misassignment, and misinterpretation of read assignments into something meaningful - phenomenon's that are particularly associated with low biomass samples generating low numbers of reads; Afshinnekoo et al. 2015 [302] and Poore et al 2020 [303] fell victim to this, identifying Y*ersinia pestis* and *Bacillus anthracis* from a New York city subway, and extremely low abundant "contaminant" microbes in cancer patients tissue and blood microbiomes, both of which caused a wave of news stories.

Skin microbiome studies commonly use traditional 16S rRNA gene amplicon sequencing methods to taxonomically classify microbiota to a genus level; some variations in this method can identify species level, but cannot provide information about strain variations or functional capacities in biological processes [288]. Amplicon sequencing targets specific genomic regions (16S rRNA, 18S rRNA and Internal Transcribed Spacer (ITS) genes) [304-306]. The bacteriome is more commonly studied, therefore, the 16S rRNA gene is most commonly targeted as it is ubiquitous and highly conserved between different bacterial species [288, 295, 304, 305]. There are nine hypervariable regions (V1–V9) in the 16S rRNA gene with certain regions being more useful for determining lower- and higher-ranking taxa [305, 307, 308]. Skin microbiome research projects tend to target the V2-V3 and

V3–V4 regions of the 16S rRNA gene [308]. Analysis of the 16S rRNA gene sequence cannot provide information on other skin microbiota such as fungi and viruses; the 18S rRNA and ITS genes are specific targets used for fungi, but viruses lack a universally conserved target therefore pathogen specific primers are required [306, 309, 310]. For these amplicon sequencing techniques Polymerase Chain Reaction (PCR) is required to amplify the targeted variable region of choice [288]. Low Molecular Weight (LMW) DNA is typically used as Illumina technologies are used to sequence this DNA [311].

SMS is an emerging method in microbiome studies, as this method sequences all genetic material in a sample (bacteria, archaea, viruses and fungi), without being limited to amplified regions [13, 304, 312]. This has major advantages, reducing bias from selective amplification efficiency and providing information about the rest of the genomes present [311]. As the entire genetic material in a sample is captured, large volumes of data are generated; this is harder to analyse and requires filtering to remove large volumes of human DNA, thereby becoming more costly [33, 313, 314]. SMS has the further advantage of being able to taxonomically classify bacteria up to strain level [13]. This method can also yield insight into the functional diversity within ecosystems as well as identifying the presence of mobile elements such as plasmids [315, 316]. Liu et al. [317] stated the importance of the research question in determining which sequencing methods to be used, and SMS is advisable if higher taxonomic resolution and functional information are required.

SMS can be performed using multiple technologies, including Illumina, Oxford Nanopore (ONT) and PacBio Single Molecule Real-Time (SMRT) platforms [311, 318]. Illumina uses multiple sequencers (NextSeq, NovaSeq, iSeq, MiniSeq, MiSeq) and a clonal amplification method (which uses individual fluorescent molecule tags attached to the ends of DNA strands) to detect sequences [319]. ONT uses three sequencers (MinION, GridION and PromethION), which measure the ionic fluctuations when single-stranded nucleic acids pass through the nanopore [320, 321]. Within the pore, different nucleotides provide different electrical resistances as they pass through a pore allowing their sequence to be determined [318]. SMRT also uses three sequencers (RSII, Sequel and Sequel II), which detect fluorescence events that correspond to the addition of one specific nucleotide by a polymerase tethered to the bottom of a micro well [322, 323].

The Illumina technology can input High Molecular Weight (HMW) and LMW DNA, which becomes further fragmented, resulting in the sequencing of LMW DNA; Illumina therefore produces short reads with read lengths limited to about 300-500 base pairs (bp) [311, 319, 324]. In comparison to the Illumina technology, the ONT and SMRT technologies input HMW DNA and result in long reads, ranging from read lengths of 500bp to several Megabase pairs (Mbp) [311, 325, 326]. Longer reads facilitate the generation of higher-quality MAGs, used for higher taxonomic resolution and functional information [311, 317, 327]. SMS works by piecing these long and short genome read fragments together [13]. Therefore, long read SMS technologies can reconstruct more complete genomes more easily than the Illumina short read technology due to its preference for HMW input and direct sequencing method [311, 328].

In recent years, more integrative approaches are being used to study the infant skin microbiome and its function, using combinations of culturomics, metagenomics, metatranscriptomics, metabolomics and metaproteomics; Culturomics identifies culturable microbiota, metagenomics provides taxonomic profiles, metatranscriptomics distinguishes functional profiles through transcribed genes, and metabolomics and metaproteomics determine microbe metabolite and protein byproducts respectively [13, 20, 22, 39-43, 329].

1.4 Technical Challenges and Unmet Research Questions

With an increasing popularity in infant skin microbiome research comes several technical and study design limitations that may impact the associations and conclusions drawn. Many studies are longitudinal or randomised controlled trials that focus on multiple timepoints closely clustered temporally, rather than taking a longer-term outlook into infancy [9, 11, 14, 16, 17, 22, 39, 40, 42, 72-75, 97, 98, 123, 126, 127]. Furthermore, given the frequent use of 16S rRNA gene amplicon sequencing methods, many studies are limited due to either low microbial taxonomic resolution and/or lack of microbial functional information [13]. The resolution and robustness of sequencing methods is further impacted by the way skin samples are collected; for example, research has identified the deeper layers of skin tissue to hold stable viable bacterial populations compared with the skin surface [330]. Contrary to this, multiple studies using skin scrapings (which can access these deeper tissues) resulted in reduced DNA quantity and quality inadequate for 16S

rRNA gene sequencing, and elevated host DNA contamination (reviewed in Manos [73]); furthermore, skin scrapings are invasive causing neonatal skin trauma.

Another technical challenge is the use of commercialised DNA extraction methods that are not optimised for skin, to study the infant skin microbiome (we summarised methods in our review [45]). The skin is a low-biomass niche with difficult to lyse microbial components, such as the mycobiome and these affect DNA quantity and quality, thus influencing sequencing outcomes [331]. For example, ITS sequencing, used to profile fungal communities, was identified as not yielding enough reads for analysis [75]. Furthermore, even with rigorous controls in place, low-biomass niches are at a high risk of contamination [41, 43]. New methodologies optimised for skin microbiome research are required to address these technical limitations and improve research outcomes.

Evidentially, there is a pressing need for more longitudinal (following infants from birth into infancy) and higher-resolution infant skin microbiome studies; we highlighted common study types in our review [45]. SMS will provide higher resolution over targeted gene sequencing, allowing species and strain-level profiling as well as functional information to be obtained [13]. To date, there are very few skin microbiome studies that utilise SMS [16, 19, 22, 126, 332]. A consideration for these types of studies is the ability to extract enough high-quality DNA, but ongoing studies in this area are addressing this need [19]. Profiling using SMS will enable researchers to robustly study the infant skin microbiome, including the mycobiome, virome and archaeome, which are significantly understudied in infants when compared with the bacteriome [45]. There is also a research deficit on how the whole microbiome (viruses, fungi, bacteria and archaea) interacts and functions as a community, and how this is associated with skin health or disease [22]. Although there has been a research focus on infant skin microbiome in relation to certain skin diseases, as of yet, no one has been able to distinguish between microbiome-driven and microbiome-independent cases [14, 22, 24, 72-75, 121, 123]. This is a crucial research gap that needs to be addressed for improved treatment interventions [14]. Clinical studies have used culturomic and viromic methods to diagnose skin disease in independent cases and examine certain factors such as the effects of cleansing and moisturising products on skin bacterial colonisation [21, 72, 124, 128]. However, to improve the robustness of these studies, multi-omics techniques should be more widely employed, as currently there are few and far-between [20, 24, 41, 42]. These multi-omics studies would help to identify perturbation-associated microbiota

changes and fully understand their functional role in infant skin health and immunity, before and after birth, which is lacking [45]. This would also accelerate development of improved maintenance, preventative and treatment options that will improve standard of care of full-term and pre-term neonates, therefore optimising their long-term health [45].

Current therapeutic intervention studies underway include oral and topical probiotic use, development of LBPs and improved skin transplantation and wound dressing techniques [72, 96]. Presently, the vast majority of infant skin microbiome studies have only focused on a few HICs, but skin health interventions would be significantly benefitted by diversifying the cohorts that are sampled to a wider geographical and socio-economic reach [45]. This would also help establish a baseline for a healthy infant skin microbiome composition, the influence of factors modulating the skin microbiota and how skin conditions are linked to microbial compositional changes [45]. Certain interventional practices, such as kangaroo mother care (KMC) (skin-to-skin contact between a mother and infant) and breastfeeding, are known to improve neonatal health outcomes, and their practice is advised globally by the WHOs updated 2022 recommendations [269]. The presence of *M. tuberculosis* and *M. abscessus* on Indian neonates diminished with KMC, demonstrating its importance in neonatal health outcomes [127]. Other practices worldwide may also influence seeding of the infant skin microbiota and should be explored further to understand the effects that practices have on birth outcomes and neonatal microbiomes, which could have lasting health implications [333].

Overall, infant skin microbiome research has focused on the bacteriome and investigated a more detailed interaction of certain factors that affect it. Infant skin exposed to these factors and afflicted with skin disease (mainly AD) has been assessed and profiled (mainly in HICs), which has led to initial investigations of probiotic skin health interventions. These research bottlenecks have left open questions on how the whole infant skin microbiome interacts and functions in infant skin health, before and after birth. These could be addressed by profiling cohorts across the globe, and overcoming technical limitations, which include lack of integrative approaches used, overuse of targeted gene sequencing and unoptimised skin collection and DNA extraction methods used. This would improve neonatal skin health interventions and optimise their long-term health.

This PhD project therefore aimed to use SMS and a healthy participant cohort to analyse infant skin microbial composition changes (with a focus on the bacteriome) over time from birth into infancy; also considering microbiome influencing factors (timepoint, pregnancy term, delivery mode and feeding mode). Multi body site (skin, gut and vagina) sharing of microbiota was compared over time between mothers and their infants to investigate skin microbiome establishment. The functional capacity of these shared microbiota in skin health was also investigated.

1.5 Hypothesis

An improved understanding of the development and evolution of the infant skin microbiome will provide insight into the infant skin composition, ecosystem dynamics, and functional roles in skin health within the first year of life. This will further help prediction of probiotics to treat and prevent skin disease in early life and maintain skin health.

1.6 Aims

1. Investigate the development and evolution of a healthy infant skin microbiome over time.
2. Understand the factors that influence the infant skin microbiome development, evolution and skin health over time.
3. Investigate sharing of microbes between mothers and their infants over time that may influence skin health.
4. Identify a PEARL genome MAGs catalogue, which will provide future insight into the functional importance of the shared microbes in infant skin health over time and help formulate the next generation of oral probiotic and LBP interventions to help predict how to alter a microbiome to prevent and treat skin disease and maintain skin health across the life course.

1.7 Objectives

1. Develop an optimised high-throughput automated DNA extraction method, for recovery of bacterial DNA from skin swabs, that results in DNA with high yield and molecular weight, suitable for SMS.
2. Analyse changes to the composition of the developing infant skin microbiome to establish a baseline for a healthy skin microbiome.
3. Compare the developing infant skin microbiome to the maternal skin, vaginal and gut microbiome at birth and within the first year of life, to investigate microbiome establishment and associated influencing factors.

4. Assess changes to the infant gut microbiome and its impact on the skin microbiome development; which could help determine nutritional interventions to treat and prevent skin disease in early life and maintain skin health during the life course.
5. Identify shared microbes between mothers and their infants over time that can be used in future work to elucidate their functional importance in infant skin health; which could help predict how to alter a microbiome to prevent and treat skin disease and maintain skin health in infants.

CHAPTER 2

METHODOLOGY

CHAPTER 2 OVERVIEW

With this chapter, the methodology and cohorts (adult volunteers and the Pregnancy and Early Life (PEARL) Study) used throughout this thesis are introduced. The optimised extraction method developed (Chapter 3) was the basis for all data generated in this project to explore the infant skin microbiome using the PEARL cohort. In this chapter, cohort specific details are covered for ethics, study design, sample collection and handling, and ethical and regulatory considerations for data protection and donor confidentiality. Methods are also outlined for sample processing, culturing and bacterial DNA extractions for Shotgun Metagenomic Sequencing (SMS), relevant assays, SMS, statistical analysis and data analysis. I performed the vast majority of laboratory and bioinformatics work for my thesis described in this chapter: (1) I optimised a human skin swab specific microbial DNA extraction method through performing a series of initial diluent and lysis laboratory experiments using infant skin derived bacterial isolates and bioinformatics analysis; (2) I then applied for, and obtained, Faculty of Medicine and Health Sciences (FMH) ethical approval, and prepared a human tissue study protocol for my clinical study using adult volunteer skin swab samples, to validate the skin swab specific microbial DNA extraction method that I optimised; (3) I composed an SOP for the optimised skin swab microbial DNA extraction method; (4) Using the SOP that I composed, I then applied the validated optimised DNA extraction method to the PEARL study cohort consisting of mothers and their infants; (5) I processed skin and vaginal swabs and extracted DNA from them, and performed culturing of these samples alongside the DNA extraction, which I also extracted using the optimised method; (6) I performed the bioinformatics analysis for the skin, vaginal, stool and culture metagenomes that resulted from the shotgun metagenomic sequencing data (taxonomic profiling, Metagenome Assembled Genomes (MAGS), data visualisation and statistical analysis). Exceptions of work that were performed by other scientists included the stool DNA extractions (performed by the QIB PEARL study team), the shotgun metagenomics library preparation and sequencing (performed by the QIB sequencing team and Genewiz), the data clean up and strain level analysis (performed by the QIB bioinformatics support team), the archaeal taxonomic profiling (performed by Charlotte Neumann; I visualised and performed statistical analysis for the archaeal taxonomic profiles), and the optimisation of the generation of MAGs in Chapter 3 (performed by Dr. Raymond Kiu).

2.1 Adult Volunteer and PEARL Study Cohorts

2.1.1 Ethical Statement and Regulatory Considerations for Data Protection and Donor Confidentiality

I obtained ethical approval for the adult volunteer recruitment, skin swab sampling and processing from the University of East Anglia (UEA) Faculty of Medicine and Health Sciences (FMH) ethics committee (FMH ethical approval reference: 2020/21-065). The recruitment, sampling and processing were performed under the NNUH Biorepository ethics (Research Ethics Committee (REC)); the biorepository team recruited and collected samples, and I processed all samples. Documents relating to the FMH ethical approval, including the ethics approval letter, my application form, the human tissue study protocol that I composed, and my registration forms, can be viewed in the Appendices (**Appendix 1.**; **Appendix 2.**; **Appendix 3.**; **Appendix 4.**). The Biorepository ethics covered the volunteer recruitment, sampling and processing. All samples were anonymised from the researcher, and I discarded the swabs after use. I bioinformatically identified and deleted all host DNA sequences (i.e., DNA reads mapping to the human genome). DNA preparations were only kept for 12 months, in case any further methodological verification, in terms of sequence library preparation was needed. DNA preparations were then destroyed according to the established waste disposal procedures within QIB. Bacterial DNA sequences were kept on secure servers within QIB and were anonymised from the donor.

The PEARL study [334] is registered in the ClinicalTrials.gov Database with ID: NCT03916874. Ethical approval for this study was set up by the chief investigator Professor Lindsay Hall and the PEARL study team and was approved by the Quadram Institute Bioscience (QIB), Human Research Governance Committee (HRGC), the Health Research Authority (HRA) and the NHS London-Dulwich REC, obtained for the NNUH biorepository that I was working under (HRA REC NHS London-Dulwich reference: 18/LO/1703). This study was conducted in accordance with the principles of the Declaration of Helsinki, the conditions and principles of the International Conference on Harmonisation Good Clinical Practice (ICH GCP), and in compliance with the UK national law. The research meets the requirements of the EU General Data Protection Regulation (GDPR), UK Data Protection Act 2018 and the relevant sponsor's policies (IRAS number: 241880). For further details and documents relating to the PEARL study approval see Phillips et al. [334]. The PEARL study complies with EU GDPR, which came into force in the UK on 25 May 2018 and the UK Data Protection Act 2018, which requires data to be anonymised

as soon as it is practical to do so. Personal data were subjected to appropriate safeguards and a specific Standard Operating Procedure (SOP) was established with the Norwich Biosciences Institute (NBI) data protection officer. This ensured that the study team had appropriate organisational and technical measures in place to ensure that data were processed lawfully, fairly, in a transparent manner and were kept to a minimum and secure in the research context.

2.1.2 Study Design

For the adult volunteer cohort, the Biorepository recruited and consented adult volunteer participants ranging between the ages of 23 and 65. There was no contact between the researcher and participant to ensure anonymity. Volunteers were recruited from QIB via an email from the Biorepository. To be eligible, volunteer participants should have had no current skin conditions or been prescribed antibiotics over the last 3 months. As this study was carried out during the COVID-19 pandemic, 12 volunteer participants were recruited, consented, and self-swabbed remotely at one timepoint, to adhere to the COVID-19 regulations reducing person-to-person contact. The volunteer participants were provided with the established Biorepository Participant Information Sheets (PIS) and were consented with the established Biorepository Consent Forms (CF) (**Appendix 5.**). The Biorepository contact information was also provided to the volunteer participants for them to express any questions or concerns regarding the study. The Biorepository could then contact the researcher for further clarification.

For the PEARL study cohort, there was a maximum of 12 collection timepoints for each PEARL study participant over approximately 31 months. There are 250 PEARL mother participants, which were representative healthy pregnant women recruited from across Norfolk. The PEARL study is based on a healthy cohort, therefore, early pre-term births were excluded from the PEARL study. Participants had planned to give birth at the NNUH National Health Service (NHS) Trust or have a home birth with maternity care provided by the NNUH. The study was divided into three collection phases. Phase 1 was during each pregnancy trimester, phase 2 was from birth to 3 weeks post birth, and phase 3 was 4-24 months after birth; there was no follow up past 24 months. Participants were asked to collect appropriate samples (skin, stool/meconium, vaginal, urine, colostrum/breastmilk, blood/cord blood and questionnaires on lifestyle, health and medical conditions) from themselves and their infants during each phase; maternal samples were only collected up to month 4 due to budget and feedback from focused discussion groups with mothers

regarding pragmatism. Samples included in this thesis were skin, gut, vaginal and questionnaires. Most of the sample collection was at home using collection packs provided to the participants. Participants were provided with a small benchtop freezer where they stored their samples until an authorised member of the research team collected samples from the participant's home. The only exception would be if the participant or their infant was an inpatient at NNUH where they were able to collect their samples at the hospital. If they wished, these could be given to their midwife/obstetrician to put in the PEARL study freezer located at the hospital. The duration of the recruitment period was 18-24 months from the project starting. Recruitment was mainly opportunistic at NNUH antenatal clinics and midwife led community clinics in General Practice (GP) surgeries and health clinics. Recruitment was facilitated by the close links with the NNUH and community midwife network. In response to the COVID-19 pandemic, digital appointments were put in place for pregnant women that replace the community clinics in GP surgeries and health clinics. The informed consent procedure took place up to the 22-week gestation point in a confidential space in the NNUH, at GP practices/health clinics where a community midwife clinic is held, at the QIB Clinical Research Facility (CRF), or via virtual means (video/telephone call). The individuals obtaining informed consent were GCP trained, suitably qualified and experienced, have been authorised to do so by the chief/principal investigator and appear in the delegation of duties log. Participants were consented with established participant information sheets and consent forms. A unique PEARL study identification number was provided to each participant. For further details on the PEARL study protocol, see Phillips et al. [334]. An overview of the PEARL study design is summarised in **Table 2.1**.

Table 2.1. Overview of the PEARL Study Design. Extracted from Phillips et al. [334].

		Samples collected from participants' homes by research team at these 4 timepoints											
		Phase 1 (pregnancy)			Phase 2 (birth)			Phase 3 (following birth)					
		Trimester 1 (week 12)	Trimester 2 (week 23)	Trimester 3 (week 30)	At birth	Week 1	Week 3	Month 4	Month 8	Month 12	Month 16	Month 20	Month 24
Informed consent up to 22 weeks													
Mother's samples	Urine Stool Low vaginal swab Skin swab	✓	✓	✓	✓	✓	✓	✓					
	CDCP questionnaire		✓										
	Health questionnaire		✓										✓
	Dietary questionnaire (optional)		✓					✓					✓
	Colostrum/ breast milk (optional)				✓	✓	✓	✓					
	Blood samples (optional)	Please provide samples at 12, 16, 20 and 28 weeks if you can											
Baby's samples	Stool, meconium and skin swabs				✓	✓	✓	✓	✓	✓	✓	✓	✓
	CDCP questionnaire						✓	✓	✓	✓	✓	✓	✓
	Cord blood (optional)				✓								

2.1.3 Sample Collection and Handling

Skin Swabs

The adult volunteer consent and swab sampling were condensed into one day. Black top charcoal amies transport medium cotton swabs (M40-A2, Technical Service Consultants Ltd.) were used in the study. The 12 volunteer participants were provided with swabs, which were already labelled with the date of sample collection, and sample IDs; these were completely unlinked to the participants personal information to eliminate the chance of identifiable information being given to the researcher. The volunteer participants self-swabbed at the QIB, using the PEARL study self-swabbing instructions as follows.

1. To perform swabbing, pull and twist the swab out of its tube taking care not to touch the tip (swab end) with anything other than the area to be swabbed.
2. Holding the swab in one hand, roll the swab end along the inside of the forearm from the wrist to the elbow pit ensuring that the whole of the swab end has come in to contact with the volunteer participant skin.
3. Replace the swab back into the tube ensuring the lid is on tight.
4. The swab may leave a grey residue on the forearm. Once sampling is complete, wash the area with soap and water.

The date of sample collection and Biorepository sample ID were written on the container in pen. Mother and infant pairs were allocated the same sample IDs and twins were allocated with either “1” or “2” after the sample ID to indicate twin 1 or twin 2 (e.g. 2-1 and 2-2). PEARL samples were frozen and the volunteer samples refrigerated until collected for testing.

The optimisation study volunteer participants swabbed themselves twice, one swab on each arm, to produce 24 biological samples in total. The participants then returned their samples to the Biorepository staff the same day, who ensured anonymity prior to swab transfer to the researcher. I collected and tested the samples the same day swabbing was performed (24.08.2021).

The PEARL mother participants swabbed only one arm of themselves and their infants following the PEARL study self-swabbing instructions. The PEARL swabs were labelled with a yellow sticker stating that they were skin swabs. Mothers should have avoided showering, bathing, or using any perfume or creams on their arms or infants’ arms in the 8-10 hours before taking the swab and advised to take the swab

first thing in the morning. Mothers were told to use the same forearm each time they swabbed theirs and their infants' skin.

Vaginal Swabs

The same type of swab was used to collect the skin swabs, but with a blue top lid (M40-A, Technical Service Consultants Ltd.). Swabs were labelled with a yellow sticker stating that they were low vaginal swab. The maternal vaginal swabs were self-collected by the mothers following the PEARL study self-swabbing instructions, as follows.

1. Pull and twist the swab out of its tube taking care not to touch the tip (swab end) with anything.
2. Find a comfortable position either sitting, squatting, or standing.
3. Insert the tip of the cotton swab about 2 cm (about 1 finger joint) inside the vagina.
4. Rotate the swab for 10 seconds and then hold the swab still for a further 10 seconds inside the vagina.
5. Remove the swab.
6. Replace the swab back into the tube ensuring the lid is on tight.
7. Write the date of sample collection on the sample container in pen.
8. Freeze the sample in the sample freezer provided.

Stool Samples

The mothers collected their own stool samples and infant stool samples following the PEARL stool collection instructions.

Maternal Sample

1. Empty your bladder before collecting the stool sample, to prevent accidental contamination of the sample.
2. Ensure the toilet seat is dry by wiping with toilet paper.
3. Peel both sides of the provided faeces catcher out in the directions of the arrows.
4. Stick the faeces catcher to the rim of the toilet seat.
5. Produce stool onto the faeces catcher and ensure used toilet paper goes into the toilet water so as not to cover the stool sample.
6. Unscrew the cap of the stool collection pot provided.

7. Hold the collection pot close to the sample on the faeces catcher with one hand and, with your other hand use the scoop attached to the lid of the sample container to fill the collection pot until it is about half full.
8. Screw the lid back on the collection pot securely ensuring the lid is not cross-threaded.
9. Flush the faeces catcher once sample has been collected (the faeces catcher is fully biodegradable).
10. Write the date of sample collection on the sample container in pen.
11. Freeze the sample in the sample freezer provided.

Infant Sample

1. Open the infant's nappy.
2. Unscrew the cap from collection pot provided and using the scoop, transfer as much of the stool as possible into the collection pot, or until it is about half full.
3. Replace the cap taking care that is not cross-threaded.
4. Write the date of sample collection on the sample container in pen.
5. Freeze the sample in the sample freezer provided.

Questionnaires

PEARL study mother participants completed questionnaires for health, lifestyle, and diet during the study. Questionnaires specifically asked for:

- Pregnancy, breastfeeding, and infant feeding practices;
- Information on lifestyle, diet, medication and medical history;
- Dietary perceptions and preferences.

The research team also collected clinical information from GP and hospital records that related to:

- Changes in the mother's or her infant's medical history during the study;
- Routine data collected by health visitors, which included information on fine and gross motor skills and problem solving;
- Routine pregnancy scans and tests.

For further details on the questionnaire and template see Phillips et al. [334]. For details on the subset of PEARL samples used in this thesis see **Table 2.2** for the PEARL infant participant metadata and **Table 2.3** for PEARL participant sample

collection timepoints. The sample subsets used for this thesis included timepoints from birth up to 12 months post birth. Adult volunteer metadata for the optimisation study was fully anonymised and therefore not provided.

Table 2.2. PEARL Infant Participant Metadata

Sample ID	Gestational Age (Weeks)	Pregnancy Term	Delivery Mode	Feeding Mode (Birth - Month 12)	Skin Disease
1	37.57	Late Preterm	Spontaneous Vertex (VD)	Breast	Not Detected
2-1	37	Late Preterm	Elective CS	Breast/Formula	Not Detected
2-2	37	Late Preterm	Elective CS	Breast/Formula	Not Detected
3	36.57	Late Preterm	Spontaneous Vertex (VD)	Breast	Not Detected
4	37.57	Late Preterm	Emergency CS	Breast/Formula	Not Detected
5	39.14	Full Term	Spontaneous Vertex (VD)	Breast	Not Detected
6	40.71	Full Term	Spontaneous Vertex (VD)	Breast	Not Detected
7	39.14	Full Term	Elective CS	Breast/Formula	Not Detected
8	39.29	Full Term	Spontaneous Vertex (VD)	Breast	Not Detected
9	40	Full Term	Spontaneous Vertex (VD)	Breast/Formula	Not Detected
10	39.86	Full Term	Elective CS	Breast	Not Detected
11	38.14	Full Term	Elective CS	Breast	Not Detected
12	38.57	Full Term	Spontaneous Vertex (VD)	Breast/Formula	Not Detected
13	39.43	Full Term	Elective CS	Breast	Not Detected
14	40.14	Full Term	Spontaneous Vertex (VD)	Breast	Not Detected

Note: “Not Detected” under the Skin Disease section signifies that no skin disease was present in neonates from birth into infancy up to 12 months of age (the period of study for this thesis)

Table 2.3. PEARL Participant Sample Collection Timepoints

Mother	Sample ID	Trimester 1	Trimester 2	Trimester 3	Birth	Week 1	Week 3	Month 4	Month 8	Month 12
	1	N/A	05/06/2019 16:09	24/07/2019 16:11	14/09/2019 16:13	07/10/2019 16:18	22/10/2019 16:16	17/01/2020 16:19	N/A	N/A
	2	N/A	19/09/2019 9:51	04/11/2019 9:53	Not Provided	31/12/2019 10:02	13/01/2020 10:00	26/04/2020 12:55	N/A	N/A
	3	03/10/2019 10:50	Not Provided	Not Provided	21/03/2020 0:00	27/03/2020 0:00	10/04/2020 0:00	17/07/2020 8:30	N/A	N/A
	4	N/A	Not Provided	Not Provided	24/01/2020 0:00	30/01/2020 11:20	14/02/2020 11:22	19/05/2020 13:48	N/A	N/A
	5	01/11/2019 14:37	19/01/2020 14:40	09/03/2020 15:35	10/05/2020 10:44	Not Provided	Not Provided	11/09/2020 16:15	N/A	N/A
	6	N/A	26/01/2020 14:50	17/03/2020 12:10	Not Provided	09/06/2020 15:20	18/06/2020 21:15	29/09/2020 10:20	N/A	N/A
	7	N/A	10/03/2020 9:00	01/05/2020 8:00	08/07/2020 13:52	17/07/2020 11:02	30/07/2020 9:00	04/11/2020 7:00	N/A	N/A
	8	N/A	24/11/2020 14:45	17/01/2021	17/03/2021	25/03/2021 13:00	08/04/2021 17:45	19/07/2021 7:30	N/A	N/A
	9	N/A	19/10/2020 8:50	24/11/2020 13:15	16/02/2021 20:00	23/02/2021 8:30	07/03/2021 10:00	13/06/2021	N/A	N/A
	10	N/A	22/12/2020 19:30	09/02/2021 10:30	17/04/2021 21:00	21/04/2021	07/05/2021 10:45	17/08/2021 6:45	N/A	N/A
	11	N/A	21/10/2020	12/12/2020	04/02/2021	11/02/2021 15:30	25/02/2021 19:00	04/06/2021 9:50	N/A	N/A
	12	N/A	22/10/2020	15/12/2020	15/02/2021	20/02/2021	06/03/2021	13/06/2021	N/A	N/A
	13	N/A	29/10/2020 18:15	23/12/2020	24/02/2021 7:00	02/03/2021 8:45	17/03/2021 11:45	21/06/2023 19:45	N/A	N/A
	14	N/A	08/03/2020 15:48	Not Provided	08/07/2020 12:51	Not Provided	29/07/2020 9:00	06/11/2020 13:00	N/A	N/A
Baby	**Sample ID**	**Trimester 1**	**Trimester 2**	**Trimester 3**	**Birth**	**Week 1**	**Week 3**	**Month 4**	**Month 8**	**Month 12**
	1	N/A	N/A	N/A	14/09/2019 16:22	07/10/2019 16:26	22/09/2019 16:24	15/01/2020 16:27	Not Provided	15/09/2020 10:00
	2-1	N/A	N/A	N/A	19/12/2019 10:14	31/12/2019 10:17	13/01/2020 10:21	26/04/2020 13:00	24/08/2020 8:15	23/12/2020 12:35
	2-2	N/A	N/A	N/A	19/12/2019 10:23	31/12/2019 10:26	13/01/2020 10:25	26/04/2020 11:00	24/08/2020 9:30	24/12/2020 17:45
	3	N/A	N/A	N/A	21/03/2020 0:00	26/03/2020 0:00	09/04/2020 11:33	17/07/2020 11:45	16/11/2020 13:00	17/03/2021 9:00
	4	N/A	N/A	N/A	23/0/2020 11:27	29/01/2020 11:29	13/02/2020 11:30	18/05/2020 14:00	16/09/2020	16/01/2021
	5	N/A	N/A	N/A	Not Provided	Not Provided	Not Provided	11/09/2020 13:50	09/01/2021 13:30	08/05/2021 2:45
	6	N/A	N/A	N/A	28/05/2020 15:41	Not Provided	18/06/2020 21:15	29/09/2020	27/01/2021 17:12	27/05/2021 19:15
	7	N/A	N/A	N/A	Not Provided	17/07/2020 11:45	29/07/2020 00:30	05/11/2020 14:00	19/03/2021 8:00	09/07/2021
	8	N/A	N/A	N/A	18/03/2021	25/03/2021 11:30	08/04/2021 17:00	19/07/2021 11:30	20/11/2021 13:00	18/03/2022 18:30
	9	N/A	N/A	N/A	17/02/2021 8:00	23/02/2021 7:00	07/03/2021 10:30	13/06/2021	12/10/2021	14/02/2022
	10	N/A	N/A	N/A	16/04/2021 14:30	21/04/2021	07/05/2021 21:00	16/08/2021 6:30	20/12/2021 14:25	14/04/2022 13:50
	11	N/A	N/A	N/A	04/02/2021	11/02/2021 13:00	25/02/2021 17:00	03/06/2021 16:00	04/08/2021	07/02/2022 8:30
	12	N/A	N/A	N/A	10/02/2021	21/02/2021	05/03/2021	13/06/2021	11/10/2021	12/02/2022
	13	N/A	N/A	N/A	17/02/2021 16:00	02/03/2021 16:00	15/03/2021 15:00	22/06/2021 20:00	21/10/2021 18:00	17/02/2022 7:00
	14	N/A	N/A	N/A	04/07/2020 13:03	11/07/2020 11:30	24/07/2020 5:13	07/11/2020 11:30	02/03/2021 14:00	20/07/2021 8:30

Note:
N/A under trimester 1 in the mothers' data signifies no sample was expected due to some mothers not being recruited at this timepoint
N/A under month 8-12 in the mothers' data signifies no sample was expected due to sample collection being stopped at month 4 for mothers
N/A under trimesters 1-3 in the neonate/infant data signifies no sample was expected due to the unborn neonate

2.2 Sample Processing, Culturing and DNA Extraction Method for Shotgun Metagenomic Sequencing

I transferred skin and vaginal swabs into Phosphate Buffered Saline (PBS) and cultured them aerobically (using the Walker class II microbiological safety cabinet and Binder static incubator BD 115) and anaerobically (using Baker Ruskinn Concept 400 or 1000 anaerobic cabinets and the central QIB triple mix gas supply - 5% CO2, 10% Hydrogen, 85% Nitrogen) in parallel with the swab DNA extraction. I scraped these aerobic and anaerobic cultures into 40% glycerol stocks; I performed sweep-plate metagenomics to characterize the cultivated skin microbiota. I extracted the DNA from the processed swabs and cultures using the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit protocol, optimised for low biomass swabs in Chapter 3. This extraction protocol was optimised in Chapter 3 and included an overnight heated lysis, an offboard heated lysis and a bead beat lysis (using the Qiagen tissue lyser). Stool samples were extracted by the PEARL study team using the Promega Maxwell RSC Pure Food GMO kit and Promega Maxwell® RSC 48 instrument. A cell-free, diluent-only sample was included as a negative control on each extraction run to ensure no contamination had occurred. An established commercial skin mock community (American Type Culture Collection (ATCC) skin microbiome whole cell mix) was included as a positive control [335] for the skin and vaginal runs, and a stool mock community (M004Base and M008Base) was created by the PEARL study team and used as a positive control for the stool runs (**Appendix 6.**), to determine if lack of DNA yield from swabs was due to absence of material or an extraction problem. I also prepared dilutions of the positive control sample to validate skin DNA extraction efficiency and identify a cut-off point of starting material needed for SMS (Chapter 3). See Appendices for details on sample processing, culturing and DNA extraction protocols for swabs, stools, equipment and reagents used (**Appendix 7.**; **Appendix 8.**).

2.3 Qubit Assay

I performed High Sensitivity (HS) and Broad Range (BR) Qubit assays to quantify DNA extracts according to manufacturer's instructions, using the Qubit 2.0 fluorometer instrument.

2.4 Chapter 3

2.4.1 Bacterial Strains

I optimised a skin microbial DNA extraction method in Chapter 3, which involved the use of the bacterial strains listed below. I chose the strains based on their source from neonatal skin at the NICU department at the Norfolk and Norwich University Hospital (NNUH), and the requirement for a good representation of Gram-negative and Gram-positive skin bacteria.

Staphylococcus aureus (Sample ID. 21 / Isolate No. 60) (*S. aureus*)

Staphylococcus epidermidis (Sample ID. 274 / Biorepository No. 16TB0533) (*S. epidermidis*)

Pseudomonas aeruginosa PA14 (Sample ID. 29) (*P. aeruginosa*)

Escherichia coli EC18PR-0166-1 (BW25113) (Sample ID. 129) (*E. coli*)

The *Staphylococcus* species were isolated by Sethi et al. [336] from skin swabs from healthy neonates on Colombia Blood Agar (CBA). Colonies consistent with Staphylococci were isolated onto Mannitol Salt Agar before being sequenced [336]. These can be obtained from Professor Mark Webber on request. The *Pseudomonas* and *Eshcerichia* species are both reference strains and can be obtained from the National Collection of Type Cultures.

2.4.2 Overnight Culture Methods

I prepared overnight cultures by inoculating from bacterial glycerol stocks into either 10 ml of Luria Bertani (LB) broth in a falcon tube or onto LB agar plates using the quad-streak method (media prepared by QIB media kitchen). I incubated them for 24 hours at 37°C (for protocols see **Appendix 9.** and **Appendix 10.**).

2.4.3 Tapestation Assay

To measure molecular weight, I performed Tapestation assays on DNA extracts from the culture lysis experiments and adult skin validation experiments. I used the D5000 and HS D5000 Tapestation kits accordingly, using the Agilent 2200 instrument as per the manufacturer protocols.

2.4.4 Illumina Library Preparation for Shotgun Metagenomic Sequencing

Sequencing libraries were prepared and sequenced by the QIB sequencing team for both Illumina and ONT sequencing of the skin and culture samples. Libraries were sequenced on the Illumina NextSeq500 and ONT minION platforms for SMS.

For Illumina library preparation, the tagmentation-based library construction with the Illumina DNA Prep Kit (Illumina Catalogue No. 20018704) was used. Where possible, samples were normalised to 5 ng/µl with PCR-grade water. 10% of skin swabs and 70% of skin cultures were normalised. The remaining skin swab and cutlure DNA concentrations were lower than 5 ng/µl and were used at their existing concentration. A bead tagmentation of metagenomic DNA was then performed whereby transposomes are randomly inserted into the genomic DNA adding Illumina adapters enabling barcoding and amplification of the libraries for multiplexing on an Illumina run. This tagmentation was a reduced reaction (20-fold dilution) and was prepared by mixing 0.5 µl of Tagmentation buffer (TB1), 0.5 µl Bead Linked Transposomes (BLT) (Illumina Catalogue No. 20018704) and 4 µl PCR grade water. This master mix (5 µl) was then added to a 96-well plate. Normalised DNA (10 ng) was pipette mixed with the tagmentation mix and heated (55 ºC, 15 min) in a PCR block. For samples containing less than 5 ng/µl, a separate tagmentation master mix was used with no water added and instead 6 µl input DNA volume. Tagmented samples were barcoded by setting up the following PCR master mix: 10 µl KAPA 2G Fast Hot Start Ready Mix (Merck Catalogue No. KK5601) and 2 µl PCR grade water per sample. The master mix was added to 1 µl of 10 µM custom made primer stock containing both P7 and P5 Illumina 9 bp barcodes (Illumina Catalogue No. FC-131-2001 to 2004). Finally, the 7 µl of tagmentation mix was added and mixed. The PCR was run using the following protocol: 72 ºC 3 min, 95ºC 1 min, 14 cycles of (95ºC 10 s, 55 ºC 20 s, 72 ºC 3 min).

Individual libraries were quantified using the Promega QuantiFluor® dsDNA System (Catalogue No. E2670) and run on a GloMax® Discover Microplate Reader. They were also run on the Agilent Tapestation 4200 instrument using a D5000 screentape (Agilent Catalogue No. 5067-5579) to check the library size distribution. Libraries were pooled based on DNA concentration. The pooled library was double-SPRI (Solid Phase Reversible Immobilization) size selected between 0.5 and 0.7X bead volumes using sample purification beads supplied in the Illumina DNA prep kit (Illumina® DNA Prep, (M) Tagmentation (96 Samples, IPB), 20060059). The final Illumina pooled library was then quantified using the Qubit 3.0 fluorometer and run on the Agilent 4200 Tapestation instrument. Both the concentration and average insert size were used to estimate the molarity of the final pool. The final pool was denatured and diluted to 1.5 pM. Libraries were loaded on the QIB in-house NextSeq500 Instrument using a Mid-output 300 cycle kit (Illumina Catalogue FC-

404-2003) with paired-end 150 bp chemistry. The Illumina recommended denaturation and loading recommendations were followed, which included a 1% PhiX spike in (PhiX Control v3 Illumina Catalogue FC-110-3001). Approximately 300 Mbp data was generated per sample. Conversion of the raw data into Fast Quality (FastQ) file format was performed by the Bioinformatics team at QIB.

2.4.5 Oxford Nanopore Library Preparation for Shotgun Metagenomic Sequencing

Oxford Nanopore libraries were prepared and sequenced in-house at QIB for the Chapter 3 optimisation samples only, due to cost limitations. A novel modified Illumina DNA Prep ultra-cost-effective method (Illumina® DNA Prep, Tagmentation 20018704) was used. Primers were designed with a 3' end compatible with the Nextera transposon insert, a 24 bp barcode at the 5' end and included a pad and spacer. The same barcode was used at each end using up to 96 combinations. Primer pairs were mixed at a concentration of 10 µM. The tagmentation reaction as described above was set up. Where possible, DNA was normalised to 50 ng/µl and 2 µl was used in the tagmentation reaction. 30% of skin swab cultures were normalised. For the remaining skin swab culture and swab DNA that could not be normalised, 6 µl DNA was used and the water was omitted from the master mix; 90% of skin swab DNA was below 6 µl DNA and their existing concentration was used. A PCR master mix was prepared using 10 µl NEB LongAmp Taq 2x Master Mix (Catalog no. M0287S) and 2 µl PCR grade water per sample. The appropriate primer pair (1 µl at 10µM) was added to each well. Finally, the tagmentation reaction was added and mixed. The PCR was run as follows: 94 ºC 1 min, 14 cycles (94 ºC 30 s, 55 ºC 1 min, 65 ºC30 min).

Following the PCR, libraries were quantified using the Promega Quantifluor dsDNA system (Catalogue No. E2670) and run on a Promega Glomax plate reader. Equal amounts of libraries were pooled together and the SPRI size was selected at 0.5X bead volume using sample purification beads (Illumina® DNA Prep, (M) Tagmentation (96 Samples, IPB), 20060059) with a final elution in 30 µl. The pooled library was prepared for sequencing on a MinION flow cell (R9.4.1) following the Nanopore SQK-LSK109 protocol with the following modifications: The DNA end-repair reaction was extended to 15 min, the formalin-fixed paraffin-embedded (FFPE) repair buffer and enzyme mix were omitted, instead 7 µl Ultra II End-prep buffer and 3 µl Ultra II End-prep enzyme mix (cat no. E7546L) were used. The library was loaded on the MinION in accordance with manufacturer's instructions. The data

coverage varied from 300 Mbp to 10 Gbp per sample. The Bioinformatics team at QIB converted the raw data into FastQ file format.

2.4.6 Data Clean Up

I obtained the SMS data through the Integrated Rapid Infectious Disease Analysis (IRIDA) platform (irida-19.09.2) [337], and uploaded the data to the QIB Galaxy platform (V.19.05) [338, 339]. Here, I cleaned the data by removing human (classified) reads using the Kraken2 tool (V.2.1.1+galaxy0) [340] (confidence level set to 0.1) to map human reads to a human database (human_20200311). This human database was custom built by the QIB core bioinformatics team using a human genome available on the Reference Sequence (RefSeq) database (available from 11.03.2020), and the custom database manual [341] (see Github link for details: Manual · DerrickWood/kraken2 Wiki · GitHub). I filtered adaptors and checked the quality of sequence reads using the tools Fastp (V.0.19.5+galaxy1) [342] (all default settings used, except for quality phred, which was set to 20) and Fast Quality Control (FastQC) (V.0.72+galaxy1) (default settings used) [343] respectively.

2.4.7 Taxonomic Profiling

I generated taxonomic profiles using the Galaxy tools Kraken2 (V.2.1.1+galaxy0) [340] and Bracken (V.2.2) [344]. I mapped microbial reads to genus and species using a microbial metagenomic database (Kraken2_db_20190111) in Kraken2 (V.2.1.1+galaxy0) [340] (confidence level set to 0.1). This microbial database was custom built by the QIB core bioinformatics team using RefSeq genomes (bacteria, viral, archaea, fungi, human GRCh38) (available from 11.01.2019), and a custom database manual [341] (see Github link for details: Manual · DerrickWood/kraken2 Wiki · GitHub). I then ran the Kraken2 output report through Bracken (V.2.2) [344] (default settings used and species level selected) to obtain taxonomic profiles of microbiota. I plotted the taxonomic profiles to show the relative abundance (%) and number of reads of the ten most abundant species (for simplified visualisation) either within each individual or across the sample set. This identified common species within each individual and across the whole population to allow easier comparisons of the abundance of major organisms.

2.4.8 Assembly of Metagenome Assembled Genomes

The MAGs were assembled by Dr. Raymond Kiu. The shotgun metagenome raw reads (Fastq.gz) were trimmed, adaptor-removed and quality-filtered using Fastp

(V.0.20.0) (-q 20) [342]. Subsequently, host-associated sequences were removed via Kneaddata (V.0.10.0) [345] with human genome (GRCh38.p13 - no ALT) bowtie2 index file retrieved from https://benlangmead.github.io/aws-indexes/bowtie (--bypass-trim and –reorder) to generate clean FastQ reads. These metagenome reads were then co-assembled using the HPC. MEGAHIT (V.1.2.9) [346] was used prior to extraction of MAGs to produce a collection of output assemblies per sample. The Metagenomic Wrapper (MetaWRAP) (V.1.3.2) pipeline [347] was then used to extract MAGs based upon the metagenome assemblies generated. Metagenome clean reads were produced via three binning software packages: Metagenome Binning with Abundance and Tetra-nucleotide frequencies (MetaBAT) (V.2.12.1) [348], MAXBIN2 (V.2.2.6) [349] and Clustering cONtigs with COverage and ComposiTion (CONCOCT) (V.1.1.0) [350], using sub-module metaWRAP binning. The MAGs were further refined using sub-module metaWRAP bin_refinement to select the high-quality bins from each sample with completeness >80% and contamination <10% via CheckM (V.1.1.3) [351]. Finally, MAGs were taxonomically ranked using the Genome Database Taxonomy Toolkit (GTDB-Tk) (V.2.1.0) via module GTDB-Tk classify_wf [352].

2.4.9 Data Visualisation and Statistical Testing

I used R (V.4.1.2) (R Core Team, 2021) and ggplot2 (Wickham, 2016) to plot data and perform statistics. I used GraphPad Prism (V.5.04) (GraphPad Software, 2021) to generate jitter plots and perform Unpaired *t tests* (significance level ≤ 0.05). Access to Bioinformatics scripts is restricted because they hold identifiable information embedded within the scripts. Full statistical test reports (containing actual P values and coefficients) can be viewed in Github: https://github.com/ilianaserghiou/Iliana-Rosa-Serghiou-Thesis-2024

2.5 Chapter 4

2.5.1 Illumina Library Preparation for Shotgun Metagenomic Sequencing

Sequencing libraries were prepared by the QIB sequencing team for Illumina sequencing of the skin, vaginal, stool and culture samples. Libraries were prepared for each sample type separately to reduce the chance of contamination occurring from different sources that varied in biomass and extraction efficiency. Libraries were sent externally to Genewiz for SMS on the Illumina iSeq100 platform.

For Illumina library preparation, the tagmentation-based library construction using the Illumina DNA Prep Kit (Illumina Catalogue No. 20018704) was used. Where possible, samples were normalised to 5 ng/µl with PCR-grade water. 20% of skin swabs, 75% of skin cultures, 96% of vaginal swabs and 90% of vaginal cultures were normalised. The remaining swab and culture DNA concentrations were lower than 5 ng/µl and were used at their existing concentration. A bead tagmentation of metagenomic DNA was then performed whereby transposomes are randomly inserted into the genomic DNA adding Illumina adapters enabling barcoding and amplification of the libraries for multiplexing on an Illumina run. This tagmentation was a reduced reaction (20-fold dilution) and was prepared by mixing 0.5 µl of Tagmentation buffer (TB1), 0.5 µl Bead Linked Transposomes (BLT) (Illumina Catalogue No. 20018704) and 4 µl PCR grade water. This master mix (5 µl) was then added to a 96-well plate. Normalised DNA (10 ng) was pipette mixed with the tagmentation mix and heated (55 ºC, 15 min) in a PCR block. For samples containing less than 5 ng/µl, a separate tagmentation master mix was used with no water added and instead 6 µl input DNA volume. Tagmented samples were barcoded by setting up the following PCR master mix: 10 µl KAPA 2G Fast Hot Start Ready Mix (Merck Catalogue No. KK5601) and 2 µl PCR grade water per sample. The master mix was added to 1 µl of 10 µM custom made primer stock containing both P7 and P5 Illumina 9 bp barcodes (Illumina Catalogue No. FC-131-2001 to 2004). Finally, the 7 µl of tagmentation mix was added and mixed. The PCR was run using the following protocol: 72 ºC 3 min, 95ºC 1 min, 18 cycles of (95ºC 10 s, 55 ºC 20 s, 72 ºC 3 min).

Individual libraries were quantified using the Promega QuantiFluor® dsDNA System (Catalogue No. E2670) and run on a GloMax® Discover Microplate Reader. They were also run on the Agilent Tapestation 4200 instrument using a D5000 screentape (Agilent Catalogue No. 5067-5579) to check the library size distribution. Libraries were pooled based on DNA concentration. The pooled library was double-SPRI (Solid Phase Reversible Immobilization) size selected between 0.5 and 0.7X bead volumes using sample purification beads supplied in the Illumina DNA prep kit (Illumina® DNA Prep, (M) Tagmentation (96 Samples, IPB), 20060059). The final Illumina pooled library was then quantified using the Qubit 3.0 fluorometer and run on the Agilent 4200 Tapestation instrument. Both the concentration and average insert size were used to estimate the molarity of the final pool. The final pool was denatured and diluted to 1.5 pM. Libraries were then sent to Genewiz for SMS. A

target of 8 Gb data per sample was established, and 8-11 Gb data were obtained per sample. Conversion of the raw data into Fast Quality (FastQ) file format was performed by the Bioinformatics team at QIB.

2.5.2 Data Clean Up

The QIB bioinformatics team uploaded all PEARL data to the High Performance Computing (HPC), where they cleaned the data by removing human reads using their clean up pipeline, which used the tool Sequence Fastx Utilities (seqfu) (V.0.23.2) [353] to obtain read counts for each sample. Samples that obtained < 1000 reads were not used as this threshold defines the samples that are not biologically relevant for analysis [353]. Host removal was then performed using the Kraken2 tool (V.2.1.0) [340] to map human (classified) reads to a host contaminant database [354], built upon a viral sequence masked human reference database (Virus (Sars-Cov-2 and PhiX) + Sequence Masked Human Reference Genome (hg19)) [355]; Contaminant sequences (Sars-Cov-2, PhiX) were also removed using the same reference database and Burrows-Wheeler Aligner (BWA) mapping tool (V.0.7.17-r1188). Adapter filtering was performed using the tool Fast Pre-processor (Fastp) (v.0.23.2) [342] (quality filtering setting was disabled). Fast profiling using the Kraken2 tool (V.2.1.0) [340] and microbial database (gutcheck-db) [354] was used to evaluate the fluctuations in the microbial (unclassified) reads. Multi-Quality Control (MultiQC) was run to identify samples that passed the Quality Control (QC), and also provided general statistics (percentage (%) of classified (host), raw sequence and cleaned reads after host removal) (**Appendix 11.**). See GitHub links for the clean up pipeline and further details on databases used:

https://github.com/telatin/cleanup

https://github.com/telatin/cleanup/blob/main/md/databases.md

https://zenodo.org/records/7050266

2.5.3 Taxonomic Profiling

I generated the microbial taxonomic profiles to genus and species level using the HPC and the optimised PEARL pipeline. The pipeline used Kraken2 (V.2.1.2) [340] and Bracken (V.2.8) [344] tools. I mapped the microbial reads to a microbial metagenomic database (StandardDB) (containing RefSeq archaea, bacteria, viral, plasmid, human, UniVec_Core genomes; available from: https://benlangmead.github.io/aws-indexes/k2) in Kraken2 (V.2.1.2) [340]

(confidence level set to 0.1). I then ran the Kraken2 output report through Bracken (V.2.8) [344] (default settings used, genus and species levels selected) to obtain taxonomic profiles. Human reads may still be detected at low levels (Kraken2 database consists of human genomes), therefore I filtered any remaining *Homo sapiens* reads from the Bracken output using the KrakenTools tool (V.1.1) [356]. I combined all filtered Bracken outputs for each sample into a single output. I plotted the taxonomic profiles to show the relative abundance percentage and number of reads of the most abundant genera and species either within each individual or across the sample set. This identified common species within each individual and across the whole population, to allow easier comparisons of the abundance of major organisms. Due to the magnitude of samples and data generated, when visualising the profiles, the five most abundant genera and species were identified within each sample as this reduced complexity to enable simplified visualisation.

Due to the lack of knowledge on the infant skin archaeome in the microbiome research field, and the absence of archaea in the skin taxonomic profiling results, a specialised archaeal taxonomic profiling pipeline was run to species level on the infant and maternal skin datasets by Dr. Charlotte Neumann from the Med Uni Graz in Austria. ATLAS (V.2.18.0) [357] was run initially for an extra QC step to ensure all human and contaminant reads were removed; human reads were mapped to a host human database (chm13v2.0.fa) and a contaminant database (phix174_virus.fa). Kraken2 (V.2.1.2) [340] and Bracken (V.2.7) [344] tools were then run on the QC files. Kraken2 (V.2.1.2) [340] mapped microbial reads to a microbial database (Unified Human Gastrointestinal Genome collection [UHGG]), which implemented archaeal species from the genome catalogue published by Chibani et al. [358]. The Kraken2 output report was run through Bracken (V.2.7) [344] to obtain taxonomic profiles.

Strainphlan was used to perform metagenomic strain-level population genomics on this large metagenomic dataset by profiling microbes from known species with strain level resolution and providing comparative and phylogenetic analysis of strains retrieved from metagenomic samples. Strainphlan took a set of metagenomic samples and reference genomes and reconstructed the sample-specific strains of all species present in the sample and the genome-specific strains from the reference genomes. Profiling for strain level detection and transmission analysis was first performed by Dr. Alise Ponsero from the QIB core bioinformatics team using the HPC. The already quality controlled samples were annotated by Metagenomic

Phylogenetic Analysis (MetaPhlAn4) (V.4.0.1) against the Jun23_CHOCOPhlAnSGB_202307 database to identify whether the samples contained enough marker genes to be used in the strain analysis and to perform profiling of the global community composition. Samples that were too shallow in terms of sequencing depth (i.e., did not appropriately cover any marker genes) were removed from the strain level analysis. These samples that were removed included 4 metagenomes with a sequencing depth between 79252 and 524690 reads. Notably, appropriate coverage for relevant marker genes could be detected in samples with a minimum of 20205 reads, in particular the culture and meconium samples, in which the microbial community is composed of few organisms. Where several metagenomes were available for one sample, the replicate with the highest number of reads was kept. Multiple samples were generated for only 10 samples in the dataset for internal control purposes and protocol development. While pooling metagenomes may have allowed for an increased microbial coverage, it would have risked introducing methodological inconsistencies. Having different effective sequencing depths between pooled and non-pooled samples would have risked introduction of systematic biases in the downstream analyses. It was therefore decided to maintain consistent methodology across all samples rather than improving the sequencing depth of these few samples.

This also identified the species and clades that will be used for the strain level analysis. Core genera of interest (*Streptococcus*, *Micrococcus*, *Staphylococcus* and *Bifidobacterium*) were identified first from the taxonomic profiling analysis, which allowed the species and clades to be identified for the strain level analysis. Sample types (i.e. infant skin swab, infant skin culture, maternal skin swab, maternal skin culture, infant stool, maternal stool, maternal vaginal swab, and maternal vaginal culture) that did not identify presence of species and clades relating to the core genera of interest were not included in the strain level analysis.

Strain detection and threshold for strain transmission events was determined using strain level profiling with Strain Phylogenetic Analysis (StrainPhlAn4) (V.4.1.1), against the strainphlan-vJun23 database [359]. The ChocoPhlAn species markers were extracted using the sample2markers.py and extract_markers.py scripts and the strainphlan.py scripts from StrainPhlAn [360]. The markers were used to build multiple sequence alignments using the following parameters: --markers_in_n_samples 1 --sample_with_n_markers 10 --phylophan_mode accurate. Pairwise phylogenetic distances between samples were used to assess

strain transmission events as defined previously [361]. SGB-specific Normalized Phylogenetic Distance (nGD) was calculated as leaf-to-leaf branch lengths normalized by total tree branch length in phylogenetic trees produced by StrainPhlAn4 (V.4.1.1); which were built on marker gene alignments on positions with at least 1% variability. The precomputed strain identity thresholds defined previously [361] were used with the nGD corresponding to the 3rd percentile of the unrelated individual nGD distribution. For SGBs detected in at least 50 pairs of same-individual samples, nGD thresholds were defined based on maximizing Youden's index [362]. For SGBs detected in fewer than 50 same-individual close pairs, the nGD corresponding to the 3rd percentile of the unrelated individual nGD distribution was used. Clades detected in only one family pair were excluded from the analysis due to the impossibility to reliably compute the appropriate thresholds. Gtree (V.3.10.1) was used to identify whether strain transmission events occurred.

Based on the strain transmission events identified, changes in species transmissibility over time was assessed between different maternal body sites (skin, stool and vagina) and infant body sites (skin and stool). Changes in species transmissibility according to specific factors (delivery mode, pregnancy term and feeding mode) were assessed between different maternal body sites (skin, stool and vagina) and infant body sites (skin and stool). To avoid confusion with potential strain persistency effect, when the species or strain was shared between maternal-infant sample pairs in multiple time points, the transmission events were counted only once. Finally, strains shared between any maternal-infant sample pairs were assessed for their presence in the culture samples to determine cultivability.

2.5.4 Assembly of Metagenome Assembled Genomes

I co-assembled the metagenome reads using the HPC and the PEARL pipeline. I used MEGAHIT (V.1.2.9) [346] prior to extraction of MAGs to produce a collection of output assemblies per sample. I then used the Metagenomic Wrapper (MetaWRAP) (V.1.3.2) pipeline [347] to extract MAGs based upon the metagenome assemblies generated. I generated metagenome clean reads via three binning software packages: Metagenome Binning with Abundance and Tetra-nucleotide frequencies (MetaBAT) (V.2.12.1) [348], MAXBIN2 (V.2.2.6) [349] and Clustering cONtigs with COverage and ComposiTion (CONCOCT) (V.1.1.0) [350], using sub-module metaWRAP binning. I further refined the MAGs using sub-module metaWRAP bin_refinement to select the high-quality bins from each sample with completeness >80% and contamination <10% via CheckM (V.1.1.3) [351]. Finally, I

taxonomically ranked the MAGs using the Genome Database Taxonomy Toolkit (GTDB-Tk) (V.2.1.0) via module GTDB-Tk classify_wf [352].

2.5.5 Data Visualisation and Statistical Testing

I used R (V.4.2.2) (R Core Team, 2021) and ggplot2 (Wickham, 2016) to plot the data and perform statistics. Access to Bioinformatics scripts is restricted because they hold identifiable information embedded within the scripts. Full statistical test reports (containing actual P values and coefficients) can be viewed in Github: https://github.com/ilianaserghiou/Iliana-Rosa-Serghiou-Thesis-2024

CHAPTER 3

METHOD DEVELOPMENT FOR HIGH MOLECULAR WEIGHT BACTERIAL DNA EXTRACTION SUITABLE FOR SHOTGUN METAGENOMICS FROM SKIN SWABS

CHAPTER 3 OVERVIEW

In this chapter, optimisation of a skin specific high-throughput bacterial DNA extraction method is presented, resulting in DNA yield and molecular weight suitable for both short and long read Shotgun Metagenomic Sequencing (SMS) methods. This method optimisation was necessary due to limited commercialised kit protocols available that produce High Molecular Weight (HMW) DNA from skin in sufficient quantities for SMS. Culture and DNA extraction-based experiments were used to optimise the initial diluent and lysis procedure. Initial experiments included measurement of DNA yield recovered from swabs spiked with bacterial colonies that were processed in Phosphate-Buffered Saline (PBS) or Ultrapure Milli-Q Water (UMQW) as diluents, and measure of cell viability in both diluents. Lysis procedure experiments included quantification of DNA yield and molecular weight extracted from overnight liquid cultures and swabs (spiked with bacterial colonies), processed in the chosen diluent, under six different lysis variations. The finalised method was validated on adult volunteer and the Pregnancy and Early Life (PEARL) study cohort skin swabs and sequenced using SMS on short (Illumina) and long (Oxford Nanopore Technologies (ONT)) read platforms. Sequence data was bioinformatically analysed to produce taxonomic profiles and investigated skin composition, platform efficiency (in terms of composition and AT and GC content differences), and species and MAGs recovery from samples sequenced at 1 Gb to 5 Gb depth. I performed all laboratory experiments and bioinformatics analysis (including visualisations and statistical tests) in this chapter, except for the SMS library preparation and sequencing (performed by the QIB sequencing team) and the assembly of the Metagenome Assembled Genomes (MAGs) (performed by Dr. Raymond Kiu). This optimised method was the basis for all the subsequent data that I generated (Chapter 4) to explore the infant skin microbiome using the PEARL study cohort.

Data in this chapter has been previously published:

Serghiou, I. R, Baker, D., Evans, R., Dalby, M. J., Kiu, R., Trampari, E., Phillips, S., Watt, R., Atkinson, T., Murphy, B., Hall, L. J., & Webber, M. A. (2023). An efficient method for high molecular weight bacterial DNA extraction suitable for shotgun metagenomics from skin swabs. Microbial Genomics,9(7).

https://doi.org/10.1099/mgen.0.001058

3.1 Introduction

Both targeted 16S rRNA gene amplicon sequencing and SMS methods are commonly used to study the skin microbiome [13, 19, 288]. The 16S rRNA gene amplicon sequencing method typically uses Illumina sequencing technology, which results in short reads for taxonomic classification to genus level; though some variations of this method can identify species level [19, 311]. This targeted sequencing method is limited to PCR amplification of targeted regions in bacteria and archaea (Prokaryotes) where the 16S rRNA gene is found; whereas SMS sequences all genetic material, without being limited to targeted regions [13, 19, 304, 312]. Therefore, in SMS selective amplification efficiency bias is reduced, whole genomes are taxonomically classified to strain level, and microbiome functions can be identified [13, 19, 288, 317]. SMS uses Illumina, ONT and PacBio SMRT technologies for taxonomic classification [19, 311, 318]. Compared to Illumina, the ONT and PacBio SMRT technologies can produce long reads as they directly sequence HMW DNA [19, 311, 328, 363]. In contrast, Illumina technology requires fragmentation of HMW DNA, resulting in LMW DNA being sequenced [19, 319]. Compared to Illumina, ONT and PacBio SMRT can generate more complete genomes and high-quality MAGs, used for higher taxonomic resolution and functional information [19, 311, 317, 327].

Skin has a relatively low bacterial biomass, challenging the extraction of sufficient DNA quantities for SMS [19, 331, 364]; especially for long read technologies where more input material is needed and sequencing is direct rather than via amplification [19, 363]. There are limited commercialised kit protocols available that produce HMW DNA from low biomass samples, such as skin, in sufficient quantities for SMS. To address this, an optimised high-throughput automated DNA extraction method for recovery of microbial DNA from skin is required.

3.2 Aims

Work presented in this chapter aimed to complete Objective 1 (section **1.7**) of this thesis, to develop an optimised high-throughput automated DNA extraction method for recovery of microbial DNA from skin swabs; and a culture collection alongside. Yielding sufficient DNA yield and molecular weight suitable for SMS was a key part of this objective.

3.3 Objectives

1. Compare the effectiveness of 1X PBS and Milli-Q water in recovering extracted bacterial DNA, to determine an effective initial diluent for recovering skin microbial cells for use in DNA extraction.
2. Compare bacterial cell viability from cultures processed and grown in 1X PBS and Milli-Q water, to determine an effective initial diluent for recovering viable skin microbial cells for culturing alongside DNA extraction.
3. Compare different bacterial DNA extraction lysis methods, using the automated Maxwell® RSC 48 Instrument and the Maxwell® RSC Blood DNA Kit, to identify a method that results in sufficient DNA yield and molecular weight suitable for SMS, and effectively extracts both Gram-positive and Gram-negative bacterial DNA.
4. Validate the optimised protocol was compatible with PEARL swab samples using adult volunteer and baby skin swabs. The resulting DNA extracts were sequenced using SMS on Illumina and ONT platforms to verify a microbial community profile could be identified.

3.4 Experimental Methods

To optimise extraction of microbial DNA from skin swabs, I evaluated different initial diluents and lysis procedures for their effectiveness to recover DNA yield and molecular weight suitable for short and long read SMS. After I performed the dilution and lysis, I heated the samples following the RSC Blood DNA Kit protocol and loaded them onto the Promega Maxwell RSC 48 instrument for the automated extraction (**Figure 3.1**). Once I tested all variables, I finalised a method and validated it using adult volunteer and PEARL study infant skin swabs. I obtained FMH Ethical approval, which was approved to test this method on volunteer skin swabs (section **2.1.1**).

I chose to use the Promega Maxwell RSC 48 instrument and RSC Blood DNA Kit as they were already in use within the PEARL study, and their protocols (**Appendix 12.**) were therefore optimised in this study. The RSC Blood DNA kit that was optimised, has previously been shown to produce HMW DNA [365, 366], and has a higher binding capacity and cleaner eluate than traditional silica-based DNA purification systems [367-370]. The Promega Maxwell RSC 48 instrument is also suitable for high-throughput automated genomic DNA isolation capable of

processing 48 samples in 40 min [370], making it compatible with larger sample sets, such as the PEARL study.

The focus of this work was to develop an efficient, optimised, and validated lysis method, with an appropriate initial diluent, for recovering DNA from skin microbiota. I developed this protocol using a Promega Maxwell RSC 48 instrument, but the method should be applicable to any other automated and column-based kits on the market. The aim was not to benchmark extraction kits, which was beyond the scope of producing this protocol.

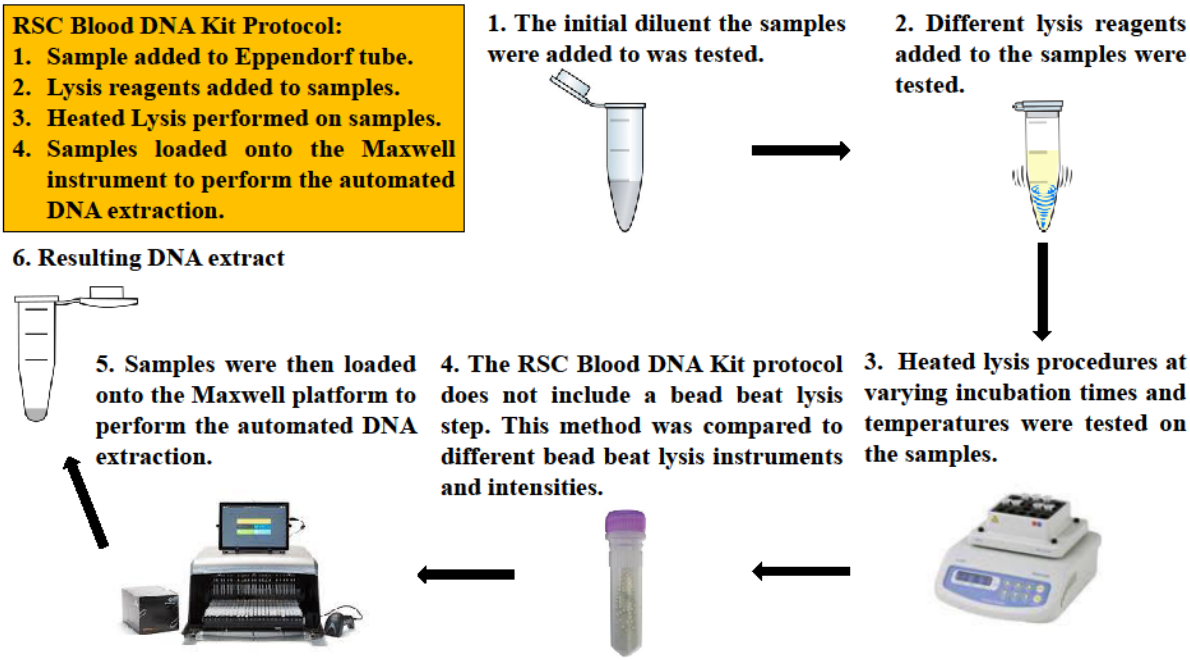


Figure 3.1. RSC Blood DNA Kit Protocol (Yellow Box) and Modifications to Test Initial Diluents and Lysis Procedures. The yellow box presents the RSC Blood DNA kit protocol. Steps 1-6 outside of the yellow box identify the modifications included at each step, and the modifications introduced as an extra step (step 4). Extracted from Serghiou et al. [19]

3.4.1 Initial Diluent Testing Experiments

A protocol where a skin swab could be processed allowing both DNA extraction and, in parallel, culture of organisms was desired. It was therefore necessary to remove material from the swab into a diluent, with 1x PBS and UMQW, which I selected for initial trials. I compared their impact on bacterial viability and ability to extract DNA by measuring extracted bacterial DNA recovered from swabs spiked with bacterial colonies, and testing recovered bacterial cell viability through recovery from liquid and plate cultures.

I selected this initial diluent variable due to a pilot study, conducted by the QIB PEARL team, for optimisation of a skin specific microbial DNA extraction method. Initial diluents (UMQW, Lysis Buffer, Proteinase K, MT Buffer, and Sodium Phosphate) were tested, and DNA recovery compared between the automated Promega Maxwell RSC 48 instrument with the RSC Blood DNA kit (**Appendix 12.**) and the column-based FastPrep MPBio kit (QIB established method – **Appendix 13.**). UMQW was the only diluent from which DNA could be recovered (**Table 3.1**) and therefore taken forward. PBS was also selected, as previous research has shown the suitability of both diluents in long-term preservation of viable bacteria [371-378], therefore may be effective with protocol alterations. Both kit protocols tested in the pilot study recovered similar DNA concentrations, and have previously been shown to recover a good representation of mock communities [367], justifying their continued use.

Table 3.1. Pilot Study DNA Yield

ID	Platform & Protocol	Volume (ml) & Initial Diluent	DNA Concentration (ng/µl)
1	Maxwell	1 UMQW	13.4
2	Maxwell	0.5 UMQW	18.4
3	Maxwell	0.99 Lysis Buffer + Proteinase K	Too Low
4	Maxwell	0.33 Lysis Buffer + Proteinase K	Too Low
5	FastPrep	1 UMQW	11.8
6	FastPrep	0.5 UMQW	Too Low
7	FastPrep	1 UMQW	Too Low
8	FastPrep	0.5 UMQW	12.2
9	FastPrep	0.98 Sodium Phosphate + 0.12 MT Buffer (PBS)	Too Low
10	FastPrep	0.49 Sodium Phosphate + 0.06 MT Buffer (PBS)	Too Low

Method for Measuring Extracted Bacterial Quantities

I spiked fourty-four sterile charcoal amies agar cotton swabs with a single colony collected from an agar plate containing *S. aureus* NCTC 8532 (*S. aureus*) (section **2.4.1; 2.4.2**). I included two negative controls containing only 1x PBS or UMQW. I removed the spiked swab heads into 1.5 ml Microcentrifuge tubes containing 1 ml

of either PBS (n=22) or UMQW (n=22). I then extracted DNA from these samples following the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit protocol (**Appendix 12.**), modified to adapt to the use of swabs. The swab heads were present throughout the entirety of the extraction procedure. Modifications included the initial vortex of swabs at full speed for 2 minutes, followed by centrifugation (14,000 × *g*, 15 min) to pellet the cells. I then removed the supernatant and resuspended cells in 300 µl of 1x PBS or UMQW. Steps 4 and 6-8 of the RSC protocol were then followed (**Appendix 12.**). I next performed a bead beating step using the MP Biomedicals FastPrep instrument (3 min, 6.0 m/s), and centrifuged the samples again as above. I loaded the sample supernatants onto the Maxwell instrument and started the extraction following steps 9-21 of the RSC protocol. I performed a Qubit assay to quantify the resulting DNA (section **2.3**).

Method for Measuring Cell Viability

I grew overnight liquid cultures from isolates of three species (*S. aureus* NCTC 8532, *P. aeruginosa* PA14 and *E. coli* EC18PR-0166–1, a food isolate of ST10), with three replicates for each (section **2.4.1; 2.4.2**). For each isolate replicate, I transferred 1 ml into x3 15 ml falcon tubes, and pelleted cells by centrifugation (14,000 × *g*, 15 min). I resuspended each isolate replicate in 200 µl of Luria Bertani (LB), 1 x PBS or UMQW and incubated for 1 h at ambient temperature (to account for future sample processing time of skin swabs), with LB acting as a control. I used three biological replicates for each diluent, and nine for each isolate. I performed serial dilutions (10^{-1} to 10^{-8}) for each resuspended replicate, and plated onto drug free LB agar using the Miles and Misra spot method [379]. I used a set of 18 LB agar plates to plate dilutions for each diluent (x6 plates for LB, x6 plates for PBS and x6 plates for UMQW) per isolate replicate; I spotted each dilution three times onto plates that were split into quadrants. After an overnight incubation, I counted the number of colonies formed as a proxy for viable bacterial cells, then calculated the CFU/ml. In total, I recorded nine data points for each diluent per isolate replicate, totalling 27 condition suspensions assessed per isolate replicate. Negative control liquid and plate cultures were included in this experiment. This experiment is summarised in **Figure 3.2**.

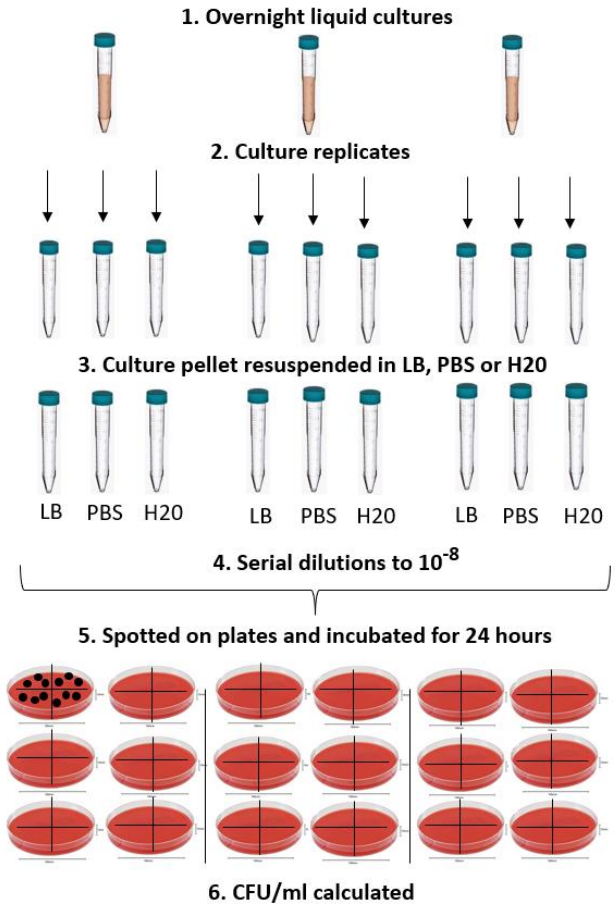


Figure 3.2. Workflow Diagram for Measuring Cell Viability. This figure shows how *E. coli*, *P. aeruginosa* and *S. aureus* isolates were cultured and processed and how serial dilutions were performed and plated. The diagram represents one species isolate. For each species isolate overnight liquid cultures were grown, with three replicates for each (step 1). For each isolate replicate, 1 ml was transferred into x3 15 ml falcon tubes (Step 2), and cells pelleted by centrifugation (14,000 × g, 15 min). Each isolate replicate was resuspended in 200 µl of LB, 1 x PBS or UMQW (step 3) and incubated for 1 h at ambient temperature, with LB acting as a control. Three biological replicates were performed for each diluent, and nine for each isolate. Serial dilutions (10^{-1} to 10^{-8}) were performed for each resuspended replicate (step 4) and plated onto drug free LB agar using the Miles and Misra spot method (step 5); a set of 18 plates were used to plate dilutions for each diluent (x6 plates for LB, x6 plates for PBS and x6 plates for UMQW) per isolate replicate. Each dilution was spotted three times onto plates that were split into quadrants. After an overnight incubation, the number of colonies formed were counted as a proxy for viable bacterial cells, then the CFU/ml was calculated (step 6). Nine data points were recorded for each diluent per isolate replicate, totalling 27 condition suspensions assessed per isolate replicate.

3.4.2 Lysis Method Experiments

I compared six lysis variations to identify the best method for high yields of HMW DNA from both Gram-negative and Gram-positive bacteria. For each method, factors from common lysis methods used in commercial kits for research were varied – heat, chemical, enzymatic, and mechanical [380, 381]. **Table 3.2** lists the differences between the six methods.

Table 3.2. Comparison of Extraction Methods. Extracted from Serghiou et al., [19]

	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6
Heated Lysis Step	Yes	Yes	Yes	Yes	Yes	Yes
Time	20 min	20 min	18 h	18 h	18 h	18 h
Temperature	56 °C	56 °C	37 °C	37 °C	37 °C	37 °C
Reagents	Proteinase K + Lysis Buffer	Proteinase K + Lysis Buffer	Epicentre ready-lyse lysozyme	Epicentre ready-lyse lysozyme	Thermo Fisher lysozyme	Thermo Fisher lysozyme
Agitation	No	No	300 rpm	300 rpm	300 rpm	300 rpm
Bead Beat Step	Yes	Yes	Yes	Yes	Yes	Yes
Instrument	FastPrep	Tissue Lyser	FastPrep	Tissue Lyser	FastPrep	Tissue Lyser
Settings	3 min 6.0 m/s	3 min 20 Hz	3 min 6.0 m/s	3 min 20 Hz	3 min 6.0 m/s	3 min 20 Hz
Heated Offboard Lysis Step	No	No	Yes	Yes	Yes	Yes
Temperature	N/A	N/A	68 °C	68 °C	68 °C	68 °C
Time	N/A	N/A	15 min	15 min	15 min	15 min
Reagents	N/A	N/A	Proteinase K + buffer ATL & ACL + carrier RNA	Proteinase K + buffer ATL & ACL + carrier RNA	Proteinase K + buffer ATL & ACL + carrier RNA	Proteinase K + buffer ATL & ACL + carrier RNA
Agitation	N/A	N/A	300 rpm	300 rpm	300 rpm	300 rpm

The gold standard bacterial lysis method is the mechanical approach to release genomic material, which results in an unbiased lysis of Gram-positive and Gram-negative bacteria. Other approaches such as enzymatic lysis have shown bias when used alone [382, 383]. Gram-positive bacteria are harder to lyse than Gram-negative bacteria due to their thicker and more rigid cell walls [384, 385]. Therefore, Gram-positive resident skin microbiota, such as *Staphylococcus* species [2, 15], will require a stronger lysis method using a high velocity mechanical approach [382]. Studies have also shown the importance of an enzyme lysis approach for the lysis of Gram-positive bacteria as it helps breakdown the thick peptidoglycan layer of the Gram-positive cell wall [386-389]. Thus, a combined method was explored.

Within the six lysis methods, I compared two high velocity bead beating instruments available, MP Bio FastPrep and QIAGEN Tissue Lyser. I selected the MP Bio Lysing Matrix E tubes to use on both instruments due to their advertised compatibility. These tubes also allow efficient lysis of diverse human and environmental sample types due the range of bead sizes contained in the matrix (1.4 mm ceramic spheres, 0.1 mm silica spheres and one 4 mm glass bead) [367]. I also compared an overnight heated lysis and off-board lysis method (using the Epicentre Ready-Lyse lysozyme), with the Promega Maxwell lysis protocol. Furthermore, I compared the effectiveness between the Epicentre Ready-Lyse lysozyme with a more cost-effective lysozyme.

I tested all six lysis methods on overnight liquid cultures. I validated the best method on sterile swab heads inoculated with harvested bacteria from overnight plate cultures. I used 1 x PBS as the initial diluent for all DNA extractions based on initial diluent testing experiment results (section **3.6.1**), which I completed on the automated Maxwell® RSC 48 Instrument.

Method for DNA Extraction on Overnight Liquid Cultures

I grew duplicate 10 ml overnight liquid cultures from isolates of three species (*S. aureus* NCTC 8532, *P. aeruginosa* PA14 and *E. coli* EC18PR-0166–1 (section **2.4.1; 2.4.2**). I tested methods 1 and 2 using 300 µl overnight culture, and methods 3-6 using 400 µl. I performed all trials in biological duplicate for all three tested strains. I started extractions following the Promega Maxwell® RSC 48 instrument and RSC Blood DNA Kit protocol (**Appendix 12.**) with the listed changes to the lysis procedure for each of the six methods tested (**Table 3.2**). I first vortexed all samples (full speed, 2 min) and centrifuged (14000 ×*g*, 15 min) then to pellet the cells. I

removed the supernatants and resuspended the pellets in 300 µl (methods 1 or 2) or 400 µl (methods 3–6) of 1 x PBS.

For method 1 and 2 samples, I added 30 µl of Proteinase K and 300 µl of Lysis Buffer (both reagents included in the RSC Blood DNA Kit) to the 300 µl sample suspensions. I then incubated the samples in a heat block (56 °C, 20 min). For methods 3 and 4, I added 3 µl Epicentre Ready-Lyse lysozyme (250 U/µl in Tris-EDTA Buffer [Sigma-Aldrich]) to the 400 µl sample suspensions. For methods 5 and 6, I added 3 µl Thermo Fisher lysozyme (250 U/µl in Tris-EDTA Buffer [Sigma-Aldrich]) to the 400 µl sample suspensions. I then incubated samples from methods 3–6 with agitation (300 rpm., 37 °C, 18 h). I then performed a bead beating step on all samples. For methods 1, 3 and 5, I used the MP Biomedicals FastPrep instrument (3 min, 6.0 m/s). For methods 2, 4 and 6, I used the Qiagen Tissue Lyser instrument (3 min, 20Hz). This was to compare the impact of a less intense bead beating step. I also performed an off-board lysis step for methods 3–6, where I added 40 µl Proteinase K, 165 µl Buffer ATL, 120 µl Carrier RNA (1 µg/µl stock in Buffer AVE), and 315 µl Buffer ACL (all off-board lysis reagents from Qiagen) into the 400 µl sample suspensions. I incubated these offboard lysis samples (68 °C, 15 min).

I centrifuged the samples (14000 ×*g*, 15 min) to pellet cells, and I loaded the supernatants onto the Promega Maxwell® RSC 48 instrument (≈ 630µl; the total recommended loading volume of sample and reagents onto the instrument [370]). I started the extraction following steps 9–21 of the initial RSC Blood DNA Kit protocol. I included a cell-free, diluent-only sample as a negative control to ensure no contamination had occurred. I performed Qubit and Tapestation assays on the resulting DNA extracts (sections **2.3; 2.4.3**).

Method for DNA extraction on spiked swabs

Method 6 was found to be most effective (section **3.6.2**), therefore I carried this method forward for validation using swab samples spiked with colonies from overnight plate cultures using the same three isolates (*S. aureus* NCTC 8532, *P. aeruginosa* PA14 and *E. coli* EC18PR-0166–1) (section **2.4.1; 2.4.2**). In this experiment, I tested the effect of an additional Agarase step, as the swab preservation agar was suspected to have reduced DNA yield in the initial diluent DNA extraction optimisation.

For validation, I spiked 24 sterile charcoal cotton swabs with one colony from overnight plates of each of the three isolates, with eight independent swabs

processed per species. I added the swab heads into 1.5 ml microcentrifuge tubes containing 1 ml 1 x PBS. I processed the swabs following the method 6 protocol outlined above. During the overnight lysis step, I added 1 µl β-Agarase-I to half of the spiked swabs, resulting in four replicates per species, per condition. I included a cell-free, diluent-only sample as a negative control to ensure no contamination had occurred. I performed a Qubit assay on extracted DNA to measure concentration (section **2.3**).

3.4.3 DNA Extraction Validation with Adult Volunteer and PEARL Study Skin Swabs

I validated the finalised method 6, described in the previous section, using skin swabs from adult volunteer and PEARL study cohorts (section **2.1**). I used both Illumina and ONT platforms for SMS to validate sequencing effectiveness on short and long read SMS platforms (section **2.4.4; 2.4.5**). I used culturing in parallel with the skin swab DNA extraction to ensure culturable organisms were also represented in the SMS. I cultured swabs aerobically and anaerobically on Columbia Blood Agar (CBA + 5% horse blood - prepared by the QIB media kitchen), as this is a known non-selective media for culturing a diverse range of bacteria [390], and has been successful in other studies in obtaining a good representation of skin bacteria [391]. I verified the successful bacterial DNA extraction through investigating the relative skin taxonomic profiles for each swab and culture metagenomes obtained using SMS (section **2.4.7**). This will provide greater insights into community compositions of the skin microbiome using the optimised extraction method. I completed this validation experiment by investigating SMS platform bias in species identification and determined the depth of sequence data required to provide optimal phylogenetic resolution and MAGs.

Method

I performed optimised bacterial DNA extractions, with culturing in parallel, on 24 adult volunteer and 10 PEARL infant skin swabs that were collected at 4 months old (see **Table 3.3** for the PEARL study infant metadata).

Table 3.3. PEARL Study Infant Metadata. Extracted from Serghiou et al. [19]

Sample ID	Age	Birth Health Status	Birth Breast Feed	Feed at Week 3	Feed at Month 4
Baby 1	Month 4	Healthy	No	Missing Data	Missing Data
Baby 2	Month 4	Healthy	No	Missing Data	Formula
Baby 3	Month 4	Healthy	Yes	Breast	Breast
Baby 5 twin 1	Month 4	Healthy	No	Breast/formula	Breast/formula
Baby 5 twin 2	Month 4	Healthy	No	Breast/formula	Breast/formula
Baby 6	Month 4	Healthy	Yes	Breast	Breast
Baby 7	Month 4	Healthy	Yes	Breast	Breast
Baby 8	Month 4	Healthy	Yes	Breast/formula	Formula
Baby 9	Month 4	Healthy	No	Missing Data	Missing Data
Baby 10	Month 4	Healthy	Breast/formula	Breast/formula	Formula

For each skin swab, I grew cells on CBA aerobically and anaerobically and harvested colonies into one glycerol stock to investigate the skin swab culture metagenome. On each cohort extraction run, I included a negative control (cell-free, diluent-only sample) and a positive control (established commercial mock community - ATCC skin microbiome whole cell mix) [335]. I prepared dilutions of the positive control microbiome mix to validate extraction efficiency and identify the amount of starting material needed for SMS (section **2.2**). I used Qubit assays to quantify all skin swab and culture DNA extracts, and I used Tapestation assays on the volunteer DNA to confirm appropriate molecular weight DNA was extracted (sections **2.3; 2.4.3**).

All DNA extracts were sent to the QIB sequencing team for SMS using the Illumina and ONT platforms. A lower sequencing depth of 300 Mbp was sought, which was more cost effective for optimisation work (section **2.4.4; 2.4.5**). I bioinformatically analysed the SMS data to identify the taxonomic profiles of the samples, to assess successful harvesting of bacterial DNA and gain insight into the skin community composition (sections **2.4.6; 2.4.7; 2.4.9**). Some swabs did not yield DNA using the optimised method, which led me to further investigate using another ten extracted skin swabs from PEARL study infants. I used different overnight lysis incubation times; I randomly incubated samples for either 18, 20 or 22 hours. I then investigated platform bias through bioinformatic comparison of resulting SMS AT and GC content (%). I finalised this method optimisation through determination of sequence data depth required to provide optimal phylogenetic resolution and MAGs (section **2.4.8; 2.4.9**). This was determined through sequencing the PEARL study infant skin swab DNA at 1 Gb and 5 Gb sequencing depth, and I bioinformatically compared MAGs and species recovery at both sequence depths. I compared species recovery outcomes using the 5 Gb per sample dataset and subsampling down to 1 Gb of data.

3.5 Data Analysis

3.5.1 Initial Diluent Testing

I used jitter plots to display the extracted DNA concentrations and CFU/ml recovered from each isolate processed in different diluents. I performed statistical analysis using an unpaired *t test* (appropriate for comparing two independent groups), to determine if there was a significant difference in concentrations (ng/µl) between samples processed in both diluents, and the CFU/ml between PBS and UMQW (section **2.4.9**).

3.5.2 Lysis Method Experiments

I used jitter plots to display DNA yield and molecular weight (bp) produced from all six extraction methods, and DNA yield from spiked swabs extracted with and without Agarase. I performed statistical analysis using unpaired *t tests* to determine if there were significant differences: (i) in the yield and molecular weight of DNA recovered using each of the six DNA extraction methods, (ii) between DNA recovery from each species, regardless of the method used, (iii) in the yield between the swabs, and each species, extracted with and without Agarase (section **2.4.9**).

3.5.3 DNA Extraction Validation with Adult Volunteer and PEARL Study Skin Swabs

I cleaned all the SMS data and taxonomically classified the data using the Galaxy platform (V.19.05) [338]. The main taxonomic classification methods are mapping microbial reads to a reference database, and assembly and binning of reads [392]. Here, I used the mapping method for taxonomic classification/profiling using the commonly used Centrifuge tool (V.1.0.3) [392-394], using the default settings. However, the resulting Centrifuge taxonomic profiles exhibited consistent discrepancies, with the Relative Abundance (RA) not matching read numbers, replicate species and a low percentage or zero human reads present in the samples. Low or zero human reads is ideal, however isn't a realistic result for skin swabs. Due to these discrepancies, I ran the data through two other commonly used taxonomic classification tools, Kraken2 (V.2.1.1+galaxy0) [340] and Bracken (V.2.2) [392, 394, 395]; the Kraken2 output was run through Bracken. For Kraken2, I used a confidence level of 0.1 and the 'Kraken2_benlangmead_standard' database. I used all default settings for Bracken. The discrepancies were resolved, and I established a pipeline to clean and analyse the data to obtain taxonomic profiles (sections **2.4.6; 2.4.7**).

I visualised the taxonomic profiles of the skin microbiome as stacked bar plots. I then used alluvial plots to explore whether the Illumina and ONT platforms identified comparable percentage total counts of the common abundant species (>0.5%) using the infant swab dataset. The adult volunteer dataset was not used for this comparison, as data were not generated on both platforms for all samples. I used a box plot to investigate the AT and GC content differences of samples sequenced on the Illumina and ONT platforms. I used box and spaghetti (rarefaction) plots to investigate species and MAGs recovery from samples sequenced at 1 Gb to 5 Gb depth. Finally, I performed statistical analysis using unpaired *t tests* to determine whether significant differences were present between: (i) the DNA yield from swabs incubated at different overnight lysis times; (ii) the AT and GC content % of samples sequenced on the Illumina and ONT platforms; (iii) species and MAGs recovery from samples sequenced at 1 Gb-5 Gb of data (section **2.4.9**).

3.6 Results

3.6.1 Initial Diluent Testing Experiments

DNA was obtained from all samples with no significant difference in DNA yield from spiked swabs processed in either diluent (**Table 3.4**; **Figure 3.3**; P > 0.05 *[0.4014]*). *S. aureus*, *P aeruginosa* and *E. coli* were all recovered in high numbers and showed no significant differences in viable numbers (CFU/ml) recovered after suspension in either diluent. *P. aeruginosa* had the lowest CFU/ml and *E. coli* had the highest CFU/ml (**Table 3.5**; **Figures 3.4-3.5**, P > 0.05 *[0.5339]*).

Table 3.4. DNA Yield from Spiked Swabs

Sample ID	Diluent Used	Qubit (ng/µl)
1	1x PBS	0.801
2	1x PBS	1.18
3	1x PBS	2.97
4	1x PBS	0.56
5	1x PBS	1.67
6	1x PBS	0.662
7	1x PBS	1.22
8	1x PBS	1.19
9	1x PBS	2.09
10	1x PBS	1.99
11	1x PBS	0.462
12	1x PBS	0.839
13	1x PBS	3.04
14	1x PBS	2.17
15	1x PBS	2.22
16	1x PBS	2.15
17	1x PBS	2.45
18	1x PBS	2.24
19	1x PBS	3.47
20	1x PBS	1.81
21	1x PBS	1.88
22	1x PBS	2.56
Negative Control	1x PBS	0.055
23	Milli-Q Water	1.28
24	Milli-Q Water	0.975
25	Milli-Q Water	2.41
26	Milli-Q Water	4.41
27	Milli-Q Water	3.26
28	Milli-Q Water	2.64
29	Milli-Q Water	1.49
30	Milli-Q Water	3.7
31	Milli-Q Water	1.93
32	Milli-Q Water	3.02
33	Milli-Q Water	1.52
34	Milli-Q Water	2.21
35	Milli-Q Water	3.4
36	Milli-Q Water	1.92
37	Milli-Q Water	1.21
38	Milli-Q Water	1.39
39	Milli-Q Water	1.67
40	Milli-Q Water	1.41
41	Milli-Q Water	1.99
42	Milli-Q Water	0.487
43	Milli-Q Water	0.804
44	Milli-Q Water	1.36
Negative Control 2	Milli-Q Water	0.146

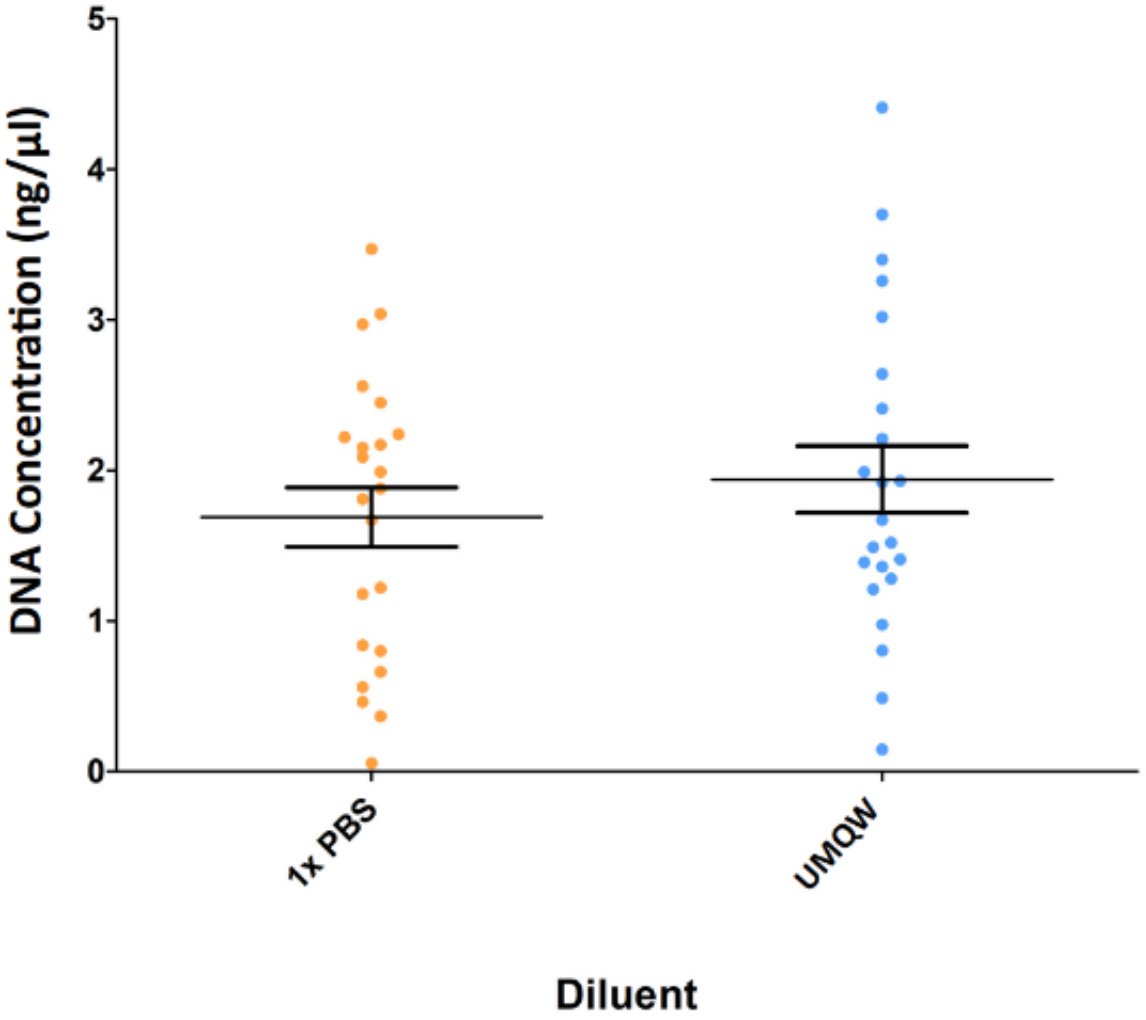


Figure 3.3. Total DNA Yield (ng/µl) of Spiked Swabs Processed in 1xPBS and UMQW. This figure identifies results for the comparison of the extracted bacterial quantities from swabs processed in 1x PBS and UMQW used as an initial diluent. Fouty four sterile charcoal amies agar cotton swabs were spiked with a single colony collected from an agar plate containing *S. aureus*. Half the spiked swabs were processed in 1x PBS (n=22) and half in UMQW (n=22). DNA was then extracted from these samples following the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit protocol, modified to adapt to the use of swabs. The swab heads were present throughout the entirety of the extraction procedure. Modifications included the initial vortex of swabs at full speed for 2 minutes, followed by centrifugation (14,000 × g, 15 min) to pellet the cells. The supernatant was then removed, and cells resuspended in 300 µl of 1x PBS or UMQW. Steps 4 and 6-8 of the RSC protocol were then followed. A bead beating step was performed using the MP Biomedicals FastPrep instrument (3 min, 6.0 m/s). Samples were centrifuged again as above. Sample supernatants were loaded onto the Maxwell instrument and the extraction started following steps 9-21 of the RSC protocol. A Qubit assay was used to quantify the resulting DNA. Horizontal bars on each plot show the means, vertical bars show the standard error of the mean (SEM). Unpaired *t tests* were performed to identify significant differences ($P < 0.05$) in DNA yield from spiked swabs processed in either diluent; no significant results identified. Extracted from Serghiou et al. [19].

Table 3.5. CFU/ml for Each Culture Condition Suspension Replicate

Condition Suspension Diluent	*E. coli* CFU/ml	*S. aureus* CFU/ml	*P. aeruginosa* CFU/ml
LB	195,000,000,000	60,000,000,000	85,000,000,000
LB	205,000,000,000	35,000,000,000	120,000,000,000
LB	170,000,000,000	75,000,000,000	115,000,000,000
LB	180,000,000,000	95,000,000,000	26,000,000,000
LB	190,000,000,000	105,000,000,000	29,500,000,000
LB	175,000,000,000	55,000,000,000	22,500,000,000
LB	220,000,000,000	95,000,000,000	20,500,000,000
LB	135,000,000,000	125,000,000,000	23,000,000,000
LB	135,000,000,000	115,000,000,000	29,500,000,000
1x PBS	145,000,000,000	70,000,000,000	30,500,000,000
1x PBS	15,000,000,000	70,000,000,000	27,500,000,000
1x PBS	140,000,000,000	35,000,000,000	27,000,000,000
1x PBS	185,000,000,000	70,000,000,000	24,000,000,000
1x PBS	175,000,000,000	55,000,000,000	21,500,000,000
1x PBS	160,000,000,000	125,000,000,000	13,000,000,000
1x PBS	315,000,000,000	95,000,000,000	22,000,000,000
1x PBS	245,000,000,000	205,000,000,000	23,000,000,000
1x PBS	230,000,000,000	165,000,000,000	28,500,000,000
Milli-Q Water	90,000,000,000	55,000,000,000	22,500,000,000
Milli-Q Water	95,000,000,000	85,000,000,000	30,000,000,000
Milli-Q Water	175,000,000,000	45,000,000,000	25,000,000,000
Milli-Q Water	120,000,000,000	110,000,000,000	23,000,000,000
Milli-Q Water	110,000,000,000	100,000,000,000	24,000,000,000
Milli-Q Water	115,000,000,000	120,000,000,000	19,000,000,000
Milli-Q Water	175,000,000,000	70,000,000,000	17,500,000,000
Mill-Q Water	140,000,000,000	70,000,000,000	26,000,000,000
Milli-Q Water	140,000,000,000	140,000,000,000	27,500,000,000
Average CFU/ml Per Isolate Per Suspension			
LB	178,333,333,333	84,444,444,444	52,333,333,333
1x PBS	178,888,888,889	98,888,888,889	24,111,111,111
Milli-Q Water	128,888,888,889	88,333,333,333	23,833,333,333
Average CFU/ml Per Isolate Log Fold Change			
1x PBS vs Milli-Q Water	50,000,000,000	10,555,555,556	277,777,778
Unpaired t *test* P Values Per Isolate			
1x PBS vs Milli-Q Water	0.3618	0.7599	0.9254
Unpaired *t test* P Value Across all Isolates			
1x PBS vs Milli-Q Water		0.5339	

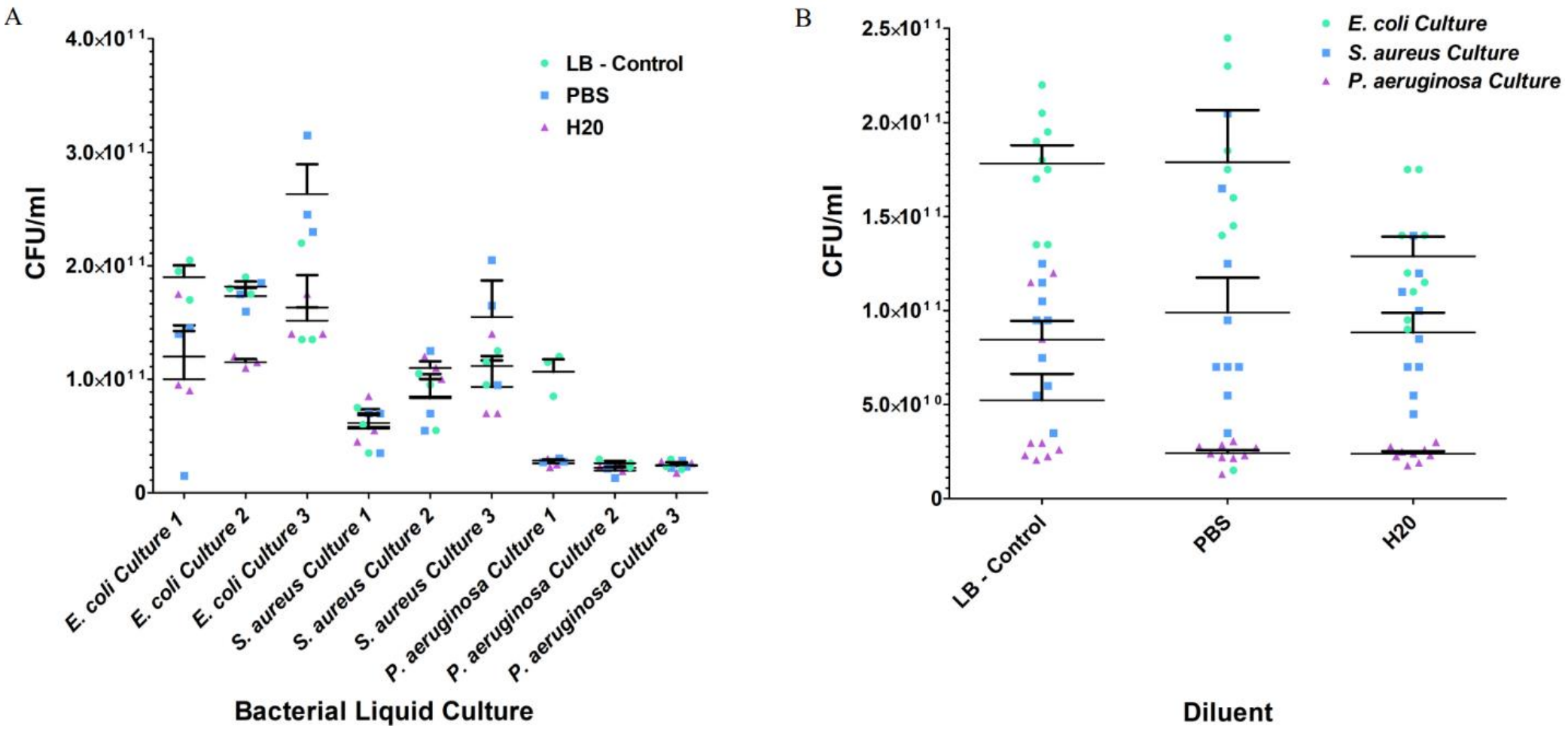

A
4.0×10^11
3.0×10^11
2.0×10^11
1.0×10^11
0
CFU/ml
LB - Control
PBS
H20
E. coli Culture 1
E. coli Culture 2
E. coli Culture 3
S. aureus Culture 1
S. aureus Culture 2
S. aureus Culture 3
P. aeruginosa Culture 1
P. aeruginosa Culture 2
P. aeruginosa Culture 3
Bacterial Liquid Culture
B
2.5×10^11
2.0×10^11
1.5×10^11
1.0×10^11
5.0×10^10
0
CFU/ml
E. coli Culture
S. aureus Culture
P. aeruginosa Culture
LB - Control
PBS
H20
Diluent

Figure 3.4. Bacterial Growth for *E. coli*, *S. aureus* and *P. aeruginosa* that were Processed and Grown in 1x PBS and UMQW. This figure identifies results for the comparison of cell viability from overnight liquid cultures that were grown for *E. coli*, *S. aureus* and *P. aeruginosa* in either 1x PBS or UMQW (and LB as a control), with three replicates for each isolate (n=9). For each isolate replicate, 1 ml was transferred into x3 15 ml falcon tubes, cells were pelleted by centrifugation and resuspended in 200 µl of LB, 1 x PBS or UMQW. These were incubated for 1 h at ambient temperature, with LB acting as a control. Three biological replicates were performed for each diluent, and nine for each isolate. Serial dilutions (10^{-1} to 10^{-8}) were performed for each resuspended replicate and plated onto drug free LB agar using the Miles and Misra spot method; a set of 18 plates were used to plate dilutions for each diluent (x6 plates for LB, x6 plates for PBS and x6 plates for UMQW) per isolate replicate. Each dilution was spotted three times onto plates that were split into quadrants. After an overnight incubation, the number of colonies formed were counted as a proxy for viable bacterial cells, then the CFU/ml was calculated. Nine data points were recorded for each diluent per isolate replicate, totalling 27 condition suspensions assessed per isolate replicate. **A** and **B:** CFU/ml recovered for each isolate replicate cultured in LB broth (n=9), 1x PBS (n=9) and Milli-Q water (n=9), and the reproducibility of replicates in each diluent condition. Horizontal bars on each plot show means and the vertical bars show the SEM. Unpaired *t tests* were performed to identify significant differences ($P < 0.05$) in viable numbers (CFU/ml) recovered after suspension in either diluent for all isolates; no significant results identified. Extracted from Serghiou et al. [19].

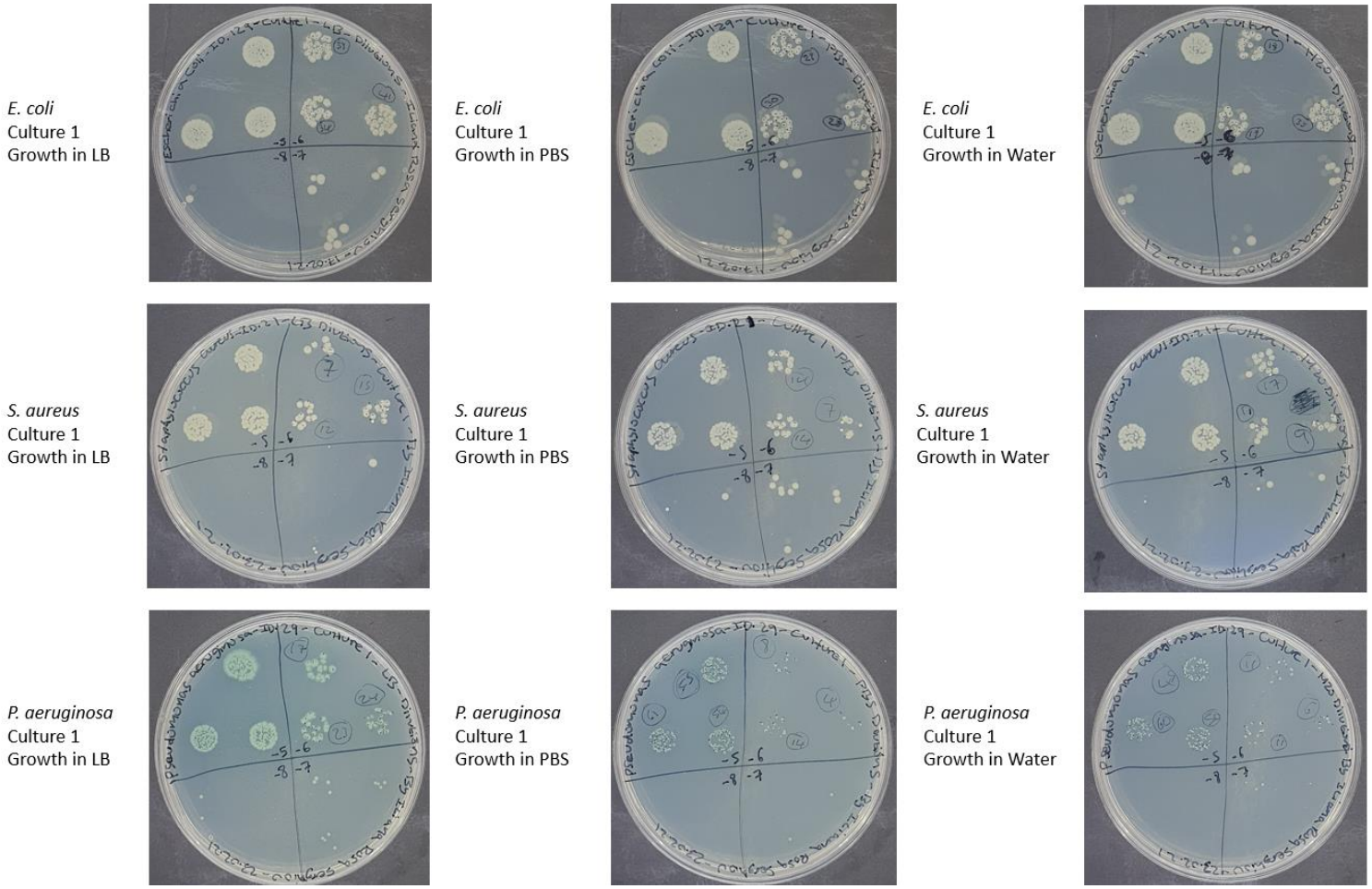


Figure 3.5. Example Photos of *E. coli*, *S. aureus* and *P. aeruginosa* Colony Growth. This figure displays example photos for colony growth from overnight liquid cultures that were grown for *E. coli* (top row), *S. aureus* (middle row) and *P. aeruginosa* (bottom row) in either 1x PBS or UMQW (and LB as a control), for the comparison of cell viability. Three replicate cultures were grown for each isolate (n=9) but only replicate 1 results of each isolate is shown.

3.6.2 Lysis Method Experiments

Comparison of six lysis variations was used to identify the best method for obtaining high yields of HMW DNA from both Gram-negative and Gram-positive bacteria. For each method, factors from common lysis methods (heat, chemical, enzymatic, and mechanical) used in commercial kits for research were varied (section **3.4.2**). Methods 5 and 6 yielded the most DNA from liquid cultures (40.9–97.7 ng/µl and 37–104 ng/µl respectively). There was a significant difference in the quantity of DNA recovered using methods 5 and 6 versus methods 1 and 2 (**Table 3.6**; **Figure 3.6**; $P < 0.05$). Methods 5 and 6 yielded higher DNA concentrations for *P. aeruginosa* (**Figure 3.6**) compared to other species, but there was no statistically significant difference in extraction efficiency between each bacterial species ($P > 0.05$). Methods 2, 4 and 6 extracted DNA of a higher molecular weight than the other methods, with the averages ranging from 20232 to 31786 bp (**Table 3.7**). Together, these results demonstrated that method 6 produced the most DNA of highest molecular weight; the molecular weight produced from method 6 was significantly higher than method 5 ($P < 0.05$). Method 6 was also the most cost effective due to the cheaper Themo Fisher lysozyme used and was chosen for further validation. This method included overnight lysis with lysozyme, a further heated offboard lysis step and a bead beating lysis using the Qiagen Tissue Lyser. DNA extractions from sterile charcoal cotton swabs spiked with independent colonies were successful, with DNA concentrations averaged at 22.1 ng/µl (**Table 3.8**; **Figure 3.7**). There was no significant difference in DNA yield between all swabs, and each species, extracted with and without Agarase ($P > 0.05$).

Table 3.6. Bacterial DNA Yield (ng/µl) from the Six DNA Extraction Methods. Extracted from Serghiou et al. [19]

Sample	Method 1 (ng/µl)	Method 2 (ng/µl)	Method 3 (ng/µl)	Method 4 (ng/µl)	Method 5 (ng/µl)	Method 6 (ng/µl)
E. coli	19.9	11.5	36.8	47.2	40.9	46.8
E. coli	33.2	11.3	46.5	55.3	54.9	38.5
P. aeruginosa	8.68	18.6	63.9	30.8	88.4	96.7
P. aeruginosa	7.96	12.2	33.5	42.7	97.7	104
S. aureus	9.7	5.18	32.8	29.7	51.2	37.8
S. aureus	8.76	4.16	28.7	39.7	49.8	45.6

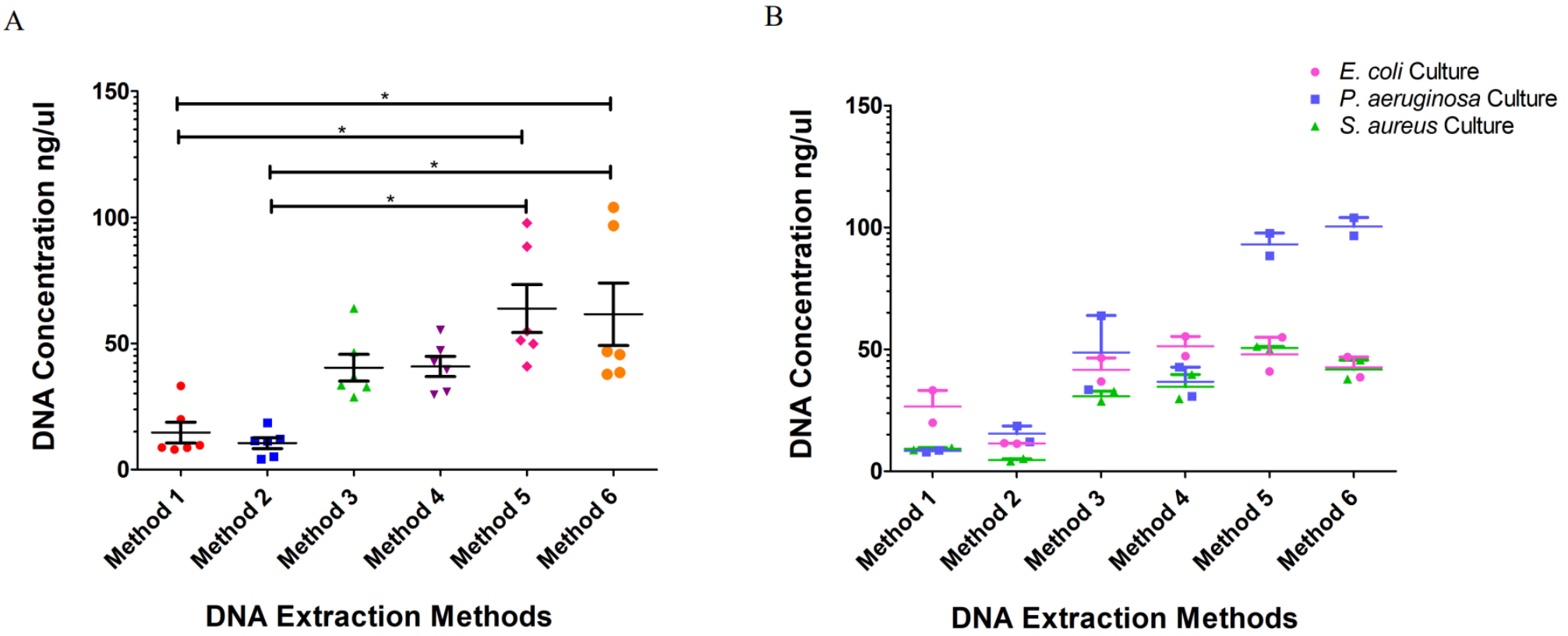


Figure 3.6. Total DNA Yield (ng/µl) Observed for each DNA Extraction Method for *E. coli*, *S. aureus* and *P. aeruginosa*. Duplicate overnight liquid cultures were grown from isolates of three species (*E. coli*, *S. aureus* and *P. aeruginosa*). Six lysis methods were tested on each duplicate overnight liquid culture: *E. coli* (n=12) *S. aureus* (n=12) and *P. aeruginosa* (n=12). All DNA extractions used 1 x PBS as the initial diluent and were completed on the automated Maxwell® RSC 48 Instrument. A Qubit assay was used to quantify the resulting DNA. **A.** Total DNA yield (ng/µl) obtained from each DNA extraction method. **B.** Total DNA yield (ng/µl) per species for each method. Horizontal bars on each plot show means, vertical bars show the SEM and the lines with an asterisk (*) show significant differences ($P < 0.05$). The lines with an asterisk (*) identify the unpaired *t test* significant differences in the quantity of DNA recovered between methods. Extracted from Serghiou et al. [19].

Table 3.7. Average Molecular Weight (bp) of DNA Extracted. Extracted from Serghiou et al. [1]

Sample	Method 1 (bp)	Method 2 (bp)	Method 3 (bp)	Method 4 (bp)	Method 5 (bp)	Method 6 (bp)
E. coli	0	25321	1735	21693	1909	24219
E. coli	0	25697	2057	20232	1877	22730
P. aeruginosa	0	23771	1763	23163	1678	29459
P. aeruginosa	0	23826	1631	23742	1538	31786
S. aureus	0	25085	1831	23954	1668	22118
S. aureus	0	22011	1762	24084	1711	22904

Table 3.8. DNA Yield (ng/µl) of Spiked Swabs Extracted with and without Agarase

Sample	Condition	(ng/µl)
E. coli	Without Agarase	8.81
E. coli	Without Agarase	25.4
E. coli	Without Agarase	30.2
E. coli	Without Agarase	18.9
S. aureus	Without Agarase	9.28
S. aureus	Without Agarase	26.7
S. aureus	Without Agarase	5.6
S. aureus	Without Agarase	4.82
P. aeruginosa	Without Agarase	26.2
P. aeruginosa	Without Agarase	46
P. aeruginosa	Without Agarase	47.4
P. aeruginosa	Without Agarase	89.9
E. coli	With Agarase	18.8
E. coli	With Agarase	10.9
E. coli	With Agarase	24.7
E. coli	With Agarase	29.6
S. aureus	With Agarase	6.81
S. aureus	With Agarase	3.82
S. aureus	With Agarase	6.02
S. aureus	With Agarase	4.04
P. aeruginosa	With Agarase	29.2
P. aeruginosa	With Agarase	11.5
P. aeruginosa	With Agarase	20.7
P. aeruginosa	With Agarase	24.4
Negative Control	N/A	2.31

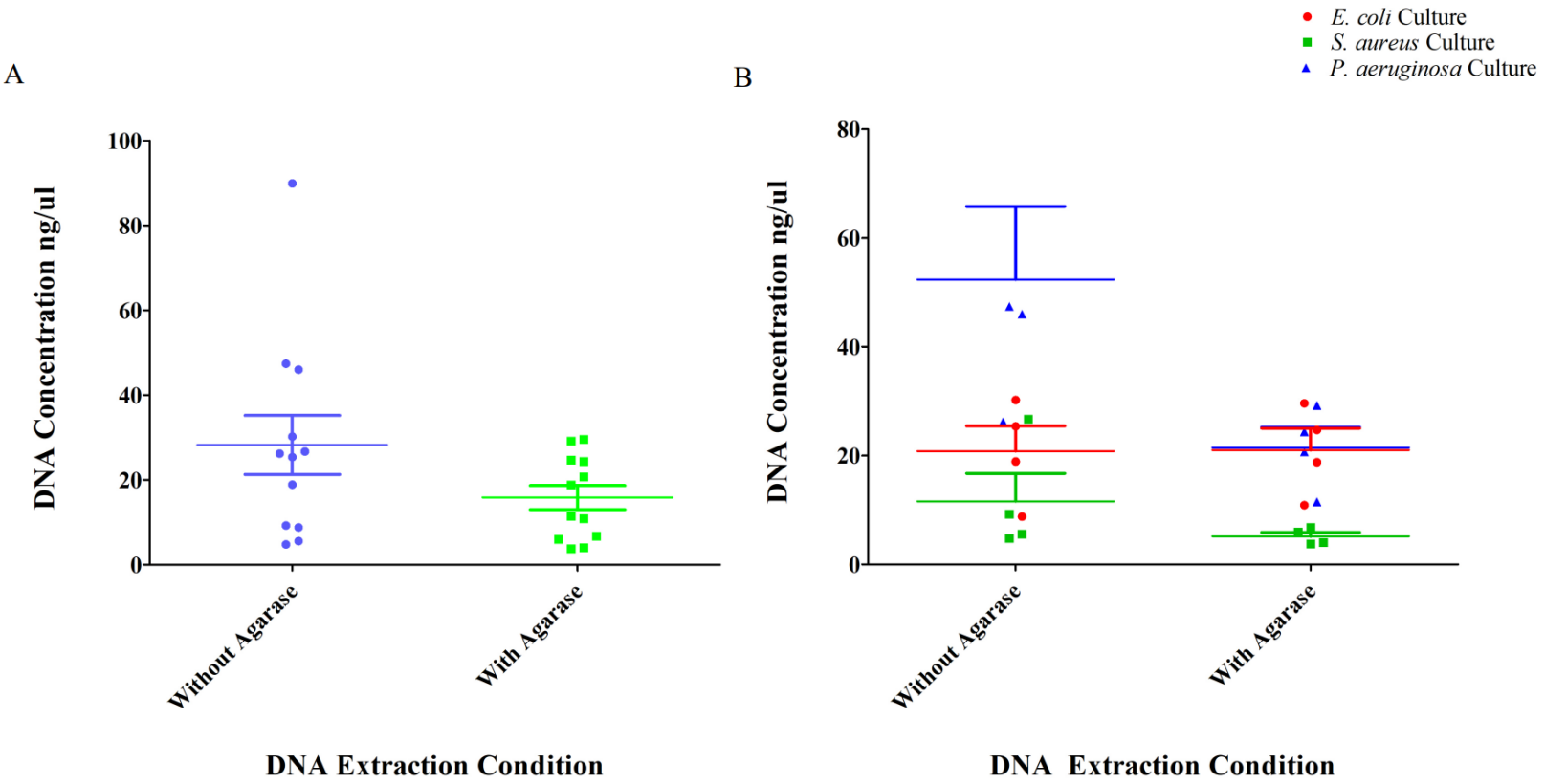


Figure 3.7. DNA Yield (ng/µl) from Spiked Swabs (with *E. coli*, *S. aureus* and *P. aeruginosa*) Extracted with and without Agarase. 24 sterile charcoal cotton swabs were each spiked with one colony from overnight plate cultures of three isolates: *S. aureus* (n=8), *P. aeruginosa* (n=8) and *E. coli* (n=8). Swabs were processed and DNA extracted following the optimised method 6 protocol. During the overnight lysis step, 1 µl β-Agarase-I was added to half of the spiked swabs (n=12), resulting in four replicates per species, per condition. **A.** Total DNA yield (ng/µl) obtained from spiked swabs extracted with and without Agarase. **B.** Total DNA yield (ng/µl) per species extracted with and without Agarase. Horizontal bars on each plot show means, vertical bars show the SEM. Unpaired *t tests* were performed to identify significant differences ($P < 0.05$) in DNA yield between all swabs, and each species, extracted with and without Agarase; no significant results identified.

3.6.3 DNA Extraction Validation with Adult Volunteer and PEARL Study Skin Swabs

DNA concentrations from adult and infant skin swabs, extracted using the optimised method, ranged from <0.50 (no detected DNA) to 10.5 ng/µl (**Table 3.9**). DNA was successfully extracted from all PEARL infant samples, but only 15/24 adult volunteer samples. This result could be due to the adult volunteers having to self-swab, which would be more difficult than being swabbed by another person (as were the infants by their mothers). Cultured plates recovered bacteria from all adult skin swabs, however recovery of cultures from the infant samples was only successful for 4/10 swabs (**Table 3.10**; **Figure 3.8-3.9**). This result was likely due to long term storage of the PEARL study swabs in a -80°C freezer reducing microbial cell viability, whereas the adult volunteer swabs were fresh samples. Concentrations of DNA extracted from cultured bacteria averaged 79.9 ng/µl. The positive control DNA concentrations showed a clear decrease with the dilutions, as expected (**Table 3.9**). Molecular weights of extracted DNA were successfully quantified for 4/24 adult swab extractions (**Table 3.9**) and 17/24 adult swab culture extractions (**Table 3.10**), ranging from 1000 to 6000 bp; molecular weights were only quantified for the adult swabs as a validation that the extraction method was able to produce appropriate molecular weight for real human swabs. These results were due to low, or no DNA detected by the Tapestation to calculate the molecular weight. A QC summary of the total number of reads, and human and microbial reads generated from each sample can also be viewed in Tables 3.9 and 3.10. Results for the DNA yield at different overnight lysis incubation times identified a significant difference between 18 and 20 h, and 18 and 22 h *(*$P < 0.05$*)*, but no significant difference was observed between 20 and 22 h *(*$P > 0.05$*)* (**Figure 3.10**). The adult volunteer swabs that did not yield detectable amounts of DNA were those incubated for 18 h, therefore subsequent samples were incubated between 20 and 22 h to obtain higher DNA yield.

Table 3.9. Skin Swab DNA Yield/Weight and Sequencing QC Summary. Extracted from Serghiou et al. [19]

Sample ID	DNA Yield (ng/µl)	Average Molecular Weight (bp)	Illumina Total Reads	Illumina Total Human Reads	Illumina Total Microbial Reads	Nanopore Total Human Reads	Nanopore Total Microbial Reads	Nanopore Total Reads
Volunteer 1 – Swab 1	< 0.50	Not Detected	149280	69698	79582	10	51	61
Volunteer 1 – Swab 2	< 0.50	Not Detected	474556	158980	315576	15	89	104
Volunteer 2 – Swab 1	< 0.50	Not Detected	296986	91226	205760	9	21	30
Volunteer 2 – Swab 2	< 0.50	Not Detected	444346	219776	224570	27	111	138
Volunteer 3 – Swab 1	< 0.50	Not Detected	1476424	813892	662532	60	463	523
Volunteer 3 – Swab 2	< 0.50	Not Detected	3743138	1300116	2443022	59	315	374
Volunteer 4 – Swab 1	< 0.50	3563	3822458	43180	3779278	7	614	621
Volunteer 4 – Swab 2	< 0.50	Not Detected	221642	173798	47844	9	37	46
Volunteer 5 – Swab 1	1.09	Not Detected	802948	557172	245776	30	87	117
Volunteer 5 – Swab 2	2.16	Not Detected	534748	388310	146438	47	39	86
Volunteer 6 – Swab 1	3.31	Not Detected	4425280	1005038	3420242	59	999	1058
Volunteer 6 – Swab 2	1.67	3934	911920	89916	822004	19	290	309
Volunteer 7 – Swab 1	0.68	Not Detected	188944	149928	39016	21	83	104
Volunteer 7 – Swab 2	0.477	Not Detected	2030116	1021646	1008470	42	180	222
Volunteer 8 – Swab 1	< 0.50	Not Detected	2874830	249134	2625696	17	162	179
Volunteer 8 – Swab 2	0.156	4885	2134824	21944	2112880	682	24350	25032
Volunteer 9 – Swab 1	0.837	Not Detected	160574	88216	72358	18	98	116
Volunteer 9 – Swab 2	0.809	Not Detected	1879600	1348416	531184	66	137	203
Volunteer 10 – Swab 1	6.52	4554	11690	6404	5286	433	5305	5738
Volunteer 10 – Swab 2	9.73	Not Detected	145694	63050	82644	7	33	40
Volunteer 11 – Swab 1	10.5	Not Detected	37436	21794	15642	296	3574	3870
Volunteer 11 – Swab 2	5	Not Detected	64242	16252	47990	328	3933	4261
Volunteer 12 – Swab 1	1.69	Not Detected	1011326	770872	240454	29	73	102
Volunteer 12 – Swab 2	3.63	Not Detected	205884	25522	180362	510	6161	6671
Positive Control 1 (Neat)	2.47	510	1783516	118	1783398	374	18674	19048
Positive Control 2 (10^{-1})	1.1	Not Detected	320322	47872	272450	1172	14868	16040
Positive Control 3 (10^{-2})	0.41	Not Detected	20022	142	19880	279	5922	6201
Positive Control 4 (10^{-3})	0.261	Not Detected	3738	318	3420	346	6467	6813
Negative Control	0.064	Not performed	5216	62	5154	177	3759	3936
Baby 1 – Month 4 - Swab	5.1	Not performed	1551032	649294	901738	19363	34305	53668
Baby 2 – Month 4 - Swab	4.05	Not performed	46308	26690	19618	3510	5552	9062
Baby 3 – Month 4 - Swab	4.48	Not performed	369470	304628	64842	16167	5700	21867
Baby 5 – Twin 1 – Month 4 - Swab	4.07	Not performed	2389960	2330166	59794	112683	10920	123603
Baby 5 – Twin 2 – Month 4 - Swab	3.71	Not performed	169008	87880	81128	3124	4323	7447
Baby 6 – Month 4 - Swab	4.47	Not performed	878478	694382	184096	34750	18487	53237
Baby 7 – Month 4 - Swab	3.61	Not performed	327370	247800	79570	22096	9363	31459
Baby 8 – Month 4 - Swab	3.91	Not performed	1141178	651174	490004	14342	15360	29702
Baby 9 – Month 4 - Swab	3.22	Not performed	342902	309828	33074	31036	6296	37332
Baby 10 – Month 4 - Swab	3.46	Not performed	31536	19818	11718	4719	4095	8814
Positive Control	4.53	Not performed	7944470	186	7944284	23	425832	425855
Negative Control	2.69	Not performed	3400	72	3328	5	200	205

Table 3.10. Skin Swab Culture DNA Yield/Weight and Sequencing QC Summary

Sample ID	DNA Yield (ng/µl)	Average Molecular Weight (bp)	Illumina Total Reads	Illumina Total Human Reads	Illumina Total Microbial Reads	Nanopore Total Human Reads	Nanopore Total Microbial Reads	Nanopore Total Reads
Volunteer 1 - Swab 1 - Culture	26.9	Not Detected	4990780	92	4990688	36	3760	3796
Volunteer 1 - Swab 2 - Culture	32.4	5083	4745388	58	4745330	23	3403	3426
Volunteer 2 - Swab 1 - Culture	6.29	6007	4418158	52	4418106	27	3268	3295
Volunteer 2 - Swab 2 - Culture	57.1	5434	4847838	48	4847790	31	3452	3483
Volunteer 3 - Swab 1 - Culture	38.9	1295	3172496	52	3172444	23	1736	1759
Volunteer 3 - Swab 2 - Culture	204	Not Detected	4137884	28	4137856	21	3081	3102
Volunteer 4 - Swab 1 - Culture	91.5	4896	3475754	28	3475726	130	6535	6665
Volunteer 4 - Swab 2 - Culture	7.73	Not Detected	3856832	84	3856748	40	3931	3971
Volunteer 5 - Swab 1 - Culture	93.3	5175	4574220	40	4574180	24	4311	4335
Volunteer 5 - Swab 2 - Culture	3.43	1447	4856288	422	4855866	33	4108	4141
Volunteer 6 - Swab 1 - Culture	680	1497	2459038	22	2459016	0	283	283
Volunteer 6 - Swab 2 - Culture	178	4499	2859508	34	2859474	12	1315	1327
Volunteer 7 - Swab 1 - Culture	10.3	1382	4494442	68	4494374	29	3513	3542
Volunteer 7 - Swab 2 - Culture	77.7	4886	4612254	42	4612212	35	3637	3672
Volunteer 8 - Swab 1 - Culture	289	Not Detected	2412866	12	2412854	7	1087	1094
Volunteer 8 - Swab 2 - Culture	334	Not Detected	2499434	10	2499424	15	1494	1509
Volunteer 9 - Swab 1 - Culture	2.75	1101	3635370	116	3635254	28	2535	2563
Volunteer 9 - Swab 2 - Culture	21.4	4629	2684504	18	2684486	22	1509	1531
Volunteer 10 - Swab 1 - Culture	4.98	1124	4511712	44	4511668	29	2974	3003
Volunteer 10 - Swab 2 - Culture	2.17	Not Detected	4469876	62	4469814	25	2737	2762
Volunteer 11 - Swab 1 - Culture	14.8	1394	4240154	156	4239998	43	4129	4172
Volunteer 11 - Swab 2 - Culture	2.4	Not Detected	627278	170	627108	3	430	433
Volunteer 12 - Swab 1 - Culture	19.5	4976	3365414	26	3365388	29	2918	2947
Volunteer 12 - Swab 2 - Culture	22.8	5269	4090276	28	4090248	47	6206	6253
Positive Control	2.78	Not Detected	3249852	90	3249762	36	3273	3309
Negative Control	2.57	Not Detected	3804	1432	2372	24	1374	1398
Baby 1 – Month 4 – Swab - Culture	6.4	Not performed	3508900	120	3508780	9	378863	378872
Baby 2 – Month 4 – Swab - Culture	2.79	Not performed	362834	1212	361622	40	86034	86074
Baby 5 – Twin 1 – Month 4 – Swab - Culture	2.52	Not performed	146188	380	145808	91	65671	65762
Baby 6 – Month 4 – Swab - Culture	4.93	Not performed	3410516	528	3409988	14	349299	349313
Positive Control	5.72	Not performed	8411820	480	8411340	19	426288	426307
Negative Control	2.61	Not performed	878	118	760	10	159	169

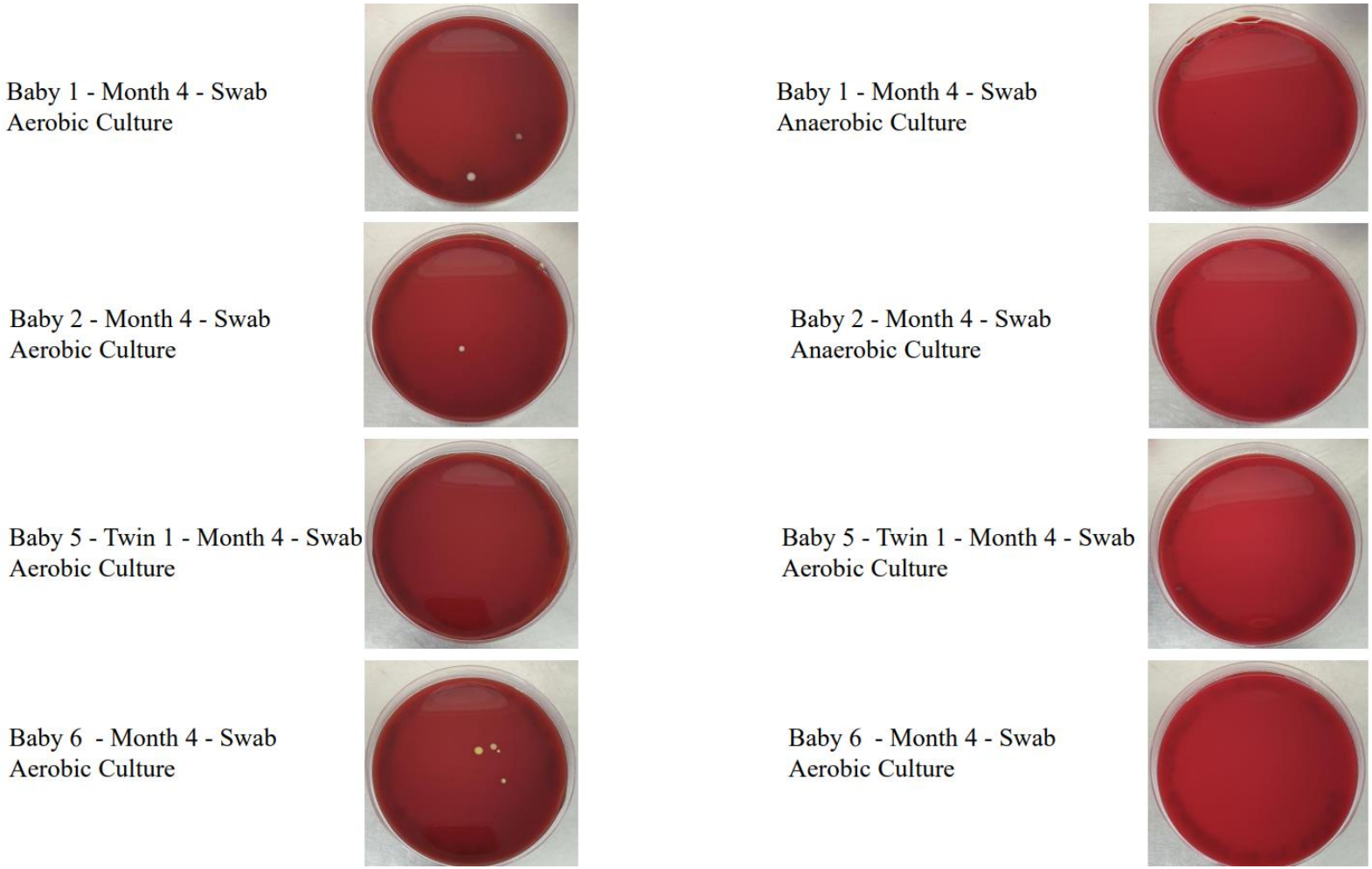


Figure 3.8. Skin Swab Aerobic and Anaerobic Cultures Obtained from PEARL Infants. The optimised bacterial DNA extraction method was performed on 10 PEARL infant skin swabs that were collected at 4 months old, with aerobic and anaerobic culturing of the swabs performed in parallel. Swabs were cultured on Columbia Blood Agar (CBA + 5% horse blood). Each row shows the aerobic and anaerobic cultures obtained from each PEARL infant from one arm swab (n=4).

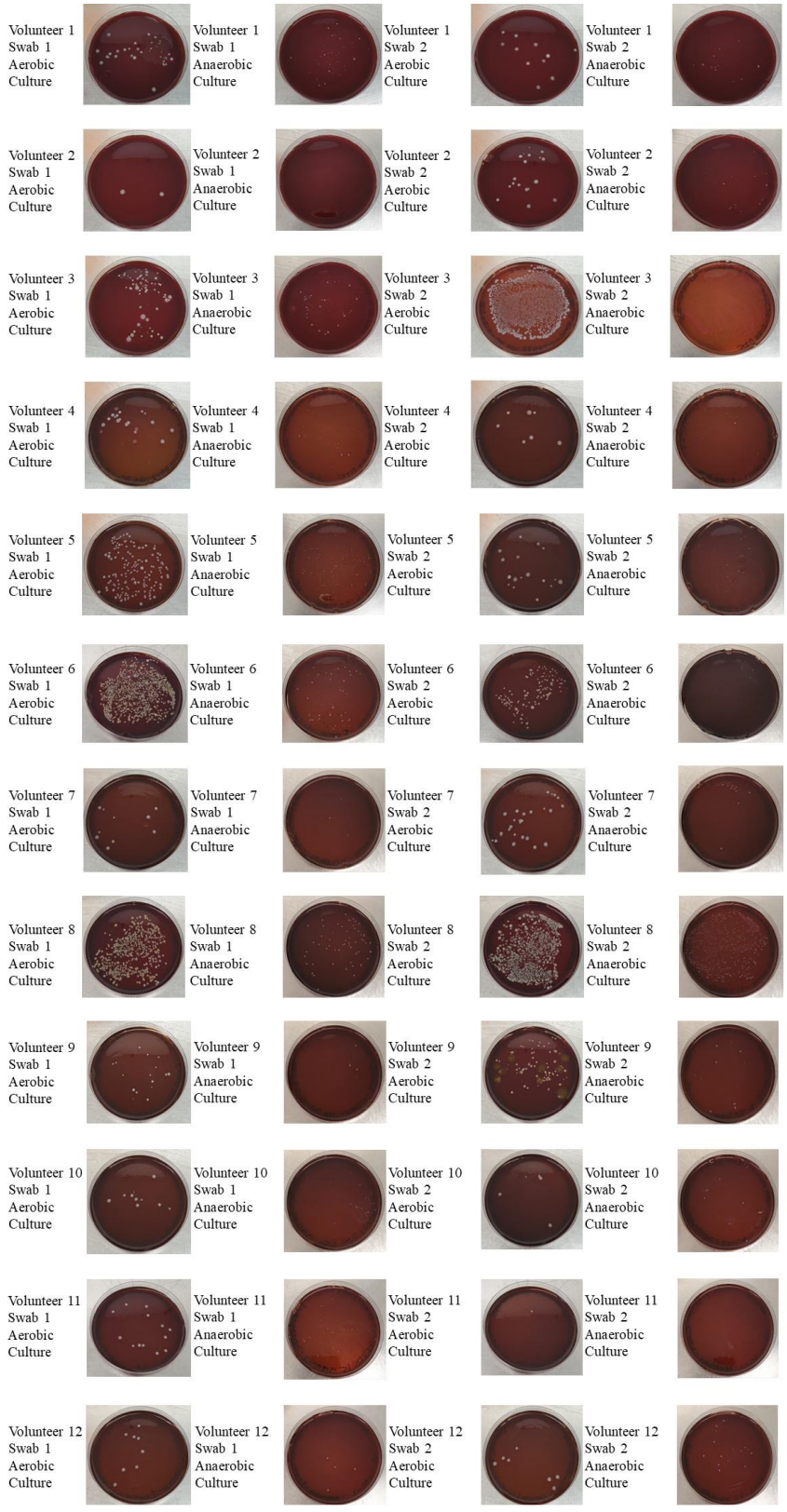


Figure 3.9. Skin Swab Aerobic and Anaerobic Cultures Obtained from Adult Volunteers. The optimised bacterial DNA extraction method was performed on 24 skin swabs from 12 adult volunteers (each volunteer had a skin swab collected from the right and left arms), with aerobic and anaerobic culturing of the swabs performed in parallel. Swabs were cultured on Columbia Blood Agar (CBA + 5% horse blood). Each row shows the aerobic and anaerobic cultures from the left and right arm swabs of each volunteer participant. It is unknown which swabs represent the left and right arms as this was anonymised.

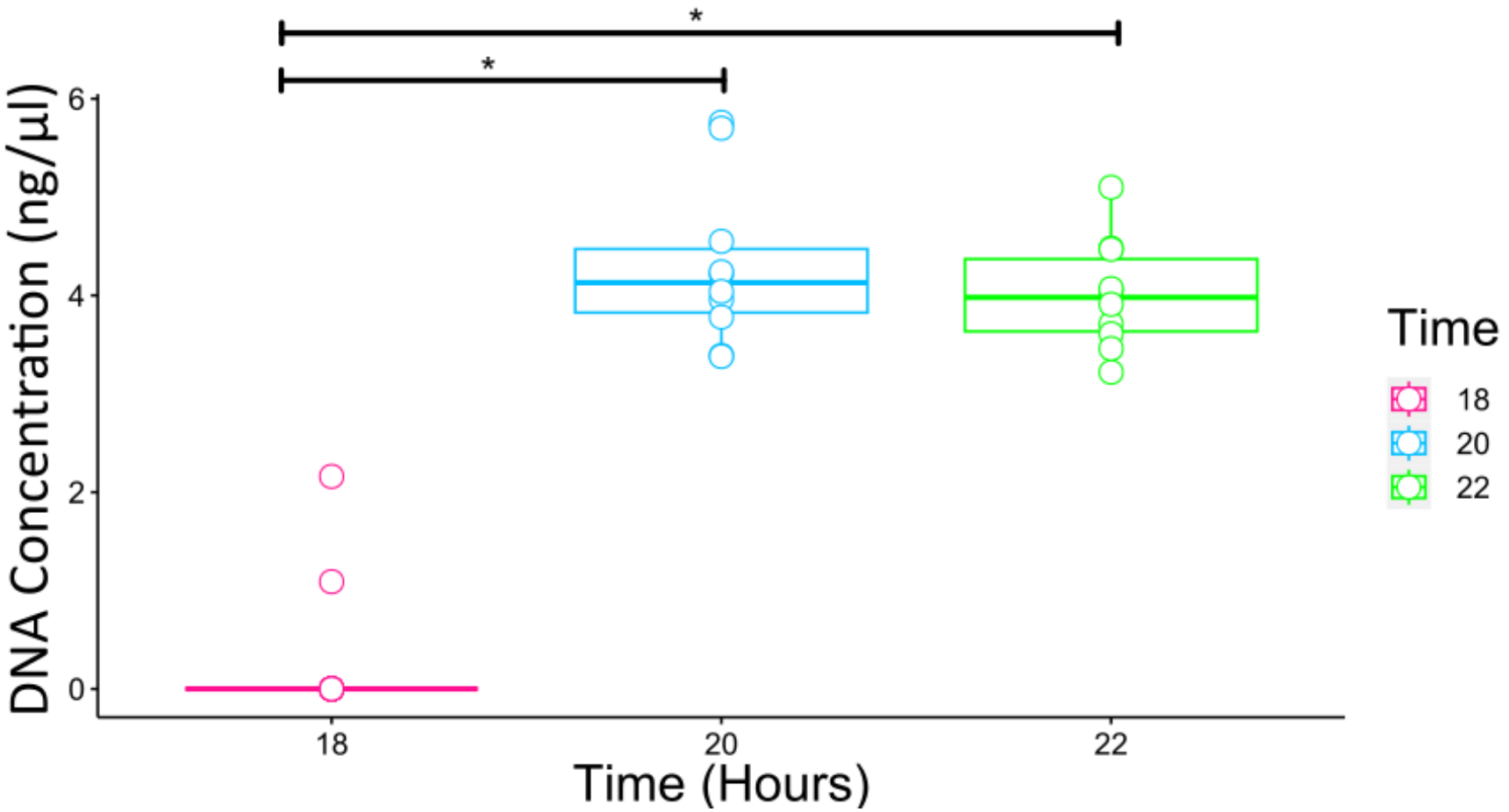


Figure 3.10. Comparison of DNA Yield (ng/µl) from Swabs Incubated for Different Periods. The box plot shows the extracted DNA concentrations (ng/µl) from PEARL study infant skin swabs (collected from one forearm from each infant) that included an offboard heated lysis incubation period at three incubation times (18, 20 and 22 hours- n=10 per incubation time). Horizontal bars on each box show the means, vertical bars show the standard error of the mean (SEM) and lines with an asterisk (*) indicate significant ($P < 0.05$) differences. The key shows incubation times, which were allocated a unique colour. Extracted from Serghiou et al. [19].

Illumina sequencing per sample yielded up to 300 Mbp data, and ONT up to 10 Gb. The taxonomic profiles generated show the microbial species diversity within each adult and infant individual for the swab and culture Illumina and ONT datasets (**Figures 3.11-3.16**). The positive controls displayed the expected microbiota in the ATCC skin microbiome whole cell mix: *Acinetobacter johnsonii*, *Corynebacterium striatum*, *Micrococcus luteus*, *Cutibacterium acnes*, *Staphylococcus epidermidis*, and *Streptococcus mitis*. There was a clear reduction in reads across the dilution series. This demonstrated that the DNA extraction method was able to effectively extract DNA from the diverse range of species present in the ATCC skin microbiome mix. Some background contamination was detected in negative controls (Illumina Median Number of Reads = 2850; Nanopore Median Number of Reads = 787; Illumina IQR of Reads = 2675; Nanopore IQR of Reads = 2387), although the number of reads was always much lower than in test samples and most DNA fragments identified in the negative controls mostly mapped against organisms not seen in the test samples (**Tables 3.9-3.10; Figures 3.13-3.14, 3.16**).

Both Illumina and ONT data indicated a typical skin microbiota from both adult and infant skin swabs and generated a sufficient number of reads to allow downstream taxonomic analysis at the species level. The adult skin swabs contained bacteria, viruses, and phages, whereas the infant swabs only displayed bacterial diversity. Infant skin swabs contained more *Streptococcus* and fewer *Staphylococcus* species (7 and 3 species respectively) when compared to adult skin swabs (6 and 7 species respectively). The infant skin swabs also contained *Bifidobacterium longum*, *Bifidobacterium breve* and *Bifidobacterium bifidum*, which are not typical skin residents, but common residents of the infant gut. This likely demonstrates transient skin contamination on the infants [396, 397]. Importantly, data were generated for skin swabs that had no DNA detected by the Qubit (< 0.50).

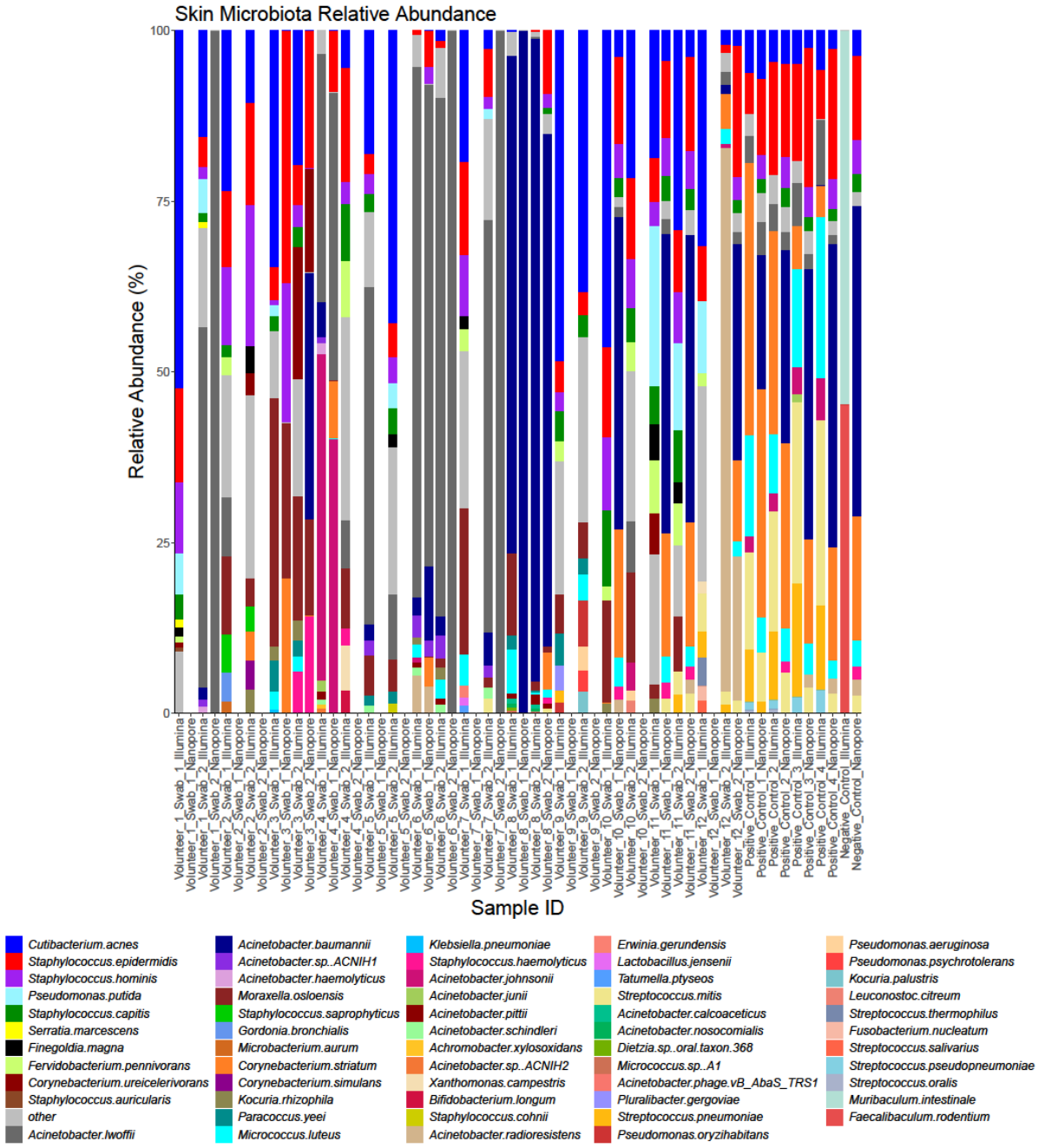


Figure 3.11. Taxonomic Profiles of Skin Swabs from Adult Volunteers Generated using Illumina and ONT Data. Composition data was generated from adult volunteer skin swabs using shotgun metagenomics sequencing data, generated on the Illumina and Oxford Nanopore platforms. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bar plots display the relative abundance data of the 10 most abundant species that occur within each sample. The origin swab samples were collected from adult volunteers (n=12) between the ages of 23 and 65, with one swab collected from each forearm from each volunteer (n=24). The key allocates a unique colour to the species represented in the stacked bars. Extracted from Serghiou et al. [19].

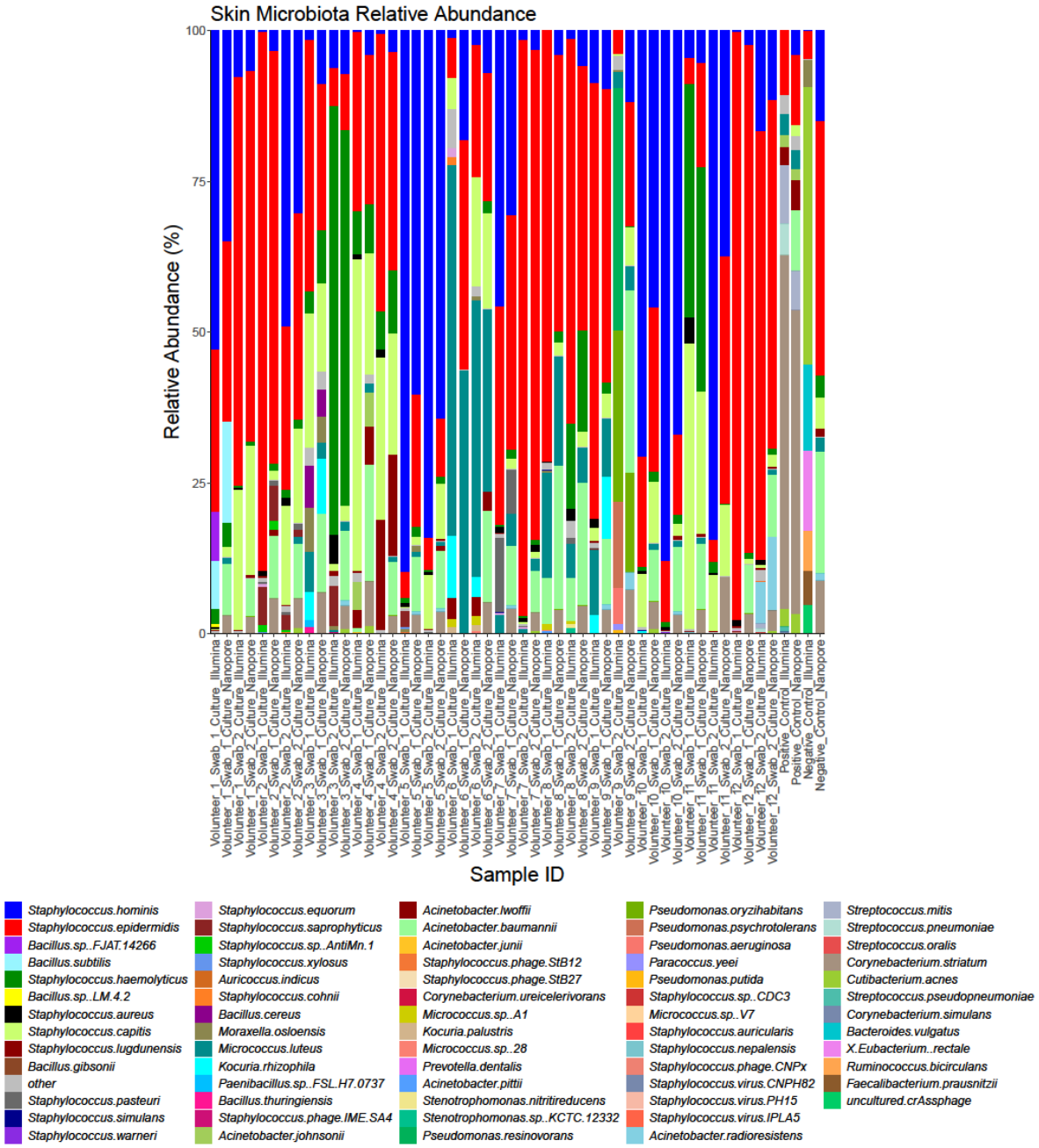


Figure 3.12. Taxonomic Profiles of Skin Swab Cultures from Adult Volunteers Generated using Illumina and ONT Data. Composition data was generated from adult volunteer skin swab cultures using shotgun metagenomics sequencing data, generated on the Illumina and Oxford Nanopore platforms. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bar plots display the relative abundance data of the 10 most abundant species that occur within each sample. The origin swab samples were collected from adult volunteers (n=12) between the ages of 23 and 65, with one swab collected from each forearm from each volunteer (n=24). The key allocates a unique colour to the species represented in the stacked bars. Extracted from Serghiou et al. [19].

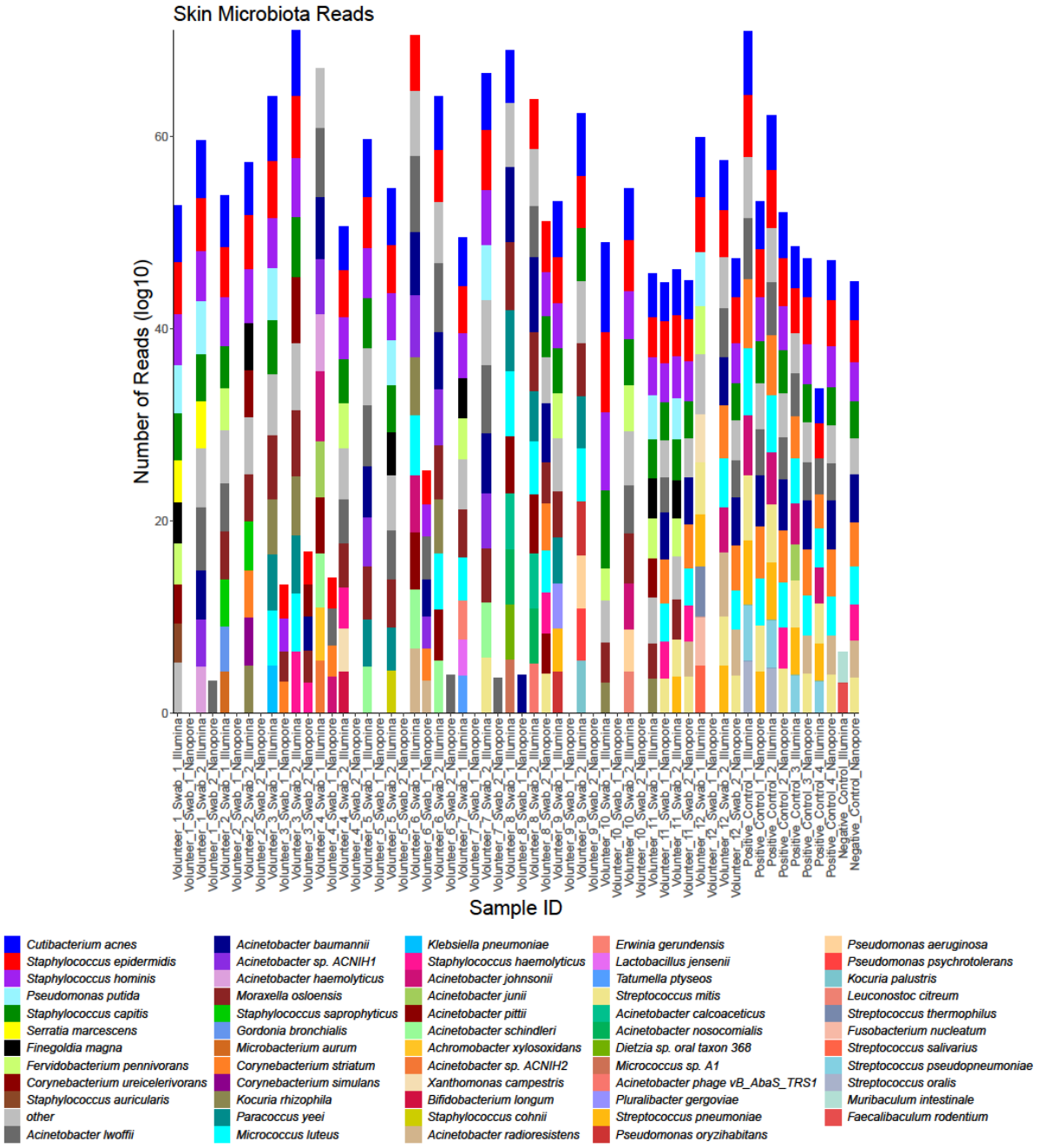


Figure 3.13. Taxonomic Profile Reads of Skin Swabs from Adult Volunteers Generated using Illumina and ONT Data. Composition data was generated from adult volunteer skin swabs using shotgun metagenomics sequencing data, generated on the Illumina and Oxford Nanopore platforms. Read numbers of the taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bar plots display the number of reads of the 10 most abundant species that occur within each sample. The origin swab samples were collected from adult volunteers (n=12) between the ages of 23 and 65, with one swab collected from each forearm from each volunteer (n=24). The key allocates a unique colour to the species represented in the stacked bars.

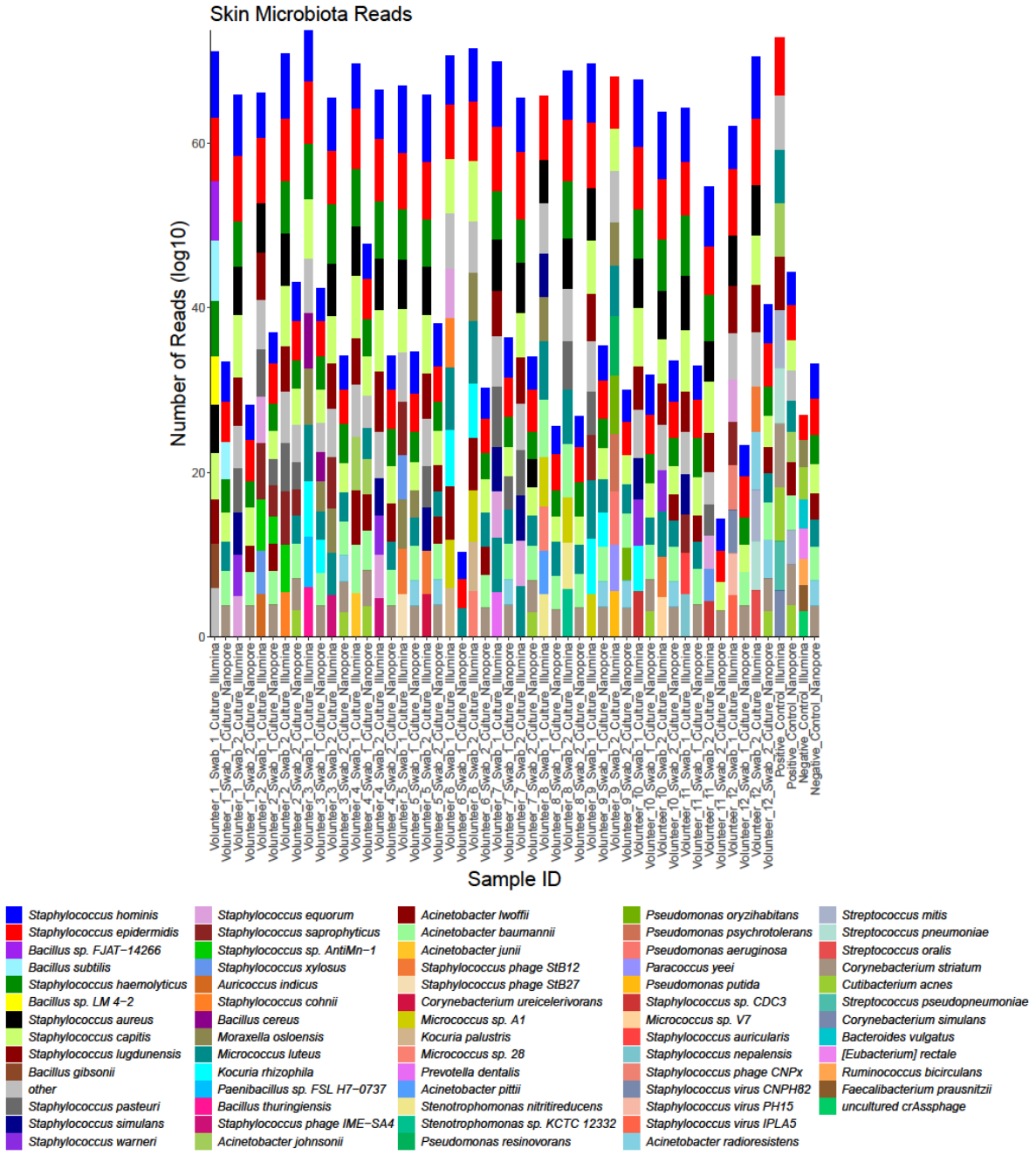


Figure 3.14. Taxonomic Profile Reads of Skin Swab Cultures from Adult Volunteers Generated using Illumina and ONT Data. Composition data was generated from adult volunteer skin swab cultures using shotgun metagenomics sequencing data, generated on the Illumina and Oxford Nanopore platforms. Read numbers of the taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bar plots display the number of reads of the 10 most abundant species that occur within each sample. The origin swab samples were collected from adult volunteers (n=12) between the ages of 23 and 65, with one swab collected from each forearm from each volunteer (n=24). The key allocates a unique colour to the species represented in the stacked bars.

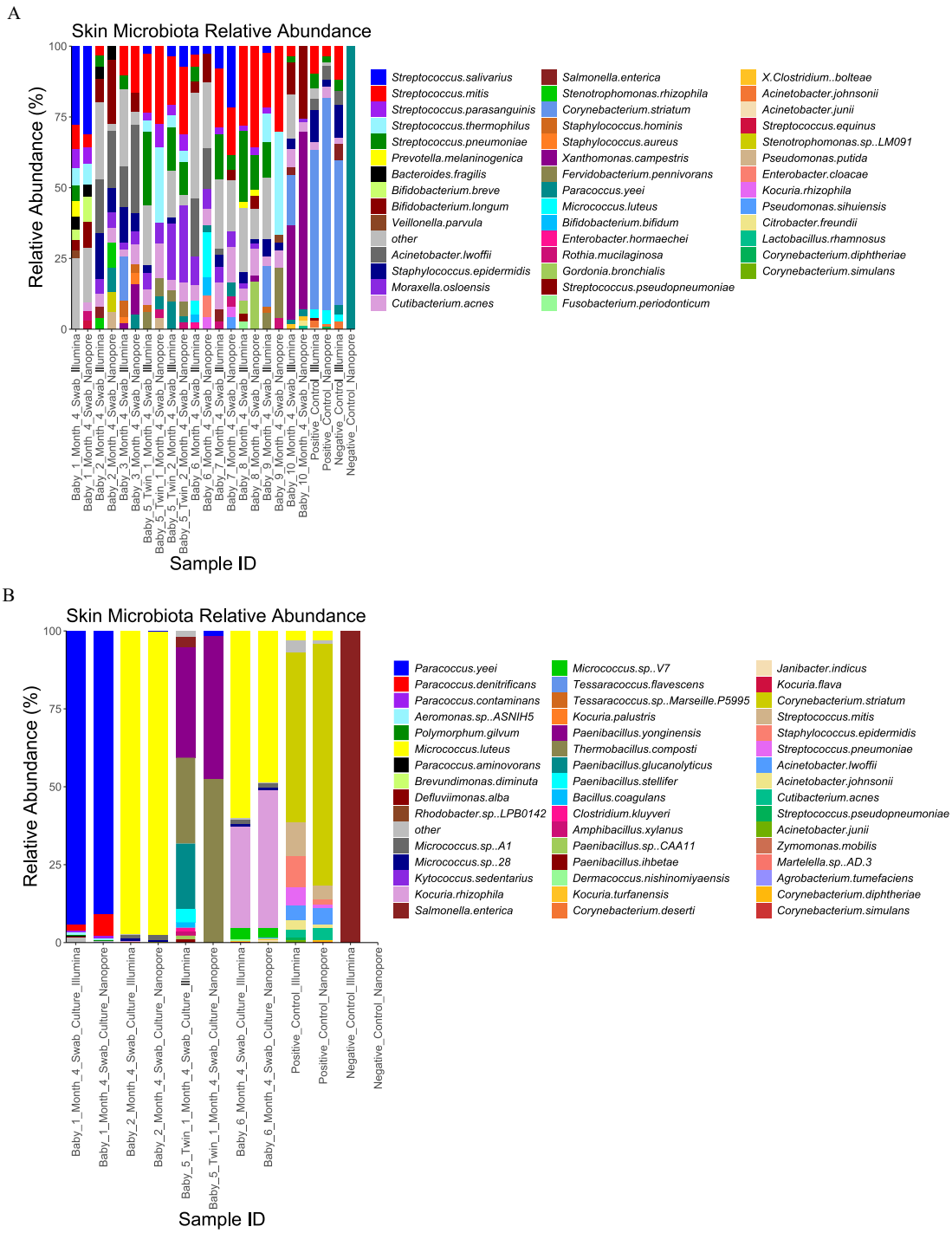


Figure 3.15. Taxonomic Profiles of the Skin Swabs and Cultures from PEARL Study Infants Generated using Illumina and ONT Data. Composition data was generated from PEARL study infant skin swabs and cultures using shotgun metagenomics sequencing data, generated on the Illumina and Oxford Nanopore platforms. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bar plots displays the relative abundance data of the 10 most abundant species that occur within each sample. The origin swab samples were collected from PEARL study infants at 4 months old with one swab collected from one forearm (n=10). The key allocates a unique colour to the species represented in the stacked bars. **A.** Illumina and ONT skin swab data. **B.** Illumina and ONT skin culture data. The key allocates a unique colour to the species represented in the stacked bars. Extracted from Serghiou et al. [19].

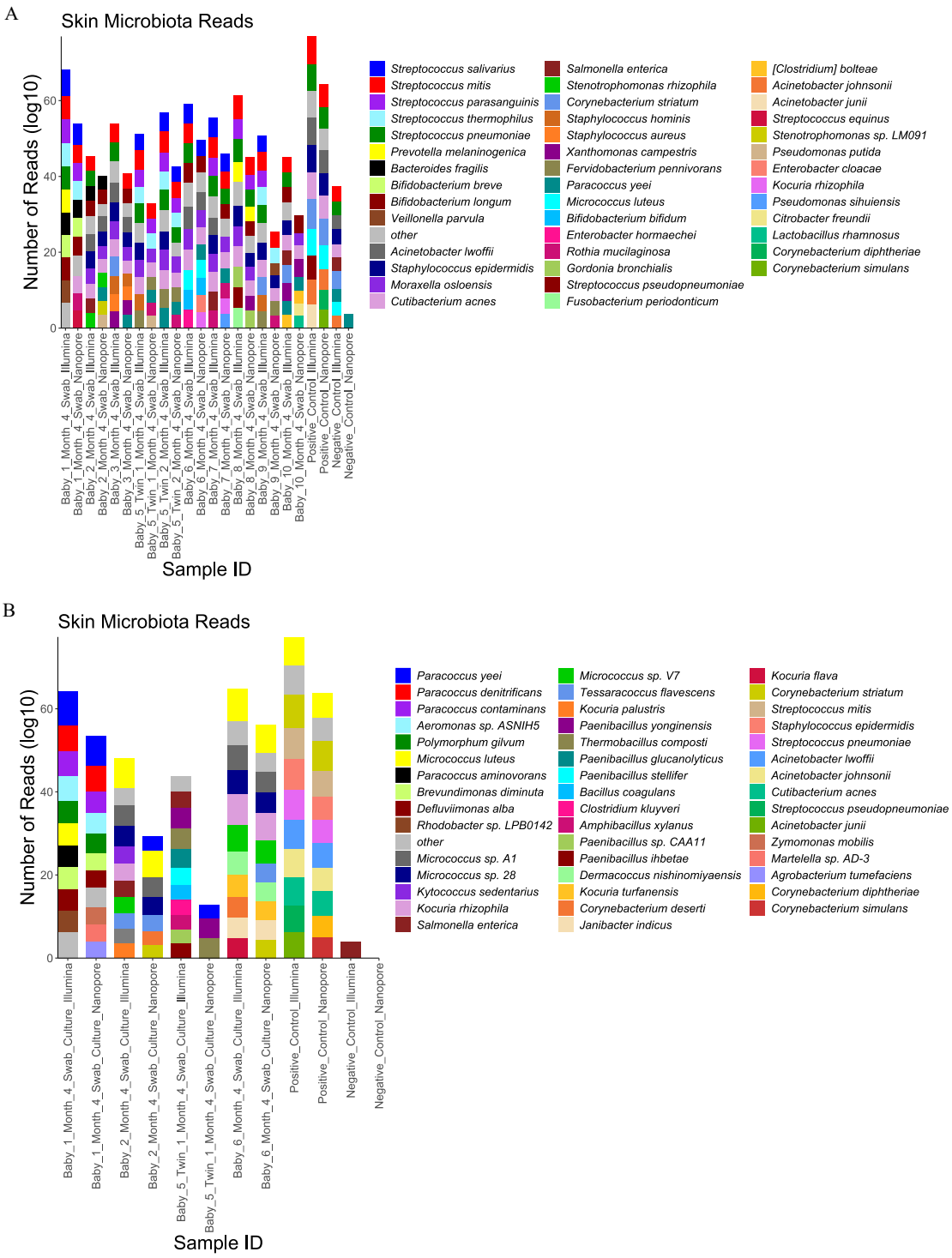


Figure 3.16. Taxonomic Profile Reads of the Skin Swabs and Cultures from PEARL Study Infants Generated using Illumina and ONT Data. Composition data was generated from PEARL study infant skin swabs and cultures using shotgun metagenomics sequencing data, generated on the Illumina and Oxford Nanopore platforms. Read numbers of the taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bar plots display the number of reads of the 10 most abundant species that occur within each sample. The origin swab samples were collected from PEARL study infants at 4 months old with one swab collected from one forearm (n=10). The key allocates a unique colour to the species represented in the stacked bars. **A.** Illumina and ONT skin swab data. **B.** Illumina and ONT skin culture data.

Illumina and ONT platforms identified very similar microbiota profiles for both skin swabs and cultures, with comparable percentage total counts of the most abundant species (those representing more than 0.5 % of each sample) (**Figure 3.17**).

There was a significant difference identified between the AT and GC content of samples sequenced on both platforms (**Figure 3.18**; $P < 0.05$), indicating platform sequencing bias. However, due to both platforms identifying similar taxonomic profiles, this difference in AT and GC content was not viewed to cause bias in species identification and could be mitigated by using a combined sequencing approach.

Species recovery identified more species, as more data was used. There was a significant difference ($P < 0.05$) in species recovery between 1 Gb and 5 Gb, though no significant difference was found between 2.5 and 5 Gb of data (**Figure 3.19. A**). Recovery of MAGs was also higher from samples where 5 Gb of data were used than 1Gb, although this difference was not found to be statistically significant (**Figure 3.19. B**). MAGs predicted to be from diverse genera were isolated, showing the method can isolate and assemble DNA from different parts of the phylogeny of the skin microbiota (**Table 3.11**). Based on this analysis, 5 Gb of data per sample appeared to be adequate for phylogenetic analysis of the skin microbiota using this method, whilst providing enough data for generation of MAGs. This allowed detailed study of the functional capacity of the skin microbiome based on these results.

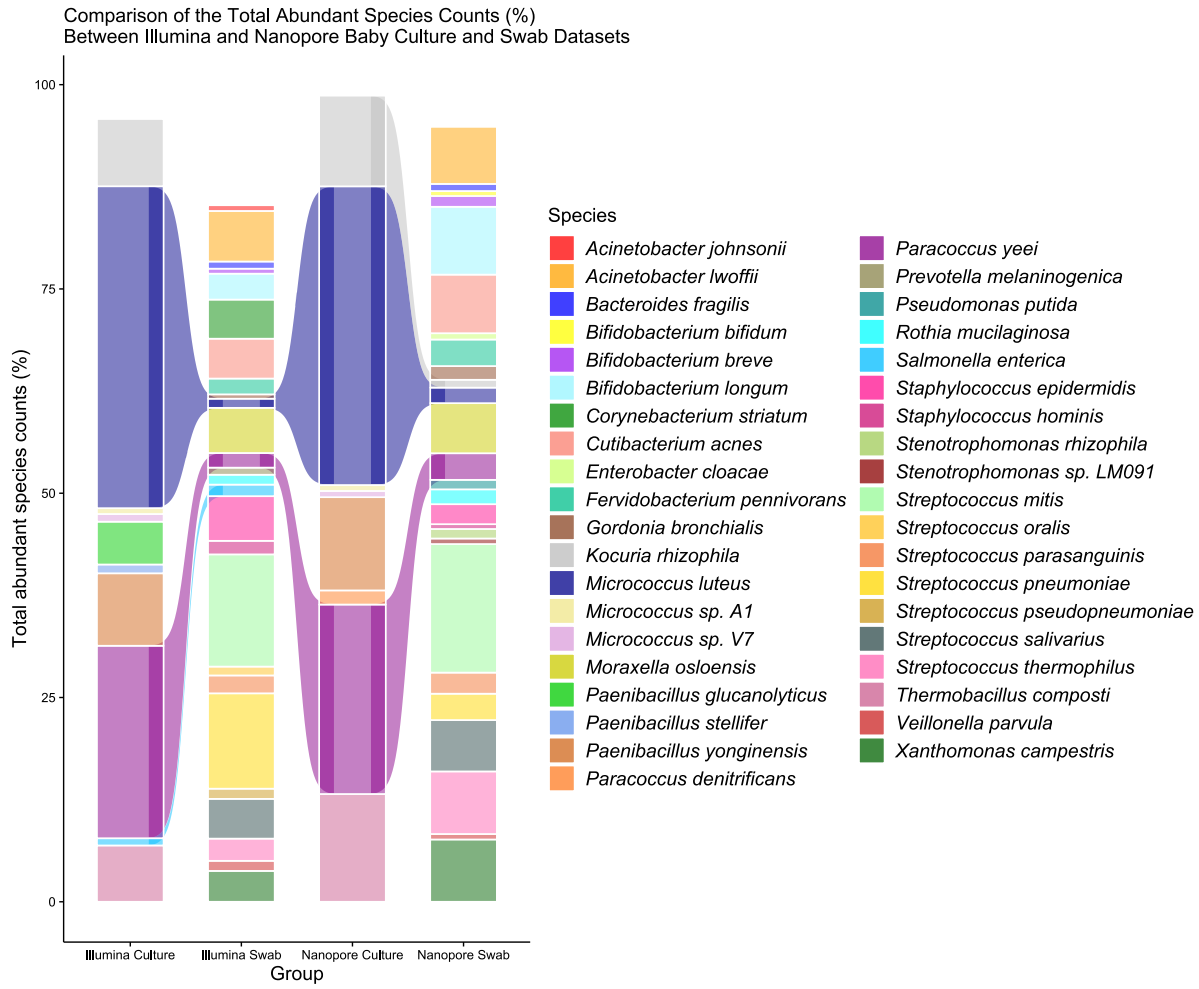


Figure 3.17. Alluvial Plot of the Percentage Total Abundant Species Counts between Illumina and ONT Shotgun Metagenomics data from the PEARL Study Infant Skin Swab and Culture Datasets. Composition data was generated from PEARL study infant skin swabs and cultures using shotgun metagenomics sequencing data, generated on the Illumina and Oxford Nanopore platforms. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. The origin swab samples were collected from 10 PEARL study infants at 4 months old with one swab collected from one forearm (n=10). Counts were calculated by adding the total counts of each species, which were converted into a percentage of the total species count. Only species with >0.5% of counts were used. Adjoining lines show the abundant skin species that were detected by both platforms. The key allocates a unique colour to the species represented in the stacked bars. Extracted from Serghiou et al. [19].

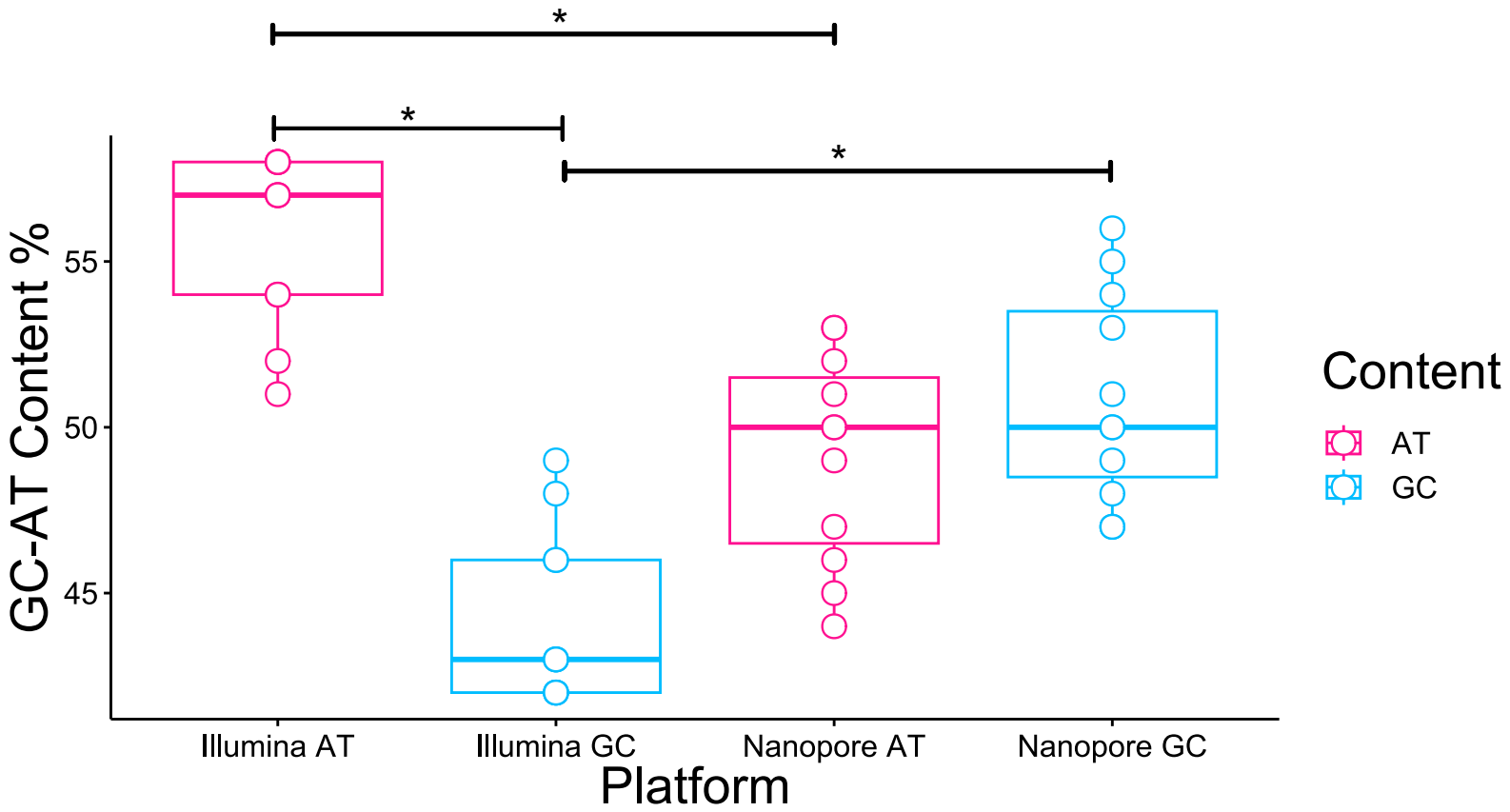


Figure 3.18. GC-AT Content (%) of the Illumina and Nanopore Swab Datasets from the PEARL Study Infants. Composition data was generated from PEARL study infant skin swabs using shotgun metagenomics sequencing data, generated on the Illumina and Oxford Nanopore platforms. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. The origin swab samples were collected from PEARL study infants at 4 months old with one swab collected from one forearm (n=10). The box plot shows the AT and GC content (%) of the shotgun metagenomics profiles from the infant samples sequenced on both platforms. Horizontal bars on each plot show averages, vertical bars show the SEM and the lines with an asterisk (*) show significant differences ($P < 0.05$). Extracted from Serghiou et al. [19].

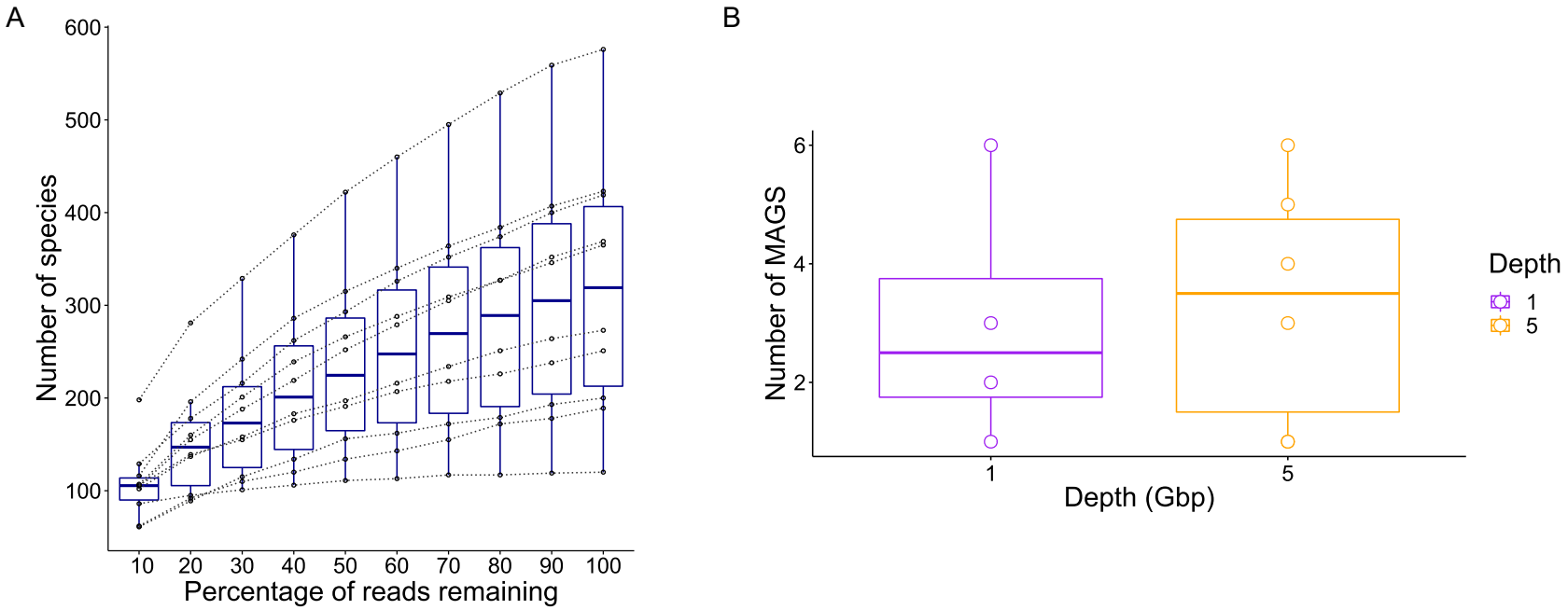


Figure 3.19. Species and MAGs Recovery from PEARL Infant Skin Swabs. Composition data was generated from PEARL study infant skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Read taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. The shotgun metagenomics data was run through a MetaWRAP three binning (MetaBAT, MAXBIN2 and CONCOCT) pipeline to generate high quality MAGs assessed using the criteria from Kiu et al. [398] (completeness >80% and contamination <10%). The origin swab samples were collected from PEARL study infants at 4 months old with one swab collected from one forearm (n=10). **A.** Species recovery at 5 Gb sequencing depth and subsampled depths. The horizontal bars give the average number of species recovered at each level of reads, and the dots and lines show how each sample changes with reduced sequencing depth. 10%-100% reads represent 0.5 Gb-5Gb at increments of 0.5 Gb. **B.** MAGs recovery at 1 Gb and 5 Gb sequencing depths. Horizontal bars give the average number of MAGs at each sequencing depth. The key allocates a unique colour to the sequencing depth. The vertical lines on each box plot show the SEM. Unpaired *t tests* identified significant differences ($P < 0.05$) between species and MAGs recovery from samples sequenced at 1 Gb to 5 Gb of data. Extracted from Serghiou et al. [19].

Table 3.11. MAGs Recovered from Ten Infant Skin Swabs at 1 Gb and 5 Gb Sequencing Depth. Extracted from Serghiou et al. [19]

Sequence Depth (GB)	Sample	Number of MAGs Recovered	Predicted Genus
1GB	Baby-1-B	0	N/A
	Baby-1-W1	0	N/A
	Baby-2-1-B	1	*Paracoccus*
	Baby-2-2-B	3	*Bacillus* *Moraxella* *Pseudomonas*
	Baby-2-2-W3	0	N/A
	Baby-2-2-M8	2	*Streptococcus* *Neisseria*
	Baby-4-B	0	N/A
	Baby-5-B	0	N/A
	BSS-5-W3	0	N/A
	Positive Control	6	*Streptococcus* *Micrococcus* *Cutibacterium* *Staphylococcus* *Acinetobacter* *Corynebacterium*
5GB	Baby-1-B	0	N/A
	Baby-1-W1	0	N/A
	Baby-2-1-B	3	*Paracoccus* *Paracoccus* *Paracoccus*
	Baby-2-2-B	4	*Moraxella* *Pseudomonas* *Bacillus* *Psychrobacter*
	Baby-2-2-W3	1	*Cutibacterium*
	Baby-2-2-M8	5	*Rothia* *Veillonella* *Neisseria* *Granulicatella* *Prevotella*
	Baby-4-B	0	N/A
	Baby-5-B	1	*Streptococcus*
	Baby-5-W3	0	N/A
	Positive Control	6	*Streptococcus* *Staphylococcus* *Corynebacterium* *Cutibacterium* *Acinetobacter* *Micrococcus*

3.7 Discussion

3.7.1 Initial Diluent Testing Experiments

There was effective cell suspension and DNA recovery from both diluents tested. PBS is more widely used than water in molecular techniques, due to more longitudinal research on its use [371, 376]. DNA yield was lower than expected for spiked swabs and may be due to presence of the swab transport medium (charcoal amies agar) during the DNA extraction, which is known to inhibit molecular techniques. This could be overcome by including a prolonged heated lysis step during the DNA extraction [399, 400], which was subsequently explored (section **3.4.2**). Negative controls resulted in low DNA recovery (lower than the test samples), suggesting the presence of a ‘kitome’ - background DNA contaminants found in DNA extraction kit reagents and consumables [331, 401-404]. This was monitored

in subsequent DNA extractions on adult volunteer and PEARL study skin swabs (section **3.4.3**), as this could have interfered with microbiome data interpretation from low biomass samples [298, 403, 405-409].

In agreement with other studies, there was effective viable cell recovery from both diluents, with no significant loss of viable numbers [371-375, 377, 378]. This highlights the suitability of both diluents to collect viable bacterial cells for culturing alongside the DNA extraction. Results further demonstrated better survivability of *E. coli* compared to *P. aeruginosa* and *S. aureus* in all diluents. As there was no observed significant difference in DNA extraction and culturing effectiveness between PBS and UMQW, and PBS appears more widely used in molecular techniques [371], PBS was selected for lysis optimisation.

3.7.2 Lysis Method Experiments

Comparison of a variety of lysis procedures identified the effectiveness of method 6, which incorporated a combined approach using both overnight heated enzymatic (lysozyme) and mechanical (bead beat) lysis methods. This resulted in sufficient yield of HMW DNA from both Gram-positive and Gram-negative bacteria, and was the most cost effective due to the cheaper lysozyme used. Previous work has indicated that the type of enzyme and mechanical intensity is important for lysis of certain bacterial species [387, 388, 410]; however, the combined use of both mechanical and enzymatic lysis resulted in effective DNA extraction from both Gram-positive and Gram-negative bacteria. The addition of Agarase did not significantly alter DNA recovery. Overall, the Promega Maxwell® RSC 48 instrument with method 6 was able to reliably extract DNA with a high enough yield and molecular weight for both long and short read SMS. Other platforms are available, and this method is likely to be transferable. This study did not aim to benchmark extraction kits, but this could be done in the future to refine this method further. The final optimised method included an overnight lysis with lysozyme, a further heated offboard lysis step and a bead beating lysis using a Tissue Lyser. This method was next validated on low biomass skin swabs from adults and infants.

3.7.3 DNA Extraction Validation with Adult Volunteer and PEARL Study Skin Swabs

Given the low biomass of skin microbiota, some adult skin swabs yielded very low/absent DNA concentrations and paired cultures also indicated low bacterial burden. Individual variations when swabbing (pressure, direction, frequency) can

affect the yield of DNA and viable bacteria, but these are difficult to control [13]. Alternative skin collection methods include pad scrubbing, cylinder suspension and tape-stripping. Some of these methods have previously been shown to improve DNA yield and viable bacteria compared to the swabbing method [391, 411-413]. However, the swabbing method was used here as the PEARL study had already established this method within their sample collection protocol. Swabbing was initially selected due to its commonality for collecting skin samples and non-invasive nature, suitable for neonates who have an underdeveloped skin structure [235, 237]. The lack of cultures grown from some infant swabs may be due to the long-term storage of them in the -80°C freezer, which can reduce bacterial cell viability [414]. However, we were still able to determine skin swab taxonomic profiles using SMS, allowing microbiota to be identified. A difference in sensitivity was identified between platforms for samples with low amounts of DNA. Some adult swabs did not produce data using the ONT platform but did generate bacterial cultures. As the ONT platform requires more input DNA to generate data than Illumina platforms [363], the inability to generate data for some samples was not surprising as skin swabs can be low biomass [331, 364]. This indicated that ONT was less sensitive than Illumina. However, increasing the overnight incubation time did improve DNA yield, and Illumina sequencing generated data for all samples. Overall, the results show a complimentary pairing of sequencing and cultures to identify skin microbiota from skin swabs.

Data were generated for most samples from both Illumina and ONT platforms, which presented similar microbiota profiles from skin swabs and cultures. Typical adult skin microbiota (phyla Pseudomonadota, Actinomycetota, and Bacillota) [7, 15, 37] and infant skin microbiota (phyla Bacillota, Actinomycetota, Pseudomonadota, and Bacteroidota) [1] were detected. Infant profiles demonstrated less microbial diversity than adults, as has been shown in previous studies [18]. Infant skin did exhibit more *Streptococcus* species than adult skin, which agrees with previous work demonstrating a predominance of *Streptococcus* species in early age, decreasing over time [1, 18]. Interestingly, sequencing of swabs from infant skin identified *Bifidobacterium* species, which are not typical skin residents, but rather maternal and infant gut residents also found in breast milk [396, 397]. Given the paired cultures did not result in any *Bifidobacterium* isolates, this is likely to indicate transient transfer to the infants' skin through breast feeding. The infants with

available metadata that showed *Bifidobacterium* presence on the skin were all breast fed at some point between birth and month 4.

Whilst skin is a relatively low biomass environment, we did not need to include any methods to mechanically deplete human DNA or selectively enrich microbial DNA before SMS, which have been needed in some other studies on low biomass samples [300]. These enrichment approaches do not reliably target all species, will not remove all human reads [300], and can skew the resulting genomic profiles [301]. Depletion can also result in some loss of bacteria [300], and further steps are required for downstream analysis. In this described method, enough microbial data were generated, and human DNA reads were depleted computationally. This precluded the need for any additional steps that may introduce biases and skew skin microbiota profiles.

Both Illumina and ONT sequence data allowed identification of all ATCC positive control species, with a clear reduction in read number across the dilution series. These results further demonstrate the effectiveness of both the extraction method and utilisation of both sequencing platforms. Inclusion of a commercially available mixed community positive control, with a known cell concentration, is important for standardising the extraction process, and serial diluting the positive control can determine the limit of detection [415]. This was also helpful when comparing different sequencing runs and sample sets, allowing more robust comparisons to be made. Definition of a limit of detection was attempted for DNA concentration and read number required for effective SMS. Several swab samples did not yield quantifiable DNA, but usable reads were produced for taxonomic profiling. Therefore, no obvious cut-off for a limit of detection was determined, and indeed there is also no ‘defined’ limit identified in the literature for low biomass samples, such as skin swabs.

Some background contamination was identified in the negative controls. Contamination commonly occurs in metagenomic studies, especially those with low biomass samples [416]. Several studies have identified contamination sources occurring from neighbouring samples and the ‘kitome’ [405, 416]. Contamination within a dataset can be identified and removed using bioinformatic techniques [407, 409, 417-419] although low biomass samples have a higher risk of true microbial microbiota members being removed [420]. Given the background contaminants in the controls were at a very low level and mostly represented species not seen in the

test samples, they were not removed as they had a negligible impact on the profiles produced. To address contamination in subsequent work, the Illumina and ONT libraries for low biomass samples were prepared separately from high biomass samples as recommended in the literature [405, 416].

Although this work identified a significant difference in AT and GC content of samples sequenced on both platforms, from a profiling perspective, the data indicated that they were highly comparable, agreeing with other work [421]. Some studies have demonstrated platform biases in the recovery of GC and AT rich microbial species [422, 423]. The Illumina platform can effectively determine diversity within samples, but the associated short reads reduce the chance of complete genomes being reconstructed [424]. In contrast, the ONT platform is less sensitive and has an increased chance of high raw error rates [395], but can result in the reconstruction of complete genomes [424]. Therefore, using a combination of the two should improve complete reconstruction of genomes, microbial identification and reduce biases.

Generation of 5 Gb Illumina data per skin swab was determined to be suitable for microbial species profiling, producing a limited number of MAGs. MAGs are important for in-depth functional information [327] and indicate genome quality [351, 425-427]. They can be used to identify novel taxa and allow further comparison with whole genome sequence data from isolates. This method was found to be compatible with both Illumina and ONT platforms. Combining a higher sequencing depth with ONT, there is potential to improve the number and quality of MAGs recovered [327, 428, 429].

3.8 Conclusion

An optimised medium-throughput DNA extraction, SMS, and analysis approach effectively characterised the skin microbiota from both adults and infants. This method can be applied for in-depth analysis of cohort studies (PEARL study), allowing identification of taxonomic and functional temporal changes of mothers and infants with comparisons to other body sites (gut and vagina); work that can be found in Chapter 4. Robust microbial profiling, particularly in less well studied niches such as the skin, is important for the development of methods to alter microbiome compositions, with a goal to maintain skin health and prevent and treat skin disease.

CHAPTER 4

THE COMPOSITION OF THE DEVELOPING INFANT SKIN MICROBIOME

CHAPTER 4 OVERVIEW

In this chapter, the optimised skin microbial DNA extraction and Shotgun Metagenomics Sequencing (SMS) method from Chapter 3 were applied to samples from a healthy participant cohort (the Pregnancy and Early Life (PEARL) study) to analyse the infant skin composition (bacteriome focused with an insight into the archaeome) and temporal changes, diversity changes, and microbiome influencing factors (pregnancy term, delivery mode, feeding mode). Multi body site (skin, stool and vaginal) comparisons were made for microbial sharing over time between mothers and their infants, to investigate skin microbiome establishment. Genus, species and strain level analysis was conducted. Metagenome Assembled Genomes (MAGs) from both maternal and infant datasets were identified, which generated a PEARL genome catalogue. The strain level data formed the basis for subsequent future work discussed in Chapter 5 that will explore the potential functional capacities of important skin microbiota members. These data and findings may help guide targeted probiotic interventions to prevent and treat skin disease in early life and maintain skin health. I performed the vast majority of laboratory experiments and bioinformatics analysis in this chapter: I processed skin and vaginal swabs and extracted DNA from them using the extraction protocol that I optimised in Chapter 3; I performed culturing of these samples alongside the DNA extraction, which I also extracted using the optimised extraction method; I performed the bioinformatics analysis for the skin, vaginal, stool and culture metagenomes that resulted from the SMS data (taxonomic profiling, MAGS, data visualisation and statistical analysis). Exceptions of work that were performed by other scientists included the stool DNA extractions (performed by the QIB PEARL study team), the shotgun metagenomics library preparation and sequencing (performed by the QIB sequencing team and Genewiz), the data clean up and strain level analysis (performed by the QIB bioinformatics support team), the archaeal taxonomic profiling (performed by Dr. Charlotte Neumann; I visualised and performed statistical analysis for the archaeal taxonomic profiles).

Due to the large and complex dataset curated, some figures have been adapted with an interactive element, which can be viewed in the supplementary files provided; a hyperlink to the interactive versions are also available within the figure legends.

4.1 Introduction

The composition and functional role of the infant skin in maintaining skin health is an emerging area of research. While age has been identified as a key factor in the establishment of the skin microbiome, it is also transiently influenced by multiple intrinsic and extrinsic factors including delivery mode, labour, vaginal seeding, antibiotic use, parental factors, environment, skincare, probiotics, neonatal skin structure and function, diet, breastfeeding, genetics, and gender [1, 11, 12, 14, 122, 129-131]. Despite these influences, research indicates that community assembly is largely driven by stochastic processes [123]. Profiling of the skin microbiome has identified many of these factors as important for skin health, while others have been linked to perturbations and skin disease [14, 16, 17, 22, 23, 39, 40, 121, 122, 124, 125, 127, 128]. However, recent research has focused on diseased infant skin [9, 14, 22, 24, 72-75], leaving gaps in our understanding of the healthy infant skin microbiota, particularly understudied skin communities, such as the archaeome [12]. Furthermore, few studies have used integrative approaches combining multiple techniques (such as culturomics, metagenomics and metabolomics) to robustly study the infant skin microbiome composition and function [20, 22, 39-43].

There is a pressing need for research to use an integrative approach to establish a baseline composition of the healthy infant skin microbiome, and to clarify the roles of these microbes in the skin health. This will aid in identifying microbiota altered by various intrinsic and extrinsic factors and deepen our understanding of their functional role in infant skin health and disease prevention.

4.2 Aims

The work presented in this chapter addresses thesis Objectives 2-5 (section **1.7**), enabling investigation into the development and evolution of the healthy infant skin microbiome over time. It also explores the factors influencing the microbiome, the microbial sharing between mothers and their infants over time, and identifies MAGs, which may influence skin health; in turn addressing thesis Aims 1-4 (section **1.6**).

1. To establish a baseline for a healthy skin microbiome by analysing changes to the composition of the developing infant skin microbiome.
2. Investigate microbiome evolution and associated influencing factors by comparing the developing infant skin microbiome to the maternal skin, vaginal and gut microbiome at birth and into infancy.

3. Assess changes to the infant gut microbiome and its impact on the skin microbiome development; which could help determine nutritional interventions to treat and prevent skin disease in early life and maintain skin health during the life course.
4. Identify shared microbes between mothers and their infants over time.
5. Identify a PEARL genome MAGs catalogue, which will provide future insight into the functional importance of the shared microbes in infant skin health over time, and help formulate the next generation of oral probiotic, and LBP interventions to help predict how to alter a microbiome to prevent and treat skin disease and maintain skin health across the life course.

4.3 Objectives

1. Identify changes in the skin microbiome of infants over time from birth up to 1 year of age.
2. Identify common genera and species between the skin, gut and vaginal microbiome taxonomic profiles of mother and infant pairs over time that indicate microbe sharing.
3. Investigate the skin, gut, and vaginal microbiome compositional diversity changes over time, to genus and species level, across the maternal and infant pair datasets; whilst accounting for factors (delivery and feeding mode, and pregnancy term) that may significantly alter the diversity over time.
4. Investigate the skin, gut, and vaginal microbiome compositional relative abundances of the identified shared genera and species over time across the maternal and infant pair datasets; whilst accounting for factors (delivery and feeding mode, and pregnancy term) that may significantly alter their relative abundances over time.
5. Identify strain level sharing of the identified shared genera and species between mother and infant pairs.
6. Identify the MAGs generated from the eight maternal and infant pair datasets.

4.4 Experimental Methods

I performed my optimised skin microbial DNA extraction method on PEARL maternal and infant paired skin swabs (n=85 and n=89 respectively), and maternal vaginal swabs (n=85). Maternal and infant paired stool samples (n=81 and n=80 respectively) were DNA extracted by the PEARL team at QIB using their protocol (section **2.2**). This sample set was collected from 14 mothers and 15 infants (one

set of twins). The infant skin swab and stool samples selected for extraction were collected at Birth (B), 1 week (W1), 3 weeks (W3), 4 months (M4), 8 months (M8), and 12 months (M12) after birth. The maternal skin and vaginal swabs selected for extraction were collected at Trimesters 1-3 (T1-T3), B, W1, W3, and M4; the maternal samples were only collected up to month 4 due to budget and feedback from focused discussion groups with mothers regarding pragmatism. For PEARL cohort and study details see section **2.1**.

I processed all skin and vaginal swabs into 1 x PBS (due to previous results from initial diluent optimisation experiments in Chapter 3) and cultured them aerobically and anaerobically in parallel with the swab DNA extraction, which were scraped into combined glycerol stocks. Culturing was not performed alongside the stool DNA extractions as this was not part of the PEARL team protocol (section **2.2**). This resulted in 48 infant and 56 maternal skin cultures, and 69 maternal vaginal cultures. Skin, vaginal and culture samples totalled 432 samples (plus 15 positive and 15 negative controls = 462 samples). Infant and maternal stool samples totalled 161 samples (plus 2 positive and 16 negative controls = 179 samples). This resulted in 593 samples (plus 48 controls), across eight datasets (Baby Skin Swab [BSS], Mother Skin Swab [MSS], Baby Skin Culture [BSC], Mother Skin Culture [MSC], Baby Stool [BST], Mother Stool [MST], Mother Vaginal Swab [MVS], and Mother Vaginal Culture [MVC]). See **Figure 4.1** for a summary of the sample numbers collected and sequenced, for each maternal and infant group, across the skin, stool and vaginal body sites. I performed a Qubit assay (section **2.3**) on all DNA extracts before SMS Illumina libraries were prepared by the QIB sequencing team and samples sent to Genewiz for SMS on the Illumina platform. Libraries were prepared for each sample type separately to reduce the chance of contamination occurring from different sources that vary in biomass and extraction efficiency. The libraries generated for the skin and vaginal samples were split across two library pools, and the stool samples were split across eight library pools (section **2.5.1**).

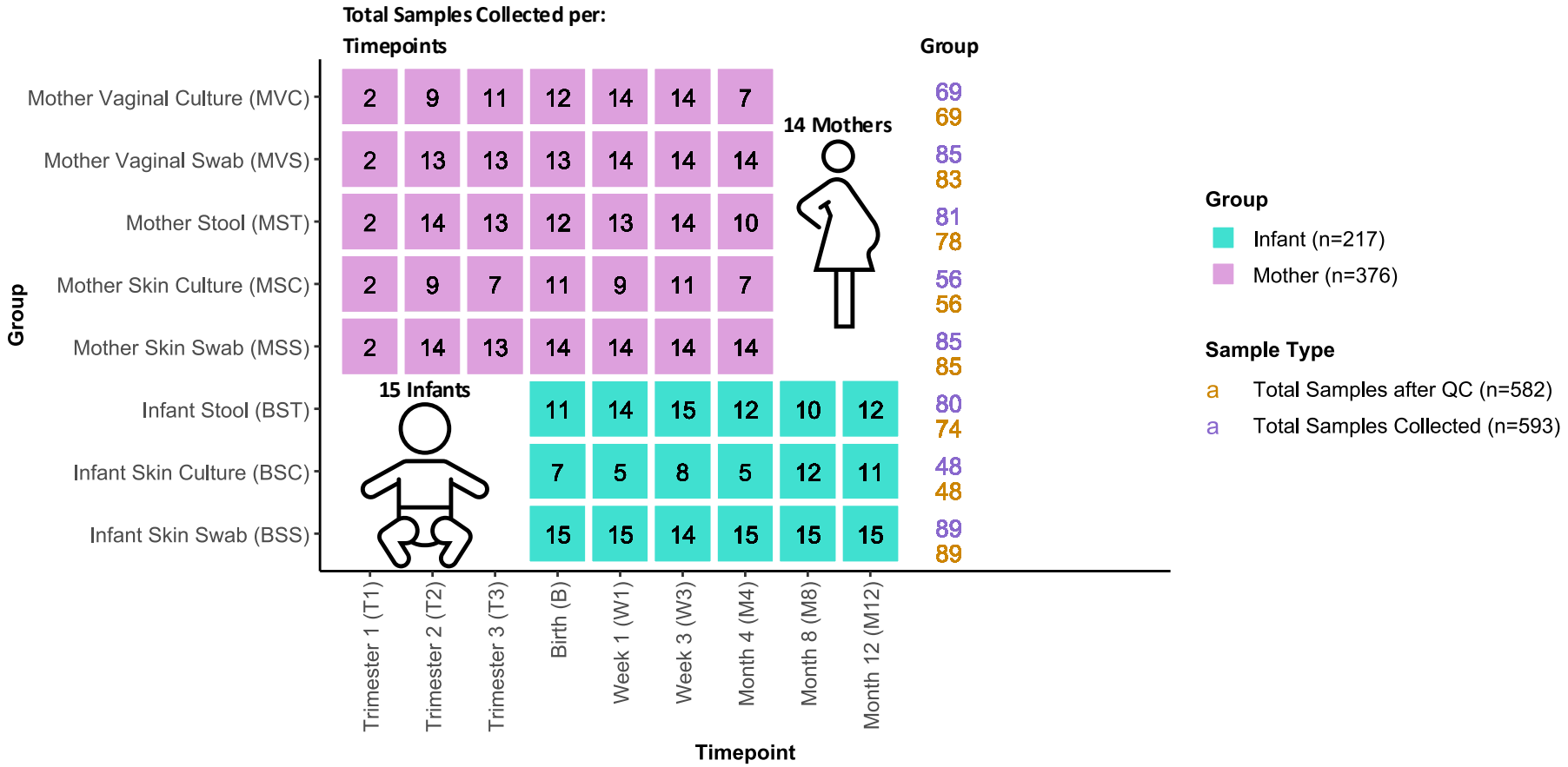


Figure 4.1 Summary of the PEARL Study Total Sample Numbers Collected and Sequenced. This Figure displays: the total number of mother and infant participants (n=14 and n=15 respectively) as coded in the animated icons; the total number of samples collected for mothers and infants (n=376 and n=217 respectively) as colour coded in the key under “Group”; the total number of samples collected before and after sequencing Quality Control (QC) (n=593 and n=582 respectively) as colour coded in the key under “Sample Type”; the total number of samples collected for each maternal and infant group, across the skin, stool and vaginal body sites, at each timepoint.

4.5 Data Analysis

The QIB core bioinformatics team cleaned the SMS data, and I performed taxonomic profiling to genus and species level, whilst the core bioinformatics team generated the strain level data. I assembled the MAGs and visualised the data using a variety of plot styles. The infant archaeome was also explored; Dr. Charlotte Neumann generated the archaeome profiles and I visualised and performed statistical analysis for the archaeal taxonomic profiles (sections **2.5.2; 2.5.3; 2.5.4; 2.5.5**).

4.5.1 Genus and Species Level Analysis

I plotted rarefaction curves to determine whether the sequence depth allowed each sample to reach a 'full' microbial profile. I visualised taxonomic profiles of the eight datasets as stacked bar plots and pie charts to show microbiome compositions and identify sharing of microbes between datasets. The stacked bar plots provided an overview of the top five genera and species (plotted as relative abundance (RA)) within individuals from the eight maternal and infant pair datasets and the content of the controls. There were no species detected in 13 samples (MVS-14-M4, MVS-13-W3, MVS-13-W1, MVS-13-T3, MVS-12-M4, MVS-11-T2, MVS-8-M4, MVS-6-B, MVS-3-M4, MVS-2-W1, MVS-1-W3, MVC-14-M4, BSC-2-2-M8) and were therefore removed from the plots; these samples were mostly vaginal samples that presented with blood and discharge, which is known to introduce background contaminants that interfere with sequencing. I displayed the common microbes across the maternal and infant datasets and within each dataset over time in pie charts (relative abundance counts were used from all individuals that were combined to calculate the mean relative abundance for each of the common genera and species at each timepoint across all datasets and within each dataset). I displayed the archaeal profiles in bubble plots to show the relative abundance of species within individuals from the infant and maternal skin swab datasets, and changes over time within and across the infant and maternal skin swab datasets.

I examined the relative abundant differences of the shared genera and species between maternal and infant paired datasets, accounting for individual differences and factors over time. I accounted for certain factors including pregnancy term (Full Term [FT] and Late Pre-Term [LP]), delivery mode (Vaginal Delivery [VD] and Caesarean Section [CS]) and feeding mode (Breast [BR] and Breast/Formula [BR/FO] milk), which I examined over time and visualised using spaghetti plots. I

examined the overall compositional diversity differences within and between the eight maternal and infant pair datasets (accounting for individual differences and pregnancy term, delivery and feeding mode factors) over time to genus and species level, measuring alpha (within sample diversity using the Richness and Shannon indexes) [430, 431], beta (between sample diversity using the Bray-Curtis index) [430, 432], and differential abundance (abundance changes using Microbiome Multivariable Association with Linear Models [MaAsLin2]) [433] diversity metrics. I used the R Vegan (V.2.6-6.1) [434] and MaAsLin2 [433] packages for this analysis. I chose the Alpha diversity indexes for their ability to encompass community richness and community diversity. I chose the beta diversity index for its ability to account for both the presence of each genus and its abundance percentage. I chose differential abundance test as it accommodates microbiome longitudinal data and adjusts for multiple samples whilst accounting for individual variation (random effects). I visualised this diversity analysis using box-violin (alpha diversity), Non-metric Multi-dimensional Scaling (NMDS) plots (beta diversity), and heatmaps (differential abundance).

I performed statistical analysis to determine whether there was a significant difference in the relative abundance of shared microbiota and compositional diversity (alpha, beta, differential abundance) over time within the eight maternal and infant paired datasets; also accounting for pregnancy term, delivery and feeding mode factors. I performed the Analysis of Variance (ANOVA) test [435] in R using the Vegan package (V.2.6-6.1) [434] to identify whether there was an overall significant difference in the: (i) relative abundance of the shared microbiota (to genus and species level) over time, (ii) relative abundance of archaea (to species level) over time, (iii) alpha diversity over time; within each dataset, whilst accounting for individual variation and pregnancy term, delivery and feeding mode factors. I further adjusted the ANOVA test [435] accordingly for multiple testing using the Estimated Marginal Means (EMMeans) R package (V.1.10.4) [436], which is a posthoc test allowing for multiple comparisons to be made between individual timepoints, microbes, and factors accounted for (pregnancy term, delivery and feeding mode). As an overarching view of microbial changes over time was warranted, the ANOVA test seemed most appropriate than other tests, such as the Mixed Model for Repeated Measures (MMRM) test (using the R MMRM package) which is advantageous for larger and more complex datasets [437].

I performed the Permutational Multivariate ANOVA (PERMANOVA) test [438] with Adonis2 (a statistical function used to perform the PERMANOVA test) and Bray-Curtis using the R Vegan package (V.2.6-6.1) [434] to identify whether there was an overall significant difference over time in beta diversity between each dataset, whilst accounting for individual variation and factors. I performed the PairwiseAdonis package (V.0.4) [439] as a posthoc test alongside PERMANOVA to allow for multiple comparisons to be made between individual timepoints and factors (pregnancy term, delivery and feeding mode) accounted for. I chose PERMANOVA for its robustness in testing the difference between groups based on measures of beta-diversity, and adjustment for covariates.

MaAsLin2 is a multivariable regression and mixed effect model [433] that I performed to test significant positive and negative associations (providing directionality and strength of associations) between microbial abundance over time, whilst accounting for individual variation, and allowing for multiple comparisons to be made across timepoints and factors accounted for (pregnancy term, delivery and feeding mode).

4.5.2 Strain Level Analysis

The QIB core bioinformatics team used Strainphlan to perform and visualise metagenomic strain-level population genomics on this large metagenomic dataset by profiling microbes from known species with strain level resolution and providing comparative and phylogenetic analysis of strains retrieved from metagenomic samples. Strainphlan took a set of metagenomic samples and reference genomes and then reconstructed the sample-specific strains of all species present in the sample and the genome-specific strains from reference genomes. Strain taxonomic profiling and sharing of microbiota between mother and infant pairs were visualised using tile plots, dot plots, stacked bar plots and tables. Samples available for strain level profiling were visualised using a tile plot. Dot plots were used to visualise the number of strains shared between mother and infant pairs over time across the eight datasets. Stacked bar plots were used to display the overall strain (same or different strain of a species) transmissibility between mother and infant pairs across the different body sites, and the strain transmissibility accounting for time, pregnancy term, delivery and feeding mode. To avoid confusion with potential strain persistency effect, when the species or strain was shared between maternal and infant sample pairs in multiple timepoints, the transmission events were counted only once. Finally, I used tables to display the strains that were shared between any

mother and infant pairs and present in the culture samples. A strain sharing event was defined as the presence of at least one shared strain between pairs of samples from the same individual or from the same family unit. Species transmissibility was defined as the ratio of the number of sharing events with strain level sharing events over the number of families with species level sharing events. Fisher tests were performed to identify whether there were significant differences overall in the strain transmissibility between mother and infant pairs at different body sites (BSS, BST, MSS, MST and MVS); also accounting for factors (timepoint, pregnancy term, delivery mode and feeding mode). This test was chosen for its ability to compare distributions more robustly on smaller sample sizes.

4.5.3 MAGs Analysis

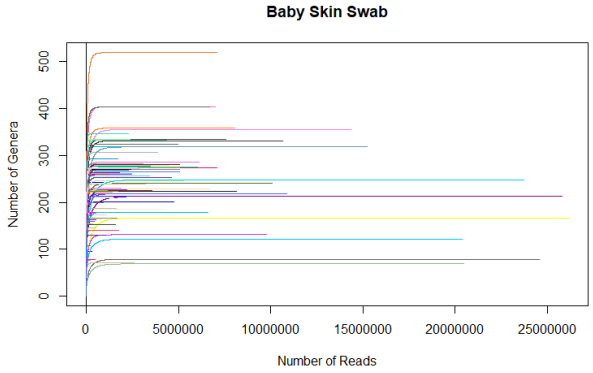
I ran the MAGs generated through Mashtree (V.1.2.0) [440], which uses Mash (V.2.0) [441] to generate PHYLogeny Inference Package (PHYLIP) [442] distance matrices to be used for neighbour function in PHYLIP software to generate trees. PHYLIP is a computational phylogenetics package for inferring evolutionary trees (phylogeny). Mashtree (V.1.2.0) [440] demonstrates clustering of genera together but it does not infer phylogeny. I used Mashtree (V.1.2.0) [440] here to purely display the genera in the dataset and identify whether the MAGs composition was in line with the taxonomic profiles. I uploaded the Mashtree output to Interactive Tree of Life (ITOl) (V.6.9.1) [443, 444] where the Tree of Life (V.1.0) [445] function was used to generate a tree of all the MAGs.

4.6 Results

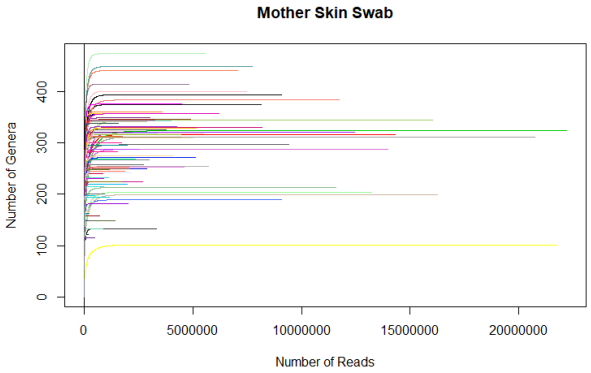
Sequencing resulted in between 8 and 11Gb of data per sample. Of the 641 samples, 29 failed QC, and were not used for further analysis; 18/29 failed samples were negative controls, 2/29 were vaginal swabs, and 9/29 were stools. Overall, 96% of samples generated useable data (582 samples plus 30 controls). Rarefaction curves of the eight maternal and infant pair datasets, to genus and species level, showed that all samples plateaued, suggesting sequencing had captured the complete microbiome composition. Thus, the sequencing depth was sufficient to generate comprehensive data for the skin, stool, vaginal and cultures samples (**Figure 4.1-4.2**).

Mother and Infant Genus Rarefaction Curve

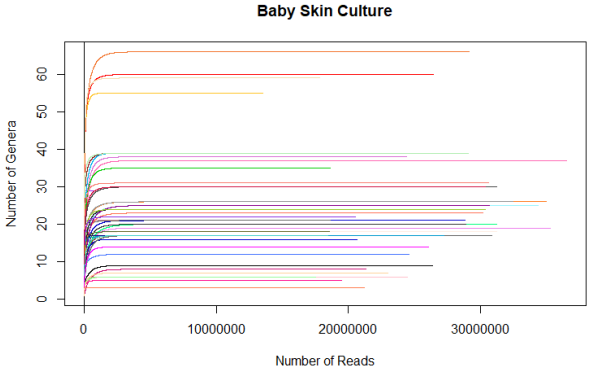
A



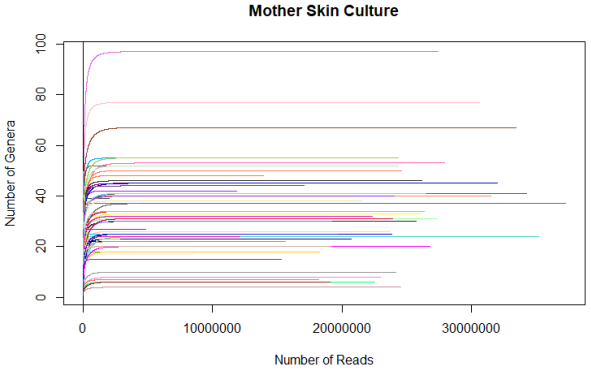
B



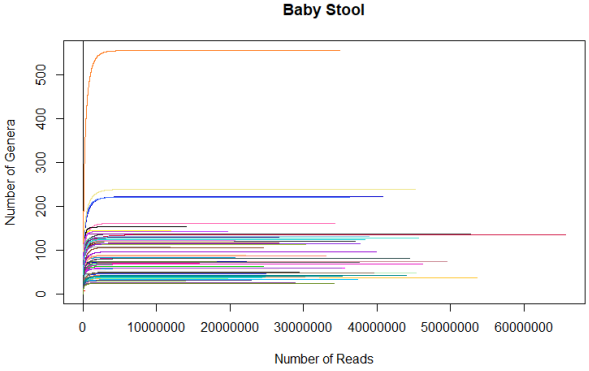
C



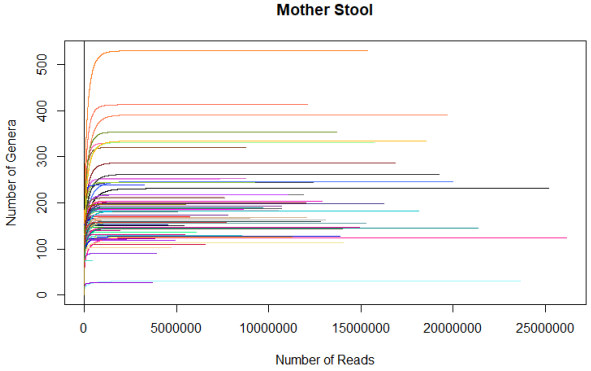
D



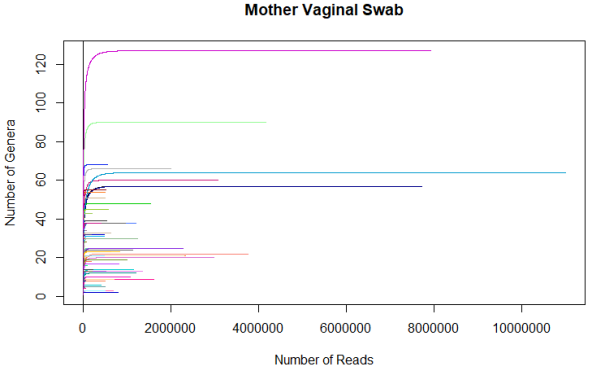
E



F



G



H

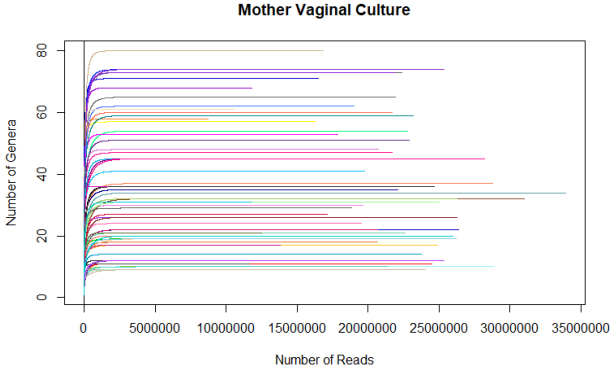
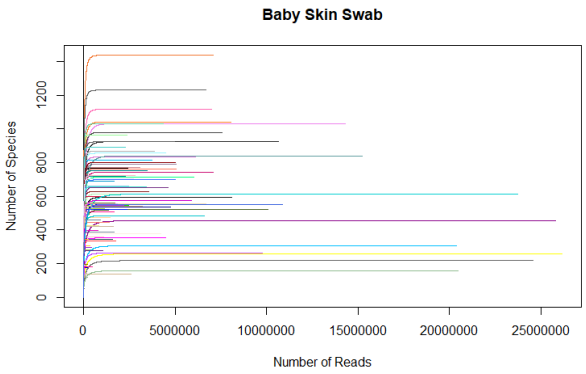


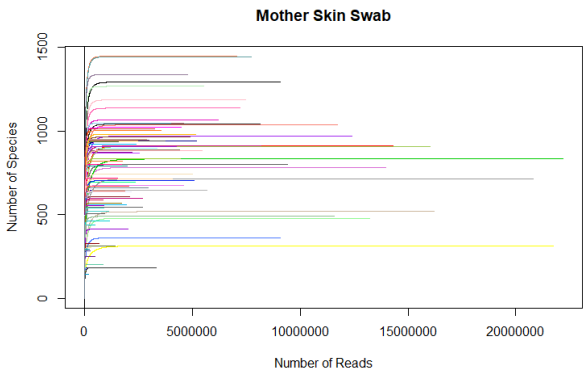
Figure 4.2. Rarefaction Curves of the Maternal and Infant Datasets at Genus Level. Shotgun metagenomics data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures) using shotgun metagenomics sequencing, generated on the Illumina platform. Number of reads of the taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Using the number of read values, rarefaction curves were generated. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. **A.** Number of genera detected reads for the infant skin swabs. **B.** Number of genera detected reads for the maternal skin swabs. **C.** Number of genera detected reads for the infant skin cultures. **D.** Number of genera detected reads for the maternal skin culture. **E.** Number of genera detected reads for the infant stool. **F.** Number of genera detected reads for the maternal stools. **G.** Number of genera detected reads for the maternal vaginal swabs. **H.** Number of genera detected reads for the maternal vaginal cultures.

Mother and Infant Species Rarefaction Curve

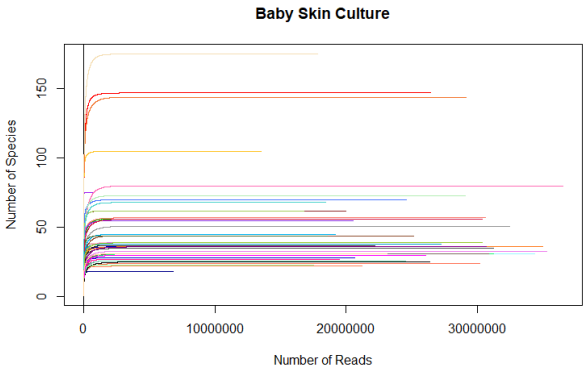
A



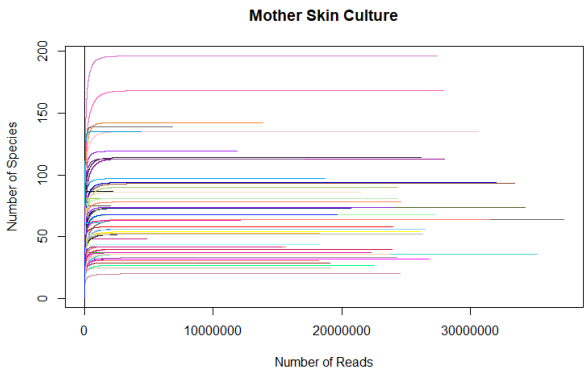
B



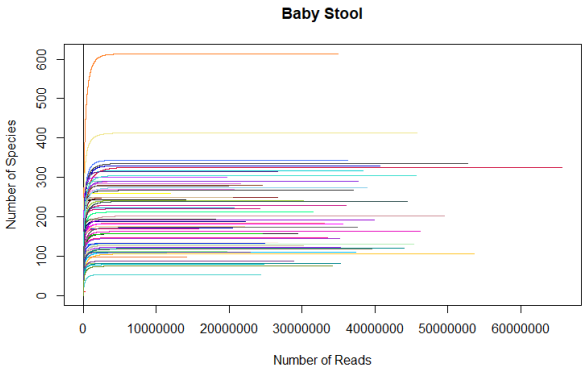
C



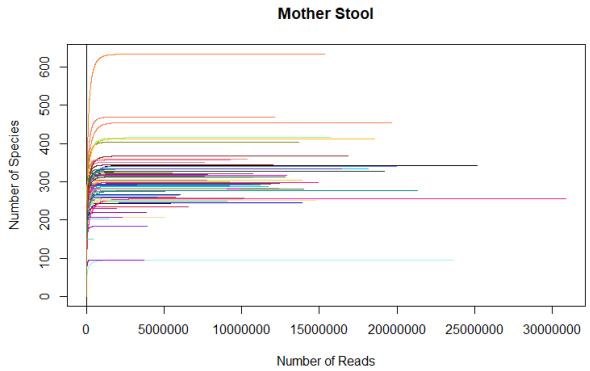
D



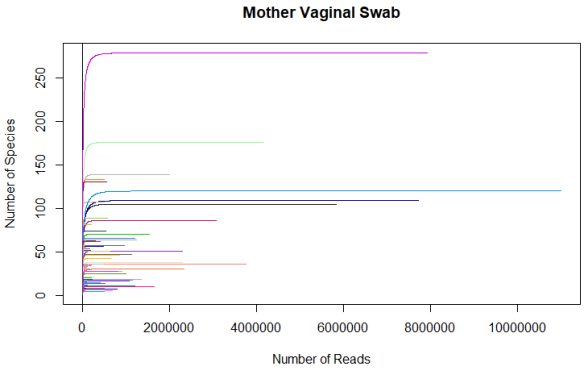
E



F



G



H

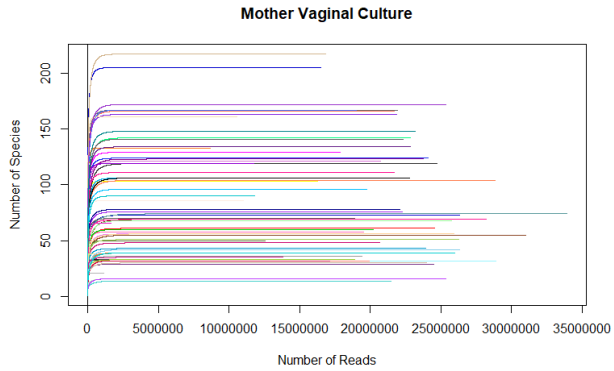


Figure 4.3. Rarefaction Curves of the Maternal and Infant Datasets at Species Level. Shotgun metagenomics data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures) using shotgun metagenomics sequencing, generated on the Illumina platform. Number of reads of the taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Using the number of read values, rarefaction curves were generated. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. **A.** Number of species detected reads for the infant skin swabs. **B.** Number of species detected reads for the maternal skin swabs. **C.** Number of species detected reads for the infant skin cultures. **D.** Number of species detected reads for the maternal skin culture. **E.** Number of species detected reads for the infant stool. **F.** Number of species detected reads for the maternal stools. **G.** Number of species detected reads for the maternal vaginal swabs. **H.** Number of species detected reads for the maternal vaginal cultures.

Positive controls displayed the expected microbiota from the ATCC skin microbiome whole cell mix (*A. johnsonii*, *Cor. striatum*, *M. luteus*, *Cut. acnes*, *Sta. epidermidis*, and *Str. mitis*) and the M004Base and M008Base mix (*Akkermansia muciniphila*, *Bacteroides caccae*, *Bacteroides uniformis*, *Faecalibacterium prausnitzii*, *Alistipes shahii*, *Lachnospira elegans*, *Parabacteroides sp.* CT06, *Phocaeicola coprophilus*, *Phocaeicola dorei*, *Segatella copri* (formerly *Prevotella copri*), *Ruminococcus champanellensis* and *Ruthenibacterium lactatiformans*) (**Figure 4.3**). Some background contamination was detected in negative controls, although the number of reads was much lower than in test samples and most DNA fragments mapped against organisms not seen in test samples (**Figure 4.4**). One negative control to genus level and three to species level generated no profiles and were therefore removed from plots.

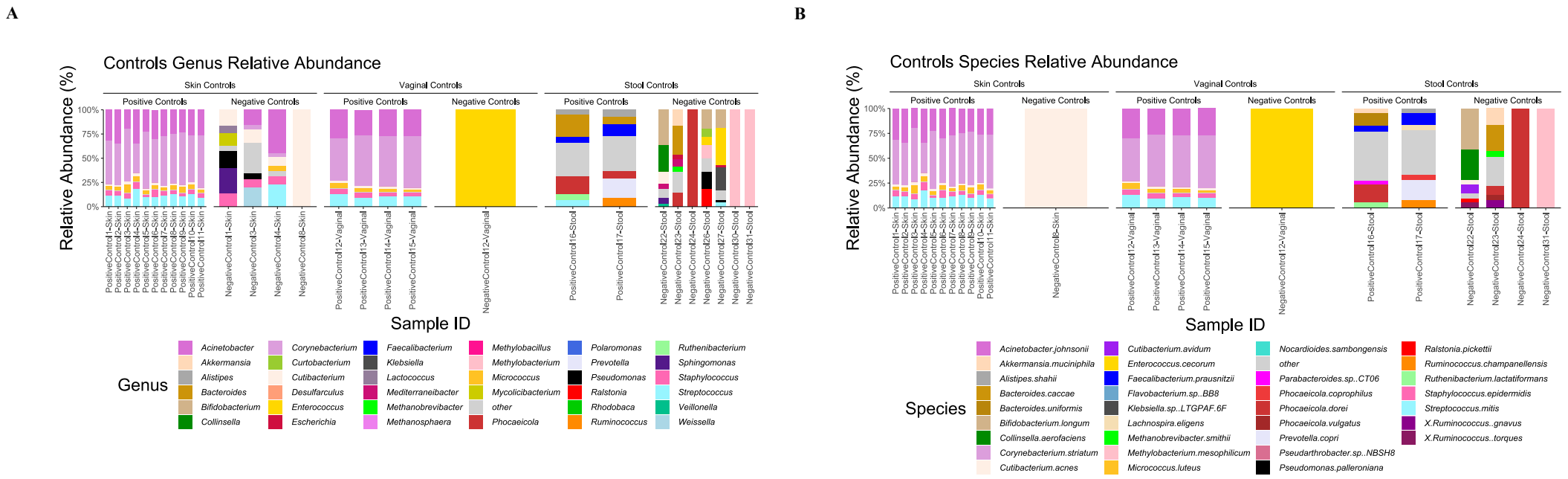
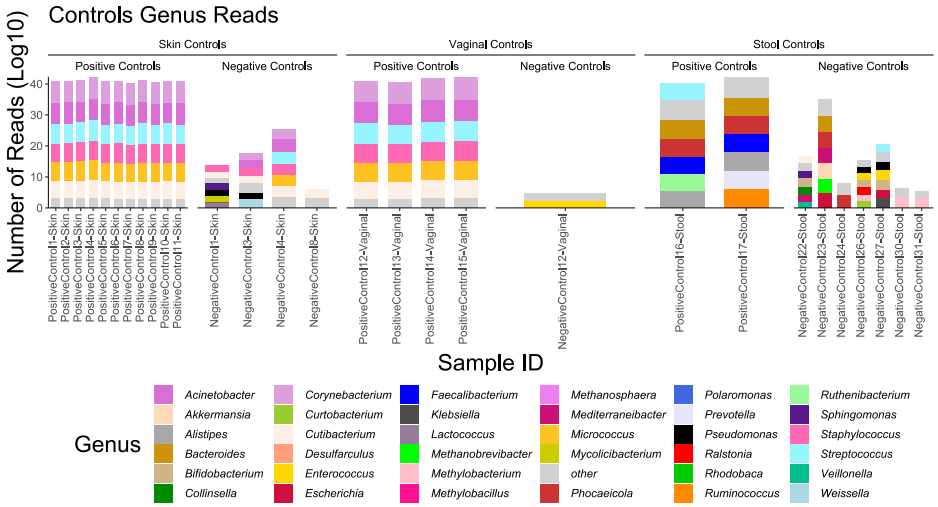


Figure 4.4. Taxonomic Relative Abundance Illumina Profiles of Controls to Genus and Species) Level. Shotgun metagenomics data was generated from the positive and negative control samples (skin positive (n=11) and negative (n=4) controls, vaginal positive (n=4) and negative (n=1) controls, stool positive (n=2) and negative (n=7) controls), that were run alongside the maternal and infant test samples, using shotgun metagenomics sequencing, generated on the Illumina platform. The relative abundance (%) taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bar plots were used to display the relative abundance (%) values, and the key allocates a unique colour to the genus and species represented in the stacked bars. **A.** Profiles show the relative abundance of the genera within each control sample. **B.** Profiles show the relative abundance of the species within each control sample.

A



B

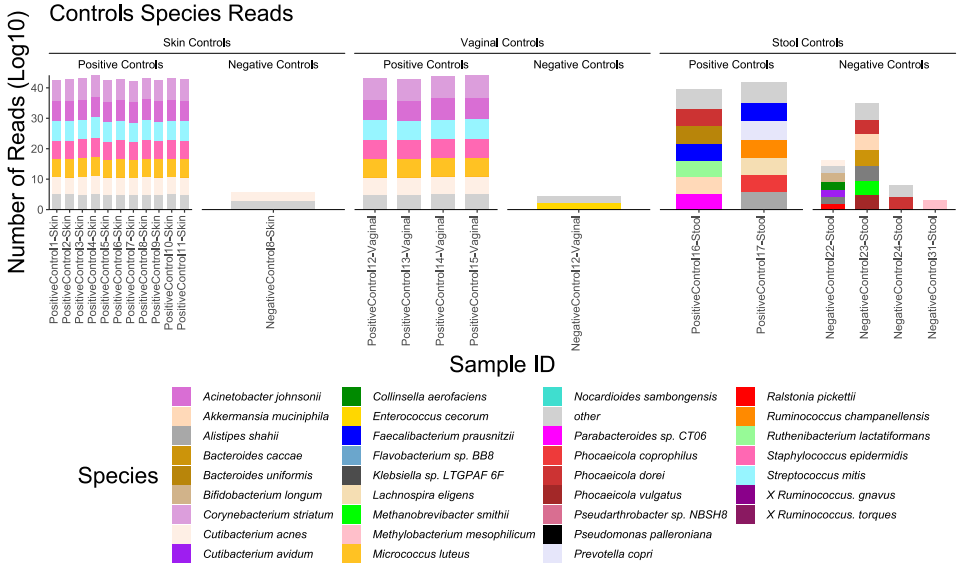


Figure 4.5. Taxonomic Read Illumina Profiles of Controls to Genus and Species Level. Shotgun metagenomics data was generated from the positive and negative control samples (skin positive (n=11) and negative (n=4) controls, vaginal positive (n=4) and negative (n=1) controls, stool positive (n=2) and negative (n=7) controls), that were run alongside the maternal and infant test samples, using shotgun metagenomics sequencing, generated on the Illumina platform. The number of reads of the taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bar plots were used to display the number of read values, and the key allocates a unique colour to the genus and species represented in the stacked bars. **A.** Profiles show the number of reads of the genera within each control sample. **B.** Profiles show the number of reads of the species within each control sample.

4.6.1 The Composition and Baseline of a Healthy Infant Skin Microbiome Over Time

The 'common' or 'core' microbes across the infant skin swab dataset over time were identified to genus and species level (**Figures 4.6-4.7**). All infants included in this analysis were healthy, providing a baseline for the healthy infant skin microbiome from Birth (B) to Month 12 (M12). The common genera and species identified within the infant skin swab dataset over time included *Acinetobacter* (*A. baumannii*, *A. johnsonii* and *A. lwoffii*), *Bifidobacterium* (*B. longum*), *Corynebacterium*, *Cutibacterium* (*C. acnes*), *Epilithonimonas* (*E. vandammei*), *Fervidobacterium* (*F. pennivorans*), *Gemella* (*G. haemolysans*), *Gordonia*, Haemophilus, *Micrococcus* (*M. luteus*), *Moraxella* (*Moraxella osloensis*), *Neisseria*, *Paracoccus* (*P. yeei*), *Pseudomonas* (*P. putida*), *Rothia*, *Staphylococcus* (*S. epidermidis* and *S. hominis*) and *Streptococcous* (*S. mitis*, *S. pneumoniae* and *S. salivarius*). Relative abundance trends of these genera and species were identified. These common shared genera and species over time at each timepoint represented between 78% and 86% and 52% and 63% respectively of the microbiota across all individuals, leaving between 14% and 22% and 43% and 48% respectively representing the "Other" microbes that were not common across all timepoints.

The relative abundance of *Acinetobacter* genus and species across all infants (*A. baumannii*, *A. johnsonii* and *A. lwoffii*) fluctuated over time and was highest at M12. The *Acinetobacter* genus (and to species level, *A. johnsonii*) was in higher abundance at M12 compared to the other common microbes. *Staphylococcus* genus and species (*S. epidermidis* and *S. hominis*) relative abundance fluctuated over time and was higher between B and Week 1 (W1), before decreasing over subsequent timepoints to Month 8 (M8), when a slight increase was observed at M12. *Streptococcus* genus and species (*S. mitis*, *S. pneumoniae* and *S. salivarius*) relative abundance was found to increase from B to Month 4 (M4), at which point it decreased up to M12. There appeared to be a relationship between *Staphylococcus* and *Streptococcus* (particularly *S. epidermidis* and *S. mitis*). When *Staphylococcus* (*S. epidermidis*) relative abundance was higher from B to W1, *Streptococcus* (*S. mitis*) was lower. This relationship reversed from Week 3 (W3). Overall, to genus and species level, *Staphylococcus* started off in higher relative abundance than *Streptococcus*, which then reversed from W3. But at M12 *Streptococcus* relative abundance rapidly decreased and *Staphylococcus* started to increase again. There was a significant difference in relative abundance changes over time of

Streptococcus (*S. mitis* and *S. salivarius*) and *Staphylococcus* on infant skin, and between *Str. mitis* and *Sta. hominis*, *Str. pneumoniae* and *Sta. epidermidis*, and *Str. salivarius* and *Sta. epidermidis* (**Figures 4.8. A.** & **B.**). These relative abundance trends correspond to the individual infant taxonomic profiles (to genus and species level) over time (**Figures 4.9-4.10**). The trends at M12 of *Acinetobacter*, *Staphylococcus* and *Streptococcus* genus and species on infant skin, resembled the individual maternal skin taxonomic profiles, suggesting a starting point for transition to an adult microbiome. Overall, *Acinetobacter* and *Staphylococcus* was higher in abundance on the mothers' skin than infants, and *Streptococcus* was lower on the mothers' skin than infants. Other skin microbes also suggesting this transition point across the infants are *Epilithonimonas* (*E. vandammei*), *Paracoccus* (*P. yeei*) and *Pseudomonas* (*P. putida*), which exhibited relative abundance fluctuations over time being highest from M8 to M12.

Abundant culturable microbes were also identified from both maternal and infant skin, such as *Micrococcus* (*M. luteus*) and *Staphylococcus* (*S. hominis*, *S. warneri*, *S. epidermidis* and *S. pasteuri* and *S. capitis*) (**Appendix 14.; Appendix 15.)**. These higher abundance species were present with other lower abundance species of the same genus over time in the infant and maternal skin (*Micrococcus sp. KBS0714*, *Micrococcus sp. A1*, *Micrococcus sp. A7*, *Micrococcus sp. MG2010D12*, *Micrococcus sp. 28*, *Micrococcus sp. V7*, *Staphylococcus lugdunensis*, *Sta. haemolyticus*, *Sta. aureus* and *Staphylococcus sp. M0911*).

Over time, the infant skin microbiome appeared most influenced by the maternal skin as identified in the individual taxonomic profiles (**Figures 4.9-4.10**). But there were some associations with the infant skin and gut microbiome, and other maternal body sites, i.e. vaginal (*Lactobacillus crispatus*, *Bif. breve, Bif. longum* and *Bif. bifidum*) and stool (*Bif. longum and Bif. bifidum*), as well as associated factors, such as feeding (*Bif. breve* was identified in infant stools, but not maternal stools, suggesting transfer to infant gut via breastfeeding) and delivery (*L. crispatus*, *Bif. breve, Bif. longum* and *Bif. bifidum*) mode in the individual mother and infant taxonomic profiles (**Appendix 14.; Appendix 15.)** and confirmed metadata (**Table 2.2** in section **2.1.3**).

For example, one infant skin profile (BSS-1-B) identified *L. crispatus,* which was likely a transient acquisition from their maternal vaginal microbiome (MVS-1-B) at B where *L. crispatus* was in high abundance. Metadata (**Table 2.2** in section **2.1.3**)

identified a VD for this mother and infant pair. Some infant skin profiles identified *Bif. breve* (BSS-7-W1), *Bif. longum* (BSS-9-W3) and *Bif. bifidum*, which may be transiently acquired from the maternal stool (MST-9-W3,) and vaginal microbiomes (MVS -7-W1) where these *Bifidobacterium* species were present; metadata (**Table 2.2** in section **2.1.3**) for these mother and infant pairs identified different birthing modes (VD and CS), suggesting vertical transfer from breast milk instead, as these infants were fed breast milk from B to M12. Individual infant skin (BSS-14-W1) and mother vaginal (MVS-14-W1) taxonomic profiles of mother and infant pairs identified potential *S. epidermidis* transfer from the mothers' vagina to the infants' skin during delivery. Although, other mother (MVS-4-B) and infant (BSS-4-B) pairs underwent a CS, suggesting transfer via other means; these mother skin taxonomic profiles presented with *S. epidermidis*, suggesting main transfer was likely via skin-to-skin contact. *Bifidobacterium* (*B. longum*) relative abundance fluctuated over time across the infant skin, being highest at W3 and lowest at M12. This trend of *Bifidobacterium* being highest at W3 could suggest transient species from breast milk, as all infants were fed breast milk at some point between B and M12. However, there were no significant differences in the *Bifidobacterium* (*B. longum*) relative abundance changes over time when accounting for feeding mode (BR and BR/FO) (**Figures 4.8 C.** & **D.**). The relative abundance fluctuations could be indicative of infants that were fed a combination of breast and formula milk. The drop in relative abundance at M12 could indicate infants being weaned off breast milk. Some mother-infant pair individual taxonomic profiles (BSS-9-W3 and MST-9-W3) indicated that the initial transfer of *B. longum* may have occurred via vertical transmission from the mothers' stool to the infants' skin, likely during the birthing process where a VD occurred. There was also potential sharing of *B. longum* within individual infants between their skin and stools (BSS-12-W1 and BST-12-W1; BSS-9-W3 and BST-9-W3; BSS-14-W3 and BST-14-W3; BSS-11-M4 and BST-11-M4; BSS-14-M4 and BST-14-M4). However, this may also indicate transient and vertical transfer from breast milk to the skin and gut as all infants were fed breast milk from B to M12.

Oral specific microbiota (*Gemella* [*G. haemolysans*], *Neisseria* and *Rothia*) were also identified across infant skin over time and remained in low consistent relative abundance over time with slight fluctuations (*Neisseria* was highest at B, *Rothia* was highest at M4 and *Gemella* was lowest at M12). This trend could relate to consistent breast feeding from B to M12 where the transfer of oral microbes to skin occurred; *Neisseria* relative abundance was highest at birth, which may relate to the first

breast feed at birth. *Gemella* relative abundance was lowest at M12 and is consistent with infants being weaned off breast feeding.

Fervidobacterium genus and species (*F. pennivorans*) relative abundance on infant skin decreased gradually from B to M12. This trend could be suggestive of environmental hospital acquired microbiota due to the highest relative abundance at B and consistent gradual decrease over time. The relative abundance of the genera and species for *Cutibacterium* (*C. acnes*) and *Moraxella* (*M. osloensis*) fluctuated over time being highest at B, which could be suggestive of hospital acquired microbiota from skin-to-skin contact during heavy handling of the infant in the hospital setting. *Gordonia*, *Haemphilus*, *Micrococcus* (*M. luteus*), and *Corynebacterium* relative abundance were low and consistent over time with slight fluctuations, which may indicate stable low relative abundant microbes over time.

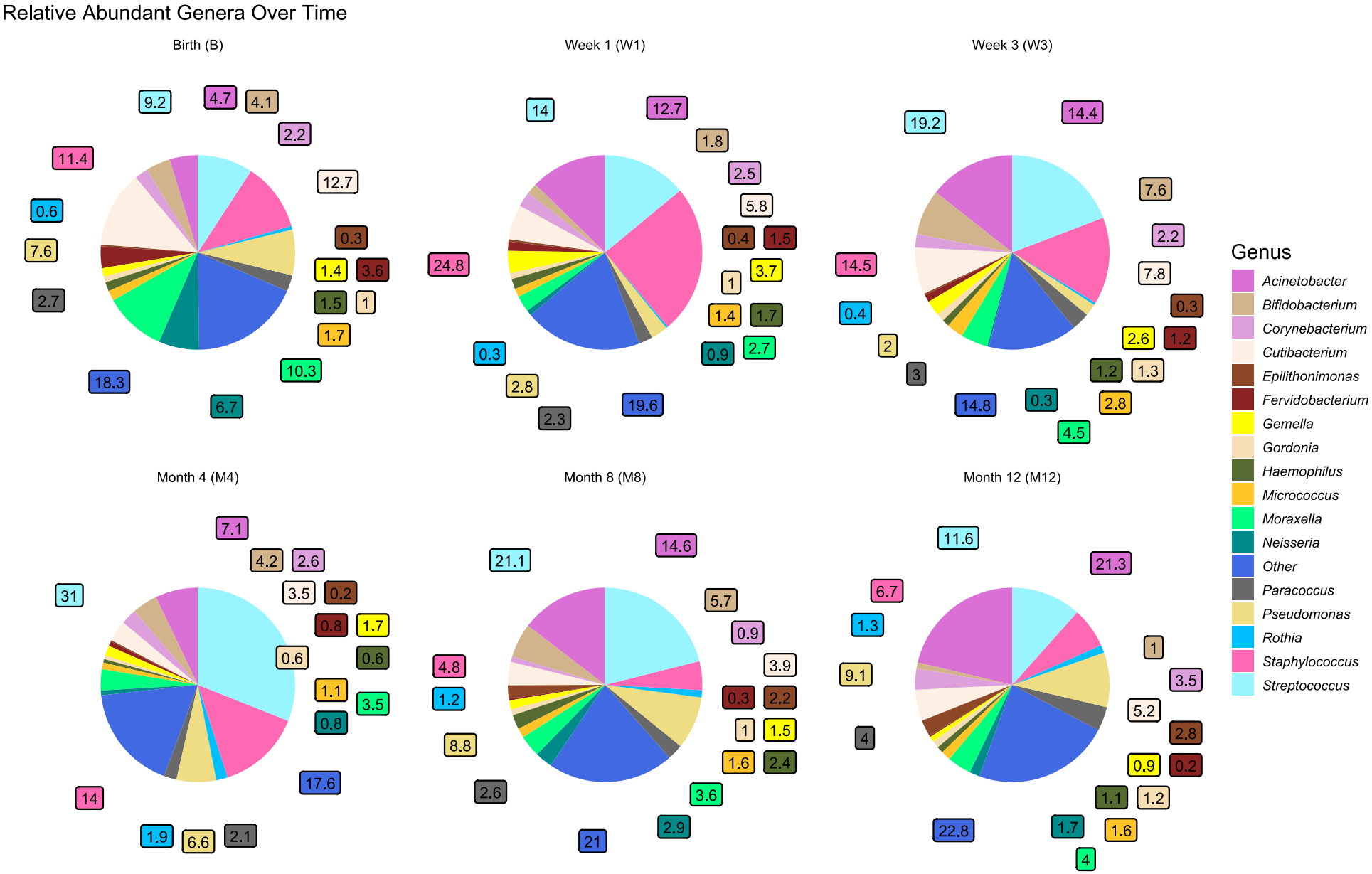

Relative Abundant Genera Over Time
Birth (B)
9.2
4.7
4.1
2.2
11.4
12.7
0.6
0.3
7.6
1.4
3.6
2.7
1.5
1
1.7
10.3
18.3
6.7
Week 1 (W1)
14
12.7
1.8
2.5
5.8
0.4
1.5
24.8
3.7
1
1.4
1.7
0.9
2.7
0.3
2.8
2.3
19.6
Week 3 (W3)
19.2
14.4
7.6
2.2
14.5
7.8
0.4
0.3
2
2.6
1.2
3
1.2
1.3
14.8
0.3
2.8
4.5
Month 4 (M4)
7.1
4.2
2.6
31
3.5
0.2
0.8
1.7
0.6
0.6
1.1
3.5
0.8
17.6
14
1.9
6.6
2.1
Month 8 (M8)
21.1
14.6
5.7
0.9
3.9
4.8
0.3
2.2
1.2
1
1.5
8.8
1.6
2.4
2.6
3.6
2.9
21
Month 12 (M12)
11.6
21.3
6.7
1.3
1
9.1
3.5
4
5.2
2.8
0.9
0.2
1.1
1.2
22.8
1.7
1.6
4
Genus
Acinetobacter
Bifidobacterium
Corynebacterium
Cutibacterium
Epilithonimonas
Fervidobacterium
Gemella
Gordonia
Haemophilus
Micrococcus
Moraxella
Neisseria
Other
Paracoccus
Pseudomonas
Rothia
Staphylococcus
Streptococcus

Figure 4.6. Common Relative Abundant Skin Genera Across all Infants Over Time. Composition data was generated from infant skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Pie charts displayed the relative abundance data. All individuals were combined to calculate the mean relative abundance (%) for each of the common genera (represented in the pie chart slices) at each timepoint (represented in each pie chart) in the infant population. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89). Timepoints investigated included Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12). The coloured labels corresponding to the pie chart slices indicates the genus relative abundance (%), and the Genus key allocates a unique colour to the genera represented in the pie charts and labels.

Relative Abundant Species Over Time

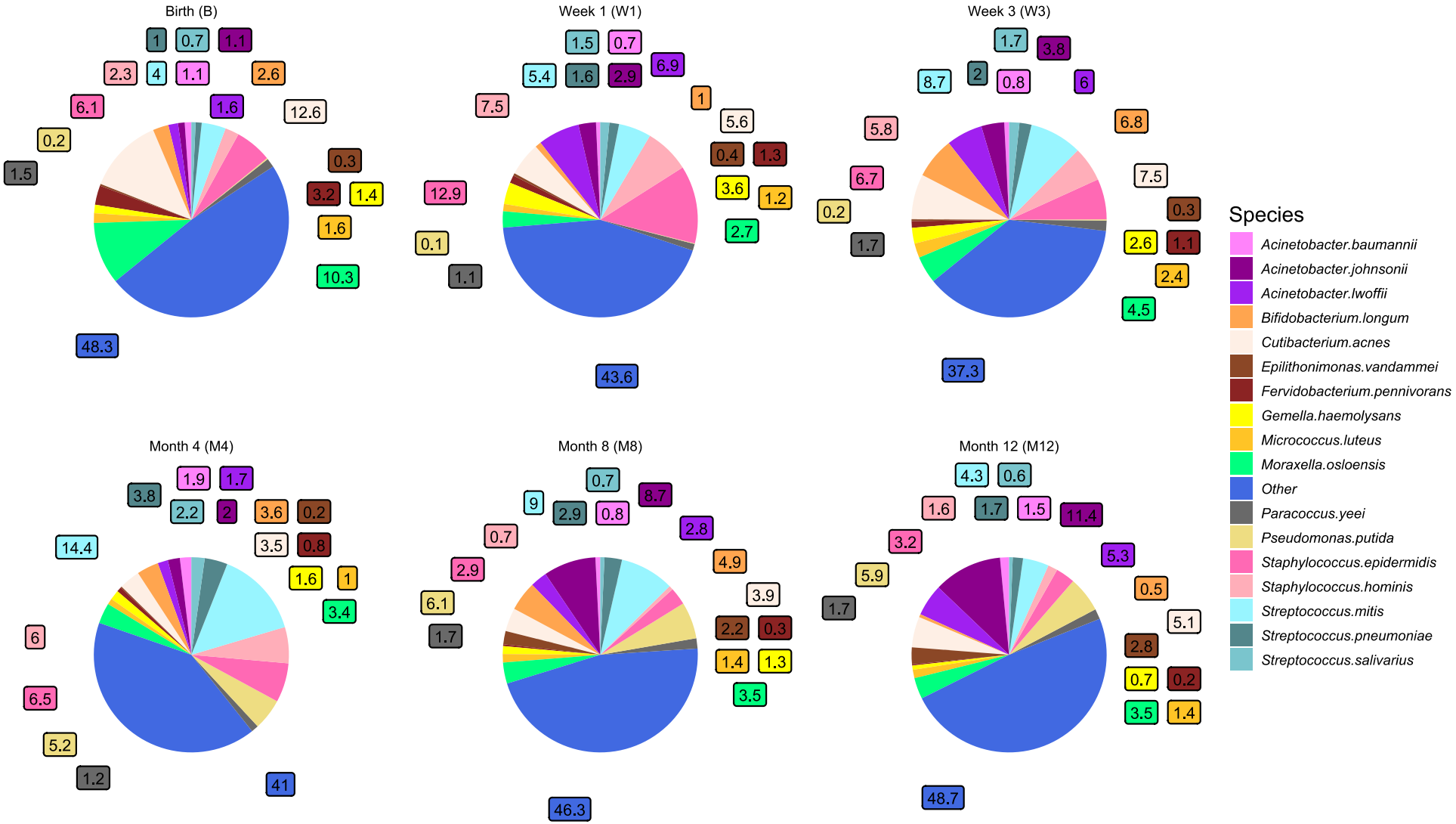
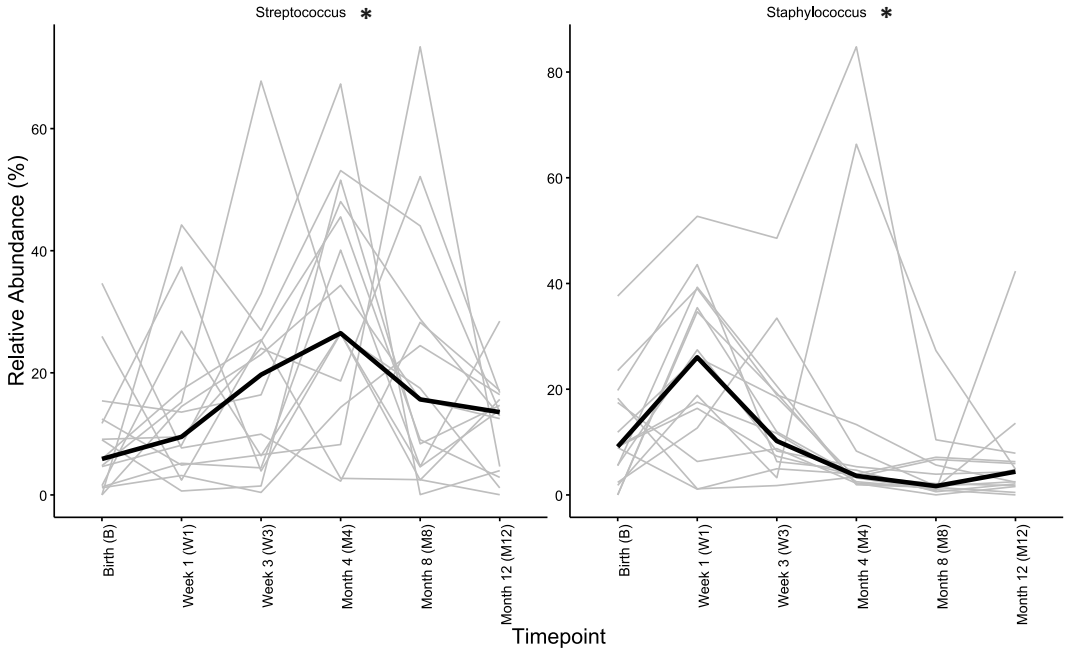
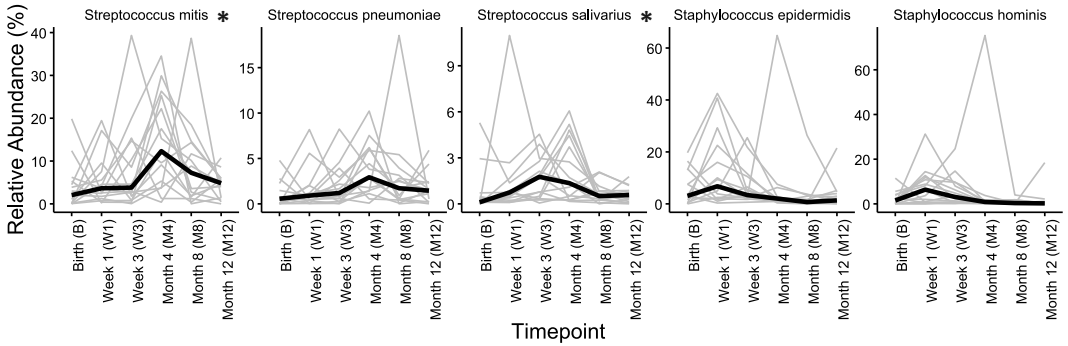


Figure 4.7. Common Relative Abundant Skin Species Across all Infants Over Time. Composition data was generated from infant skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Pie charts displayed the relative abundance data. All individuals were combined to calculate the mean relative abundance (%) for each of the common species (represented in the pie chart slices) at each timepoint (represented in each pie chart) in the infant population. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89). Timepoints investigated included Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12). The coloured labels corresponding to the pie chart slices indicates the species relative abundance (%), and the Species key allocates a unique colour to the species represented in the pie charts and labels.

A. Staphylococcus and Streptococcus Genus Level Relative Abundance Over Time in Infants



B. Staphylococcus and Streptococcus Species Level Relative Abundance Over Time in Infants

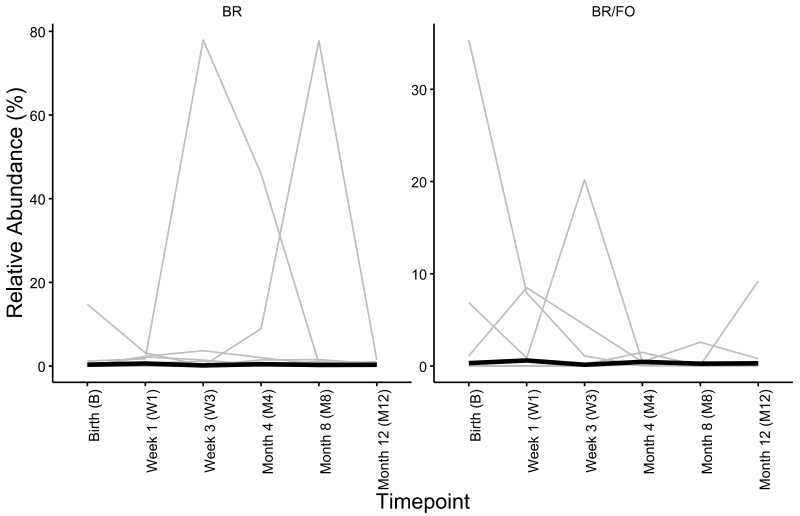


Streptococcus mitis - Staphylococcus hominis *

Streptococcus pneumoniae - Staphylococcus epidermidis *

Streptococcus salivarius - Staphylococcus epidermidis *

C. Bifidobacterium Genus Level Relative Abundance Over Time in Infants



D. Bifidobacterium longum Relative Abundance Over Time in Infants

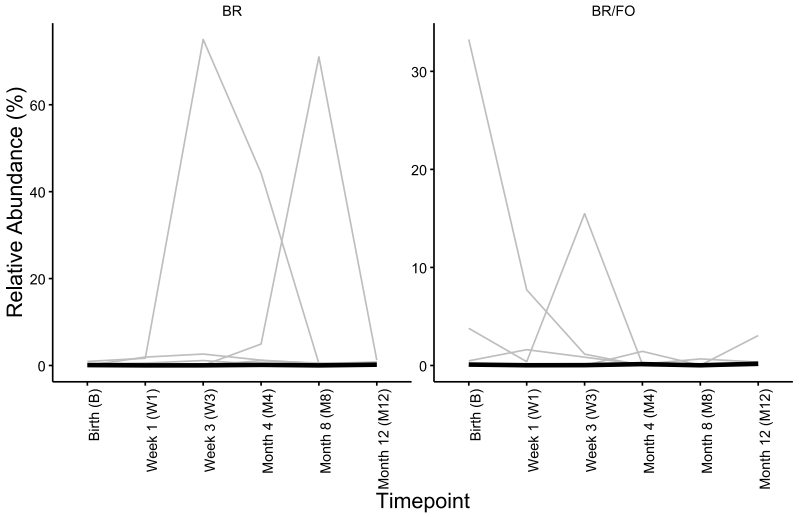


Figure 4.8. Significant Genus and Species Level Relative Abundance Changes Over Time on Infant Skin. Relative abundance data was generated from infant skin swabs using shotgun metagenomics sequencing, generated on the Illumina platform; and cleaned read sequences were run through Kraken2 and Bracken, which are read mapping methods. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89). Timepoints investigated included Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12). The black line intersecting each plot shows the average relative abundance over time. An asterisk (*) indicates whether the genera and species had ANOVA significant differences ($P < 0.05$). **A.** *Staphylococcus* and *Streptococcus* genus level relative abundance changes over time. **B.** *Streptococcus mitis*, *Streptococcus pneumoniae*, *Streptococcus salivarius, Staphylococcus epidermidis and Staphylococcus hominis* species level relative abundance changes over time. **C.** *Bifidobacterium* genus level relative abundance changes over time accounting for feeding mode – Breast (BR) and Breast/Formula (BR/FO) fed. **D.** *Bifidobacterium longum* species level relative abundance changes over time accounting for feeding mode – Breast (BR) and Breast/Formula (BR/FO) fed.

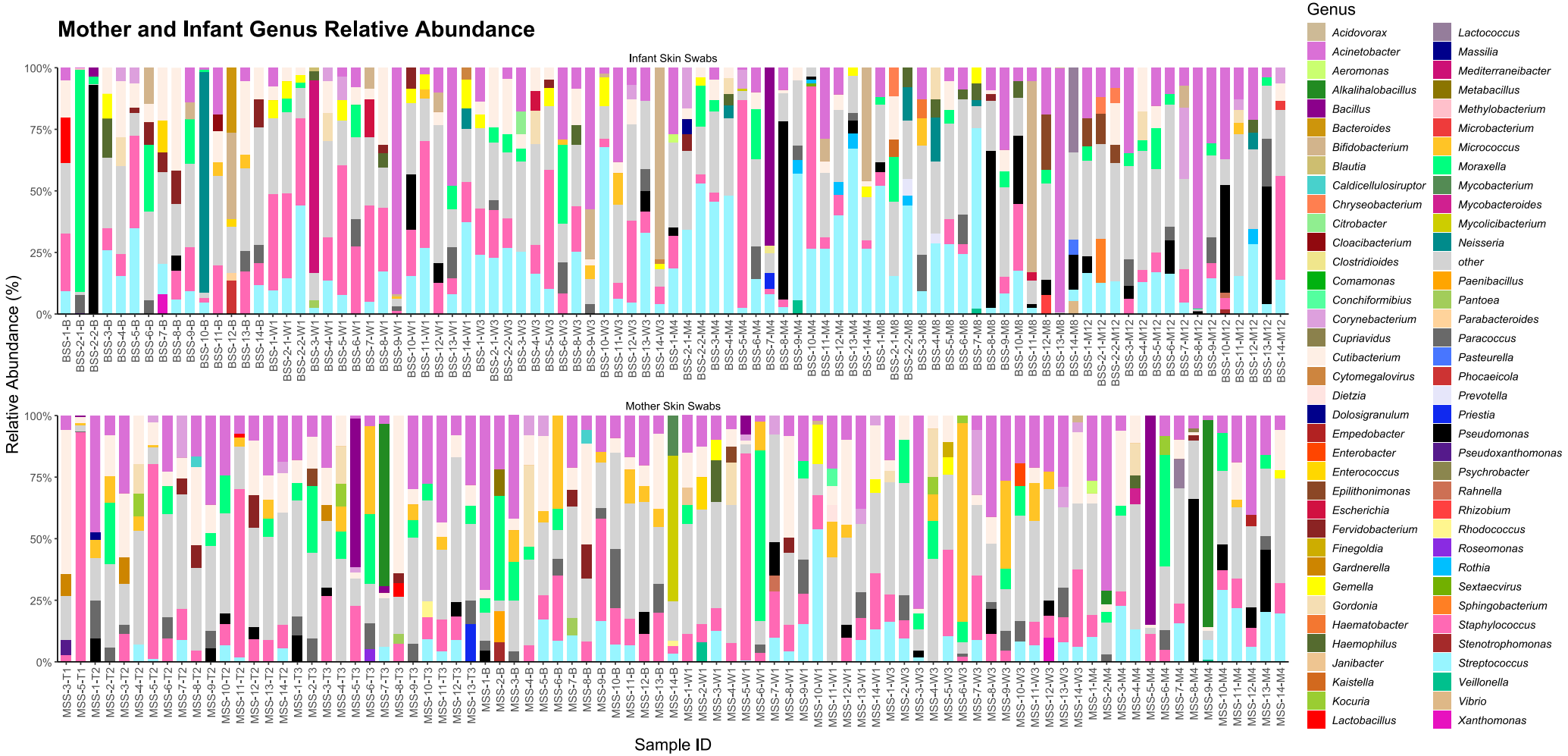

Mother and Infant Genus Relative Abundance
Infant Skin Swabs
Mother Skin Swabs
Relative Abundance (%)
100%
75%
50%
25%
0%
Sample ID
Genus
Acidovorax
Acinetobacter
Aeromonas
Alkalihalobacillus
Bacillus
Bacteroides
Bifidobacterium
Blautia
Caldicellulosiruptor
Chryseobacterium
Citrobacter
Cloacibacterium
Clostridioides
Comamonas
Conchiformibius
Corynebacterium
Cupriavidus
Cutibacterium
Cytomegalovirus
Dietzia
Dolosigranulum
Empedobacter
Enterobacter
Enterococcus
Epilithonimonas
Escherichia
Fervidobacterium
Finegoldia
Gardnerella
Gemella
Gordonia
Haematobacter
Haemophilus
Janibacter
Kaistella
Kocuria
Lactobacillus
Lactococcus
Massilia
Mediterraneibacter
Metabacillus
Methylobacterium
Microbacterium
Micrococcus
Moraxella
Mycobacterium
Mycobacteroides
Mycolicibacterium
Neisseria
other
Paenibacillus
Pantoea
Parabacteroides
Paracoccus
Pasteurella
Phocaeicola
Prevotella
Priestia
Pseudomonas
Pseudoxanthomonas
Psychrobacter
Rahnella
Rhizobium
Rhodococcus
Roseomonas
Rothia
Sextaecvirus
Sphingobacterium
Staphylococcus
Stenotrophomonas
Streptococcus
Veillonella
Vibrio
Xanthomonas

Figure 4.9. Individual Infant and Maternal Skin Relative Abundant Genus Level Taxonomic Profiles Over Time. Composition data was generated from infant and maternal skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bars displayed the relative abundance data. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89) and mother (n=85). Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Each stacked bar represents the relative abundance (%) of the top 5 genera in an individual infant or mother over time. The Genus key allocates a unique colour to the genera represented in the stacked bars.

Mother and Infant Species Relative Abundance

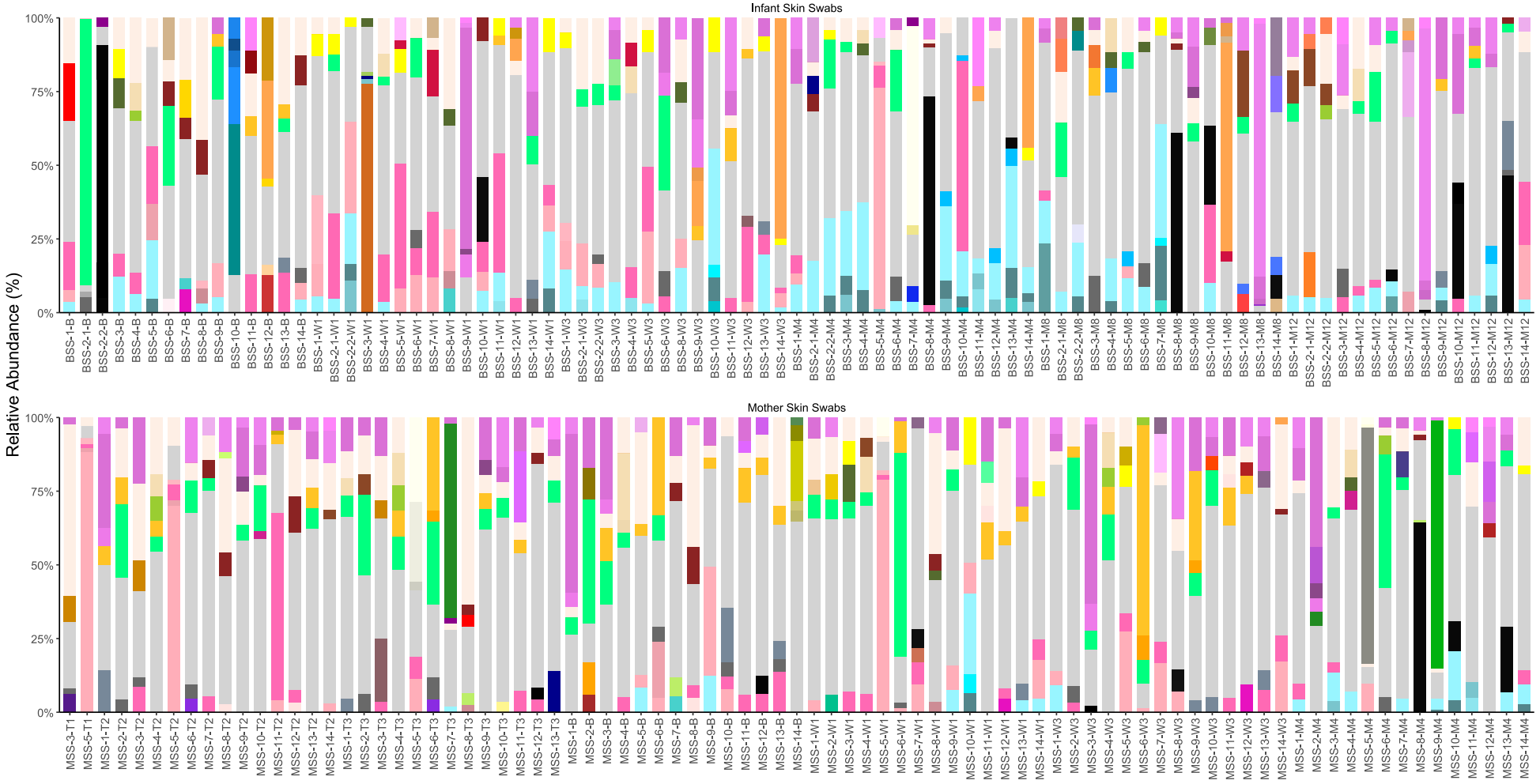




Figure 4.10. Individual Infant and Maternal Skin Relative Abundant Species Level Taxonomic Profiles Over Time. Composition data was generated from infant and maternal skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bars displayed the relative abundance data. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89) and mother (n=85). Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Each stacked bar represents the relative abundance (%) of the top 5 species in an individual infant or mother over time. The Species key allocates a unique colour to the species represented in the stacked bars.

Genus and species level significant differential abundance associations over time across infant skin were identified, also accounting for individual variation (**Figure 4.11**). Multiple genera (*Acinetobacter*, *Epilithonimonas*, *Fervidobacterium* and *Streptococcus*) and species (*A. johnsonii* and *S. salivarius*) that were previously identified as common across all infant skin over time (**Figures 4.6-4.7**), were identified as significantly differentially abundant with positive and negative associations temporally on infant skin.

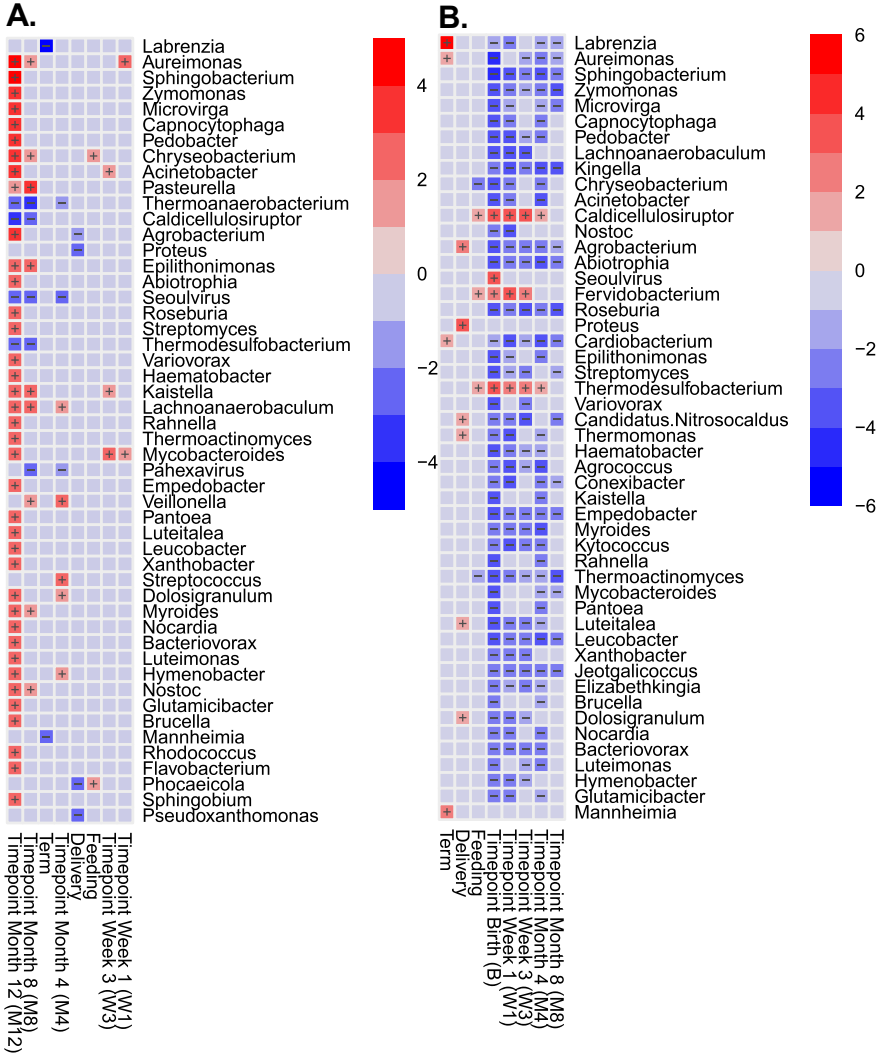
With B as a reference timepoint (**Figure 4.11. A.**), *Acinetobacter* had a weaker positive association on infant skin at W3 and a stronger positive association at M12. *Streptococcus* had a strong positive association on infant skin at M4. *Epilithonimonas* had a strong positive association on infant skin at M8 and M12. To species level (**Figure 4.11. C.**), *A. johnsonii* had a weaker positive association at M8 and a stronger positive association at M12 on infant skin. *S. salivarius* had a weaker positive association at W1 and W3, and a stronger positive association at M4 on infant skin.

With M12 as a reference timepoint (**Figure 4.11. B.**), *Acinetobacter* had a stronger negative association at B and M4, and a weaker negative association at W1 on infant skin. *Epilithonimonas* had a stronger negative association at B, and weaker negative association at W1 and M4 on infant skin. *Fervidobacterium* had a weaker positive association at B and W3, and a stronger positive association at W1 on infant skin. To species level (**Figure 4.11. D.**), *A. johnsonii* had a stronger negative association at B and M4, and a weaker negative association at W1 and W3 on infant skin.

These results correlate with **Figures 4.6-4.7** where certain microbes were in higher abundance at timepoints, such as *Acinetobacter* (*A. johnsonii*) at M12, *Streptococcus* (*S. salivarius*) at M4, *Epilithonimonas* at M8 and M12, and *Fervidobacterium* between B and W3. There was no association of these genera or species with pregnancy term, delivery or feeding mode on infant skin. But *Fervidobacterium* had a weak positive association with feeding mode over time, which is in line with *Fervidobacterium* relative abundance being higher between B to W3, when breast feeding would have been more prevalent. Timepoint, thus age, was influential on the microbial abundance on infant skin, whilst pregnancy term, delivery and feeding mode were not. Feeding mode appeared to have some influence on *Fervidobacterium* abundance at earlier timepoints.

Top 50 features with significant associations (−log(qval)*sign(coeff))

Infant Skin Swab Differential Abundant Genera



Infant Skin Swab Differential Abundant Species

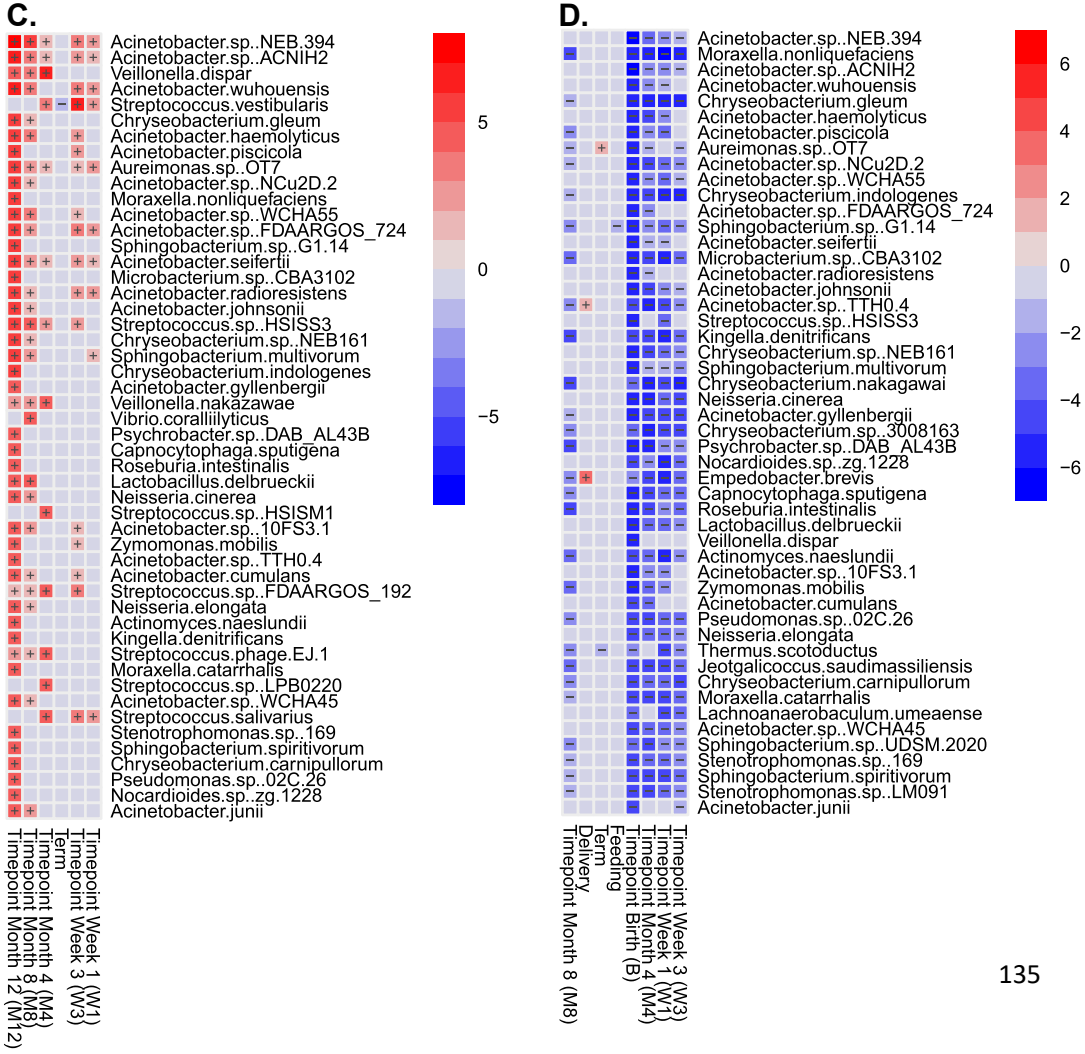


Figure 4.11. Top 50 Differential Abundant Genera and Species Significant Associations Over Time on Infant Skin. Relative abundance data was generated from infant skin swabs using shotgun metagenomics sequencing, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. The relative abundance values were run through MaAsLin2 to obtain significant differential abundant genera and species associations over time, accounting for individual variation and multiple factors. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89). Timepoints investigated included Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12). Factors accounted for included pregnancy term (Late Preterm [LP] and Full Term [FT]), delivery mode (Vaginal Delivery [VD] and Caesarean Section [CS]) and feeding mode (Breast [BR] and Breast/Formula [BR/FO]). The key allocates a significance scale (-log[q-value]*sign[coefficient]) with a unique colour and direction to positive (red [+]) and negative (blue [-]) microbial associations. **A.** Infant skin swab differential abundant genera significant associations over time using birth as a reference timepoint. **B.** Infant skin swab differential abundant genera significant associations over time using month 12 as a reference timepoint. **C.** Infant skin swab differential abundant species significant associations over time using birth as a reference timepoint. **D.** Infant skin swab differential abundant species significant associations over time using month 12 as a reference timepoint.

4.6.2 Maternal and Infant Skin Archaeal Composition Over Time

Archaeal profiles were successfully generated from 67/174 maternal and infant skin swab samples, but not the skin culture datasets demonstrating that archaea were unculturable from the skin; as only 67 samples generated archaeal profiles, timepoints were lost. The archaeal taxonomic profiles within each individual, in the maternal and infant datasets, identified 13 common species (*Halorubrum lipolyticum*, *Haloferax massiliensis*, *Methanobacterium sp.* 000499765, *Methanobrevibacter_A oralis*, *Methanobrevibacter_A sp.* 900766745, *Methanobrevibacter_A sp.* 900769095, *Methanobrevibacter_A smithii*, *Methanobrevibacter_A smithii_A*, *Methanobrevibacter_A woesei*, *Methanosphaera sp.* 900322125, *Methanosphaera stadtmanae*, MGYG000003962 and UBA71 *sp.* 006954425) (**Figure 4.12**). The maternal and infant skin exhibited similar archaeal profiles, which were in very low abundance in the skin (1% of each sample) relative to the other microbial communities present (99% of each sample); “Other” represents the other microbial communities identified within the samples.

Species level relative abundance changes of archaea over time across all maternal and infant datasets were identified, accounting for individual variation (**Figure 4.13**). Significant differences were identified in the relative abundance changes between archaeal species over time ($P < 0.001$); in the relative abundance changes of archaeal species between the maternal and infant skin swab datasets over time ($P < 0.05$). These trends were still apparent when accounting for pregnancy term (**Figure 4.14**), delivery (**Figure 4.15**) and feeding mode (**Figure 4.16**). There were additional significant differences identified in the relative abundance changes of archaeal species between mothers that solely BR and BR/FO fed in the maternal skin swab dataset ($P < 0.05$).

This appears to be the first descriptive account of archaea (sequenced using SMS) identified on infant skin within an infant cohort. These archaeal results indicate that timepoint (age) did not influence the archaeal relative abundance within each infant and maternal dataset but did influence relative abundance between all archaeal species across and between maternal and infant skin. Pregnancy term (gestational age), delivery and feeding (breast feeding) mode, did not appear to directly influence the establishment and development of the infant skin microbiome.

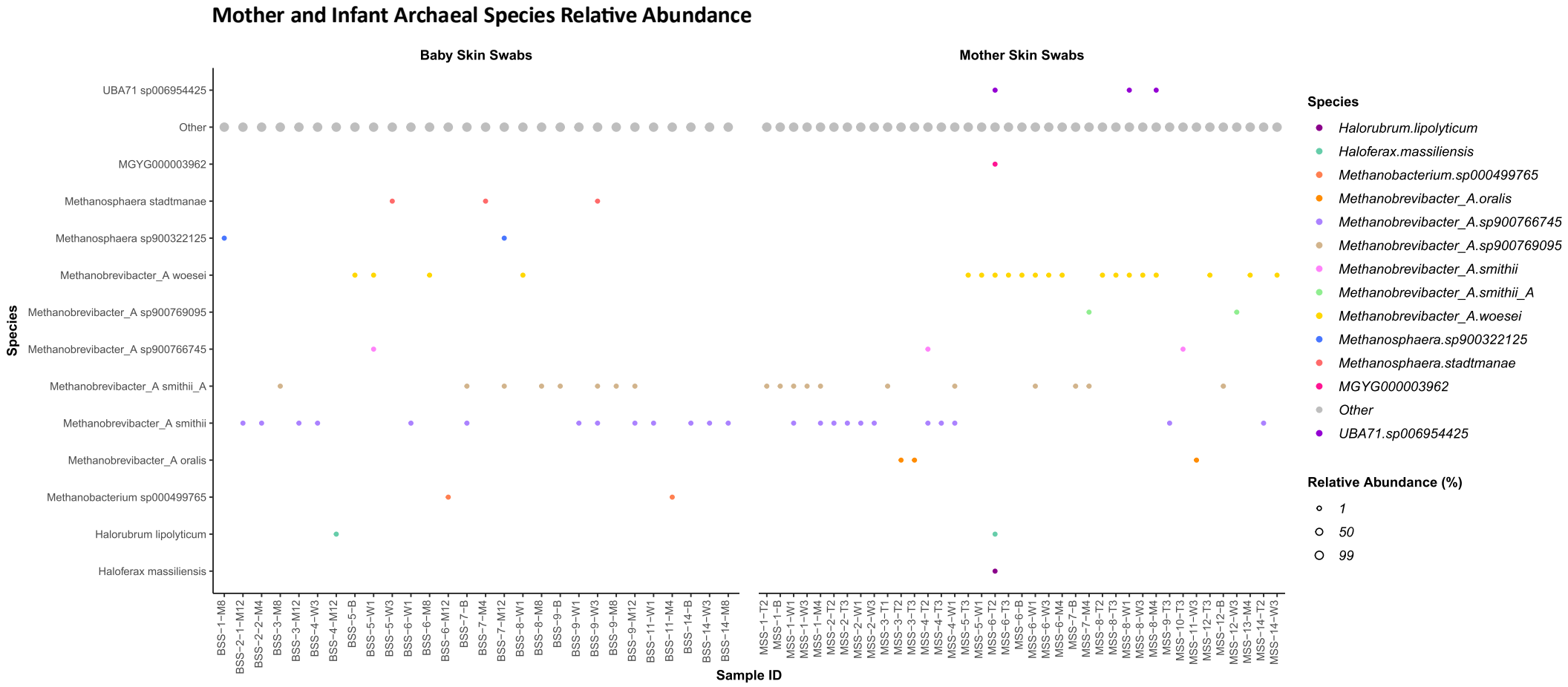
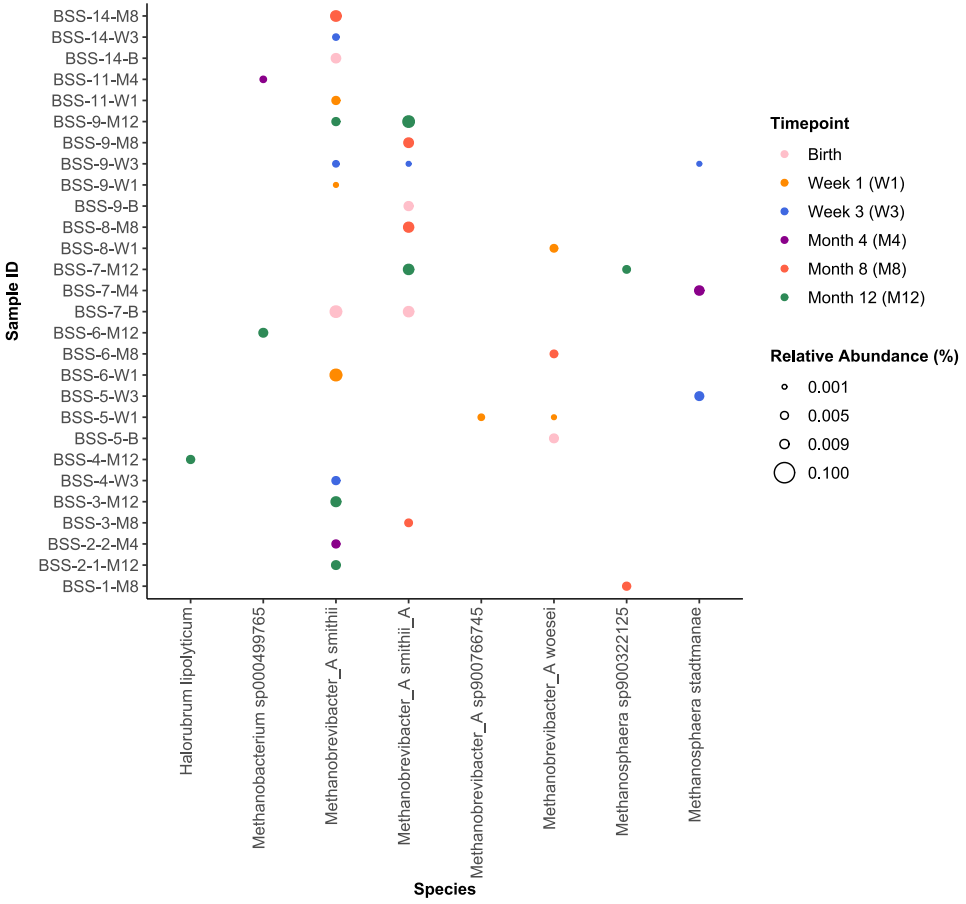

Mother and Infant Archaeal Species Relative Abundance
Baby Skin Swabs
Mother Skin Swabs
Species
UBA71 sp006954425
Other
MGYG000003962
Methanosphaera stadtmanae
Methanosphaera sp900322125
Methanobrevibacter_A woesei
Methanobrevibacter_A sp900769095
Methanobrevibacter_A sp900766745
Methanobrevibacter_A smithii_A
Methanobrevibacter_A smithii
Methanobrevibacter_A oralis
Methanobacterium sp000499765
Halorubrum lipolyticum
Haloferax massiliensis
BSS−1−M8
BSS−2−1−M12
BSS−2−2−M4
BSS−3−M8
BSS−3−M12
BSS−4−W3
BSS−4−M12
BSS−5−B
BSS−5−W1
BSS−5−W3
BSS−6−W1
BSS−6−M8
BSS−6−M12
BSS−7−B
BSS−7−M4
BSS−7−M12
BSS−8−W1
BSS−8−M8
BSS−9−B
BSS−9−W1
BSS−9−W3
BSS−9−M8
BSS−9−M12
BSS−11−W1
BSS−11−M4
BSS−14−B
BSS−14−W3
BSS−14−M8
MSS−1−T2
MSS−1−B
MSS−1−W1
MSS−1−W3
MSS−1−M4
MSS−2−T2
MSS−2−T3
MSS−2−W1
MSS−2−W3
MSS−3−T1
MSS−3−T2
MSS−3−T3
MSS−4−T2
MSS−4−T3
MSS−4−W1
MSS−5−T3
MSS−5−W1
MSS−6−T2
MSS−6−T3
MSS−6−B
MSS−6−W1
MSS−6−W3
MSS−6−M4
MSS−7−B
MSS−7−M4
MSS−8−T2
MSS−8−T3
MSS−8−W1
MSS−8−W3
MSS−8−M4
MSS−9−T3
MSS−10−T3
MSS−11−W3
MSS−12−T3
MSS−12−B
MSS−12−W3
MSS−13−M4
MSS−14−T2
MSS−14−W3
Sample ID
Species
Halorubrum.lipolyticum
Haloferax.massiliensis
Methanobacterium.sp000499765
Methanobrevibacter_A.oralis
Methanobrevibacter_A.sp900766745
Methanobrevibacter_A.sp900769095
Methanobrevibacter_A.smithii
Methanobrevibacter_A.smithii_A
Methanobrevibacter_A.woesei
Methanosphaera.sp900322125
Methanosphaera.stadtmanae
MGYG000003962
Other
UBA71.sp006954425
Relative Abundance (%)
1
50
99

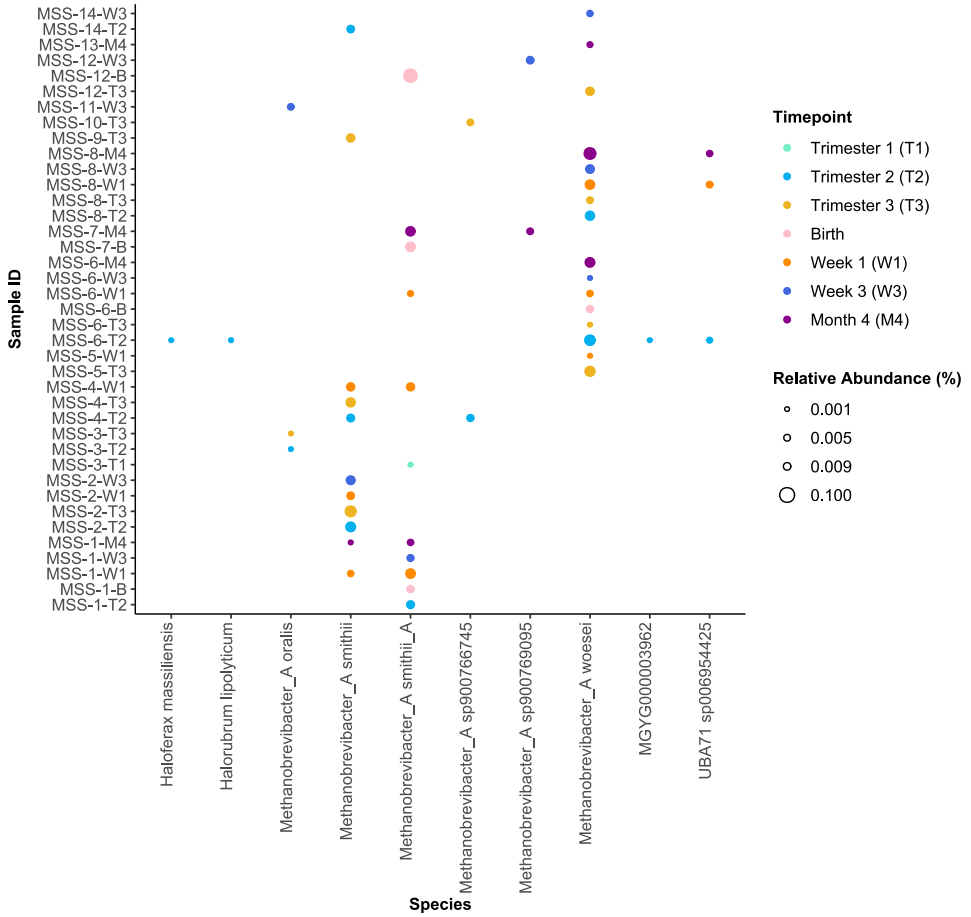
Figure 4.12. Archaeal Species Relative Abundance Profiles in Maternal and Infant Skin Swab Datasets. Shotgun metagenomics data was generated from infant and maternal skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with ATLAS, Kraken2 and Bracken. Bubbles display the relative abundance data of species. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89) and mother (n=85). Only 67 samples generated archaeal profiles from mothers (n=39) and infants (n=28). Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. The species were allocated a unique colour and the relative abundance (%) is represented by the bubble size as identified in the key.

Mother and Infant Archaeal Species Relative Abundance Changes Over Time

A



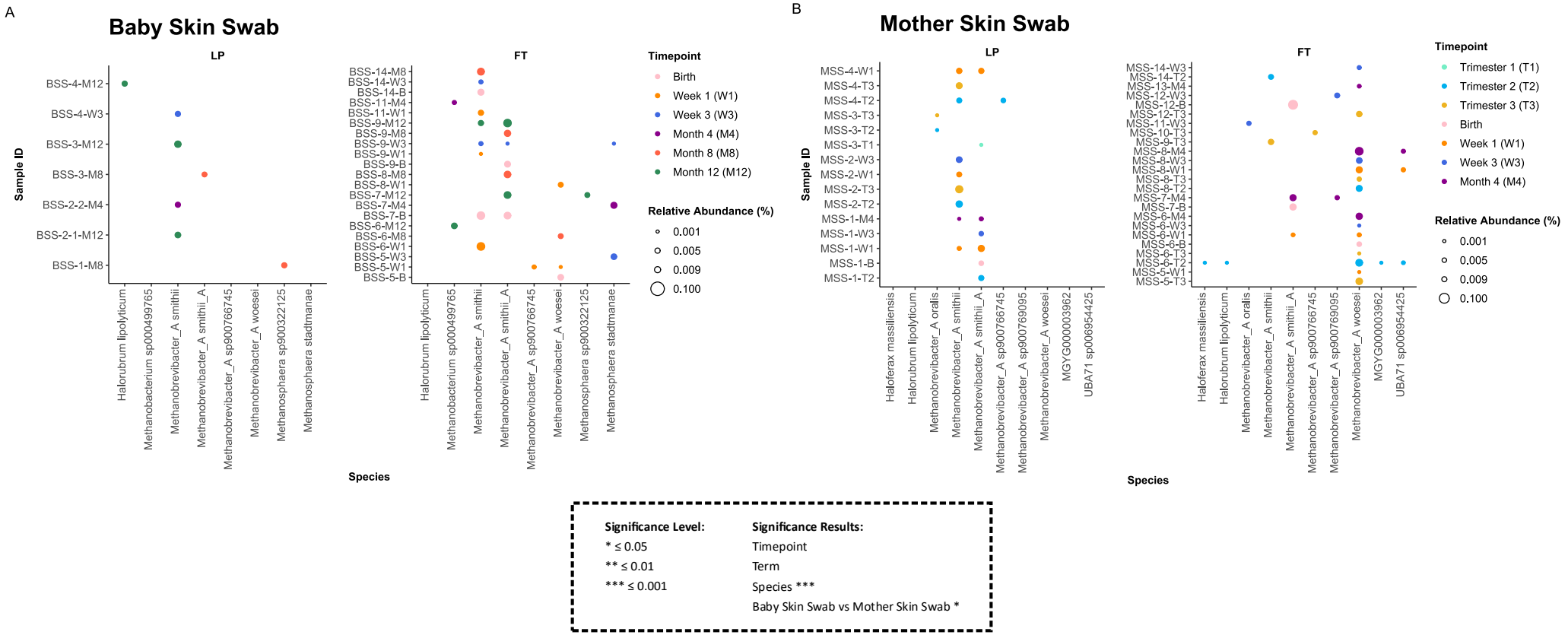
B



Significance Level:	Significance Results:
* ≤ 0.05	Timepoint
** ≤ 0.01	Species ***
*** ≤ 0.001	Baby Skin Swab vs Mother Skin Swab *

Figure 4.13. Archaeal Species Relative Abundance Changes Over Time in Maternal and Infant Skin Swab Datasets. Shotgun metagenomics data was generated from infant and maternal skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with ATLAS, Kraken2 and Bracken. Bubbles display the relative abundance data of species. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89) and mother (n=85). Only 67 samples generated archaeal profiles from mothers (n=39) and infants (n=28). Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. The timepoints were allocated a unique colour as identified in the Timepoint key, the relative abundance (%) is represented by the bubble size as identified in the Relative Abundance (%) key, and the ANOVA significant differences are indicated with an asterisk (*) as identified in the Significance Level and Significance Results key.

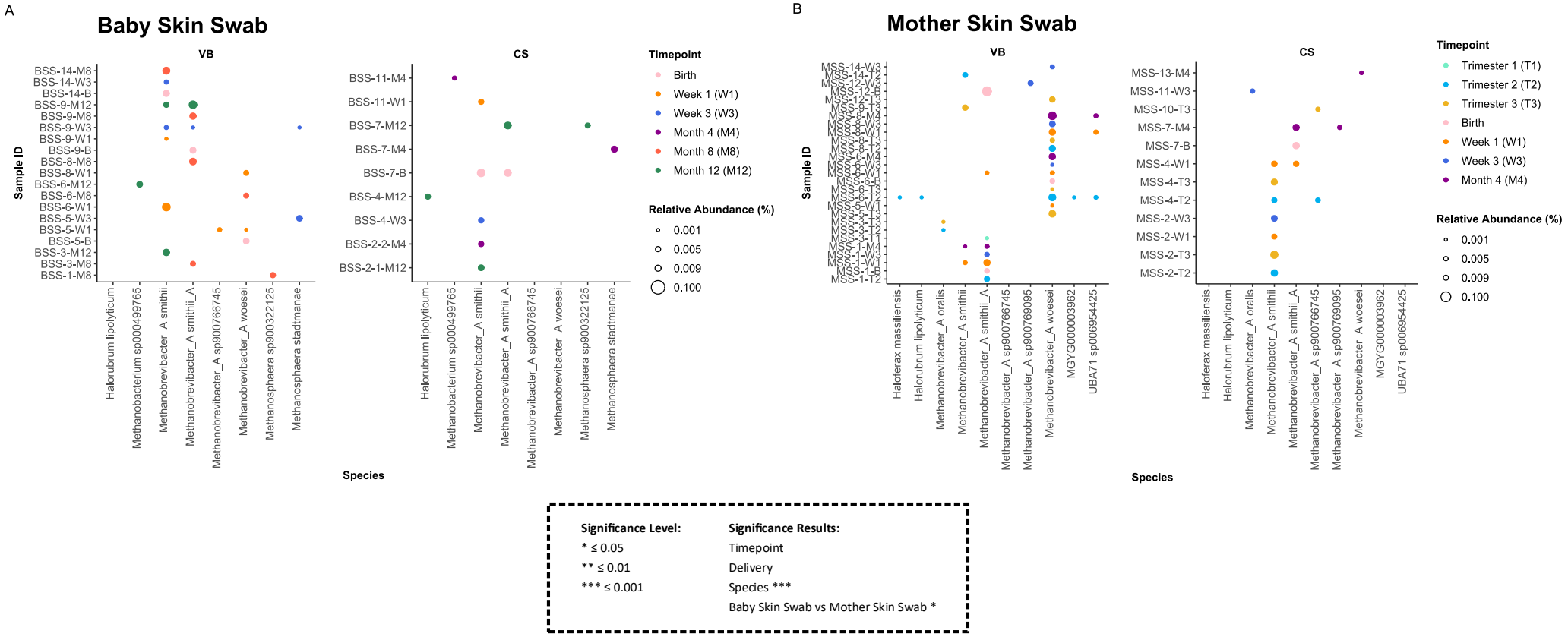
Mother and Infant Archaeal Species Relative Abundance Changes Over Time Accounting for Pregnancy Term



Significance Level:	Significance Results:
* ≤ 0.05	Timepoint
** ≤ 0.01	Term
*** ≤ 0.001	Species ***
	Baby Skin Swab vs Mother Skin Swab *

Figure 4.14. Archaeal Species Relative Abundance Changes Over Time in Maternal and Infant Skin Swab Datasets Accounting for Pregnancy Term. Shotgun metagenomics data was generated from infant and maternal skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with ATLAS, Kraken2 and Bracken. Bubbles display the relative abundance data of species. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89) and mother (n=85). Only 67 samples generated archaeal profiles from mothers (n=39) and infants (n=28). Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Pregnancy term accounted for was split into Late Pre-Term (LP) and Full Term (FT). The timepoints were allocated a unique colour as identified in the Timepoint key, the relative abundance (%) is represented by the bubble size as identified in the Relative Abundance (%) key, and the ANOVA significant differences are indicated with an asterisk (*) as identified in the Significance Level and Significance Results key.

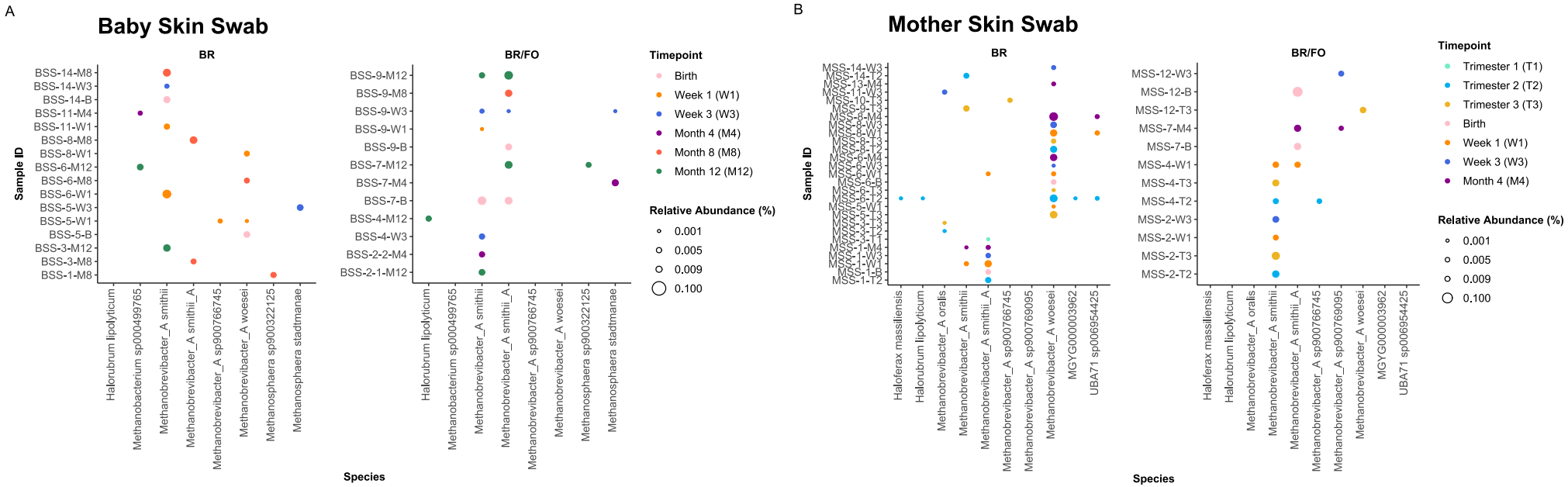
Mother and Infant Archaeal Species Relative Abundance Changes Over Time Accounting for Delivery Mode



Significance Level:	Significance Results:
* ≤ 0.05	Timepoint
** ≤ 0.01	Delivery
*** ≤ 0.001	Species ***
	Baby Skin Swab vs Mother Skin Swab *

Figure 4.15. Archaeal Species Relative Abundance Changes Over Time in Maternal and Infant Skin Swab Datasets Accounting for Delivery Mode. Shotgun metagenomics data was generated from infant and maternal skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with ATLAS, Kraken2 and Bracken. Bubbles display the relative abundance data of species. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89) and mother (n=85). Only 67 samples generated archaeal profiles from mothers (n=39) and infants (n=28). Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Delivery mode accounted for was split into Caesarean Section (CS) and Vaginal Delivery (VD). The timepoints were allocated a unique colour as identified in the Timepoint key, the relative abundance (%) is represented by the bubble size as identified in the Relative Abundance (%) key, and the ANOVA significant differences are indicated with an asterisk (*) as identified in the Significance Level and Significance Results key.

Mother and Infant Archaeal Species Relative Abundance Changes Over Time Accounting for Feeding Mode



Significance Level:	Significance Results:
* ≤ 0.05	Timepoint
** ≤ 0.01	Feeding: Baby Skin Swab
*** ≤ 0.001	Feeding: Mother Skin Swab *
	Species ***
	Baby Skin Swab vs Mother Skin Swab *

Figure 4.16. Archaeal Species Relative Abundance Changes Over Time in Maternal and Infant Skin Swab Datasets Accounting for Feeding Mode. Shotgun metagenomics data was generated from infant and maternal skin swabs using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with ATLAS, Kraken2 and Bracken. Bubbles display the relative abundance data of species. The origin swab samples were collected at multiple timepoints, with one swab collected from one forearm from each infant (n=89) and mother (n=85). Only 67 samples generated archaeal profiles from mothers (n=39) and infants (n=28). Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Feeding mode accounted for was split into Breast (BR) and Breast/Formula combined (BR/FO) fed. The timepoints were allocated a unique colour as identified in the Timepoint key, the relative abundance (%) is represented by the bubble size as identified in the Relative Abundance (%) key, and the ANOVA significant differences are indicated with an asterisk (*) as identified in the Significance Level and Significance Results key.

4.6.3 Maternal and Infant Skin, Stool and Vaginal Composition Diversity Changes Over Time

Significant alpha diversity genus and species level changes over time across infant skin (and all other maternal and infant body sites) were identified, also accounting for individual variation (**Figures 4.17-4.20; Table 4.1**). Alpha diversity significant differences indicated that timepoint influenced within sample microbial diversity in the infant skin swabs and infant stools to genus and species level, mother vaginal swabs to genus level and mother skin cultures to species level; between the infant skin swab and other maternal and infant body sites to genus and species level (infant and maternal skin swab, infant skin swab and infant stool, infant skin swab and maternal stool, infant skin swab and maternal vaginal swab, infant skin swab and maternal vaginal culture, infant skin swab and infant skin culture, and infant skin swab and maternal skin culture); and overall across all maternal and infant body sites to genus and species level. All P values can be seen in **Table 4.1**.

These trends were still apparent when accounting for pregnancy term (**Table 4.2**), delivery mode (**Table 4.3**), and feeding mode (**Table 4.4**). There was a significant difference in the within sample microbial diversity over time between LP and FT infants and mothers in the infant and mother skin swabs to species level, infant stools to genus and species level, and maternal stools to genus level. There was a significant difference in the within sample microbial diversity over time between CS and VB infants in infant stools to genus and species level, and maternal skin and vaginal swabs to genus level. There was a significant difference in the within sample microbial diversity over time between BR and BR/FO fed infants in the infant stools to genus level and the maternal skin to genus and species level. All P values can be seen in **Tables 4.2-4.4**.

With a focus on the infant microbiome, these trends indicate age as a defining factor for the within sample microbial diversity changes on infant skin and within infant stools. The diversity differences between pregnancy term, delivery and feeding modes in infant stools, and pregnancy term on infant skin could indicate that gestational age, delivery mode and breast feeding were involved in the within sample microbial changes in infants over time. Thus, age and gestational age appear important in the establishment and development of the infant skin microbiome, whilst age, gestational age, delivery and feeding mode appear important in the establishment and development of the infant gut microbiome.

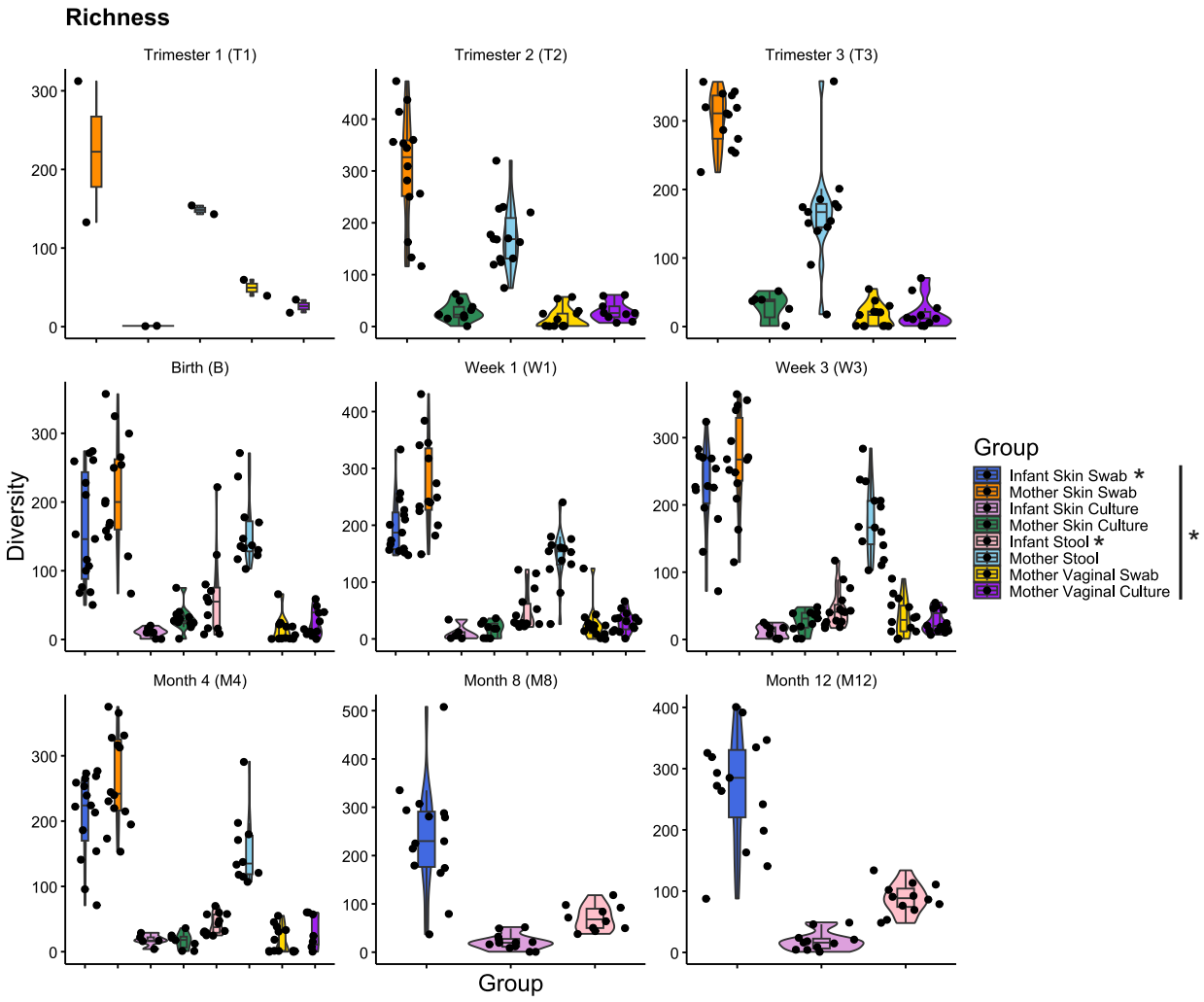


Figure 4.17. Mother and Infant Genus Level Richness Alpha Diversity Changes Over Time at Multiple Body Sites. Diversity data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures) using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Using the relative abundance values, alpha diversity data was generated and displayed in box/violin plots. The Richness metric was used to measure alpha diversity. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Horizontal bars on each plot show averages and vertical bars show the SEM. The Group key allocates a unique colour to indicate the maternal and infant body sites represented in the boxes/violins. An asterisk (*) indicates the infant and mother body sites with ANOVA significant differences ($P < 0.05$) in alpha diversity changes over time, and a line with an asterisk (*) indicates a significant difference ($P < 0.05$) over time across all the groups combined.

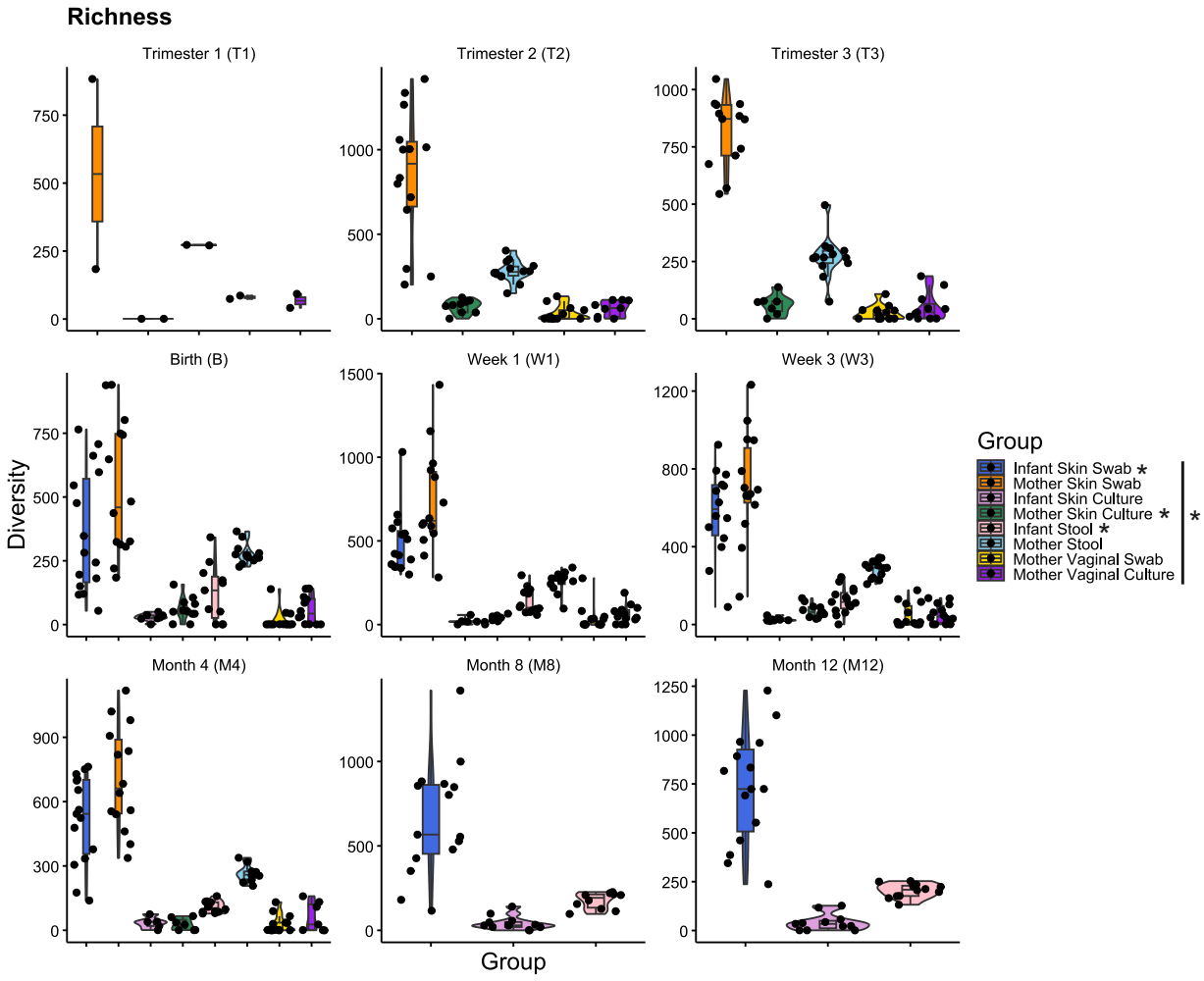


Figure 4.18. Mother and Infant Species Level Richness Alpha Diversity Changes Over Time at Multiple Body Sites. Diversity data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures) using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Using the relative abundance values, alpha diversity data was generated and displayed in box/violin plots. The Richness metric was used to measure alpha diversity. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Horizontal bars on each plot show averages and vertical bars show the SEM. The Group key allocates a unique colour to indicate the maternal and infant body sites represented in the boxes/violins. An asterisk (*) indicates the infant and mother body sites with ANOVA significant differences ($P < 0.05$) in alpha diversity changes over time, and a line with an asterisk (*) indicates a significant difference ($P < 0.05$) over time across all the groups combined.

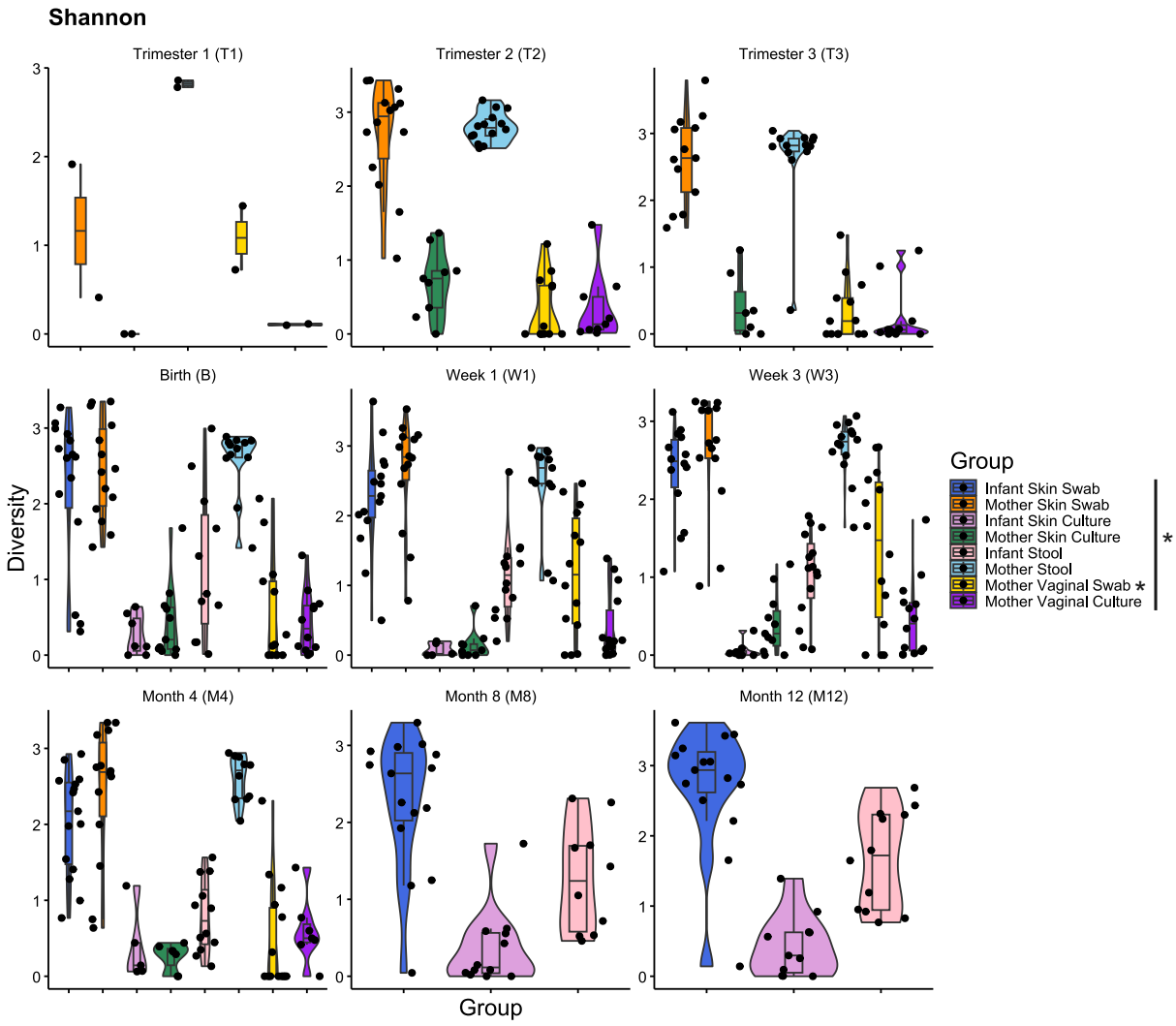


Figure 4.19. Mother and Infant Genus Level Shannon Alpha Diversity Changes Over Time at Multiple Body Sites. Diversity data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures) using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Using the relative abundance values, alpha diversity data was generated and displayed in box/violin plots. The Shannon metric was used to measure alpha diversity. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Horizontal bars on each plot show averages and vertical bars show the SEM. The Group key allocates a unique colour to indicate the maternal and infant body sites represented in the boxes/violins. An asterisk (*) indicates the infant and mother body sites with ANOVA significant differences ($P < 0.05$) in alpha diversity changes over time, and a line with an asterisk (*) indicates a significant difference ($P < 0.05$) over time across all the groups combined.

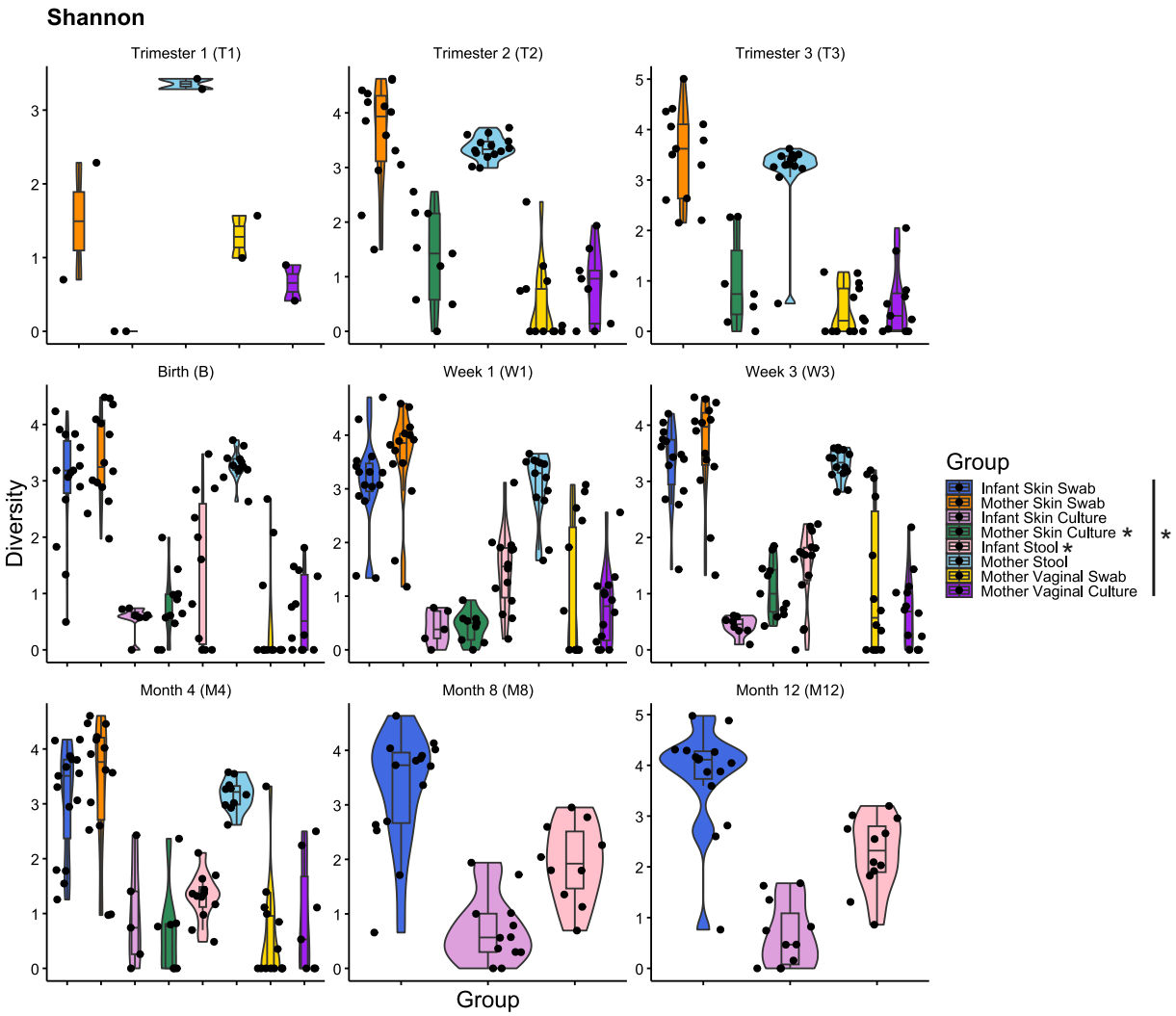


Figure 4.20. Mother and Infant Species Level Shannon Alpha Diversity Changes Over Time at Multiple Body Sites. Diversity data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures) using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Using the relative abundance values, alpha diversity data was generated and displayed in box/violin plots. The Shannon metric was used to measure alpha diversity. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Horizontal bars on each plot show averages and vertical bars show the SEM. The Group key allocates a unique colour to indicate the maternal and infant body sites represented in the boxes/violins. An asterisk (*) indicates the infant and mother body sites with ANOVA significant differences ($P < 0.05$) in alpha diversity changes over time, and a line with an asterisk (*) indicates a significant difference ($P < 0.05$) over time across all the groups combined.

Table 4.1. Alpha Diversity Index ANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time

Datasets	Genus P Values		Species P Values	
Individual	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS	*P≤0.01*		*P≤0.01*	
MSS				
BST	*P≤0.05*		*P≤0.01*	*P≤0.05*
MST				
MVS		*P≤0.01*		
MVC				
BSC				
MSC			*P≤0.05*	*P≤0.05*
Between	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS-MSS	*P≤0.001*	*P≤0.05*	*P≤0.001*	
BSS-BST	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MST	*P≤0.001*	*P≤0.001*	*P≤0.001*	
BSS-MVS	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MVC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-BSC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MSC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
Across All	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.01*

Table 4.2. Alpha Diversity Index ANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Pregnancy Term

Datasets	Genus P Values		Species P Values	
Individual	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS	*P≤0.01*		*P≤0.01*	
MSS				
BST	*P≤0.01*	*P≤0.05*	*P≤0.01*	*P≤0.05*
MST				
MVS		*P≤0.01*		
MVC				
BSC				
MSC			*P≤0.05*	
Between	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS-MSS	*P≤0.001*	*P≤0.05*	*P≤0.001*	
BSS-BST	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MST	*P≤0.001*	*P≤0.001*	*P≤0.001*	
BSS-MVS	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MVC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-BSC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MSC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
Across All	*P≤0.001*	*P≤0.05*	*P≤0.001*	*P≤0.01*
FT vs LP	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS			*P≤0.05*	
MSS			*P≤0.01*	
BST	*P≤0.01*	*P≤0.05*	*P≤0.05*	
MST	*P≤0.01*			
MVS				
MVC				
BSC				
MSC				

Table 4.3. Alpha Diversity Index ANOVA P Value for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Delivery Mode

Datasets	Genus P Values		Species P Values	
Individual	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS	*P≤0.01*		*P≤0.01*	
MSS				
BST	*P≤0.01*		*P≤0.01*	*P≤0.05*
MST				
MVS		*P≤0.01*		
MVC				
BSC				
MSC			*P≤0.05*	
Between	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS-MSS	*P≤0.001*	*P≤0.05*	*P≤0.001*	
BSS-BST	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MST	*P≤0.001*	*P≤0.001*	*P≤0.001*	
BSS-MVS	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MVC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-BSC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MSC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
Across All	*P≤0.001*	*P≤0.05*	*P≤0.001*	*P≤0.01*
CS vs VD	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS				
MSS		*P≤0.01*		*P≤0.01*
BST	*P≤0.01*	*P≤0.05*	*P≤0.05*	
MST				
MVS	*P≤0.05*			
MVC				
BSC				
MSC				

Table 4.4. Alpha Diversity Index ANOVA P Value for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Feeding Mode

Datasets	Genus P Values		Species P Values	
Individual	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS	*P≤0.01*		*P≤0.01*	
MSS				
BST	*P≤0.01*		*P≤0.01*	*P≤0.05*
MST				
MVS		*P≤0.01*		
MVC				
BSC				
MSC			*P≤0.05*	
Between	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS-MSS	*P≤0.001*	*P≤0.01*	*P≤0.001*	
BSS-BST	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MST	*P≤0.001*	*P≤0.001*	*P≤0.001*	
BSS-MVS	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MVC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-BSC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
BSS-MSC	*P≤0.001*	*P≤0.001*	*P≤0.001*	*P≤0.001*
Across All	*P≤0.001*	*P≤0.05*	*P≤0.001*	*P≤0.01*
BR vs BR/FO	**Richness**	**Shannon**	**Richness**	**Shannon**
BSS				
MSS		*P≤0.01*		*P≤0.05*
BST	*P≤0.05*	*P≤0.05*		
MST				
MVS				
MVC				
BSC				
MSC				

Beta diversity genus and species level changes over time across infant skin (and all other maternal and infant datasets) were identified, also accounting for individual variation (**Figures 4.21-4.22; Table 4.5**). Significant differences indicated that timepoint influenced the between sample microbial diversity in the infant skin swab, infant stool and mother vaginal swab body sites to genus and species level, and the infant skin cultures to species level; between the infant skin swab and certain maternal and infant body sites (infant and maternal skin swab, infant skin and infant stool, infant skin and maternal stool, infant skin and maternal vaginal swab, infant skin and maternal vaginal culture, infant skin and infant skin culture, infant skin and mother skin culture) to genus and species level; overall across all maternal and infant datasets to genus and species level. P values can be seen in **Table 4.5**.

These trends were still apparent when accounting for pregnancy term (**Table 4.6**), delivery (**Table 4.7**) and feeding mode (**Table 4.8**). Pregnancy term also influenced the between sample microbial diversity in the maternal skin cultures to species level over time. There was a significant difference in the between sample microbial diversity over time between LP and FT infants in the infant skin swabs to genus and species level, and mothers in the mother skin swab and stools to genus and species level, and mother vaginal swab and cultures to species level. There was a significant difference in the between sample microbial diversity over time between CS and VB infants in the infant skin swabs to species level, and mothers in the mother skin swab, stool and vaginal swabs to genus and species level, and the mother vaginal cultures to species level. There was a significant difference in the between sample microbial diversity over time between BR and BR/FO fed infants in the infant skin swabs to species level, infant stools to genus and species level, and mothers in the mother skin swab, stool and vaginal swab and culture body sites to genus and species level. P values can be seen in **Tables 4.6-4.8**.

With a focus on the infant microbiome, these trends indicate age as a defining factor for the between sample microbial diversity changes on infant skin and in infant stool. The diversity differences between pregnancy term, delivery and feeding mode on infant skin and feeding mode in infant stool, indicates that gestational age, delivery mode and breast feeding were involved in the between sample microbial diversity changes in infants. Thus, age, gestational age, delivery and feeding mode appear important in the establishment and development of the infant skin microbiome, and age and feeding mode appear important in the establishment and development of the infant gut microbiome.

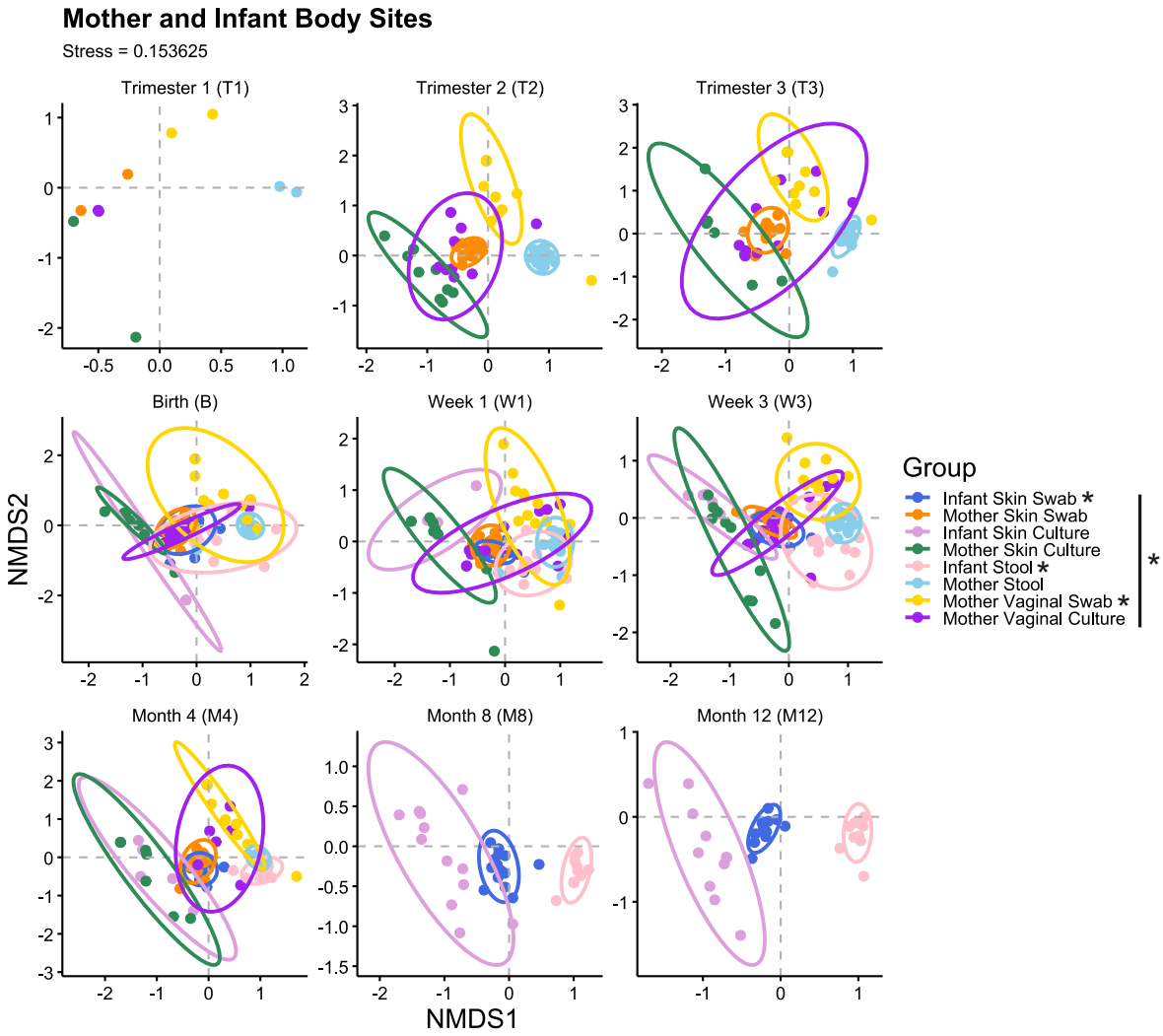


Figure 4.21. Mother and Infant Genus Level Beta Diversity Changes Over Time at Multiple Body Sites. Diversity data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures) using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Using the relative abundance values, beta diversity data was generated and displayed in NMDS plots. The Bray-Curtis metric was used to measure beta diversity. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. The Group key allocates a unique colour to indicate the maternal and infant body sites represented in the ellipses. An asterisk (*) indicates the infant and mother body sites with PERMANOVA significant differences ($P < 0.05$) in beta diversity changes over time, and a line with an asterisk (*) indicates a significant difference ($P < 0.05$) over time across all the groups combined. The stress level is measure of how well the clustering has worked and should be 0.2 or less for a good result.

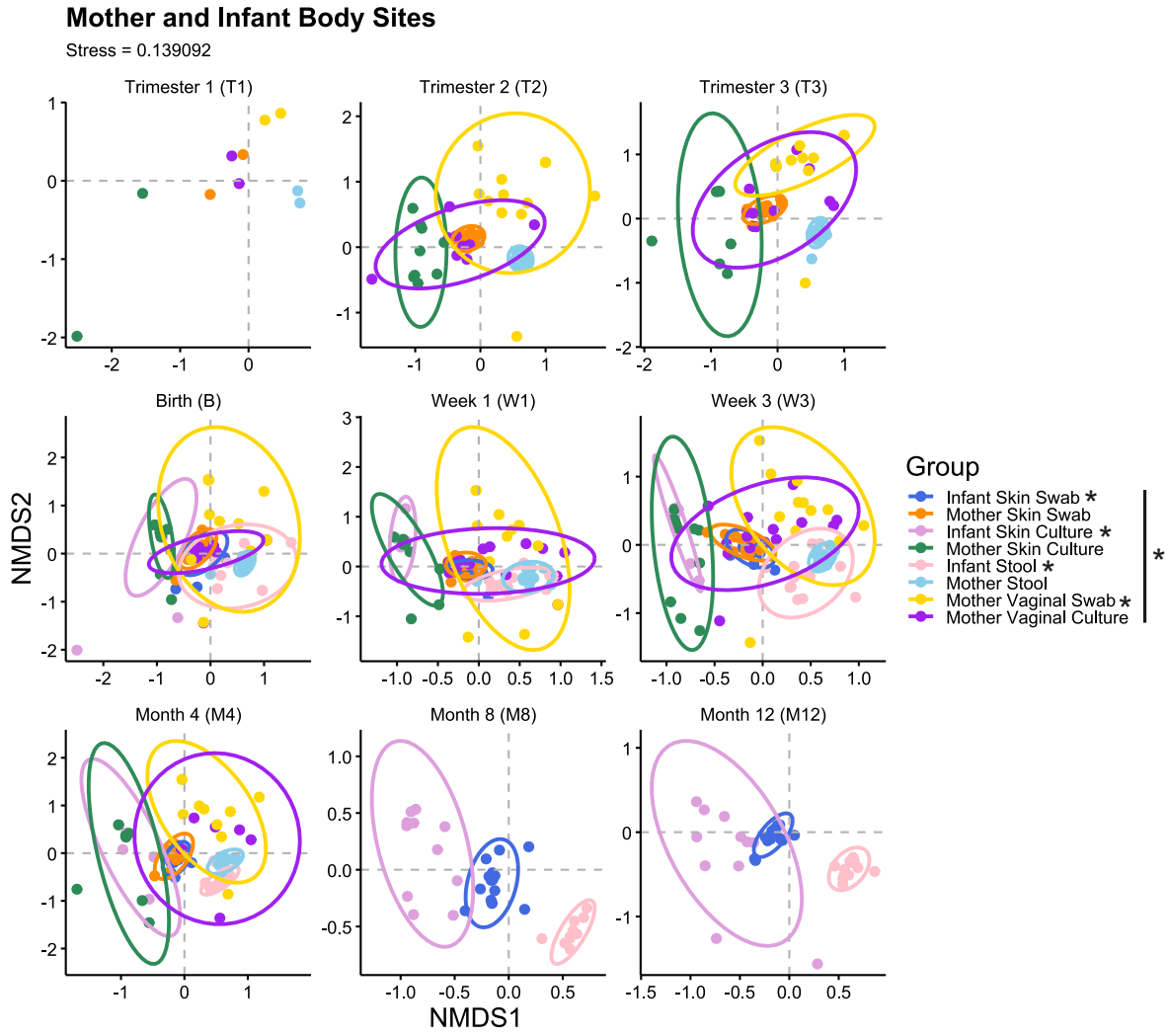


Figure 4.22. Mother and Infant Species Level Beta Diversity Changes Over Time at Multiple Body Sites. Diversity data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures) using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Using the relative abundance values, beta diversity data was generated and displayed in NMDS plots. The Bray-Curtis metric was used to measure beta diversity. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. The Group key allocates a unique colour to indicate the maternal and infant body sites represented in the ellipses. An asterisk (*) indicates the infant and mother body sites with PERMANOVA significant differences ($P < 0.05$) in beta diversity changes over time, and a line with an asterisk (*) indicates a significant difference ($P < 0.05$) over time across all the groups combined. The stress level is measure of how well the clustering has worked and should be 0.2 or less for a good result.

Table 4.5. Beta Diversity Index PERMANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time

Datasets	Genus P Values	Species P Values
Individual	**Bray-Curtis**	**Bray-Curtis**
BSS	*P≤0.001*	*P≤0.001*
MSS		
BST	*P≤0.001*	*P≤0.001*
MST		
MVS	*P≤0.001*	*P≤0.05*
MVC		
BSC		*P≤0.05*
MSC		
Between	**Bray-Curtis**	**Bray-Curtis**
BSS-MSS	*P≤0.01*	*P≤0.01*
BSS-BST	*P≤0.01*	*P≤0.01*
BSS-MST	*P≤0.01*	*P≤0.01*
BSS-MVS	*P≤0.01*	*P≤0.01*
BSS-MVC	*P≤0.01*	*P≤0.01*
BSS-BSC	*P≤0.01*	*P≤0.01*
BSS-MSC	*P≤0.01*	*P≤0.01*
Across All	*P≤0.001*	*P≤0.001*

Table 4.6. Beta Diversity Index PERMANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Pregnancy Term

Datasets	Genus P Values	Species P Values
Individual	**Bray-Curtis**	**Bray-Curtis**
BSS	*P≤0.001*	*P≤0.001*
MSS		
BST	*P≤0.001*	*P≤0.001*
MST		
MVS	*P≤0.001*	*P≤0.05*
MVC		
BSC		*P≤0.05*
MSC		*P≤0.05*
Between	**Bray-Curtis**	**Bray-Curtis**
BSS-MSS	*P≤0.01*	*P≤0.01*
BSS-BST	*P≤0.01*	*P≤0.01*
BSS-MST	*P≤0.01*	*P≤0.01*
BSS-MVS	*P≤0.01*	*P≤0.01*
BSS-MVC	*P≤0.01*	*P≤0.01*
BSS-BSC	*P≤0.01*	*P≤0.01*
BSS-MSC	*P≤0.01*	*P≤0.01*
Across All	*P≤0.001*	*P≤0.001*
LP vs FT	**Bray-Curtis**	**Bray-Curtis**
BSS	*P≤0.01*	*P≤0.001*
MSS	P≤0.001	*P≤0.001*
BST		
MST	*P≤0.01*	*P≤0.001*
MVS		*P≤0.001*
MVC		*P≤0.05*
BSC		
MSC		

Table 4.7. Beta Diversity Index PERMANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Delivery Mode

Datasets	Genus P Values	Species P Values
Individual	**Bray-Curtis**	**Bray-Curtis**
BSS	*P≤0.001*	*P≤0.001*
MSS		
BST	*P≤0.001*	*P≤0.001*
MST		
MVS	*P≤0.001*	*P≤0.05*
MVC		
BSC		*P≤0.05*
MSC		
Between	**Bray-Curtis**	**Bray-Curtis**
BSS-MSS	*P≤0.01*	*P≤0.01*
BSS-BST	*P≤0.01*	*P≤0.01*
BSS-MST	*P≤0.01*	*P≤0.01*
BSS-MVS	*P≤0.01*	*P≤0.01*
BSS-MVC	*P≤0.01*	*P≤0.01*
BSS-BSC	*P≤0.01*	*P≤0.01*
BSS-MSC	*P≤0.01*	*P≤0.01*
Across All	*P≤0.001*	*P≤0.01*
CS vs VB	**Bray-Curtis**	**Bray-Curtis**
BSS		*P≤0.05*
MSS	P≤0.01	*P≤0.01*
BST		
MST	*P≤0.05*	*P≤0.05*
MVS	*P≤0.05*	*P≤0.05*
MVC		*P≤0.01*
BSC		
MSC		

Table 4.8. Beta Diversity Index PERMANOVA P Values for Significant Changes in the Maternal and Infant Datasets Over Time Accounting for Feeding Mode

Datasets	Genus P Values	Species P Values
Individual	**Bray-Curtis**	**Bray-Curtis**
BSS	*P≤0.001*	*P≤0.001*
MSS		
BST	*P≤0.001*	*P≤0.001*
MST		
MVS	*P≤0.001*	*P≤0.05*
MVC		
BSC		*P≤0.05*
MSC		
Between	**Bray-Curtis**	**Bray-Curtis**
BSS-MSS	*P≤0.01*	*P≤0.01*
BSS-BST	*P≤0.01*	*P≤0.01*
BSS-MST	*P≤0.01*	*P≤0.01*
BSS-MVS	*P≤0.01*	*P≤0.01*
BSS-MVC	*P≤0.01*	*P≤0.01*
BSS-BSC	*P≤0.01*	*P≤0.01*
BSS-MSC	*P≤0.01*	*P≤0.01*
Across All	*P≤0.001*	*P≤0.001*
BR vs BR/FO	**Bray-Curtis**	**Bray-Curtis**
BSS		*P≤0.05*
MSS	P≤0.01	*P≤0.01*
BST	*P≤0.05*	*P≤0.05*
MST	*P≤0.01*	*P≤0.01*
MVS	*P≤0.05*	*P≤0.05*
MVC	*P≤0.01*	*P≤0.01*
BSC		
MSC		

4.6.4 Shared Microbiota Over Time between Maternal and Infant Skin, Stool and Vaginal Body Sites

The common genera and species across all the maternal and infant body sites (skin, stool, vaginal and culture), combining all individuals, were also identified (**Figures 4.23-4.24**); the cultures were included in this analysis to enable identification of culturable microbiota of interest. Common genera identified across all the datasets over time included *Staphylococcus*, *Streptococcus*, *Micrococcus* and *Bifidobacterium*. Species of interest included *S. epidermidis* and *M. luteus.* Of these common shared genera and species identified across all datasets over time, trends were identified. From T1 to T3, across the maternal datasets, there was an overall trend of *Staphylococcus* (*S. epidermidis*) relative abundance decreasing and *Bifidobacterium*, *Streptococcus* and *Micrococcus* (*M. luteus*) increasing. From B to M4, across the maternal and infant datasets, *Staphylococcus* (*S. epidermidis*) decreased, which then increased from M8 to M12. From B to M8 the abundance of *Streptococcus* and *Bifidobacterium* increased and then decreased at M12. These trends are likely indicative of weaning the infants off breast milk and a possible start point for transition from an infant microbiome to an adult microbiome. *Micrococcus* (*M. luteus*) abundance slightly fluctuated over time across the maternal and infant datasets. These common microbiota over time across all maternal and infant datasets may indicate the occurrence of microbial sharing between mother and infant pairs across different body sites (skin, gut and vagina) over time (all of which should be culturable from the culture collection); and may have a functional importance in the developing infant microbiomes.

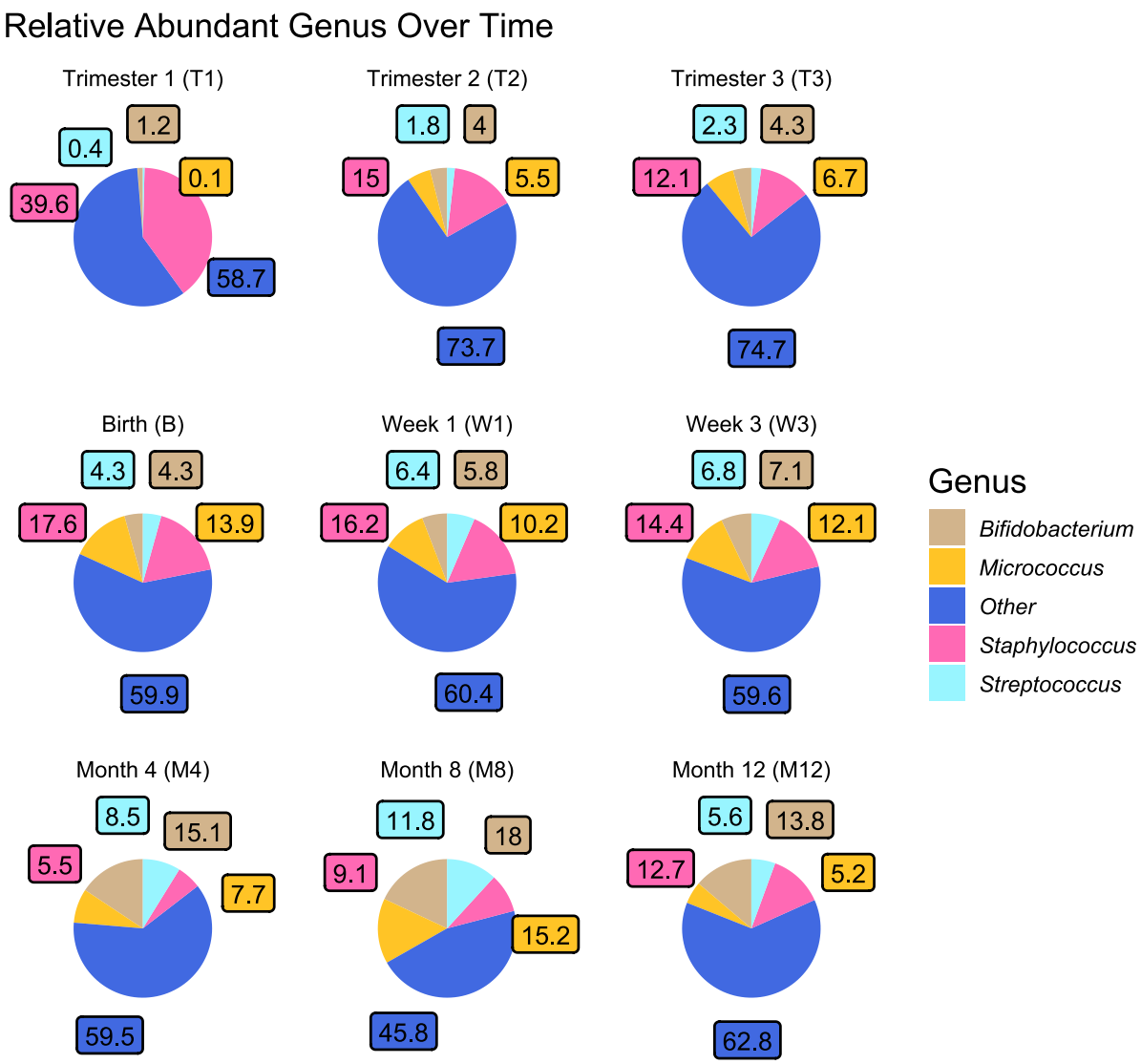


Figure 4.23. Common Shared Relative Abundant Genera Over Time Across all Maternal and Infant Datasets. Relative abundance data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures), using shotgun metagenomics sequencing, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. All individuals and body sites (skin, stool, vaginal and cultures) were combined to calculate the mean relative abundance (%) for each of the common genera (represented in the pie chart slices) at each timepoint (represented in each pie chart) in the maternal and infant population; cultures were included to identify what microbiota are culturable from these body sites. The coloured labels matching the pie chart slices indicates the genus relative abundance (%). The Genus key allocates a unique colour to the genera represented in the pie charts.

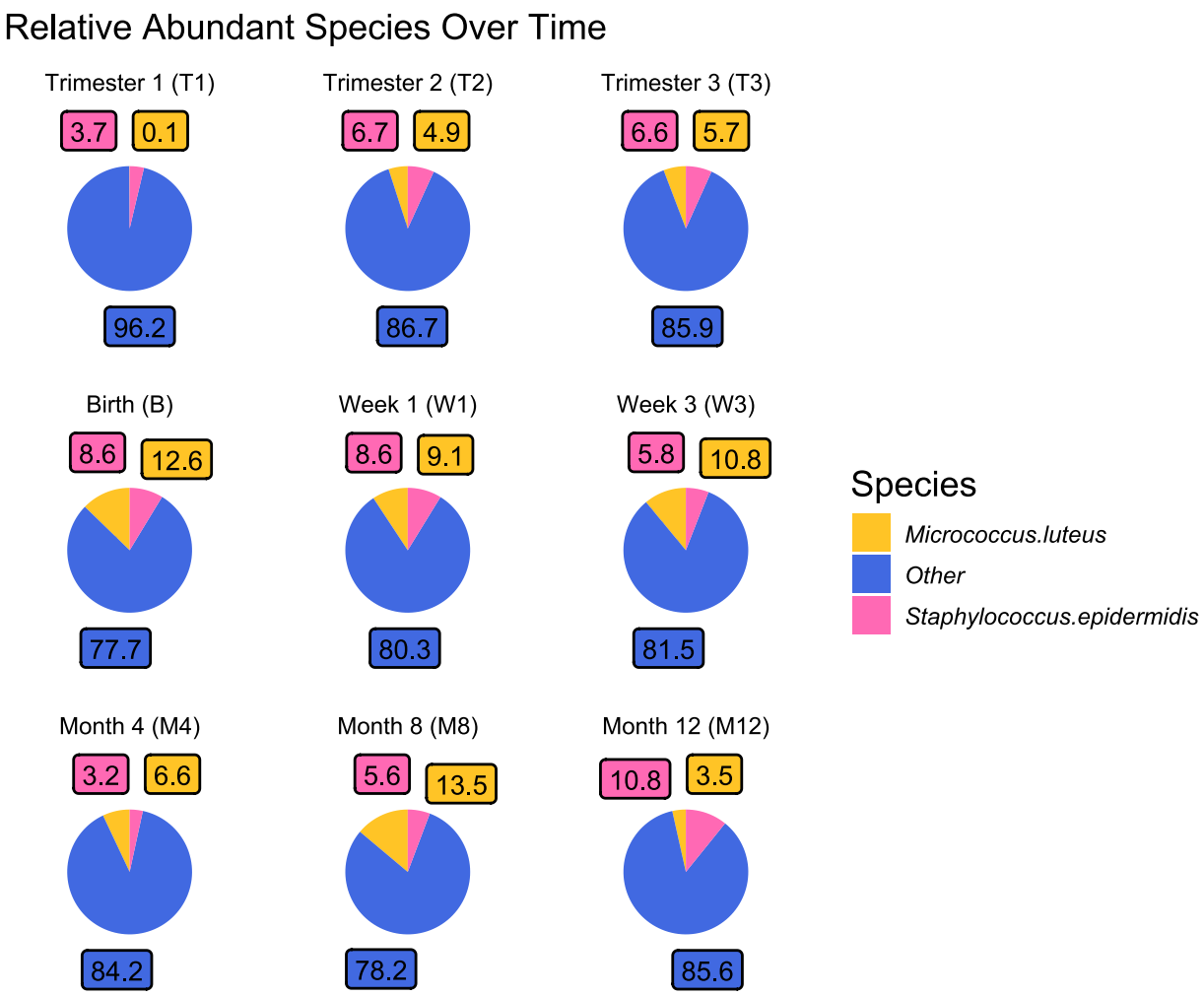


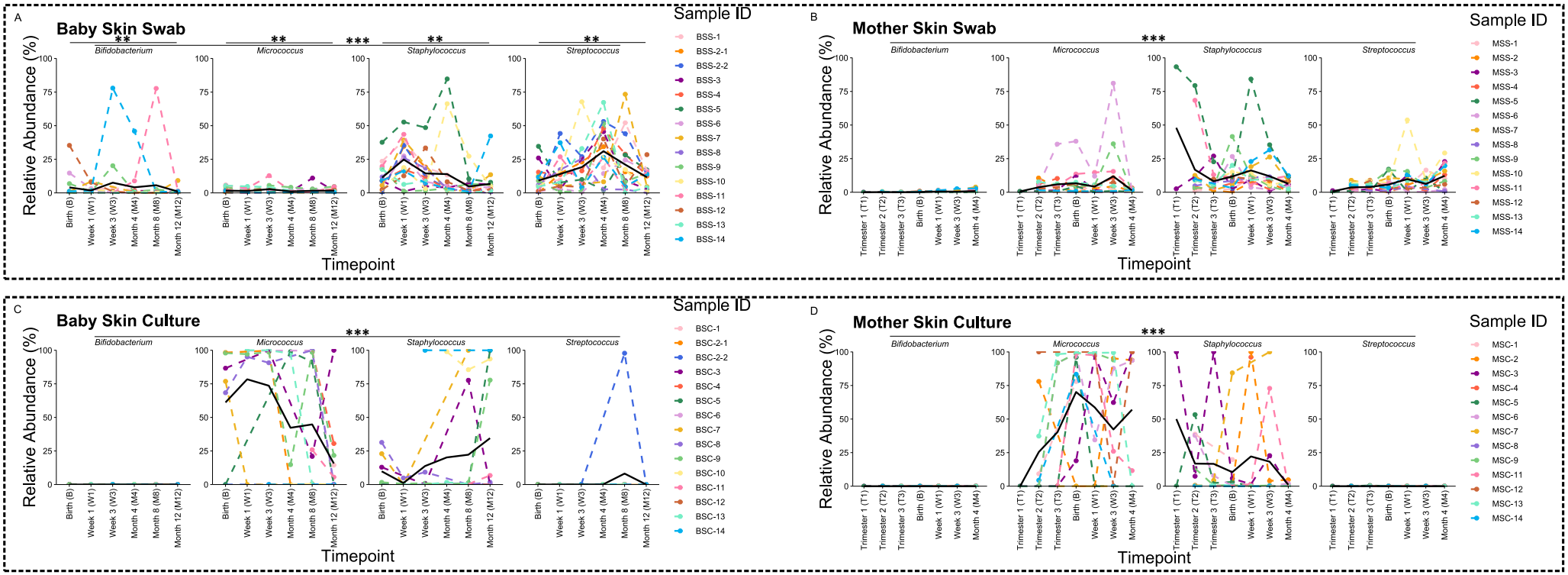
Figure 4.24. Common Shared Relative Abundant Species Over Time Across all Maternal and Infant Datasets. Relative abundance data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures), using shotgun metagenomics sequencing, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. All individuals and body sites (skin, stool, vaginal and cultures) were combined to calculate the mean relative abundance (%) for each of the common species (represented in the pie chart slices) at each timepoint (represented in each pie chart) in the maternal and infant population; cultures were included to identify what microbiota are culturable from these body sites. The coloured labels matching the pie chart slices indicates the species relative abundance (%). The Species key allocates a unique colour to the species represented in the pie charts.

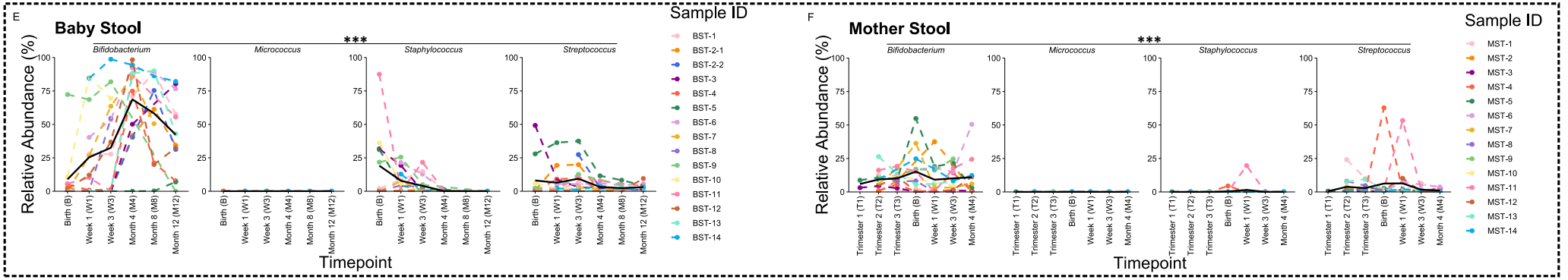
Relative abundance changes over time for each of the common shared genera (*Staphylococcus*, *Streptococcus*, *Micrococcus* and *Bifidobacterium*) and species (*S. epidermidis* and *M. luteus*) across all maternal and infant datasets were identified, accounting for individual variation (**Figure 4.25-4.26**). Significant differences were found in the relative abundance of each common genera over time in the infant skin swab dataset ($P < 0.01$), and common species in the infant stool dataset ($P < 0.01$); relative abundance between the common genera in all maternal and infant datasets over time ($P < 0.001$; maternal vaginal swab dataset $P < 0.05$), and between the common species in the infant skin swab, maternal skin culture, maternal vaginal culture ($P < 0.001$), infant skin culture and infant stool datasets ($P < 0.01$) over time; relative abundance of the common genera over time between the infant skin swab and maternal stool datasets, and the infant skin swab and maternal vaginal swab datasets ($P < 0.001$); and the overall relative abundance of the common genera and species over time across all maternal and infant skin, stool, vaginal and culture datasets ($P < 0.001$).

These trends were still apparent when accounting for pregnancy term, delivery and feeding mode. When accounting for delivery mode (VD vs CS), there were additional significant differences in the relative abundance of each common genera over time in the maternal skin swabs ($P < 0.01$). There was a significant difference in the overall relative abundance of the common genera over time between LP and FT mothers and infants in the maternal skin swabs and infant stools ($P < 0.01$), and common species over time between LP and FT mothers in the maternal vaginal cultures ($P < 0.01$). There was a significant difference in the overall relative abundance of the common genera over time between solely BR and BR/FO combined feeding mothers and infants in the maternal skin swabs ($P < 0.05$), maternal vaginal cultures ($P < 0.01$) and infant stools ($P < 0.05$), and of the common species over time between solely BR feeding and BR/FO combined feeding infants in the infant stools ($P < 0.05$).

With a focus on the infant microbiome, these trends indicate age as a defining factor for the relative abundance changes of the common genera and species on infant skin and within infant stools. The influence of pregnancy term and feeding mode on the relative abundance changes of the shared microbes over time in infant stools could indicate the involvement of gestational age and breast feeding in the establishment and development of the infant skin and gut microbiomes. Delivery mode did not appear to have an influence.

Mother and Infant Genus Relative Abundance Changes Over Time

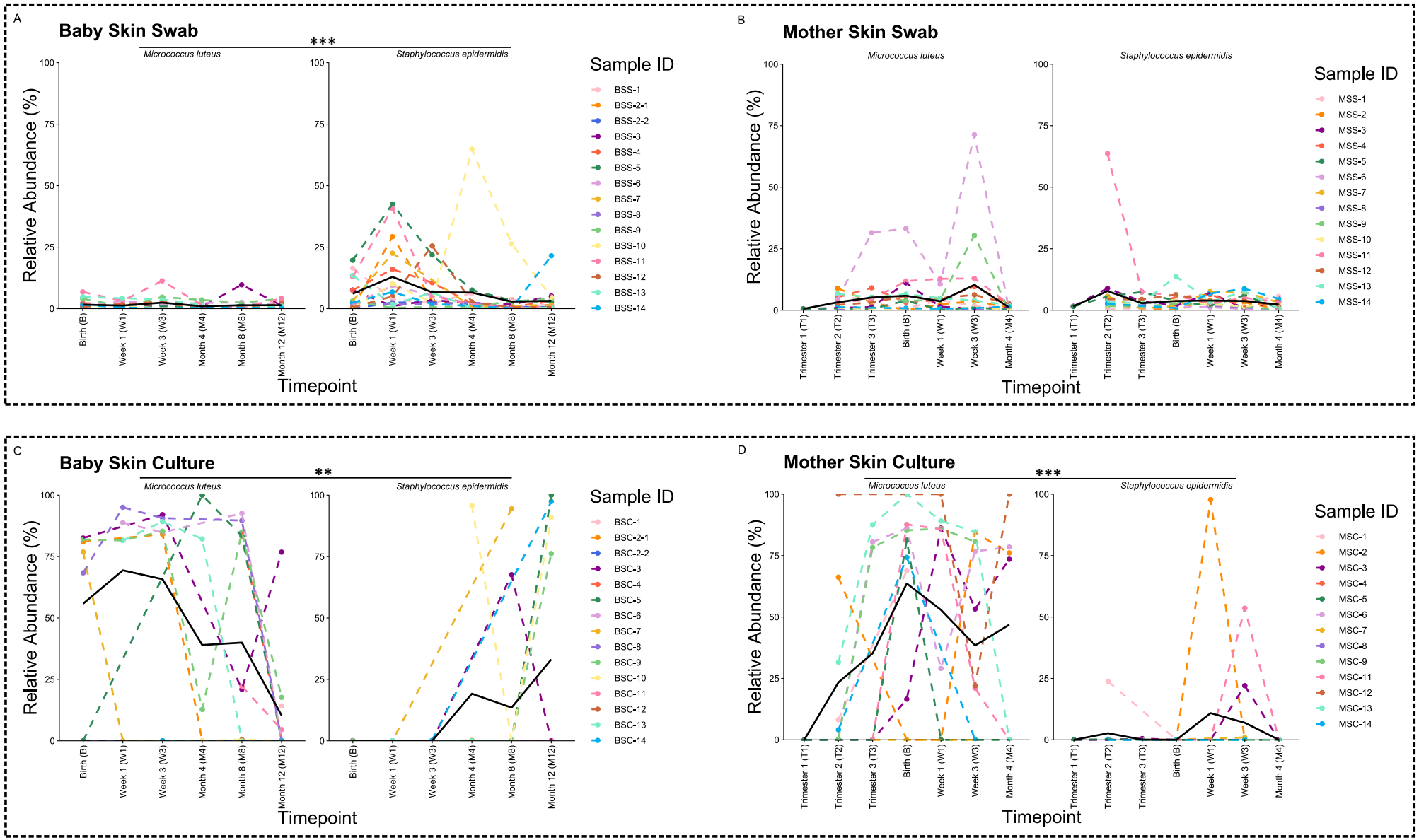


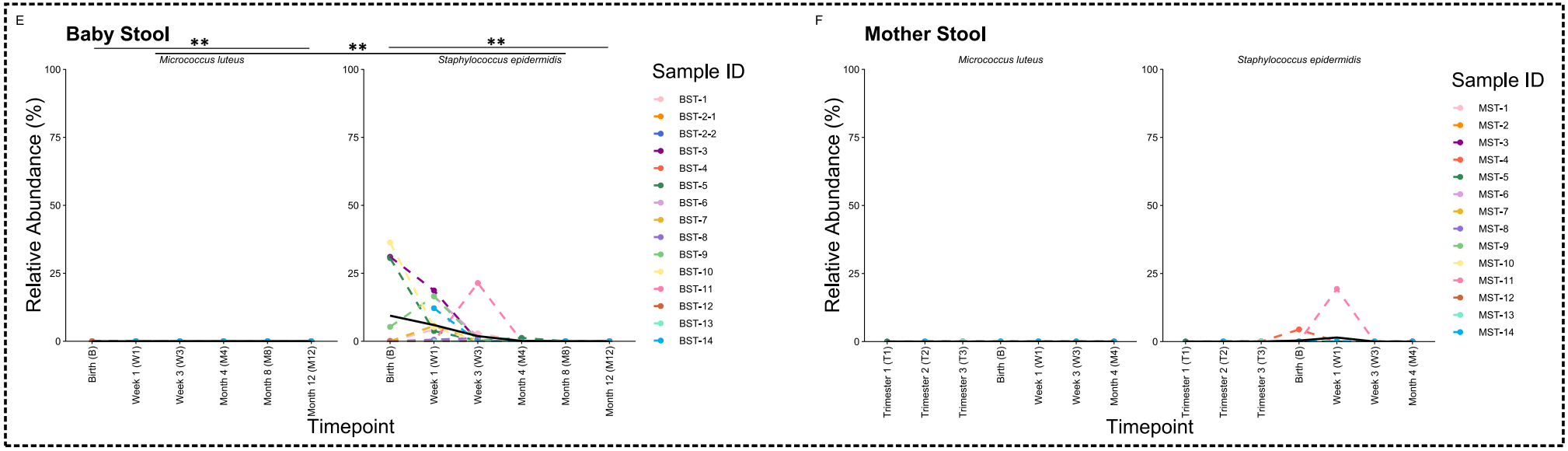


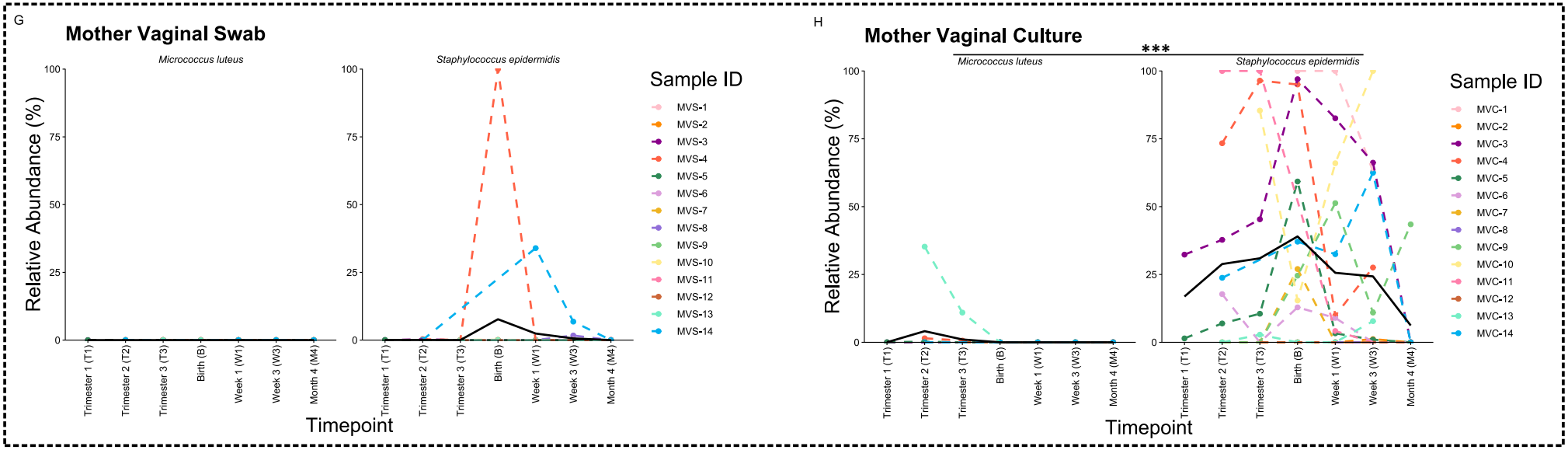
Baby Skin Swab vs Mother Skin Swab
Baby Skin Swab vs Baby Stool
Baby Skin Swab vs Mother Stool ***
Baby Skin Swab vs Mother Vaginal Swab ***

Figure 4.25. Individual Variation of the Relative Abundance Changes Over Time of the Common Shared Genera in Maternal and Infant Datasets. Relative abundance data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures), using shotgun metagenomics sequencing, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. The relative abundance (%) for each of the common genera at each timepoint, within each mother and infant group, were plotted as spaghetti plots. The black line intersecting each plot shows the average, each sample ID has been allocated a unique colour and the lines with an asterisk (*) show the ANOVA significant differences (* = ≤ 0.05; ** = ≤ 0.01; *** = ≤ 0.001).

Mother and Infant Species Relative Abundance Changes Over Time







Baby Skin Swab vs Mother Skin Swab
Baby Skin Swab vs Baby Stool
Baby Skin Swab vs Mother Stool
Baby Skin Swab vs Mother Vaginal Swab

Figure 4.26. Individual Variation of the Relative Abundance Changes Over Time of the Common Shared Species in Maternal and Infant Datasets. Relative abundance data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures), using shotgun metagenomics sequencing, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. The relative abundance (%) for each of the common species at each timepoint, within each mother and infant group, were plotted as spaghetti plots. The black line intersecting each plot shows the average, each sample ID has been allocated a unique colour and the lines with an asterisk (*) show ANOVA significant differences (* = ≤ 0.05; ** = ≤ 0.01; *** = ≤ 0.001).

Genus and species level top 50 significant differential abundance associations over time across infant skin (and all other maternal and infant datasets) were identified, also accounting for individual variation (**Figure 4.27**). With a focus on the common shared genera (*Staphylococcus*, *Streptococcus*, *Micrococcus* and *Bifidobacterium*) and species (*S. epidermidis* and *M. luteus*) identified previously between mothers and infants, only *Streptococcus* was identified in the top 50 significant associations when B and infant skin were used as the reference timepoint and body site (**Figure 4.27. A.**).

Streptococcus was weakly negatively associated with infant and maternal stool, and strongly negatively associated with maternal and infant skin cultures, and maternal vaginal swab and cultures. *Streptococcus* was not identified in the top 50 significant associations with maternal skin or pregnancy term, delivery or feeding mode. But *Streptococcus* was weakly positively associated with certain timepoints (W1, W3 and M4) across the maternal and infant body sites. *Streptococcus* was also weakly positively associated with M8 and M12, which are timepoints only found across the infant body sites (skin and stool). When using M4 and maternal skin as the reference timepoint and body site, there were some timepoint specific associations with other genera within and across maternal and infant body sites. But there were no specific maternal and infant body site associations with genera in the top 50 significant associations identified (**Figure 4.27. B.**).

There were also multiple *Streptococcus* species (*S. mitis, S. oralis, Streptococcus gwangjuense, Streptococcus sp. oral. taxon. 431, S. pseudopneumoniae, Streptococcus australis, S. pneumoniae*, and *S. sanguinis*) identified in the top 50 significant associations within and across the maternal and infant body sites over time (**Figure 4.27. C.** & **D.**). When using B and infant skin as the reference timepoint and body site (**Figure 4.27. C.**), some *Streptococcus* species identified were weakly positively associated with timepoints W1 (*S. mitis*, *S. oralis, Streptococcus sp. oral. taxon. 431,* S. *gwangjuense*, *S. pseudopneumoniae, S. australis* and *S. pneumoniae*) and W3 (*S. mitis*, *S. oralis,* S. *gwangjuense*, *S. pseudopneumoniae, S. australis* and *S. pneumoniae*) in both mothers and infants, and more strongly positively associated at M4 in both groups; the positive association signal for *S. pneumoniae, S. oralis* and *Streptococcus sp. oral. taxon. 431* at M4 was weaker than the other species. Other *Streptococcus* species identified were weakly positively associated with timepoints M8 (*S. mitis*, *Streptococcus sp. oral. taxon. 431,* S. *gwangjuense*, *S. pseudopneumoniae, S. australis* and *S. pneumoniae*) and

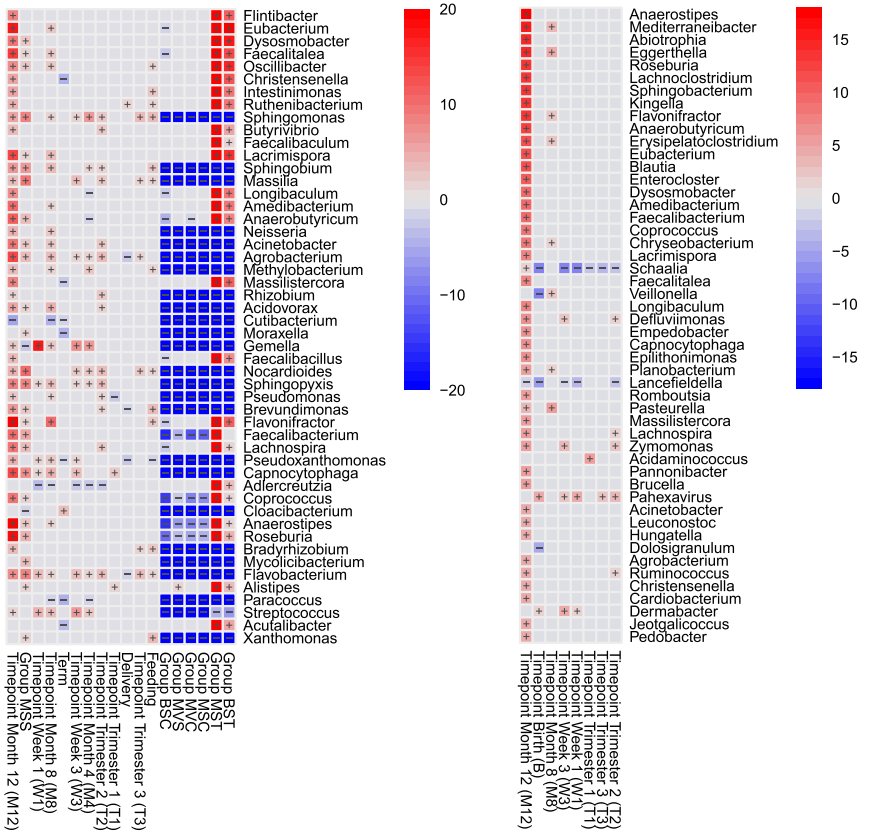
M12 (*Streptococcus sp. oral. taxon. 431, S. mitis, S. pseudopneumoniae* and *S. australis*), which are infant specific timepoints, and T2 (*S. australis*) and T3 (*Streptococcus sp. oral. taxon. 431*), which are maternal specific timepoints. All these *Streptococcus* species were strongly negatively associated with all maternal and infant body sites, except for maternal skin swab where, *S. mitis*, S. *gwangjuense*, *Streptococcus sp. oral. taxon. 431*, *S. pseudopneumoniae*, S. *australis*, and *S. pneumoniae* had a weak negative association. Some Streptococcus species were not identified in the top 50 significant associations for pregnancy term (*S. mitis*, *S. oralis*, *Streptococcus sp. oral. taxon. 431, S. pseudopneumoniae*, and *S. pneumoniae*), delivery (*S. mitis*, *S. oralis*, *S. pseudopneumoniae*, S. *gwangjuense* and *S. pneumoniae*) or feeding mode (*S. mitis*, *S. oralis*, *Streptococcus sp. oral. taxon. 431, S. pseudopneumoniae*, S. *gwangjuense* and *S. pneumoniae*). *Streptococcus sp. oral. taxon. 431* and *S. australis* were weakly positively associated with delivery mode. *S. australis* and S. *gwangjuense* were also weakly negatively associated with pregnancy term.

When using M4 and maternal skin as the reference timepoint and body site (**Figure 4. 27. D.**), some *Streptococcus* species identified were strongly (*S. mitis*) and weakly (*S. oralis* and *Streptococcus sp. oral. taxon. 431*) negatively associated with the timepoint B, W3 (*S. mitis, S. oralis* and *Streptococcus sp. oral. taxon. 431*), and W1 (*S. mitis*) in both mothers and infants. *S. sanguinis* was weakly positively associated with M12, which is an infant specific timepoint. Some species were also weakly negatively associated with maternal specific timepoints T1 (*S. sanguinis, S. mitis, S. oralis* and *Streptococcus sp. oral. taxon. 431*), T2 (*S. mitis, S. oralis* and *Streptococcus sp. oral. taxon. 431*) and T3 (*S. mitis*). *S. mitis*, *S. oralis* and *Streptococcus sp. oral. taxon. 431* followed the same group associations as identified when using B and infant skin as the reference timepoint and body site. *S. mitis* and *Streptococcus sp. oral. taxon. 431* had an additional weak positive association found with infant skin, while *S. oralis* and *S. sanguinis* had no association on infant skin. *S. mitis, S. oralis*, *S. sanguinis* had no association with pregnancy term, feeding and delivery mode. *Streptococcus sp. oral. taxon. 431* also had no associations with these factors except for a weak negative association with delivery mode.

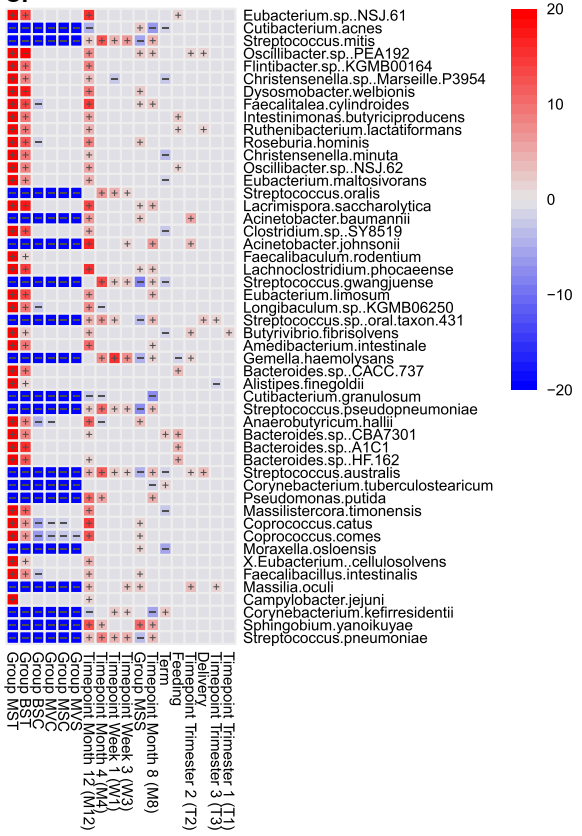
These results suggest that timepoint (age) was a defining factor on *Streptococcus* abundance on the infant skin and gut microbiome, whilst pregnancy term (gestational age), delivery and feeding mode were not largely influential.

Top 50 features with significant associations (−log(qval)*sign(coeff))

Mother and Infant Differential Abundant Genera



Mother and Infant Differential Abundant Species



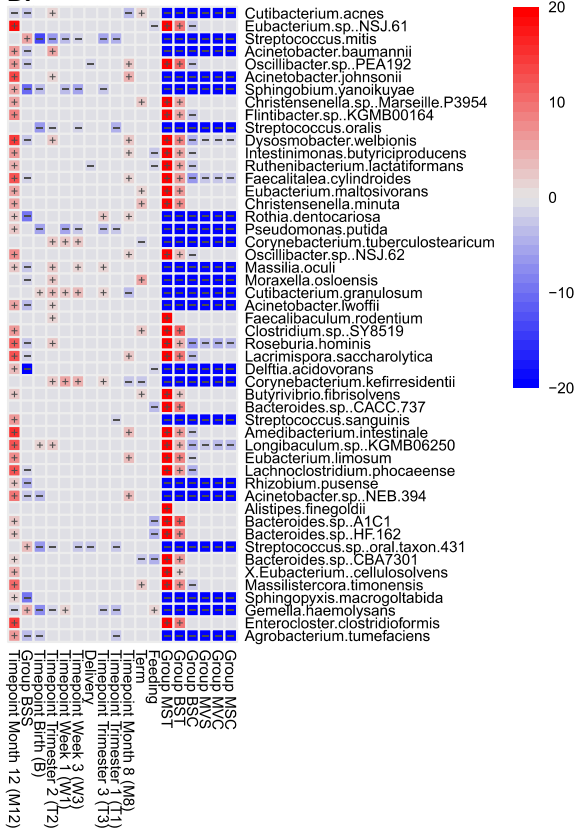


Figure 4.27. Top 50 Differential Abundant Genera and Species Significant Associations Over Time Within and Across Mother and Infant Body Sites. Relative abundance data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, skin cultures, stools, and maternal vaginal swabs and cultures), using shotgun metagenomics sequencing, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. The relative abundance values were run through MaAsLin2 to obtain significant differential abundant genera and species associations over time, accounting for individual variation and multiple factors. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Factors accounted for included pregnancy term (Late Preterm [LP] and Full Term [FT]), delivery mode (Vaginal Delivery [VD] and Caesarean Section [CS]) and feeding mode (Breast [BR] and Breast/Formula [BR/FO]). The key allocates a significance scale (-log[q-value]*sign[coefficient]) with a unique colour and direction to positive (red [+]) and negative (blue [-]) microbial associations. **A.** Mother and infant multi body site differential abundant genera significant associations over time using birth as a reference timepoint. **B.** Mother and infant multi body site differential abundant genera significant associations over time using month 4 as a reference timepoint. **C.** Mother and infant multi body site differential abundant species significant associations over time using birth as a reference timepoint. **D.** Mother and infant multi body site differential abundant species significant associations over time using month 4 as a reference timepoint.

4.6.5 Strain Level Sharing of the Common Microbiota between Maternal and Infant Skin, Stool and Vaginal Body Sites Over Time

For the strain level analysis 4/582 samples were excluded from the analysis (BSS-12-W1, BSS-12-W3, MVS-9-W3, BSC-2-2-M12) as the sequencing was not deep enough to correctly cover any markers. In case of when several metagenomes were available for one sample, the replicate with the higher number of reads were kept, with 566 samples included in the analysis (**Figure 4.28**), allowing full maternal and infant paired sample sets to be defined.

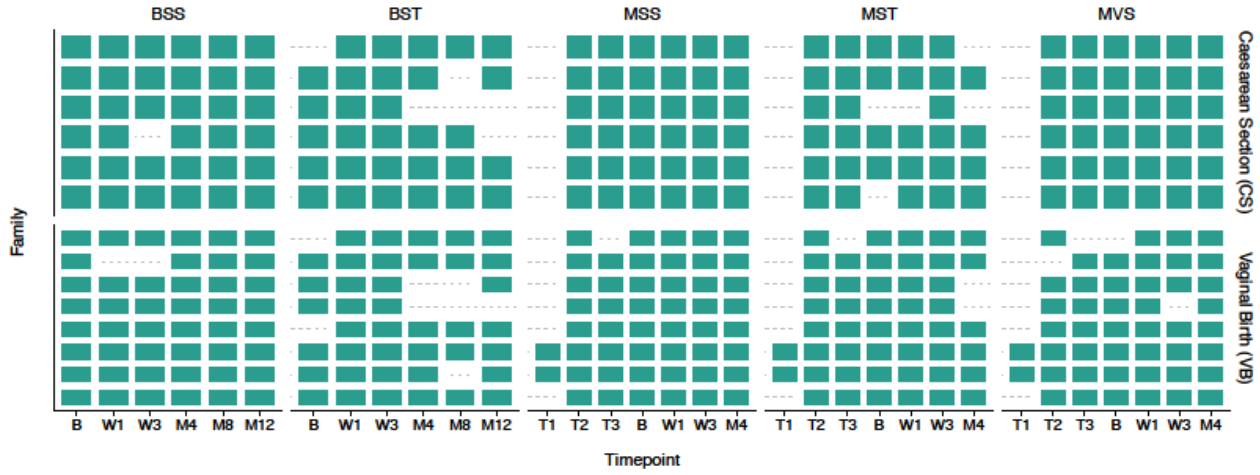


Figure 4.28. Sample Availability Overview. Green tiles represent samples available for the analysis. Shotgun metagenomics data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, stools, and maternal vaginal swabs) using shotgun metagenomics sequencing, generated on the Illumina platform. The shotgun metagenomics data were annotated by MetaPhlAn4 to identify whether the samples contained enough marker genes to be used in the strain analysis. Samples that didn't meet the sequencing depth threshold to cover enough markers for analysis were excluded, leaving a total of n=566 samples, allowing full maternal and infant paired sample sets to be defined. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Factors accounted for included pregnancy term (Late Preterm [LP] and Full Term [FT]), delivery mode (Vaginal Delivery [VD] and Caesarean Section [CS]) and feeding mode (Breast [BR] and Breast/Formula [BR/FO]).

Of the four genera of interest, strain detection identified seven clades of *Bifidobacterium*, one clade of *Micrococcus*, nine clades of *Streptococcus* and ten clades of *Staphylococcus* (**Table 4.9**); see **Appendix 16** for further details on which infant and mother samples these strains were found in, and which were skin specific.

A total of 27 strains and 109 sharing events between mother and infant pairs were identified (**Figure 4.29**).

Table 4.9. Clade ID and nGD Threshold to Detect and Identify Strains

Species Name	Clade ID	nGD Threshold
Micrococcus luteus	*t__SGB16676*	*0.03*
Bifidobacterium dentium	*t__SGB17234*	*0.05*
Bifidobacterium pseudocatenulatum	*t__SGB17237*	*0.05*
Bifidobacterium catenulatum	*t__SGB17241*	*0.03*
Bifidobacterium adolescentis	*t__SGB17244*	*0.00542012*
Bifidobacterium breve	*t__SGB17247*	*0.03*
Bifidobacterium longum	*t__SGB17248*	*0.03413473*
Bifidobacterium bifidum	*t__SGB17256*	*0.0204073*
Staphylococcus simulans	*t__SGB7832*	*0.03*
Staphylococcus lugdunensis	*t__SGB7850*	*0.03*
Staphylococcus aureus	*t__SGB7852*	*0.03*
Staphylococcus warneri	*t__SGB7857*	*0.03*
Staphylococcus hominis	*t__SGB7858*	*0.03*
Staphylococcus haemolyticus	*t__SGB7861*	*0.03*
Staphylococcus caprae	*t__SGB7862*	*0.03*
Staphylococcus capitis	*t__SGB7863*	*0.03*
Staphylococcus epidermidis	*t__SGB7865*	*0.03*
Streptococcus thermophilus	*t__SGB8002*	*0.05*
Streptococcus salivarius	*t__SGB8005*	*0.03*
Streptococcus salivarius	*t__SGB8007_group*	*0.05*
Streptococcus gallolyticus	*t__SGB8018*	*0.03*
Streptococcus lutetiensis	*t__SGB8021*	*0.03*
Streptococcus anginosus	*t__SGB8028_group*	*0.03*
Streptococcus australis	*t__SGB8059_group*	*0.03*
Streptococcus lactarius	*t__SGB8065*	*0.03*
Streptococcus parasanguinis	*t__SGB8071*	*0.03607692*
Streptococcus mitis	*t__SGB8168*	*Present in positive controls only*

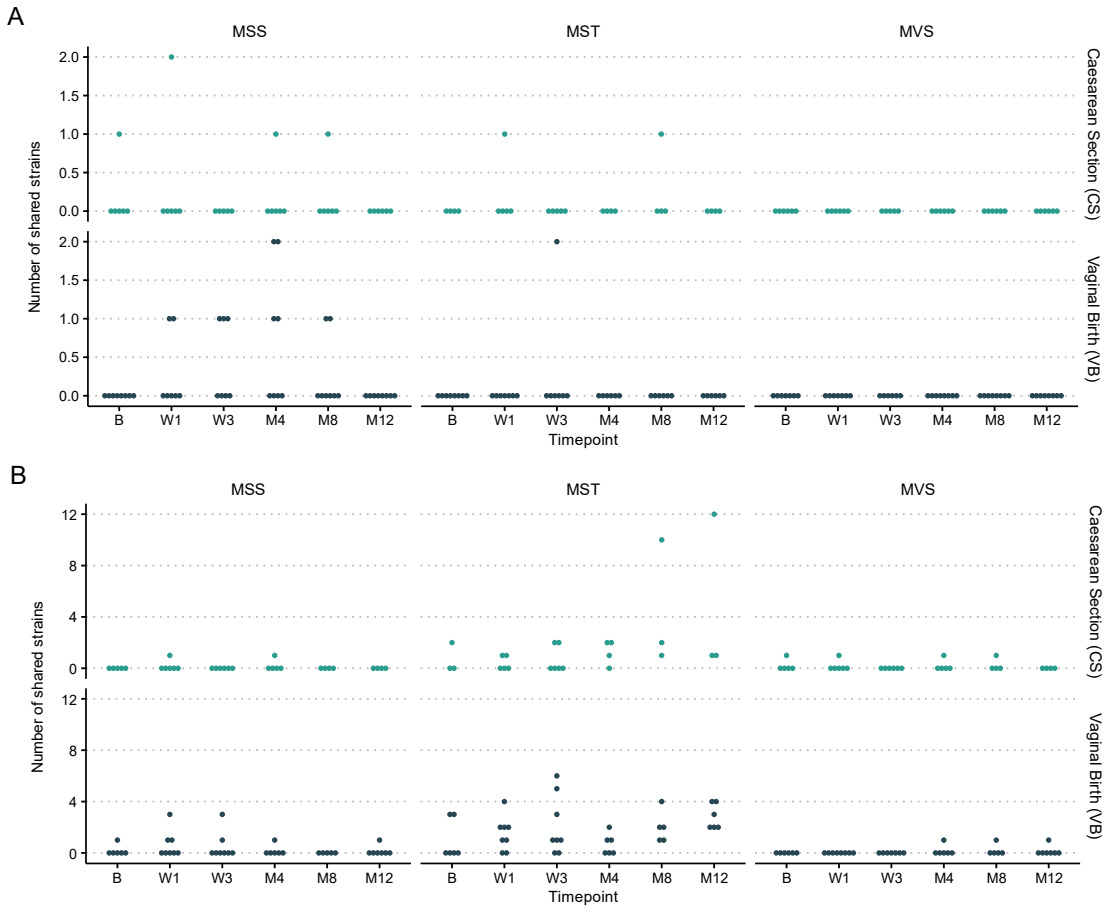


Figure 4.29. Strains Shared between Mother and Infant Pairs. Strain level data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, stools, and maternal vaginal swabs) using shotgun metagenomics sequencing data, generated on the Illumina platform. The shotgun metagenomics data were annotated by MetaPhlAn4 to identify whether the samples contained enough marker genes to be used in the strain analysis. Samples that didn't meet the sequencing depth threshold to cover enough markers for analysis were excluded, leaving a total of n=566 samples, allowing full maternal and infant paired sample sets to be defined. Strains were generated from the shotgun metagenomics data using a read mapping method with MetaPhlAn4 and StrainPhlan identifying 27 strains and 109 sharing events for the four genera of interest (*Bifidobacterium*, *Micrococcus*, *Streptococcus* and *Staphylococcus*). The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Factors accounted for included pregnancy term (Late Preterm [LP] and Full Term [FT]), delivery mode (Vaginal Delivery [VD] and Caesarean Section [CS]) and feeding mode (Breast [BR] and Breast/Formula [BR/FO]). **A.** Number of strains shared between maternal and infant skin samples over time, for the four genera of interest. **B.** Number of strains shared between maternal and infant stool samples over time, for the four genera of interest.

The overall strain transmissibility sharing events of the four genera of interest were identified between the maternal body sites and infant skin, and the maternal body sites and infant stool (focusing on 109 strain level sharing events) (**Figure 4.30**). Sharing events labelled as “Maternal related strain shared” indicated mother and infant pairs containing the same strains from the same species, and those labelled as “Maternal unrelated strain shared” indicated mother and infant pairs containing different strains from the same species.

Sharing events (**Figure 4.30. A.**) were identified between infant and maternal skin swabs (Maternal related strain shared: *Str. salivarius*, *Streptococcus lactarius*, *Sta. hominis*; Maternal unrelated strain shared: *Sta. hominis*, *Sta. epidermidis*, *Mic. luteus*); infant skin swab and maternal stool (Maternal related strain shared: *Sta. epidermidis*, *Bif. longum* and *Bif. bifidum*; Maternal unrelated strain shared: *Bif. longum*); infant stool and maternal skin swab (Maternal related strain shared: *Str. salivarius*, *Str. lactarius*, *Sta. hominis* and *Bif. longum*; Maternal unrelated strain shared: *Sta. hominis* and *Sta. epidermidis*); infant stool and maternal stool (Maternal related strain shared: *Str. salivarius*, *Str. parasanguinis*, *Bif. longum*, *Bifidobacterium dentium*, *Bif. bifidum* and *Bif. adolescentis*; Maternal unrelated strain shared: *Str. salivarius*, *Str. parasanguinis*, *Bif. longum*, *Bif. dentium*, *Bif. bifidum* and *Bif. adolescentis*); infant stool and maternal vaginal swab (Maternal related strain shared: *Bif. breve*; Maternal unrelated strain shared: *Sta. epidermidis* and *Bif. breve*); and no sharing events were identified between infant skin swab and maternal vaginal swab. Fisher tests identified no significant differences ($P > 0.05$) in strain transmissibility in mother and infant pairs between the maternal body sites and infant skin swab, and maternal body sites and infant stool ($P > 0.05$). When accounting for timepoint (**Figure 4.30. B.**), pregnancy term (**Figure 4.30. C.**), delivery (**Figure 4.30. D.**) and feeding mode (**Figure 4.30. E.**), the same sharing events and Fisher test results were identified.

The absence of significant differences in strains detected in mother and infant pairs, was likely due to the small number of mother and infant pairs (15 pairs), even though these pairs represented large sample sets and datasets. There is generally a low strain transmissibility (where the mother and infant pairs contain different strains of the same species) of *Sta. epidermidis* and *Mic. luteus* between maternal and infant body sites; these trends were still apparent when accounting for the factors timepoint, pregnancy term, delivery mode and feeding mode. There was a low strain transmissibility of *Sta. epidermidis* specifically between maternal and infant skin

from W1 to M8, and within the FT, LP, CS, VD, BR/FO and BR groups; between maternal skin and infant stool from W1 to W3, and within the FT, LP, VB, and BR groups (there was no transmission in the CS and BR/FO groups); between maternal vaginal swab and infant stool at W1, and within the FT, VB and BR groups (there was no transmission in the LP, CS and BR/FO groups). There was a low strain transmissibility of *Mic. luteus* specifically between maternal and infant skin from W1 to W3, and within the FT, CS, VD, BR/FO and BR groups.

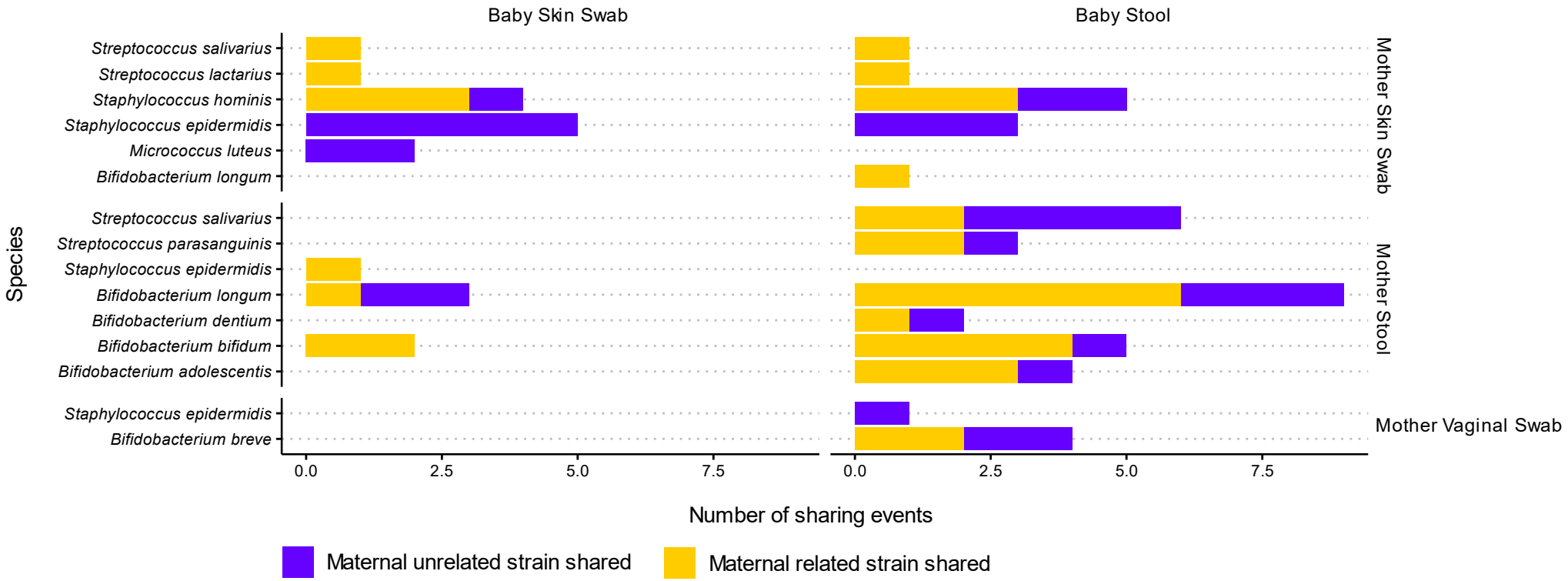
There were also factor specific trends identified. There were no transmission events at B and M12 between maternal and infant body sites, M8 between maternal skin swab and infant stool or B and W3 between maternal vaginal swab and infant stool. There was low strain transmissibility between maternal and infant body sites at particular timepoints (*Sta. hominis* between maternal and infant skin swabs at M4; *Bif. longum* between maternal stool and infant skin swabs at M4 and M8; *Str. salivarius* between maternal and infant stool at M4; *Str. parasanguinis* between maternal and infant stool at M4 and M8; *Bif. dentium* between maternal and infant stool at M12; *Bif. breve* between maternal vaginal swab and stool at W1). When accounting for pregnancy term, there were no transmission events between maternal stool and infant skin swabs in the LP group, and low transmissibility of *Str. salivarius* and *Bif. longum* between maternal and infant stool in the LP group. When accounting for delivery mode, there was low transmissibility between maternal stool and infant skin swabs of *Bif. longum* in the CS group, maternal skin swabs and infant stool of *Sta. hominis* in the CS group. Low transmissibility was also observed between maternal and infant stool of *Str. salivarius* in CS group, *Str. parasanguinis* in the VB group, and *Bif. adolescentis* in the CS group. When accounting for feeding mode, there was low transmissibility of *Bif. longum* between maternal stool and infant skin swabs in the BR group, and *Str. salivarius* between maternal and infant stool in the BR/FO group.

These transmissibility results overall suggest species specific sharing between certain body sites. *Streptococcus*, *Staphylococcus* and *Micrococcus* species were shared between maternal and infant skin. *Staphylococcus* and *Bifidobacterium* species were shared between maternal stool and infant skin however no species-specific sharing was observed between infant skin and maternal vaginal body sites. *Streptococcus*, *Staphylococcus*, and *Bifidobacterium* species were shared between infant stool and maternal skin, *Streptococcus* and *Bifidobacterium* species were shared between infant and maternal stool, and *Staphylococcus* and *Bifidobacterium*

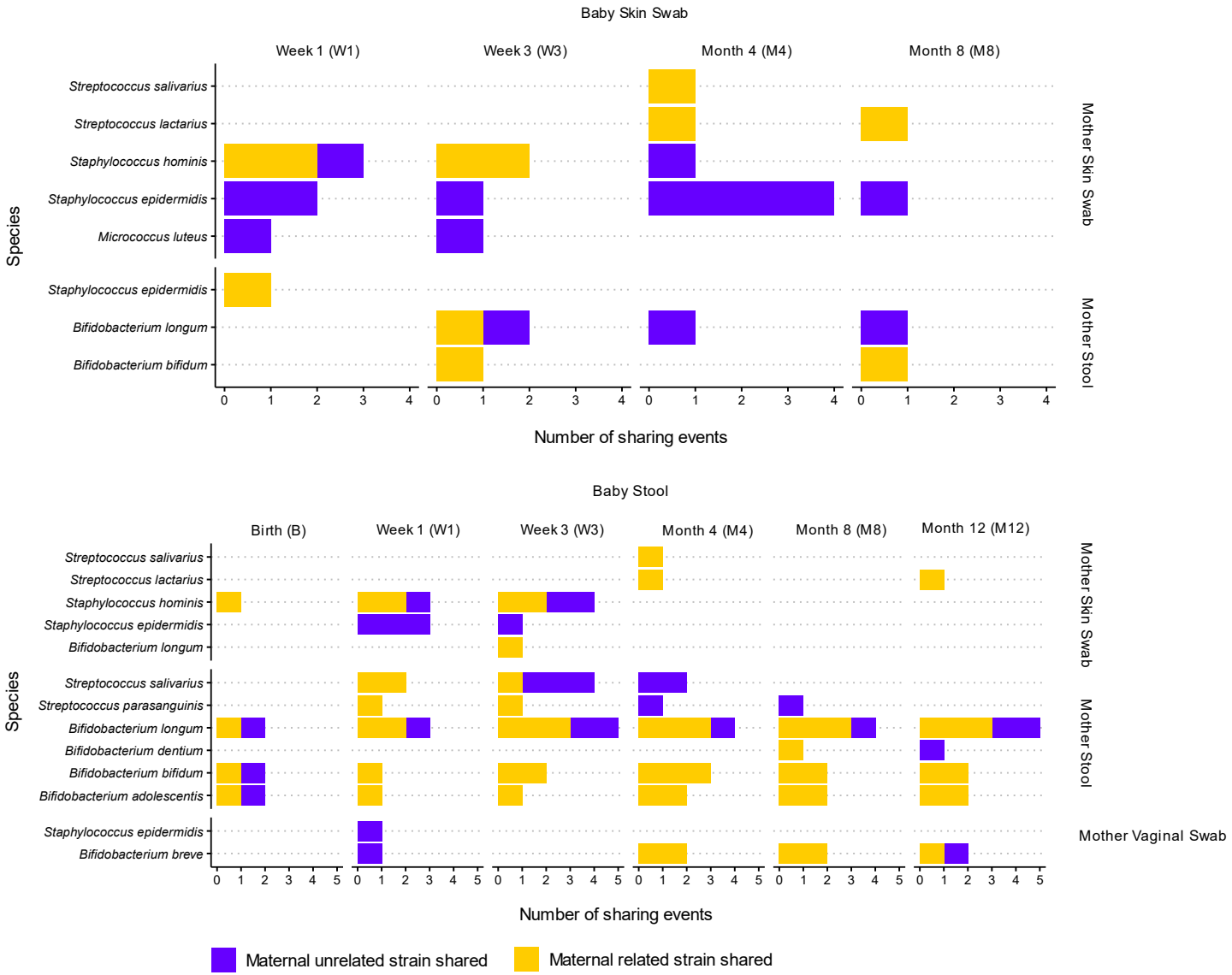
species were shared between infant stool and maternal vaginal body sites. These body site specific species sharing events appear to correlate with the taxonomic profiles previously discussed identifying “typical” maternal and infant skin and gut microbes (*Streptococcus*, *Staphylococcus*, *Micrococcus, Bifidobacterium*) and maternal vaginal microbes (*Staphylococcus* and *Bifidobacterium*). The maternal skin appears to be the main source of microbial transfer to infant skin, followed by the maternal gut, whereas the maternal vaginal body site did not appear to play a part in microbial transfer to the infant skin. The maternal skin also appeared to be the main source of microbial transfer to the infant gut, followed by the maternal gut and vaginal body sites.

Finally, the strains that were shared between mother and infant pairs, were also identified in some culture samples (**Table 4.10**). In total, 27 strains were identified in one or more maternal skin or vaginal culture samples and were found in at least one paired infant sample. Six species of the 27 strains were identified as culturable from the culture metagenome samples (*Bif. breve*, *Sta. aureus*, *Sta. epidermidis*, *Sta. haemolyticus*, *Sta. hominis* and *Str. anginosus*).

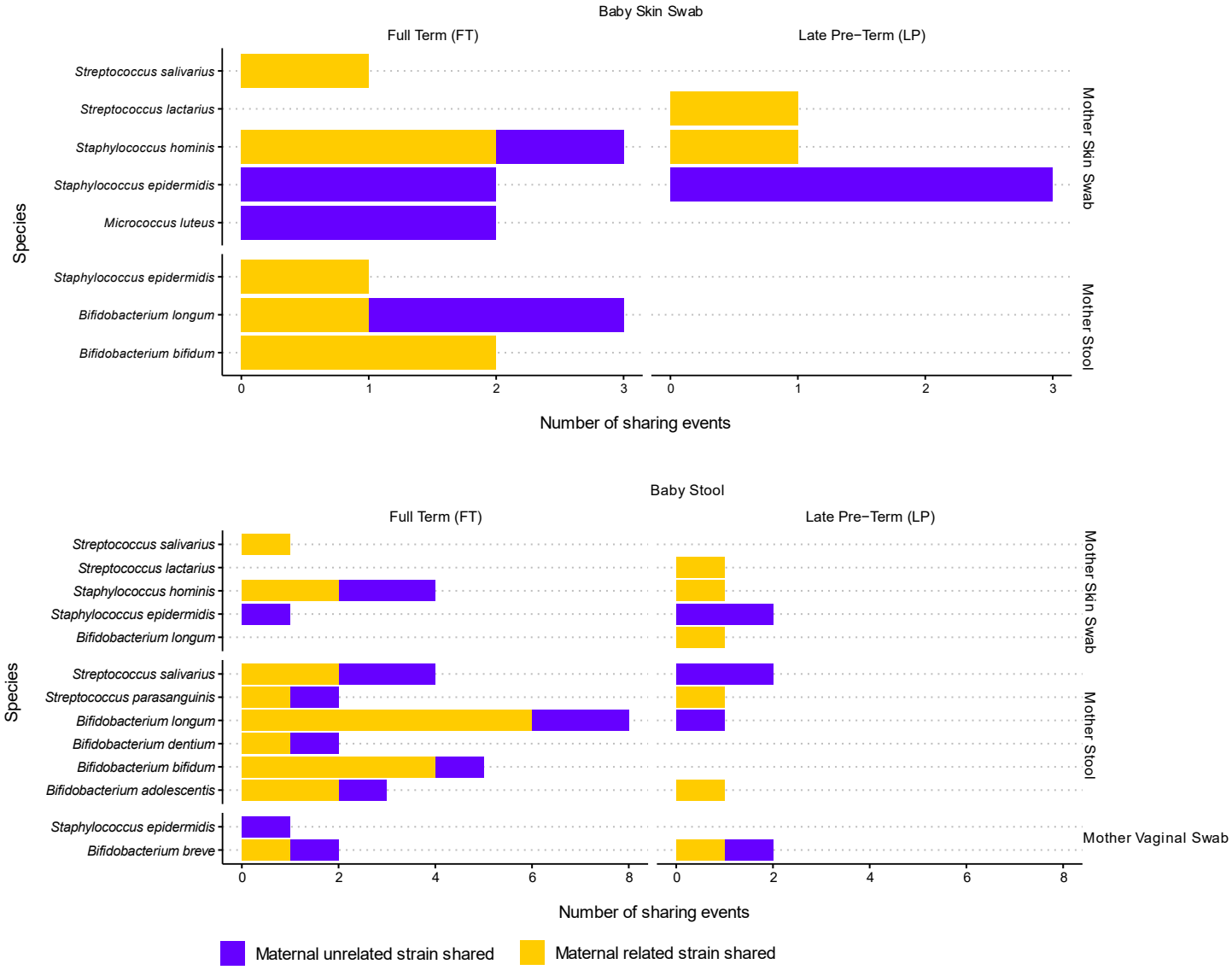
A. Overall Strain Transmissibility between Mothers and Infants



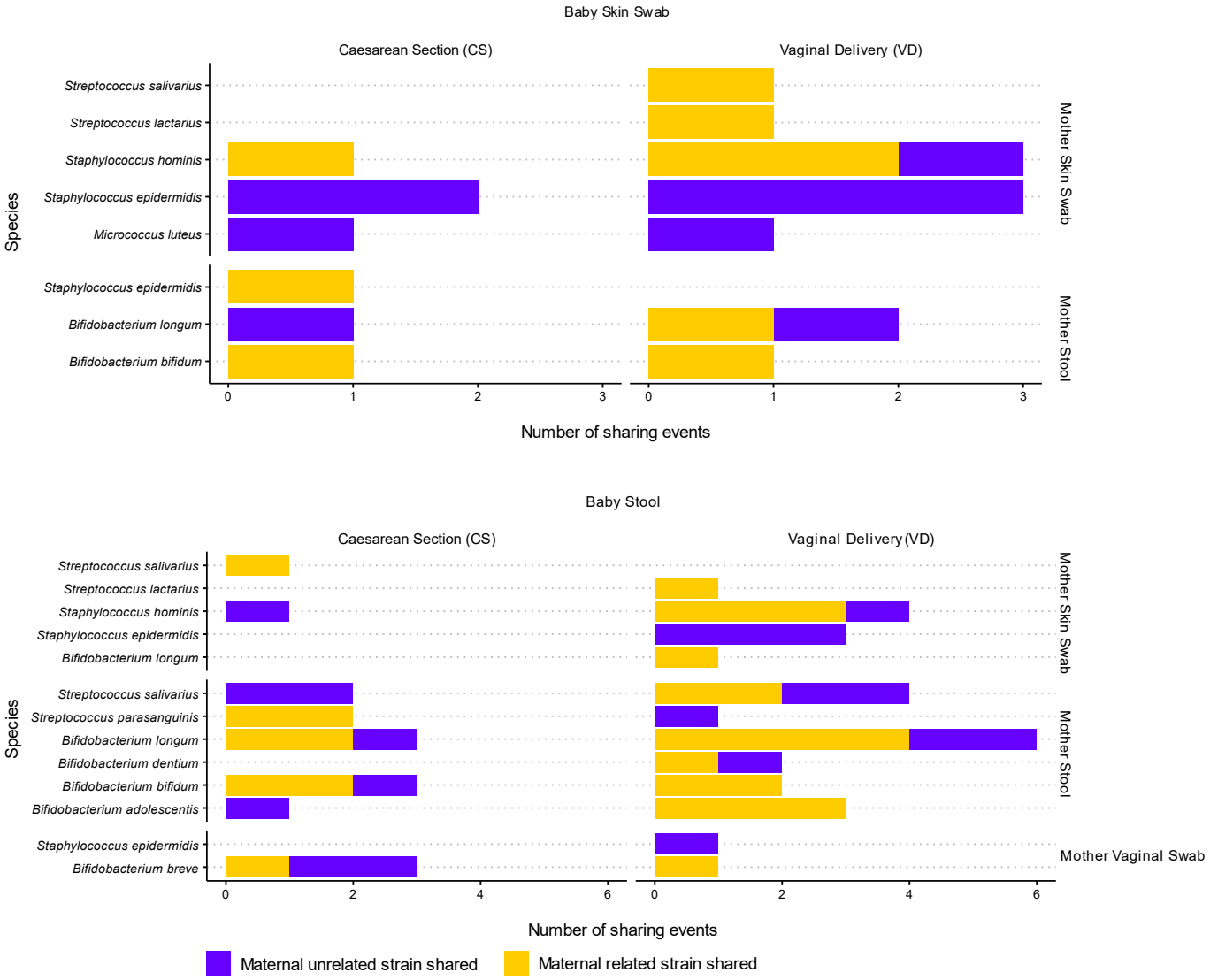
B. Strain Transmissibility Over Time between Mothers and Infants



C. Strain Transmissibility between Mothers and Infants During Pregnancy Term



D. Strain Transmissibility between Mothers and Infants During Delivery Mode



E. Strain Transmissibility between Mothers and Infants During Feeding Mode

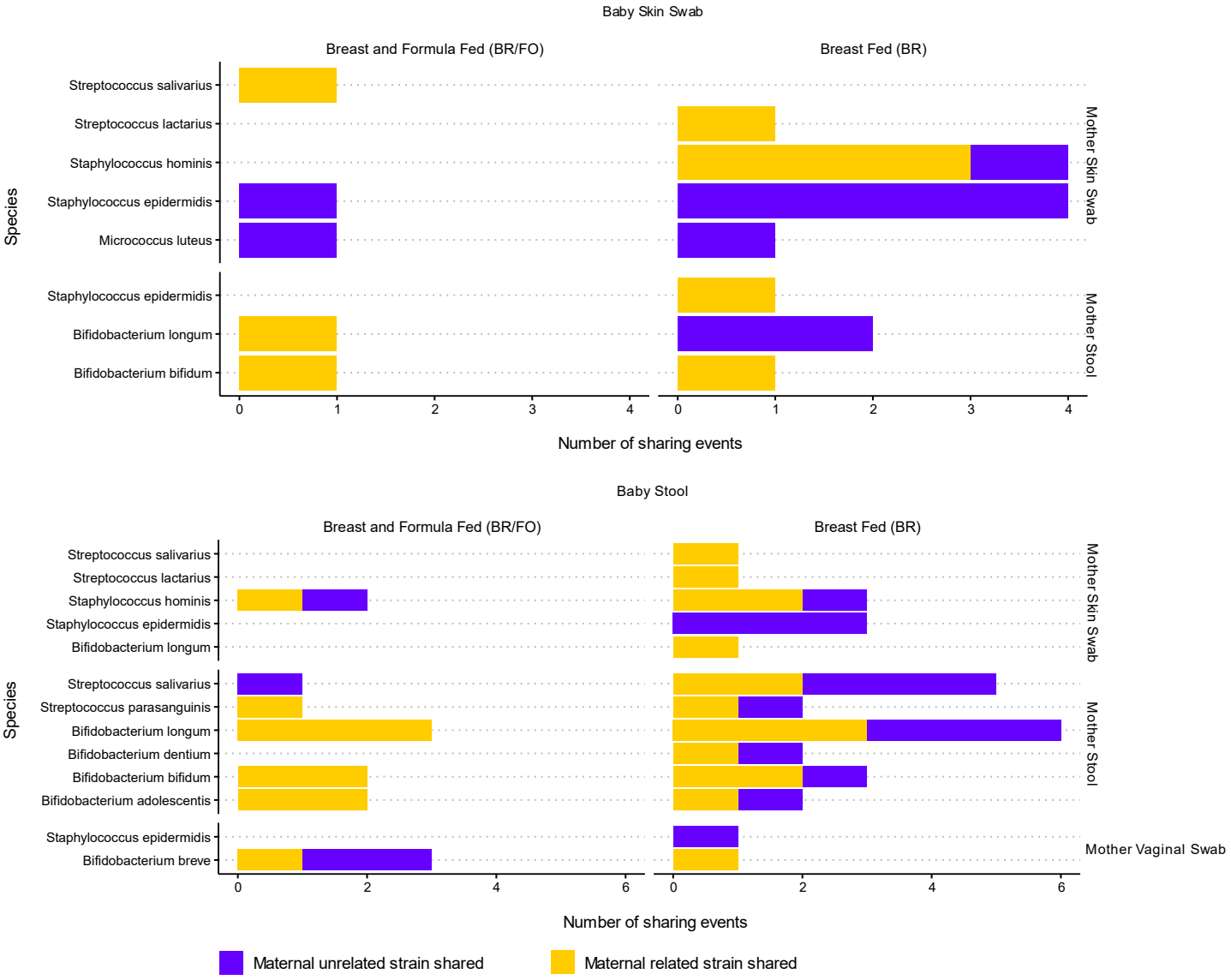


Figure 4.30. Strain Transmissibility of the Shared Common Microbiota Between Maternal and Infant Pairs. Strain level data was generated from maternal and infant samples from the skin, stool and vaginal body sites (infant and maternal skin swabs, stools, and maternal vaginal swabs) using shotgun metagenomics sequencing data, generated on the Illumina platform. The shotgun metagenomics data were annotated by MetaPhlAn4 to identify whether the samples contained enough marker genes to be used in the strain analysis. Samples that didn't meet the sequencing depth threshold to cover enough markers for analysis were excluded, leaving a total of n=566 samples, allowing full maternal and infant paired sample sets to be defined. Strains were generated from the shotgun metagenomics data using a read mapping method with MetaPhlAn4 and StrainPhlan identifying 27 strains and 109 sharing events for the four genera of interest (*Bifidobacterium*, *Micrococcus*, *Streptococcus* and *Staphylococcus*), and strain transmissibility was assessed between infant and maternal body sites. The origin samples were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Factors accounted for included pregnancy term (Late Preterm [LP] and Full Term [FT]), delivery mode (Vaginal Delivery [VD] and Caesarean Section [CS]) and feeding mode (Breast [BR] and Breast/Formula [BR/FO]). The strain transmissibility data displayed in stacked bar plots. The yellow bars correspond to the number of maternal and infant pairs sharing the same strain of each species (maternal related strains), while the purple bars correspond to the number of maternal and infant pairs carrying a different strain of the same species (maternal unrelated strains). Charted figures show the strain transmissibility between the maternal body sites and the infant skin (top chart) and infant stool (bottom chart). Fisher tests were used to test for statistically significant differences ($P < 0.05$) in strain transmissibility in maternal and infant pairs between the maternal body sites and infant skin swab, and maternal body sites and infant stool (also accounting for timepoint, pregnancy term, delivery and feeding mode); there were no significant differences identified. **A.** Overall Strain Transmissibility Between Maternal and Infant Pairs. **B.** Strain Transmissibility Over Time Between Maternal and Infant Pairs. **C.** Strain Transmissibility Between Maternal and Infant Pairs During Pregnancy Term. **D.** Strain Transmissibility Between Maternal and Infant Pairs During Delivery Mode. **E.** Strain Transmissibility Between Maternal and Infant Pairs During Feeding Mode.

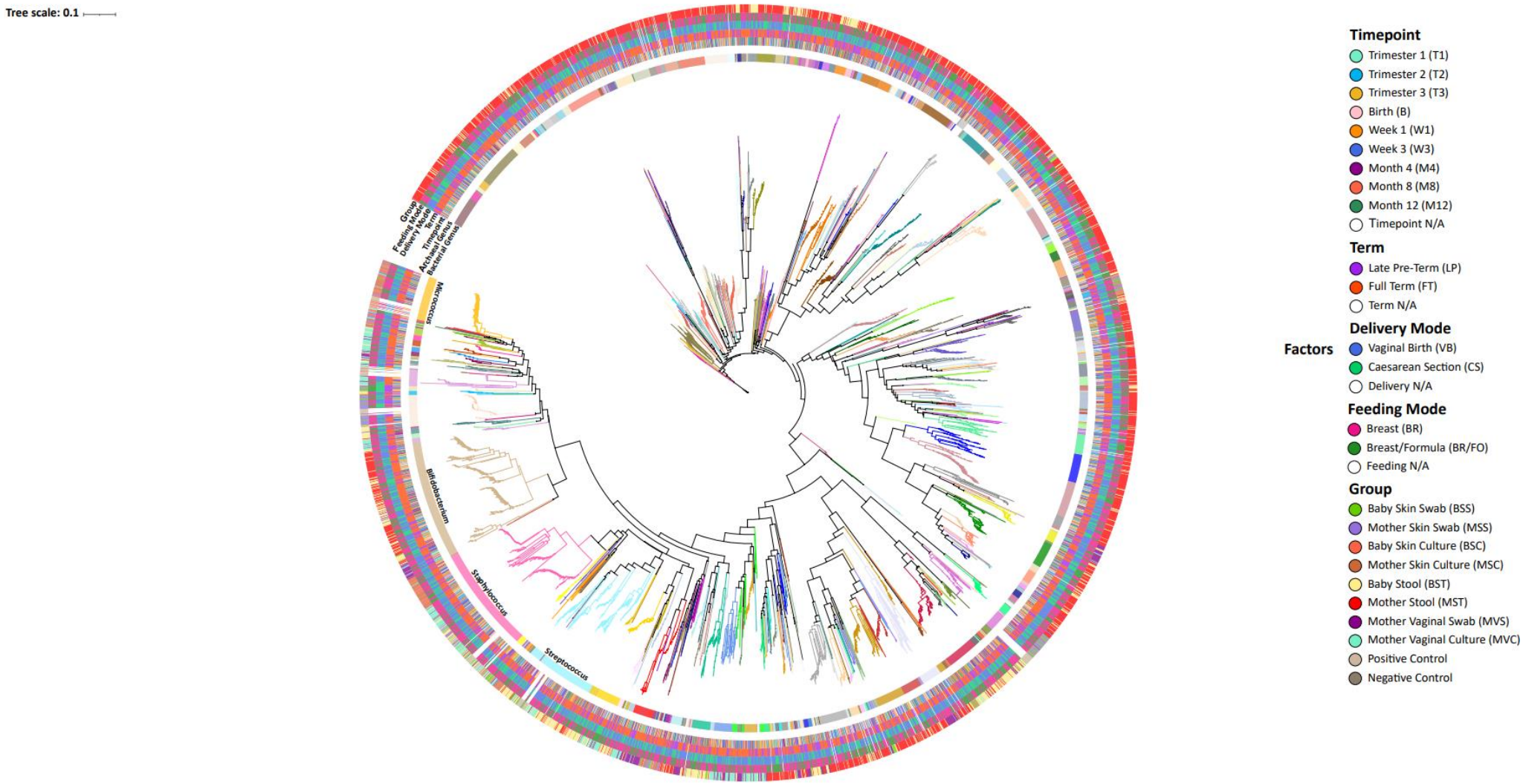
Table 4.10. List of Species of Strains Identified in One Maternal Culture Sample and At Least One Infant Sample

Species	Family ID	Mother Culture ID	Infant Sample ID
Bifidobacterium breve	6	MVC-6-T3	BST-6-M12; BST-6-M4; BST-6-W3; BST-6-M8; BST-6-W1
Staphylococcus aureus	1	MVC-1-W3	BST-1-W3
Staphylococcus epidermidis	1	MSC-1-T2	BST-1-W1
Staphylococcus epidermidis	5	MVC-5-T2	BST-5-B; BSS-5-W1
Staphylococcus epidermidis	5	MVC-5-W1	BST-5-W1
Staphylococcus epidermidis	7	MVC-7-B	BST-7-W1
Staphylococcus epidermidis	10	MVC-10-B	BST-10-B
Staphylococcus epidermidis	10	MVC-10-W1	BST-10-B
Staphylococcus epidermidis	10	MVC-10-W3	BST-10-B
Staphylococcus epidermidis	14	MVC-14-T2	BST-14-W1
Staphylococcus haemolyticus	5	MVC-5-B	BST-5-B
Staphylococcus haemolyticus	5	MVC-5-T3	BST-5-B
Staphylococcus haemolyticus	5	MVC-5-W1	BST-5-B
Staphylococcus haemolyticus	6	MVC-6-W1	BST-6-W1
Staphylococcus hominis	5	MSC-5-T2	BSS-5-W3
Staphylococcus hominis	5	MVC-5-T2	BSS-5-W1; BST-5-W1; BST-5-W3
Staphylococcus hominis	5	MVC-5-T3	BSS-5-W1; BST-5-W1; BST-5-W3
Staphylococcus hominis	5	MVC-5-W1	BSS-5-W3
Staphylococcus hominis	11	MVC-11-B	BSS-11-B; BST-11-W1
Staphylococcus hominis	11	MVC-11-W1	BSS-11-B; BST-11-W1
Streptococcus anginosus	1	MVC-1-W3	BST-1-M4; BST-1-W3
Streptococcus anginosus	5	MVC-5-B	BST-5-W3; BST-5-M4; BST-5-W1; BST-5-B
Streptococcus anginosus	5	MVC-5-T1	BST-5-W3; BST-5-M4; BST-5-W1; BST-5-B
Streptococcus anginosus	5	MVC-5-T2	BST-5-W3; BST-5-M4; BST-5-W1; BST-5-B
Streptococcus anginosus	5	MVC-5-T3	BST-5-W3; BST-5-M4; BST-5-W1; BST-5-B
Streptococcus anginosus	5	MVC-5-W1	BST-5-W3; BST-5-M4; BST-5-W1; BST-5-B
Streptococcus anginosus	5	MVC-5-W3	BST-5-W3; BST-5-M4; BST-5-W1; BST-5-B

4.6.6 Generation of a PEARL Study MAGS Genome Catalogue

A PEARL genome catalogue was generated, comprising 4,408 MAGs across various body sites: 110 infant skin swabs, 98 infant skin cultures, 871 infant stools, 130 maternal skin cultures, 166 maternal skin swabs, 2319 maternal stools, 315 maternal vaginal cultures, 230 maternal vaginal swabs and 169 from controls. All MAGs were of high quality (assessed using the criteria from Kiu et al. [398] with completeness >80% and contamination <10%), ranging in size from 1 to 7MB. Notably, 119 MAGs were unassigned at the species level and 7 MAGs were unassigned at the genus level, suggesting the potential discovery of novel species and genera. A phylogenomic tree provided an overview of the genus-level MAGs, highlighting the complexity of the data, and accounting for multifactorial differences (timepoint, pregnancy term, delivery and feeding mode) in microbial changes over time across eight maternal and infant datasets (covering skin, stool and vaginal body sites) (**Figure 4.31**). The tree also demonstrated alignment between MAGs composition and the taxonomic profiles presented earlier in this chapter (**Figures 4.9; Appendix 14**). There were more MAGs generated for the previously identified core genera (*Bifidobacterium*, *Micrococcus*, *Streptococcus* and *Staphylococcus*) (**Figure 4.18**) than other genera. These core genera were highlighted in the inner circle of the tree (**Figure 4.31**) and zoomed in on in a pruned version of the tree (**Figure 4.32**). Overall, these results support earlier compositional and microbial sharing findings, confirming the sharing of core genera between maternal and infant pairs across multiple body sites over time. Additionally, the trees indicate that the composition remained relatively consistent over time, as no clustering was identified at specific timepoints. An overview of the species-level MAGs and their counts revealed greater species diversity in the maternal skin and stool sites compared to their infants (**Figure 4.33**).

Mother and Infant Genus MAGs Generated



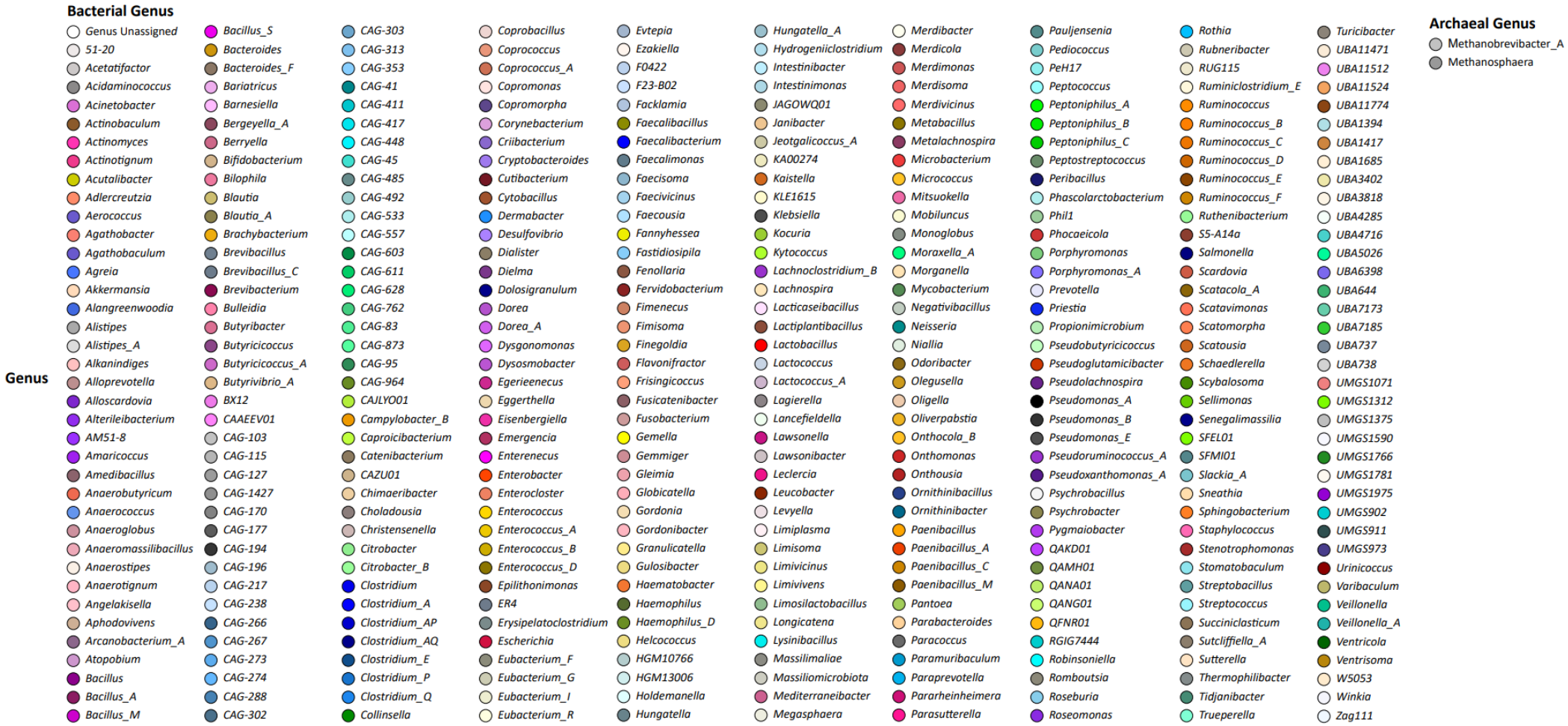

Bacterial Genus
Genus
Genus Unassigned
51-20
Acetatifactor
Acidaminococcus
Acinetobacter
Actinobaculum
Actinomyces
Actinotignum
Acutalibacter
Adlercreutzia
Aerococcus
Agathobacter
Agathobaculum
Agreia
Akkermansia
Alangreenwoodia
Alistipes
Alistipes_A
Alkanindiges
Alloprevotella
Alloscardovia
Alterileibacterium
AM51-8
Amaricoccus
Amedibacillus
Anaerobutyricum
Anaerococcus
Anaeroglobus
Anaeromassilibacillus
Anaerostipes
Anaerotignum
Angelakisella
Aphodovivens
Arcanobacterium_A
Atopobium
Bacillus
Bacillus_A
Bacillus_M
Bacillus_S
Bacteroides
Bacteroides_F
Bariatricus
Barnesiella
Bergeyella_A
Berryella
Bifidobacterium
Bilophila
Blautia
Blautia_A
Brachybacterium
Brevibacillus
Brevibacillus_C
Brevibacterium
Bulleidia
Butyribacter
Butyricicoccus
Butyricicoccus_A
Butyrivibrio_A
BX12
CAAEEV01
CAG-103
CAG-115
CAG-127
CAG-1427
CAG-170
CAG-177
CAG-194
CAG-196
CAG-217
CAG-238
CAG-266
CAG-267
CAG-273
CAG-274
CAG-288
CAG-302
CAG-303
CAG-313
CAG-353
CAG-41
CAG-411
CAG-417
CAG-448
CAG-45
CAG-485
CAG-492
CAG-533
CAG-557
CAG-603
CAG-611
CAG-628
CAG-762
CAG-83
CAG-873
CAG-95
CAG-964
CAJLYO01
Campylobacter_B
Caproicibacterium
Catenibacterium
CAZU01
Chimaeribacter
Choladousia
Christensenella
Citrobacter
Citrobacter_B
Clostridium
Clostridium_A
Clostridium_AP
Clostridium_AQ
Clostridium_E
Clostridium_P
Clostridium_Q
Collinsella
Coprobacillus
Coprococcus
Coprococcus_A
Copromonas
Copromorpha
Corynebacterium
Criibacterium
Cryptobacteroides
Cutibacterium
Cytobacillus
Dermabacter
Desulfovibrio
Dialister
Dielma
Dolosigranulum
Dorea
Dorea_A
Dysgonomonas
Dysosmobacter
Egerieenecus
Eggerthella
Eisenbergiella
Emergencia
Enterenecus
Enterobacter
Enterocloster
Enterococcus
Enterococcus_A
Enterococcus_B
Enterococcus_D
Epilithonimonas
ER4
Erysipelatoclostridium
Escherichia
Eubacterium_F
Eubacterium_G
Eubacterium_I
Eubacterium_R
Evtepia
Ezakiella
F0422
F23-B02
Facklamia
Faecalibacillus
Faecalibacterium
Faecalimonas
Faecisoma
Faecivicinus
Faecousia
Fannyhessea
Fastidiosipila
Fenollaria
Fervidobacterium
Fimenecus
Fimisoma
Finegoldia
Flavonifractor
Frisingicoccus
Fusicatenibacter
Fusobacterium
Gemella
Gemmiger
Gleimia
Globicatella
Gordonia
Gordonibacter
Granulicatella
Gulosibacter
Haematobacter
Haemophilus
Haemophilus_D
Helcococcus
HGM10766
HGM13006
Holdemanella
Hungatella
Hungatella_A
Hydrogeniiclostridium
Intestinibacter
Intestinimonas
JAGOWQ01
Janibacter
Jeotgalicoccus_A
KA00274
Kaistella
KLE1615
Klebsiella
Kocuria
Kytococcus
Lachnoclostridium_B
Lachnospira
Lacticaseibacillus
Lactiplantibacillus
Lactobacillus
Lactococcus
Lactococcus_A
Lagierella
Lancefieldella
Lawsonella
Lawsonibacter
Leclercia
Leucobacter
Levyella
Limiplasma
Limisoma
Limivicinus
Limivivens
Limosilactobacillus
Longicatena
Lysinibacillus
Massilimaliae
Massiliomicrobiota
Mediterraneibacter
Megasphaera
Merdibacter
Merdicola
Merdimonas
Merdisoma
Merdivicinus
Metabacillus
Metalachnospira
Microbacterium
Micrococcus
Mitsuokella
Mobiluncus
Monoglobus
Moraxella_A
Morganella
Mycobacterium
Negativibacillus
Neisseria
Niallia
Odoribacter
Olegusella
Oligella
Oliverpabstia
Onthocola_B
Onthomonas
Onthousia
Ornithinibacillus
Ornithinibacter
Paenibacillus
Paenibacillus_A
Paenibacillus_C
Paenibacillus_M
Pantoea
Parabacteroides
Paracoccus
Paramuribaculum
Paraprevotella
Pararheinheimera
Parasutterella
Pauljensenia
Pediococcus
PeH17
Peptococcus
Peptoniphilus_A
Peptoniphilus_B
Peptoniphilus_C
Peptostreptococcus
Peribacillus
Phascolarctobacterium
Phil1
Phocaeicola
Porphyromonas
Porphyromonas_A
Prevotella
Priestia
Propionimicrobium
Pseudobutyricicoccus
Pseudoglutamicibacter
Pseudolachnospira
Pseudomonas_A
Pseudomonas_B
Pseudomonas_E
Pseudoruminococcus_A
Pseudoxanthomonas_A
Psychrobacillus
Psychrobacter
Pygmaiobacter
QAKD01
QAMH01
QANA01
QANG01
QFNR01
RGIG7444
Robinsoniella
Romboutsia
Roseburia
Roseomonas
Rothia
Rubneribacter
RUG115
Ruminiclostridium_E
Ruminococcus
Ruminococcus_B
Ruminococcus_C
Ruminococcus_D
Ruminococcus_E
Ruminococcus_F
Ruthenibacterium
S5-A14a
Salmonella
Scardovia
Scatcola_A
Scatavimonas
Scatomorpha
Scatousia
Schaedlerella
Scybalosoma
Sellimonas
Senegalimassilia
SFEL01
SFMI01
Slackia_A
Sneathia
Sphingobacterium
Staphylococcus
Stenotrophomonas
Stomatobaculum
Streptobacillus
Streptococcus
Succiniclasticum
Sutcliffiella_A
Sutterella
Thermophilibacter
Tidjanibacter
Trueperella
Turicibacter
UBA11471
UBA11512
UBA11524
UBA11774
UBA1394
UBA1417
UBA1685
UBA3402
UBA3818
UBA4285
UBA4716
UBA5026
UBA6398
UBA644
UBA7173
UBA7185
UBA737
UBA738
UMGS1071
UMGS1312
UMGS1375
UMGS1590
UMGS1766
UMGS1781
UMGS1975
UMGS902
UMGS911
UMGS973
Urinicoccus
Varibaculum
Veillonella
Veillonella_A
Ventricola
Ventrisoma
W5053
Winkia
Zag111
Archaeal Genus
Methanobrevibacter_A
Methanosphaera

Figure 4.31. Genus Level Tree for Maternal and Infant MAGs Generated. This tree provides an overview of genus level MAGs (n=4,408) identified in mothers and infants at multi body sites. MAGs were generated from infant (n=110) and maternal (n=166) skin swabs, infant (n=98) and maternal skin swab cultures (n=130), infant (n=871) and maternal (n=2319) stools, and maternal vaginal swab (n=230) and cultures (n=315) and controls (n=169) using the shotgun metagenomics sequencing data on the Illumina platform. The shotgun metagenomics data was run through a MetaWRAP three binning (MetaBAT, MAXBIN2 and CONCOCT) pipeline to generate high quality MAGs assessed using the criteria from Kiu et al. [398] (completeness >80% and contamination <10%). MAGs were run through Mashtree to generate a tree in ITOL. The origin samples (these MAGs were identified in) were collected at multiple timepoints, one swab was collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab was collected from each mother (n=83), a stool sample was collected from each mother (78) and infant (n=74), and cultures were prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Other factors accounted for included: Pregnancy Term - Late Preterm (LP) and Full Term (FT); Delivery Mode - Vaginal Delivery (VD) and Caesarean Section (CS); Feeding Mode - Breast (BR) and Breast/Formula (BR/FO) milk. The Factor and Genus key allocates a unique colour to each factor and genus identified in the tree clades and rings. The tree scale indicates the branch length distance between the clades.

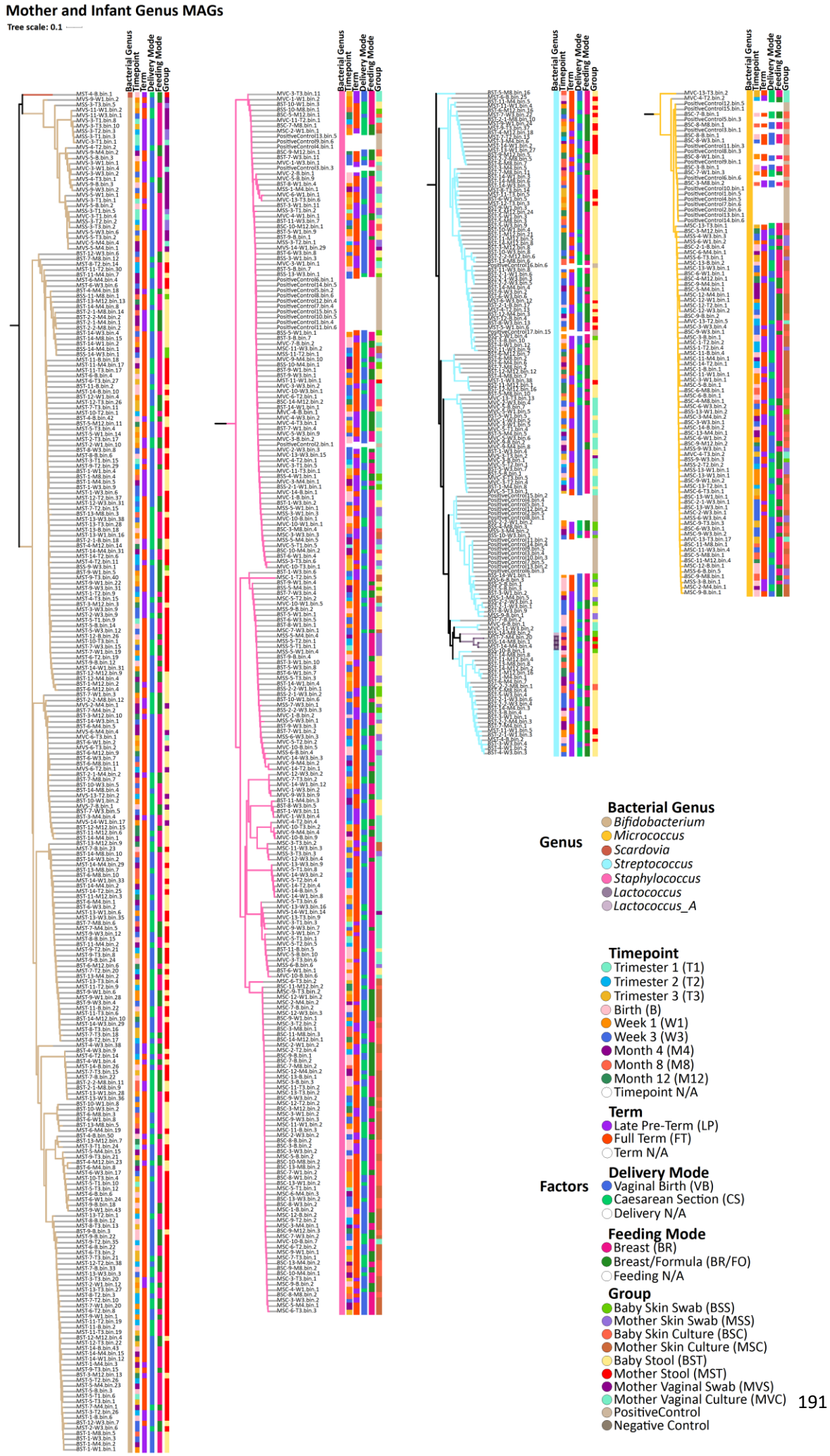
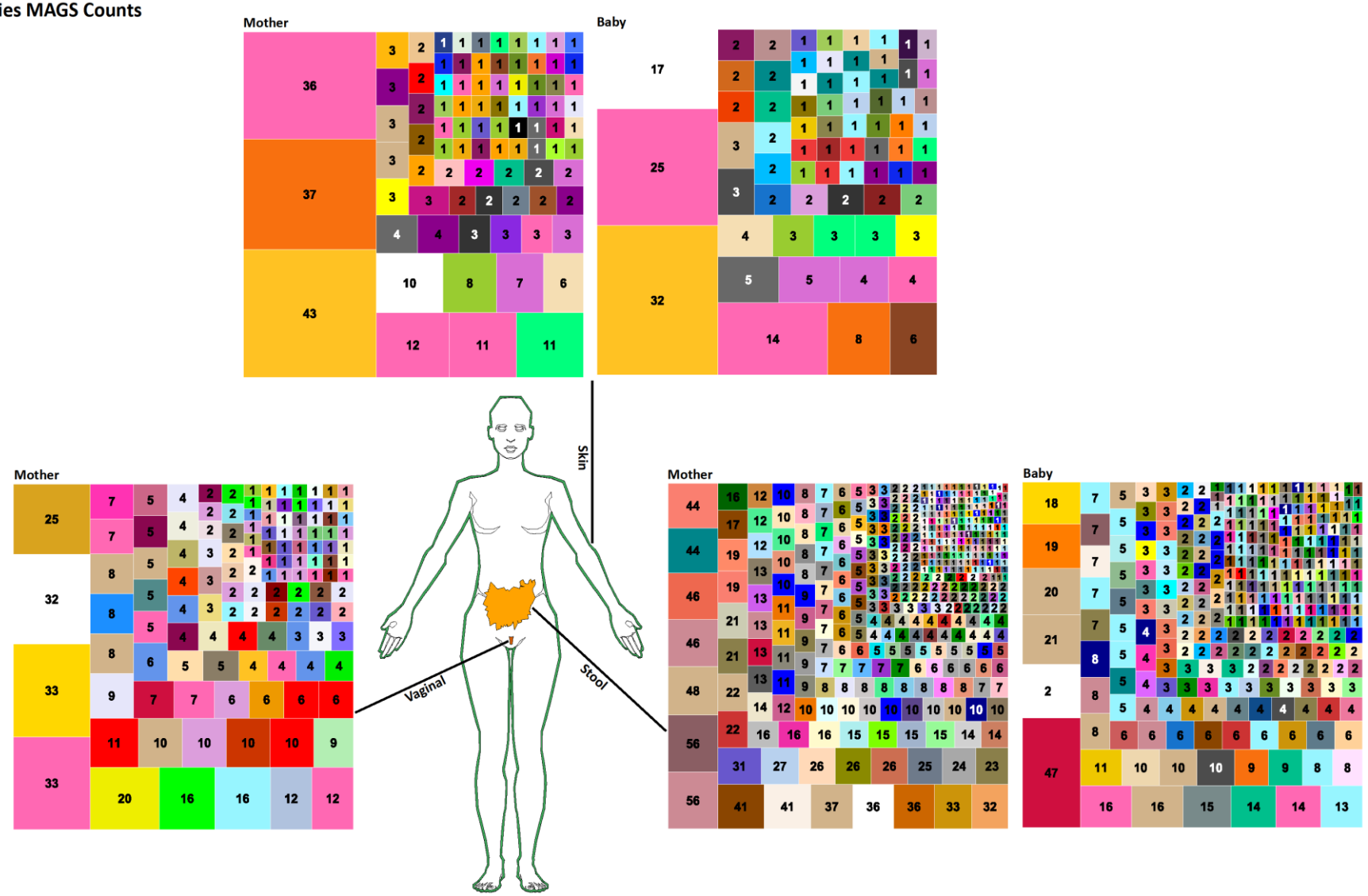

Mother and Infant Genus MAGs
Tree scale: 0.1
Bacterial Genus
Timepoint
Term
Delivery Mode
Feeding Mode
Group
Genus
Bifidobacterium
Micrococcus
Scardovia
Streptococcus
Staphylococcus
Lactococcus
Lactococcus_A
Timepoint
Trimester 1 (T1)
Trimester 2 (T2)
Trimester 3 (T3)
Birth (B)
Week 1 (W1)
Week 3 (W3)
Month 4 (M4)
Month 8 (M8)
Month 12 (M12)
Timepoint N/A
Term
Late Pre-Term (LP)
Full Term (FT)
Term N/A
Factors
Delivery Mode
Vaginal Birth (VB)
Caesarean Section (CS)
Delivery N/A
Feeding Mode
Breast (BR)
Breast/Formula (BR/FO)
Feeding N/A
Group
Baby Skin Swab (BSS)
Mother Skin Swab (MSS)
Baby Skin Culture (BSC)
Mother Skin Culture (MSC)
Baby Stool (BST)
Mother Stool (MST)
Mother Vaginal Swab (MVS)
Mother Vaginal Culture (MVC)
PositiveControl
Negative Control

Figure 4.32. Pruned Tree for Maternal and Infant MAGs Generated for the Core Genera. MAGs were generated using the shotgun metagenomics data, generated from the Illumina platform, from infants and mothers at multiple body sites (skin, stool and vagina). The shotgun metagenomics data was run through a MetaWRAP three binning (MetaBAT, MAXBIN2 and CONCOCT) pipeline to generate high quality MAGs assessed using the criteria from Kiu et al. [398] (completeness >80% and contamination <10%). MAGs were run through Mashtree to generate a tree in ITOL. This tree provides an overview of the genus level MAGs for the four shared genera - *Staphylococcus* (n=228), *Streptococcus* (n=94), *Micrococcus* (n=94) and *Bifidobacterium* (n=200) identified in mothers and infants at multiple body sites - Baby Skin Swabs (BSS), Baby Skin Cultures (BSC), Mother Skin Swabs (MSS), Mother Skin Cultures (MSC), Baby Stools (BST), Mother Stool (MST), Mother Vaginal Swabs (MVS) and Mother Vaginal Cultures (MVC). The origin samples (where these MAGs were identified in) were collected at multiple timepoints, with one swab collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab collected from each mother (n=83), a stool sample collected from each mother (78) and infant (n=74), and cultures prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Other factors accounted for included: Pregnancy Term - Late Preterm (LP) and Full Term (FT); Delivery Mode - Vaginal Delivery (VD) and Caesarean Section (CS); Feeding Mode - Breast (BR) and Breast/Formula (BR/FO) milk. The Factor and Genus key allocates a unique colour to each factor and genus identified in the tree clades and rings. The tree scale indicates the branch length distance between the clades.

Mother and Infant Species MAGS Counts




Species
51−20 sp001917175
Acetatifactor intestinalis
Acetatifactor sp900066565
Acidaminococcus intestini
Acinetobacter fasciculus
Acinetobacter johnsonii
Acinetobacter parvus
Acinetobacter pittii
Acinetobacter schindleri
Acinetobacter sp002296655
Actinobaculum massiliense
Actinomyces urogenitalis
Actinotignum timonense
Acutalibacter ornithocaccae
Adlercreutzia celatus_A
Adlercreutzia equolifaciens
Aerococcus christensenii
Aerococcus sanguinicola
Aerococcus urinae_B
Aerococcus viridans
Aerococcus viridans_B
Agathobacter faecis
Agathobacter rectalis
Agathobacter sp000434275
Agathobaculum butyriciproducens
Agathobaculum intestinigallinarum
Agathobaculum sp003481705
Agathobaculum sp900291975
Agreia pratensis
Akkermansia muciniphila
Akkermansia muciniphila_B
Akkermansia sp001580195
Akkermansia sp004167605
Alangreenwoodia sp905206455
Alistipes avistercoris
Alistipes communis
Alistipes dispar
Alistipes excrementigallinarum
Alistipes finegoldii
Alistipes onderdonkii
Alistipes putredinis
Alistipes shahii
Alistipes_A ihumii
Alistipes_A indistinctus
Alkanindiges illinoisensis_A
Alloprevotella sp015259235
Alloprevotella sp905371275
Alloscardovia omnicolens
AM51−8 sp900761925
Amedibacillus dolichus
Anaerobutyricum hallii
Anaerobutyricum soehngenii
Anaerococcus hydrogenalis
Anaerococcus jeddahensis
Anaerococcus lactolyticus
Anaerococcus mediterraneensis
Anaerococcus obesiensis
Anaerococcus octavius
Anaerococcus tetradius
Anaerococcus vaginalis_B
Anaerococcus vaginimassiliensis
Anaeroglobus micronuciformis
Anaeromassilibacillus sp001305115
Anaeromassilibacillus stercoravium
Anaerostipes caccae
Anaerostipes hadrus
Anaerostipes hadrus_A
Anaerotignum faecicola
Anaerotignum lactatifermentans
Angelakisella sp900547385
Arcanobacterium_A urinimassiliense
Atopobium deltae
Bacillus altitudinis
Bacillus licheniformis
Bacillus subtilis
Bacillus velezensis
Bacillus_A cereus
Bacillus_A cereus_U
Bacillus_A mobilis
Bacillus_A wiedmannii
Bacillus_M haloduranss
Bacillus_S marmarensis
Bacteroides caccae
Bacteroides cellulosilyticus
Bacteroides clarus
Bacteroides eggerthii
Bacteroides faecis
Bacteroides fragilis
Bacteroides ovatus
Bacteroides stercoris
Bacteroides thetaiotaomicron
Bacteroides uniformis
Bacteroides xylanisolvens
Bacteroides_F pectinophilus
Bariatricus comes
Barnesiella intestinihominis
Berryella sp001552935
Berryella sp900604965
Bifidobacterium adolescentis
Bifidobacterium animalis
Bifidobacterium bifidum
Bifidobacterium breve
Bifidobacterium catenulatum
Bifidobacterium dentium
Bifidobacterium infantis
Bifidobacterium kashiwanohense
Bifidobacterium longum
Bifidobacterium piotii
Bifidobacterium pseudocatenulatum
Bifidobacterium pullorum_B
Bifidobacterium swidsinskii
Bifidobacterium vaginale
Bilophila wadsworthia
Blautia hansenii
Blautia ornithocaccae
Blautia pseudococcoides
Blautia sp001304935
Blautia sp900541955
Blautia sp900555025
Blautia sp900753905
Blautia stercoris
Blautia_A caecimuris
Blautia_A faecis
Blautia_A hydrogenotrophica
Blautia_A massiliensis
Blautia_A obeum
Blautia_A schinkii
Blautia_A sp000285855
Blautia_A sp000436615
Blautia_A sp000066205
Blautia_A sp000066335
Blautia_A sp000066355
Blautia_A sp000066505
Blautia_A sp000551715
Blautia_A wexlerae
Brachybacterium muris
Brachybacterium paraconglomeratum
Brachybacterium sp015209585
Brevibacillus borstelensis
Brevibacillus_C massiliensis
Brevibacterium casei
Brevibacterium ravenspurgense
Brevibacterium sp001815905
Bulleidia sp900539965
Butyribacter intestini
Butyribacter sp003149875
Butyribacter sp003524945
Butyribacter sp003529475
Butyricicoccus pullicaecorum
Butyricicoccus_A sp002395695
Butyrivibrio_A crossotus
CAJLYO01 sp905210885
Campylobacter_B ureolyticus
CAAEEV01 sp900754955
CAG−103 sp000432375
CAG−103 sp900317855
CAG−103 sp900543625
CAG−115 sp003531585
CAG−127 sp900319515
CAG−1427 sp000434775
CAG−1427 sp000435475
CAG−1427 sp000435675
CAG−1427 sp000436075
CAG−1427 sp900542265
CAG−1427 sp900544455
CAG−1427 sp900554685
CAG−170 sp000432135
CAG−170 sp002404795
CAG−170 sp003516765
CAG−170 sp900545925
CAG−177 sp003514385
CAG−177 sp003538135
CAG−194 sp000432915
CAG−196 sp002102975
CAG−217 sp000436335
CAG−238 sp900551415
CAG−266 sp000436095
CAG−267 sp001917135
CAG−273 sp000435755
CAG−273 sp003534295
CAG−274 sp000432155
CAG−274 sp900545305
CAG−288 sp000437395
CAG−302 sp000431795
CAG−302 sp001916775
CAG−303 sp000437755
CAG−313 sp000433035
CAG−353 sp900066885
CAG−41 sp900066215
CAG−411 sp000437275
CAG−417 sp000432835
CAG−448 sp003150135
CAG−45 sp000438375
CAG−485 sp900761855
CAG−492 sp000434335
CAG−533 sp000434495
CAG−533 sp900553855
CAG−557 sp000435275
CAG−603 sp900066105
CAG−611 sp000434175
CAG−628 sp000438415
CAG−628 sp003524085
CAG−762 sp900551315
CAG−83 sp000431575
CAG−83 sp000435555
CAG−83 sp000435975
CAG−83 sp001916855
CAG−83 sp003487665
CAG−83 sp900545585
CAG−83 sp900547745
CAG−83 sp900548615
CAG−83 sp900549395
CAG−873 sp001701165
CAG−95 sp000438155
CAG−964 sp000435335
Cutibacterium acnes
Cutibacterium avidum
Dermabacter hominis
Desulfovibrio legallii
Desulfovibrio sp900556755
Dialister hominis
Dialister invisus
Dialister sp001553355
Dialister sp002320515
Dialister sp900541485
Dialister succinatiphilus
Dielma fastidiosa
Dolosigranulum pigrum
Dorea formicigenerans
Dorea_A longicatena
Dorea_A longicatena_B
Dysgonomonas mossii
Dysosmobacter avistercoris
Dysosmobacter sp900544615
Dysosmobacter welbionis
Egerieenecus merdigallinarum
Eggerthella lenta
Eisenbergiella intestinigallinarum
Eisenbergiella porci
Eisenbergiella sp900066775
Eisenbergiella sp900539715
Eisenbergiella tayi
Enterenecus faecium
Enterenecus sp900549885
Enterobacter hormaechei_A
Enterocloster aldenensis
Enterocloster clostridioformis
Enterocloster sp000431375
Enterocloster sp001517625
Enterocloster sp900538485
Enterocloster sp900541315
Enterococcus faecalis
Enterococcus_A avium
Enterococcus_A raffinosus
Enterococcus_B durans
Enterococcus_B faecium
Enterococcus_B lactis
Coprococcus eutactus
Coprococcus eutactus_A
Coprococcus sp000154245
Coprococcus sp000066115
Coprococcus_A catus
Copromonas sp000435795
Copromorpha excrementipullorum
Copromorpha sp900549125
Corynebacterium accolens
Corynebacterium amycolatum
Corynebacterium amycolatum_A
Corynebacterium aurimucosum_C
Corynebacterium coyleae
Corynebacterium minutissimum_A
Corynebacterium sp001807205
Corynebacterium sp001875725
Corynebacterium striatum
Corynebacterium tuscaniense
Criibacterium bergeronii
Cryptobacteroides sp000432515
Cryptobacteroides sp000434935
Cutibacterium acnes
Cutibacterium avidum
Dermabacter hominis
Desulfovibrio legallii
Desulfovibrio sp900556755
Dialister hominis
Dialister invisus
Dialister sp001553355
Dialister sp002320515
Dialister sp900541485
Dialister succinatiphilus
Dielma fastidiosa
Dolosigranulum pigrum
Dorea formicigenerans
Dorea_A longicatena
Dorea_A longicatena_B
Dysgonomonas mossii
Dysosmobacter avistercoris
Dysosmobacter sp900544615
Dysosmobacter welbionis
Egerieenecus merdigallinarum
Eggerthella lenta
Eisenbergiella intestinigallinarum
Eisenbergiella porci
Eisenbergiella sp900066775
Eisenbergiella sp900539715
Eisenbergiella tayi
Enterenecus faecium
Enterenecus sp900549885
Enterobacter hormaechei_A
Enterocloster aldenensis
Enterocloster clostridioformis
Enterocloster sp000431375
Enterocloster sp001517625
Enterocloster sp900538485
Enterocloster sp900541315
Enterococcus faecalis
Enterococcus_A avium
Enterococcus_A raffinosus
Enterococcus_B durans
Enterococcus_B faecium
Enterococcus_B lactis
Enterococcus_B pernyi
Enterococcus_D casseliflavus
Enterococcus_D gallinarum
Enterococcus_D sp002850555
Epilithonimonas hominis
Epilithonimonas vandammei
ER4 sp000765235
ER4 sp900317525
Erysipelatoclostridium ramosum
Erysipelatoclostridium saccharogumia
Erysipelatoclostridium spiroforme
Escherichia coli
Escherichia marmotae
Eubacterium_F sp000433735
Eubacterium_F sp003491505
Eubacterium_F sp900539115
Eubacterium_G sp000432355
Eubacterium_G sp000434315
Eubacterium_G sp900435815
Eubacterium_G sp900548465
Eubacterium_G ventriosum
Eubacterium_I ramulus
Eubacterium_I sp900546495
Eubacterium_R faecale
Eubacterium_R sp000433975
Eubacterium_R sp000434995
Eubacterium_R sp000436835
Eubacterium_R sp003526845
Eubacterium_R sp009774735
Eubacterium_R sp014465955
Eubacterium_R sp900539775
Evtepia excrementipullorum
Evtepia faecavium
Evtepia gabavorous
Evtepia sp004556345
Ezakiella coagulans
F0422 sp001553345
F0422 sp003992015
F0422 sp900544365
F23−B02 sp001916715
F23−B02 sp002472405
Facklamia hominis
Faecalibacillus faecis
Faecalibacillus intestinalis
Faecalibacterium prausnitzii
Faecalibacterium prausnitzii_A
Faecalibacterium prausnitzii_C
Faecalibacterium prausnitzii_D
Faecalibacterium prausnitzii_E
Faecalibacterium prausnitzii_G
Faecalibacterium prausnitzii_J
Faecalibacterium sp900539945
Faecalimonas phoceensis
Faecalimonas umbilicata
Faecisoma sp000432255
Faecivicinus sp900553265
Faecousia gallistercoris
Faecousia sp000434635
Faecousia sp000435995
Faecousia sp003525905
Faecousia sp900546075
Fannyhessea vaginae
Fastidiosipila sanguinis
Fenollaria massiliensis
Fervidobacterium pennivorans
Fimenecus sp000432435
Fimisoma sp000435715
Finegoldia magna_F
Finegoldia magna_H
Flavonifractor plautii
Fusicatenibacter intestinigallinarum
Fusicatenibacter saccharivorans
Fusicatenibacter sp900543115
Fusobacterium animalis
Gemella haemolysans_B
Gemmiger formicilis
Gemmiger qucibialis
Gemmiger sp900539695
Gemmiger sp900540775
Gleimia europaea_A
Gleimia europaea_B
Globicatella sanguinis
Gordonia bronchialis
Gordonia otitidis
Gordonibacter pamelaeae
Gulosibacter sp904848455
Haematobacter massiliensis
Haemophilus seminalis
Haemophilus_D parainfluenzae_L
Haemophilus_D parainfluenzae_M
Haemophilus_D sp001815355
Haemophilus_D sp900756155
Helcococcus massiliensis
HGM10766 sp900757295
HGM13006 sp900756575
HGM13006 sp900757695
Holdemanella biformis
Holdemanella porci
Holdemanella sp003458715
Hungatella effluvii
Hungatella_A hathewayi_A
Hydrogeniiclostridium mannosilyticum
Intestinibacter bartlettii
Intestinimonas massiliensis
Janibacter indicus
Jeotgalicoccus_A schoeneichii
KA00274 sp001552885
KA00274 sp902373515
Kaistella haifensis
KLE1615 sp900066985
Klebsiella grimontii
Klebsiella michiganensis
Klebsiella pneumoniae
Kocuria carniphila
Kocuria marina
Kocuria palustris
Kocuria palustris_A
Kocuria rhizophila
Kytococcus sedentarius
Lachnoclostridium_B phocaeense
Lachnoclostridium_B sp000765215
Lachnoclostridium_B sp900066555
Lachnospira eligens_A
Lachnospira rogosae_A
Lachnospira sp000436475
Lachnospira sp003451515
Lachnospira sp900316325
Lacticaseibacillus paracasei
Lacticaseibacillus rhamnosus
Lactiplantibacillus plantarum
Lactobacillus acidophilus
Lactobacillus crispatus
Lactobacillus gasseri
Lactobacillus iners
Lactobacillus jensenii
Lactobacillus johnsonii
Lactobacillus paragasseri
Lactococcus cremoris
Lactococcus lactis
Lactococcus_A raffinolactis
Lancefieldella parvula
Lancefieldella sp000564995
Lawsonella clevelandensis_A
Lawsonibacter asaccharolyticus
Lawsonibacter sp000177015
Lawsonibacter sp900066825
Leclercia adecarboxylata
Leucobacter sp001373475
Levyella massiliensis
Limiplasma merdipullorum
Limiplasma sp900548145
Limisoma sp000437795
Limivicinus sp003150355
Limivicinus sp900547315
Limivicinus sp900549645
Limivivens sp900543575
Limivivens sp905214955
Limosilactobacillus fermentum
Longicatena caecimuris
Massiliomicrobiota merdigallinarum
Mediterraneibacter caccavium
Mediterraneibacter excrementipullorum
Mediterraneibacter faecigallinarum
Mediterraneibacter faecipullorum
Mediterraneibacter faecis
Mediterraneibacter lactaris
Mediterraneibacter merdigallinarum
Mediterraneibacter torques
Megasphaera lornae
Megasphaera sp900417505
Merdibacter merdigallinarum
Merdicola sp001915925
Merdicola sp018371855
Merdimonas faecis
Merdisoma faecalis
Merdivicinus sp900555435
Metabacillus sp002577655
Metalachnospira sp900548595
Methanobrevibacter_A smithii
Methanobrevibacter_A smithii_A
Methanosphaera stadtmanae
Microbacterium aurum
Microbacterium lacticum
Microbacterium oxydans
Micrococcus luteus
Micrococcus lylae
Mitsuokella jalaludinii
Mobilunculs curtisii
Monoglobus pectinilyticus
Moraxella_A cinereus
Moraxella_A osloensis
Moraxella_A sp002478835
Morganella morganii
Negativibacillus massiliensis
Neisseria cinerea
Neisseria flavescens_A
Neisseria sp000186165
Neisseria subflava_C
Niallia sp900199695
Odoribacter splanchnicus
Olegusella massiliensis
Oligella urethralis
Oliverpabstia faecicola
Oliverpabstia intestinalis
Onthocola_B sp900544625
Onthomonas sp900545815
Onthomonas sp900752305
Onthousia sp000438295
Ornithinibacillus bavariensis
Paenibacillus cookii
Paenibacillus polymyxa_B
Paenibacillus sp000499205
Paenibacillus sp018310695
Paenibacillus sp018403285
Paenibacillus_A barengoltzii
Paenibacillus_M ginsengihumi
Pantoea agglomerans
Pantoea brenneri
Pantoea septica
Parabacteroides distasonis
Parabacteroides johnsonii
Parabacteroides merdae
Paracoccus aerius
Paracoccus haematequi
Paracoccus marcusii
Paracoccus yeei
Paramuribaculum sp000431155
Paraprevotella xylaniphila
Pararheinheimera tangshanensis
Parasutterella excrementihominis
Parasutterella gallistercoris
Pauljensenia pyogenes_A
Pauljensenia radingae_A
Pauljensenia sp000308055
Pauljensenia sp000411415
Pauljensenia sp000466265
Pauljensenia sp001838165
Pauljensenia sp018375675
Pauljensenia sp018382595
Pauljensenia turicensis
Pediococcus acidilactici
PeH17 sp000435055
Peptococcus niger
Peptoniphilus_A grossensis
Peptoniphilus_A harei_A
Peptoniphilus_A lacrimalis
Peptoniphilus_A senegalensis
Peptoniphilus_A vaginalis
Peptoniphilus_B duerdenii
Peptoniphilus_B sp000478985
Peptoniphilus_C coxii
Peptostreptococcus anaerobius
Phascolarctobacterium faecium
Phil1 sp001940855
Phocaeicola coprocola
Phocaeicola coprophilus
Phocaeicola dorei
Phocaeicola gallinaceus
Phocaeicola vulgatus
Porphyromonas pasteri
Porphyromonas sp003640335
Porphyromonas sp900548415
Porphyromonas uenonis_A
Porphyromonas_A somerae
Porphyromonas_A sp001808555
Prevotella bergensis
Prevotella bivia
Prevotella buccae
Prevotella buccalis
Prevotella colorans
Prevotella copri
Prevotella disiens
Prevotella melaninogenica
Prevotella mizrahii
Prevotella sp000479005
Prevotella sp002265625
Prevotella sp003447235
Prevotella timonensis
Prevotella timonensis_A
Priestia aryabhattai
Priestia flexa
Priestia megaterium
Propionimicrobium lymphophilum
Pseudobutyricicoccus lothianensis
Pseudoglutamicibacter albus_A
Pseudoglutamicibacter cumminsii
Pseudomonas_A stutzeri
Pseudomonas_B oryzihabitans
Pseudomonas_B oryzihabitans_B
Pseudomonas_E putida_E
Pseudomonas_E putida_H
Pseudoruminococcus_A massiliensis
Pseudoxanthomonas_A mexicana
Psychrobacillus sp008807035
Psychrobacter sanguinis
Pygmaiobacter massiliensis
QAKD01 sp003343965
QAMH01 sp900544245
QFNR01 sp003248485
Robinsoniella peoriensis
Romboutsia timonensis
Roseburia hominis
Roseburia intestinalis
Roseburia inulinivorans
Roseburia sp900753715
Roseomonas mucosa
Rothia dentocariosa
Rothia mucilaginosa
Rothia sp001808955
Rubneribacter badeniensis
RUG115 sp900066395
Ruminiclostridium_E siraeum
Ruminiclostridium_E sp900539195
Ruminococcus sp900540005
Ruminococcus_B gnavus
Ruminococcus_C callidus
Ruminococcus_C sp000433635
Ruminococcus_C sp000437255
Ruminococcus_C sp000980705
Ruminococcus_C sp900545285
Ruminococcus_D bicirculans
Ruminococcus_E bromii_B
Ruminococcus_E sp003526955
Ruminococcus_E sp900314705
Ruminococcus_F champanellensis
Ruthenibacterium lactatiformans
Ruthenibacterium merdipullorum
Ruthenibacterium sp018380115
S5−A14a sp000758905
Salmonella enterica
Scardovia wiggsiae
Scatacola_A faecigallinarum
Scatavimonas sp900540275
Scatomorpha sp900545405
Scatomorpha sp900546485
Scatousia sp900554385
Schaedlerella glycyrrhizinilytica_A
Scybalosoma faecavium
Sellimonas avistercoris
Sellimonas intestinalis
Senegalimassilia anaerobia
SFEL01 sp004557245
SFMI01 sp004556155
Slackia_A isoflavoniconvertens
Snaethia sanguinegens
Snaethia vaginalis
Species Unassigned
Sphingobacterium lactis
Staphylococcus aureus
Staphylococcus capitis
Staphylococcus caprae
Staphylococcus epidermidis
Staphylococcus equorum
Staphylococcus haemolyticus
Staphylococcus hominis
Staphylococcus lugdunensis
Staphylococcus pasteuri
Staphylococcus simulans
Staphylococcus warneri
Stenotrophomonas maltophilia
Stenotrophomonas maltophilia_AK
Stomatobaculum sp002892395
Streptococcus anginosus
Streptococcus dysgalactiae
Streptococcus lactarius
Streptococcus lutetiensis
Streptococcus mitis
Streptococcus mitis_A
Streptococcus mitis_N
Streptococcus mitis_P
Streptococcus mitis_Q
Streptococcus mutans
Streptococcus parasanguinis_E
Streptococcus parasanguinis_F
Streptococcus pasteurianus
Streptococcus salivarius
Streptococcus sp000187445
Streptococcus sp001556435
Streptococcus sp001813295
Streptococcus sp900766505
Streptococcus sp901875575
Streptococcus sp902373455
Streptococcus thermophilus
Streptococcus vestibularis
Streptococcus xiaochunlingii
Succiniclasticum sp900314855
Sutterella sp900762445
Sutterella wadsworthensis_A
Thermophilibacter provencensis
Tidjanibacter inops_A
Trueperella bernardiae
Turicibacter sanguinis
UBA11471 sp000434215
UBA11512 sp003522145
UBA11524 sp000437595
UBA11774 sp003507655
UBA1394 sp900066845
UBA1394 sp900538575
UBA1417 sp003531055
UBA1417 sp900549945
UBA1685 sp002320595
UBA3402 sp003478355
UBA3818 sp900557155
UBA3818 sp904397845
UBA4285 sp900542465
UBA4285 sp900545225
UBA4716 sp905215805
UBA5026 sp900552335
UBA6398 sp900550635
UBA644 sp002299265
UBA644 sp900547165
UBA7173 sp900546835
UBA7185 sp900756555
UBA737 sp900549755
UBA738 sp003522945
UMGS1071 sp900542375
UMGS1312 sp900550625
UMGS1375 sp900066615
UMGS1590 sp900553245
UMGS1766 sp900554855
UMGS1781 sp900553695
UMGS1975 sp900546685
UMGS902 sp003343845
UMGS911 sp900545935
UMGS973 sp900547295
Urinicoccus massiliensis
Varibaculum cambriense
Varibaculum cambriense_A
Varibaculum massiliense
Veillonella atypica
Veillonella nakazawae
Veillonella parvula
Veillonella rogosae
Veillonella sp900550455
Veillonella_A seminalis
Ventricola sp900542395
Ventricola sp900542445
Ventricola sp900548125
Ventrisoma faecale
Winkia neuii
Winkia sp002849225
Zag111 sp900555325

Figure 4.33. Treemap of the Maternal and Infant Species MAGs Counts. This treemap provides an overview of the species counts for the MAGs generated (n=4,408) from mothers and infants at multi body sites (skin, stool and vagina). MAGs were generated from infant (n=110) and maternal (n=166) skin swabs, infant (n=98) and maternal skin swab cultures (n=130), infant (n=871) and maternal (n=2319) stools, and maternal vaginal swab (n=230) and cultures (n=315) and controls (n=169) using shotgun metagenomics sequencing data, generated on the Illumina platform. The shotgun metagenomics data was run through a MetaWRAP three binning (MetaBAT, MAXBIN2 and CONCOCT) pipeline to generate high quality MAGs assessed using the criteria from Kiu et al. [398] (completeness >80% and contamination <10%). MAGs species counts were displayed in a treemap. The origin samples (these MAGs were identified in) were collected at multiple timepoints, one swab was collected from one forearm from each mother (n=85) and infant (n=89), one vaginal swab was collected from each mother (n=83), a stool sample was collected from each mother (78) and infant (n=74), and cultures were prepared for maternal (56) and infant (n=48) skin swabs and maternal vaginal swab cultures (n=69); total n=582. Timepoints investigated included: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Each species were allocated a unique colour as identified in the key.

4.7 Discussion

The Composition and Baseline of a Healthy Infant Skin Microbiome Over Time

The dominant microbes in varying abundances identified in infant and maternal skin (*Streptococcus*, *Staphylococcus* and *Acinetobacter*), and associations of certain infant microbes (*Lactobacillus* and *Bifidobacterium*) with other maternal body sites (vaginal and stool) and factors (feeding and delivery mode) correlated with the dominant adult and infant phyla identified in other studies [1, 7, 14, 15, 446-448]. The higher abundances of *Streptococcus spp.*, such as *S. mitis* and *S. salivarius* identified in infant skin compared to their mothers, align to other research; indicating that increased *Streptococcus* abundance contributes to the acidification of infant skin in the months following birth, contributing to the formation of the acid mantle [449]. This protective layer, which begins forming during the last trimester of pregnancy, creates an environment that supports the growth of lactic-acid producing *Streptococcus spp.* [449, 450]. This may explain the steady increase in *Streptococcus* abundance on infant skin from B to M8, followed by a decrease at M12 as the acid mantle becomes fully developed. In parallel, a higher abundance of *Staphylococcus* on infant skin from B to W3, may be attributed to the vernix caseosa, known as a "natural moisturiser", that makes skin greasy and moist, providing an ideal environment for *Staphylococcus*, as seen on adult skin [37, 238-240]. After W3, the abundance of *Staphylococcus* decreases significantly, but begins to rise again at M12. This resurgence may coincide with development of the acid mantle by *Streptococcus*, leading to a more stable and acidic skin barrier, that supports the growth of *Staphylococcus*; ultimately contributing to a shift towards a more adult-like skin microbiome, where *Staphylococcus* plays a key role in maintaining skin barrier integrity [44]. This trend aligns with the consistent presence of *Staphylococcus* on maternal skin across all timepoints. *Acinetobacter* showed higher abundance on maternal skin compared to infants, with its presence on infant skin increasing by M12, consistent with other studies showing a gradual increase in *Acinetobacter* abundance from B to M12 [14]. Additionally, oral related taxa, such as *S. oralis*, *G. haemolysans*, *Neisseria* and *Rothia*, were identified on infant skin, likely transferred through breast feeding [449, 451]. This is supported by previous studies showing that breast milk influences the development of the infant oral microbiome [451], suggesting a plausible transfer between the infant skin, breast milk and oral microbiome.

Breast milk derived species, such as *B. breve* [397, 452], were identified exclusively in infant stools and not in maternal stools, suggesting potential vertical transfer from breast milk, as supported by previous studies [397, 452]. Other *Bifidobacterium* species present in mother and infant included *B. longum* and *B. bifidum*. These species have also been detected in both breast milk and faeces of mothers and infants [397]. Research has shown *Bifidobacterium* to dominate the infant gut microbiome, which aligns with my findings [446]. This may be due to the infant's milk based diet, which promotes a saccharolytic metabolism of human milk oligosaccharides characterised by smaller sugar like glucose - primarily driven by *Bifidobacterium* spp., such as *B. bifidum*, *B. breve*, *B. longum* [446]. In contrast, the adult gut microbiome supports a metabolic profile with lower sugar levels, typically sustained by *F. prausnitzii* [446], which may explain the decrease of *Bifidobacterium* in infant stools at M12 identified; coinciding when infants stop consuming any breast milk.

From a multi-body site microbial sharing perspective, the infant skin microbiome over time appears to be most influenced by the maternal skin, with limited contributions from other maternal body sites. However, some sharing of vaginal associated bacteria such as *L. crispatus*, and stool and vaginal associated bacteria such as *B. breve*, *B. longum* and *B. bifidum* from maternal to infant skin was observed, as shown in other studies [129]. Mothers with *Bifidobacterium* in both vaginal and stool profiles gave birth via different delivery modes, suggesting that these microbes may be transferred vertically through breast milk. Additionally, there was potential microbial sharing within individual infants between their skin and stool, specifically for *Bifidobacterium spp.*, indicating transient and vertical transfer from breast milk to the skin and gut. This transfer between body sites has been supported by previous literature [397, 451, 452].

Across all infants, the common bacteria present on the skin from B to M12 included *Acinetobacter* (*A. baumannii*, *A. johnsonii* and *A. iwoffii*), *Bifidobacterium* (*B. longum*), *Corynebacterium*, *Cutibacterium* (*C. acnes*), *Epilithonimonas* (*E. vandammei*), *Fervidobacterium* (*F. pennivorans*), *Gemella* (*G. haemolysans*), *Gordonia*, *Haemophilus*, *Micrococcus* (*M. luteus*), *Moraxella* (*M. osloensis*), *Neisseria*, *Paracoccus* (*P. yeei*), *Pseudomonas* (*P. putida*), *Rothia*, *Staphylococcus* (*S. epidermidis*, *S. hominis*) and *Streptococcous* (*S. mitis*, *S. pneumoniae* and *S. salivarius*), which agreed with 16S rRNA gene amplicon profiles from a healthy cohort of infants (using leg skin samples) by Murphy et al. [449]. This further

suggests that this study may provide the first potential baseline for a healthy infant skin microbiome generated using SMS.

Maternal and Infant Skin Archaeal Composition Over Time

Though the focus of this thesis was on the bacteriome, the infant skin archaeome was explored to address a research gap in the microbiome field. Using a specialised archaeome pipeline, the first species-level skin archaeome was identified in an infant cohort. The same archaeal species were found on both maternal and infant skin including *Halorubrum lipolyticum*, *Haloferax massiliensis*, *Methanobacterium sp. 000499765*, and various *Methanobrevibacter* and *Methanosphaera* species. All archaeal species were present at very low abundance, consistent with other research on the human skin archaeome [30, 453-456]. While timepoint did not appear to drive archaeal abundance on maternal and infant skin at the individual species level, it did influence archaeal abundance across and between species in the combined maternal and infant datasets. These trends were still apparent when accounting for pregnancy term, delivery mode and feeding mode, although feeding mode specifically influenced the abundance of archaea on the maternal skin.

Maternal and Infant Skin, Stool and Vaginal Diversity Changes Over Time

With a focus on the infant microbiome, these trends indicate age as a defining factor for the within (alpha) and between (beta) sample microbial diversity changes on infant skin and within infant stools; thus, the establishment and development of the infant skin and gut microbiome. From an alpha diversity perspective, gestational age exhibited additional influences on the infant skin and gut microbiome, with delivery and feeding mode having additional influences on the gut microbiome. From a beta diversity perspective, gestational age, delivery and feeding mode had additional influences on the infant skin microbiome and feeding mode had additional influences on the infant gut microbiome. The consistency in diversity analysis trends over time aligns with findings from several studies that highlight that the skin microbiome is defined by age [1, 4, 14, 18, 56, 175, 176].

Shared Microbiota Over Time between Maternal and Infant Skin, Stool and Vaginal Body Sites

The common microbiota (*Bifidobacterium*, *Micrococcus*, *Staphylococcus* and *Streptococcus*) over time across all maternal and infant dataset indicate the occurrence of microbial sharing between mother and infant pairs across different

body sites (skin, gut and vagina) over time (all of which should be culturable from the culture collection); and may have a functional importance in the developing infant microbiomes. Timepoint emerged as the primary driver of abundance changes of the shared microbes across all maternal and infant body sites. With a focus on the infant microbiome, the trends indicate age as a defining factor for the relative abundance changes of the common genera and species on infant skin and within infant stools. The influence of pregnancy term and feeding mode on the relative abundance changes of the shared microbes over time in infant stools could indicate the involvement of gestational age and breast feeding in the establishment and development of the infant skin and gut microbiomes. Delivery mode did not appear to have an influence. The consistency in results for the shared microbial abundance and differential associations across and within maternal and infant body sites, suggests timepoint is the main determinant of the infant skin microbiome development. This aligns with findings from several studies that highlight that the skin microbiome is defined by age [1, 4, 14, 18, 56, 175, 176].

Strain Level Sharing of the Common Microbiota between Maternal and Infant Skin, Stool and Vaginal Body Sites Over Time

Species specific strain level sharing was observed between certain body sites including maternal and infant skin swab (*Streptococcus*, *Staphylococcus* and *Micrococcus*), maternal stool and infant skin swab (*Staphylococcus* and *Bifidobacterium*), infant stool and maternal skin swab (*Streptococcus*, *Staphylococcus*, and *Bifidobacterium*), infant and maternal stool (*Streptococcus* and *Bifidobacterium*) and infant stool and maternal vaginal swab body sites (*Staphylococcus* and *Bifidobacterium*). Infants displayed a mix of maternal related (same strains shared between mother and infant) and unrelated strains, which has been previously highlighted [457]. *S. epidermidis* and *M. luteus* exhibited high levels of unrelatedness between mothers and infants, indicating low vertical transmission, and suggesting that they may have been acquired by other means. Research suggests that vertically transmitted strains are more likely to persist in the infant microbiome and may play a role in its development [458, 459].

No strain level sharing was detected between infant skin and maternal vaginal sites, suggesting that the maternal vaginal microbiome does not contribute significantly to microbial transfer to the infant skin. Instead, the maternal skin appeared to be the main source of microbial transfer to infant skin, followed by the maternal gut.

Similarly, maternal skin was the main source of microbial transfer to the infant gut, followed by the maternal gut and vaginal sites. While research on strain level sharing between maternal and infant skin is limited, with only one study demonstrating *C. acnes* sharing between maternal and infant skin [31], more studies have focused on gut microbiome sharing between maternal body sites and infant gut [361, 457-464].

In contrast to this study's findings, other research has shown a smaller contribution of maternal skin to infant stool transfer [461]. The mother vaginal site was not significant for microbial transfer to infant skin, though other studies have shown its' importance in microbial transfer to the infant gut. VD infants were previously found to share more gut strains with their mothers than CS infants, suggesting delivery mode as a stronger factor for gut strain sharing [461, 462]. This pattern was also reflected in this study, with fewer species and strain transmission between maternal and infant stools in CS infants. One study contradicted this by suggesting that the vaginal microbiome is unlikely to be a major contributor of maternal to infant gut microbial transmission [464]. More strains of *Bifidobacterium* were shared between maternal and infant stools in infants delivered via VD than CS, consistent with previous research [457]. CS deliveries may disrupt maternal-infant microbial seeding, as demonstrated previously [462]. This 'disruption' could potentially be mitigated by exclusive breastfeeding, as more *Bifidobacterium* strains were shared between maternal and infant pairs (maternal and infant stool, maternal skin swab and infant stool, and maternal vaginal swab and infant stool) that solely breastfed compared to those that had a combination of breast and formula milk. Similar findings have been reported where *Bifidobacterium* strains were identified in both breast milk and infant stool of FT infants [458]. In this study, FT infants shared more *Bifidobacterium* strains than LP infants. Strain sharing between maternal and infant skin was not detected at B or at M12, although rapid microbial influx has been identified by others between the infant and maternal gut, and other maternal body sites at birth [458]. Strain sharing between maternal and infant stools was present at B and M12, aligning with previous research that showed maternally shared strains contributed substantial fractions to the infant microbiome during the first two years of life [461].

Generation of a PEARL Study MAGS Genome Catalogue

Overall, these results support earlier compositional and microbial sharing findings, confirming the sharing of core genera between maternal and infant pairs across

multiple body sites over time. Additionally, the trees indicate that the composition remained relatively consistent over time, as no clustering was identified at specific timepoints. An overview of the species-level MAGs and their counts revealed greater species diversity in the maternal skin and stool sites compared to their infants.

Limitations

This study involved a multifactorial analysis, accounting for timepoint, pregnancy term, delivery and feeding mode, highlighting the complexity of the data. However, some limitations were present. This study focused on the bacteriome. However, an overarching view of all microbiome communities within mothers and infants, including fungal, archaeome and viral communities with the bacteriome, will enable a more comprehensive analysis of all microbiome communities and provide a clearer picture of how these ecosystems interact in the establishment and development of the infant skin microbiome, and consequently its impact on skin health.

Furthermore, this study included several body sites from mothers and infants, including the skin, stool and vagina. However, understanding the parental and family contributions to the early life microbiome establishment would be beneficial, as explored in other studies [461, 463]. The inclusion of samples from different skin sites, oral and breast milk samples, would further provide a more complete understanding of the infant skin microbiome, as different body environments support distinct microbial communities [2, 15, 36, 37]. The detection of oral related microbiota on infant skin further emphasises the need for these additional sample types.

Expanding the study to include more mother-infant pairs, later timepoints up to 2 years of age, as well as early pre-term, formula fed infants, and both emergency and elective CS maternal and infant pairs, would provide a deeper insight into the influence of timepoint, pregnancy term, and feeding and delivery mode on the infant skin microbiome establishment and development. This would also enhance the statistical power of strain level sharing analysis. However, due to the ongoing recruitment in the PEARL study, there were not enough mother-infant pairs meeting these criteria at the time of analysis to ensure sufficient statistical power.

Moreover, incorporating factors, such as antibiotic use and environmental (e.g. hospital and home) influences, would provide a fuller understanding of delivery

mode effects on the infant skin microbiome. Previous research has shown antibiotics can markedly influence the infant gut microbiota fitness and dynamics [463], and environmental factors may explain the presence of maternally unrelated strains identified in this study. Hospital environments warrant further exploration, as they have been identified as reservoirs for early colonisers (*Sta. epidermidis*, *Sta. haemolyticus*, *Ent. faecalis*, *Psuedomonas aeruginosa*, *Aeromonas hydrophilia*, *Cutibacterium avidum*, *Esc. coli* and *Klebsiella pneumoniae*), which may pose risks to pre-term infant health [465].

There was some contamination of the negative controls. Contamination commonly occurs in metagenomic studies, especially those with low biomass samples [416]. Several studies have identified contamination sources occurring from neighbouring samples and the 'kitome' [405, 416]. Contamination within a dataset can be identified and removed using bioinformatic techniques [407, 409, 417-419] although low biomass samples have a higher risk of true microbial microbiota members being removed [420]. Given the background contaminants in the controls were at a very low level and mostly represented species not seen in the test samples, they were not removed as they had a negligible impact on the profiles produced.

Finally, this chapter aimed to explore the diversity of genomes present in the samples rather than a full functional analysis, therefore, a quality assessment of the MAGs was run following a set of parameters (>80 % completeness, <10 % contamination) established in Kiu et al. [398]. As future work will identify the functional capacities of generated MAGs for the shared common microbiota between mothers and infants, with a focus on providing new insight into their role in skin health, further analysis will be undertaken for the quality assessment using the MIMAG criteria set out in Bowers et al. [426]. In addition to providing the completeness and contamination information, other information will be highlighted (species name, length of MAG sequence in nt, number of contigs and N50 values) for each MAG. The MAGs will then be annotated using Prokka or Bakta and the number of coding sequences, rRNAs and tRNAs will be reported. Then using the criteria set out in Bowers et al. [426] it can be determined whether the MAGs represent draft genomes.

4.8 Conclusion

This study provides the first baseline for a healthy infant skin microbiome using SMS. A consensus of microbial sharing was observed between body sites, with

common microbiota identified across both maternal and infant samples over time. Notably, this is also the first study to characterise the infant skin archaeome to species level using SMS. From a multi body site microbial sharing perspective, the infant skin microbiome was primarily influenced by the maternal skin, with minimal contributions from other body sites. Timepoint was the main determinant of skin microbiome development. No strain level sharing was observed between infant skin and the maternal vaginal microbiome, suggesting that this does not play a major part in microbial transfer to infant skin. Instead, the maternal skin was the main source of microbial transfer to infant skin, followed by the maternal gut. Similarly, the maternal skin was the main source of microbial transfer to the infant gut, followed by the maternal gut and vagina. Vertical transmission of *S. epidermidis* and *M. luteus* strains between mothers and infants was low, as most strains identified were different between mothers and infants, indicating transmission from non-maternal sources. As strain level sharing was observed for some species, but not for others, further investigation is needed to fully understand microbial transmission pathways between mothers, infants, and other potential sources. Subsequent studies should expand on these findings, to explore species not covered within the scope of this study. Furthermore, a PEARL genome catalogue was generated, containing 4,408 high quality MAGs. The MAGs generated support earlier compositional and microbial sharing findings between maternal and infant pairs across multiple body sites over time. Additionally, the trees indicate that the composition remained relatively consistent over time, as no clustering was identified at specific timepoints. An overview of the species-level MAGs and their counts revealed greater species diversity in the maternal skin and stool sites compared to their infants. Future work will focus on identifying the functional capacities of generated MAGs for the shared common microbiota across maternal and infant body sites. This will provide new insight into their role in skin health.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

CHAPTER 5 OVERVIEW

In this chapter, the main conclusions from the thesis are summarised, along with limitations and suggestions for future work that could enhance and expand the robustness of this complex, multifactorial dataset involving mother and infant pairs.

The overarching conclusions from this research highlight the success in developing an optimised medium-throughput DNA extraction, Shotgun Metagenomic Sequencing (SMS), and analysis approach, which effectively characterised the skin microbiota of both adults and infants. This method was successfully applied to the Pregnancy and Early Life (PEARL) cohort, allowing for an in-depth analysis of temporal changes in the skin microbiome of mothers and infants, as well as comparisons with gut and vaginal microbiota. The research identified microbial sharing between body sites and accounted for key factors influencing the infant skin microbiome, such as pregnancy term (late-preterm vs full term), delivery (caesarean section vs vaginal delivery) and feeding mode (breast vs breast/formula combination fed). Timepoint was the main determinant linked to changes in the skin microbiome development. Core genera - *Bifidobacterium, Micrococcus, Streptococcus, and Staphylococcus* - were consistently shared between mothers and infants across multiple body sites, with the maternal skin being the primary source of microbial transfer to the infant's skin and gut. The maternal vaginal site was not a major source of microbial transfer to the infant skin. Strain-level sharing of *S. epidermidis* and *M. luteus* between mothers and infants was minimal, suggesting other non-maternal sources of transmission. This study also produced the first potential baseline for a healthy infant skin microbiome using SMS. It further provided the first indication of the infant skin archaeome at the species level. A PEARL genome catalogue with 4,408 high-quality MAGs was generated, showing consistency with taxonomic profiles and confirming the shared genera between mother-infant pairs over time.

This study had multiple strengths, including its multifactorial analysis involving infants and matched maternal samples, to explore the influence of maternal factors on the establishment and development of the infant skin microbiome. This study accounted for maternal influential factors including pregnancy term, delivery and feeding mode. Several body sites from mothers and infants, including the skin, stool and vagina were investigated, providing a thorough picture of microbial sharing between mothers and infants, and the influence of the mothers' microbiome on the establishment and development of the infant skin microbiome. This study was also longitudinal following infants at multiple timepoints from birth up to 1 year of age. This study provided a robust investigation of the infant skin microbiome to genus, species and strain level using shotgun metagenomics. It also enables the future investigation of the function of core members of the infant skin microbiome. Though the bacteriome was of focus, Shotgun metagenomics also enabled the exploration

of other low biomass communities within infant skin, such as the archaeome, and will enable further investigation of other communities of interest in the future. This robust compositional investigation would have been restricted with traditional 16s rRNA amplicon sequencing, which would have removed the future ability to explore the microbiome functions [13, 19, 288, 317]. This study also used integrative approaches to study the infant skin microbiome in more detail utilising shotgun metagenomics and culturomics. There are few studies utilising an integrative approach, which is necessary to robustly study the infant skin microbiome composition and function [20, 22, 39-43]. Current research in this field focuses on diseased infant skin [9, 14, 22, 24, 72-75], whilst this study focused on a healthy cohort of infants with healthy mothers. This research, therefore, satisfied a pressing need to use an integrative approach to establish a baseline composition of the healthy infant skin microbiome longitudinally, which will help clarify the roles of microbes in infant skin health. This work also identified the various intrinsic (age, gestational age) and extrinsic (delivery and feeding mode) factors, that alter the infant skin microbiome establishment and development, and identified specific microbes of interest, that may be important to infant skin health. This will aid future work in deepening the understanding of the functional role of these microbes in infant skin health and disease prevention. Finally, this work further provided the first indication of the infant skin archaeome at the species level, which is a highly understudied community within infant skin [12].

However, some limitations were present. This study focused on the bacteriome with an insight into the archaeome. Therefore, future work would be to explore other communities, including the mycobiome and virome. Having an overarching view of all microbiome communities within mothers and infants, will enable a more comprehensive analysis of all microbiome communities and provide a clearer picture of how these ecosystems interact in the establishment and development of the infant skin microbiome, and consequently its impact on skin health. Furthermore, these microbial communities in skin can differ across different skin body sites, due to differing body conditions, which support distinct microbial communities [2, 15, 36, 37].Therefore, it would have been beneficial to have skin samples from multiple body sites, however, the PEARL study only collected skin samples from the forearm. Furthermore, the detection of oral related microbiota on infant skin in this thesis, emphasises the need for additional sample types (such as oral and breast milk) that unfortunately weren’t collected in the PEARL cohort. Though this study explored the

influence of maternal factors on the infant skin microbiome establishment and development, and understanding of both parental (including paternal) and other family contributions, to the early life microbiome establishment would be beneficial, as explored in other studies [461, 463].

Expanding the study to include a larger number of mother-infant pairs with samples and metadata collected from pairs at later timepoints (up to 2 years of age), as well as early pre-term, formula only fed, and both emergency and elective caesarean pairs, would provide a deeper insight into the influence of the investigated factors (timepoint, pregnancy term, feeding and delivery mode) on the infant skin microbiome establishment and development. This would also enhance the statistical power of strain level sharing analysis. However, as this study was conducted during the PEARL study recruitment phase, during the COVID period, which delayed PEARL recruitment and sample collection, this study was delayed, and there were not enough mother-infant pairs meeting these additional criteria at the time of this studies commencement.

Moreover, incorporating factors, such as antibiotic use and environmental (e.g. hospital and home) influences, would provide a fuller understanding of delivery mode effects on the infant skin microbiome. Previous research has shown antibiotics can markedly influence the infant gut microbiota fitness and dynamics [463], and environmental factors may explain the presence of maternally unrelated strains identified in this study. Hospital environments warrant further exploration, as they have been identified as reservoirs for early colonisers (*Sta. epidermidis*, *Sta. haemolyticus*, *Ent. faecalis*, *Psu. aeruginosa*, *Aer. hydrophilia*, *Cut. avidum*, *Esc. coli* and *Kle. pneumoniae*), which may pose risks to pre-term infant health [465].

Finally, the culturomics work in this thesis used a suitable media (CBA) to encompass a broad range of microbes [390], which has been successful in other studies in obtaining a good representation of skin bacteria [391]. However, it is acknowledged that different communities (archaeome, mycobiome and virome) and certain microbes require special conditions to facilitate their recovery (such as, *Corynebacteria* that require Tween 80 to be added to medium as they need access to lipids) [466], which could potentially underestimate their prevalence. Nevertheless, culturomics was not the focus of this study, and was only used to identify what was culturable, and whether microbes of interest from the shotgun data (sequenced directly from skin swabs) could be cultivated for future functional work.

Leading on from this thesis, future research could expand on these findings by investigating the functional capacities of the shared microbiota and their relationship to skin health. This could be investigated by functionally characterising MAGs of interest (i.e., *Streptococcus*, *Micrococcus*, *Staphylococcus* and *Bifidobacterium*), that were generated in the final chapter of this thesis, using the Distilled and Refined Annotation of Metabolism (DRAM) tool [467], which would infer the main metabolic functions of interest detected in the MAGs, relating to skin that support skin microbial development and stability. Prior to the use of DRAM for functional characterisation, a thorough quality assessment of the MAGs of interest would be required using a set of parameters and criteria (>80 % completeness, <10 % contamination, species name, length of MAG sequence in nt, number of contigs and N50 values) established in Kiu et al. [398] and the MIMAG recommendations in Bowers et al. [426]. The MAGs should then be annotated using Prokka or Bakta, and the number of coding sequences, rRNAs and tRNAs be reported. Then using the criteria set out in Bowers et al. [426] it can be determined whether the MAGs represent draft genomes.

Additionally, a more in-depth strain level analysis could be performed to identify all strains being shared between mother and infant pairs for a thorough investigation to compare what is present in infant and maternal samples. Exploring microbial transmission pathways and the mechanisms of whether maternal strains are more likely to persist in infants than non-maternal strains could be examined, as this may be important in the infant skin microbiome establishment [457]. Prediction of these microbiome interactions using bioinformatic network analysis and more complex computational modelling such as Minimal Microbial Consortia (MiMiC), as shown in some studies [449, 468], would aid the design of probiotic interventions for maintaining healthy skin or preventing skin disease in early life. The MiMiC predictions could then identify culturable microbiota within this study to be applied on *in vitro* 3D skin models to mimic a functioning, healthy infant skin microbiome, and study the effects of individual taxa and the microbial consortia in skin health, in a controlled context, as utilised in some studies with *Streptococcus*, *Staphylococcus* and *Micrococcus* [469]. This could provide critical insights into how microbiota influence skin processes and immune responses, enabling the development of precision probiotics for skin health. *Micrococcus* has already been trialled in skincare for therapeutic use, with promising results in promoting skin health [470]. Additionally, procedures such as vaginal seeding and Faecal Microbiota

Transplantation (FMT) could be explored to replicate and restore maternal microbial seeding in caesarean delivered infants, to prevent skin disease in early life, and has been successfully performed in some studies [463]. MiMiC predictions could select the appropriate microbiota for use in FMT trials. Though, further research is needed to address associated safety concerns [159, 213].

In conclusion, this research presents an optimised method for skin microbiota characterisation, establishes a baseline composition for the healthy infant skin microbiome, and opens new avenues for probiotic development based on evidence from compositional analyses.

REFERENCES

1. Capone, K.A., et al., *Diversity of the Human Skin Microbiome Early in Life.* Journal of Investigative Dermatology, 2011. **131**(10): p. 2026-2032.
2. Grice, E.A. and J.A. Segre, *The skin microbiome.* Nature Reviews Microbiology, 2011. **9**(4): p. 244-253.
3. Smythe, P. and H.N. Wilkinson, *The Skin Microbiome: Current Landscape and Future Opportunities.* International Journal of Molecular Sciences, 2023. **24**(4): p. 3950.
4. Huttenhower, C., et al., *Structure, function and diversity of the healthy human microbiome.* Nature, 2012. **486**(7402): p. 207-214.
5. Arumugam, M., et al., *Enterotypes of the human gut microbiome.* Nature, 2011. **473**(7346): p. 174-80.
6. Aagaard, K., et al., *A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy.* PLoS One, 2012. **7**(6): p. e36466.
7. Costello, E.K., et al., *Bacterial community variation in human body habitats across space and time.* Science, 2009. **326**(5960): p. 1694-7.
8. National Academies of Sciences, E., et al., *The National Academies Collection: Reports funded by National Institutes of Health*, in *Environmental Chemicals, the Human Microbiome, and Health Risk: A Research Strategy*. 2017, National Academies Press (US)

Copyright 2018 by the National Academy of Sciences. All rights reserved.: Washington (DC).

9. Townsend, E.C. and L.R. Kalan, *The dynamic balance of the skin microbiome across the lifespan.* Biochemical Society Transactions, 2023. **51**(1): p. 71-86.
10. Cho, I. and M.J. Blaser, *The human microbiome: at the interface of health and disease.* Nature Reviews Genetics, 2012. **13**(4): p. 260-270.
11. Dwyer, L.R. and T.C. Scharschmidt, *Early life host-microbe interactions in skin.* Cell Host Microbe, 2022. **30**(5): p. 684-695.
12. Skowron, K., et al., *Human Skin Microbiome: Impact of Intrinsic and Extrinsic Factors on Skin Microbiota.* Microorganisms, 2021. **9**(3): p. 543.
13. Sfriso, R., et al., *Revealing the secret life of skin - with the microbiome you never walk alone.* International Journal of Cosmetic Science, 2020. **42**(2): p. 116-126.
14. Rapin, A., et al., *The skin microbiome in the first year of life and its association with atopic dermatitis.* Allergy, 2023.
15. Grice, E.A., et al., *Topographical and Temporal Diversity of the Human Skin Microbiome.* Science, 2009. **324**(5931): p. 1190-1192.
16. Hellmann, K.T., et al., *Viability-Resolved Metagenomics Reveals Antagonistic Colonization Dynamics of Staphylococcus epidermidis Strains on Preterm Infant Skin.* mSphere, 2021. **6**(5): p. e0053821.
17. Schoch, J.J., et al., *Skin microbiome sampling in the preterm neonate.* Pediatr Dermatol, 2023. **40**(1): p. 129-131.
18. Zhu, T., et al., *Age and Mothers: Potent Influences of Children's Skin Microbiota.* Journal of Investigative Dermatology, 2019. **139**(12): p. 2497-2505.e6.
19. Serghiou, I.R., et al., *An efficient method for high molecular weight bacterial DNA extraction suitable for shotgun metagenomics from skin swabs.* Microbial Genomics, 2023. **9**(7).
20. Saheb Kashaf, S., et al., *Integrating cultivation and metagenomics for a multi-kingdom view of skin microbiome diversity and functions.* Nature Microbiology, 2022. **7**(1): p. 169-179.
21. Hayes, C.K., et al., *Herpes Simplex Virus-2 Variation Contributes to Neurovirulence During Neonatal Infection.* J Infect Dis, 2022. **226**(9): p. 1499-1509.
22. Nguyen, U.T. and L.R. Kalan, *Forgotten fungi: the importance of the skin mycobiome.* Current Opinion in Microbiology, 2022. **70**: p. 102235.
23. Wang, Y.R., et al., *Infant Mode of Delivery Shapes the Skin Mycobiome of Prepubescent Children.* Microbiol Spectr, 2022. **10**(5): p. e0226722.
24. Tao, R., et al., *Altered skin fungal and bacterial community compositions in tinea capitis.* Mycoses, 2022. **65**(8): p. 834-840.
25. Shah, P.A., et al., *Exploring the Skin Mycobiome in Very Preterm babies during the early neonatal period in a Neonatal Intensive Care Unit of India.* Trop Doct, 2022. **52**(2): p. 362-364.
26. Harding, K.R., N. Kyte, and P.C. Fineran, *Jumbo phages.* Current Biology, 2023. **33**(14): p. R750-R751.
27. Kuehnast, T., et al., *Exploring the human archaeome: its relevance for health and disease, and its complex interplay with the human immune system.* The FEBS Journal. **n/a**(n/a).

28. Probst, A.J., A.K. Auerbach, and C. Moissl-Eichinger, *Archaea on Human Skin.* PLOS ONE, 2013. **8**(6): p. e65388.
29. Mahnert, A., et al., *Molecular Tracking and Cultivation Reveal Ammonia Oxidizing Archaea as Integral Members of the Human Skin Microbiome.* bioRxiv, 2024: p. 2024.08.05.606590.
30. Moissl-Eichinger, C., et al., *Human age and skin physiology shape diversity and abundance of Archaea on skin.* Sci Rep, 2017. **7**(1): p. 4039.
31. Shen, Z., et al., *A genome catalog of the early-life human skin microbiome.* Genome Biology, 2023. **24**(1): p. 252.
32. Belkaid, Y. and J.A. Segre, *Dialogue between skin microbiota and immunity.* Science, 2014. **346**(6212): p. 954-9.
33. Oh, J., et al., *Biogeography and individuality shape function in the human skin metagenome.* Nature, 2014. **514**(7520): p. 59-64.
34. Balato, A., et al., *Human Microbiome: Composition and Role in Inflammatory Skin Diseases.* Arch Immunol Ther Exp (Warsz), 2019. **67**(1): p. 1-18.
35. Grice, E.A., et al., *A diversity profile of the human skin microbiota.* Genome Res, 2008. **18**(7): p. 1043-50.
36. Pammi, M., et al., *Development of the cutaneous microbiome in the preterm infant: A prospective longitudinal study.* PLoS One, 2017. **12**(4): p. e0176669.
37. Byrd, A.L., Y. Belkaid, and J.A. Segre, *The human skin microbiome.* Nature Reviews Microbiology, 2018. **16**(3): p. 143-155.
38. Meisel, J.S., et al., *Commensal microbiota modulate gene expression in the skin.* Microbiome, 2018. **6**(1): p. 20.
39. Orchanian, S.B., et al., *Multiomic Analyses of Nascent Preterm Infant Microbiomes Differentiation Suggest Opportunities for Targeted Intervention.* Adv Biol (Weinh), 2022. **6**(8): p. e2101313.
40. Capone, K.A., et al., *A randomized clinical study on the effects of emollient use on the developing infant skin microbiome and metabolome.* Experimental Dermatology, 2023. **32**(1): p. 75-77.
41. Mishra, A., et al., *Microbial exposure during early human development primes fetal immune cells.* Cell, 2021. **184**(13): p. 3394-3409.e20.
42. Roux, P.F., T. Oddos, and G. Stamatas, *Deciphering the Role of Skin Surface Microbiome in Skin Health: An Integrative Multiomics Approach Reveals Three Distinct Metabolite–Microbe Clusters.* J Invest Dermatol, 2022. **142**(2): p. 469-479.e5.
43. Kennedy, K.M., et al., *Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies.* Nature, 2023. **613**(7945): p. 639-649.
44. Hu, X., et al., *Changes in the skin microbiome during male maturation from 0 to 25 years of age.* Skin Res Technol, 2023. **29**(9): p. e13432.
45. Serghiou, I.R., M.A. Webber, and L.J. Hall, *An update on the current understanding of the infant skin microbiome and research challenges.* Current Opinion in Microbiology, 2023. **75**: p. 102364.
46. Ridaura, V.K., et al., *Gut microbiota from twins discordant for obesity modulate metabolism in mice.* Science, 2013. **341**(6150): p. 1241214.
47. Qin, J., et al., *A metagenome-wide association study of gut microbiota in type 2 diabetes.* Nature, 2012. **490**(7418): p. 55-60.
48. Morgan, X.C., et al., *Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment.* Genome Biol, 2012. **13**(9): p. R79.
49. Schulz, M.D., et al., *High-fat-diet-mediated dysbiosis promotes intestinal carcinogenesis independently of obesity.* Nature, 2014. **514**(7523): p. 508-512.
50. Hooper, L.V. and J.I. Gordon, *Commensal host-bacterial relationships in the gut.* Science, 2001. **292**(5519): p. 1115-8.
51. Turnbaugh, P.J., et al., *An obesity-associated gut microbiome with increased capacity for energy harvest.* Nature, 2006. **444**(7122): p. 1027-31.
52. Ravel, J., et al., *Vaginal microbiome of reproductive-age women.* Proc Natl Acad Sci U S A, 2011. **108 Suppl 1**(Suppl 1): p. 4680-7.
53. Yan, J., L. Sheng, and H. Li, *Akkermansia muciniphila: is it the Holy Grail for ameliorating metabolic diseases?* Gut Microbes, 2021. **13**(1): p. 1984104.
54. Lyu, Q., et al., *Bifidobacterium alleviate metabolic disorders via converting methionine to 5'-methylthioadenosine.* Gut Microbes, 2024. **16**(1): p. 2300847.
55. Zhou, X., et al., *Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women.* The ISME Journal, 2007. **1**(2): p. 121-133.
56. Chu, D.M., et al., *Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery.* Nat Med, 2017. **23**(3): p. 314-326.

57. Croitoru, D.O. and V. Piguet, *A Mother's Touch: Emerging Roles in Development of the Cutaneous Microbiome.* J Invest Dermatol, 2019. **139**(12): p. 2414-2416.
58. Kong, H.H., *Skin microbiome: genomics-based insights into the diversity and role of skin microbes.* Trends in Molecular Medicine, 2011. **17**(6): p. 320-328.
59. Marchini, G., et al., *Erythema toxicum neonatorum is an innate immune response to commensal microbes penetrated into the skin of the newborn infant.* Pediatr Res, 2005. **58**(3): p. 613-6.
60. Cogen, A.L., V. Nizet, and R.L. Gallo, *Skin microbiota: a source of disease or defence?* Br J Dermatol, 2008. **158**(3): p. 442-55.
61. Scharschmidt, T.C., et al., *A Wave of Regulatory T Cells into Neonatal Skin Mediates Tolerance to Commensal Microbes.* Immunity, 2015. **43**(5): p. 1011-1021.
62. Lai, Y., et al., *Commensal bacteria regulate Toll-like receptor 3-dependent inflammation after skin injury.* Nat Med, 2009. **15**(12): p. 1377-82.
63. Scharschmidt, T.C., *Establishing Tolerance to Commensal Skin Bacteria: Timing Is Everything.* Dermatol Clin, 2017. **35**(1): p. 1-9.
64. Scharschmidt, T.C., et al., *Commensal Microbes and Hair Follicle Morphogenesis Coordinately Drive Treg Migration into Neonatal Skin.* Cell Host Microbe, 2017. **21**(4): p. 467-477.e5.
65. Findley, K. and E.A. Grice, *The skin microbiome: a focus on pathogens and their association with skin disease.* PLoS Pathog, 2014. **10**(10): p. e1004436.
66. Huang, J.T., et al., *Treatment of Staphylococcus aureus colonization in atopic dermatitis decreases disease severity.* Pediatrics, 2009. **123**(5): p. e808-14.
67. McDowell, A., et al., *A novel multilocus sequence typing scheme for the opportunistic pathogen Propionibacterium acnes and characterization of type I cell surface-associated antigens.* Microbiology (Reading), 2011. **157**(Pt 7): p. 1990-2003.
68. Oh, J., et al., *Shifts in human skin and nares microbiota of healthy children and adults.* Genome Medicine, 2012. **4**(10): p. 77.
69. O'Connor, N.R., M.R. McLaughlin, and P. Ham, *Newborn skin: Part I. Common rashes.* Am Fam Physician, 2008. **77**(1): p. 47-52.
70. Tanaka, A., et al., *Comprehensive pyrosequencing analysis of the bacterial microbiota of the skin of patients with seborrheic dermatitis.* Microbiol Immunol, 2016. **60**(8): p. 521-6.
71. Kong, H.H., et al., *Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis.* Genome Res, 2012. **22**(5): p. 850-9.
72. McKeown, K.E. and R.B. Baker, *Staphylococcal Scalded Skin Syndrome, Identification, and Wound Care: A Case Report Series.* Advances in Neonatal Care, 2022. **22**(4): p. 325-332.
73. Manos, J., *The human microbiome in disease and pathology.* APMIS, 2022. **130**(12): p. 690-705.
74. Olesen, C.M., et al., *Altered Maturation of the Skin Microbiome in Infants with Atopic Dermatitis.* Acta Derm Venereol, 2022. **102**: p. adv00817.
75. Teufel, A., et al., *Characterization of the microbiome in the infant diapered area: Insights from healthy and damaged skin.* Exp Dermatol, 2021. **30**(10): p. 1409-1417.
76. Monteagudo, B., et al., *Prospective study of erythema toxicum neonatorum: epidemiology and predisposing factors.* Pediatr Dermatol, 2012. **29**(2): p. 166-8.
77. Prescott, S.L., et al., *The skin microbiome: impact of modern environments on skin ecology, barrier integrity, and systemic immune programming.* World Allergy Organ J, 2017. **10**(1): p. 29.
78. Kaiser, C., et al., *Microbial community dynamics alleviate stoichiometric constraints during litter decay.* Ecology Letters, 2014. **17**(6): p. 680-690.
79. Foster, J.A., S.M. Krone, and L.J. Forney, *Application of ecological network theory to the human microbiome.* Interdiscip Perspect Infect Dis, 2008. **2008**: p. 839501.
80. Prosser, J.I., et al., *The role of ecological theory in microbial ecology.* Nat Rev Microbiol, 2007. **5**(5): p. 384-92.
81. Marshall, B.M., D.J. Ochieng, and S.B. Levy. *Commensals: Underappreciated Reservoir of Antibiotic Resistance Probing the role of commensals in propagating antibiotic resistance should help preserve the efficacy of these critical drugs*. 2009.
82. Schommer, N.N. and R.L. Gallo, *Structure and function of the human skin microbiome.* Trends Microbiol, 2013. **21**(12): p. 660-8.
83. Moskovicz, V., A. Gross, and B. Mizrahi, *Extrinsic Factors Shaping the Skin Microbiome.* Microorganisms, 2020. **8**(7).
84. Candela, M., et al., *Unbalance of intestinal microbiota in atopic children.* BMC Microbiology, 2012. **12**(1): p. 95.
85. Penders, J., et al., *Gut microbiota composition and development of atopic manifestations in infancy: the KOALA Birth Cohort Study.* Gut, 2007. **56**(5): p. 661-7.

86. Kalliomäki, M., et al., *Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing.* J Allergy Clin Immunol, 2001. **107**(1): p. 129-34.
87. Wang, M., et al., *Reduced diversity in the early fecal microbiota of infants with atopic eczema.* J Allergy Clin Immunol, 2008. **121**(1): p. 129-34.
88. He, F., et al., *Intestinal Bifidobacterium species induce varying cytokine production.* Journal of Allergy and Clinical Immunology, 2002. **109**(6): p. 1035-1036.
89. Chua, H.H., et al., *Intestinal Dysbiosis Featuring Abundance of Ruminococcus gnavus Associates With Allergic Diseases in Infants.* Gastroenterology, 2018. **154**(1): p. 154-167.
90. Li, L., et al., *Probiotic Supplementation for Prevention of Atopic Dermatitis in Infants and Children: A Systematic Review and Meta-analysis.* Am J Clin Dermatol, 2019. **20**(3): p. 367-377.
91. Krieger, Y., et al., *Meconium-stained amniotic fluid as a protective factor against childhood dermatitis and skin rash-related hospitalization in the offspring - a population-based cohort analysis.* J Eur Acad Dermatol Venereol, 2020. **34**(2): p. 319-324.
92. Salem, I., et al., *The Gut Microbiome as a Major Regulator of the Gut-Skin Axis.* Front Microbiol, 2018. **9**: p. 1459.
93. Schoch, J.J., et al., *The infantile cutaneous microbiome: A review.* Pediatr Dermatol, 2019. **36**(5): p. 574-580.
94. Mahmud, M.R., et al., *Impact of gut microbiome on skin health: gut-skin axis observed through the lenses of therapeutics and skin diseases.* Gut Microbes, 2022. **14**(1): p. 2096995.
95. Łoś-Rycharska, E., et al., *The microbiome and its impact on food allergy and atopic dermatitis in children.* Postepy Dermatol Alergol, 2020. **37**(5): p. 641-650.
96. Andrade, P.D.S.M.A.d., et al., *Efficacy of Probiotics in Children and Adolescents With Atopic Dermatitis: A Randomized, Double-Blind, Placebo-Controlled Study.* Frontiers in Nutrition, 2022. **8**.
97. Pan, H. and J. Su, *Association of Probiotics with Atopic Dermatitis among Infant: A Meta-analysis of Randomized Controlled Trials.* Oxidative Medicine and Cellular Longevity, 2022. **2022**: p. 5080190.
98. Carucci, L., et al., *Therapeutic effects elicited by the probiotic Lacticaseibacillus rhamnosus GG in children with atopic dermatitis. The results of the ProPAD trial.* Pediatric Allergy and Immunology, 2022. **33**(8): p. e13836.
99. Łoś-Rycharska, E., et al., *A Combined Analysis of Gut and Skin Microbiota in Infants with Food Allergy and Atopic Dermatitis: A Pilot Study.* Nutrients, 2021. **13**(5).
100. Raimer, S.S., *Managing Pediatric Atopic Dermatitis.* Clinical Pediatrics, 2000. **39**(1): p. 1-14.
101. Kennedy, E.A., et al., *Skin microbiome before development of atopic dermatitis: Early colonization with commensal staphylococci at 2 months is associated with a lower risk of atopic dermatitis at 1 year.* J Allergy Clin Immunol, 2017. **139**(1): p. 166-172.
102. Weidinger, S. and N. Novak, *Atopic dermatitis.* The Lancet, 2016. **387**(10023): p. 1109-1122.
103. Tian, J., et al., *Global epidemiology of atopic dermatitis: a comprehensive systematic analysis and modelling study.* British Journal of Dermatology, 2023. **190**(1): p. 55-61.
104. Chiesa Fuxench, Z.C., *Atopic Dermatitis: Disease Background and Risk Factors*, in *Management of Atopic Dermatitis: Methods and Challenges*, E.A. Fortson, S.R. Feldman, and L.C. Strowd, Editors. 2017, Springer International Publishing: Cham. p. 11-19.
105. Zheng, T., et al., *The atopic march: progression from atopic dermatitis to allergic rhinitis and asthma.* Allergy Asthma Immunol Res, 2011. **3**(2): p. 67-73.
106. Dom, S., et al., *Pre- and post-natal exposure to antibiotics and the development of eczema, recurrent wheezing and atopic sensitization in children up to the age of 4 years.* Clin Exp Allergy, 2010. **40**(9): p. 1378-87.
107. Droste, J.H., et al., *Does the use of antibiotics in early childhood increase the risk of asthma and allergic disease?* Clin Exp Allergy, 2000. **30**(11): p. 1547-53.
108. Powers, C.E., et al., *Microbiome and pediatric atopic dermatitis.* J Dermatol, 2015. **42**(12): p. 1137-42.
109. Kobayashi, T., et al., *Dysbiosis and Staphylococcus aureus Colonization Drives Inflammation in Atopic Dermatitis.* Immunity, 2015. **42**(4): p. 756-66.
110. Moosavi, Z. and T. Hosseini, *One-year survey of cutaneous lesions in 1000 consecutive Iranian newborns.* Pediatr Dermatol, 2006. **23**(1): p. 61-3.
111. Reginatto, F.P., et al., *Epidemiology and Predisposing Factors for Erythema Toxicum Neonatorum and Transient Neonatal Pustular: A Multicenter Study.* Pediatric Dermatology, 2017. **34**(4): p. 422-426.
112. Marchini, G., et al., *Increased Expression of HMGB-1 in the Skin Lesions of Erythema Toxicum.* Pediatric Dermatology, 2007. **24**(5): p. 474-482.
113. Marchini, G., et al., *AQP1 and AQP3, psoriasin, and nitric oxide synthases 1-3 are inflammatory mediators in erythema toxicum neonatorum.* Pediatr Dermatol, 2003. **20**(5): p. 377-84.

114. Sarin, A., J. Bhatia, and R. Sarkar, *Diaper Dermatitis - A Narrative Review.* Asian Journal of Pediatric Dermatology, 2023. **1**(1): p. 19-23.
115. Bień, N., et al., *Infantile seborrheic dermatitis differential diagnosis based on case report.* Forum Dermatologicum, 2023. **9**(3): p. 123-125.
116. Gupta, A.K., et al., *Skin diseases associated with Malassezia species.* J Am Acad Dermatol, 2004. **51**(5): p. 785-98.
117. Soares, R.C., et al., *Malassezia intra-specific diversity and potentially new species in the skin microbiota from Brazilian healthy subjects and seborrheic dermatitis patients.* PLoS One, 2015. **10**(2): p. e0117921.
118. Ro, B.I. and T.L. Dawson, *The role of sebaceous gland activity and scalp microfloral metabolism in the etiology of seborrheic dermatitis and dandruff.* The journal of investigative dermatology. Symposium proceedings, 2005. **10**(3): p. 194-197.
119. Nagata, R., et al., *Transmission of the major skin microbiota, Malassezia, from mother to neonate.* Pediatr Int, 2012. **54**(3): p. 350-5.
120. Michaels, B.D. and J.Q. Del Rosso, *Tinea capitis in infants: recognition, evaluation, and management suggestions.* J Clin Aesthet Dermatol, 2012. **5**(2): p. 49-59.
121. Gołębiewski, M., et al., *Mother's Milk Microbiome Shaping Fecal and Skin Microbiota in Infants with Food Allergy and Atopic Dermatitis: A Pilot Analysis.* Nutrients, 2021. **13**(10).
122. Hourigan, S.K., M.G. Dominguez-Bello, and N.T. Mueller, *Can maternal-child microbial seeding interventions improve the health of infants delivered by Cesarean section?* Cell Host Microbe, 2022. **30**(5): p. 607-611.
123. Zioutis, C., et al., *Ecological Processes Shaping Microbiomes of Extremely Low Birthweight Infants.* Front Microbiol, 2022. **13**: p. 812136.
124. Sharma, A., et al., *Aqueous chlorhexidine 1% versus 2% for neonatal skin antisepsis: a randomised non-inferiority trial.* Archives of Disease in Childhood - Fetal and Neonatal Edition, 2021. **106**(6): p. 643-648.
125. Song, S.J., et al., *Naturalization of the microbiota developmental trajectory of Cesarean-born neonates after vaginal seeding.* Med, 2021. **2**(8): p. 951-964.e5.
126. Xu, Y., et al., *Antibiotic exposure prevents acquisition of beneficial metabolic functions in the preterm infant gut microbiome.* Microbiome, 2022. **10**(1): p. 103.
127. Shah, P.A., et al., *Mycobacterium Species on the Cutaneous Microbiome of Very Preterm Neonates.* J Trop Pediatr, 2022. **68**(2).
128. Dramowski, A., et al., *Impact of 1% chlorhexidine gluconate bathing and emollient application on bacterial pathogen colonization dynamics in hospitalized preterm neonates - A pilot clinical trial.* EClinicalMedicine, 2021. **37**: p. 100946.
129. Dominguez-Bello, M.G., et al., *Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns.* Proceedings of the National Academy of Sciences, 2010. **107**(26): p. 11971-11975.
130. Brooks, B., et al., *Microbes in the neonatal intensive care unit resemble those found in the gut of premature infants.* Microbiome, 2014. **2**(1): p. 1.
131. Grice, E.A. and J.A. Segre, *The human microbiome: our second genome.* Annu Rev Genomics Hum Genet, 2012. **13**: p. 151-70.
132. Perez-Muñoz, M.E., et al., *A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: implications for research on the pioneer infant microbiome.* Microbiome, 2017. **5**(1): p. 48.
133. Aagaard, K., et al., *The placenta harbors a unique microbiome.* Sci Transl Med, 2014. **6**(237): p. 237ra65.
134. Collado, M.C., et al., *Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid.* Scientific Reports, 2016. **6**(1): p. 23129.
135. Antony, K.M., et al., *The preterm placental microbiome varies in association with excess maternal gestational weight gain.* Am J Obstet Gynecol, 2015. **212**(5): p. 653.e1-16.
136. DiGiulio, D.B., et al., *Microbial prevalence, diversity and abundance in amniotic fluid during preterm labor: a molecular and culture-based investigation.* PLoS One, 2008. **3**(8): p. e3056.
137. Steel, J.H., et al., *Bacteria and inflammatory cells in fetal membranes do not always cause preterm labor.* Pediatr Res, 2005. **57**(3): p. 404-11.
138. Jiménez, E., et al., *Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section.* Curr Microbiol, 2005. **51**(4): p. 270-4.
139. Ardissone, A.N., et al., *Meconium microbiome analysis identifies bacteria correlated with premature birth.* PLoS One, 2014. **9**(3): p. e90784.

140. Hansen, R., et al., *First-Pass Meconium Samples from Healthy Term Vaginally-Delivered Neonates: An Analysis of the Microbiota.* PLoS One, 2015. **10**(7): p. e0133320.
141. Jiménez, E., et al., *Is meconium from healthy newborns actually sterile?* Res Microbiol, 2008. **159**(3): p. 187-93.
142. Chu, D.M., et al., *The early infant gut microbiome varies in association with a maternal high-fat diet.* Genome Medicine, 2016. **8**(1): p. 77.
143. Nikolovski, J., et al., *Barrier function and water-holding and transport properties of infant stratum corneum are different from adult and continue to develop through the first year of life.* J Invest Dermatol, 2008. **128**(7): p. 1728-36.
144. Visser, G.H.A., et al., *FIGO position paper: how to stop the caesarean section epidemic.* Lancet, 2018. **392**(10155): p. 1286-1287.
145. Sandall, J., et al., *Short-term and long-term effects of caesarean section on the health of women and children.* Lancet, 2018. **392**(10155): p. 1349-1357.
146. Renz-Polster, H., et al., *Caesarean section delivery and the risk of allergic disorders in childhood.* Clin Exp Allergy, 2005. **35**(11): p. 1466-72.
147. Sevelsted, A., et al., *Cesarean section and chronic immune disorders.* Pediatrics, 2015. **135**(1): p. e92-8.
148. Mitselou, N., et al., *Cesarean delivery, preterm birth, and risk of food allergy: Nationwide Swedish cohort study of more than 1 million children.* J Allergy Clin Immunol, 2018. **142**(5): p. 1510-1514.e2.
149. Feng, X.L., et al., *Factors influencing rising caesarean section rates in China between 1988 and 2008.* Bull World Health Organ, 2012. **90**(1): p. 30-9, 39a.
150. Osterman, M.J. and J.A. Martin, *Trends in low-risk cesarean delivery in the United States, 1990-2013.* Natl Vital Stat Rep, 2014. **63**(6): p. 1-16.
151. Almqvist, C., et al., *The impact of birth mode of delivery on childhood asthma and allergic diseases--a sibling study.* Clin Exp Allergy, 2012. **42**(9): p. 1369-76.
152. Black, M., et al., *Planned Repeat Cesarean Section at Term and Adverse Childhood Health Outcomes: A Record-Linkage Study.* PLOS Medicine, 2016. **13**(3): p. e1001973.
153. Stinson, L.F., M.S. Payne, and J.A. Keelan, *A Critical Review of the Bacterial Baptism Hypothesis and the Impact of Cesarean Delivery on the Infant Microbiome.* Front Med (Lausanne), 2018. **5**: p. 135.
154. Mueller, N.T., et al., *Does cesarean delivery impact infant weight gain and adiposity over the first year of life?* Int J Obes (Lond), 2019. **43**(8): p. 1549-1555.
155. Huh, S.Y., et al., *Delivery by caesarean section and risk of obesity in preschool age children: a prospective cohort study.* Arch Dis Child, 2012. **97**(7): p. 610-6.
156. Darmasseelane, K., et al., *Mode of delivery and offspring body mass index, overweight and obesity in adult life: a systematic review and meta-analysis.* PLoS One, 2014. **9**(2): p. e87896.
157. Li, H.T., Y.B. Zhou, and J.M. Liu, *The impact of cesarean section on offspring overweight and obesity: a systematic review and meta-analysis.* Int J Obes (Lond), 2013. **37**(7): p. 893-9.
158. Kuhle, S., O.S. Tong, and C.G. Woolcott, *Association between caesarean section and childhood obesity: a systematic review and meta-analysis.* Obes Rev, 2015. **16**(4): p. 295-303.
159. Mueller, N.T., et al., *Bacterial Baptism: Scientific, Medical, and Regulatory Issues Raised by Vaginal Seeding of C-Section-Born Babies.* J Law Med Ethics, 2019. **47**(4): p. 568-578.
160. Renz, H., et al., *An exposome perspective: Early-life events and immune development in a changing world.* J Allergy Clin Immunol, 2017. **140**(1): p. 24-40.
161. Rautava, S. and W.A. Walker, *Commensal bacteria and epithelial cross talk in the developing intestine.* Curr Gastroenterol Rep, 2007. **9**(5): p. 385-92.
162. Azad, M.B., et al., *Gut microbiota of healthy Canadian infants: profiles by mode of delivery and infant diet at 4 months.* Cmaj, 2013. **185**(5): p. 385-94.
163. Malamitsi-Puchner, A., et al., *The influence of the mode of delivery on circulating cytokine concentrations in the perinatal period.* Early Hum Dev, 2005. **81**(4): p. 387-92.
164. Black, M., et al., *Planned Cesarean Delivery at Term and Adverse Outcomes in Childhood Health.* Jama, 2015. **314**(21): p. 2271-9.
165. O'Neill, S.M., et al., *Birth by Caesarean Section and the Risk of Adult Psychosis: A Population-Based Cohort Study.* Schizophr Bull, 2016. **42**(3): p. 633-41.
166. Mårild, K., et al., *Pregnancy outcome and risk of celiac disease in offspring: a nationwide case-control study.* Gastroenterology, 2012. **142**(1): p. 39-45.e3.
167. Caughey, A.B., et al., *Safe prevention of the primary cesarean delivery.* Am J Obstet Gynecol, 2014. **210**(3): p. 179-93.
168. Barber, E.L., et al., *Indications contributing to the increasing cesarean delivery rate.* Obstet Gynecol, 2011. **118**(1): p. 29-38.

169. Hannah, M.E., *Planned elective cesarean section: a reasonable choice for some women?* Cmaj, 2004. **170**(5): p. 813-4.
170. Loke, A.Y., L. Davies, and Y.W. Mak, *Is it the decision of women to choose a cesarean section as the mode of birth? A review of literature on the views of stakeholders.* BMC Pregnancy Childbirth, 2019. **19**(1): p. 286.
171. Gao, Z., et al., *Molecular analysis of human forearm superficial skin bacterial biota.* Proc Natl Acad Sci U S A, 2007. **104**(8): p. 2927-32.
172. Yassour, M., et al., *Natural history of the infant gut microbiome and impact of antibiotic treatment on bacterial strain diversity and stability.* Sci Transl Med, 2016. **8**(343): p. 343ra81.
173. Bokulich, N.A., et al., *Antibiotics, birth mode, and diet shape microbiome maturation during early life.* Sci Transl Med, 2016. **8**(343): p. 343ra82.
174. Yatsunenko, T., et al., *Human gut microbiome viewed across age and geography.* Nature, 2012. **486**(7402): p. 222-227.
175. Costello, E.K., et al., *Microbiome assembly across multiple body sites in low-birthweight infants.* mBio, 2013. **4**(6): p. e00782-13.
176. Luna, P.C., *Skin Microbiome as Years Go By.* Am J Clin Dermatol, 2020. **21**(Suppl 1): p. 12-17.
177. Dimitriu, P.A., et al., *New Insights into the Intrinsic and Extrinsic Factors That Shape the Human Skin Microbiome.* mBio, 2019. **10**(4).
178. Lehtimäki, J., et al., *Patterns in the skin microbiota differ in children and teenagers between rural and urban environments.* Sci Rep, 2017. **7**: p. 45651.
179. Shibagaki, N., et al., *Aging-related changes in the diversity of women's skin microbiomes associated with oral bacteria.* Sci Rep, 2017. **7**(1): p. 10567.
180. Dréno, B., et al., *Cutibacterium acnes (Propionibacterium acnes) and acne vulgaris: a brief look at the latest updates.* J Eur Acad Dermatol Venereol, 2018. **32 Suppl 2**: p. 5-14.
181. Dréno, B., et al., *Microbiome in healthy skin, update for dermatologists.* J Eur Acad Dermatol Venereol, 2016. **30**(12): p. 2038-2047.
182. Cox, L.M., et al., *Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences.* Cell, 2014. **158**(4): p. 705-721.
183. Nogacka, A., et al., *Impact of intrapartum antimicrobial prophylaxis upon the intestinal microbiota and the prevalence of antibiotic resistance genes in vaginally delivered full-term neonates.* Microbiome, 2017. **5**(1): p. 93.
184. Azad, M.B., et al., *Impact of maternal intrapartum antibiotics, method of birth and breastfeeding on gut microbiota during the first year of life: a prospective cohort study.* Bjog, 2016. **123**(6): p. 983-93.
185. Mazzola, G., et al., *Early Gut Microbiota Perturbations Following Intrapartum Antibiotic Prophylaxis to Prevent Group B Streptococcal Disease.* PLOS ONE, 2016. **11**(6): p. e0157527.
186. Aloisio, I., et al., *Influence of intrapartum antibiotic prophylaxis against group B Streptococcus on the early newborn gut composition and evaluation of the anti-Streptococcus activity of Bifidobacterium strains.* Appl Microbiol Biotechnol, 2014. **98**(13): p. 6051-60.
187. Fallani, M., et al., *Intestinal microbiota of 6-week-old infants across Europe: geographic influence beyond delivery mode, breast-feeding, and antibiotics.* J Pediatr Gastroenterol Nutr, 2010. **51**(1): p. 77-84.
188. Fouhy, F., et al., *High-throughput sequencing reveals the incomplete, short-term recovery of infant gut microbiota following parenteral antibiotic treatment with ampicillin and gentamicin.* Antimicrob Agents Chemother, 2012. **56**(11): p. 5811-20.
189. Pammi, M., et al., *Intestinal dysbiosis in preterm infants preceding necrotizing enterocolitis: a systematic review and meta-analysis.* Microbiome, 2017. **5**(1): p. 31.
190. Zhang, M., et al., *Oral antibiotic treatment induces skin microbiota dysbiosis and influences wound healing.* Microb Ecol, 2015. **69**(2): p. 415-21.
191. Dellacecca, E.R., et al., *Antibiotics Drive Microbial Imbalance and Vitiligo Development in Mice.* J Invest Dermatol, 2020. **140**(3): p. 676-687.e6.
192. Prince, T., A.J. McBain, and C.A. O'Neill, *Lactobacillus reuteri protects epidermal keratinocytes from Staphylococcus aureus-induced cell death by competitive exclusion.* Appl Environ Microbiol, 2012. **78**(15): p. 5119-26.
193. Mohammedsaeed, W., et al., *Lactobacillus rhamnosus GG inhibits the toxic effects of Staphylococcus aureus on epidermal keratinocytes.* Appl Environ Microbiol, 2014. **80**(18): p. 5773-81.
194. Panduru, M., et al., *Probiotics and primary prevention of atopic dermatitis: a meta-analysis of randomized controlled studies.* J Eur Acad Dermatol Venereol, 2015. **29**(2): p. 232-242.
195. Baquerizo Nole, K.L., E. Yim, and J.E. Keri, *Probiotics and prebiotics in dermatology.* J Am Acad Dermatol, 2014. **71**(4): p. 814-21.

196. Lapin, B., et al., *Relationship between prenatal antibiotic use and asthma in at-risk children.* Ann Allergy Asthma Immunol, 2015. **114**(3): p. 203-7.
197. Metsälä, J., et al., *Prenatal and post-natal exposure to antibiotics and risk of asthma in childhood.* Clin Exp Allergy, 2015. **45**(1): p. 137-45.
198. Örtqvist, A.K., et al., *Antibiotics in fetal and early life and subsequent childhood asthma: nationwide population based study with sibling analysis.* BMJ : British Medical Journal, 2014. **349**: p. g6979.
199. Stensballe, L.G., et al., *Use of antibiotics during pregnancy increases the risk of asthma in early childhood.* J Pediatr, 2013. **162**(4): p. 832-838.e3.
200. McKeever, T.M., et al., *The importance of prenatal exposures on the development of allergic disease: a birth cohort study using the West Midlands General Practice Database.* Am J Respir Crit Care Med, 2002. **166**(6): p. 827-32.
201. Wohl, D.L., et al., *Intrapartum antibiotics and childhood atopic dermatitis.* J Am Board Fam Med, 2015. **28**(1): p. 82-9.
202. Mueller, N.T., et al., *Prenatal exposure to antibiotics, cesarean section and risk of childhood obesity.* Int J Obes (Lond), 2015. **39**(4): p. 665-70.
203. Dinić, M., et al., *Postbiotic lactobacilli induce cutaneous antimicrobial response and restore the barrier to inhibit the intracellular invasion of Staphylococcus aureus in vitro and ex vivo.* The FASEB Journal, 2024. **38**(14): p. e23801.
204. Sinha, T., et al., *Maternal antibiotic prophylaxis during cesarean section has a limited impact on the infant gut microbiome.* Cell Host & Microbe, 2024. **32**(8): p. 1444-1454.e6.
205. Haahr, T., et al., *Vaginal seeding or vaginal microbial transfer from the mother to the caesarean-born neonate: a commentary regarding clinical management.* BJOG: An International Journal of Obstetrics & Gynaecology, 2018. **125**(5): p. 533-536.
206. Practice, C.o.O., *Committee Opinion No. 725: Vaginal Seeding.* Obstet Gynecol, 2017. **130**(5): p. e274-e278.
207. Zhou, L., et al., *Effects of vaginal microbiota transfer on the neurodevelopment and microbiome of cesarean-born infants: A blinded randomized controlled trial.* Cell Host & Microbe, 2023. **31**(7): p. 1232-1247.e5.
208. Dominguez-Bello, M.G., et al., *Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer.* Nat Med, 2016. **22**(3): p. 250-3.
209. Mueller, N.T., et al., *Maternal Bacterial Engraftment in Multiple Body Sites of Cesarean Section Born Neonates after Vaginal Seeding—a Randomized Controlled Trial.* mBio, 2023. **14**(3): p. e00491-23.
210. Liu, Y., et al., *Effects of vaginal seeding on gut microbiota, body mass index, and allergy risks in infants born through cesarean delivery: a randomized clinical trial.* Am J Obstet Gynecol MFM, 2023. **5**(1): p. 100793.
211. Wilson, B.C., et al., *Oral administration of maternal vaginal microbes at birth to restore gut microbiome development in infants born by caesarean section: A pilot randomised placebo-controlled trial.* eBioMedicine, 2021. **69**.
212. Huynh, J., P. Palasanthiran, and B. McMullan, *Potential Transmission of Herpes Simplex Virus Via Vaginal Seeding.* Pediatr Infect Dis J, 2018. **37**(11): p. e278.
213. Lokugamage, A.U. and S.D.C. Pathberiya, *The microbiome seeding debate – let's frame it around women-centred care.* Reproductive Health, 2019. **16**(1): p. 91.
214. Olsen, R., et al., *Prophylactic Probiotics for Preterm Infants: A Systematic Review and Meta-Analysis of Observational Studies.* Neonatology, 2016. **109**(2): p. 105-12.
215. Martín, R. and P. Langella, *Emerging Health Concepts in the Probiotics Field: Streamlining the Definitions.* Front Microbiol, 2019. **10**: p. 1047.
216. Lolou, V. and M.I. Panayiotidis, *Functional Role of Probiotics and Prebiotics on Skin Health and Disease.* Fermentation, 2019. **5**(2): p. 41.
217. Knackstedt, R., T. Knackstedt, and J. Gatherwright, *The role of topical probiotics in skin conditions: A systematic review of animal and human studies and implications for future therapies.* Exp Dermatol, 2020. **29**(1): p. 15-21.
218. Kalliomäki, M., et al., *Probiotics during the first 7 years of life: a cumulative risk reduction of eczema in a randomized, placebo-controlled trial.* J Allergy Clin Immunol, 2007. **119**(4): p. 1019-21.
219. Rosenfeldt, V., et al., *Effect of probiotics on gastrointestinal symptoms and small intestinal permeability in children with atopic dermatitis.* J Pediatr, 2004. **145**(5): p. 612-6.
220. Jung, G.W., et al., *Prospective, randomized, open-label trial comparing the safety, efficacy, and tolerability of an acne treatment regimen with and without a probiotic supplement and minocycline in subjects with mild to moderate acne.* J Cutan Med Surg, 2013. **17**(2): p. 114-22.
221. Nakatsuji, T., et al., *Antimicrobials from human skin commensal bacteria protect against Staphylococcus aureus and are deficient in atopic dermatitis.* Sci Transl Med, 2017. **9**(378).

222. Myles, I.A., et al., *First-in-human topical microbiome transplantation with Roseomonas mucosa for atopic dermatitis.* JCI Insight, 2018. **3**(9).
223. Hendricks, A.J., B.W. Mills, and V.Y. Shi, *Skin bacterial transplant in atopic dermatitis: Knowns, unknowns and emerging trends.* J Dermatol Sci, 2019. **95**(2): p. 56-61.
224. Locker, J., et al., *Microbiological insights and dermatological applications of live biotherapeutic products.* Journal of Applied Microbiology, 2024. **135**(8).
225. Arboleya, S., et al., *Establishment and development of intestinal microbiota in preterm neonates.* FEMS Microbiol Ecol, 2012. **79**(3): p. 763-72.
226. Hill, C.J., et al., *Evolution of gut microbiota composition from birth to 24 weeks in the INFANTMET Cohort.* Microbiome, 2017. **5**(1): p. 4.
227. La Rosa, P.S., et al., *Patterned progression of bacterial populations in the premature infant gut.* Proc Natl Acad Sci U S A, 2014. **111**(34): p. 12522-7.
228. Delnord, M., et al., *Varying gestational age patterns in cesarean delivery: an international comparison.* BMC Pregnancy Childbirth, 2014. **14**: p. 321.
229. Patel, K., et al., *Trends and determinants of gastric bacterial colonization of preterm neonates in a NICU setting.* PLOS ONE, 2015. **10**(7): p. e0114664.
230. Outcomes, I.o.M.U.C.o.U.P.B.a.A.H., *The National Academies Collection: Reports funded by National Institutes of Health*, in *Preterm Birth: Causes, Consequences, and Prevention*, R.E. Behrman and A.S. Butler, Editors. 2007, National Academies Press (US) Copyright © 2007, National Academy of Sciences: Washington (DC).
231. Moster, D., R.T. Lie, and T. Markestad, *Long-term medical and social consequences of preterm birth.* N Engl J Med, 2008. **359**(3): p. 262-73.
232. Marchant, E.A., et al., *Attenuated innate immune defenses in very premature neonates during the neonatal period.* Pediatric Research, 2015. **78**(5): p. 492-497.
233. Melville, J.M. and T.J. Moss, *The immune consequences of preterm birth.* Front Neurosci, 2013. **7**: p. 79.
234. Hardman, M.J., et al., *Barrier formation in the human fetus is patterned.* J Invest Dermatol, 1999. **113**(6): p. 1106-13.
235. Narendran, V., et al., *Biomarkers of Epidermal Innate Immunity in Premature and Full-Term Infants.* Pediatric Research, 2010. **67**(4): p. 382-386.
236. Walker, V.P., et al., *Host defense proteins on the surface of neonatal skin: implications for innate immunity.* J Pediatr, 2008. **152**(6): p. 777-81.
237. Chiou, Y.B. and U. Blume-Peytavi, *Stratum Corneum Maturation.* Skin Pharmacology and Physiology, 2004. **17**(2): p. 57-66.
238. Rissmann, R., et al., *New insights into ultrastructure, lipid composition and organization of vernix caseosa.* J Invest Dermatol, 2006. **126**(8): p. 1823-33.
239. Yoshio, H., et al., *Antimicrobial Polypeptides of Human Vernix Caseosa and Amniotic Fluid: Implications for Newborn Innate Defense.* Pediatric Research, 2003. **53**(2): p. 211-216.
240. Tollin, M., et al., *Proteome analysis of vernix caseosa.* Pediatr Res, 2006. **60**(4): p. 430-4.
241. Visscher, M.O., et al., *Newborn infant skin: physiology, development, and care.* Clin Dermatol, 2015. **33**(3): p. 271-80.
242. Visscher, M.O., et al., *Vernix caseosa in neonatal adaptation.* J Perinatol, 2005. **25**(7): p. 440-6.
243. Marchini, G., et al., *The newborn infant is protected by an innate antimicrobial barrier: peptide antibiotics are present in the skin and vernix caseosa.* British Journal of Dermatology, 2002. **147**(6): p. 1127-1134.
244. Akinbi, H.T., et al., *Host defense proteins in vernix caseosa and amniotic fluid.* Am J Obstet Gynecol, 2004. **191**(6): p. 2090-6.
245. Tollin, M., et al., *Vernix caseosa as a multi-component defence system based on polypeptides, lipids and their interactions.* Cellular and Molecular Life Sciences CMLS, 2005. **62**(19): p. 2390-2399.
246. Baker, S.M., N.N. Balo, and F.T. Abdel Aziz, *Is vernix caseosa a protective material to the newborn? A biochemical approach.* Indian J Pediatr, 1995. **62**(2): p. 237-9.
247. Stamatas, G.N., et al., *Infant skin microstructure assessed in vivo differs from adult skin in organization and at the cellular level.* Pediatr Dermatol, 2010. **27**(2): p. 125-31.
248. Stoll, B.J., et al., *Late-onset sepsis in very low birth weight neonates: the experience of the NICHD Neonatal Research Network.* Pediatrics, 2002. **110**(2 Pt 1): p. 285-91.
249. Round, J.L., R.M. O'Connell, and S.K. Mazmanian, *Coordination of tolerogenic immune responses by the commensal microbiota.* J Autoimmun, 2010. **34**(3): p. J220-5.
250. Naik, S., et al., *Compartmentalized control of skin immunity by resident commensals.* Science, 2012. **337**(6098): p. 1115-9.
251. von Mutius, E., *The shape of the microbiome in early life.* Nat Med, 2017. **23**(3): p. 274-275.

252. Lax, S., et al., *Longitudinal analysis of microbial interaction between humans and the indoor environment.* Science, 2014. **345**(6200): p. 1048-52.
253. Barberán, A., et al., *The ecology of microscopic life in household dust.* Proc Biol Sci, 2015. **282**(1814).
254. Prescott, S.L. and A.C. Logan, *Transforming Life: A Broad View of the Developmental Origins of Health and Disease Concept from an Ecological Justice Perspective.* International Journal of Environmental Research and Public Health, 2016. **13**(11): p. 1075.
255. van Mierlo, M.M.F., et al., *The influence of treatment in alpine and moderate maritime climate on the composition of the skin microbiome in patients with difficult to treat atopic dermatitis.* Clin Exp Allergy, 2019. **49**(11): p. 1437-1445.
256. Dominguez-Bello, M.G., et al., *Role of the microbiome in human development.* Gut, 2019. **68**(6): p. 1108-1114.
257. Bach, J.F., *Infections and autoimmune diseases.* J Autoimmun, 2005. **25 Suppl**: p. 74-80.
258. Semenkovich, C.F., et al., *American Diabetes Association and JDRF Research Symposium: Diabetes and the Microbiome.* Diabetes, 2015. **64**(12): p. 3967-77.
259. Blustein, J. and J. Liu, *Time to consider the risks of caesarean delivery for long term child health.* BMJ : British Medical Journal, 2015. **350**: p. h2410.
260. Nodake, Y., et al., *Pilot study on novel skin care method by augmentation with Staphylococcus epidermidis, an autologous skin microbe--A blinded randomized clinical trial.* J Dermatol Sci, 2015. **79**(2): p. 119-26.
261. Lopes, E.G., et al., *Topical application of probiotics in skin: adhesion, antimicrobial and antibiofilm in vitro assays.* J Appl Microbiol, 2017. **122**(2): p. 450-461.
262. Shokryazdan, P., et al., *Probiotic potential of Lactobacillus strains with antimicrobial activity against some human pathogenic strains.* Biomed Res Int, 2014. **2014**: p. 927268.
263. Khmaladze, I., et al., *Lactobacillus reuteri DSM 17938-A comparative study on the effect of probiotics and lysates on human skin.* Exp Dermatol, 2019. **28**(7): p. 822-828.
264. Paetzold, B., et al., *Skin microbiome modulation induced by probiotic solutions.* Microbiome, 2019. **7**(1): p. 95.
265. Nakatsuji, T., et al., *426 Clinical improvement in atopic dermatitis following autologous application of microbiome therapy targeting Staphylococcus aureus.* Journal of Investigative Dermatology, 2018. **138**(5): p. S72.
266. Rogier, E.W., et al., *Secretory antibodies in breast milk promote long-term intestinal homeostasis by regulating the gut microbiota and host gene expression.* Proc Natl Acad Sci U S A, 2014. **111**(8): p. 3074-9.
267. Schwartz, S., et al., *A metagenomic study of diet-dependent interaction between gut microbiota and host in infants reveals differences in immune response.* Genome Biol, 2012. **13**(4): p. r32.
268. Praveen, P., et al., *The role of breast-feeding in infant immune system: a systems perspective on the intestinal microbiome.* Microbiome, 2015. **3**(1): p. 41.
269. WHO, *WHO recommendations for care of the preterm or low birth weight infant.* 2022, Geneva: World Health Organization: Licence: CC BY-NC-SA 3.0 IGO.
270. Evans, K.C., et al., *Effect of caesarean section on breast milk transfer to the normal term newborn over the first week of life.* Arch Dis Child Fetal Neonatal Ed, 2003. **88**(5): p. F380-2.
271. Dewey, K.G., et al., *Risk factors for suboptimal infant breastfeeding behavior, delayed onset of lactation, and excess neonatal weight loss.* Pediatrics, 2003. **112**(3 Pt 1): p. 607-19.
272. Bai, D.L., K.M. Wu, and M. Tarrant, *Association between intrapartum interventions and breastfeeding duration.* J Midwifery Womens Health, 2013. **58**(1): p. 25-32.
273. Gaitanis, G., et al., *Variation of cultured skin microbiota in mothers and their infants during the first year postpartum.* Pediatric Dermatology, 2019. **36**(4): p. 460-465.
274. Cordain, L., *Implications for the role of diet in acne.* Semin Cutan Med Surg, 2005. **24**(2): p. 84-91.
275. Bowe, W.P., S.S. Joshi, and A.R. Shalita, *Diet and acne.* J Am Acad Dermatol, 2010. **63**(1): p. 124-41.
276. Clemente, J.C., et al., *The microbiome of uncontacted Amerindians.* Sci Adv, 2015. **1**(3).
277. Moestrup, K.S., et al., *Dietary Control of Skin Lipid Composition and Microbiome.* J Invest Dermatol, 2018. **138**(5): p. 1225-1228.
278. Nguyen, T.L., et al., *How informative is the mouse for human gut microbiota research?* Dis Model Mech, 2015. **8**(1): p. 1-16.
279. Summerfield, A., F. Meurens, and M.E. Ricklin, *The immunology of the porcine skin and its value as a model for human skin.* Molecular Immunology, 2015. **66**(1): p. 14-21.
280. Beckers, K.F., J.P. Flanagan, and J.L. Sones, *Microbiome and pregnancy: focus on microbial dysbiosis coupled with maternal obesity.* International Journal of Obesity, 2024. **48**(4): p. 439-448.
281. Tun, H.M., et al., *Roles of Birth Mode and Infant Gut Microbiota in Intergenerational Transmission of Overweight and Obesity From Mother to Offspring.* JAMA Pediatr, 2018. **172**(4): p. 368-377.

282. Marchi, J., et al., *Risks associated with obesity in pregnancy, for the mother and baby: a systematic review of reviews.* Obes Rev, 2015. **16**(8): p. 621-38.
283. Carpenter, J.R., *Intrapartum Management of the Obese Gravida.* Clin Obstet Gynecol, 2016. **59**(1): p. 172-9.
284. Cabrera-Rubio, R., et al., *The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery.* Am J Clin Nutr, 2012. **96**(3): p. 544-51.
285. Si, J., et al., *Genetic associations and shared environmental effects on the skin microbiome of Korean twins.* BMC Genomics, 2015. **16**: p. 992.
286. Turnbaugh, P.J., et al., *A core gut microbiome in obese and lean twins.* Nature, 2009. **457**(7228): p. 480-4.
287. Ying, S., et al., *The Influence of Age and Gender on Skin-Associated Microbial Communities in Urban and Rural Human Populations.* PLoS One, 2015. **10**(10): p. e0141842.
288. Jo, J.-H., E.A. Kennedy, and H.H. Kong, *Research Techniques Made Simple: Bacterial 16S Ribosomal RNA Gene Sequencing in Cutaneous Research.* Journal of Investigative Dermatology, 2016. **136**(3): p. e23-e27.
289. Wang, W.L., et al., *Application of metagenomics in the human gut microbiome.* World J Gastroenterol, 2015. **21**(3): p. 803-14.
290. Rosenthal, M., et al., *Skin microbiota: microbial community structure and its potential association with health and disease.* Infect Genet Evol, 2011. **11**(5): p. 839-48.
291. Goodrich, J.K., et al., *Conducting a microbiome study.* Cell, 2014. **158**(2): p. 250-262.
292. Mardis, E.R., *Next-generation sequencing platforms.* Annu Rev Anal Chem (Palo Alto Calif), 2013. **6**: p. 287-303.
293. Stinson, L.F., J.A. Keelan, and M.S. Payne, *Identification and removal of contaminating microbial DNA from PCR reagents: impact on low-biomass microbiome analyses.* Lett Appl Microbiol, 2019. **68**(1): p. 2-8.
294. Santiago-Rodriguez, T.M., et al., *The Skin Microbiome: Current Techniques, Challenges, and Future Directions.* Microorganisms, 2023. **11**(5).
295. Kong, H.H., et al., *Performing Skin Microbiome Research: A Method to the Madness.* J Invest Dermatol, 2017. **137**(3): p. 561-568.
296. Grogan, M.D., et al., *Research Techniques Made Simple: Profiling the Skin Microbiota.* J Invest Dermatol, 2019. **139**(4): p. 747-752.e1.
297. Tanner, M.A., et al., *Specific ribosomal DNA sequences from diverse environmental settings correlate with experimental contaminants.* Appl Environ Microbiol, 1998. **64**(8): p. 3110-3.
298. Kim, D., et al., *Optimizing methods and dodging pitfalls in microbiome research.* Microbiome, 2017. **5**(1): p. 52.
299. Islam Sajib, M.S., et al., *Advances in Host Depletion and Pathogen Enrichment Methods for Rapid Sequencing–Based Diagnosis of Bloodstream Infection.* The Journal of Molecular Diagnostics, 2024. **26**(9): p. 741-753.
300. Marquet, M., et al., *Evaluation of microbiome enrichment and host DNA depletion in human vaginal samples using Oxford Nanopore's adaptive sequencing.* Scientific Reports, 2022. **12**(1): p. 4000.
301. Hammond, M., et al., *Picodroplet partitioned whole genome amplification of low biomass samples preserves genomic diversity for metagenomic analysis.* Microbiome, 2016. **4**(1): p. 52.
302. Afshinnekoo, E., et al., *Geospatial Resolution of Human and Bacterial Diversity with City-Scale Metagenomics.* Cell Systems, 2015. **1**(1): p. 72-87.
303. Poore, G.D., et al., *RETRACTED ARTICLE: Microbiome analyses of blood and tissues suggest cancerdiagnostic approach.* Nature, 2020. **579**(7800): p. 567-574.
304. Kuczynski, J., et al., *Experimental and analytical tools for studying the human microbiome.* Nature Reviews Genetics, 2012. **13**(1): p. 47-58.
305. Ranjan, R., et al., *Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing.* Biochem Biophys Res Commun, 2016. **469**(4): p. 967-77.
306. Liu, J., et al., *Comparison of ITS and 18S rDNA for estimating fungal diversity using PCR-DGGE.* World J Microbiol Biotechnol, 2015. **31**(9): p. 1387-95.
307. Chakravorty, S., et al., *A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria.* J Microbiol Methods, 2007. **69**(2): p. 330-9.
308. Bukin, Y.S., et al., *The effect of 16S rRNA region choice on bacterial community metabarcoding results.* Scientific Data, 2019. **6**(1): p. 190007.
309. Maina, S., L. Zheng, and B.C. Rodoni, *Targeted Genome Sequencing (TG-Seq) Approaches to Detect Plant Viruses.* Viruses, 2021. **13**(4).
310. Houldcroft, C.J., M.A. Beale, and J. Breuer, *Clinical and biological insights from viral genome sequencing.* Nature Reviews Microbiology, 2017. **15**(3): p. 183-192.

311. Pearman, W.S., N.E. Freed, and O.K. Silander, *Testing the advantages and disadvantages of short- and long- read eukaryotic metagenomics using simulated reads.* BMC Bioinformatics, 2020. **21**(1): p. 220.
312. Allaband, C., et al., *Microbiome 101: Studying, Analyzing, and Interpreting Gut Microbiome Data for Clinicians.* Clinical Gastroenterology and Hepatology, 2019. **17**(2): p. 218-230.
313. Chng, K.R., et al. *Whole metagenome profiling reveals skin microbiome-dependent susceptibility to atopic dermatitis flare*. Nature microbiology, 2016. **1**, 16106 DOI: 10.1038/nmicrobiol.2016.106.
314. Oh, J., et al., *Temporal Stability of the Human Skin Microbiome.* Cell, 2016. **165**(4): p. 854-66.
315. Schloss, P.D. and J. Handelsman, *Metagenomics for studying unculturable microorganisms: cutting the Gordian knot.* Genome Biology, 2005. **6**(8): p. 229.
316. Keeling, P.J., et al., *The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing.* PLOS Biology, 2014. **12**(6): p. e1001889.
317. Liu, Y.-X., et al., *A practical guide to amplicon and metagenomic analysis of microbiome data.* Protein & Cell, 2020. **12**(5): p. 315-330.
318. Amarasinghe, S.L., et al., *Opportunities and challenges in long-read sequencing data analysis.* Genome Biology, 2020. **21**(1): p. 30.
319. Illumina. *Understanding the NGS workflow*. . 2023 [cited 2023 February 2023]; Available from: https://emea.illumina.com/science/technology/next-generation-sequencing/beginners/ngs-workflow.html.
320. Jain, M., et al., *The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community.* Genome Biology, 2016. **17**(1): p. 239.
321. Rang, F.J., W.P. Kloosterman, and J. de Ridder, *From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy.* Genome Biology, 2018. **19**(1): p. 90.
322. Merker, J.D., et al., *Long-read genome sequencing identifies causal structural variation in a Mendelian disease.* Genetics in medicine : official journal of the American College of Medical Genetics, 2018. **20**(1): p. 159-163.
323. Roberts, R.J., M.O. Carneiro, and M.C. Schatz, *The advantages of SMRT sequencing.* Genome Biology, 2013. **14**(6): p. 405.
324. Stevens, B.M., et al., *Comparison of Oxford Nanopore Technologies and Illumina MiSeq sequencing with mock communities and agricultural soil.* Scientific Reports, 2023. **13**(1): p. 9323.
325. Payne, A., et al., *BulkVis: a graphical viewer for Oxford nanopore bulk FAST5 files.* Bioinformatics, 2019. **35**(13): p. 2193-2198.
326. Logsdon, G.A., M.R. Vollger, and E.E. Eichler, *Long-read human genome sequencing and its applications.* Nature Reviews Genetics, 2020. **21**(10): p. 597-614.
327. Singleton, C.M., et al., *Connecting structure to function with the recovery of over 1000 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing.* Nature Communications, 2021. **12**(1): p. 2009.
328. Nanopore. *Nanopore DNA sequencing.* 2023 [cited 2023 February 2023]; Available from: https://nanoporetech.com/applications/dna-nanopore-sequencing.
329. Aguiar-Pulido, V., et al., *Metagenomics, Metatranscriptomics, and Metabolomics Approaches for Microbiome Analysis.* Evol Bioinform Online, 2016. **12**(Suppl 1): p. 5-16.
330. Acosta, E.M., et al., *Bacterial DNA on the skin surface overrepresents the viable skin microbiome*. 2023, eLife Sciences Publications, Ltd.
331. Bjerre, R.D., et al., *Effects of sampling strategy and DNA extraction on human skin microbiome investigations.* Scientific Reports, 2019. **9**(1): p. 17287.
332. Fleming, E., et al., *Cultivation of common bacterial species and strains from human skin, oral, and gut microbiota.* BMC Microbiology, 2021. **21**(1): p. 278.
333. Nsereko, E., et al., *Consideration of Cultural Practices When Characterizing the Vaginal Microbiota Among African and African American Women.* Biol Res Nurs, 2021. **23**(1): p. 91-99.
334. Phillips, S., et al., *The Pregnancy and EARly Life study (PEARL) - a longitudinal study to understand how gut microbes contribute to maintaining health during pregnancy and early life.* BMC Pediatrics, 2021. **21**(1): p. 357.
335. ATCC. *ATCC Skin Microbiome Whole Cell Mix*. 2022; Available from: https://www.atcc.org/products/msa-2005.
336. Sethi, D.K., et al., *Chlorhexidine gluconate usage is associated with antiseptic tolerance in staphylococci from the neonatal intensive care unit.* JAC Antimicrob Resist, 2021. **3**(4): p. dlab173.
337. Matthews, T.C., et al., *The Integrated Rapid Infectious Disease Analysis (IRIDA) Platform.* bioRxiv, 2018: p. 381830.

338. Afgan, E., et al., *The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update.* Nucleic Acids Research, 2018. **46**(W1): p. W537-W544.
339. Le Viet T., A.N.F.a.P.A., *Infrastructure for Healthier Living – Galaxy at Quadram Institute Bioscience [version 1; not peer reviewed].* F1000Research, 2019. **8**(1663 (slides)).
340. Wood, D.E., J. Lu, and B. Langmead, *Improved metagenomic analysis with Kraken 2.* Genome Biology, 2019. **20**(1): p. 257.
341. Wood, D., *Custom Database Manual*. 2019.
342. Chen, S., et al., *fastp: an ultra-fast all-in-one FASTQ preprocessor.* Bioinformatics, 2018. **34**(17): p. i884-i890.
343. Andrews, S. *FastQC: A Quality Control Tool for High Throughput Sequence Data.* . 2010 [cited 2023 January 2023]; Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.
344. Jennifer Lu, F.P.B., Peter Thielen, Steven L. Salzberg, *Bracken: estimating species abundance in metagenomics data.* PeerJ Computer Science, 2017. **3**: p. e104.
345. Huttenhower. *KneadData*. [cited 2023 April]; Available from: https://huttenhower.sph.harvard.edu/kneaddata/.
346. Li, D., et al., *MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph.* Bioinformatics, 2015. **31**(10): p. 1674-1676.
347. Uritskiy, G.V., J. DiRuggiero, and J. Taylor, *MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis.* Microbiome, 2018. **6**(1): p. 158.
348. Dongwan D. Kang, J.F., Rob Egan, Zhong Wang, *MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities.* . PeerJ, 2015. **3**: p. e1165.
349. Wu, Y.-W., B.A. Simmons, and S.W. Singer, *MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets.* Bioinformatics, 2015. **32**(4): p. 605-607.
350. Alneberg, J., et al., *CONCOCT: Clustering cONtigs on COverage and ComposiTion.* arXiv: Genomics, 2013.
351. Parks, D.H., et al., *CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes.* Genome Res, 2015. **25**(7): p. 1043-55.
352. Chaumeil, P.-A., et al., *GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database.* Bioinformatics, 2019. **36**(6): p. 1925-1927.
353. Telatin, A., P. Fariselli, and G. Birolo, *SeqFu: A Suite of Utilities for the Robust and Reproducible Manipulation of Sequence Files.* Bioengineering, 2021. **8**(5): p. 59.
354. Telatin, A., *Host removal database: Homo sapiens, Sars-Cov-2, PhiX174 (1.1) [Data set]*. 2022, Zenodo: Zenodo.
355. Handley, S.A., *Virus+ Sequence Masked Human Reference Genome (hg19) (1.0) [Data set].* 2020: Zenodo.
356. Lu, J., et al., *Metagenome analysis using the Kraken software suite.* Nature Protocols, 2022. **17**(12): p. 2815-2839.
357. Kieser, S., et al., *ATLAS: a Snakemake workflow for assembly, annotation, and genomic binning of metagenome sequence data.* BMC Bioinformatics, 2020. **21**(1): p. 257.
358. Chibani, C.M., et al., *A catalogue of 1,167 genomes from the human gut archaeome.* Nature Microbiology, 2022. **7**(1): p. 48-61.
359. Beghini, F., et al., *Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3.* eLife, 2021. **10**: p. e65088.
360. Blanco-Míguez, A., et al., *Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4.* Nature Biotechnology, 2023. **41**(11): p. 1633-1644.
361. Valles-Colomer, M., et al., *The person-to-person transmission landscape of the gut and oral microbiomes.* Nature, 2023. **614**(7946): p. 125-135.
362. Youden, W.J., *Index for rating diagnostic tests.* Cancer, 1950. **3**(1): p. 32-35.
363. Wang, Y., et al., *Nanopore sequencing technology, bioinformatics and applications.* Nature Biotechnology, 2021. **39**(11): p. 1348-1365.
364. de Goffau, M.C., et al., *Recognizing the reagent microbiome.* Nature Microbiology, 2018. **3**(8): p. 851-853.
365. Bey, B.S., et al., *Extraction of high molecular weight DNA from microbial mats.* BioTechniques, 2010. **49**(3): p. 631-640.
366. Paraj V. Mandrekar, Z.M., Steve Krueger and Cristopher Cowan. *High-Concentration (>100ng/µl) Genomic DNA From Whole Blood Using the Maxwell® 16 Low Elution Volume Instrument.* 2007 2010 [cited 2020 November]; Available from: https://www.promega.co.uk/resources/pubhub/high-concentration-genomic-dna-from-whole-blood-using-the-maxwell-16-low-elution-volume-instrument.

367. Sui, H.-y., et al., *Impact of DNA Extraction Method on Variation in Human and Built Environment Microbial Community and Functional Profiles Assessed by Shotgun Metagenomics Sequencing.* Frontiers in Microbiology, 2020. **11**.
368. Moeller, J.R., et al., *Paramagnetic cellulose DNA isolation improves DNA yield and quality among diverse plant taxa.* Applications in Plant Sciences, 2014. **2**(10): p. 1400048.
369. Dunbar, J., et al., *Evaluation of DNA extraction methods to detect bacterial targets in aerosol samples.* Journal of Microbiological Methods, 2018. **153**: p. 48-53.
370. Maxwell®, P. *Promega Maxwell® RSC Blood DNA Kit*. 2020 [cited 2020 November 2020]; Available from: https://www.promega.com/-/media/files/resources/protocols/technical-manuals/101/maxwell-rsc-blood-dna-kit-protocol.pdf?la=en.
371. Liao, C.-H. and L.M. Shollenberger, *Survivability and long-term preservation of bacteria in water and in phosphate-buffered saline*.* Letters in Applied Microbiology, 2003. **37**(1): p. 45-50.
372. Kelman, A., *Factors influencing viability and variation in cultures of Pseudomonas solanacearum. .* Phytopathology, 1956. **46**: p. 16–17.
373. Wakimoto, S., Utatsu, I., Masuo, N. and Hayashi, N., *Multiplication of Pseudomonas solanacearum in pure water.* Annals of the Phytopathological Society of Japan, 1982. **48**: p. 620-627.
374. van Elsas, J.D., et al., *Effects of ecological factors on the survival and physiology of Ralstonia solanacearum bv. 2 in irrigation water.* Can J Microbiol, 2001. **47**(9): p. 842-54.
375. Banning, N., S. Toze, and B.J. Mee, *Escherichia coli survival in groundwater and effluent measured using a combination of propidium iodide and the green fluorescent protein.* Journal of Applied Microbiology, 2002. **93**(1): p. 69-76.
376. Downey, A.S., et al., *Impact of Processing Method on Recovery of Bacteria from Wipes Used in Biological Surface Sampling.* Applied and Environmental Microbiology, 2012. **78**(16): p. 5872-5881.
377. Iacobellis, N.S. and J.E. Devay, *Long-term storage of plant-pathogenic bacteria in sterile distilled water.* Appl Environ Microbiol, 1986. **52**(2): p. 388-9.
378. Schau, H.-P., *E. Mitscherlich and E. H. Marth, Microbial Survival in the Environment — Bacteria and Rickettsiae Important in Human and Animal Health. X + 802 S., 30 Abb., 328 Tab. Berlin-Heidelberg-New York-Tokyo 1984. Springer-Verlag. DM 390,00. ISBN: 3-540-13726-2.* Journal of Basic Microbiology, 1985. **25**(10): p. 674-674.
379. Miles, A.A., S.S. Misra, and J.O. Irwin, *The estimation of the bactericidal power of the blood.* Epidemiology and Infection, 1938. **38**(6): p. 732-749.
380. Gill, C., et al., *Evaluation of Lysis Methods for the Extraction of Bacterial DNA for Analysis of the Vaginal Microbiota.* PLOS ONE, 2016. **11**(9): p. e0163148.
381. Martzy, R., et al., *Simple lysis of bacterial cells for DNA-based diagnostics using hydrophilic ionic liquids.* Scientific Reports, 2019. **9**(1): p. 13994.
382. Vandeventer, P.E., et al., *Mechanical disruption of lysis-resistant bacterial cells by use of a miniature, low-power, disposable device.* J Clin Microbiol, 2011. **49**(7): p. 2533-9.
383. Maghini, D.G., et al., *Improved high-molecular-weight DNA extraction, nanopore sequencing and metagenomic assembly from the human gut microbiome.* Nat Protoc, 2021. **16**(1): p. 458-471.
384. Li, X., et al., *Efficiency of chemical versus mechanical disruption methods of DNA extraction for the identification of oral Gram-positive and Gram-negative bacteria.* J Int Med Res, 2020. **48**(5): p. 300060520925594.
385. Auer, G.K. and D.B. Weibel, *Bacterial Cell Mechanics.* Biochemistry, 2017. **56**(29): p. 3710-3724.
386. Davis, K.M. and J.N. Weiser, *Modifications to the peptidoglycan backbone help bacteria to establish infection.* Infect Immun, 2011. **79**(2): p. 562-70.
387. Schindler, C.A. and V.T. Schuhardt, *LYSOSTAPHIN: A NEW BACTERIOLYTIC AGENT FOR THE STAPHYLOCOCCUS.* Proc Natl Acad Sci U S A, 1964. **51**(3): p. 414-21.
388. Yuan, S., et al., *Evaluation of Methods for the Extraction and Purification of DNA from the Human Microbiome.* PLOS ONE, 2012. **7**(3): p. e33865.
389. Yokogawa, K., et al., *Mutanolysin, bacteriolytic agent for cariogenic Streptococci: partial purification and properties.* Antimicrob Agents Chemother, 1974. **6**(2): p. 156-65.
390. Lagier, J.C., et al., *Current and past strategies for bacterial culture in clinical microbiology.* Clin Microbiol Rev, 2015. **28**(1): p. 208-36.
391. Ogai, K., et al., *A Comparison of Techniques for Collecting Skin Microbiome Samples: Swabbing Versus Tape-Stripping.* Frontiers in Microbiology, 2018. **9**.
392. Breitwieser, F.P., J. Lu, and S.L. Salzberg, *A review of methods and databases for metagenomic classification and assembly.* Brief Bioinform, 2019. **20**(4): p. 1125-1136.
393. Kim, D., et al., *Centrifuge: rapid and sensitive classification of metagenomic sequences.* Genome Res, 2016. **26**(12): p. 1721-1729.

394. Portik, D.M., C.T. Brown, and N.T. Pierce-Ward, *Evaluation of taxonomic classification and profiling methods for long-read shotgun metagenomic sequencing datasets.* BMC Bioinformatics, 2022. **23**(1): p. 541.
395. Lu, H., F. Giordano, and Z. Ning, *Oxford Nanopore MinION Sequencing and Genome Assembly.* Genomics Proteomics Bioinformatics, 2016. **14**(5): p. 265-279.
396. Toscano, M., et al., *Role of the Human Breast Milk-Associated Microbiota on the Newborns' Immune System: A Mini Review.* Frontiers in Microbiology, 2017. **8**.
397. Yan, W., et al., *Association and Occurrence of Bifidobacterial Phylotypes Between Breast Milk and Fecal Microbiomes in Mother–Infant Dyads During the First 2 Years of Life.* Frontiers in Microbiology, 2021. **12**.
398. Kiu, R., et al., *Impact of early life antibiotic and probiotic treatment on gut microbiome and resistome of very-low-birth-weight preterm infants.* bioRxiv, 2025: p. 2024.12.18.629205.
399. Gibb, A.P. and S. Wong, *Inhibition of PCR by agar from bacteriological transport media.* J Clin Microbiol, 1998. **36**(1): p. 275-6.
400. Bourbeau, P.P., et al., *Use of CultureSwab Plus swabs with Amies gel agar for testing of naris specimens with the GeneOhm MRSA assay.* J Clin Microbiol, 2007. **45**(7): p. 2281-3.
401. Grahn, N., et al., *Identification of mixed bacterial DNA contamination in broad-range PCR amplification of 16S rDNA V1 and V3 variable regions by pyrosequencing of cloned amplicons.* FEMS Microbiology Letters, 2003. **219**(1): p. 87-91.
402. Lauder, A.P., et al., *Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota.* Microbiome, 2016. **4**(1): p. 29.
403. Salter, S.J., et al., *Reagent and laboratory contamination can critically impact sequence-based microbiome analyses.* BMC Biology, 2014. **12**(1): p. 87.
404. Glassing, A., et al., *Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples.* Gut Pathogens, 2016. **8**(1): p. 24.
405. Olomu, I.N., et al., *Elimination of "kitome" and "splashome" contamination results in lack of detection of a unique placental microbiome.* BMC Microbiology, 2020. **20**(1): p. 157.
406. Jervis-Bardy, J., et al., *Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data.* Microbiome, 2015. **3**: p. 19.
407. Davis, N.M., et al., *Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data.* Microbiome, 2018. **6**(1): p. 226.
408. Laurence, M., C. Hatzis, and D.E. Brash, *Common contaminants in next-generation sequencing that hinder discovery of low-abundance microbes.* PLoS One, 2014. **9**(5): p. e97876.
409. Minich, J.J., et al., *KatharoSeq Enables High-Throughput Microbiome Analysis from Low-Biomass Samples.* mSystems, 2018. **3**(3).
410. Albertsen, M., et al., *Back to Basics – The Influence of DNA Extraction and Primer Choice on Phylogenetic Analysis of Activated Sludge Communities.* PLOS ONE, 2015. **10**(7): p. e0132783.
411. Whyte, W., W. Carson, and A. Hambraeus, *Methods for calculating the efficiency of bacterial surface sampling techniques.* J Hosp Infect, 1989. **13**(1): p. 33-41.
412. Hambraeus, A., J. Hoborn, and W. Whyte, *Skin sampling-validation of a pad method and comparison with commonly used methods.* J Hosp Infect, 1990. **16**(1): p. 19-27.
413. Tajima, M., et al., *Molecular analysis of Malassezia microflora in seborrheic dermatitis patients: comparison with other diseases and healthy subjects.* J Invest Dermatol, 2008. **128**(2): p. 345-51.
414. Sarnaik, A., et al., *Novel perspective on a conventional technique: Impact of ultra-low temperature on bacterial viability and protein extraction.* PLoS One, 2021. **16**(5): p. e0251640.
415. Eisenhofer, R., et al., *Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations.* Trends in Microbiology, 2019. **27**(2): p. 105-117.
416. Lou, Y.C., et al., *Using strain-resolved analysis to identify contamination in metagenomics data.* bioRxiv, 2022: p. 2022.01.16.476537.
417. Zhou, Q., X. Su, and K. Ning, *Assessment of quality control approaches for metagenomic data analysis.* Scientific Reports, 2014. **4**(1): p. 6957.
418. Kostic, A.D., et al., *PathSeq: software to identify or discover microbes by deep sequencing of human tissue.* Nat Biotechnol, 2011. **29**(5): p. 393-6.
419. Marsh, R.L., et al., *How low can we go? The implications of low bacterial load in respiratory microbiota studies.* Pneumonia, 2018. **10**(1): p. 7.
420. Díaz, S., J.S. Escobar, and F.W. Avila, *Identification and Removal of Potential Contaminants in 16S rRNA Gene Sequence Data Sets from Low-Microbial-Biomass Samples: an Example from Mosquito Tissues.* mSphere, 2021. **6**(3): p. e00506-21.

421. Heikema, A.P., et al., *Comparison of Illumina versus Nanopore 16S rRNA Gene Sequencing of the Human Nasal Microbiota.* Genes, 2020. **11**(9): p. 1105.
422. Browne, P.D., et al., *GC bias affects genomic and metagenomic reconstructions, underrepresenting GC-poor organisms.* GigaScience, 2020. **9**(2).
423. Benjamini, Y. and T.P. Speed, *Summarizing and correcting the GC content bias in high-throughput sequencing.* Nucleic Acids Research, 2012. **40**(10): p. e72-e72.
424. McNaughton, A.L., et al., *Illumina and Nanopore methods for whole genome sequencing of hepatitis B virus (HBV).* Scientific Reports, 2019. **9**(1): p. 7081.
425. Nayfach, S., et al., *New insights from uncultivated genomes of the global human gut microbiome.* Nature, 2019. **568**(7753): p. 505-510.
426. Bowers, R.M., et al., *Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea.* Nature Biotechnology, 2017. **35**(8): p. 725-731.
427. Sczyrba, A., et al., *Critical Assessment of Metagenome Interpretation—a benchmark of metagenomics software.* Nature Methods, 2017. **14**(11): p. 1063-1071.
428. DeMaere, M.Z. and A.E. Darling, *bin3C: exploiting Hi-C sequencing data to accurately resolve metagenome-assembled genomes.* Genome Biology, 2019. **20**(1): p. 46.
429. Gweon, H.S., et al., *The impact of sequencing depth on the inferred taxonomic composition and AMR gene content of metagenomic samples.* Environmental Microbiome, 2019. **14**(1): p. 7.
430. Whittaker, R.H., *Vegetation of the Siskiyou Mountains, Oregon and California.* Ecological Monographs, 1960. **30**(3): p. 279-338.
431. Whittaker, R.H., *EVOLUTION AND MEASUREMENT OF SPECIES DIVERSITY.* TAXON, 1972. **21**(2-3): p. 213-251.
432. Bray, J.R. and J.T. Curtis, *An Ordination of the Upland Forest Communities of Southern Wisconsin.* Ecological Monographs, 1957. **27**(4): p. 325-349.
433. Mallick, H., et al., *Multivariable association discovery in population-scale meta-omics studies.* PLOS Computational Biology, 2021. **17**(11): p. e1009442.
434. Oksanen, J., et al., *vegan community ecology package version 2.6-2 April 2022*. 2022.
435. W. P, E., *Statistical Methods for Research Workers. By R. A. Fisher. [Pp. 239 + ix + VI Tables. Edinburgh and London: Oliver & Boyd. 1925. Price 15s.] - The Fundamentals of Statistics. By L. L. Thurstone. [Pp. 237 + xvi. New York: The Macmillan Company. 1925. Price 8s. 6d.].* Journal of the Institute of Actuaries, 1925. **56**(3): p. 326-327.
436. Lenth, R.V. *emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.10.4*. 2024 [cited 2024; Available from: https://rvlenth.github.io/emmeans/.
437. Daniel Sabanes Bove, L.L., Julia Dedic, Doug Kelkhoff, Kevin Kunzmann, Brian Matthew Lang, Christian Stock, Ya Wang, Dan James, Jonathan Sidi, Daniel Leibovitz, Daniel D. Sjoberg, Boehringer Ingelheim Ltd., Gilead Sciences, Inc., F. Hoffmann-La Roche AG, Merck Sharp & Dohme, Inc., AstraZeneca plc. *mmrm: Mixed Models for Repeated Measures. R package version 0.3.12*. 2024 [cited 2024; Available from: https://openpharma.github.io/mmrm/.
438. Anderson, M.J., *A new method for non-parametric multivariate analysis of variance.* Austral Ecology, 2001. **26**(1): p. 32-46.
439. Martinez Arbizu, P. *pairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.4*. 2020 [cited 2024; Available from: https://github.com/pmartinezarbizu/pairwiseAdonis.
440. Katz, L.S., et al., *Mashtree: a rapid comparison of whole genome sequence files.* J Open Source Softw, 2019. **4**(44).
441. Ondov, B.D., et al., *Mash: fast genome and metagenome distance estimation using MinHash.* Genome Biology, 2016. **17**(1): p. 132.
442. Felsenstein, J., *Evolutionary trees from DNA sequences: A maximum likelihood approach.* Journal of Molecular Evolution, 1981. **17**(6): p. 368-376.
443. Letunic, I. and P. Bork, *Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool.* Nucleic Acids Research, 2024. **52**(W1): p. W78-W82.
444. Letunic, I. and P. Bork, *Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation.* Bioinformatics, 2007. **23**(1): p. 127-8.
445. Ciccarelli, F.D., et al., *Toward automatic reconstruction of a highly resolved tree of life.* Science, 2006. **311**(5765): p. 1283-7.
446. Barker-Tejeda, T.C., et al., *Comparative characterization of the infant gut microbiome and their maternal lineage by a multi-omics approach.* Nature Communications, 2024. **15**(1): p. 3004.
447. Lebeer, S., et al., *A citizen-science-enabled catalogue of the vaginal microbiome and associated factors.* Nature Microbiology, 2023.

448. Hu, F., et al., *The Dynamic Changes in the Composition and Diversity of Vaginal Microbiota in Women of Different Pregnancy Periods.* Microorganisms, 2023. **11**(11).
449. Murphy, B., et al., *Compositional Variations between Adult and Infant Skin Microbiome: An Update.* Microorganisms, 2023. **11**(6).
450. Rahma, A. and M.E. Lane, *Skin Barrier Function in Infants: Update and Outlook.* Pharmaceutics, 2022. **14**(2).
451. Butler, C.A., et al., *Breastmilk influences development and composition of the oral microbiome.* J Oral Microbiol, 2022. **14**(1): p. 2096287.
452. Kordy, K., et al., *Contributions to human breast milk microbiome and enteromammary transfer of Bifidobacterium breve.* PLoS One, 2020. **15**(1): p. e0219633.
453. Umbach, A.K., A.A. Stegelmeier, and J.D. Neufeld, *Archaea Are Rare and Uncommon Members of the Mammalian Skin Microbiome.* mSystems, 2021. **6**(4): p. 10.1128/msystems.00642-21.
454. Pausan, M.R., et al., *Exploring the Archaeome: Detection of Archaeal Signatures in the Human Body.* Frontiers in Microbiology, 2019. **10**.
455. Ross, A.A., et al., *Comprehensive skin microbiome analysis reveals the uniqueness of human skin and evidence for phylosymbiosis within the class Mammalia.* Proc Natl Acad Sci U S A, 2018. **115**(25): p. E5786-e5795.
456. Ross, A.A., A.C. Doxey, and J.D. Neufeld, *The Skin Microbiome of Cohabiting Couples.* mSystems, 2017. **2**(4): p. 10.1128/msystems.00043-17.
457. Koo, H., et al., *An individualized mosaic of maternal microbial strains is transmitted to the infant gut microbial community.* Royal Society Open Science, 2020. **7**(4): p. 192200.
458. Ferretti, P., et al., *Mother-to-Infant Microbial Transmission from Different Body Sites Shapes the Developing Infant Gut Microbiome.* Cell Host Microbe, 2018. **24**(1): p. 133-145.e5.
459. Asnicar, F., et al., *Studying Vertical Microbiome Transmission from Mothers to Infants by Strain-Level Metagenomic Profiling.* mSystems, 2017. **2**(1): p. e00164-16.
460. Lou, Y.C., et al., *Infant gut strain persistence is associated with maternal origin, phylogeny, and traits including surface adhesion and iron acquisition.* Cell Reports Medicine, 2021. **2**(9): p. 100393.
461. Podlesny, D. and W.F. Fricke, *Strain inheritance and neonatal gut microbiota development: A meta-analysis.* International Journal of Medical Microbiology, 2021. **311**(3): p. 151483.
462. Yang, R., et al., *Inspecting mother-to-infant microbiota transmission: disturbance of strain inheritance by cesarian section.* Frontiers in Microbiology, 2024. **15**.
463. Dubois, L., et al., *Paternal and induced gut microbiota seeding complement mother-to-infant transmission.* Cell Host & Microbe, 2024. **32**(6): p. 1011-1024.e4.
464. Santos, S.J.D., et al., *Birth Mode Does Not Determine the Presence of Shared Bacterial Strains between the Maternal Vaginal Microbiome and the Infant Stool Microbiome.* Microbiology Spectrum, 2023. **11**(4): p. e00614-23.
465. Brooks, B., et al., *Strain-resolved analysis of hospital rooms and infants reveals overlap between the human and room microbiome.* Nature Communications, 2017. **8**(1): p. 1814.
466. Wang, C., et al., *Novel Polyoxyethylene-Containing Glycolipids Are Synthesized in Corynebacterium matruchotii and Mycobacterium smegmatis Cultured in the Presence of Tween 80.* J Lipids, 2011. **2011**: p. 676535.
467. Shaffer, M., et al., *DRAM for distilling microbial metabolism to automate the curation of microbiome function.* Nucleic Acids Research, 2020. **48**(16): p. 8883-8900.
468. Kumar, N., et al., *MiMiC: a bioinformatic approach for generation of synthetic communities from metagenomes.* Microb Biotechnol, 2021. **14**(4): p. 1757-1770.
469. Loomis, K.H., et al., *A mixed community of skin microbiome representatives influences cutaneous processes more than individual members.* Microbiome, 2021. **9**(1): p. 22.
470. McLoughlin, I.J., et al., *Cosmetic Efficacy of the Topical Probiotic Micrococcus luteus Q24 in Healthy Human Adults.* Cosmetics, 2024. **11**(4): p. 122.

APPENDICES

Appendix 1. FMH Approval Letter

Faculty of Medicine and Health Sciences Research Ethics Committee

University of East Anglia

NORWICH MEDICAL SCHOOL
Bob Champion Research & Educational Building
Rosalind Franklin Road
University of East Anglia
Norwich Research Park
Norwich NR4 7UQ

Email: fmh.ethics@uea.ac.uk
www.med.uea.ac.uk

Iliana Serghiou
Quadram Institute Bioscience
University of East Anglia
Norwich Research Park
Norwich
NR4 7UQ

29th July 2021

Dear Iliana

Project title: Method Development for Bacterial DNA Extraction from Skin Swabs

Reference: 2020/21-065

Thank you for your email of 27th July 2021 notifying us of the amendments you would like to make to your above proposal. These have been considered and I can confirm that your amendments have been approved.

Please can you ensure that any further amendments to either the protocol or documents submitted are notified to us in advance, and that any adverse events which occur during your project are reported to the Committee.

Approval by the FMH Research Ethics Committee should not be taken as evidence that your study is compliant with GDPR and the Data Protection Act 2018. If you need guidance on how to make your study GDPR compliant, please contact your institution's Data Protection Officer.

Please can you arrange to send us a report once your project is completed.

Yours sincerely

Dr Jackie Buck
Chair
FMH Research Ethics Committee

Appendix 2. FMH Application

Application Form for Ethical Approval of a Health Related Research Project

IMPORTANT: PLEASE READ BEFORE COMPLETING THIS FORM:

1. Please refer to the guidelines for Applicants conducting Research Projects in FMH when completing this form. The link to this is on the FMH Ethics intranet page.

2. **It is ESSENTIAL that you do not use complex/technical language.** This is to ensure that the objectives of your project/research and the procedures to be conducted can easily be understood by non-specialists and lay members of the Committee.

3. **All submitted applications must include an electronic version of all relevant documents.** All hard copies must be collated and fixed together in the top left hand corner.

4. **Please include your supervisor when emailing the application if you are a student.**

5. **If the project involves the use of drugs, or testing of new equipment, or research on NHS patients** it MUST be referred to an NHS Research Ethics Committee for approval.

6. **Applications cannot be accepted after the deadline** so please ensure that you read the information above and complete the checklist at the back of this form. To avoid delays it is essential that you ensure you have provided all of the required information in the requested format, on or before the deadline.

If this is related to a Research project please include the following information (if student, you may need to ask your supervisor for this):

NB I am a PhD student registered at UEA but with the NRP DTP and based in the Quadram Institute

REN project number: N/A

Name of REN Project Officer: N/A

If the project is a resubmission, please provide the FMH Ethics Reference Number:

N/A

For standard applications. Please send 2 hard copies of the application form, proposal and all other supporting documents to: FMH Research Ethics, c/o FMH RIN Administration, Research & Innovation Office, The Registry 1.14, University of East Anglia, Norwich NR4 7TJ. **Please also e-mail ONE copy of all documents to** fmh.ethics@uea.ac.uk on or before the deadline shown on the following intranet page. (https://portal.uea.ac.uk/faculty-school-intranets/fmh-intranet/ethics-committee).

If you are submitting a **Service Evaluation/Audit or Human Tissue application,** please tick one of the following relevant boxes:

1. **If the project involves the use of Human Tissue**, please complete this form and email **ONE** signed copy (if student) including all supporting documents to fmh.ethics@uea.ac.uk. These can be emailed at any time and do not usually need to go to a Committee meeting. **We do not need hard copies of Human Tissue applications.** ☐

2. **If the project is a Service Evaluation or Audit within the NHS**, please complete this form and email **ONE** signed copy (if student) including all supporting documents to fmh.ethics@uea.ac.uk. **The Chair will also require evidence of acceptance by the relevant host NHS Trust with your submission.** These can be emailed at any time and do not usually need to go to a Committee meeting. **We do not need hard copies of Service Evaluations/Audits.** ☐

For any queries please email: fmh.ethics@uea.ac.uk

Please ensure that you have completed the checklist at the end of the form before submitting and have provided all of the required information, as the Committee will return your application if the required information is not provided.

Application form

1. Name of applicant: ILIANA ROSA SERGHIOU
   (Block letters)

2. Academic address for correspondence (please do not use your home address):

   Quadram Institute of Biosciences, Norwich Research Park Post code: NR4 7UA

3. Tel No: Work - 01603 255034 / Mobile - 07738610612

4. Academic (UEA) E-mail address:

   Btm20tqu@uea.ac.uk / Iliana.serghiou2@quadram.ac.uk

5. School: Quadram Institute Biosciences

6. Status of applicant (Staff, UG or PG student - and year of course): PG PhD student - 2020

7. Supervisory arrangements for **STUDENT PROJECTS ONLY:**

   Is this study being carried out to fulfil a required part of your course? Yes

   Degree/Course: PhD Study – The Development and Evolution of the Human Infant Skin Microbiome

   School: QIB

   Name and contact details of UEA supervisor:

   Supervisor Name: Dr. Mark Webber (MED)

   Supervisor Email: Mark.Webber@quadram.ac.uk / m.webber@uea.ac.uk

Please ensure that your supervisor signs the declaration on page 4 of this document

8. Has this application gone to an Ethics Committee elsewhere? No

If YES, please indicate where and **provide copies of correspondence**:

N/A

Project details (**sections 9, 10 and 11 must be limited to a combined maximum of 3000 words**).

9. Full title: Method Development for Bacterial DNA Extraction from Skin Swabs

10. Purpose of project:

    My PhD project will study how bacteria on the skin change through early life. This will use samples from the 'PEARL' project which has already been established and has full ethical approval (Lindsay Hall, the PEARL PI is my second supervisor). However, before working with any PEARL samples I want to optimise a method for extracting bacterial DNA from swabs taken from human skin. This application is to ask for ethical permission for this work which will then allow the PEARL samples to be processed in an optimal manner.

    Currently, there is a lack of research on the human skin microbiome and an optimised bacterial DNA extraction method specific to skin swabs has not been established. Therefore, there is a need to develop a method for this.

    Due to the scale of the PEARL study, the development of a method compatible with downstream use of a high-throughput automated DNA extraction method using the Promega Maxwell® RSC Blood DNA Kit and Maxwell® RSC 48 Instrument (purchased for the study), is required.

    Therefore, the purpose of this project is to develop a high-throughput automated method, using the Promega Maxwell® RSC Blood DNA Kit with the Maxwell® RSC 48 Instrument, for bacterial DNA extraction from skin swabs that results in DNA with high yield, quality, purity, and high molecular weight suitable for metagenomics sequencing.

11. Methodology, Procedure and Analysis:

    QIB Researcher: Iliana Rosa Serghiou
    Principle Investigator: Dr. Mark Webber
    Supervisors: Dr. Mark Webber & Dr. Lindsay Hall

    <u>Study Design</u>

    Volunteer participants will be recruited by an expression of interest email sent out to QIB staff by the Biorepository. The recruiters will be trained in consenting for the Biorepository. The recruiters will then provide participants with the established Biorepository participant information sheets and consent participants with the established Biorepository consent forms. To adhere to Covid19 regulations, participants will be consented either remotely or at a socially distanced location at QIB to reduce person-person contact. 24 participants will be recruited and each swabbed once to collect one skin swab per participant. If it is difficult to recruit 24 participants, 12 participants may be swabbed at two separate body sites to produce the 24 biological samples in total. The recruiters will swab participants to collect the samples. If the participants are consented to the study in person, their swabs will be collected at the same time. If the participant has been consented remotely, the recruiters will arrange a suitable day/ time to meet the participant to collect these samples. We hope that we can condense these sampling sessions into a period of 2 days, with no more than 12 participants a day. Appointments will be staggered on each of these days to reduce social contact. All participants and those swabbing will wear a face mask, and those swabbing will also wear gloves, to reduce person-person contact. In the case of reduced recruiting staff due to covid, participants will be consented remotely, as described above, and will self-swab at the QIB, using the PEARL study self-swabbing instructions. The participants will then return the swabs to the available recruiters the same day.

Participants must have no current skin conditions and have had no skin conditions or been prescribed antibiotics over the last 3 months.

Sample Collection and Handling

As this method development procedure is being performed for the purpose of optimising the DNA extraction method of the PEARL study skin swabs, swabs will be swabbed according to the PEARL study protocol which swabs the inner (volar) forearm of either the left or right arm with Charcoal Amies Transport Medium swabs. The protocol will be as follows:

To perform swabbing, pull and twist the swab out of its tube taking care not to touch the tip (swab end) with anything other than the area to be swabbed. Holding the swab in one hand, roll the opposite end (swab end) along the inside of the forearm from the wrist to the elbow pit ensuring that the whole of the swab end has come in to contact with the volunteers skin. Replace the swab back into the tube ensuring the lid is on tight. The swab may leave a grey residue on the forearm. Once sampling is complete, wash the area with soap and water. The date of sample collection and Biorepository sample ID will be written on the container in pen. Samples will be stored in a 4°C fridge until samples are collected for testing.

Samples will be collected and tested on the same day swabbing was performed.

Experimental Methods

A review of the literature and protocols supplied by the CASE partner (Unilever) have been analysed to identify a potentially efficient DNA extraction method. The main variables to be tested are the initial diluent for swab suspension, type of 'bead beating' to disrupt bacterial cells, and different lysis protocols. The detailed protocol and variations to be tested are listed below:

Initial diluent testing
The skin swab samples that were collected and stored in Charcoal Amies Transport Media will be processed in two diluents, PBS and water, then cultured to grow bacteria in order to determine which diluent is more effective for collecting skin bacterial cells in the suspension (water may have additional lytic effect to PBS). When culturing the skin swabs, Colombia Blood Agar plates will be used as this is a known non-selective media suggested in literature for culturing a diversity of bacteria. To obtain a good representation of aerobic and anaerobic bacteria present on the skin, skin swabs will be cultured aerobically and anaerobically (Ogai et al. 2018).

Two studies by Alan and Sallene 1998 and Bourbeau et al. 2007 have indicated that the agar in the Amies transport medium swabs inhibits molecular techniques such as PCR. Skin swabs will be washed twice to reduce the chance of the presence of transport media which could potentially impact metagenomics sequencing.

The processed swabs will then go through a series of five DNA extraction procedures using the Promega Maxwell® RSC Blood DNA Kit and the Maxwell® RSC 48 Instrument. This Promega Maxwell® kit and instrument permits a high-throughput automated genomic DNA isolation from up to 48 whole blood samples in 40 minutes which is ideal for processing large sample sets efficiently. The Maxwell® RSC 48 Instrument purifies samples using a paramagnetic particle to allow sample capture, washing, and purification of gDNA, and provides a higher bind capacity and cleaner eluate than traditional silica-based DNA purification systems (Sui et al. 2020; Moeller et al. 2014; Dunbar et al. 2018).

The Promega Maxwell® Kit and instrument was purchased as it allows high throughput and has been used to produce high molecular weight DNA (Mandrekar et al. 2010) which is important in metagenomic studies such as this (Bey et al. 2010).

As the manufacturers technical manual recommends the use of up to 300µl of sample, this volume will be used in the five experimental methods below.

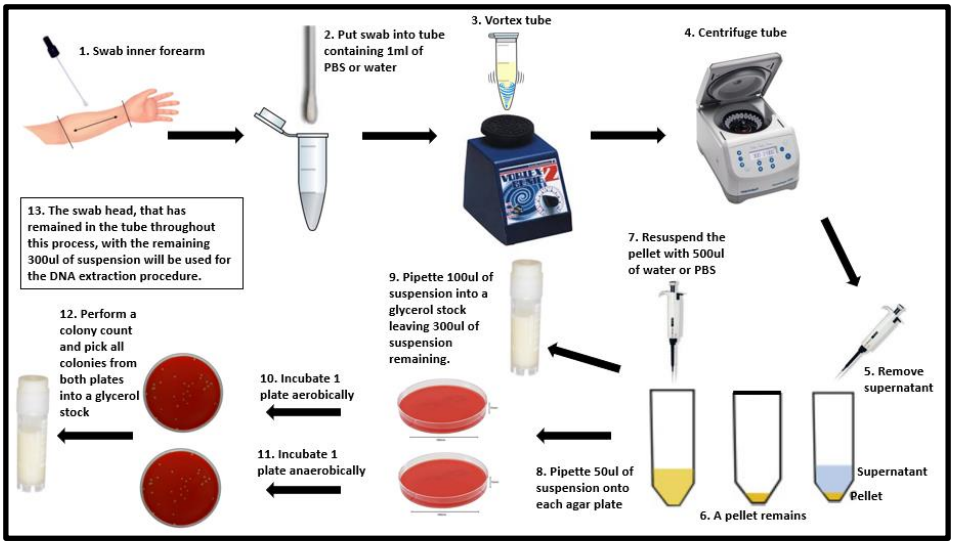
Bead beating
Previous bacterial DNA extraction methods using the Promega Maxwell® instrument have incorporated a bead beating step using e.g. the Lysing Matrix E tubes and the MP Bio FastPrep instrument which is a high velocity bead beater. This protocol has been incorporated as it is an established protocol at the QIB for extracting bacterial DNA from human samples. The bead beating step has been suggested to be important for the lysis of bacteria, and certain microorganisms, such as Staphylococcus and Streptococcus which are vital skin probiotics, require a stronger homogenization using a high velocity bead beater. The Lysing Matrix E tubes contain 1.4mm ceramic spheres, 0.1 mm silica spheres and one 4 mm glass bead, and this range of bead sizes allows efficient lysis of diverse human and environmental sample types (Sui et al. 2020).

Lysing matrix E tubes are a suitable choice for the bead beating step in this protocol as they can be used on any bead-beating device according to the manufacturer's instructions. Two different high velocity bead beating instruments will be compared, the MP Bio FastPrep and the QIAGEN Tissue Lyser, to see if there is a difference in DNA yield, quality, molecular weight, and bacterial composition recovered after each. The QIAGEN Tissue Lyser is an established instrument used for a bead beating step in Unilever's skin bacterial DNA extraction protocol which uses the QIAGEN QIAsymphony instrument with the QIAGEN QIAsymphony Virus/Bacteria Midi Kit (931055). As the CASE partner of this PhD study, Unilever have provided access to their established skin bacteria DNA extraction procedure. However, this Unilever protocol is specific for 16s sequencing which allows for more fragmented DNA, thus there is a worry that using a high velocity bead beating instrument might cause very fragmented DNA not suitable for metagenomics. If the resulting DNA is too fragmented, the bead beating instrument could be replaced with a series of vortexing steps that other high molecular weight kits perform.

Lysis
The Unilever protocol also includes, an overnight heated lysis with the Epicentre Ready-Lyse lysozyme, and the QIAGEN QIASymphony off-board lysis procedure, which will be compared to the Maxwell lysis method to see if there is a difference in DNA concentration, quality, molecular weight and bacterial composition. The overnight heated lysis with a Lysozyme is vital for the lysis of certain bacteria such as Staphylococcus and Streptococcus. The Ready-Lyse lysozyme will also be compared to the cheaper Thermofisher lysozyme to see if this affects bacteria extracted. If there is no difference the Thermofisher lysozyme will be more cost-effective to use.

For each run a negative control containing PBS will be used to ensure no contamination has occurred, and a positive control containing a spiked suspension with an established mock community will be used to demonstrate that any negative results are a lack of material rather than an extraction problem. A recent study evaluated the composition of the mock microbial communities across four DNA extraction procedures, one being the Promega Maxwell gDNA which provided the best representation of the composition of the mock community (Sui et al. 2020).

Figure 1 below summarises the protocol and variations to be tested.



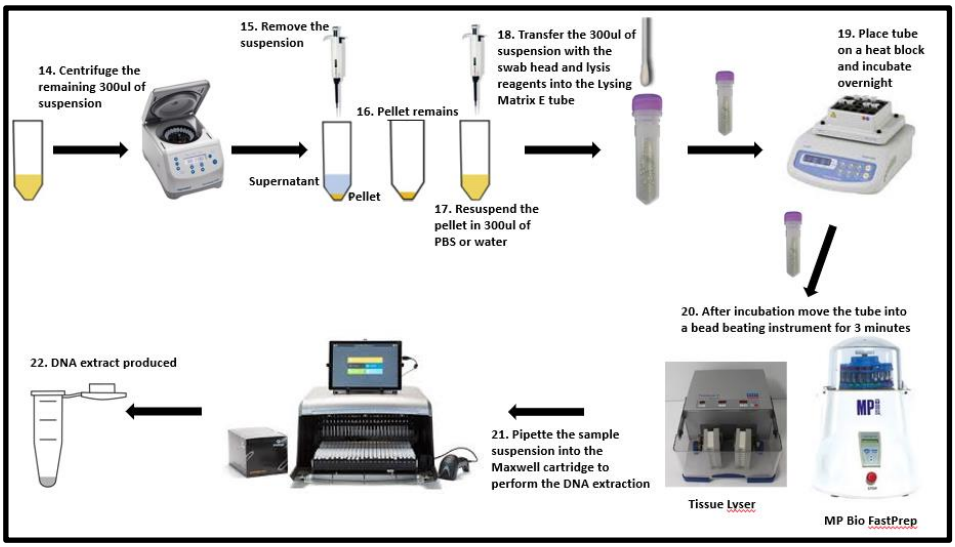


Figure 1 – Protocol Variations to be Tested

Analysis of DNA

Resulting DNA extracts will then be analysed for quantity, quality and fragment length using a QuBit and Tapestation instruments (both available in QIB). DNA will then be sequenced using an Illumina platform (NextSeq) to verify successfully harvesting of bacterial DNA from the volunteers. WGS sequencing allows the microbiota to be resolved down to a species level. Scripts already established in the Webber lab will be utilised to resolve a microbiota profile to identify the bacterial composition of the samples. This will help determine which DNA extraction method provides a better representation of the skin bacterial composition.

Statistical Analysis

As this protocol is for the purpose of method development and determining optimum conditions for DNA extraction data no statistical power 'effect' is being examined.

Storage of Processed Samples

There is no need to store any of these samples in the long term – DNA preparations will be kept for 12 months in case any further methodological verification in terms of sequence library preparation is needed but will then be destroyed according to established waste disposal procedures within QIB.
Bacterial DNA sequences will be kept on secure servers within QIB (and anonymised from donor as described below).

Data protection and Donor Confidentiality

There will be no contact between researcher and participant. The Biorepository members of staff will remain the link between the researcher and participant. The Biorepository will also label all swabs with a Biorepository sample ID completely unlinked to the participants personal information to eliminate the chance of identifiable information being given to the researcher. Samples will be collected by the Biorepository staff members who will ensure its anonymity prior to being given to the researcher. The Biorepository contact information will be provided to participants for them to express any questions or concerns regarding the study. The Biorepository can then contact the researcher for further clarification.
Any DNA reads mapping to the human genome will be identified and destroyed (deleted) bioinformatically in the initial analysis following an established procedure within QIB.

Ethical and Regulatory Considerations

As this method development will go through the Biorepository, swabbing must be anonymised, and no samples can be stored therefore swabs must be discarded after use. To adhere to this, the Biorepository will recruit, consent, and swab volunteers, and samples will be collected from the Biorepository and tested on the same day the swabbing took place. Further to this, after the aims have been answered, the host DNA and resulting host sequences will be discarded.

12. Resources required:

    The cooperation of the Biorepository will be required.

13. Source of Funding:

    CTP iCASE Project (Unilever)

14. Has this project been peer reviewed? If yes, please include details of who the project has been peer reviewed by.

    As part of the grant application this project has been peer reviewed.

15. Ethical issues:

    There are minimal ethical issues. The collection of skin swab samples is non-invasive.

16. Proposed start and finish dates:

    Start date: 31/05/2021 Finish date: 31/09/2021

17. Where will the research be carried out?

    Quadram Institute of Biosciences

18. Information sheets and consent forms must be appended (c.f. NRES site for models, http://www.hra-decisiontools.org.uk/consent/ Please ensure that participants are requested to initial the boxes on the consent forms.

19. Checklist - **please check and complete before submitting your application as incomplete applications will be returned by the Committee. Many applications come in very close to the deadline and if we have to return it you may miss that month's meeting:**

	Yes / No / N/A
Have you completed all sections of the application in language which will be understood by lay people?	Yes
If student, has your supervisor signed the form?	Yes
If student, please provide name and email address for your supervisor	Yes
Have you included your academic address (not your home address)?	Yes
Have you included a header and footer on each page with your name, date of submission, version number and page number?	Yes
Have you included the following documents, if applicable?	
• Protocol. It is recommended that a protocol is always submitted as it facilitates a comprehensive review of the project	Yes
• Gatekeeper consent	Yes
• Participant information sheets (using NRES format)	N/A
• Consent forms	N/A
• Letters to participants	N/A
• Copies of questionnaires	N/A
• Copies of correspondence from other ethics committees	N/A
• Copies of all recruitment letters, emails, posters and adverts	N/A
• Research Safety Checklist **(please complete even if no risks are identified)**	Yes
Have you proof-read your application to check for typographical and grammatical errors?	Yes
Have you included 5 hard copies of your application and all supporting documents (**collated and attached together**) and e-mailed a copy of all documents to fmh.ethics@uea.ac.uk?	N/A
If this is a Service Evaluation/Audit have you emailed ONE signed copy of your application and all supporting documents to fmh.ethics@uea.ac.uk?	N/A
If this is a Service Evaluation have you included the evidence of acceptance by the relevant host NHS Trust with your submission?	N/A
If this is a Human Tissue application have you emailed ONE signed copy of your application and all supporting documents to fmh.ethics@uea.ac.uk?	Yes

Academic Supervisor Declaration – for STUDENT PROJECTS ONLY:

I have read this application and can confirm that I am taking supervisory responsibility for this project.

In the case of a student research outside the normal course requirements I confirm that I am happy to take responsibility for the quality of protocol design, the provision of necessary resources, statistical support and usual supervision and governance of the student.

Project Supervisor's signature: Date:



8.12.20

Post Held: Group Leader

Appendix 3. Human Tissue study Protocol

Method Development for Bacterial DNA Extraction from Skin Swabs

Short title: Method Development for Bacterial DNA Extraction from Skin Swabs

PROTOCOL:

Version 1

QIB Investigator (s):

[Iliana Rosa Serghiou]

NHS Investigator(s):

[N/A]

Other Investigator(s):

[Dr. Mark Webber & Dr. Linsay Hall]

List of contents

Executive Summary
1. Introduction
1.1. Purpose
1.2. Aims and Objectives
2. Study Design
2.1. Recruitment policy
2.2. Eligibility criteria
2.2.1. Basic inclusion criteria
2.2.2. Basic exclusion criteria
3. Sample Collection and Handling
4. Experimental Methods
5. Statistical Analysis
6. Storage of Processed Samples
7. Data protection and Donor Confidentiality
8. Ethical and Regulatory Considerations
9. Reference

Executive Summary

1. Introduction

My PhD project will study how bacteria on the skin change through early life. This will use samples from the 'PEARL' project which has already been established and has full ethical approval (Lindsay Hall, the PEARL PI is my second supervisor). However, before working with any PEARL samples I want to optimise a method for extracting bacterial DNA from swabs taken from human skin. This application is to ask for ethical permission for this work which will then allow the PEARL samples to be processed in an optimal manner.

Currently, there is a lack of research on the human skin microbiome and an optimised bacterial DNA extraction method specific to skin swabs has not been established. Therefore, there is a need to develop a method for this.

Due to the scale of the PEARL study, the development of a method compatible with downstream use of a high-throughput automated DNA extraction method using the Promega Maxwell® RSC Blood DNA Kit and Maxwell® RSC 48 Instrument (purchased for the study), is required.

Therefore, the purpose of this project is to develop a high-throughput automated method, using the Promega Maxwell® RSC Blood DNA Kit with the Maxwell® RSC 48 Instrument, for bacterial DNA extraction from skin swabs that results in DNA with high yield, quality, purity, and high molecular weight suitable for metagenomics sequencing.

1.1. Purpose

The purpose of this project is to develop a high-throughput automated method, using the Promega Maxwell® RSC Blood DNA Kit with the Maxwell® RSC 48 Instrument, for bacterial DNA extraction from skin swabs that results in DNA with high yield, quality, purity, and high molecular weight suitable for metagenomics sequencing.

1.2. Aims and Objectives

Aims

- To determine which established diluent, PBS or water, is more effective for collecting live skin bacteria from skin swabs for use in DNA extraction.
- To develop a high-throughput automated method for bacterial DNA extraction from skin swabs that results in DNA with high yield, quality, purity, and high molecular weight, suitable for metagenomics sequencing.

Objectives

- Compare bacterial growth from skin swabs processed and grown in PBS and water to determine which diluent is more effective for collecting live skin bacteria.
- Compare five bacterial DNA extraction methods from skin swabs, using the automated Maxwell® RSC 48 Instrument and the Maxwell® RSC Blood DNA Kit, to identify which method results in the highest DNA yield, quality, purity, and the highest molecular weight DNA.
- Analyse resulting DNA extracts for quantity, quality, purity and fragment length (molecular weight) using a QuBit and Tapestation instruments.
- DNA will be sequenced using an Illumina platform (NextSeq) to verify successful harvesting of bacterial DNA from the volunteers.
- Microbiota profiles will be resolved to identify the bacterial composition of samples using Whole Genome Sequencing (WGS) and established scripts in the Webber Lab. This will help determine which DNA extraction method provides a better representation of the skin bacterial composition.

2. Study Design

2.1. Recruitment policy

Volunteer participants will be recruited by an expression of interest email sent out to QIB staff by the Biorepository. The recruiters will be trained in consenting for the Biorepository. The recruiters will then provide participants with the established Biorepository participant information sheets and consent participants with the established Biorepository consent forms. To adhere to Covid19 regulations, participants will be consented either remotely or at a socially distanced location at QIB to reduce person-person contact. 24 participants will be recruited and each swabbed once to collect one skin swab per participant. If it is difficult to recruit 24 participants, 12 participants may be swabbed at two separate body sites to produce the 24 biological samples in total. The recruiters will swab participants to collect the samples. If the participants are consented to the study in person, their swabs will be collected at the same time. If the participant has been consented remotely, the recruiters will arrange a suitable day/time to meet the participant to collect these samples. We hope that we can condense these sampling sessions into a period of two days, with no more than 12 participants a day. Appointments will be staggered on each of these days to reduce social contact. All participants and those swabbing will wear a face mask, and those swabbing will also wear gloves,

to reduce person-person contact. In the case of reduced recruiting staff due to covid, participants will be consented remotely, as described above, and will self-swab at the QIB, using the PEARL study self-swabbing instructions. The participants will then return the swabs to the available recruiters the same day.

2.2. Eligibility criteria

2.2.1. Basic inclusion criteria

Participants must have no current skin conditions and have had no skin conditions or been prescribed antibiotics over the last 3 months.

2.2.2. Basic exclusion criteria

Participants with current skin conditions and have had skin conditions and been prescribed antibiotics over the last 3 months.

3. Sample collection and Handling

As this method development procedure is being performed for the purpose of optimising the DNA extraction method of the PEARL study skin swabs, swabs will be swabbed according to the PEARL study protocol which swabs the inner (volar) forearm of either the left or right arm with Charcoal Amies Transport Medium swabs. The protocol will be as follows:

To perform swabbing, pull and twist the swab out of its tube taking care not to touch the tip (swab end) with anything other than the area to be swabbed. Holding the swab in one hand, roll the opposite end (swab end) along the inside of the forearm from the wrist to the elbow pit ensuring that the whole of the swab end has come in to contact with the volunteers skin. Replace the swab back into the tube ensuring the lid is on tight. The swab may leave a grey residue on the forearm. Once sampling is complete, wash the area with soap and water. The date of sample collection and Biorepository sample ID will be written on the container in pen. Samples will be stored in a 4°C fridge until samples are collected for testing.

Samples will be collected and tested on the same day swabbing was performed.

4. Experimental methods

A review of the literature and protocols supplied by the CASE partner (Unilever) have been analysed to identify a potentially efficient DNA extraction method. The main variables to be tested are the initial diluent for swab suspension, type of ‘bead beating’ to disrupt bacterial cells, and different lysis protocols. The detailed protocol and variations to be tested are listed:

Initial diluent testing

The skin swab samples that were collected and stored in Charcoal Amies Transport Media will be processed in two diluents, PBS and water, then cultured to grow bacteria in order to determine which diluent is more effective for collecting skin

bacterial cells in the suspension (water may have additional lytic effect to PBS). When culturing the skin swabs, Colombia Blood Agar plates will be used as this is a known non-selective media suggested in literature for culturing a diversity of bacteria. To obtain a good representation of aerobic and anaerobic bacteria present on the skin, skin swabs will be cultured aerobically and anaerobically (Ogai et al. 2018). Figure 1 below demonstrates the initial diluent testing.

Two studies by Alan and Sallene 1998 and Bourbeau et al. 2007 have indicated that the agar in the Amies transport medium swabs inhibits molecular techniques such as PCR. Skin swabs will be washed twice to reduce the chance of the presence of transport media which could potentially impact metagenomics sequencing.

The processed swabs will then go through a series of five DNA extraction procedures using the Promega Maxwell® RSC Blood DNA Kit and the Maxwell® RSC 48 Instrument. This Promega Maxwell® kit and instrument permits a high-throughput automated genomic DNA isolation from up to 48 whole blood samples in 40 minutes which is ideal for processing large sample sets efficiently. The Maxwell® RSC 48 Instrument purifies samples using a paramagnetic particle to allow sample capture, washing, and purification of gDNA, and provides a higher bind capacity and cleaner eluate than traditional silica-based DNA purification systems (Sui et al. 2020; Moeller et al. 2014; Dunbar et al. 2018).

The Promega Maxwell® Kit and instrument was purchased as it allows high throughput and has been used to produce high molecular weight DNA (Mandrekar et al. 2010) which is important in metagenomic studies such as this (Bey et al. 2010).

As the manufacturers technical manual recommends the use of up to 300µl of sample, this volume will be used in the five experimental methods below.

Bead beating

Previous bacterial DNA extraction methods using the Promega Maxwell® instrument have incorporated a bead beating step using e.g. the Lysing Matrix E tubes and the MP Bio FastPrep instrument which is a high velocity bead beater. This protocol has been incorporated as it is an established protocol at the QIB for extracting bacterial DNA from human samples. The bead beating step has been suggested to be important for the lysis of bacteria, and certain microorganisms, such as Staphylococcus and Streptococcus which are vital skin probiotics, require a stronger homogenization using a high velocity bead beater. The Lysing Matrix E tubes contain 1.4mm ceramic spheres, 0.1 mm silica spheres and one 4 mm glass bead, and this range of bead sizes allows efficient lysis of diverse human and environmental sample types (Sui et al. 2020).

Lysing matrix E tubes are a suitable choice for the bead beating step in this protocol as they can be used on any bead-beating device according to the manufacturer's instructions. Two different high velocity bead beating instruments will be compared, the MP Bio FastPrep and the QIAGEN Tissue Lyser, to see if there is a difference in DNA yield, quality, molecular weight, and bacterial composition recovered after each. The QIAGEN Tissue Lyser is an established instrument used for a bead beating step in Unilever's skin bacterial DNA extraction protocol which uses the QIAGEN QIAsymphony instrument with the QIAGEN QIAsymphony Virus/Bacteria Midi Kit (931055). As the CASE partner of this PhD study, Unilever have provided access to their established skin bacteria DNA extraction procedure. However, this

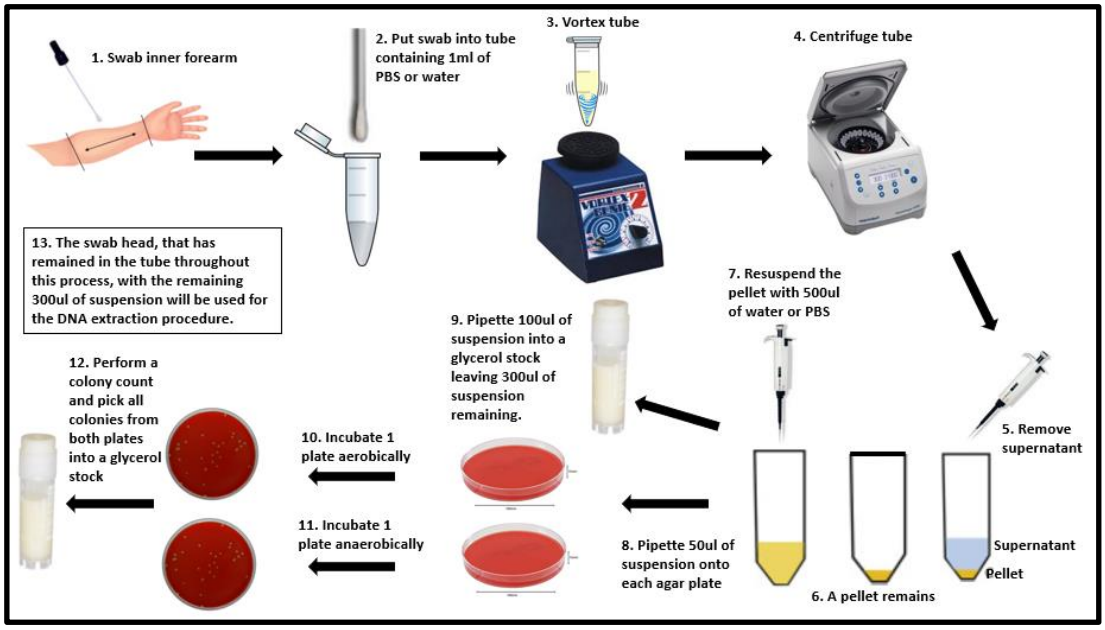
Unilever protocol is specific for 16s sequencing which allows for more fragmented DNA, thus there is a worry that using a high velocity bead beating instrument might cause very fragmented DNA not suitable for metagenomics. If the resulting DNA is too fragmented, the bead beating instrument could be replaced with a series of vortexing steps that other high molecular weight kits perform.

Lysis

The Unilever protocol also includes, an overnight heated lysis with the Epicentre Ready-Lyse lysozyme, and the QIAGEN QIASymphony off-board lysis procedure, which will be compared to the Maxwell lysis method to see if there is a difference in DNA concentration, quality, molecular weight and bacterial composition. The overnight heated lysis with a Lysozyme is vital for the lysis of certain bacteria such as Staphylococcus and Streptococcus. The Ready-Lyse lysozyme will also be compared to the cheaper Thermofisher lysozyme to see if this affects bacteria extracted. If there is no difference the Thermofisher lysozyme will be more cost-effective to use.

For each run a negative control containing PBS will be used to ensure no contamination has occurred, and a positive control containing a spiked suspension with an established mock community will be used to demonstrate that any negative results are a lack of material rather than an extraction problem. A recent study evaluated the composition of the mock microbial communities across four DNA extraction procedures, one being the Promega Maxwell gDNA which provided the best representation of the composition of the mock community (Sui et al. 2020).

Figure 1 below summarises the protocol and variations to be tested.



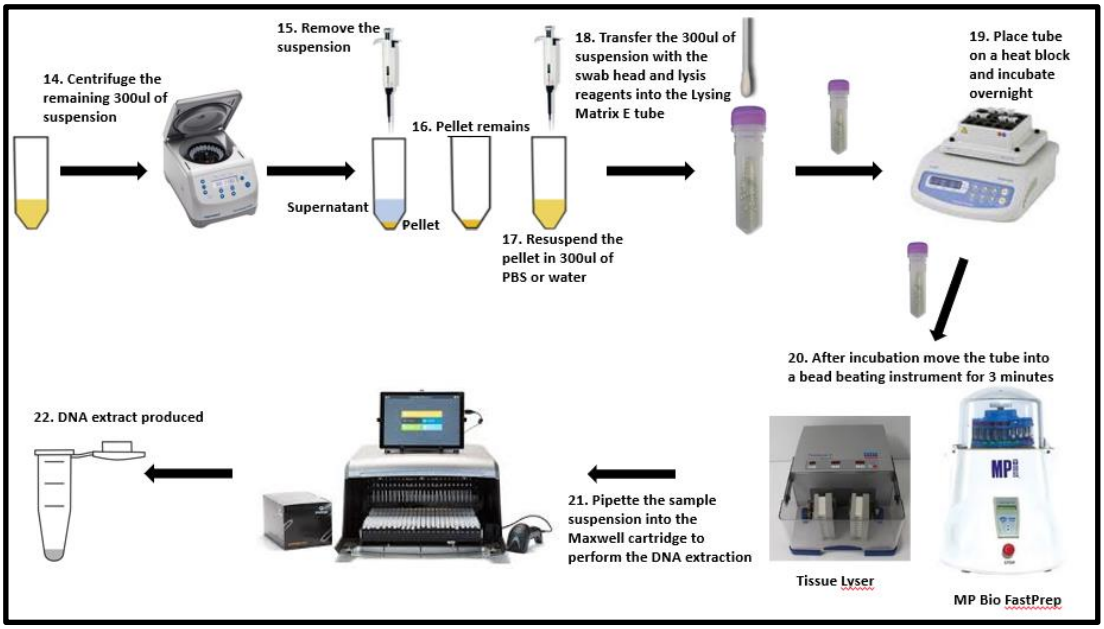


Figure 1 – Protocol and Variations to be Tested

Analysis of DNA

Resulting DNA extracts will then be analysed for quantity, quality and fragment length using a QuBit and Tapestation instruments (both available in QIB). DNA will then be sequenced using an Illumina platform (NextSeq) to verify successfully harvesting of bacterial DNA from the volunteers. WGS sequencing allows the microbiota to be resolved down to a species level. Scripts already established in the Webber lab will be utilised to resolve a microbiota profile to identify the bacterial composition of the samples. This will help determine which DNA extraction method provides a better representation of the skin bacterial composition.

Step by Step Procedure for Processing Skin Swabs and Performing a DNA Extraction

This procedure should be performed aseptically i.e. all samples, sterile reagents and equipment should be handled and processed in the Class II Microbiological Safety Cabinet, new pipette tips, spreaders, loops, scalpels & forceps should be used for each sample, good personal hygiene should be carried out such as washing hands before and after working in the lab, appropriate PPE worn such as orange shield gloves and lab coat, and work surfaces cleaned with Bio shield and then 70% ethanol before use, between samples, and after use, and sterile equipment used.
Processing Skin Swabs for DNA Extraction and Preparing Glycerol Stocks

1. For the 24 swabs taken, 12 will be processed in 1X PBS and 12 in Milli-Q water.
2. Appropriately label two Colombia Blood Agar plates for each swab (one plate for anaerobic culturing and one plate for aerobic culturing).
3. Appropriately label one 1.5ml Eppendorf tube for each swab.
4. Prepare the appropriate volume of 1X PBS from the 10X PBS stock solution using Milli-Q water. Add either 1ml of 1X PBS or Milli-Q water into the 1.5ml Eppendorf tubes.
5. Process the swabs by cutting each swab head into a 1.5ml Eppendorf tube containing either 1ml of 1X PBS or Milli-Q water.

6. Vortex the tubes @ full speed for 2 minutes.
7. Centrifuge tubes at 14,000 x g for 15 minutes to pellet the cells and remove the suspension.
8. For each pelleted sample, resuspend the cells in 500µl of 1X PBS or Milli-Q water.
9. Pipette 50µl onto the labelled aerobic agar plate, and 50µl onto the paired anaerobic plate.
10. Incubate the Aerobic plate in the 37° incubator and the Anaerobic plate in the 37° anaerobic cabinet for 3 days.
11. From the remaining 400µl of suspension, pipette 100µl into 1ml of 30% glycerol TSB prefilled in a 2ml cryovial to form a glycerol stock of the original processed skin cells. Freeze the glycerol stock in the -80°C freezer to cryopreserve the skin cells.
12. Centrifuge the remaining 300µl of suspension at 14,000 x g for 15 minutes to pellet the cells. This pellet will be used for DNA extraction. Proceed to the Skin Swab DNA Extraction Procedure below to begin the DNA extraction.
13. After incubation, perform a plate count by counting all bacterial colonies observed on each plate. Work out the CFU/ml by multiplying the number of colonies by the dilution factor (Neat in this case), then multiply by 10.
14. Once plate count has been performed, combine all colonies from the paired aerobic and anaerobic agar plates into a glycerol stock to cryopreserve the cultured skin microbiome. This can be performed by aseptically picking all colonies from both agar plates using a 10µl loop and inoculating them into 1ml of 30% glycerol TSB prefilled in a 2ml cryovial. Freeze the glycerol stock in the -80°C freezer.

Skin Swab DNA Extraction Procedure Using the Maxwell Blood DNA Kit and Instrument

NB: The total volume of reagents (330µl) + sample (300µl) used in the Maxwell manufacturers manual to process the samples before loading onto the instrument equals to 630µl. As DNA extraction methods 2, 3, 4 & 5 below use a total volume of reagents (643µl) + sample (400µl) totalling to 1043µl, which is much higher than the recommended Maxwell input volume, 630µl of the 1043µl centrifuged sample suspension will be added to well 1 of each cartridge. This leaves 413µl of sample leftover. Rather than discarding this, another extraction will be performed.

Method 1:

1. Resuspend the swab pellet in 300µl of PBS (using MilliQ H20 prepare the readymade 10x PBS to 1x PBS) or Milli-Q water and transfer the 300µl of sample with the original swab head into a 2ml Lysing Matrix E tube.
2. Add 30µl of Proteinase K and 300µl of Lysis Buffer to each sample and incubate in a heating block (set to 56°C) for 20 minutes.
3. Then perform a bead beating step using the FastPrep instrument programme 1 and run three times (3mins @ 6.0 FastPrep).
4. Proceed to step 5.

Method 2:

1. Resuspend the swab pellet in 300µl of PBS (using MilliQ H20 prepare the readymade 10x PBS to 1x PBS) or Milli-Q water and transfer the 300µl of sample with the original swab head into a 2ml Lysing Matrix E tube.

2. Add 30µl of Proteinase K and 300µl of Lysis Buffer to each sample and incubate in a heating block (set to 56°C) for 20 minutes.
3. Then perform a bead beating step using the Tissue Lyser instrument programme 1 and run three times (3mins @ 6.0 FastPrep).
4. Proceed to step 5.

Method 3:

1. Resuspend the swab pellet in 400µl of PBS (using MilliQ H20 prepare the readymade 10x PBS to 1x PBS) or water and transfer the 400µl of sample with the original swab head into a 2ml Lysing Matrix E tube, then add 3µl of Ready-Lyse lysozyme (Epicentre, 250U/µl; dilute to 250U/µl with TES buffer) and incubate with agitation at 300rpm, 37oC for 18 hours.
2. Following incubation, perform a bead beating step using the FastPrep instrument programme 1 and run three times (3mins @ 6.0 FastPrep) followed by an off-board lysis.
3. Off-board lysis is performed by adding 40 µl proteinase K, 165 µl Buffer ATL, 120 µl Carrier RNA, and 315 µl Buffer ACL into the 2 ml Lysing Matrix E tubes containing 400ul of sample and incubate the samples at 68 °C for 15 minutes.
4. Proceed to step 5; 630µl of the sample suspension will be added to well 1 of each cartridge. The remaining 413µl will be added to well 1 of a second cartridge.

Method 4:

1. Resuspend the swab pellet in 400µl of PBS (using MilliQ H20 prepare the readymade 10x PBS to 1x PBS) or Milli-Q water and transfer the 400µl of sample with the original swab head into a 2ml Lysing Matrix E tube, then add 3µl of Ready-Lyse lysozyme (Epicentre, 250U/µl; Manufacturer's instructions dilute to 250U/µl with TES buffer) and incubate with agitation at 300rpm, 37oC for 18 hours.
2. Following incubation, perform a bead beating step using a Tissue Lyser instrument for 3 minutes at 20 Hz followed by an off-board lysis.
3. Off-board lysis is performed by adding 40 µl proteinase K, 165 µl Buffer ATL, 120 µl Carrier RNA, and 315 µl Buffer ACL into the 2 ml Lysing Matrix E tubes containing 400ul of sample and incubate the samples at 68 °C for 15 minutes.
4. Proceed to step 5; 630µl of the sample suspension will be added to well 1 of each cartridge. The remaining 413µl will be added to well 1 of a second cartridge.

Method 5:

1. Resuspend the swab pellet in 400µl of PBS (using MilliQ H20 prepare the readymade 10x PBS to 1x PBS) or Milli-Q and transfer the 400µl of sample with the original swab head into a 2ml Lysing Matrix E tube, then add 3µl of Thermofisher Lysozyme and incubate with agitation at 300rpm, 37oC for 18 hours.
2. Following incubation, perform a bead beating step using the FastPrep instrument programme 1 and run three times (3mins @ 6.0 FastPrep) followed by an off-board lysis.
3. Off-board lysis is performed by adding 40 µl proteinase K, 165 µl Buffer ATL, 120 µl Carrier RNA, and 315 µl Buffer ACL into the 2 ml Lysing Matrix E

tubes containing 400ul of sample and incubate the samples at 68 °C for 15 minutes.

4. Proceed to step 5; 630µl of the sample suspension will be added to well 1 of each cartridge. The remaining 413µl will be added to well 1 of a second cartridge.

Method 6:

1. Resuspend the swab pellet in 400µl of PBS (using MilliQ H20 prepare the readymade 10x PBS to 1x PBS) or Milli-Q water and transfer the 400µl of sample with the original swab head into a 2ml Lysing Matrix E tube, then add 3µl of the Thermofisher Lysozyme and incubate with agitation at 300rpm, 37oC for 18 hours.
2. Following incubation, perform a bead beating step using a Tissue Lyser instrument for 3 minutes at 20 Hz followed by an off-board lysis.
3. Off-board lysis is performed by adding 40 µl proteinase K, 165 µl Buffer ATL, 120 µl Carrier RNA, and 315 µl Buffer ACL into the 2 ml Lysing Matrix E tubes containing 400ul of sample and incubate the samples at 68 °C for 15 minutes.
4. Proceed to step 5; 630µl of the sample suspension will be added to well 1 of each cartridge. The remaining 413µl will be added to well 1 of a second cartridge.

- - - -

14. When the automated purification run is complete, the LCD screen will display a message that the method has ended.
15. Follow on-screen instructions at the end of the method to open door.
16. Verify that plungers are located in well 8 of the cartridges at the end of the run. If plungers are not removed from the magnetic plunger bar, push them down gently by hand to remove them.
17. Press the Run/Stop button to extend the platform out of the instrument.
18. Remove the Maxwell® RSC Cartridge Rack from the instrument, remove Elution Tubes containing DNA, cap tubes, and store in -80°C freezer.

5. Statistical Analysis

As this protocol is for the purpose of method development and determining optimum conditions for DNA extraction data no statistical power 'effect' is being examined.

6. Storage of Processed Samples

There is no need to store any of these samples in the long term – DNA preparations will be kept for 12 months in case any further methodological verification in terms of sequence library preparation is needed but will then be destroyed according to established waste disposal procedures within QIB.
Bacterial DNA sequences will be kept on secure servers within QIB (and anonymised from donor as described below).

7. Data protection and Donor confidentiality

There will be no contact between researcher and participant. The Biorepository members of staff will remain the link between the researcher and participant. The Biorepository will also label all swabs with a Biorepository sample ID completely unlinked to the participants personal information to eliminate the chance of identifiable information being given to the researcher. Samples will be collected by the Biorepository staff members who will ensure its anonymity prior to being given to the researcher. The Biorepository contact information will be provided to participants for them to express any questions or concerns regarding the study. The Biorepository can then contact the researcher for further clarification.
Any DNA reads mapping to the human genome will be identified and destroyed (deleted) bioinformatically in the initial analysis following an established procedure within QIB.

8. Ethical and regulatory considerations

As this method development will go through the Biorepository, swabbing must be anonymised, and no samples can be stored therefore swabs must be discarded after use. To adhere to this, the Biorepository will recruit, consent, and swab volunteers, and samples will be collected from the Biorepository and tested on the same day the swabbing took place. Further to this, after the aims have been answered, the host DNA and resulting host sequences will be discarded.

9. Reference

Bey, S.B., Fichot, E.B., Dayama, G., Decho, A.W. and Norman, R.S. (2010). Extraction of High Molecular Weight DNA from Microbial Mats. BioTechniques 49(3):631-640.

Bourbeau, P.P., Riley, J.A., Shoemaker, B.C., and Jones, K.S. (2007). Use of CultureSwab Plus Swabs with Amies Gel Agar for Testing of Naris Specimens with the GeneOhm MRSA Assay. Journal of Clinical Microbiology 45(7): 2281-2283.

Dunbara, J., Gallegos-Gravesa, LV., Gansa, J., Morseb, S.A., Pillaic, S., Andersond, K. and Hodged, D.R. (2018). Evaluation of DNA Extraction Methods to Detect Bacterial Targets in Aerosol Samples. Journal of Microbiological Methods 153:48–53.

Gibb, P.A. and Wong, S. (1998). Inhibition of PCR by Agar from Bacteriological Transport Media. Journal of Clinical Microbiology 36(1):275-276.

American Medical Association, Manual of Style, 10th edition, 2007.
Mandrekar, P. V., Ma, Z., Krueger, S. and Cowan, C. High-Concentration (>100ng/µl) Genomic DNA From Whole Blood Using the Maxwell® 16 Low Elution Volume Instrument. Promega Corporation Web site.
https://www.promega.co.uk/resources/pubhub/high-concentration-genomic-dna-from-whole-blood-using-the-maxwell-16-low-elution-volume-instrument
Updated 2010. Accessed: November 29th 2020.

Moeller, J.R., Moehn, N.R., Waller, D.M. and Givnish, T.J. (2014). Paramagnetic Cellulose DNA Isolation Improves DNA Yield and Quality Among Diverse Plant Taxa. Applications in Plant Sciences 2(10): 1400048.

Ogai K., Nagase S., Mukai K., Iuchi T., Mori Y., Matsue M., Sugitani K., Sugama J. and Okamoto S. (2018). A Comparison of Techniques for Collecting Skin Microbiome Samples: Swabbing Versus Tape-Stripping. Front. Microbiol. 9:2362.

Sui, H., Weil, A.A., Nuwagira, E., Qadri, F., Ryan, E.T., Mezzari, M.P., Phipatanakul, W. and Lai, P.S. (2020). Impact of DNA Extraction Method on Variation in Human and Built Environment Microbial Community and Functional Profiles Assessed by Shotgun Metagenomics Sequencing. Front. Microbiol. 11:953.

Promega Maxwell® RSC Blood DNA Kit.
https://www.promega.com/-/media/files/resources/protocols/technical-manuals/101/maxwell-rsc-blood-dna-kit-protocol.pdf?la=en

DNA Extraction Using MPbio Fast DNA Spin Kit for Soil QIB Hall Lab SOP

Unilevers Skin Bacteria DNA Extraction Protocol

QIAsymphony® SP Protocol Sheet - Complex400_OBL_V4_DSP protocol (user-validated). This document is the Complex400_OBL_V4_DSP QIAsymphony SP Protocol Sheet, R2, for QIAsymphony Virus/Bacteria Midi Kit.
https://www.qiagen.com/us/resources/resourcedetail?id=9176b1ba-60ac-4399-834f-015dca2091a2&lang=en

Epicentre Ready-Lyse™ Lysozyme Solution Manufacturer Instruction (Cat. Nos. R1802M, R1804M, and R1810M) –

https://us.vwr.com/assetsvc/asset/en_US/id/22292169/contents

Appendix 4. HRGC Human Tissue Registration Form

Quadram Institute
Science • Health • Food • Innovation

Human Research Governance Committee (HRGC) Registration Form for Use of Human Tissue Samples Provided by Norwich Research Park Biorepository

1. DETAILS OF STUDY	
Study Title:	Method Development for Bacterial DNA Extraction from Skin Swabs
Summary of research question and relevant background information:	My PhD project will study how bacteria on the skin change through early life. This will use samples from the 'PEARL' project which has already been established and has full ethical approval (Lindsay Hall, the PEARL PI is my second supervisor). However, before working with any PEARL samples I want to optimise a method for extracting bacterial DNA from swabs taken from human skin. Currently, there is a lack of research on the human skin microbiome and an optimised bacterial DNA extraction method specific to skin swabs has not been established. Therefore, there is a need to develop a method for this. Due to the scale of the PEARL study, the development of a method compatible with downstream use of a high-throughput automated DNA extraction method using the Promega Maxwell® RSC Blood DNA Kit and Maxwell® RSC 48 Instrument (purchased for the study), is required. Therefore, the purpose of this project is to develop a high-throughput automated method, using the Promega Maxwell® RSC Blood DNA Kit with the Maxwell® RSC 48 Instrument, for bacterial DNA extraction from skin swabs that results in DNA with high yield, quality, purity, and high molecular weight suitable for metagenomics sequencing.
Planned start date:	May 2021
Planned end date:	May 2022
Funding source:	CTP iCASE Project (Unilever)
Quadram (QIB) Programme	QIB Project number if available: 42971-000-B

2. DETAILS OF QIB PRINCIPAL INVESTIGATOR (PI)	
Name:	Dr. Mark Webber
Telephone:	01603 255233
E-mail:	Mark.Webber@quadram.ac.uk
Person running study if not PI:	Iliana Rosa Serghiou

3. HUMAN TISSUE SAMPLE INFORMATION	
Type of sample:	Blood Urine Tissue – Skin Swabs Faeces Saliva

	Primary cell culture Other (*please give details*) Fresh ☐ Archived ☐ Frozen ☐
4. HUMAN DATA INFORMATION	
Type of data:	• Non-clinical data. No data from the participants is needed for this method validation study
Give details of how data will be accessed and custodial arrangements.	There will be no contact between researcher and participant. The Biorepository members of staff will remain the link between the researcher and participant. The Biorepository will also label all swabs with a Biorepository sample ID completely unlinked to the participants personal information to eliminate the chance of identifiable information being given to the researcher. Samples will be collected by the Biorepository staff members who will ensure its anonymity prior to being given to the researcher. The Biorepository contact information will be provided to participants for them to express any questions or concerns regarding the study. The Biorepository can then contact the researcher for further clarification. Host DNA and any DNA reads mapping to the human genome will be identified and destroyed in the initial analysis following an established procedure within QIB.

5. NRP BIOREPOSITORY INFORMATION	
Name and address of the organisation providing samples/data:	Norwich Research Park Biorepository Norfolk and Norwich University Hospital Bob Champion Research & Education Building Norwich NR4 7UQ
Point contact:	Roxanne Brunton-Sim
E-mail/telephone:	roxanne.brunton-sim@nnuh.nhs.uk/1603 289221
Other organisations involved in this study:	N/A
How and by whom will potential participants first be approached?	Participants will first be approached by an expression of interest email sent out to QIB staff by the Biorepository. Biorepository staff contact details: • Roxanne Brunton-Sim - roxanne.brunton-sim@nnuh.nhs.uk/1603 289221 • Rachael Stanley - rachael.stanley@nnuh.nhs.uk/1603 289221
Will you obtain informed consent from research participants? *If you are not obtaining consent, please explain why not.*	To ensure anonymity between researcher and participants, the Biorepository will consent participants with the established Biorepository consent forms. Biorepository staff contact details: Roxanne Brunton-Sim - roxanne.brunton-sim@nnuh.nhs.uk/1603 289221 Rachael Stanley - rachael.stanley@nnuh.nhs.uk/1603 289221
Give details of where the samples will be processed, who will have access and the custodial arrangements.	Samples will be processed on the day of sample collection in the Tissue Processing Laboratory (LG65) at the Quadram Institute. The sample collection and processing will be condensed into a period of 2 days. Contact details of person responsible for sample handling and processing: Iliana Rosa Serghiou – Iliana.serghiou2@quadram.ac.uk/01603 255034
Will the research involve the analysis or use of human DNA in the samples?	If yes, please confirm compliance with consent requirements The research will involve the analysis and use of bacterial DNA from the human skin swabs. Some human DNA is likely to be present although this will be removed when analysing sequence

	data, or directly removed using a host depletion method. After the aims have been answered, all samples will be discarded.
What will happen to the samples at the end of the research?	Disposal in accordance with the Human Tissue Authority's Code of Practice
Have you discussed this proposal with the NRP Biorepository Manager?	**Authorisation on behalf of NRP Biorepository** Signature of NRP Biorepository Manager:  Date: 11.03.21

ETHICS APPROVAL	
Name of the Research Ethics Committee (REC) approving sample collection and use for this study:	• UEA FMH *REC reference number - 2020/21-065*
Date of favourable ethical opinion:	18.01.2021
End date of approval:	N/A

6. Training requirements

Team Member Name:	Allocated Time to this research (FTE)	QIB Human Studies Training Programme	Good Clinical Practice (GCP) Training	Human Sample Handling Training	Informed Consent Training	Other
Iliana Rosa Serghiou		Yes/No/NA Date:	Yes/No/NA Date: 28/10/2020	Yes/No/NA Date: 28/10/2020	Yes/No/NA Date: 28/10/2020	
		Yes/No/NA Date:	Yes/No/NA Date:	Yes/No/NA Date:	Yes/No/NA Date:	

All named staff must attend relevant modules of the Human Studies Training Course and prepare the necessary SOPs prior to commencing work on the study. Please contact the QIB Human Study Lead for further details.

7. DECLARATION	
Signature of QIB Human Study Lead:	
Date:	11th March 2021
Signature of PI:	
Date of Submission:	11.3.21

Please submit the completed form to the HRGC Chair along with copies of the UEA FMH or HRA approval documents and study protocol, information sheet(s) and consent form(s).

HRGC ACKNOWLEDGEMENT

This study and associated tissue samples have been registered with the QIB Human Research Governance Committee

Signature of HRGC Chair: .....

Print name: Dr Antonietta Hayhoe

Date: 11th March 2021

Comments: QIB HRGC reference number QIB-T45

FMH has approved two amendments (N1, March 2021; N2 July 2021).

Appendix 5. PIS and CF

Our Vision
To provide every patient with the care we want for those we love the most

Norfolk and Norwich University Hospitals NHS
NHS Foundation Trust

The Norwich Biorepository

The donation, collection, storage and use of samples of tissue and/or fluids and/or other material from a healthy adult donor for research

Information sheet for healthy donors - Version 15 (21 February 2014)

Thank you for considering giving a sample for biomedical research. This information sheet provides a brief summary to help you to understand what this means and involves.

There is a consent form after the information sheet. It is important that you complete and sign it, if you decide to give a sample. Please complete all parts of the consent form.

Doctors and other health professionals may take samples (which may be blood, other fluids, small biopsies, or something else) from patients to help make a diagnosis and decide how best to treat them.

They may also ask to take similar samples from healthy donors, like you, specifically for research purposes. They use those samples to learn more about health and/or illness, how disease happens and how to treat it, and sometimes to help develop new medicines.

For the purposes of medical research you do not need to be in perfect health. We may need to ask you some questions to confirm that you do NOT suffer from the medical condition which is being studied.

Samples from healthy volunteers, like you, will include only those that can be obtained externally, or by normal routes (e.g., saliva, urine, faeces), or by minimal invasion involving very little risk to the donor (e.g., blood taken from a vein close to the surface, or a throat swab). Such minimally invasive procedures include those that might be undertaken during a routine visit to a general practitioner. Nothing more invasive than that is permitted using this consent form.

Samples donated (given) to the Norwich Biorepository are not:

- Normally used in animal research. It will be made clear to you if animal research is an integral part of the project for which we are seeking a donation.

Continued.....

- Used in cloning experiments. However, the Biorepository would consider the use of donations in non-reproductive cloning experiments based on their scientific value and in the context of prevailing law and ethical standards. It will be made clear to you if cloning experiments are part of the research project for which we are seeking a donation.

To undertake research on the sample(s) that you are considering donating, we need your permission and signed consent.

If you give permission for a sample to be taken –

- The Hospital will own the sample
- The sample may be stored, usually in a deep freezer, until it is used. The freezer is referred to as a tissue bank in the consent form
- Nobody involved in the research will know where the sample has come from.
- The sample will be used only in experiments that are ethical and to help other people. Please see the section entitled 'Scientific and ethical approval' below to understand what we mean by *ethical*.
- Your donated sample(s) and any genetic material derived from it (them) may be stored indefinitely for future research projects, which may include whole genome sequencing. *Whole genome sequencing* means reading your total DNA code (your genetic blueprint) in a single assessment.
- We might give some or all of the sample to other doctors or researchers for their experiments, if they are ethical and to help other people. Some of these people might work in companies in this country or abroad
- Data derived from your sample(s) may be placed anonymously in an international database for future research. While we will take all possible steps to maintain your anonymity and protect your privacy, there is a very small risk that genetic information produced in the research and stored on databases could lead to your identification by being linked to other stored information.
- We will keep some facts about you on our Biorepository database
- Although these facts might be given to the research doctors or scientists to help their experiments, we will NOT tell them your name or other details that would let them know who you are
- Doctors in the Hospital might also read your hospital records to help them understand what the doctors or researchers find out in the experiments. This is possible because your hospital records can be linked to the anonymous research sample without loss of confidentiality as far as the researchers are concerned. If the research results are important for you, it might be possible using this linkage to feed back the information to your doctor, so that any appropriate action can be considered.

The next sections give more detailed information. If you have any questions, please ask the person who is asking for your consent.

Continued.....

WHAT WILL HAPPEN

Tissue, blood or other samples taken from you will be sent either to the Norwich Biorepository or to a research laboratory. **Only as much sample as is needed for research will be removed.**

The donation of a sample for research is designed to be as safe as possible. It is most unlikely that you will come to any harm as a result of this donation (though we cannot give any guarantees). Potential physical problems depend on the type of donation. They might include, for example, discomfort or pain, bleeding, or infection. **The person collecting the donation will be properly trained in the procedure. You will be asked to give your written consent before donation.**

While **we will take all possible steps to maintain your anonymity and protect your privacy**, there is a very small risk that genetic information produced in the research and stored on databases could lead to your identification by being linked to other stored information.

MEDICAL RESEARCH AND WHY THIS PROGRAMME IS IMPORTANT

Some of your sample, or material extracted from it, will be stored in a local tissue bank. This is part of a research programme which now includes the Norfolk and Norwich University Hospitals NHS Foundation Trust, the James Paget University Hospitals NHS Foundation Trust, the University of East Anglia (UEA), and the Institute of Food Research (IFR). Your donation may be used by ourselves or by researchers from other centres at a later date. Some of this research may involve an assessment of genetic material (DNA and/or RNA) to help us understand the genetic basis of health and disease.

The purpose of this research is to understand more about human health and disease. It may also allow us to develop new methods of disease prevention or new treatments for the benefit of future patients. Some of these research programmes could lead to the development of new products and processes, which may be developed commercially for the improvement of patient care. In that case, there would be no financial benefit to you.

Medically qualified doctors or other suitably qualified staff at the hospital may need to review your hospital case records, including hospital notes, if any, to help understand how the research findings made by other doctors or researchers using your donated samples fit with what is known of your medical history. It may be important to be able see how their research findings relate to past events in your health record. The hospital doctors will not give your name to those doing the research.

The research may also involve training doctors and scientists in scientific medicine, and may lead to higher qualifications for them (e.g., PhD or MD degrees). This is important for future research into diseases and for looking for new, more effective, treatments for them.

LINKS WITH OTHER ORGANISATIONS

If you agree, we may send stored samples or products derived from them to other approved tissue banks or companies in this country or abroad. This would be to support their research programmes or the research programmes of those companies' clients.

Continued……

Such outside organisations will provide financial support for the Norwich Biorepository (our tissue bank), to help it recover its operating costs. We are not, however, allowed to sell the stored samples in order to make any financial profit from these commercial links.

SCIENTIFIC AND ETHICAL APPROVAL

The Norwich Biorepository acts as a custodian of the samples it holds. It releases them only to individuals or organisations that have an acceptable scientific background and work to high ethical standards. We require that all such medical research has been approved by a properly constituted Research Governance Committee before it starts. It must also be approved by a Research Ethics Committee or on behalf of the Research Ethics Committee that oversees the work of the Norwich Biorepository under the terms of the Biorepository's own Research Ethics Committee approval. That committee is the Cambridge East Research Ethics Committee. These committees look particularly at the purpose and validity of the research proposal, the welfare of any participants and issues of consent and confidentiality. We will release samples to commercial companies only if they work to appropriate ethical and scientific standards.

YOUR RIGHTS

When your samples are obtained, some information about you will be kept on a computer in the Norwich Biorepository. This will help us understand how what we find in the laboratory relates to you as a person. You are entitled to ask to see what is recorded about you by applying to the Chairman of the Norwich Biorepository Committee, Norfolk & Norwich University Hospital, c/o Dept. of Histopathology, The Cotman Centre, Colney Lane, Norwich, NR4 7UB. No one other than you has the right to see these records. Any information needed for research purposes will be made anonymous before it is given to the researcher.

The researchers will not be able to find out your name or any personal details about you from the information that they receive.

You will have the opportunity to discuss with a doctor or researcher issues relating to the possible use of your samples for research purposes now or at the time the sample(s) is (are) collected. He or she will answer any questions you may have.

MAKING A DONATION (GIFT) OF A SAMPLE FOR RESEARCH

If you decide that you want your sample to be donated for research purposes, you will be asked to sign a special consent form. This will confirm your decision and state that you have read and understood this information sheet. When you sign the form you will give the ownership of the sample(s) to the Norfolk & Norwich University Hospitals NHS Foundation Trust. The sample(s) will then belong to the Trust. It will store the sample(s) for an indefinite period of time and will able to decide how it (they) should best be used for research. It will also have the right to dispose of unused stored material in an appropriate legal and ethical manner following normal procedures.

If you do not want to donate a sample to be stored in the tissue bank or used for research, please tell us and do NOT sign the special Consent Form. If you do not sign this form, the donation will not proceed.

Continued.....

Complete the following details:

Donor's name.....

Date of birth.....

Address .....

.....

.....

The Norwich Biorepository

Consent for the collection, storage and release of human samples for research

I agree (Please initial small box) that the following sample(s) may be used for research, including genetic (DNA and/or RNA) studies and for the possible development of commercial products for the improvement of patient care, from which I would receive no financial benefit:

List sample(s) for research:	
	☐

I also agree that (Please initial small boxes, as appropriate):

These samples become the property of the Norfolk & Norwich University Hospitals NHS Foundation Trust ("the Trust") ☐

The Trust may store these samples in a tissue bank / biorepository ☐

The Trust may use these samples at its discretion in properly approved research programmes ☐

The Trust may pass on these samples to other approved tissue banks and/or companies, which may be in this country or abroad, in properly approved research programmes

Yes ☐
No ☐

My genetic material and donated sample(s) may be stored for an indefinite amount of time for future research projects, which may include whole genome sequencing ☐

Information about my case may be kept on the Norwich Biorepository database ☐

Anonymous data derived from my sample(s) may be placed in an international database for future research ☐

Such information may be passed in an anonymous form to persons outside the Trust In connection with research and may be published with any research findings ☐

I agree that appropriately qualified staff employed by the Trust may review my hospital records, including case notes, as appropriate, for the purposes of research using the donated samples ☐

Continued.....

Complete the following details:

Donor's name.....

Date of birth.....

Address .....

.....

.....

The Norwich Biorepository

Consent for the collection, storage and release of human samples for research

I also agree that (Please initial small boxes, as appropriate):

These samples may be used in ethically approved animal research	Yes / No
These samples may be used in ethically approved cloning research	Yes / No

I confirm that:

1) **I have read and understand the Information Sheet for healthy donors, Version 15, dated 21 February 2014**
2) **The issues have been explained to me, and that I have had the opportunity to ask questions.**

Signed _____ **(Donor)** **Date** _____

I have explained the request for sample for research purposes and have answered such questions as the donor has asked.

Signed_____ **Print name** _____

Doctor / Nursing Practitioner / Researcher / *Other*_____
(Please delete as appropriate / indicate Other status)

Date _____

Appendix 6. Stool Positive Controls

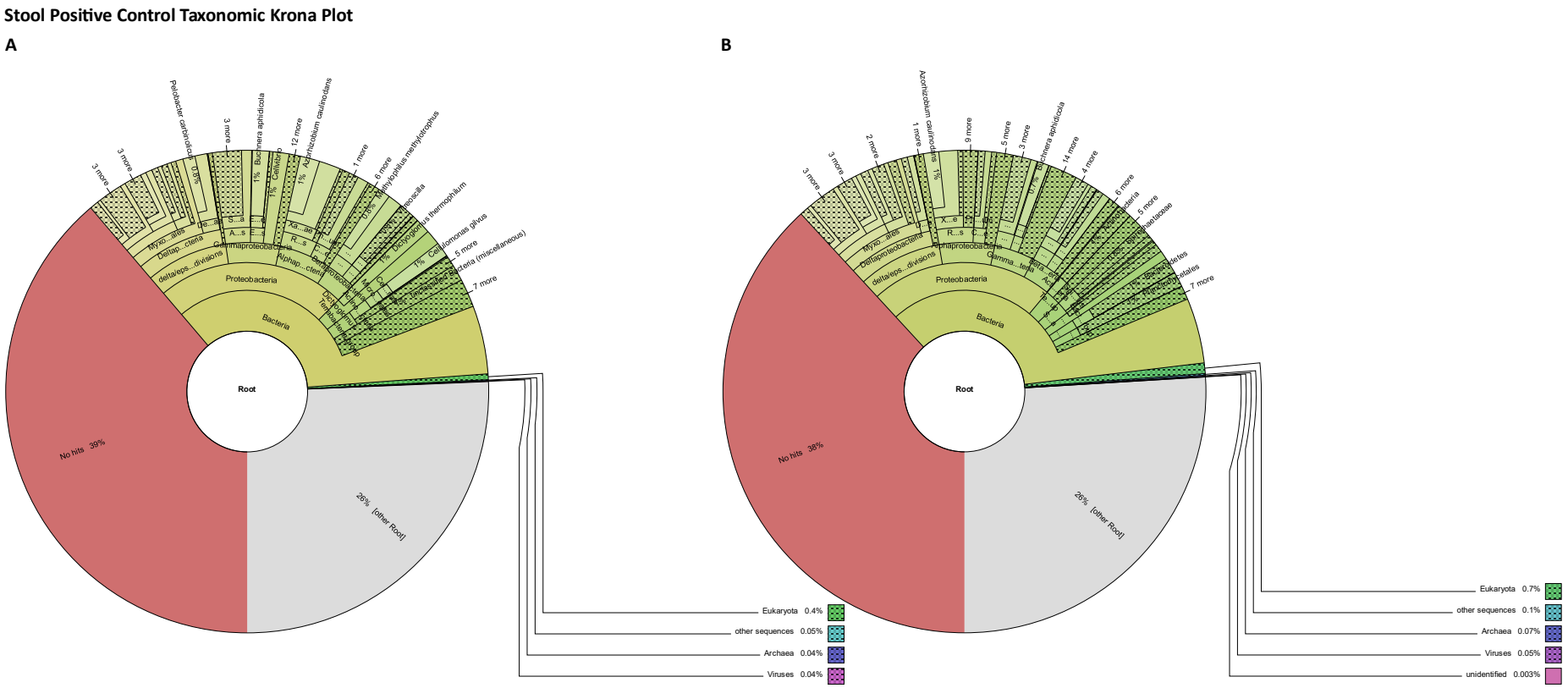


Figure Appendix 6. Stool Positive Control Taxonomic Krona Plot. Taxonomic profiles of the PEARL study stool positive controls. A: Taxonomic profile for the M004Base positive control. B: Taxonomic profile for the M008Base positive control. For the interactive version view links: M004Base - Krona - Root and M008Base - Krona - Root (the interactive version is also available in the supplementary files).

Appendix 7. Processing, Culturing and Extracting Bacterial DNA from PEARL Study Skin Swab Samples

QUADRAM INSTITUTE BIOSCIENCE

STANDARD OPERATING PROCEDURE

Quadram Institute
Science • Health • Food • Innovation

Title: Processing, Culturing and Extracting Bacterial DNA from PEARL Study Skin Swab Samples Using the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit

***Applies to staff in**: PEARL Study*	***Reference number:** 241880-LH-033*
***Author**: Iliana Rosa Serghiou*	***Approver:** Dr. Lindsay Hall*
***Issue date**: 28 January 2022*	***Review date**: 27 January 2024*
QA and H&S Review:	***Version Number:** 001*

AUTHORISATION: I confirm that this SOP is an accurate description of the procedure or method in use.

DECLARATION. The safety information contained in this document is to the best of my knowledge an accurate statement of the known and foreseeable hazards to health and of the safety precautions, which are to be followed to minimise the risks to health in the process. All users of this SOP are adequately trained to be competent in the process.

Approved by ISP/NC Leader, or Research Leader, or Senior Scientist, or Senior Administrator

Signed:	Print: Dr. Lindsay Hall	Date: 28 January 2022

TITLE: PROCESSING, CULTURING AND EXTRACTING BACTERIAL DNA FROM PEARL STUDY SKIN SWAB SAMPLES USING THE PROMEGA MAXWELL® RSC 48 INSTRUMENT AND RSC BLOOD DNA KIT

1. PURPOSE OF PROCEDURE/METHOD AND ITS SCOPE

- Perform a bacterial DNA extraction on skin swabs.
- Prepare glycerol stocks of the skin cells collected off the swabs.
- Culture skin microbiota from the skin swabs and prepare glycerol stocks of the cultured skin microbiota.
- Perform a DNA extraction on the cultured skin microbiota glycerol stocks.

List of current studies:

- PEARL: Skin swab collection was approved by Research Ethics Committee (REC reference: 18/LO/1703) under the IRAS approval reference number 241880. (SOP: 241880-LH-007).

2. EQUIPMENT AND REAGENTS NEEDED

Equipment

- Class II Microbiological Safety Cabinet (follow manufacturer instructions on how to use)
- Vortex (follow manufacturer instructions on how to use)
- Centrifuge (follow manufacturer instructions on how to use)
- 10µl Pipette and filter tips
- 20-200µl Pipette and filter tips
- 100-1000µl (1ml) Pipette and filter tips
- 10µl Loops
- Spreaders

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

- Scalpels
- 1.5ml Eppendorf Tubes
- 2ml Lysing Matrix E Tubes
- 2ml Cryovials
- Tube Rack
- Plate Rack
- -80°C Freezer
- Maxwell® RSC 48 Instrument (follow the Maxwell RSC 48 Instrument Operating Manual TM510 – View Related Procedures 1 at the end of the document)
- Qiagen Tissue Lyser (follow manufacturer instructions on how to use)
- Heating Block (follow manufacturer instructions on how to use)
- 37°C Anaerobic Cabinet (follow manufacturer instructions on how to use)
- 37°C Incubator (follow manufacturer instructions on how to use)

Reagents

- Colombia Blood Agar + 5% Horse Blood
- 10X PBS Stock Solution
- Milli-Q Water
- 40% Glycerol

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

- Skin Swabs
- Maxwell® RSC Blood DNA Kit (follow the Maxwell RSC Blood DNA Kit TM419 Full Protocol and Technical Manual / Processing Blood Samples Using the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit SOP – View Related Procedures 2 and 3 at the end of the document)
- Thermo Fischer Lysozyme (at 250U/ul)
- SIGMA-ALDRICH Tris-EDTA Buffer Solution
- Carrier RNA and AVE Buffer from QIAamp Viral RNA Mini Accessory Set
- QIAGEN Buffer ATL
- QIAGEN Buffer ACL
- ATCC MSA-2005 - Skin Microbiome Whole Cell Mix

3. STEPS IN PROCEDURE (INCLUDING SAFETY INFORMATION)

NB:

- This SOP was composed using the "Maxwell RSC Blood DNA Kit TM419 Full Protocol and Technical Manual" (view Related Procedures 2 at the end of the document) and "Processing Blood Samples Using the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit SOP" (view Related Procedures 3 at the end of the document) with alterations to enhance the extraction of bacterial DNA from skin swabs.
- Before beginning this procedure, please refer to the 'QIB Biosafety Handbook' and 'Code of Practice for Work with Human Body Tissues, Fluids and Other Samples of Human Origin' for information on local procedures for working with biological samples.
- All work should be conducted in a microbiological safety cabinet within a containment level 2 laboratory.
- Hands should be washed before and after working in the lab.
- Approved disinfectants should be available in case of a leak or any spillage:
    - Work surfaces should be cleaned with Bio Shield and then 70% Ethanol before and after use.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

- Specimen or reagent spills on any part of the Maxwell Instrument should be cleaned with a detergent-water solution, followed by a bactericidal spray or wipe, then water. Do not use bleach on any instrument parts.

- After each run, the maxwell should be cleaned and a UV cycle run.
- A lab coat and protective gloves should be worn whenever working in the laboratory with human samples.

Ethics Clearance and QIB Governance:

- The proposed use of human samples for this method will be registered with the QIB Human Research Governance Committee before samples are received on site. The QIB Investigator will be responsible for the completion of the relevant QIB Human Research Governance Committee Registration Form for Use of Human Tissue Samples.
- Long-term storage will be allowed only for samples with appropriate NHS HRA approval.
- Personnel at QIB working with human material should have Hepatitis B immunisation and other relevant vaccinations following NBI Occupational Health Advice.
- Personnel at QIB working with human material should also have attended training on working with human samples through the QIB Human Studies Training program.
- Sample disposal will be carried out in accordance with the HTA's Codes of Practice and relevant QIB policy (reference to 'QIB Biological Handbook', 'Code of Practice for Work with Human Body Tissues, Fluids and Other Samples of Human Origin).
- Clinical waste should be triple bagged, with the use of a yellow outer bag; A new clinical waste bag should be set up before work is started. Non-clinical biological waste should be double bagged.
- All pipette tips and other sharps should be put into a clinical waste sharps bin.
- All waste should be sealed with the room number marked on the bag/bin and should be sent for autoclaving at the end of the working day.

Processing Skin Swabs for DNA Extraction, Culturing, and Preparing Glycerol Stocks:

1. Appropriately label one 1.5ml Eppendorf tube for each swab.
2. Prepare the appropriate volume of 1x PBS from the 10x PBS stock solution using Milli-Q water. Add 1ml of 1x PBS to the labelled 1.5ml Eppendorf tubes for each swab.
3. For each run use a negative control containing PBS to ensure no contamination has occurred and use the ATCC Skin Microbiome Whole Cell Mix as a positive control, which is an established mock community, to demonstrate that any negative results are a lack of material rather than an extraction problem.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

4. Prepare the ATCC Skin Microbiome Whole Cell Mix using the manufacturer instructions (view Related Procedures 4 at the end of the document):
    a. Add 1 mL of chilled 1x PBS to the lyophilized pellet (chill the PBS in the fridge before beginning).
    b. Allow the pellet to dissolve for a minimum of two minutes; keep the vial on ice to prevent cell lysis (this vial contains $1.2x10^{-8}$ cells).
    c. Mix the suspension gently by pipetting up and down a few times and aliquot 200µL into five Eppendorf tubes (to work out the number of cells in each of the five vials, divide $1.2x10^{-8}$ by 5 = 24,000,000 cells).
    d. Centrifuge the tubes for 10 min at 10,000 x g at 4°C and carefully discard the supernatant, then freeze the pellets in the -20°C freezer for future use; Each pellet should contain 24,000,000 cells.

5. Resuspend one of the frozen pellets with 1ml of chilled 1x PBS; If dilutions of the positive control are required, resuspend the frozen pellet with 200µl of 1x chilled PBS, then prepare dilutions up to 10^{-3} and work out the number of cells in the dilutions:
    a. No. Cells in 10^{-1}: 24,000,000÷10 = 2,400,000 cells
    b. No. Cells in 10^{-2} = 24,000,000÷100 = 240,000 cells
    c. No. Cells in 10^{-3} = 24,000,000÷1000 = 24,000 cells

6. Process the swabs by cutting each swab head, using a disposable, sterile scalpel, into a 1.5ml Eppendorf tube containing 1ml of 1X PBS; If positive control dilutions were made, top up the positive controls to 1ml using 1x PBS.
7. Vortex the tubes @ full speed for 2 minutes.
8. Centrifuge tubes at 14,000 x g for 15 minutes to pellet the cells and remove the supernatant.
9. Resuspend the cells in 500µl of 1x PBS.
10. Appropriately label two Colombia Blood Agar plates for each swab (one plate for anaerobic culturing and one plate for aerobic culturing).
11. Pipette 50µl onto the labelled aerobic agar plate, and 50µl onto the paired anaerobic plate.
12. Incubate the aerobic plate in the 37° incubator and the anaerobic plate in the 37° anaerobic cabinet for 48-72 hours.
13. From the remaining 400µl of suspension, pipette 100µl into 1ml of 40% glycerol, prefilled in a 2ml cryovial, to form a glycerol stock of the original skin cells. Freeze the glycerol stock in the -80°C freezer to cryopreserve the skin cells.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

14. Centrifuge the remaining 300µl of suspension at 14,000 x g for 15 minutes to pellet the cells; remove the supernatant and keep the pellet for DNA extraction. Proceed to section: Skin Swab DNA Extraction Procedure Using the Maxwell Blood DNA Kit (Page 6) to begin the DNA extraction procedure.
15. After the 48h incubation, combine all colonies from the paired aerobic and anaerobic agar plates into a glycerol stock to cryopreserve the cultured skin microbiome. This can be performed by aseptically picking all colonies from both agar plates using a 10µl loop and inoculating them into 1ml of 40% glycerol, prefilled in a 2ml cryovial. Freeze the glycerol stock in the -80°C freezer. If master and working cell banks are required, follow the SOP "Preparation and Storage of Bacterial Master and Working Stocks SOP" (view Related Procedures 5 at the end of the document).
16. A DNA extraction should also be performed on these cultures following section: Skin Swab Culture DNA Extraction Procedure Using the Maxwell Blood DNA Kit (Page 8).

Skin Swab DNA Extraction Procedure Using the Maxwell Blood DNA Kit:

1. Resuspend the swab pellet in 400µl of 1x PBS and transfer the 400µl of sample with the original swab head into a 2ml Lysing Matrix E tube, then add 3µl of the Thermo Fischer Lysozyme (dilute to 250U/µl with Tris-EDTA Buffer; For ease of use prepare a 300 µl working stock of the lysozyme at 250U/µl by adding 75µl of the Lysozyme to 225µl of Tris-EDTA Buffer) and incubate with agitation at 300rpm, 37oC for 18-22 hours.
2. Following incubation, perform a bead beating step using a Tissue Lyser instrument for 3 minutes at 20 Hz followed by an off-board lysis.
3. Off-board lysis is performed by adding 40 µl proteinase K, 165 µl Buffer ATL, 120 µl Carrier RNA (resuscitate each vial containing 310µg of lyophilised Carrier RNA with 310µl of Buffer AVE to make a 1µg/µl solution; obtained from the QIAMP® Viral RNA Mini Handbook found under Related Procedures 6 at the end of the document), and 315 µl Buffer ACL into the 2 ml Lysing Matrix E tubes containing 400ul of sample and incubate the samples at 68 °C for 15 minutes; The off-board lysis protocol was obtained from the QIASymphony SP Protocol Sheet - Complex400 OBL V4 DSP Virus Bacteria found under Related Procedures 7 at the end of the document.
4. Centrifuge samples at 14,000 x g for 15 minutes to pellet the cells.
5. Whilst the samples are centrifuging, label the Elution Tubes with Sample IDs and add 50µl of Elution Buffer to the bottom of each 0.5ml Elution Tube; Change gloves before handling Maxwell® RSC Cartridges, RSC Plungers and Elution Tubes.
6. Turn on the Maxwell® RSC 48 Instrument and proceed through a self-check.

7. Select the RSC Blood DNA method, then select the cartridge positions to be used for this extraction run and the door should automatically open.
8. For each cartridge, vortex the well containing beads to resuspend them, then prepare the cartridges on the cartridge rack (Refer to Figure 1).
9. Remove the cartridge rack from the Maxwell instrument and prepare cartridges in the rack; Place each cartridge in the rack with the label identifier side facing away from the 0.5ml Elution Tubes (Refer to Figure 1); The cartridges will only fit in the instrument in this orientation. Press down on the cartridge to snap it into position. Carefully peel back the seal so that all plastic comes off the top of the cartridge - ensure that all sealing tape and any residual adhesive are removed.
10. After the samples have been centrifuged, add 630µl of the supernatant to well #1 of each cartridge and place one plunger into well #8 of each cartridge; Well #1 is the well closest to the label identifier side, and well #8 is closest to the elution tube (Refer to Figure 1).
11. Transfer the prepared cartridges in the rack to the Maxwell® RSC 48 Instrument platform. Ensure that the cartridge rack is level on the instrument platform.
12. Place Elution Tubes containing the Elution Buffer into the Elution Tube position for each cartridge (with the Elution Tube lids open, facing away from the cartridge - Refer to Figure 1).

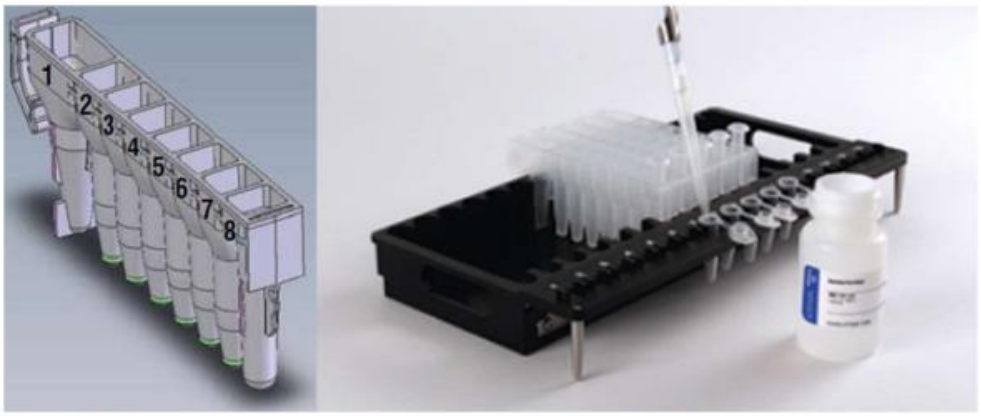


Figure 1 – Preparing Cartridges and Elution Tubes on the Cartridge Rack

(Obtained from the Promega Maxwell RSC Blood DNA Kit Manual, which can be viewed under the Related Procedures 2)

13. Press the Run/Stop button. The platform will retract, and the door will close.
14. The Maxwell® RSC 48 Instrument will immediately begin the purification run. The screen will display the steps performed and the approximate time remaining in the run.
15. When the automated purification run is complete, the LCD screen will display a message that the run has ended.
16. Follow on-screen instructions at the end of the run to open the door.
17. Verify that the plungers are located in well #8 of the cartridges at the end of the run. If plungers are not removed from the magnetic plunger bar, push them down gently by hand to remove them.
18. Remove Elution Tubes containing DNA and cap the tubes.
19. Remove the Maxwell® RSC Cartridge Rack from the instrument and dispose of the cartridges into the Maxwell Waste Sharps Bin.
20. DNA extracts can now be used for downstream use.

Skin Swab Culture DNA Extraction Procedure Using the Maxwell Blood DNA Kit:

1. Appropriately label one 1.5ml Eppendorf tube for each culture.
2. Prepare the appropriate volume of 1x PBS from the 10x PBS stock solution using Milli-Q water. Add 800µL of the 1x PBS to the labelled 1.5ml Eppendorf tubes.
3. Add 1ml of 1x PBS to a 1.5ml Eppendorf tube for the Negative control and 1ml of chilled 1x PBS to one of the Eppendorf tubes containing the previously frozen ATCC Skin Microbiome Whole Cell Mix pellet, as described above in step 5 under section: Processing Skin Swabs for DNA Extraction and Preparing Glycerol Stocks (Page 5).
4. For each run a negative control containing PBS must be used to ensure no contamination has occurred and the ATCC Skin Microbiome Whole Cell Mix must be used as a positive control, which is an established mock community, to demonstrate that any negative results are a lack of material rather than an extraction problem.
5. Thaw the culture glycerol stocks on ice, pipette up and down a few times to mix the suspension, and add 200µL of the stocks to the appropriately labelled tubes containing 800µL of the 1x PBS, to add up to 1ml.
6. Vortex the tubes @ full speed for 2 minutes.
7. Centrifuge the tubes at 14,000 x g for 15 minutes to pellet the cells and remove the supernatant.
8. Resuspend the culture pellet in 400µl of 1x PBS and transfer the 400µl of sample into a 2ml Lysing Matrix E tube, then add 3µl of the Thermo Fischer Lysozyme (dilute to 250U/µl with Tris-EDTA Buffer; For ease of use prepare a 300 µl working stock of the lysozyme at 250U/µl by adding 75µl of the Lysozyme to 225µl of Tris-EDTA Buffer) and incubate with agitation at 300rpm, 37oC for 18-22 hours.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

9. Following incubation, perform a bead beating step using a Tissue Lyser instrument for 3 minutes at 20 Hz followed by an off-board lysis.
10. Off-board lysis is performed by adding 40 µl proteinase K, 165 µl Buffer ATL, 120 µl Carrier RNA (resuscitate each vial containing 310µg of lyophilised Carrier RNA with 310µl of Buffer AVE to make a 1µg/µl solution; obtained from the QIAMP® Viral RNA Mini Handbook found under Related Procedures 6 at the end of the document), and 315 µl Buffer ACL into the 2 ml Lysing Matrix E tubes containing 400ul of sample and incubate the samples at 68 °C for 15 minutes (the off-board lysis protocol was obtained from the QIASymphony SP Protocol Sheet - Complex400 OBL V4 DSP Virus Bacteria found under Related Procedures 7 at the end of the document).
11. Centrifuge the samples at 14,000 x g for 15 minutes to pellet the cells.
12. Follow steps 5-20 under section: Skin Swab DNA Extraction Procedure Using the Maxwell Blood DNA Kit (Page 7 & 8).

COVID19 Precautionary Measures:

The handling of these samples has been assessed by the NBIP HSEQA team who have concluded that samples from these individuals would be very low risk of being contaminated with COVID-19 because the participants represent a vulnerable group within the community taking a lot of care to avoid contacts outside the home. The following guidance has been requested, to be conducted, in order to process these samples at the Quadram Institute:

- Before collecting any samples from participants, please ask if they, or anyone in their household has knowingly had COVID-19. Any participants that have had coronavirus should not donate samples until an individual risk assessment can be conducted.
- Samples should all be handled in a designated Class II MSC until they would be classified as biologically inactive in a side room off the main lab. This is currently either the tissue processing laboratory on the lower ground floor or the microbiology laboratory on level 1. It is essential that anyone conducting work with these samples are particularly cautious over the use of gloves, regular hand washing and not touching the face.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

SAFETY INFORMATION

Principal Hazards:

- Biohazards from human skin swab material
- Chemical hazards from small volume of kit materials
- Sharps – scalpels

Method Steps Involving Higher Risk (and precautions to minimise risk):

- Wear a lab coat, orange shield-skin gloves and safety glasses
- Ensure human samples are handled as per the handling human tissue COP and HG2 RA
- Dispose of scalpels immediately after use in a sharps (cin) bin, keep fingers away from the blade.

Essential Precautions:

- All work must be performed in a microbiological safety cabinet within a containment level 2 laboratory
- Appropriate PPE (lab coat, gloves and safety glasses) should be worn when handling human samples
- Personnel at QIB working with human material should have Hepatitis B immunisation and other relevant vaccinations following NBI Occupational Health Advice

Personal Protective Clothing:

- Lab coat
- Orange shield-skin gloves
- Safety glasses

Any Restrictions on Use:

- Trained personnel only
- Hepatitis B vaccination recommended

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

- Not to be carried out by under 18s

Exceptions or Warnings:

- Should be included at relevant step in the procedure

4. RISK STATEMENT

The overall risk is LOW if the guidelines in this SOP are followed carefully.

All individuals using this procedure will be shown the risk assessment and given appropriate information, instruction and training in the risks and precautions necessary, including the use of any personal protective equipment required.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

SOP HEALTH RISK ASSESSMENT			
[1] Activity:	Processing, Culturing and Extracting Bacterial DNA from PEARL Study Skin Swab Samples Using the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit		
[2] Location of activity:	QIB Laboratories		
[3] Who is involved:	Trained PEARL Study Staff		
[4] Frequency of activity:	When required – approximately once a week		
[5] Duration of Activity:	Approximately 4 hours		
[6] Chemical Hazard Name:	**Hazard Statements**	**Route of Exposure***	**Quantity Used**
Proteinase K	H317: May cause an allergic skin reaction H334: May cause allergy or asthma symptoms or breathing difficulties if inhaled	S, B	40ul per sample
ACL	H302: Harmful if swallowed H314: Causes severe skin burns and eye damage H318: Causes serious eye damage H412: Harmful to aquatic life with long lasting effects EUH032: Contact with acids liberates very toxic gas	S, I, B	315ul per sample
[7] Details of biological agents of risk to human health:			**GMRA Number**
Hazard Group 2 bacteria from human skin swab samples All such specimens must be treated as though they are infected with bacterial infectious agents Low risk of samples being contaminated by COVID-19 as participants are in a vulnerable group and have been careful to distance from people outside their homes			

[8] Other Hazards:	Please ■ as necessary				
Hot or Cold Burns	☐	Ionising Radiation**	☐	Ultra Violet or Infra Red	☐

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

Dust	☐	Noise	☐	Pollen Sensitizer	☐
Repetitive Action	☐	Extreme Cold Environment (< 0°C)	☐	Lifting / Manual Handling	☐
Asphyxiation	☐	Cuts	■	Electrical	☐
Slips / trips / falls	☐	Display Screen Equipment	☐	Nanomaterials (man-made particles, tubes, rods, or fibers ≤100nm in size)	☐
Other (give details)	☐				
[9] Control Measures:		Please ■ as necessary			
Fume Cupboard	☐	Microbiological Safety Cabinet	■	Total Containment Cabinet	☐
Ventilated Bench	☐	Spill Tray	☐	Trained personnel only	■
Signs	☐	Reduce frequency/alternate activity	☐	Reduce duration of activity	☐
Sub divide a load	☐	2 persons lift of equipment	☐	Not for more than 1 hour	☐
Regular, short breaks	☐	Alternate activities	☐	Other (give details)	☐
[10] Personal Protection:		Please ■ as necessary			
Lab coat	■	Safety Glasses	■	Face Shield	☐
Goggles	☐	Gloves *(specify)*	■	Protective Gloves *(specify)*	☐
Ear defenders	☐	Other (give details)	☐		
[11] Is personal monitoring and/or health surveillance required?			Yes	■	No ☐
Details:	Hepatitis B vaccination recommended.				
[12] Restrictions:		Please ■ as necessary			
No lone working	☐	Not to be left unattended	☐	Named persons only	☐
In restricted area	☐	Risk Assessment Required for New or Expectant Mothers	■	Under constant supervision	☐
Not by under 18's	■	Other (give details)	☐		

[13] Level of Residual Risk:		Please ■ as necessary				
Low	■	Medium	☐	High		☐
Name of Assessor:		Iliana Rosa Serghiou			Date:	26/11/2021

* Route of exposure; S = skin, I = ingestion, B = inhalation

5. DOCUMENTS

Good Laboratory Practice in the Use of Chemicals: http://intranet/cms/2199

Link to Laboratory Waste Disposal: http://intranet/cms/968

Biosafety Handbook: http://intranet/cms/970

6. RELATED PROCEDURES

1. Maxwell RSC 48 Instrument Operating Manual TM510
   (Double click the icon below to open)



Maxwell RSC 48
Instrument Operatir

2. Maxwell RSC Blood DNA Kit TM419 Full Protocol and Technical Manual
   (Click the link below to open)



Maxwell RSC Blood
DNA Kit TM419 - Ful

3. Processing Blood Samples Using the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit
   (Click the link below to open)



Processing Blood
Samples Using the F

4. ATCC Skin Microbiome Whole Cell Mix Manufacturer Instructions
   (Click the link below to open)

Skin Microbiome Whole Cell Mix | ATCC

5. Preparation and Storage of Bacterial Master and Working Stocks SOP
   (Double click the icon below to open)



Preparation and
Storage of Bacterial

6. QIAMP® Viral RNA Mini Handbook
   (Double click the icon below to open)



HB-0354-007_HB_Q
A_Viral_RNA_Mini_0

7. QIASymphony SP Protocol Sheet - Complex400 OBL V4 DSP Virus Bacteria
   (Double click the icon below to open)



HB-0678-S04-001_P
S_QIASymphony_Co

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

8. Use of Class II Microbiological Safety cabinet – GHFS-LH-004

7. DOCUMENT HISTORY

Issue Number	Reason for update
001	First Issue

8. APPENDIX

Staff Training record for SOP (Link to: Processing, Culturing and Extracting Bacterial DNA from PEARL Study Skin Swab Samples Using the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit Training. doc).

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 28 January 2022.

Appendix 8. 241617-SC-024 MOTION study Stool Sample DNA Extraction Using the Promega Maxwell V3

QUADRAM INSTITUTE BIOSCIENCE

STANDARD OPERATING PROCEDURE

Quadram Institute
Science Health Food Innovation

Title: MOTION study_ Stool sample DNA extraction using the Promega Maxwell

Applies to staff in: MOTION study	***Reference number: 241617-SC-024***
Author: Sarah Phillips	***Approver: Simon Carding***
Issue date: 27 May 2022	***Review date: 26 May 2024***
QA and H&S Approval: R Walding	***Version Number: 003***

AUTHORISATION: I confirm that this SOP is an accurate description of the procedure or method in use.

DECLARATION. The safety information contained in this document is to the best of my knowledge an accurate statement of the known and foreseeable hazards to health and of the safety precautions, which are to be followed to minimise the risks to health in the process. All users of this SOP are adequately trained to be competent in the process.

Approved by ISP/NC Leader, or Research Leader, or Senior Scientist, or Senior Administrator

Signed:	Print: S Carding	Date: 27 May 2022

1. PURPOSE OF PROCEDURE/METHOD AND ITS SCOPE

Use of Promega Maxwell 48 to perform DNA extraction from faecal samples.

Work under ethics approval: (IRAS project ID 241617)

2. EQUIPMENT AND REAGENTS NEEDED

Equipment

- Promega Maxwell 48
- MPbio fastprep 24
- Microfuge
- Vortex
- 1ml Gilson pipette and tips
- Hot Block 95^{o}C
- Hot Block 70^{o}C

Reagents

- 2ml Lysing matrix E centrifuge tubes from Fisher Scientific item 11452420
- Promega Maxwell RSC PureFood GMO and Authentication kit item #AS1600

3. STEPS IN PROCEDURE (INCLUDING SAFETY INFORMATION)

Before beginning this procedure, please refer to the 'QIB Biosafety Handbook' and 'Code of Practice for work with human body tissues, fluids and other samples of human origin' for information on local procedures for working with biological samples.

All work should be conducted in a microbiological safety cabinet within a containment level 2 laboratory. Approved disinfectants should be available in case of a leak or any spillage. A lab coat and protective gloves should be worn whenever working in the laboratory with human samples. Personnel at QIB working with human material should have Hepatitis B immunisation and should have attended training on working with human samples.

A new clinical waste bag (double bagged) should be set up before work is started. All pipette tips or sharps should be put into a clinical waste sharps bin. Non-clinical biological waste should be sealed, room number marked on the bag and should be sent for autoclaving at the end of the working day.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 27 May 2022.

All areas should be disinfected with an approved disinfectant after work has been completed.

Note: A negative control sample should be run with each extraction run. A negative control should follow the exact process as other samples, with the only difference being described in step 1 of the protocol. Each negative control should be given a unique reference number.

1. Add 200 mg of faeces sample to a lysing matrix E tube with screw top lid. For a negative control sample, 200ul molecular grade H2O should be added to a lysing matrix E tube in place of the faecal sample.
2. Add 1ml CTAB to each sample and vortex for 30 seconds.
3. Heat sample in Hot Block 95°C for 5 minutes.
4. Vortex for 1 minute.
5. Homogenize in the FastPrep Instrument for 45 seconds at a speed setting of 6.0 m/s.
6. Pulse centrifuge the samples for 3 seconds to reduce bubbles in the samples.
7. From Kit add 40µl of proteinase K and 20µl of RNase A to samples and vortex to mix.
8. Heat sample in Hot Block 70°C for 10 minutes.
9. While heating, prepare cartridges.
    a. Place cartridges in rack and remove foil seal (they will only go in one-way round).
    b. Add 100µl elution buffer into supplied elution tubes and place in rack.
    c. Place plunger into 8th well of cartridge (well closest to elution tubes).
10. Add 300µl of lysis buffer (RED label) into well 1 of cartridge (well furthest from elution tubes).
11. Centrifuge at 14,000 x g for 5 minutes to pellet debris.
12. Add 300µl of supernatant to well 1 of cartridge.
13. Run Maxwell RSC with the PureFood Protocol.
14. Samples should be quantified using the Qubit fluorometer 2.0 (see SOP GMH-LH-003), 3.0 (See SOP MFC-JO-006), or alternatively, for large sample sets the sequencing team may be able to provide assistance using a plate read method. Contact the sequencing team for more information. Results of the Qubit analysis should be recorded in the "MOTION study preliminary analyses" spreadsheet available on the network drive in the folder "internal use documents".
15. Log the analysis of these samples on achiever as follows:

Logging analysis on Achiever

- From the project screen, click on "analyses".
- Click on the "new Analysis/Process" button (white page – top right of analyses table).
- Enter the relevant details of the analysis and enter a Run ID. The Run ID should consist of the run number and the ID of the negative control used, for instance, DNA run 1 using Negative Control 1 should have a Run ID of "DNA1NEG1".
- Set status as pending.
- Click save.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 27 May 2022.

Adding samples to analysis

- To add samples to analysis, select "samples" from the left-hand details column
- Click on the workflow drop down (…) at the top of the samples grid and select add samples.
- Search for all samples of the run and toggle the tick boxes. Once all samples have been selected, click "add" in the top left corner.

Adding consumables (reagents) to analysis

- Click on the consumables tab on the analysis screen.
- Click on the "new consumables" button (white page – top right of the consumables grid)
- Search for and toggle the tick box of the consumables (reagents) used for the run. Click "add selected consumables". **Be careful to check that the correct batch number has been added.**
- If you are opening a new kit or consumable for the run this will need to first be added in the "manage consumables section" before it can be selected for a run. To do this, go back to the admin tab, and under General Administration, select manage consumables.
- Click on the "new consumable maintenance record" button (white page – top right of the manage consumables grid)
- Add all information of the reagent/consumables and click save and close. This will now be available to add to your analysis record.

16. DNA extracts should be stored at -80°C in study specific boxes. Sample position should be recorded on achiever following the instructions provided below:

- Once the run is completed, the samples will need to be split into DNA on achiever and stored. To do this, select all the samples on the run but toggling the tick boxes.
- Select split samples from the workflow (…) drop down box in the top left of the samples grid.
- Select the "deplete original sample" box.
- Sample type – Frozen, Tissue type – whichever samples type has been extracted (stool, skin, vaginal) Tissue subtype – DNA.
- Click "split samples"
- These can then be stored in the correct location in achiever as in SOP 241617-SC-006

WASTE DISPOSAL ARRANGEMENTS

Biological waste (including Maxwell cartridges) to be disposed of in the white autoclave bins lined with bags.

Clinical waste to be disposed of in the yellow bags for incineration, ensure double bags and no free liquids.

COVID19 precautionary measures

The handling of these samples has been assessed by the NBIP HSEQA team who have concluded that samples from these individuals would be very low risk of being contaminated with COVID-19 because the participants represent a vulnerable group within the community taking a lot of care to avoid contacts outside the home.

The following guidance has been requested to be conducted in order to process these samples at the Quadram institute:

- Before collecting any samples from participants, please ask if they, or anyone in their household has knowingly had COVID-19. Any participants that have had coronavirus should not donate samples until an individual risk assessment can be conducted.
- Samples should all be handled in a designated Class II MSC until they would be classified as biologically inactive in a side room off the main lab. This is currently either the tissue processing laboratory on the lower ground floor or the microbiology laboratory on level 1.
- It is essential that anyone conducting work with these samples are particularly cautious over the use of gloves, regular hand washing and not touching the face.

SAFETY INFORMATION

Principal Hazards: Biohazards from human faecal material. Chemical hazards from small volume of kit materials, hot burns

Method Steps involving Higher Risk (and precautions to minimise risk): Wear lab coat, Sheildskin and safety glasses. Ensure human samples are handled as per the handling human tissue COP and HG2 RA.

Essential Precautions: All work must be performed in a microbiological safety cabinet within a containment level 2 laboratory. Appropriate PPE (labcoat, gloves, protective glasses) should be worn when handling human samples. Personnel at QIB working with human material should have Hepatitis B immunisation.

Personal Protective Clothing: Wear lab coat, Shieldskin gloves and safety glasses.

Any restrictions on use: Trained personnel only. Hep B Vaccination recommended. Not to be carried out by under 18s

Exceptions or Warnings: should be included at relevant step in the procedure

4. RISK STATEMENT

The overall risk is LOW if the guidelines in this SOP are followed carefully.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 27 May 2022.

All individuals using this procedure will be shown the risk assessment and given appropriate information, instruction and training in the risks and precautions necessary, including the use of any personal protective equipment required

SOP HEALTH RISK ASSESSMENT			
[1] Activity:	MOTION study_ Stool sample DNA extraction using the Promega Maxwell		
[2] Location of activity:	QIB laboratories		
[3] Who is involved:	Trained MOTION staff		
[4] Frequency of activity:	When required – approximately once a week		
[5] Duration of Activity:	Approximately 3 hours		
[6] Chemical Hazard Name:	**Hazard Statements**	**Route of Exposure***	**Quantity Used**
Hexadecyltrimethylammonium bromide CTAB	H302 Harmful if swallowed. H315 Causes skin irritation. H318 Causes serious eye damage. H335 May cause respiratory irritation. H400 Very toxic to aquatic life	S, I, B	1ml per sample
Proteinase K 20mg/ml	H317 May cause an allergic skin reaction. H334 May cause allergy or asthma symptoms or breathing difficulties if inhaled.	S, B	40ul per sample
Ethanol	H225 Highly flammable liquid and vapour H319 Causes serious eye irritation	S, I, B	1ml per sample
Isopropanol	H225 Highly flammable liquid and vapour H319 Causes serious eye irritation H336 May cause drowsiness or dizziness.	S, I, B	1ml per sample
Lysis Buffer	H302 Harmful if swallowed. H315 Causes skin irritation. H318 Causes serious eye damage.	S, I, B	1ml per sample
[7] Details of biological agents of risk to human health:			**GMRA Number**

Hazard Group 2 bacteria (Human Faeces) **All such specimens must be treated as though they are infected with Blood Bourne Viruses (BBV) or other infectious agents.** **Low risk of samples being contaminated by COVID-19 as participants are in a vulnerable group which has been careful to distance from people outside their homes.**					
[8] Other Hazards:		Please ■ as necessary			
Hot Burns	■	Ionising Radiation**	☐	Ultra Violet or Infra Red	☐
Dust	☐	Noise	☐	Pollen Sensitizer	☐
Repetitive Action	☐	Extreme Cold Environment (< 0°C)	☐	Lifting / Manual Handling	☐
Asphyxiation	☐	Cuts	☐	Electrical	☐
Slips / trips / falls	☐	Display Screen Equipment	☐	Nanomaterials (man-made particles, tubes, rods, or fibers ≤100nm in size)	☐
Other (give details)	☐				
[9] Control Measures:		Please ■ as necessary			
Fume Cupboard	☐	Microbiological Safety Cabinet	■	Total Containment Cabinet	☐
Ventilated Bench	■	Spill Tray	☐	Trained personnel only	■
Signs	☐	Reduce frequency/alternate activity	☐	Reduce duration of activity	☐
Sub divide a load	☐	2 persons lift of equipment	☐	Not for more than 1 hour	☐
Regular, short breaks	☐	Alternate activities	☐	Other (give details)	☐
[10] Personal Protection:		Please ■ as necessary			
Lab coat	■	Safety Glasses	■	Face Shield	☐
Goggles	☐	Gloves *(skinshield)*	■	Protective Gloves *(specify)*	☐
Ear defenders	☐	Other (give details)	☐		
[11] Is personal monitoring and/or health surveillance required?			Yes	■	No ☐

Details:	Hep B vaccination recommended.				
[12] Restrictions:		Please ■ as necessary			
No lone working	☐	Not to be left unattended	☐	Named persons only	☐
In restricted area	☐	Risk Assessment Required for New or Expectant Mothers	■	Under constant supervision	☐
Not by under 18's	■	Other (give details)	☐		
[13] Level of Residual Risk:		Please ■ as necessary			
Low	■	Medium	☐	High	☐
Name of Assessor:		Sarah Phillips		Date:	06/02/2020

* Route of exposure; S = skin, I = ingestion, B = inhalation

5. DOCUMENTS

Human Tissue COP: https://intranet.nbi.ac.uk/cms/970

Good Laboratory Practice in the Use of Chemicals: http://intranet/cms/2199

Link to Laboratory Waste Disposal: http://intranet/cms/968

Biosafety Handbook: http://intranet/cms/970

6. RELATED PROCEDURES

GMH-LH-003 - Use of the Qubit 2.0 fluorometer for DNA, RNA and protein quantification

MFC-JO-006 - Use of Qubit 3.0 fluorometer for DNA and RNA quantification using Qubit BR and HS reagent kits

241617-SC-006 MOTION study_sample receipt handling storage and disposal

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 27 May 2022.

7. DOCUMENT HISTORY

Issue Number	Reason for update
001	First Issue
002	Addition of negative control, instructions for recording information on achiever system, quantification of samples. Addition of instructions for achiever
003	Waste information updated

8. APPENDIX

Staff Training record for SOP (Link to: MOTION study_ Stool sample DNA extraction using the Promega Maxwell training.doc)

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 27 May 2022.

Appendix 9. QIB SOP for General Bacteria Growth

QUADRAM INSTITUTE BIOSCIENCE

STANDARD OPERATING PROCEDURE

Quadram Institute
Science Health Food Innovation

Title: General Bacteria Growth

Applies to staff in: Level 2 labs	***Reference number: MIF-MW-014***
Author: Emma Holden	***Approver: M Webber***
Issue date: 03 February 2021	***Review date: 02 February 2023***
QA and H&S Review: R Walding	***Version Number: 002***

AUTHORISATION: I confirm that this SOP is an accurate description of the procedure or method in use.

DECLARATION. The safety information contained in this document is to the best of my knowledge an accurate statement of the known and foreseeable hazards to health and of the safety precautions, which are to be followed to minimise the risks to health in the process. All users of this SOP are adequately trained to be competent in the process.

Approved by ISP/NC Leader, or Research Leader, or Senior Scientist, or Senior Administrator

Signed:	Print: M Webber	Date: 03 Feb 2021

TITLE: GENERAL BACTERIA GROWTH

1. PURPOSE OF PROCEDURE/METHOD AND ITS SCOPE

This procedure outlines the standardised method of setting up overnight cultures

2. EQUIPMENT AND REAGENTS NEEDED

Equipment

- Toothpick
- Pipette tip
- Bunsen burner or Class II Microbiological safety cabinet

Reagents

5 mL relevant broth in 25 mL glass universal bottle (made by the media kitchen)

Freezer stock or plate of relevant bacteria

3. STEPS IN PROCEDURE (INCLUDING SAFETY INFORMATION)

All work is done under sterile conditions, under a flame or in a microbiological safety cabinet

Setting up overnight cultures from a bead stock

1. Using a loop, pick up a bead from the freezer stock
2. Open (and flame the lid of) the glass bottle and deposit the bead in the media
3. Seal the bottle and incubate at 37°C shaking at 250rpm for 16-18 hours or overnight (unless other conditions are specified).

Setting up overnight cultures from a DMSO stock

1. Using a loop, scrape some freezer stock onto the loop
2. Open (and flame the lid of) the glass bottle and put the loop in the media, mixing the loop to remove the bacteria

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 3 February 2021.

3. Seal the bottle and incubate at 37°C shaking at 250rpm for 16-18 hours or overnight (unless other conditions are specified).

Setting up overnight cultures from an agar plate

1. Using a loop, scrape some bacteria from the plate on the loop. Alternatively, use a toothpick to pick a colony from the plate
2. Open (and flame the lid of) the glass bottle and put the loop in the media, mixing the loop to remove the bacteria. If using a toothpick, drop the toothpick into the media
3. Seal the bottle and incubate at 37°C shaking at 250rpm for 16-18 hours or overnight (unless other conditions are specified).
4. If a toothpick has been used, remove the toothpick from the media after incubation before discarding the bottle

SAFETY INFORMATION

Principal Hazards: Biohazard, working with hazard group 2 organisms, burns and fire hazard from Bunsen burner

Method Steps involving Higher Risk (and precautions to minimise risk): Trained personnel to work with hazard group 2 organisms, wear outlined PPE

Essential Precautions: work under a flame or in a microbiological safety cabinet

Personal Protective Clothing: lab coat, shieldskin gloves

Any restrictions on use: Trained personnel only

Exceptions or Warnings: should be included at relevant step in the procedure

4. RISK STATEMENT

The risk of this procedure has been determined to be low.

All individuals using this procedure will be shown the risk assessment and given appropriate information, instruction and training in the risks and precautions necessary, including the use of any personal protective equipment required

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 3 February 2021.

SOP HEALTH RISK ASSESSMENT			
[1] Activity:	Making bacterial overnight cultures		
[2] Location of activity:	Level 2 labs		
[3] Who is involved:	Members of the Webber group		
[4] Frequency of activity:	Daily		
[5] Duration of Activity:	10 minutes		
[6] Chemical Hazard Name:	**Hazard Statements**	**Route of Exposure***	**Quantity Used**
N/A	N/A	N/A	N/A
[7] Details of biological agents of risk to human health:			**GMRA Number**
Bacteria in daily use include Salmonella Typhimurium, Pseudomonas aeruginosa, Staphylococcus aureus, E. coli and other hazard group 2 organisms.			

[8] Other Hazards:		Please ■ as necessary			
Hot Burns	■	Ionising Radiation**	☐	Ultra Violet or Infra Red	☐
Dust	☐	Noise	☐	Pollen Sensitizer	☐
Repetitive Action	☐	Extreme Cold Environment (< 0°C)	☐	Lifting / Manual Handling	☐
Asphyxiation	☐	Cuts	☐	Electrical	☐
Slips / trips / falls	☐	Display Screen Equipment	☐	Nanomaterials (man-made particles, tubes, rods, or fibers ≤100nm in size)	☐
Other (give details)	☐				
[9] Control Measures:		Please ■ as necessary			
Fume Cupboard	☐	Microbiological Safety Cabinet	☐	Total Containment Cabinet	☐
Ventilated Bench	☐	Spill Tray	☐	Trained personnel only	■

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 3 February 2021.

Signs	☐	Reduce frequency/alternate activity	☐	Reduce duration of activity	☐
Sub divide a load	☐	2 persons lift of equipment	☐	Not for more than 1 hour	☐
Regular, short breaks	☐	Alternate activities	☐	Other (give details)	☐
[10] Personal Protection:		Please ■ as necessary			
Lab coat	■	Safety Glasses	☐	Face Shield	☐
Goggles	☐	Gloves *(Shieldskin)*	■	Protective Gloves *(specify)*	☐
Ear defenders	☐	Other (give details)	☐		
[11] Is personal monitoring and/or health surveillance required?			Yes	☐	No ■
Details:					
[12] Restrictions:		Please ■ as necessary			
No lone working	☐	Not to be left unattended	☐	Named persons only	☐
In restricted area	☐	Risk Assessment Required for New or Expectant Mothers	☐	Under constant supervision	☐
Not by under 18's	■	Other (give details)	☐		
[13] Level of Residual Risk:		Please ■ as necessary			
Low	■	Medium	☐	High	☐
Name of Assessor:		Emma Holden		Date:	02 Feb 2021

* Route of exposure; S = skin, I = ingestion, B = inhalation

5. DOCUMENTS

Chemical Tables: http://intranet/cms/2256

Good Laboratory Practice in the Use of Chemicals: http://intranet/cms/2199

Link to Laboratory Waste Disposal: http://intranet/cms/968

Biosafety Handbook: http://intranet/cms/970

6. RELATED PROCEDURES

None

7. DOCUMENT HISTORY

Issue Number	Reason for update
001	First Issue
002	SOP reviewed; location updated

8. APPENDIX

Staff Training record for SOP (Link to: General Bacteria Growth Training.doc)

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 3 February 2021.

Appendix 10. QIB SOP Preparation and Storage of Bacterial Master and Working Stocks

QUADRAM INSTITUTE BIOSCIENCE

STANDARD OPERATING PROCEDURE

Quadram Institute
Science Health Food Innovation

Title: Preparation and Storage of Bacterial Master and Working Stocks SOP

***Applies to staff in:* QIB Level 1**	***Reference number:* GMH-LH-067**
***Author:* I.R. Serghiou**	***Approver:* L Hall**
***Issue date:* 24 April 2020**	***Review date:* 23 April 2022**
***QA and H&S Review:* Rachel Walding**	***Version Number:* 001**

AUTHORISATION: I confirm that this SOP is an accurate description of the procedure or method in use.

DECLARATION. The safety information contained in this document is to the best of my knowledge an accurate statement of the known and foreseeable hazards to health and of the safety precautions, which are to be followed to minimise the risks to health in the process. All users of this SOP are adequately trained to be competent in the process.

Approved by ISP/NC Leader, or Research Leader, or Senior Scientist, or Senior Administrator

Signed:	Print: L Hall	Date: 24 April 2020

TITLE: PREPARATION AND STORAGE OF BACTERIAL MASTER AND WORKING STOCKS SOP

1. PURPOSE OF PROCEDURE/METHOD AND ITS SCOPE

To accurately prepare and appropriately store bacterial master and working stocks so that they can be preserved and kept viable for future use (within regulatory guidelines).

Faecal samples can be used for this procedure only if they have been obtained with informed consent and approval from a research ethics committee.

List of current studies:

- BAMBI and BEAM (and other ethics-approved studies under Norwich Research Park (NRP) Biorepository): Faecal collection was approved by the Faculty of Medical and Health Sciences Ethics Committee at the University of East Anglia (UEA), and followed protocols in line with the policy of the NRP Biorepository (Human Tissue Authority-licensed tissue bank, License number 11208) under the ethics approval granted by the National Health Service Health Research Authority (NHS HRA) (East of England-Cambridge East Research Ethics Committee, REC Ref 19/EE/0089; IRAS number 259062). (SOP: GMH-LH-060)
- PEARL: Faecal collection was approved by Research Ethics Committee (REC reference: 18/LO/1703) under the IRAS approval reference number 241880. (SOP: 241880-LH-007)

2. EQUIPMENT AND REAGENTS NEEDED

Equipment

- Dry ice
- Class II Microbiological Safety Cabinet
- Incubator/Anaerobic cabinet
- Balance
- Sterile loops
- Sterile 10ml serological pipette
- 100-1000µl (1ml) pipette and sterile tips
- 20-200µl pipette and sterile tips

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

- Sterile 15ml falcon tubes
- Sterile 1.5ml & 2ml Eppendorf tubes
- Centrifuge
- Vortex
- Micro mixer
- Sterile 2ml cryovials
- Sterile spreaders
- -80 Freezer (No: 123 and 148)

Reagents

- Sterile PBS diluent
- Appropriate sterile broths and agar
- Sterile 30% Glycerol broth stock for cryopreservation
- Deionised water
- Bacterial stocks
- Faeces and bodily fluids

3. STEPS IN PROCEDURE (INCLUDING SAFETY INFORMATION)

Definitions:

Master Cell Bank (MCB) - A MCB is defined as an aliquot of a single pool of cells that generally has been derived and expanded from one-parent cell culture under defined conditions, dispensed into multiple containers, and stored under defined conditions. The MCB is used to derive all working cell banks (WCB). A MCB is produced using standardised laboratory procedures and characterized to ensure proper identity and demonstrate the cell bank is free from accidental contamination.

Working Cell Bank (WCB) - The WCB is prepared from aliquots of a homogeneous suspension of cells obtained from culturing the MCB under defined culture conditions. The WCB is prepared from one or more MCB vials and tested in similarly to the MCB before being used in production.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

- If prepared and managed correctly a MCB/WCB can provide reproducible and dependable supplies of identical cultures over many decades.
- A two-tiered cell banking system consisting of a MCB and a WCB is recommended when a cell line is to be used over many manufacturing cycles. It is recommended that production cell banks be stored in two or more widely separated areas within a production facility and at a distant site in order to avoid loss of the cell line due to local disaster or malfunction. Dual-site storage of all cells is accomplished with on-site and off-site storage.

Note: All procedures mentioned in this document must be carried out aseptically i.e. have good personal hygiene such as washing hands before and after working in the lab, wear appropriate PPE such as orange shield gloves and lab coat, make sure the work surfaces are cleaned with Bio shield and then 70% ethanol before use, between samples, and after use, and use sterile equipment and media/reagents.

For the entirety of this procedure- wear orange shield skin gloves, changing them if contaminated and a Howie-style lab coat.

Though the risk of hepatitis B through faeces is small, caution should still be taken when dealing with human tissue. Practice good hygiene and wash hands thoroughly after finishing in the lab.

Ethics clearance and QIB governance

- The proposed use of human samples (faeces) for this method will be registered with the QIB Human Research Governance Committee before samples are received on site. QIB Investigator will be responsible for the completion of the relevant QIB Human Research Governance Committee Registration Form for Use of Human Tissue Samples..
- HTA- relevant material (faeces samples) will be processed within 24 h from arrival and will not be stored in QIB laboratory facilities. Long-term storage will be allowed only for samples with appropriate NHS HRA approval.
- A log including details but not limited to Sample Shipment Date, Type of Biological Sample, Initials of Responsible person, Number of Tubes will be properly maintained by Hall's team.
- Sample disposal will be carried out in accordance with the HTA's Codes of Practice and relevant QIB policy (reference to QIB Biological Handbook, Code of practice for work with human body tissues, fluids and other samples of human origin). A log will be maintained by Hall's team to document disposal including the date, reason and method. It is good practice for human tissue to be bagged separately from

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

clinical waste. Burnable biological waste is collected for dispatch to an incinerator. Suitable waste must be triple-bagged, and the outer bag must be yellow; the bags must be sealed and labelled in such a way as to identify the originator.

- Relevant staff of the Hall's team with access to these samples will have the appropriate training through the QIB Human Studies Training program. Staff will receive relevant vaccinations following NBI Occupational Health advices.

Procedure for Preparing a Bacterial Master and Working Stock from Frozen Stocks

1. This procedure must be carried out aseptically in a Class II Microbiological Safety Cabinet.
2. Remove the required bacterial stocks from the -80°C freezer – Transport on dry ice to avoid thawing the stocks.
3. Aseptically scrape a loopful of the frozen stock using a stripette, and with a 10µl loop streak onto appropriate agar which is selective for the bacteria being cultured (To prepare agar follow the "*Media Preparation and Storage*" SOP).
4. Use the quad streak method so that distinct colonies can be obtained.
5. Appropriately grow anaerobically or aerobically at a temperature, and for a duration, specific to the bacteria being cultured.
6. Colonies can be picked for Gram staining at this stage or to purify colony growth in order to prepare liquid culture for stocks.
7. To purify colony growth, aseptically pick a single colony from the incubated agar plate using a 10µl loop, and quad streak again onto another agar plate selective to the bacteria being cultured.
8. Appropriately grow anaerobically or aerobically at a temperature, and for a duration, specific to the bacteria being cultured.
9. Colonies will then be picked for liquid culture.
10. Using a 10ml serological pipette, add 10ml of appropriate broth, which is selective for bacteria being cultured, to a 15ml falcon tube to make up the liquid culture (To prepare broths follow the "*Media Preparation and Storage SOP*").
11. Aseptically pick a single colony from your incubated agar plate using a 10µl loop and inoculate the 10ml of broth in the falcon tube with the colony.
12. Appropriately grow anaerobically or aerobically at a temperature, and for a duration, specific to the bacteria being cultured.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

13. This bacterial liquid culture can be used for a variety of downstream applications (i.e. DNA extraction or for oral gavage in mice) at this stage or prepared as stock.
14. To prepare as stock, spin down the incubated liquid culture using a centrifuge set to 4000 g x 10-15 minutes (follow the *"Use of Centrifuge - Eppendorf 5810R SOP"*).
15. Discard the supernatant and add 1ml of 30% Glycerol broth to the pellet and resuspend a few times to mix (To prepare 30% glycerol broth follow the *"Preparation of 30% Glycerol Broth for Bacterial Stock Cryopreservation SOP"*). The glycerol acts as a cryoprotectant.
16. Add the 1ml of stock into a 2ml cryovial.
17. Make sure the cryovial is labelled with a sample ID., bacteria name and type of stock (i.e. master or working) using a permanent water-resistant marker.
18. Store the isolate in the -80°C freezer for long-term storage.
19. 10 master stocks and 20 working stocks should be prepared following this method. Regulations recommend 50-200 cryovials of master stock and working stock when manufacturing a product.
20. Details of the MCB and WCB should be recorded such as:
    - ID numbers
    - Bacterial strains
    - Date first frozen
    - Storage location and temperature
    - Bank type (MCB/WCB)
    - Analysts' initials
    - Passage number
    - Date of each passage
    - Analysts initials for each passage
    - Which master stock was used to create the working stock
    - Materials used to prepare the master and working stocks
    - QC details such as morphology, cell recovery and viability before and after freezing.
21. One master and working stock should be used at a time and should be changed after a maximum of 20 passages; Regulations have not defined an acceptable maximum number of passages for each individual stock but should be set relatively low to avoid senescence and phenotypic/genotypic drift.

Procedure for Preparing a Bacterial Master and Working Stock from Faeces or Bodily Fluids

1. This procedure must be carried out aseptically in a Class II Microbiological safety cabinet.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

2. Place a 2ml Eppendorf tube on a balance and tare. Weigh 100mg (0.1g) of faeces or 100µl of bodily fluid into the tared Eppendorf tube.
3. Pipette 900µl of PBS into the same Eppendorf tube.
4. Homogenise the PBS and sample using a vortex for 30 seconds and then a micro mixer for 10-15 minutes if necessary (follow the *"Vortex Mixer SOP" and the "Micro Mixer SOP"*).
5. Perform serial 1/10 dilutions up to 10^{-4} (if necessary, prepare higher dilutions).
6. Pipette 100µl of the required dilutions, usually 10^{-3} and 10^{-4}, onto appropriate agar plates selective for the bacteria being cultured.
7. Using a spreader, spread the dilution evenly around the plate and leave to soak in for 10-20 minutes.
8. Appropriately grow anaerobically or aerobically at a temperature, and for a duration, specific to the bacteria being cultured.
9. Colonies can be picked, for Gram staining at this stage, or to purify colony growth in order to prepare liquid culture for stocks.
10. Follow steps 7-21 from the "Preparing a Bacterial Master and Working Stock from Frozen Stocks" section. Carry out step 7 & 8 up to three times if required.

Cryopreservation Check (At Least 48h After Freezing) Procedure

Growth and Purity Check:

1. Regulations recommend cryopreservation checks to be carried out for quality control (QC); Ideally 5-10% of the bank should be tested before use.
2. Transport the vial in dry ice to maintain the vial at low temperature.
3. To check growth and purity of cryopreserved stocks aseptically scrape a loopful of stock from the cryovial using a stripette, and with a 10µl loop quad streak onto agar selective to the bacteria being cultured. *Refer to Figure 1 for streaking technique.*

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

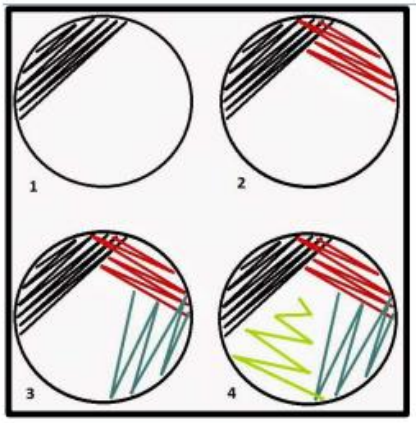


Figure1 – Streaking Technique

4. Appropriately grow anaerobically or aerobically at a temperature, and for a duration, specific to the bacteria being cultured. Check the stocks have grown and look at morphology to check purity.

Cell Recovery and Viability Check Using a CFU Count:

1. QC checks do not need to be performed on the first batch of master and working stocks prepared. When down to approximately the last three vials of stocks, prepare a new batch and carry out the quality control checks.
2. Cell recovery and viability data should be documented to demonstrate cell stability under storage conditions. For comparison, the viable cell counts before and after freezing are recorded and percentage survival is calculated.
3. Transport the vial on dry ice to maintain the vial at low temperature until you are ready to resuscitate the cells. Leave the vials on the ice to thaw.
4. Before opening the vial wipe the entire vial with a tissue soaked in 70% alcohol.
5. Centrifuge the cell suspension at approximately 14, 000 x g for 10 minutes.
6. After the centrifugation process, check the clarity of supernatant and visibility of a complete pellet and decant the supernatant without disturbing the cell pellet.
7. Then add 1ml of the required pre-warmed growth medium to the pellet to create your cell suspension and resuspend the cells to mix.
8. Perform serial 1/10 dilutions up to 10^{-6}. Prepare further dilutions if necessary.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

9. 100µl of the chosen dilutions (i.e. 10^{-4}, 10^{-5}, and 10^{-6}) should be pipetted onto appropriate agar selective for the bacteria being cultured. Spreaders should be used to spread the samples over the plates. Use a new spreader and pipette tip for each sample and dilution.
10. Plates should be labelled with the sample ID, serial dilution, date carried out, and initial of the person carrying out the procedure.
11. Appropriately grow anaerobically or aerobically at a temperature, and for a duration, specific to the bacteria being cultured.
12. After incubation, count the number of colonies using a countable dilution and work out the CFU/ml by multiplying the counted number of colonies x the dilution used x 10.

Storage of Bacterial Stocks Procedure

- Bacterial stocks should be stored in the -80 freezer in cryovials.
- Cryovials should contain 30% glycerol broth. Glycerol acts as a cryoprotectant which is important for stabilizing frozen bacteria as it prevents cell damage and keeps the cells viable. A glycerol stock of bacteria can be stably stored for many years.
- Avoid repeated thawing and refreezing of bacterial stocks as this will reduce cell viability.
- All cryovials should be labelled using a permanent water-resistant marker with the Hall Labs bacterial strain ID and individual project name and ID, date of preparation, and whether the stock is master or working stock where appropriate.
- The details stated on the cryovials should also be recorded on the Hall Lab Bacterial Archives Isolate spreadsheet as well as the freezer the stock is stored in, i.e. Freezer 123 or 148. If the stock is stored in freezer 123, the draw number (1-5), the bin (1-4) or rack number (1-5) within the draw, the tray letter (A-D) within the rack, and the box number within each tray should be recorded on the archive spreadsheet. If stocks are stored in freezer 148, the draw number (1-5), the bin number (A-D) within the draws, and the box number within the bins should be recorded on the archive spreadsheet.
- The freezers should be de-iced once a month to make sure the freezer is working appropriately.

SAFETY INFORMATION

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

Principal Hazards:

Biological Hazards – Bacterial stocks, (*Clostridium, Bacteroides* – Hazard group 2) faeces and bodily fluid samples.

Mechanical Hazards – Electrical hazard from use of centrifuge, vortex and micro mixer.

Cold burns from the -80 freezers.

Method Steps involving Higher Risk (and precautions to minimise risk): None.

Essential Precautions: All processing of biological material should be done inside the class II microbiological safety cabinet / Appropriate PPE should be worn. Refer to the Biological safety handbook and human tissue COP.

Personal Protective Clothing: Lab coat and Shieldskin gloves. Freezer gloves when using the -80 freezers

Any restrictions on use: Trained personnel only. Staff with access to human tissue samples will have the appropriate training through the QIB Human Studies Training program. Module 1 and 5 are mandatory training courses for any staff at QIB who work with human samples

Exceptions or Warnings: Should be included at relevant step in the procedure.

4. RISK STATEMENT

The overall risk is low if the guidelines are followed carefully.

All individuals using this procedure will be shown the risk assessment and given appropriate information, instruction and training in the risks and precautions necessary, including the use of any personal protective equipment required.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

SOP HEALTH RISK ASSESSMENT						
[1] Activity:	Preparation and Storage of Bacterial Master and Working Stocks					
[2] Location of activity:	QIB Level 1					
[3] Who is involved:	Trained personnel					
[4] Frequency of activity:	Daily, Weekly, or Monthly					
[5] Duration of Activity:	Short					
[6] Chemical Hazard Name:	**Hazard Statements**			**Route of Exposure***		**Quantity Used**
N/A	N/A			N/A		N/A
[7] Details of biological agents of risk to human health:						**GMRA Number**
Bacterial stocks Clostridium, Bacteroides – Hazard group 2						
Human faeces						
[8] Other Hazards:	Please ■ as necessary					
Cold Burns	■	Ionising Radiation	☐	Ultra Violet or Infra Red		☐
Dust	☐	Noise	☐	Pollen Sensitizer		☐
Repetitive Action	☐	Extreme Cold Environment (< 0°C)	☐	Lifting / Manual Handling		☐
Asphyxiation	☐	Cuts	☐	Electrical		■
Slips / trips / falls	☐	Display Screen Equipment	☐	Nanomaterials (man-made particles, tubes, rods, or fibers ≤100nm in size)		☐
Other (give details)	☐					
[9] Control Measures:	Please ■ as necessary					
Fume Cupboard	☐	Microbiological Safety Cabinet	■	Total Containment Cabinet		☐

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

Ventilated Bench	☐	Spill Tray	☐	Trained personnel only	■
Signs	☐	Reduce frequency/alternate activity	☐	Reduce duration of activity	☐
Sub divide a load	☐	2 persons lift of equipment	☐	Not for more than 1 hour	☐
Regular, short breaks	☐	Alternate activities	☐	Other (give details)	☐

[10] Personal Protection:		Please ■ as necessary			
Lab coat	■	Safety Glasses	☐	Face Shield	☐
Goggles	☐	Gloves *(Orange Shield)*	■	Protective Gloves *(Cryogenic)*	■
Ear defenders	☐	Other (give details)	☐		

[11] Is personal monitoring and/or health surveillance required?	Yes	■	No	☐
Details:	Hep B vaccination recommended.			

[12] Restrictions:		Please ■ as necessary			
No lone working	☐	Not to be left unattended	☐	Named persons only	☐
In restricted area	☐	Risk Assessment Required for New or Expectant Mothers	☐	Under constant supervision	☐
Not by under 18's	■	Other (give details)	☐		

[13] Level of Residual Risk:		Please ■ as necessary			
Low	■	Medium	☐	High	☐
Name of Assessor:		I.R. Serghiou	Date:	14 April 2020	

* Route of exposure; S = skin, I = ingestion, B = inhalation

5. DOCUMENTS

Human Tissue COP: https://intranet.nbi.ac.uk/cms/970

Chemical Tables: http://intranet/cms/2256

Good Laboratory Practice in the Use of Chemicals: http://intranet/cms/2199

Link to Laboratory Waste Disposal: http://intranet/cms/968

Biosafety Handbook: http://intranet/cms/970

6. RELATED PROCEDURES

Media Preparation and Storage SOP

Use of Centrifuge - Eppendorf 5810R SOP (GMH-LH-001)

Vortex Mixer SOP (GMH-LH-037)

Micro Mixer SOP

Preparation of 30% Glycerol Broth for Bacterial Stock Cryopreservation SOP (GMH-LH-053)

PEARL study sample MANAGEMENT: receipt, handling, storage and disposal (241880-LH-007)

BEAM Study sample collection from BioRepository (GMH-LH-060)

7. DOCUMENT HISTORY

Issue Number	Reason for update
001	First Issue

8. APPENDIX

Staff Training record for SOP (Link to: Preparation and Storage of Bacterial Master and Working Stocks Training.doc)

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 24 April 2020.

Appendix 11. PEARL Sample Sequencing MultiQC Summary

Sample	Raw Sequences	Host (%)	Non-Host Reads	Cleaned Reads
BSC-1-M12	38834238	0.00%	38832623	38806262
BSC-10-M12	16985274	0.00%	16984943	16972802
BSC-10-M4	17634086	0.00%	17633859	17626085
BSC-10-M8	15733365	0.04%	15726813	15714750
BSC-11-M12	32639104	0.00%	32637533	32612726
BSC-11-M8	21608669	0.00%	21608261	21595635
BSC-12-M8	20209729	0.00%	20209500	20168275
BSC-13-M4	31825118	0.00%	31823985	31793781
BSC-13-M8	20226896	0.00%	20226639	20217935
BSC-13-W1	32586258	0.00%	32585665	32560019
BSC-13-W3	27459421	0.00%	27458581	27434777
BSC-14-M12	19576496	0.00%	19576073	19568914
BSC-14-W3	23547017	0.00%	23546547	23538925
BSC-2-1-B	37943035	0.01%	37938255	37899793
BSC-2-1-M4	7732278	0.23%	7714565	7707980
BSC-2-1-W3	35624172	0.01%	35620250	35594524
BSC-2-2-B	24914868	0.00%	24914678	24877444
BSC-2-2-M12	30688685	0.00%	30688127	30670241
BSC-2-2-M8	249771	6.28%	234087	233842
BSC-2-2-W3	25640735	0.01%	25637708	25623466
BSC-3-B	29347722	0.01%	29345460	29323602
BSC-3-M12	37048295	0.02%	37040263	37016746
BSC-3-M8	25561966	0.00%	25561480	25544653
BSC-3-W3	31876683	0.01%	31874955	31850911
BSC-4-M12	20989562	0.00%	20989148	20977551
BSC-4-M8	20241396	0.01%	20240343	20220670
BSC-5-B	21580835	0.00%	21580481	21561023
BSC-5-M12	22500412	0.12%	22473974	22461333
BSC-5-M4	7279067	0.04%	7276121	7270471
BSC-5-M8	31464729	0.01%	31462976	31445014
BSC-6-M12	19717170	0.00%	19716968	19690180
BSC-6-M8	32426227	0.00%	32425494	32396904
BSC-6-W1	32521470	0.00%	32520359	32489680
BSC-6-W3	36094405	0.01%	36091467	36061530
BSC-7-B	30979441	0.01%	30977219	30963487
BSC-7-M8	21271547	0.00%	21271319	21260552
BSC-7-W1	26457319	0.00%	26456749	26437951
BSC-8-B	20848794	0.01%	20847052	20832285
BSC-8-M12	25928609	0.00%	25928280	25918160
BSC-8-M8	29148050	0.00%	29146864	29120127
BSC-8-W1	24743048	0.08%	24723914	24700359
BSC-8-W3	31095519	0.01%	31092616	31070542
BSC-9-B	34427392	0.00%	34426679	34391319
BSC-9-M12	27680043	0.00%	27679601	27668671
BSC-9-M4	18704937	0.00%	18704492	18676671
BSC-9-M8	31784781	0.01%	31782238	31751648
BSC-9-W1	30474619	0.00%	30474060	30438558
BSC-9-W3	32745867	0.00%	32745055	32709909
BST-1-B	1611798	85.76%	229469	219098
BST-1-M12	41311386	0.02%	41302019	40975555
BST-1-M4	36744407	0.06%	36721229	36590389
BST-1-M8	12116156	0.00%	12115860	11613250
BST-1-W1	7475801	1.83%	7338705	7316966
BST-1-W3	52161004	1.30%	51482290	51275352
BST-10-B	1214670	89.18%	131468	125976
BST-10-W1	30721109	0.01%	30719121	30497019
BST-10-W3	47408464	0.80%	47028284	46826570
BST-11-B	45273350	0.89%	44870022	44801265
BST-11-W1	97798477	0.03%	97772661	97496299
BST-11-W3	44366555	0.45%	44166996	44084774
BST-12-B	22372	0.17%	22335	20205
BST-12-M12	31138757	0.12%	31102802	31007547
BST-12-M4	42006062	0.00%	42004294	41795711
BST-12-M8	26338641	0.02%	26334469	26222003
BST-12-W1	35522938	0.13%	35477216	35400207
BST-12-W3	37183667	0.01%	37181682	36909618
BST-13-M12	29156777	0.47%	29018561	28946133
BST-13-M4	40894385	1.58%	40246640	40052247
BST-13-W1	39468276	8.09%	36274012	36173944

BST-13-W3	25429085	0.40%	25328155	25169848
BST-14-M12	55596570	11.71%	49083849	48948588
BST-14-M4	45460451	0.62%	45177273	44985701
BST-14-M8	46509604	0.06%	46480851	46310814
BST-14-W1	23943948	1.80%	23512952	23380069
BST-14-W3	35237457	0.04%	35223885	35057536
BST-2-1-B	23224603	0.01%	23221557	23015761
BST-2-1-M4	21904119	0.00%	21903982	21694611
BST-2-1-M8	38081419	0.01%	38078430	37850573
BST-2-1-M8	22267228	0.00%	22266318	22022642
BST-2-1-W1	26599364	0.01%	26595726	25987625
BST-2-1-W3	33679849	0.01%	33675165	33458751
BST-2-2-W3	31144705	0.02%	31139548	30947638
BST-3-B	24133728	1.54%	23760983	23612720
BST-3-M12	29366083	0.01%	29362188	29258358
BST-3-M4	25630277	0.04%	25620432	25301881
BST-3-W1	59026429	1.11%	58373814	57817658
BST-3-W3	32852063	0.14%	32805200	32661492
BST-4-B	88365655	0.13%	88248533	87087860
BST-4-M12	18584397	0.17%	18552765	18353516
BST-4-M4	38969040	0.00%	38968813	38772648
BST-4-M8	23130732	0.05%	23119650	22958923
BST-4-W1	63016313	0.00%	63014356	62614314
BST-4-W3	60863468	0.00%	60862537	60468599
BST-5-B	44237206	0.26%	44124033	43970966
BST-5-M12	40509271	0.04%	40491157	40482317
BST-5-M4	21807473	89.72%	2242163	2146527
BST-5-W1	22064863	0.08%	22046624	21884759
BST-5-W3	19364692	1.41%	19092088	18933828
BST-6-M12	34038769	0.03%	34029948	33942990
BST-6-M4	45827697	0.01%	45825195	45723476
BST-6-M8	34856206	0.00%	34854713	34070588
BST-6-W1	14077012	0.13%	14058627	14048314
BST-6-W3	19921497	0.07%	19907210	19844484
BST-7-B	54886674	0.13%	54817319	54678843
BST-7-M4	14866400	0.00%	14866151	14324678
BST-7-M8	39034307	0.03%	39021015	38907758
BST-7-W1	32577563	0.13%	32534283	32456829
BST-7-W3	68316877	0.01%	68307919	67898794
BST-8-W1	27291038	1.35%	26922410	26790772
BST-8-W3	42266829	0.19%	42187408	42065651
BST-8-W3	36491496	0.03%	36478968	36321043
BST-9-M12	30578108	0.06%	30558364	30415759
MSC-1-B	24801954	0.00%	24800989	24777684
MSC-1-T2	29471061	0.10%	29442953	29413573
MSC-11-B	28570762	0.00%	28570007	28543294
MSC-11-M4	19371471	0.00%	19371292	19354599
MSC-11-T3	17088653	0.01%	17087257	17074868
MSC-11-W1	24876074	0.01%	24872800	24848682
MSC-11-W3	23169681	0.00%	23169452	23153999
MSC-12-B	25326000	0.01%	25323293	25301232
MSC-12-M4	37117598	0.00%	37116178	37080500
MSC-12-T2	8974629	0.00%	8974240	8967522
MSC-12-W1	28371885	0.00%	28371030	28341018
MSC-12-W3	14710002	0.00%	14709693	14689852
MSC-13-B	25304157	0.00%	25303474	25280286
MSC-13-M4	19296106	0.01%	19295112	19285931
MSC-13-T2	23045265	0.00%	23044963	23024606
MSC-13-T3	27542774	0.01%	27541118	27516249
MSC-13-W1	20062115	0.01%	20059907	20037655
MSC-13-W3	23361411	0.01%	23359776	23335263
MSC-14-B	29458934	0.01%	29457166	29432131
MSC-14-T2	17716774	0.00%	17715942	17704079
MSC-14-W3	19836328	0.00%	19836132	19820501
MSC-2-B	14581275	0.00%	14580740	14575071
MSC-2-M4	19324821	0.06%	19312403	19294838
MSC-2-T2	32189936	0.00%	32189287	32151980
MSC-2-W1	24641840	0.00%	24641386	24621467
MSC-2-W3	27701868	0.01%	27697986	27675681
MSC-3-B	30578512	0.00%	30578128	30532049
MSC-3-M4	24593278	0.01%	24590002	24567715
MSC-3-T1	23478257	0.00%	23477815	23454474
MSC-3-T2	37865035	0.01%	37862545	37836291

MSC-3-T3	22864132	0.00%	22863847	22837530
MSC-3-W1	39016868	0.00%	39016141	38985162
MSC-3-W3	19593215	0.00%	19592589	19572699
MSC-4-W1	17176022	0.00%	17175532	17160974
MSC-5-B	27235571	0.01%	27233092	27207038
MSC-5-M4	23096699	0.00%	23096440	23072046
MSC-5-T1	19218528	0.00%	19217936	19197271
MSC-5-T2	20428419	0.00%	20428189	20410571
MSC-5-T3	26067526	0.00%	26066998	26046450
MSC-5-W1	18539192	0.00%	18539083	18469216
MSC-5-W3	20126594	0.00%	20126507	20079255
MSC-6-B	29523628	0.00%	29522877	29505574
MSC-6-M4	31267627	0.01%	31264646	31239047
MSC-6-T2	33283896	0.00%	33283677	33271304
MSC-6-T3	33491075	0.00%	33490558	33457965
MSC-6-W1	20384338	0.01%	20383066	20368474
MSC-6-W3	27649342	0.01%	27647595	27623766
MSC-7-B	24862559	0.00%	24861828	24845578
MSC-7-T3	20188587	0.00%	20188470	20169294
MSC-7-W3	24926177	0.00%	24925988	24909805
MSC-8-W3	21371395	0.00%	21371270	21353297
MSC-9-B	35814961	0.00%	35814250	35778661
MSC-9-T2	27936723	0.00%	27936361	27916458
MSC-9-T3	28410362	0.00%	28409313	28379288
MSC-9-W1	33294700	0.00%	33293459	33261948
MSC-9-W3	25863914	0.01%	25861986	25836492
MST-1-B	14305416	0.09%	14292613	14088668
MST-1-M4	23525936	0.02%	23522198	23169442
MST-1-T2	8201118	0.62%	8150470	8024392
MST-1-T3	24113579	0.02%	24108937	23886955
MST-1-W1	25182789	0.05%	25169133	24958966
MST-1-W3	23835678	0.09%	23813121	23575890
MST-10-T2	4061508	1.13%	4015655	3922292
MST-10-T3	6872022	0.05%	6868740	6785095
MST-10-W3	11336392	0.13%	11321974	10952600
MST-11-B	43786245	0.01%	43784011	43780036
MST-11-W1	4093405	0.49%	4073259	4067889
MST-11-W3	31073861	0.02%	31068547	30926635
MST-12-B	30187819	0.23%	30118317	30003755
MST-12-M4	4885011	0.42%	4864444	4803513
MST-12-T3	23374562	0.38%	23285080	23038938
MST-12-W3	33984334	0.25%	33898455	33726135
MST-13-M8	38652210	0.74%	38366644	38266714
MST-13-T2	843673	0.04%	843336	827843
MST-13-T3	24172746	0.02%	24168789	24022706
MST-14-B	39678106	0.03%	39665581	39347612
MST-14-M4	37370082	0.17%	37305538	37191262
MST-14-T2	60343592	0.01%	60336281	60043595
MST-14-W1	27192807	0.08%	27170864	27107726
MST-14-W3	18558701	0.08%	18544258	18317366
MST-2-T2	22630694	1.87%	22207361	21886912
MST-2-T3	16213874	0.75%	16091639	15942411
MST-2-W1	15293041	0.05%	15286066	14936533
MST-2-W3	19597422	0.14%	19570611	19115393
MST-3-B	19343963	0.12%	19320104	19187916
MST-3-M4	19848295	0.49%	19750446	19562244
MST-3-T1	29544833	0.09%	29518168	29330911
MST-3-T2	15373387	0.05%	15366392	15234192
MST-3-T3	19343639	0.02%	19339315	19186978
MST-3-W1	15612143	0.11%	15595159	15455149
MST-3-W3	58115449	0.05%	58088948	57849423
MST-4-B	17939925	8.50%	16415371	15937714
MST-4-M4	26195938	0.24%	26133458	25880561
MST-4-T2	41834833	0.04%	41819333	41021197
MST-4-T3	66986847	0.08%	66930605	66020704
MST-4-W1	8635760	0.10%	8626803	8530641
MST-4-W3	52524689	0.04%	52505981	52016681
MST-5-B	39687298	0.04%	39673361	39339950
MST-5-T1	46554935	0.05%	46531976	46086004
MST-5-T2	42916790	0.01%	42912727	42596898
MST-5-T3	19087241	0.02%	19083647	18842792
MST-5-W1	32271839	0.08%	32245992	31812094
MST-5-W3	24930119	0.03%	24922193	24718172

MST-6-B	23434988	0.40%	23342302	23214705
MST-6-M4	26945705	0.20%	26893048	26775380
MST-6-T2	16760741	0.07%	16748611	16357134
MST-6-T3	25133239	0.09%	25110627	24890742
MST-6-W1	25661094	0.16%	25620751	25369072
MST-6-W3	21851678	0.53%	21735841	21554567
MST-7-B	42880668	0.03%	42868293	42576916
MST-7-M4	43438266	0.15%	43375215	43241456
MST-7-T2	17275059	3.78%	16621902	16419630
MST-7-T3	9466864	0.07%	9459998	9446094
MST-7-W1	23214012	0.18%	23171120	22937141
MST-7-W3	15866911	0.11%	15849575	15755806
MST-8-B	23406541	0.04%	23396845	23299373
MST-8-T2	23848314	0.04%	23839242	23631487
MST-8-T3	27681110	0.07%	27662159	27436017
MST-8-W3	7583507	0.30%	7560841	7555071
MST-9-B	26785135	0.13%	26749066	26595952
MVC-1-B	24907201	0.00%	24906705	24891316
MVC-1-W1	20402675	3.10%	19770123	19754865
MVC-1-W3	24185647	0.01%	24184325	24166275
MVC-10-B	24197112	0.00%	24196705	24178910
MVC-10-T3	26510007	0.00%	26509684	26497282
MVC-10-W1	23526424	0.00%	23526051	23513705
MVC-10-W3	25100551	0.00%	25099983	25087712
MVC-11-T2	1183833	0.13%	1182315	1181809
MVC-11-T3	24215950	0.01%	24213232	24202329
MVC-11-W1	2558098	19.33%	2063609	2060208
MVC-11-W3	26475823	0.04%	26465298	26455455
MVC-12-B	13968223	0.00%	13967809	13955914
MVC-12-M4	26029020	0.01%	26027527	26015153
MVC-12-T3	993772	3.34%	960558	959494
MVC-12-W1	26128570	0.04%	26119224	26098418
MVC-12-W3	34381797	0.00%	34380756	34364778
MVC-13-B	25898005	0.02%	25893859	25880383
MVC-13-T2	29673814	1.44%	29245653	29225824
MVC-13-T3	22117754	0.00%	22116731	22099434
MVC-13-W1	29977684	0.00%	29976968	29948666
MVC-13-W3	30713774	0.03%	30703919	30686598
MVC-14-B	28403784	0.00%	28403148	28389841
MVC-14-M4	84397	59.22%	34417	34227
MVC-14-T2	19601051	0.03%	19594318	19581214
MVC-14-W1	25775811	0.00%	25775046	25760574
MVC-14-W3	12388869	0.00%	12388720	12381956
MVC-2- T3	11167	62.65%	4171	4090
MVC-2-B	29494759	0.01%	29492818	29470902
MVC-2-M4	45046	67.12%	14809	14646
MVC-2-W1	25051808	0.03%	25044523	25029527
MVC-2-W3	22610002	0.05%	22598978	22583681
MVC-3-B	26706416	0.11%	26677582	26658456
MVC-3-M4	20421654	0.01%	20420112	20406730
MVC-3-T1	31449430	0.02%	31443297	31421351
MVC-3-T2	23945101	0.01%	23943118	23922431
MVC-3-T3	26074159	0.02%	26069188	26044676
MVC-3-W1	20038361	0.00%	20037847	20022117
MVC-3-W3	20875393	0.03%	20868215	20850963
MVC-4-B	8030768	0.08%	8024345	8022200
MVC-4-T2	22707843	0.01%	22705113	22702651
MVC-4-T3	26642070	0.03%	26633349	26613817
MVC-4-W1	20381156	0.00%	20380934	20351280
MVC-4-W3	11992685	0.00%	11992565	11983068
MVC-5-B	27248985	0.55%	27097991	27094697
MVC-5-M4	30972136	6.17%	29060005	29038802
MVC-5-T1	24432265	0.00%	24431323	24408107
MVC-5-T2	29619124	0.17%	29568685	29534251
MVC-5-T3	23310619	0.04%	23301033	23278394
MVC-5-W1	24744724	0.12%	24713958	24698585
MVC-5-W3	21332807	0.03%	21327466	21311954
MVC-6-B	25534876	0.00%	25533916	25508016
MVC-6-M4	25371012	8.88%	23118795	23101636
MVC-6-T2	26527426	4.57%	25315168	25296622
MVC-6-T3	20446131	3.10%	19811956	19795998
MVC-6-W1	24256525	0.00%	24255928	24237659
MVC-6-W3	12762136	0.00%	12762058	12749928

MVC-7-B	19650359	1.44%	19367329	19357941
MVC-7-T2	25665829	8.35%	23522934	23505274
MVC-7-T3	24200056	15.11%	20543376	20525660
MVC-7-W1	26502336	0.00%	26502091	26454708
MVC-7-W3	26912617	0.07%	26893898	26881853
MVC-8-W1	26415397	0.00%	26414842	26390320
MVC-8-W3	21971563	0.00%	21971223	21943057
MVC-9-B	27837249	0.03%	27829395	27811676
MVC-9-M4	25179418	6.46%	23552193	23537378
MVC-9-T2	18328840	0.02%	18325136	18311634
MVC-9-T3	22122054	0.06%	22107885	22095230
MVC-9-W1	13236781	0.00%	13236225	13228594
MVC-9-W3	26311251	0.00%	26311090	26295058
Negative Control 2	1372	20.48%	1091	1059
Positive Control 10	22604104	0.00%	22603735	22588802
Positive Control 11	24349240	0.00%	24348870	24336495
Positive Control 14	32973771	0.00%	32973351	32955211
Positive Control 15	37262905	0.00%	37262044	37243296
Positive Control 2	21489979	0.01%	21488846	21477489
Positive Control 6	21162369	0.00%	21161991	21150473
Positive Control 7	19189541	0.00%	19189019	19178281

Note: Some metrics were not provided for all samples

Appendix 12. Promega Maxwell RSC 48 Blood DNA Kit Protocol

QUADRAM INSTITUTE BIOSCIENCE

STANDARD OPERATING PROCEDURE

Quadram Institute
Science Health Food Innovation

Title: Processing Blood Samples Using the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit

Applies to staff in: QIB	***Reference number: GMH-LH-082***
Author: Iliana Rosa Serghiou	***Approver: Dr. Lindsay Hall***
Issue date: 18 February 2021	***Review date: 17 February 2023***
QA and H&S Review: R Walding	***Version Number: 001***

AUTHORISATION: I confirm that this SOP is an accurate description of the procedure or method in use.

DECLARATION. The safety information contained in this document is to the best of my knowledge an accurate statement of the known and foreseeable hazards to health and of the safety precautions, which are to be followed to minimise the risks to health in the process. All users of this SOP are adequately trained to be competent in the process.

Approved by ISP/NC Leader, or Research Leader, or Senior Scientist, or Senior Administrator

Signed:	Print: Dr. Lindsay Hall	Date: 18 February 2021

TITLE: PROCESSING BLOOD SAMPLES USING THE PROMEGA MAXWELL® RSC 48 INSTRUMENT AND RSC BLOOD DNA KIT

1. PURPOSE OF PROCEDURE/METHOD AND ITS SCOPE

To perform a DNA extraction from blood samples using the Promega Maxwell® RSC 48 instrument and RSC Blood DNA Kit.

Ensure ethics approval is in place for the samples and that the handling of the samples has been risk assessed.

2. EQUIPMENT AND REAGENTS NEEDED

Equipment

- Class II Microbiological Safety Cabinet
- Vortex
- Centrifuge
- 20-200µl Pipette and filter tips
- 100-1000µl (1ml) Pipette and filter tips
- 1.5ml DNA LoBind Eppendorf Tubes
- Promega Maxwell® RSC 48 Instrument
- Heating Block

Reagents

- Promega Maxwell® RSC Blood DNA Kit
- Liquid Blood Samples

3. STEPS IN PROCEDURE (INCLUDING SAFETY INFORMATION)

NB:

- The total yield of genomic DNA from whole blood samples depends on the sample volume and number of white blood cells/ml.
- Fresh or frozen whole blood samples collected in EDTA, ACD or heparin tubes can be used. EDTA blood collection tubes are preferred if the purified DNA will be used in downstream amplification assays.
- Before beginning this procedure, please refer to the 'QIB Biosafety Handbook' and 'Code of Practice for work with human body tissues, fluids and other samples of human origin' for information on local procedures for working with biological samples.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 18 February 2021.

- During this COVID19 pandemic, additional risk assessments must be performed for any human samples being worked on at QIB, please contact H&S team to discuss any additional required measures for your sample types.
- Where appropriate, work should be conducted in a microbiological safety cabinet within a containment level 2 laboratory. Hands should be washed before and after working in the lab.
- Approved disinfectants should be available in case of a leak or any spillage:
    - Work surfaces Should be cleaned with Bioguard and then 70% ethanol before use and after use.
    - Specimen or reagent spills on any part of the Maxwell Instrument should be cleaned with a detergent-water solution, followed by a bacteriocidal spray or wipe, then water. Do not use bleach on any instrument parts.
    - After each run, the maxwell should be cleaned and a UV cycle run.
- A lab coat and protective gloves should be worn whenever working in the laboratory with human samples. Personnel at QIB working with human material are recommended to have a have Hepatitis B immunisation and should have attended training on working with human samples.
- A new clinical waste bag (double bagged) should be set up before work is started. All pipette tips or sharps should be put into a clinical waste sharps bin. Non-clinical biological waste should be sealed, room number marked on the bag and should be sent for autoclaving at the end of the working day.

DNA Extraction Procedure:

1. Mix all blood samples for at least 5 minutes at room temperature.
2. Prepare and label incubation tubes compatible with a heating block.
3. For each run use a negative control (PBS) to ensure no contamination has occurred during the run; This control will go through the same process as the blood samples.
4. Add 30µl of Proteinase K (PK) Solution to each incubation tube.
5. Add liquid blood (up to 300µl) to each incubation tube and 300µl of PBS to the negative control tube; This kit is designed to purify genomic DNA from up to 300µl of whole blood - If a larger volume is used contamination of samples could occur during the run.
6. Add 300µl of Lysis Buffer to each incubation tube.
7. Vortex each tube for 10 seconds.
8. Incubate each tube in the heating block (set to 56°C) for 20 minutes. During this incubation, prepare cartridges and turn on the Maxwell where you will proceed through a self-check. Change gloves before handling Maxwell® RSC Cartridges, RSC Plungers and Elution Tubes (0.5ml).
9. Select the RSC Blood DNA method, then select the cartridge positions to be used for this extraction run. The door should automatically open to load samples.
10. For each cartridge, vortex the well containing beads to resuspend them, then prepare the cartridges on the cartridge rack (Refer to Figure 1).
11. Place each cartridge in the rack with the label identifier side facing away from the 0.5ml Elution Tubes (Refer to Figure 1). Press down on the cartridge to snap it into position. Carefully peel back the

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 18 February 2021.

seal so that all plastic comes off the top of the cartridge - ensure that all sealing tape and any residual adhesive are removed.

12. Add 50µl of Elution Buffer to the bottom of each Elution Tube and place into the Elution Tube position for each cartridge (Refer to Figure 1).
13. Transfer each blood lysate sample from the incubation tube to well #1 of each cartridge; Well #1 is the well closest to the printed side and furthest from the elution tube (Refer to Figure 1).
14. Place one plunger into well 8 of each cartridge (Refer to Figure 1).
15. Transfer the Maxwell® RSC Cartridge Rack containing the prepared cartridges on the Maxwell® RSC 48 Instrument platform. Ensure that the cartridge rack is level on the instrument platform.

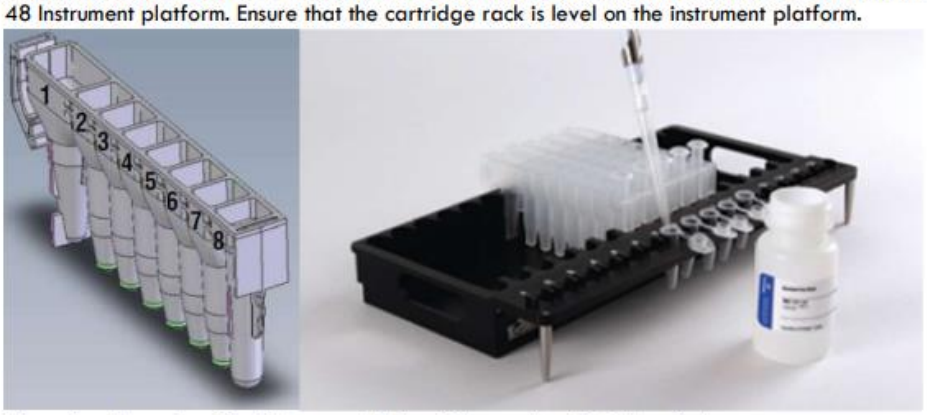


Figure 1 – Preparing Cartridges and Elution Tubes on the Cartridge Rack
(Taken from the Promega Maxwell RSC Blood DNA Kit Manual)

16. Press the Run/Stop button. The platform will retract. Close the door.
17. The Maxwell® RSC 48 Instrument will immediately begin the purification run. The screen will display the steps performed and the approximate time remaining in the run.
18. When the automated purification run is complete, the LCD screen will display a message that the method has ended.
19. Follow on-screen instructions at the end of the method to open door.
20. Verify that plungers are located in well 8 of the cartridges at the end of the run. If plungers are not removed from the magnetic plunger bar automatically, push them down gently by hand to remove them.
21. Remove the Maxwell® RSC Cartridge Rack from the instrument and dispose in the sharps bin, remove Elution Tubes containing DNA and cap the tubes.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 18 February 2021.

SAFETY INFORMATION

Principal Hazards: Biohazards from human blood material. Chemical hazards from small volume of kit materials, hot burns.

Method Steps involving Higher Risk (and precautions to minimise risk): Wear lab coat and Sheildskin gloves. Ensure human samples are handled as per the handling human tissue COP and HG2 RA.

Essential Precautions: Where appropriate, work should be performed in a microbiological safety cabinet within a containment level 2 laboratory. Appropriate PPE (laboratory coat & gloves) should be worn when handling human samples. Personnel at QIB working with human material are recommended to have Hepatitis B immunisation.

Personal Protective Clothing: Wear laboratory coat and Shieldskin gloves.

Any restrictions on use: Trained personnel only. Hep B Vaccination recommended. Not to be carried out by under 18s.

Exceptions or Warnings: Should be included at relevant step in the procedure.

4. RISK STATEMENT

The overall risk is LOW if the guidelines in this SOP are followed carefully.

All individuals using this procedure will be shown the risk assessment and given appropriate information, instruction and training in the risks and precautions necessary, including the use of any personal protective equipment required.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 18 February 2021.

SOP HEALTH RISK ASSESSMENT							
[1] Activity:	Promega Maxwell® RSC 48 DNA Extraction Using the RSC Blood DNA Kit						
[2] Location of activity:	QIB Laboratories						
[3] Who is involved:	Trained Lab Staff						
[4] Frequency of activity:	When required – approximately once a week						
[5] Duration of Activity:	Approximately 3 hours						
[6] Chemical Hazard Name:	**Hazard Statements**					**Route of Exposure***	**Quantity Used**
Proteinase K 20mg/ml	H317 May cause an allergic skin reaction. H334 May cause allergy or asthma symptoms or breathing difficulties if inhaled.					S, B	30ul per sample
Lysis Buffer – Contains guanidine hydrochloride and urea	H302 Harmful if swallowed. H315 Causes skin irritation. H318 Causes serious eye damage.					S, I, B	300ul per sample
Ethanol	H225 Highly flammable liquid and vapour H319 Causes serious eye irritation					S, I, B	Few mL for cleaning
Isopropanol	H225 Highly flammable liquid and vapour H319 Causes serious eye irritation H336 May cause drowsiness or dizziness.					S, I, B	Few mL for cleaning
[7] Details of biological agents of risk to human health:							**GMRA Number**
Hazard Group 2 bacteria (Human Blood) All such specimens must be treated as though they are infected with Blood Bourne Viruses (BBV) or other infectious agents.							
[8] Other Hazards:	Please ■ as necessary						
Hot Burns	■	Ionising Radiation**	☐	Ultra Violet or Infra Red	☐		

Dust	☐	Noise	☐	Pollen Sensitizer	☐
Repetitive Action	☐	Extreme Cold Environment (< 0°C)	☐	Lifting / Manual Handling	☐
Asphyxiation	☐	Cuts	☐	Electrical	☐
Slips / trips / falls	☐	Display Screen Equipment	☐	Nanomaterials (man-made particles, tubes, rods, or fibers ≤100nm in size)	☐
Other (give details)	☐				

[9] Control Measures:		Please ■ as necessary			
Fume Cupboard	☐	Microbiological Safety Cabinet	■	Total Containment Cabinet	☐
Ventilated Bench	■	Spill Tray	☐	Trained personnel only	■
Signs	☐	Reduce frequency/alternate activity	☐	Reduce duration of activity	☐
Sub divide a load	☐	2 persons lift of equipment	☐	Not for more than 1 hour	☐
Regular, short breaks	☐	Alternate activities	☐	Other (give details)	☐

[10] Personal Protection:		Please ■ as necessary			
Lab coat	■	Safety Glasses	☐	Face Shield	☐
Goggles	☐	Gloves *(Shieldskin)*	■	Protective Gloves *(specify)*	☐
Ear defenders	☐	Other (give details)	☐		

[11] Is personal monitoring and/or health surveillance required?	Yes	■	No	☐
Details:	Hep B vaccination recommended.			

[12] Restrictions:		Please ■ as necessary			
No lone working	☐	Not to be left unattended	☐	Named persons only	☐
In restricted area	☐	Risk Assessment Required for New or Expectant Mothers	☐	Under constant supervision	☐
Not by under 18's	■	Other (give details)	☐		

[13] Level of Residual Risk:		Please ■ as necessary			
Low	■	Medium	☐	High	☐
Name of Assessor:		Iliana Rosa Serghiou		Date:	11/02/2021

* Route of exposure; S = skin, I = ingestion, B = inhalation

5. DOCUMENTS

Chemical Tables: http://intranet/cms/2256

Good Laboratory Practice in the Use of Chemicals: http://intranet/cms/2199

Link to Laboratory Waste Disposal: http://intranet/cms/968

Biosafety Handbook: http://intranet/cms/970

6. RELATED PROCEDURES

- Use of Class II Microbiological Safety cabinet – GHFS-LH-004

7. DOCUMENT HISTORY

Issue Number	Reason for update
001	First Issue

8. APPENDIX

Staff Training record for SOP (Processing Blood Samples Using the Promega Maxwell® RSC 48 Instrument and RSC Blood DNA Kit Training.doc)

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 18 February 2021.

Appendix 13. FastPrep MPBio Kit QIB Protocol

QUADRAM INSTITUTE BIOSCIENCE

STANDARD OPERATING PROCEDURE

Quadram Institute
Science Health Food Innovation

Title: DNA extraction using MPbio fast DNA spin kit for soil

Applies to staff in: QIB Level 1	***Reference number: GMH-LH-084***
Author: Nancy Teng	***Approver: Dr L Hall***
***Issue date:* 05 November 2022**	***Review date:* 04 November 2024**
QA and H&S Review: R Walding	***Version Number: 002***

AUTHORISATION: I confirm that this SOP is an accurate description of the procedure or method in use.

DECLARATION. The safety information contained in this document is to the best of my knowledge an accurate statement of the known and foreseeable hazards to health and of the safety precautions, which are to be followed to minimise the risks to health in the process. All users of this SOP are adequately trained to be competent in the process.

Approved by ISP/NC Leader, or Research Leader, or Senior Scientist, or Senior Administrator

Signed:	Print: L Hall	Date: 05 November 2022

TITLE: DNA EXTRACTION USING MPBIO FAST DNA SPIN KIT FOR SOIL

1. PURPOSE OF PROCEDURE/METHOD AND ITS SCOPE

Use of MPbio fast DNA spin kit for soil to perform DNA extraction from faecal samples or bacterial cultures.

2. EQUIPMENT AND REAGENTS NEEDED

Equipment

- MPbio fastprep 24
- Centrifuge
- Class II safety cabinet for dealing faecal samples
- 5mL RNA free DNA free eppendorfs (Eppendorf, 0030119460)
- 2.0mL DNA Lobind Safe lock Eppendorf (Eppendorf, 0030108078)

Reagents

MPbio fast DNA spin kit for soil

Ethanol

3. STEPS IN PROCEDURE (INCLUDING SAFETY INFORMATION)

For the entirety of this procedure – wear orange shield skin gloves (or equivalent), changing them if contaminated, and a lab coat.

Human faeces, other bodily fluids and environmental samples should only be handled in the Class II microbiological safety cabinet. Please refer to COP for working with human tissues and relevant ethics agreement. Please ensure that the appropriate H&S screening has been undertaken by the worker prior to commencing work with human faeces or bodily fluids.

1. Add up to 500mg of sample to a lysing matrix E tube. (approx.. 200mg stool) **If weighing out stool samples make sure to use the Class II safety cabinet for steps 1-3.**
2. Add 980µl Sodium Phosphate Buffer to sample in Lysing Matrix E tube.
3. Add 120µl MT Buffer. **Wear safety glasses if performing this step on the bench.**
4. Homogenize in the FastPrep instrument for 3 mins at a speed setting of 6.0. **Ensure that the samples have been secured correctly before starting.**

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 5 October 2022.

5. Centrifuge at 14,000 x g for 15 minutes to pellet debris.
6. Add 250µl PPS to a clean 2.0ml microcentrifuge tube and transfer the sample supernatant to this.
7. Mix by shaking the tube by hand 10 times.
8. Centrifuge at 14,000 x g for 10 minutes to pellet the precipitate. Add 1ml of resuspended Binding Matrix to clean 5ml tube and transfer the sample supernatant to this. **Wear safety glasses during steps 8-11.**
9. Invert sample for 2 minutes allowing DNA to bind. Place the tubes in a rack and allow matrix to settle for 3 minutes.
10. Remove and discard 500µl of the supernatant – avoid the settled Binding Matrix.
11. Resuspend the Binding Matrix in the remaining supernatant. Transfer 700µl of the mixture to a SPIN filter and centrifuge at 14,000 x g for 1 minute. Empty the catch tube and transfer the left over mixture. Repeat until all the mixture has been added to the SPIN filter. Empty catch tube again. **During this step, each time a new tip is used, wipe the pipette barrel using distel wipes or blue roll soaked in ethanol to prevent cross contamination between samples.**
12. Add 500µl SEWS-M and gently resuspend the pellet using the force of the liquid from the pipette tip. **(Ensure ETHANOL has been added to SEWS-M before use.)**
13. Centrifuge at 14,000 x g for 1 minute. Empty the catch tube.
14. Centrifuge a second time (Dry spin) at 14,000 x g for 1 minute to get rid of residual wash solution. Discard the catch tube and replace with a clean Eppendorf tube.
15. Air dry the SPIN filter for 10 minutes at room temperature and for 5 minutes at 37°C
16. Add 65µl DES to the Binding Matrix
17. Incubate at room temperature for 5 minutes.
18. Centrifuge at 14,000 x g for 1 minute **(or 6600 x g for 2 minutes if spinning with lids open)** to bring eluted DNA into the Eppendorf. Discard the SPIN Filter and store at -20°C or temporarily at 4°C.

SAFETY INFORMATION

Principal Hazards:

- Exposure to biological material: faecal samples or bacterial cultures (see SOP pertaining to handling of human samples on intranet)
- Mechanical hazards, use of centrifuge and use of FastPrep 24. See relevant SOPs for safety information.
- Chemical hazards, **Wear gloves, lab coat and safety glasses when using the following reagents:**
  - MT Buffer – irritant eyes
  - Binding Matrix – harmful in contact with skin and if swallowed or inhaled – keep separate from acids. (avoid formation of aerosol by pipetting according to GLP)
  - PPS - irritant eyes and skin
  - Ethanol- highly flammable and can cause eye irritation

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 5 October 2022.

Method Steps involving Higher Risk (and precautions to minimise risk): None

Essential Precautions: All processing of biological material (human and environmental; samples) should be done inside the class II microbiological safety cabinet to prevent exposure. After the initial bacterial lysis step the procedure can be carried out on the bench.

Personal Protective Clothing: Basic safety protection should be used at all stages of the procedure i.e. lab coat and orange shield skin gloves (or equivalent). Safety specs to be worn when handling chemicals.

Any restrictions on use: Trained personnel only. Human sample training required.

Emergency Procedures: clean up any spills immediately. For biological spills, disinfect the area with Bioguard.
In an emergency dial 2222 for assistance and a first aider. If there is no response and you need an ambulance dial 999/112.
Exceptions or Warnings: should be included at relevant step in the procedure

4. RISK STATEMENT

The overall risk level is: **LOW**

All individuals using this procedure will be shown the risk assessment and given appropriate information, instruction and training in the risks and precautions necessary, including the use of any personal protective equipment required.

COVID-19 Precautionary measures

The handling of these samples has been assessed by the NBIP HSEQA team who have concluded that samples from these individuals would be very low risk of being contaminated with COVID-19 because the participants represent a vulnerable group within the community. The following guidance has been requested to be conducted in order to process these samples at the Quadram institute:

The following guidance has been requested to be conducted in order to process these samples at the Quadram institute:

- Before collecting any samples from participants, please ask if they, or anyone in their household has knowingly had COVID-19. Any participants that have had coronavirus should not donate samples until an individual risk assessment can be conducted.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 5 October 2022.

- Samples should all be handled in a designated Class II MSC until they would be classified as biologically inactive in a side room off the main lab. This is currently either the tissue processing laboratory on the lower ground floor or the microbiology laboratory on level 1.
- It is essential that anyone conducting work with these samples are particularly cautious over the use of gloves, regular hand washing and not touching the face.

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 5 October 2022.

SOP HEALTH RISK ASSESSMENT			
[1] Activity:	Use of MPbio fast DNA spin kit for soil to perform DNA extraction from faecal samples or bacterial cultures.		
[2] Location of activity:	Level 1		
[3] Who is involved:	Trained members of Level 1		
[4] Frequency of activity:	More than once per week		
[5] Duration of Activity:	4 hours		
[6] Chemical Hazard Name:	**Hazard Statements**	**Route of Exposure***	**Quantity Used**
MT Buffer	H302 Harmful if swallowed. H319 Causes serious eye irritation.	Skin	980ul per sample
Binding Matrix	H314 Causes severe skin burns and eye damage. H318 Causes serious eye damage. H302 Harmful if swallowed. H312 Harmful in contact with skin. H332 Harmful if inhaled.	Ingestion Skin Inhalation	1ml per sample
PPS	H318 Causes serious eye damage. H315 Causes skin irritation	Skin	250ul per sample
Ethanol	H225 Highly Flammable liquid vapour H319 Causes serious eye irritation	Skin	100mL
[7] Details of biological agents of risk to human health:			**GMRA Number**
Human faeces or biological cultures			

[8] Other Hazards:	Please ■ as necessary				
Hot or Cold Burns	☐	Ionising Radiation**	☐	Ultra Violet or Infra Red	☐
Dust	☐	Noise	☐	Pollen Sensitizer	☐
Repetitive Action	■	Extreme Cold Environment (< 0°C)	☐	Lifting / Manual Handling	☐

Asphyxiation	☐	Cuts	☐	Electrical	☐
Slips / trips / falls	☐	Display Screen Equipment	☐	Nanomaterials (man-made particles, tubes, rods, or fibers ≤100nm in size)	☐
Other (give details)	☐				
[9] Control Measures:		Please ■ as necessary			
Fume Cupboard	☐	Microbiological Safety Cabinet	■	Total Containment Cabinet	☐
Ventilated Bench	☐	Spill Tray	☐	Trained personnel only	■
Signs	☐	Reduce frequency/alternate activity	☐	Reduce duration of activity	☐
Sub divide a load	☐	2 persons lift of equipment	☐	Not for more than 1 hour	☐
Regular, short breaks	☐	Alternate activities	☐	Other (give details)	☐
[10] Personal Protection:		Please ■ as necessary			
Lab coat	■	Safety Glasses	■	Face Shield	☐
Goggles	☐	Gloves *(Shieldskin or equivalent)*	■	Protective Gloves *(specify)*	☐
Ear defenders	☐	Other (give details)	☐		
[11] Is personal monitoring and/or health surveillance required?			Yes	■	No ☐
Details:	Hep B Vaccination recommended All new starters working in the laboratories should have a base level lung function assessment and skin surveillance assessment with Occupational Health. Workers are reminded to do routine skin and respiratory monitoring and any concerns to skin or lung health are to be directly reported to the line manager and Occupational Health.				
[12] Restrictions:		Please ■ as necessary			
No lone working	☐	Not to be left unattended	☐	Named persons only	☐
In restricted area	☐	Risk Assessment Required for New or Expectant Mothers	☐	Under constant supervision	☐
Not by under 18's	☐	Other (give details)	☐		

[13] Level of Residual Risk:		Please ■ as necessary					
Low	■	Medium	☐	High			☐
Name of Assessor:		Nancy Teng			Date:	03 October 2022	

* Route of exposure; S = skin, I = ingestion, B = inhalation

5. DOCUMENTS

Good Laboratory Practice in the Use of Chemicals: http://intranet/cms/2199

Link to Laboratory Waste Disposal: http://intranet/cms/968

Biosafety Handbook: http://intranet/cms/970

6. RELATED PROCEDURES

- Procedure for use of Centrifuges– GHFS-LH-001
- Procedure for use of FastPrep24 – SS-LS-015
- Procedure for use of Class II Biological Safety Cabinet – GHFS- LH-004

7. DOCUMENT HISTORY

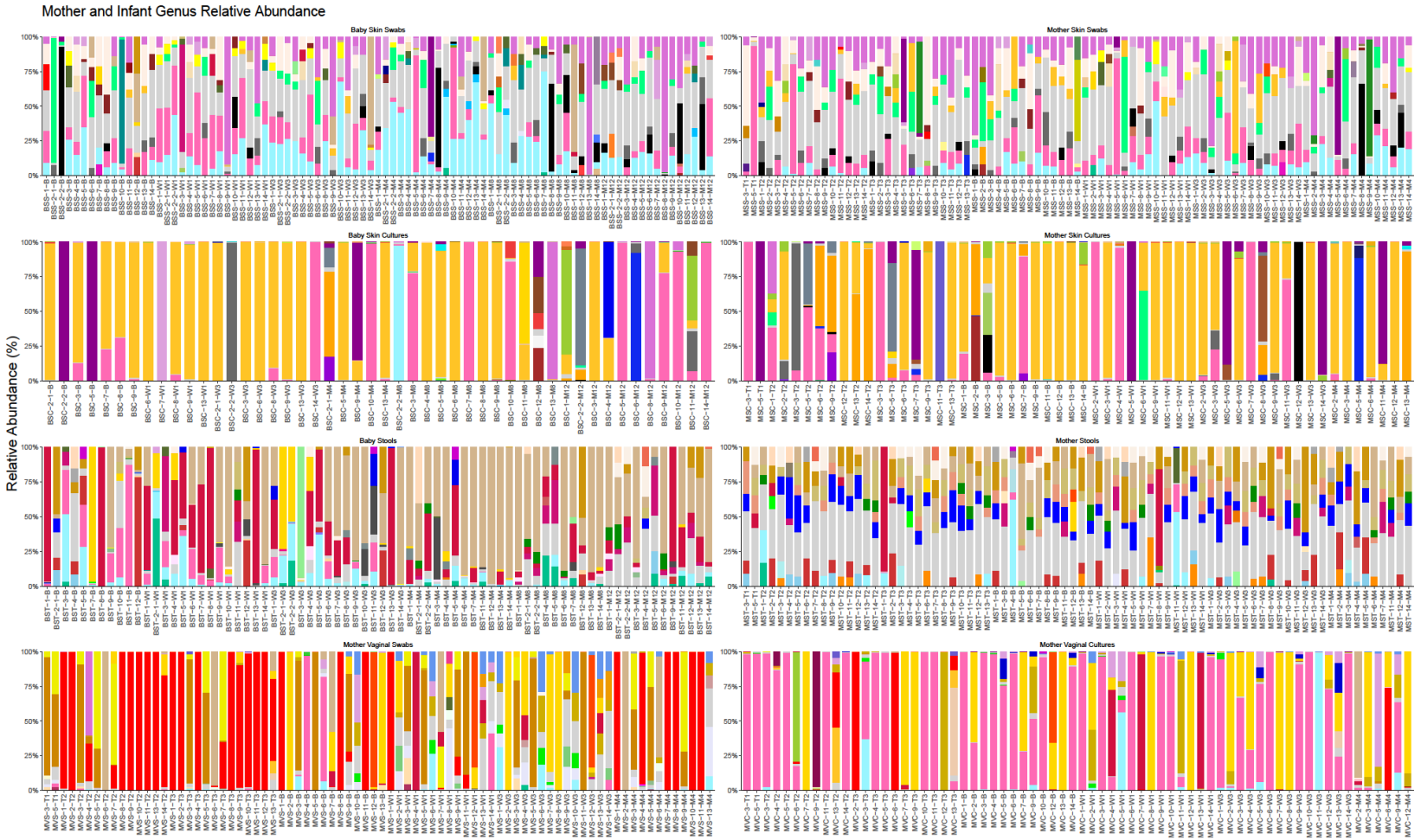
Issue Number	Reason for update
001	First issue
002	SOP reviewed; method updated

8. APPENDIX

Staff Training record for SOP (DNA extraction using MPbio fast DNA spin kit for soil Training.doc)

This is a controlled document maintained on electronic media. When appearing in paper form it should be checked against the master on the QA database to ensure that the current version is being used. This copy was printed on 5 October 2022.

Appendix 14. Maternal and Infant Genus Level Taxonomic Profiles Over Time



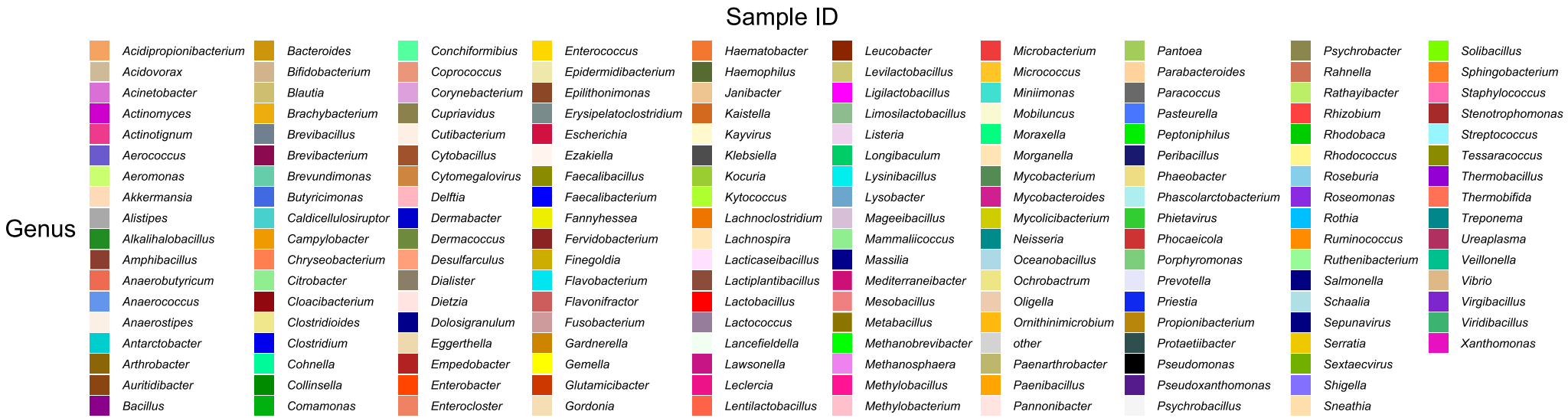
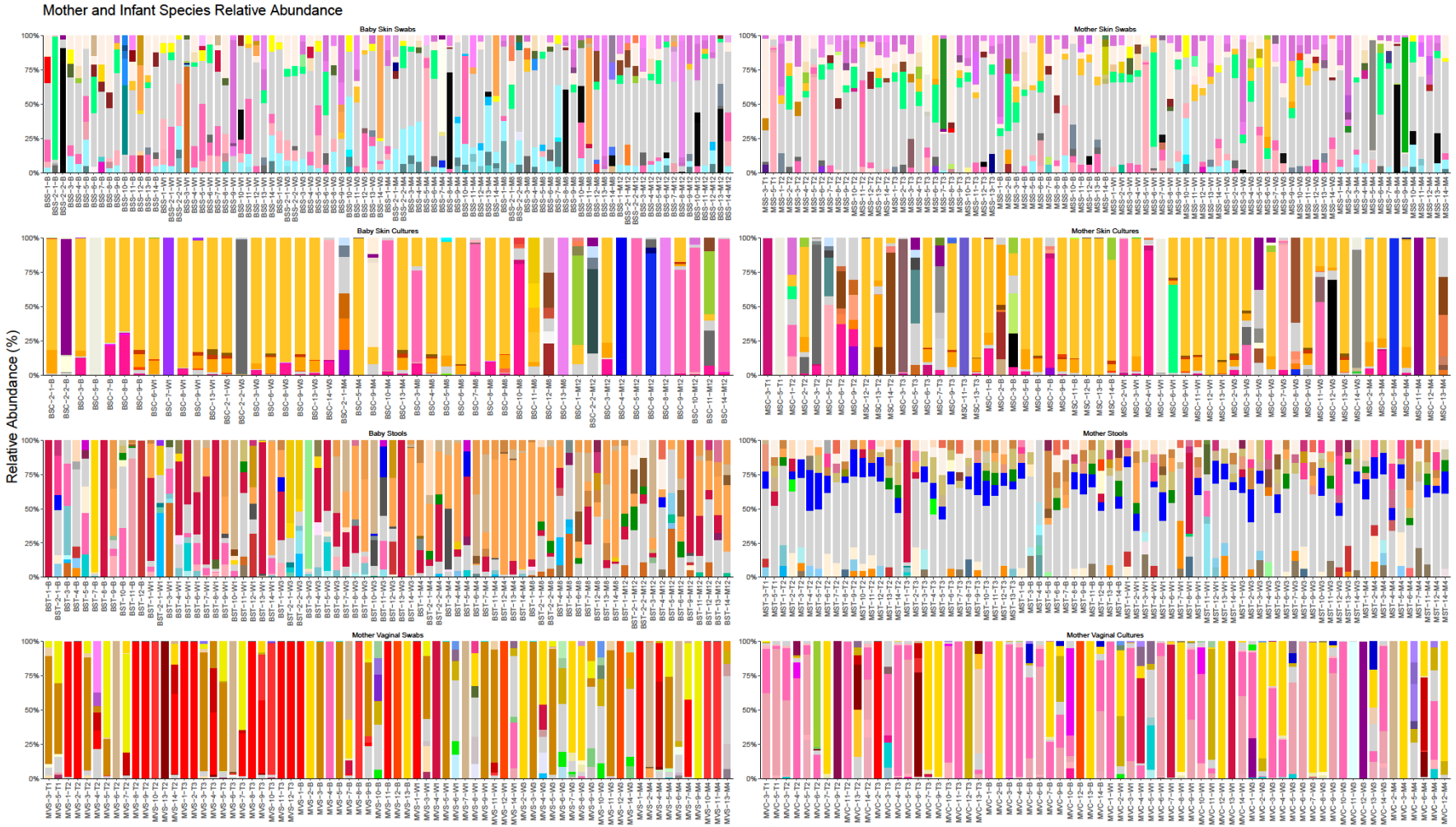


Figure Appendix 14. Maternal and Infant Taxonomic Relative Abundant Illumina Profiles to Genus Level. Composition data was generated from mothers and infants at multiple body sites (skin, stool and vaginal) using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bars displayed the relative abundance data. From each mother and infant, one stool (n=78 and n=74 respectively) and skin swab (from one forearm) (n=85 and n=89 respectively) were collected, and one vaginal swab was collected from each mother (n=85) over multiple timepoints: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Culture metagenomes from the maternal and infant skin (n=56 and n=48 respectively) and vaginal (67) swabs were also generated to identify culturable microbes. Each stacked bar represents the top 5 genera in an individual infant or mother over time. The Genus key allocates a unique colour to the genera represented in the stacked bars. This figure has an interactive element where the mouse can be hovered over the stacked bars to clearly identify the genera. **For the interactive version follow the link: Figure Appendix 14.html (the interactive version is also available in the supplementary files).**

Appendix 15. Maternal and Infant Species Level Taxonomic Profiles Over Time



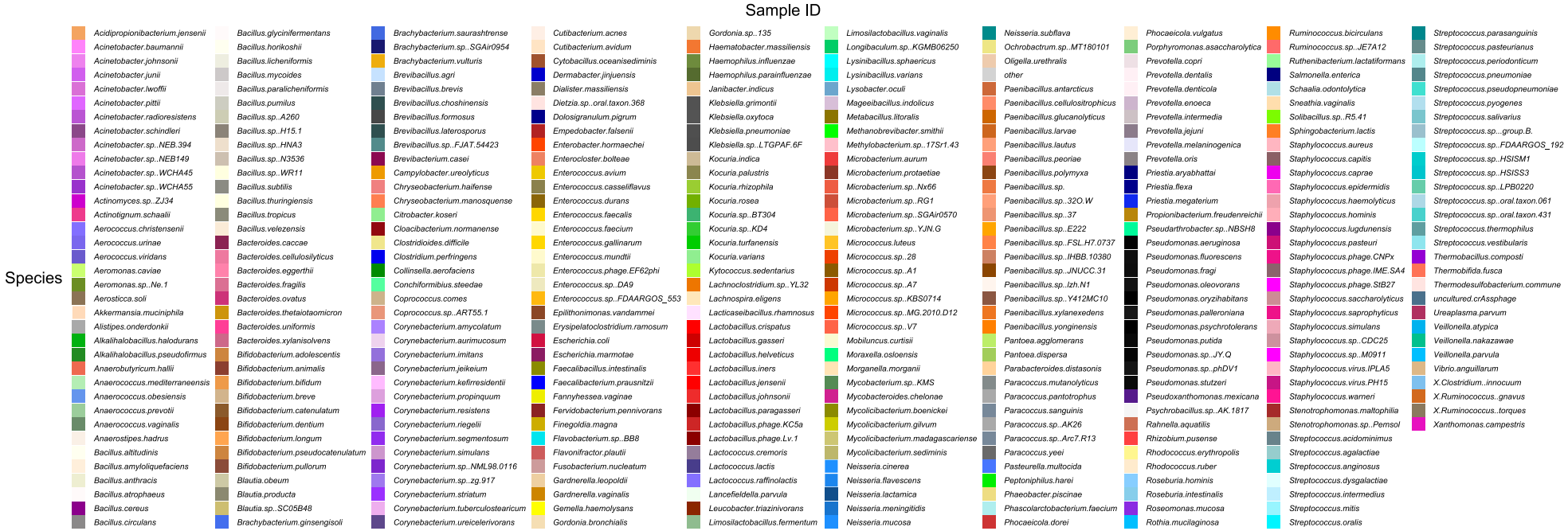


Figure Appendix 15. Maternal and Infant Taxonomic Relative Abundant Illumina Profiles to Species Level. Composition data was generated from mothers and infants at multiple body sites (skin, stool and vaginal) using shotgun metagenomics sequencing data, generated on the Illumina platform. Relative abundant taxonomic profiles were generated from the shotgun metagenomics data using a read mapping method with Kraken2 and Bracken. Stacked bars displayed the relative abundance data. From each mother and infant, one stool (n=78 and n=74 respectively) and skin swab (from one forearm) (n=85 and n=89 respectively) were collected, and one vaginal swab was collected from each mother (n=85) over multiple timepoints: Infant - Birth (B), Week 1 (W1), Week 3 (W3), Month 4 (M4), Month 8 (M8) and Month 12 (M12); Mother - Trimester 1 (T1), Trimester 2 (T2), Trimester 3 (T3), B, W1, W3 and M4. Culture metagenomes from the maternal and infant skin (n=56 and n=48 respectively) and vaginal (67) swabs were also generated to identify culturable microbes. Each stacked bar represents the top 5 species in an individual infant or mother over time. The Species key allocates a unique colour to the species represented in the stacked bars. This figure has an interactive element where the mouse can be hovered over the stacked bars to clearly identify the genera. **For the interactive version follow the link: Figure Appendix 15.html (the interactive version is also available in the supplementary files).**

Appendix 16. Maternal and Infant Sample Details for the Strain Detection and Identification Analysis

Name	Threshold	Clade ID	Family ID	Shared Strain	Mother Sample ID	Infant Sample ID	Mother Source	Infant Source	Mother Timepoint	Infant Timepoint
Micrococcus luteus	*0.03*	t__SGB16676	9	Not Shared	MSS-9-W3	BSS-9-W3	Maternal Skin Swab	Infant Skin Swab	Week 3	Week 3
Micrococcus luteus	*0.03*	t__SGB16676	9	Not Shared	MSS-9-T3	BSS-9-W3	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Week 3
Micrococcus luteus	*0.03*	t__SGB16676	13	Not Shared	MSS-13-T3	BSS-13-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Week 1
Micrococcus luteus	*0.03*	t__SGB16676	13	Not Shared	MSS-13-T2	BSS-13-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Week 1
Micrococcus luteus	*0.03*	t__SGB16676	13	Not Shared	MSS-13-W1	BSS-13-W1	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 1
Micrococcus luteus	*0.03*	t__SGB16676	3	Not Shared	MSS-3-B	BSS-3-M8	Maternal Skin Swab	Infant Skin Swab	Birth	Month 8
Micrococcus luteus	*0.03*	t__SGB16676	3	Not Shared	MSS-3-W3	BSS-3-M8	Maternal Skin Swab	Infant Skin Swab	Week 3	Month 8
Micrococcus luteus	*0.03*	t__SGB16676	13	Not Shared	MSS-13-T3	BSS-13-W3	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Week 3
Micrococcus luteus	*0.03*	t__SGB16676	13	Not Shared	MSS-13-T2	BSS-13-W3	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Week 3
Micrococcus luteus	*0.03*	t__SGB16676	13	Not Shared	MSS-13-W1	BSS-13-W3	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 3
Bifidobacterium dentium	*0.05*	t__SGB17234	6	Shared	MST-6-M4	BST-6-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Bifidobacterium dentium	*0.05*	t__SGB17234	6	Shared	MST-6-M4	BST-6-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Bifidobacterium dentium	*0.05*	t__SGB17234	6	Not Shared	MST-6-M4	BST-6-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium pseudocatenulatum	*0.05*	t__SGB17237	13	Shared	MST-13-W3	BST-13-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium pseudocatenulatum	*0.05*	t__SGB17237	13	Shared	MST-13-W1	BST-13-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium pseudocatenulatum	*0.05*	t__SGB17237	13	Shared	MST-13-W3	BST-13-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium pseudocatenulatum	*0.05*	t__SGB17237	13	Shared	MST-13-W1	BST-13-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-W1	BST-6-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-W3	BST-6-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-B	BST-6-M8	Maternal Stool	Infant Stool	Birth	Month 8
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-T3	BST-6-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-W1	BST-6-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-W3	BST-6-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-B	BST-6-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-T3	BST-6-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-W1	BST-6-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-W3	BST-6-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-B	BST-6-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium catenulatum	*0.03*	t__SGB17241	6	Not Shared	MST-6-T3	BST-6-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Not Shared	MST-11-W3	BST-11-B	Maternal Stool	Infant Stool	Week 3	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Not Shared	MST-11-B	BST-11-B	Maternal Stool	Infant Stool	Birth	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Not Shared	MST-11-T2	BST-11-B	Maternal Stool	Infant Stool	Trimester 2	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Not Shared	MST-11-T3	BST-11-B	Maternal Stool	Infant Stool	Trimester 3	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Not Shared	MST-11-M4	BST-11-B	Maternal Stool	Infant Stool	Month 4	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Not Shared	MST-1-T2	BST-1-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-B	BST-1-W3	Maternal Stool	Infant Stool	Birth	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-M4	BST-1-W3	Maternal Stool	Infant Stool	Month 4	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-W1	BST-1-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Not Shared	MST-1-T2	BST-1-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-B	BST-1-W1	Maternal Stool	Infant Stool	Birth	Week 1
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-M4	BST-1-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-W1	BST-1-W1	Maternal Stool	Infant Stool	Week 1	Week 1
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Not Shared	MST-1-T2	BST-1-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-B	BST-1-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-M4	BST-1-M4	Maternal Stool	Infant Stool	Month 4	Month 4

Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-W1	BST-1-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Not Shared	MST-1-T2	BST-1-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-B	BST-1-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-M4	BST-1-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-W1	BST-1-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Not Shared	MST-1-T2	BST-1-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-B	BST-1-M8	Maternal Stool	Infant Stool	Birth	Month 8
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-M4	BST-1-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	1	Shared	MST-1-W1	BST-1-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-M4	BST-12-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-T3	BST-12-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-T2	BST-12-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-M4	BST-12-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-T3	BST-12-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-T2	BST-12-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-M4	BST-12-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-T3	BST-12-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-T2	BST-12-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-M4	BST-12-W3	Maternal Stool	Infant Stool	Month 4	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-T3	BST-12-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	12	Shared	MST-12-T2	BST-12-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	3	Not Shared	MST-3-T3	BST-3-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	3	Not Shared	MST-3-T2	BST-3-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Shared	MST-11-W3	BST-11-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Shared	MST-11-B	BST-11-W1	Maternal Stool	Infant Stool	Birth	Week 1
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Shared	MST-11-T2	BST-11-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Shared	MST-11-T3	BST-11-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	11	Shared	MST-11-M4	BST-11-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-T2	BST-9-B	Maternal Stool	Infant Stool	Trimester 2	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-W3	BST-9-B	Maternal Stool	Infant Stool	Week 3	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-T3	BST-9-B	Maternal Stool	Infant Stool	Trimester 3	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-W1	BST-9-B	Maternal Stool	Infant Stool	Week 1	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-B	BST-9-B	Maternal Stool	Infant Stool	Birth	Birth
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-T2	BST-9-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-W3	BST-9-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-T3	BST-9-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-W1	BST-9-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Bifidobacterium adolescentis	*0.00542012*	t__SGB17244	9	Shared	MST-9-B	BST-9-W3	Maternal Stool	Infant Stool	Birth	Week 3
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T2	BST-6-M12	Maternal Vaginal Swab	Infant Stool	Trimester 2	Month 12
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T3	BST-6-M12	Maternal Vaginal Swab	Infant Stool	Trimester 3	Month 12
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-M4	BST-6-M12	Maternal Vaginal Swab	Infant Stool	Month 4	Month 12
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T2	BST-6-M4	Maternal Vaginal Swab	Infant Stool	Trimester 2	Month 4
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T3	BST-6-M4	Maternal Vaginal Swab	Infant Stool	Trimester 3	Month 4
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-M4	BST-6-M4	Maternal Vaginal Swab	Infant Stool	Month 4	Month 4
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T2	BST-6-W3	Maternal Vaginal Swab	Infant Stool	Trimester 2	Week 3
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T3	BST-6-W3	Maternal Vaginal Swab	Infant Stool	Trimester 3	Week 3
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-M4	BST-6-W3	Maternal Vaginal Swab	Infant Stool	Month 4	Week 3
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T2	BST-6-M8	Maternal Vaginal Swab	Infant Stool	Trimester 2	Month 8
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T3	BST-6-M8	Maternal Vaginal Swab	Infant Stool	Trimester 3	Month 8
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-M4	BST-6-M8	Maternal Vaginal Swab	Infant Stool	Month 4	Month 8

Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T2	BST-6-W1	Maternal Vaginal Swab	Infant Stool	Trimester 2	Week 1
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-T3	BST-6-W1	Maternal Vaginal Swab	Infant Stool	Trimester 3	Week 1
Bifidobacterium breve	*0.03*	t__SGB17247	6	Shared	MVS-6-M4	BST-6-W1	Maternal Vaginal Swab	Infant Stool	Month 4	Week 1
Bifidobacterium breve	*0.03*	t__SGB17247	2	Shared	MVS-2-M4	BST-2-1-M4	Maternal Vaginal Swab	Infant Stool	Month 4	Month 4
Bifidobacterium breve	*0.03*	t__SGB17247	2	Shared	MVS-2-M4	BST-2-1-M8	Maternal Vaginal Swab	Infant Stool	Month 4	Month 8
Bifidobacterium breve	*0.03*	t__SGB17247	2	Shared	MVS-2-M4	BST-2-2-M4	Maternal Vaginal Swab	Infant Stool	Month 4	NA
Bifidobacterium breve	*0.03*	t__SGB17247	2	Shared	MVS-2-M4	BST-2-2-M8	Maternal Vaginal Swab	Infant Stool	Month 4	NA
Bifidobacterium breve	*0.03*	t__SGB17247	2	Not Shared	MVS-2-M4	BST-2-1-M12	Maternal Vaginal Swab	Infant Stool	Month 4	Month 12
Bifidobacterium breve	*0.03*	t__SGB17247	2	Not Shared	MVS-2-M4	BST-2-2-M12	Maternal Vaginal Swab	Infant Stool	Month 4	NA
Bifidobacterium breve	*0.03*	t__SGB17247	7	Not Shared	MVS-7-W1	BST-7-M8	Maternal Vaginal Swab	Infant Stool	Week 1	Month 8
Bifidobacterium breve	*0.03*	t__SGB17247	7	Not Shared	MVS-7-B	BST-7-M8	Maternal Vaginal Swab	Infant Stool	Birth	Month 8
Bifidobacterium breve	*0.03*	t__SGB17247	13	Shared	MVS-13-T2	BST-13-M4	Maternal Vaginal Swab	Infant Stool	Trimester 2	Month 4
Bifidobacterium breve	*0.03*	t__SGB17247	13	Shared	MVS-13-T2	BST-13-M8	Maternal Vaginal Swab	Infant Stool	Trimester 2	Month 8
Bifidobacterium breve	*0.03*	t__SGB17247	7	Not Shared	MVS-7-W1	BST-7-W1	Maternal Vaginal Swab	Infant Stool	Week 1	Week 1
Bifidobacterium breve	*0.03*	t__SGB17247	7	Shared	MVS-7-B	BST-7-W1	Maternal Vaginal Swab	Infant Stool	Birth	Week 1
Bifidobacterium breve	*0.03*	t__SGB17247	7	Not Shared	MVS-7-W1	BST-7-W3	Maternal Vaginal Swab	Infant Stool	Week 1	Week 3
Bifidobacterium breve	*0.03*	t__SGB17247	7	Shared	MVS-7-B	BST-7-W3	Maternal Vaginal Swab	Infant Stool	Birth	Week 3
Bifidobacterium breve	*0.03*	t__SGB17247	7	Not Shared	MVS-7-W1	BST-7-M4	Maternal Vaginal Swab	Infant Stool	Week 1	Month 4
Bifidobacterium breve	*0.03*	t__SGB17247	7	Not Shared	MVS-7-B	BST-7-M4	Maternal Vaginal Swab	Infant Stool	Birth	Month 4
Bifidobacterium breve	*0.03*	t__SGB17247	13	Not Shared	MVS-13-T2	BST-13-M12	Maternal Vaginal Swab	Infant Stool	Trimester 2	Month 12
Bifidobacterium breve	*0.03*	t__SGB17247	14	Not Shared	MVS-14-W1	BST-14-M12	Maternal Vaginal Swab	Infant Stool	Week 1	Month 12
Bifidobacterium breve	*0.03*	t__SGB17247	14	Not Shared	MVS-14-W1	BST-14-M8	Maternal Vaginal Swab	Infant Stool	Week 1	Month 8
Bifidobacterium breve	*0.03*	t__SGB17247	14	Not Shared	MVS-14-W1	BST-14-M4	Maternal Vaginal Swab	Infant Stool	Week 1	Month 4
Bifidobacterium breve	*0.03*	t__SGB17247	14	Shared	MVS-14-W1	BST-14-W3	Maternal Vaginal Swab	Infant Stool	Week 1	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-W3	BST-1-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-T2	BST-1-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Shared	MSS-1-W3	BST-1-W1	Maternal Skin Swab	Infant Stool	Week 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	3	Shared	MST-3-T3	BST-3-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	3	Shared	MST-3-T2	BST-3-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	3	Shared	MST-3-W3	BST-3-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	3	Not Shared	MST-3-T1	BST-3-M12	Maternal Stool	Infant Stool	Trimester 1	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	3	Shared	MST-3-T3	BST-3-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	3	Shared	MST-3-T2	BST-3-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	3	Shared	MST-3-W3	BST-3-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	3	Not Shared	MST-3-T1	BST-3-M4	Maternal Stool	Infant Stool	Trimester 1	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-T3	BST-2-1-B	Maternal Stool	Infant Stool	Trimester 3	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W1	BST-2-1-B	Maternal Stool	Infant Stool	Week 1	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W3	BST-2-1-B	Maternal Stool	Infant Stool	Week 3	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-W3	BST-9-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-B	BST-9-W3	Maternal Stool	Infant Stool	Birth	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-T2	BST-9-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-T3	BST-9-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-W1	BST-9-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-W3	BST-9-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-B	BST-9-W1	Maternal Stool	Infant Stool	Birth	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-T2	BST-9-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-T3	BST-9-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-W1	BST-9-W1	Maternal Stool	Infant Stool	Week 1	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-W3	BSS-9-W3	Maternal Stool	Infant Skin Swab	Week 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-B	BSS-9-W3	Maternal Stool	Infant Skin Swab	Birth	Week 3

Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-T2	BSS-9-W3	Maternal Stool	Infant Skin Swab	Trimester 2	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-T3	BSS-9-W3	Maternal Stool	Infant Skin Swab	Trimester 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-W1	BSS-9-W3	Maternal Stool	Infant Skin Swab	Week 1	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-W3	BST-9-B	Maternal Stool	Infant Stool	Week 3	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-B	BST-9-B	Maternal Stool	Infant Stool	Birth	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-T2	BST-9-B	Maternal Stool	Infant Stool	Trimester 2	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-T3	BST-9-B	Maternal Stool	Infant Stool	Trimester 3	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	9	Shared	MST-9-W1	BST-9-B	Maternal Stool	Infant Stool	Week 1	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	4	Not Shared	MST-4-T2	BST-4-B	Maternal Stool	Infant Stool	Trimester 2	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	4	Not Shared	MST-4-T3	BST-4-B	Maternal Stool	Infant Stool	Trimester 3	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Shared	MST-13-T3	BST-13-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Shared	MST-13-B	BST-13-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Shared	MST-13-W3	BST-13-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Shared	MST-13-W1	BST-13-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Shared	MST-13-T3	BST-13-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Shared	MST-13-B	BST-13-M8	Maternal Stool	Infant Stool	Birth	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Shared	MST-13-W3	BST-13-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Shared	MST-13-W1	BST-13-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	4	Not Shared	MST-4-T2	BST-4-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	4	Not Shared	MST-4-T3	BST-4-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	4	Not Shared	MST-4-T2	BST-4-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	4	Not Shared	MST-4-T3	BST-4-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-T3	BST-2-2-M4	Maternal Stool	Infant Stool	Trimester 3	NA
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W1	BST-2-2-M4	Maternal Stool	Infant Stool	Week 1	NA
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W3	BST-2-2-M4	Maternal Stool	Infant Stool	Week 3	NA
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-T3	BST-2-1-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W1	BST-2-1-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W3	BST-2-1-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-T3	BST-2-1-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W1	BST-2-1-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W3	BST-2-1-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-T3	BST-2-2-M8	Maternal Stool	Infant Stool	Trimester 3	NA
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W1	BST-2-2-M8	Maternal Stool	Infant Stool	Week 1	NA
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Not Shared	MST-2-W3	BST-2-2-M8	Maternal Stool	Infant Stool	Week 3	NA
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Not Shared	MST-13-T3	BST-13-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Not Shared	MST-13-B	BST-13-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Not Shared	MST-13-W3	BST-13-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	13	Not Shared	MST-13-W1	BST-13-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-B	BST-11-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-T2	BST-11-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-M4	BST-11-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-W3	BST-11-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-T3	BST-11-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-B	BSS-11-M8	Maternal Stool	Infant Skin Swab	Birth	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-T2	BSS-11-M8	Maternal Stool	Infant Skin Swab	Trimester 2	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-M4	BSS-11-M8	Maternal Stool	Infant Skin Swab	Month 4	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-W3	BSS-11-M8	Maternal Stool	Infant Skin Swab	Week 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-T3	BSS-11-M8	Maternal Stool	Infant Skin Swab	Trimester 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W1	BSS-14-W3	Maternal Stool	Infant Skin Swab	Week 1	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-M4	BSS-14-W3	Maternal Stool	Infant Skin Swab	Month 4	Week 3

Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-B	BSS-14-W3	Maternal Stool	Infant Skin Swab	Birth	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-T2	BSS-14-W3	Maternal Stool	Infant Skin Swab	Trimester 2	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W3	BSS-14-W3	Maternal Stool	Infant Skin Swab	Week 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W1	BST-14-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-M4	BST-14-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-B	BST-14-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-T2	BST-14-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W3	BST-14-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W1	BST-14-W1	Maternal Stool	Infant Stool	Week 1	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-M4	BST-14-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-B	BST-14-W1	Maternal Stool	Infant Stool	Birth	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-T2	BST-14-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W3	BST-14-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W1	BST-14-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-M4	BST-14-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-B	BST-14-M8	Maternal Stool	Infant Stool	Birth	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-T2	BST-14-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W3	BST-14-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W1	BSS-14-M4	Maternal Stool	Infant Skin Swab	Week 1	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-M4	BSS-14-M4	Maternal Stool	Infant Skin Swab	Month 4	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-B	BSS-14-M4	Maternal Stool	Infant Skin Swab	Birth	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-T2	BSS-14-M4	Maternal Stool	Infant Skin Swab	Trimester 2	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W3	BSS-14-M4	Maternal Stool	Infant Skin Swab	Week 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W1	BST-14-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-M4	BST-14-W3	Maternal Stool	Infant Stool	Month 4	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-B	BST-14-W3	Maternal Stool	Infant Stool	Birth	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-T2	BST-14-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W3	BST-14-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W1	BST-14-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-M4	BST-14-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-B	BST-14-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-T2	BST-14-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	14	Not Shared	MST-14-W3	BST-14-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Shared	MST-2-T3	BST-2-2-M12	Maternal Stool	Infant Stool	Trimester 3	NA
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Shared	MST-2-W1	BST-2-2-M12	Maternal Stool	Infant Stool	Week 1	NA
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Shared	MST-2-W3	BST-2-2-M12	Maternal Stool	Infant Stool	Week 3	NA
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Shared	MST-2-T3	BST-2-1-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Shared	MST-2-W1	BST-2-1-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	2	Shared	MST-2-W3	BST-2-1-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-W3	BST-12-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T3	BST-12-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T2	BST-12-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-B	BST-12-W1	Maternal Stool	Infant Stool	Birth	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-M4	BST-12-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-W3	BST-12-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T3	BST-12-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T2	BST-12-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-B	BST-12-W3	Maternal Stool	Infant Stool	Birth	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-M4	BST-12-W3	Maternal Stool	Infant Stool	Month 4	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-W3	BST-12-M4	Maternal Stool	Infant Stool	Week 3	Month 4

Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T3	BST-12-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T2	BST-12-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-B	BST-12-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-M4	BST-12-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-W3	BST-12-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T3	BST-12-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T2	BST-12-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-B	BST-12-M8	Maternal Stool	Infant Stool	Birth	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-M4	BST-12-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-W3	BST-12-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T3	BST-12-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-T2	BST-12-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-B	BST-12-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	12	Shared	MST-12-M4	BST-12-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Not Shared	MST-7-B	BST-7-M8	Maternal Stool	Infant Stool	Birth	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-T2	BST-7-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-M4	BST-7-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-W3	BST-7-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-W1	BST-7-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-T3	BST-7-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Not Shared	MST-7-B	BST-7-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-T2	BST-7-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-M4	BST-7-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-W3	BST-7-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-W1	BST-7-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-T3	BST-7-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Not Shared	MST-7-B	BST-7-W1	Maternal Stool	Infant Stool	Birth	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-T2	BST-7-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-M4	BST-7-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-W3	BST-7-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-W1	BST-7-W1	Maternal Stool	Infant Stool	Week 1	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-T3	BST-7-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Not Shared	MST-7-B	BST-7-W3	Maternal Stool	Infant Stool	Birth	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-T2	BST-7-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-M4	BST-7-W3	Maternal Stool	Infant Stool	Month 4	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-W3	BST-7-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-W1	BST-7-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	7	Shared	MST-7-T3	BST-7-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-B	BST-11-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-T2	BST-11-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-M4	BST-11-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-W3	BST-11-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-T3	BST-11-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-B	BST-11-W1	Maternal Stool	Infant Stool	Birth	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-T2	BST-11-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-M4	BST-11-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-W3	BST-11-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Shared	MST-11-T3	BST-11-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Bifidobacterium longum	*0.03413473*	t__SGB17248	5	Shared	MST-5-B	BST-5-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	5	Shared	MST-5-T2	BST-5-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12

Bifidobacterium longum	*0.03413473*	t__SGB17248	5	Shared	MST-5-M4	BST-5-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	5	Shared	MST-5-T3	BST-5-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	5	Shared	MST-5-W1	BST-5-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	5	Shared	MST-5-T1	BST-5-M12	Maternal Stool	Infant Stool	Trimester 1	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	5	Shared	MST-5-W3	BST-5-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-B	BST-6-M8	Maternal Stool	Infant Stool	Birth	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-T2	BST-6-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-T3	BST-6-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-W3	BST-6-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-W1	BST-6-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-M4	BST-6-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-B	BST-6-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-T2	BST-6-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-T3	BST-6-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-W3	BST-6-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-W1	BST-6-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	6	Shared	MST-6-M4	BST-6-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	8	Shared	MST-8-T3	BST-8-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	8	Shared	MST-8-B	BST-8-W3	Maternal Stool	Infant Stool	Birth	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	8	Shared	MST-8-T2	BST-8-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-B	BST-11-B	Maternal Stool	Infant Stool	Birth	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-T2	BST-11-B	Maternal Stool	Infant Stool	Trimester 2	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-M4	BST-11-B	Maternal Stool	Infant Stool	Month 4	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-W3	BST-11-B	Maternal Stool	Infant Stool	Week 3	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	11	Not Shared	MST-11-T3	BST-11-B	Maternal Stool	Infant Stool	Trimester 3	Birth
Bifidobacterium longum	*0.03413473*	t__SGB17248	4	Not Shared	MST-4-T2	BST-4-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	4	Shared	MST-4-T3	BST-4-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-W3	BST-1-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-T2	BST-1-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Shared	MSS-1-W3	BST-1-M12	Maternal Skin Swab	Infant Stool	Week 3	Month 12
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-W3	BST-1-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-T2	BST-1-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Shared	MSS-1-W3	BST-1-W3	Maternal Skin Swab	Infant Stool	Week 3	Week 3
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-W3	BST-1-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-T2	BST-1-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Shared	MSS-1-W3	BST-1-M4	Maternal Skin Swab	Infant Stool	Week 3	Month 4
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-W3	BST-1-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Not Shared	MST-1-T2	BST-1-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium longum	*0.03413473*	t__SGB17248	1	Shared	MSS-1-W3	BST-1-M8	Maternal Skin Swab	Infant Stool	Week 3	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Shared	MST-7-M4	BST-7-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Shared	MST-7-T2	BST-7-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Not Shared	MST-7-B	BST-7-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Shared	MST-7-W3	BST-7-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Shared	MST-7-T3	BST-7-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	13	Shared	MST-13-W3	BST-13-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	13	Shared	MST-13-W1	BST-13-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	13	Shared	MST-13-T3	BST-13-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	13	Shared	MST-13-W3	BST-13-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	13	Shared	MST-13-W1	BST-13-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	13	Shared	MST-13-T3	BST-13-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8

Bifidobacterium bifidum	*0.0204073*	t__SGB17256	13	Shared	MST-13-W3	BST-13-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	13	Shared	MST-13-W1	BST-13-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	13	Shared	MST-13-T3	BST-13-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-W3	BST-14-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-W1	BST-14-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-M4	BST-14-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-T2	BST-14-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-B	BST-14-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-W3	BST-14-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-W1	BST-14-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-M4	BST-14-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-T2	BST-14-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-B	BST-14-M8	Maternal Stool	Infant Stool	Birth	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-W3	BST-14-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-W1	BST-14-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-M4	BST-14-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-T2	BST-14-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-B	BST-14-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-W3	BST-14-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-W1	BST-14-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-M4	BST-14-W3	Maternal Stool	Infant Stool	Month 4	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-T2	BST-14-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	14	Shared	MST-14-B	BST-14-W3	Maternal Stool	Infant Stool	Birth	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-M4	BSS-11-M8	Maternal Stool	Infant Skin Swab	Month 4	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-T3	BSS-11-M8	Maternal Stool	Infant Skin Swab	Trimester 3	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-B	BSS-11-M8	Maternal Stool	Infant Skin Swab	Birth	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-T2	BSS-11-M8	Maternal Stool	Infant Skin Swab	Trimester 2	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-W3	BSS-11-M8	Maternal Stool	Infant Skin Swab	Week 3	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-M4	BST-11-M12	Maternal Stool	Infant Stool	Month 4	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-T3	BST-11-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-B	BST-11-M12	Maternal Stool	Infant Stool	Birth	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-T2	BST-11-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-W3	BST-11-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-M4	BST-11-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-T3	BST-11-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-B	BST-11-M4	Maternal Stool	Infant Stool	Birth	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-T2	BST-11-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-W3	BST-11-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-M4	BST-11-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-T3	BST-11-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-B	BST-11-W1	Maternal Stool	Infant Stool	Birth	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-T2	BST-11-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Shared	MST-11-W3	BST-11-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-W3	BST-9-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-T3	BST-9-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-W1	BST-9-W1	Maternal Stool	Infant Stool	Week 1	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-T2	BST-9-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-B	BST-9-W1	Maternal Stool	Infant Stool	Birth	Week 1
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-W3	BST-9-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-T3	BST-9-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3

Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-W1	BST-9-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-T2	BST-9-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-B	BST-9-W3	Maternal Stool	Infant Stool	Birth	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-W3	BST-9-B	Maternal Stool	Infant Stool	Week 3	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-T3	BST-9-B	Maternal Stool	Infant Stool	Trimester 3	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-W1	BST-9-B	Maternal Stool	Infant Stool	Week 1	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-T2	BST-9-B	Maternal Stool	Infant Stool	Trimester 2	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-B	BST-9-B	Maternal Stool	Infant Stool	Birth	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-W3	BSS-9-W3	Maternal Stool	Infant Skin Swab	Week 3	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-T3	BSS-9-W3	Maternal Stool	Infant Skin Swab	Trimester 3	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-W1	BSS-9-W3	Maternal Stool	Infant Skin Swab	Week 1	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-T2	BSS-9-W3	Maternal Stool	Infant Skin Swab	Trimester 2	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	9	Shared	MST-9-B	BSS-9-W3	Maternal Stool	Infant Skin Swab	Birth	Week 3
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Not Shared	MST-11-M4	BST-11-B	Maternal Stool	Infant Stool	Month 4	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Not Shared	MST-11-T3	BST-11-B	Maternal Stool	Infant Stool	Trimester 3	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Not Shared	MST-11-B	BST-11-B	Maternal Stool	Infant Stool	Birth	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Not Shared	MST-11-T2	BST-11-B	Maternal Stool	Infant Stool	Trimester 2	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	11	Not Shared	MST-11-W3	BST-11-B	Maternal Stool	Infant Stool	Week 3	Birth
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Shared	MST-7-M4	BST-7-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Shared	MST-7-T2	BST-7-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Not Shared	MST-7-B	BST-7-M8	Maternal Stool	Infant Stool	Birth	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Shared	MST-7-W3	BST-7-M8	Maternal Stool	Infant Stool	Week 3	Month 8
Bifidobacterium bifidum	*0.0204073*	t__SGB17256	7	Shared	MST-7-T3	BST-7-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
		t__SGB17278	6	Not Shared	MST-6-T2	BST-6-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Staphylococcus hominis	*0.03*	t__SGB7858	9	Shared	MSS-9-W1	BST-9-W1	Maternal Skin Swab	Infant Stool	Week 1	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	9	Shared	MSS-9-B	BST-9-W1	Maternal Skin Swab	Infant Stool	Birth	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	9	Shared	MSS-9-W1	BST-9-W3	Maternal Skin Swab	Infant Stool	Week 1	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	9	Shared	MSS-9-B	BST-9-W3	Maternal Skin Swab	Infant Stool	Birth	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	9	Shared	MSS-9-W1	BST-9-B	Maternal Skin Swab	Infant Stool	Week 1	Birth
Staphylococcus hominis	*0.03*	t__SGB7858	9	Shared	MSS-9-B	BST-9-B	Maternal Skin Swab	Infant Stool	Birth	Birth
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-M4	BSS-5-M4	Maternal Skin Swab	Infant Skin Swab	Month 4	Month 4
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T1	BSS-5-M4	Maternal Skin Swab	Infant Skin Swab	Trimester 1	Month 4
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-W1	BSS-5-M4	Maternal Skin Swab	Infant Skin Swab	Week 1	Month 4
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T2	BSS-5-M4	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Month 4
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-W3	BSS-5-M4	Maternal Skin Swab	Infant Skin Swab	Week 3	Month 4
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T3	BSS-5-M4	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Month 4
Staphylococcus hominis	*0.03*	t__SGB7858	5	Shared	MSS-5-M4	BSS-5-W3	Maternal Skin Swab	Infant Skin Swab	Month 4	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Shared	MSS-5-T1	BSS-5-W3	Maternal Skin Swab	Infant Skin Swab	Trimester 1	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Shared	MSS-5-W1	BSS-5-W3	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Shared	MSS-5-T2	BSS-5-W3	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Shared	MSS-5-W3	BSS-5-W3	Maternal Skin Swab	Infant Skin Swab	Week 3	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T3	BSS-5-W3	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	6	Not Shared	MSS-6-T2	BSS-6-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	6	Not Shared	MSS-6-B	BSS-6-W1	Maternal Skin Swab	Infant Skin Swab	Birth	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	6	Not Shared	MSS-6-W3	BSS-6-W1	Maternal Skin Swab	Infant Skin Swab	Week 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	6	Not Shared	MSS-6-T2	BST-6-W1	Maternal Skin Swab	Infant Stool	Trimester 2	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	6	Shared	MSS-6-B	BST-6-W1	Maternal Skin Swab	Infant Stool	Birth	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	6	Shared	MSS-6-W3	BST-6-W1	Maternal Skin Swab	Infant Stool	Week 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-M4	BSS-5-W1	Maternal Skin Swab	Infant Skin Swab	Month 4	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T1	BSS-5-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 1	Week 1

Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-W1	BSS-5-W1	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T2	BSS-5-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-W3	BSS-5-W1	Maternal Skin Swab	Infant Skin Swab	Week 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T3	BSS-5-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-M4	BST-5-W1	Maternal Skin Swab	Infant Stool	Month 4	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T1	BST-5-W1	Maternal Skin Swab	Infant Stool	Trimester 1	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-W1	BST-5-W1	Maternal Skin Swab	Infant Stool	Week 1	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T2	BST-5-W1	Maternal Skin Swab	Infant Stool	Trimester 2	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-W3	BST-5-W1	Maternal Skin Swab	Infant Stool	Week 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T3	BST-5-W1	Maternal Skin Swab	Infant Stool	Trimester 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	6	Not Shared	MSS-6-T2	BST-6-W3	Maternal Skin Swab	Infant Stool	Trimester 2	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	6	Not Shared	MSS-6-B	BST-6-W3	Maternal Skin Swab	Infant Stool	Birth	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	6	Shared	MSS-6-W3	BST-6-W3	Maternal Skin Swab	Infant Stool	Week 3	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-M4	BST-5-W3	Maternal Skin Swab	Infant Stool	Month 4	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T1	BST-5-W3	Maternal Skin Swab	Infant Stool	Trimester 1	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-W1	BST-5-W3	Maternal Skin Swab	Infant Stool	Week 1	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T2	BST-5-W3	Maternal Skin Swab	Infant Stool	Trimester 2	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-W3	BST-5-W3	Maternal Skin Swab	Infant Stool	Week 3	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	5	Not Shared	MSS-5-T3	BST-5-W3	Maternal Skin Swab	Infant Stool	Trimester 3	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	13	Shared	MSS-13-W1	BSS-13-W1	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	7	Not Shared	MSS-7-W3	BSS-7-W1	Maternal Skin Swab	Infant Skin Swab	Week 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	7	Not Shared	MSS-7-W3	BST-7-W1	Maternal Skin Swab	Infant Stool	Week 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	7	Not Shared	MSS-7-W3	BST-7-W3	Maternal Skin Swab	Infant Stool	Week 3	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	14	Not Shared	MSS-14-W3	BST-14-W1	Maternal Skin Swab	Infant Stool	Week 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	1	Not Shared	MSS-1-M4	BST-1-W1	Maternal Skin Swab	Infant Stool	Month 4	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	1	Shared	MSS-1-W3	BST-1-W1	Maternal Skin Swab	Infant Stool	Week 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	1	Shared	MSS-1-W1	BST-1-W1	Maternal Skin Swab	Infant Stool	Week 1	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	1	Not Shared	MSS-1-M4	BST-1-W3	Maternal Skin Swab	Infant Stool	Month 4	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	1	Shared	MSS-1-W3	BST-1-W3	Maternal Skin Swab	Infant Stool	Week 3	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	1	Shared	MSS-1-W1	BST-1-W3	Maternal Skin Swab	Infant Stool	Week 1	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	1	Not Shared	MSS-1-M4	BSS-1-W3	Maternal Skin Swab	Infant Skin Swab	Month 4	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	1	Shared	MSS-1-W3	BSS-1-W3	Maternal Skin Swab	Infant Skin Swab	Week 3	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	1	Shared	MSS-1-W1	BSS-1-W3	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 3
Staphylococcus hominis	*0.03*	t__SGB7858	1	Not Shared	MSS-1-M4	BSS-1-W1	Maternal Skin Swab	Infant Skin Swab	Month 4	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	1	Shared	MSS-1-W3	BSS-1-W1	Maternal Skin Swab	Infant Skin Swab	Week 3	Week 1
Staphylococcus hominis	*0.03*	t__SGB7858	1	Shared	MSS-1-W1	BSS-1-W1	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 1
Staphylococcus haemolyticus	*0.03*	t__SGB7861	6	Shared	MSS-6-B	BST-6-W1	Maternal Skin Swab	Infant Stool	Birth	Week 1
Staphylococcus haemolyticus	*0.03*	t__SGB7861	6	Not Shared	MVS-6-B	BST-6-W1	Maternal Vaginal Swab	Infant Stool	Birth	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	6	Not Shared	MVS-6-B	BSS-6-W3	Maternal Vaginal Swab	Infant Skin Swab	Birth	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	6	Not Shared	MSS-6-T2	BSS-6-W3	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-B	BST-1-W1	Maternal Skin Swab	Infant Stool	Birth	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-T2	BST-1-W1	Maternal Skin Swab	Infant Stool	Trimester 2	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-M4	BST-1-W1	Maternal Skin Swab	Infant Stool	Month 4	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-W1	BST-1-W1	Maternal Skin Swab	Infant Stool	Week 1	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-W3	BST-1-W1	Maternal Skin Swab	Infant Stool	Week 3	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-W1	BST-5-W1	Maternal Skin Swab	Infant Stool	Week 1	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-M4	BST-5-W1	Maternal Skin Swab	Infant Stool	Month 4	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-T3	BST-5-W1	Maternal Skin Swab	Infant Stool	Trimester 3	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	14	Not Shared	MVS-14-W1	BST-14-W1	Maternal Vaginal Swab	Infant Stool	Week 1	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MST-4-B	BSS-4-M4	Maternal Stool	Infant Skin Swab	Birth	Month 4

Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MSS-4-M4	BSS-4-M4	Maternal Skin Swab	Infant Skin Swab	Month 4	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MST-4-B	BSS-4-W3	Maternal Stool	Infant Skin Swab	Birth	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MSS-4-M4	BSS-4-W3	Maternal Skin Swab	Infant Skin Swab	Month 4	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MST-4-B	BSS-4-W1	Maternal Stool	Infant Skin Swab	Birth	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MSS-4-M4	BSS-4-W1	Maternal Skin Swab	Infant Skin Swab	Month 4	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	6	Not Shared	MVS-6-B	BST-6-W1	Maternal Vaginal Swab	Infant Stool	Birth	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	6	Not Shared	MSS-6-T2	BST-6-W1	Maternal Skin Swab	Infant Stool	Trimester 2	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MST-4-B	BST-4-W3	Maternal Stool	Infant Stool	Birth	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MSS-4-M4	BST-4-W3	Maternal Skin Swab	Infant Stool	Month 4	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MST-4-B	BST-4-W1	Maternal Stool	Infant Stool	Birth	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	4	Not Shared	MSS-4-M4	BST-4-W1	Maternal Skin Swab	Infant Stool	Month 4	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-B	BSS-1-M8	Maternal Skin Swab	Infant Skin Swab	Birth	Month 8
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-T2	BSS-1-M8	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Month 8
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-M4	BSS-1-M8	Maternal Skin Swab	Infant Skin Swab	Month 4	Month 8
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-W1	BSS-1-M8	Maternal Skin Swab	Infant Skin Swab	Week 1	Month 8
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-W3	BSS-1-M8	Maternal Skin Swab	Infant Skin Swab	Week 3	Month 8
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-B	BSS-1-M4	Maternal Skin Swab	Infant Skin Swab	Birth	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-T2	BSS-1-M4	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-M4	BSS-1-M4	Maternal Skin Swab	Infant Skin Swab	Month 4	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-W1	BSS-1-M4	Maternal Skin Swab	Infant Skin Swab	Week 1	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-W3	BSS-1-M4	Maternal Skin Swab	Infant Skin Swab	Week 3	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	6	Not Shared	MVS-6-B	BST-6-W3	Maternal Vaginal Swab	Infant Stool	Birth	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	6	Not Shared	MSS-6-T2	BST-6-W3	Maternal Skin Swab	Infant Stool	Trimester 2	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T3	BSS-3-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-M4	BSS-3-W1	Maternal Skin Swab	Infant Skin Swab	Month 4	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T2	BSS-3-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T1	BSS-3-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 1	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-W1	BSS-3-W1	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-W3	BSS-3-W1	Maternal Skin Swab	Infant Skin Swab	Week 3	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T3	BST-3-W1	Maternal Skin Swab	Infant Stool	Trimester 3	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-M4	BST-3-W1	Maternal Skin Swab	Infant Stool	Month 4	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T2	BST-3-W1	Maternal Skin Swab	Infant Stool	Trimester 2	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T1	BST-3-W1	Maternal Skin Swab	Infant Stool	Trimester 1	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-W1	BST-3-W1	Maternal Skin Swab	Infant Stool	Week 1	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-W3	BST-3-W1	Maternal Skin Swab	Infant Stool	Week 3	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T3	BST-3-B	Maternal Skin Swab	Infant Stool	Trimester 3	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-M4	BST-3-B	Maternal Skin Swab	Infant Stool	Month 4	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T2	BST-3-B	Maternal Skin Swab	Infant Stool	Trimester 2	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T1	BST-3-B	Maternal Skin Swab	Infant Stool	Trimester 1	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-W1	BST-3-B	Maternal Skin Swab	Infant Stool	Week 1	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-W3	BST-3-B	Maternal Skin Swab	Infant Stool	Week 3	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	13	Not Shared	MSS-13-W3	BSS-13-W3	Maternal Skin Swab	Infant Skin Swab	Week 3	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	13	Not Shared	MSS-13-B	BSS-13-W3	Maternal Skin Swab	Infant Skin Swab	Birth	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	13	Not Shared	MSS-13-W1	BSS-13-W3	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-W1	BST-5-B	Maternal Skin Swab	Infant Stool	Week 1	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-M4	BST-5-B	Maternal Skin Swab	Infant Stool	Month 4	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-T3	BST-5-B	Maternal Skin Swab	Infant Stool	Trimester 3	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	11	Not Shared	MSS-11-T3	BSS-11-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	11	Not Shared	MSS-11-T2	BSS-11-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	11	Shared	MST-11-W1	BSS-11-W1	Maternal Stool	Infant Skin Swab	Week 1	Week 1

Staphylococcus epidermidis	*0.03*	t__SGB7865	11	Not Shared	MSS-11-T3	BST-11-W3	Maternal Skin Swab	Infant Stool	Trimester 3	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	11	Not Shared	MSS-11-T2	BST-11-W3	Maternal Skin Swab	Infant Stool	Trimester 2	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	11	Shared	MST-11-W1	BST-11-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	11	Not Shared	MSS-11-T3	BST-11-B	Maternal Skin Swab	Infant Stool	Trimester 3	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	11	Not Shared	MSS-11-T2	BST-11-B	Maternal Skin Swab	Infant Stool	Trimester 2	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	11	Not Shared	MST-11-W1	BST-11-B	Maternal Stool	Infant Stool	Week 1	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-W1	BSS-5-W1	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-M4	BSS-5-W1	Maternal Skin Swab	Infant Skin Swab	Month 4	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-T3	BSS-5-W1	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Week 1
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-W1	BSS-5-M4	Maternal Skin Swab	Infant Skin Swab	Week 1	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-M4	BSS-5-M4	Maternal Skin Swab	Infant Skin Swab	Month 4	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-T3	BSS-5-M4	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-B	BST-1-W3	Maternal Skin Swab	Infant Stool	Birth	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-T2	BST-1-W3	Maternal Skin Swab	Infant Stool	Trimester 2	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-M4	BST-1-W3	Maternal Skin Swab	Infant Stool	Month 4	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-W1	BST-1-W3	Maternal Skin Swab	Infant Stool	Week 1	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	1	Not Shared	MSS-1-W3	BST-1-W3	Maternal Skin Swab	Infant Stool	Week 3	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T3	BSS-3-M4	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-M4	BSS-3-M4	Maternal Skin Swab	Infant Skin Swab	Month 4	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T2	BSS-3-M4	Maternal Skin Swab	Infant Skin Swab	Trimester 2	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-T1	BSS-3-M4	Maternal Skin Swab	Infant Skin Swab	Trimester 1	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-W1	BSS-3-M4	Maternal Skin Swab	Infant Skin Swab	Week 1	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	3	Not Shared	MSS-3-W3	BSS-3-M4	Maternal Skin Swab	Infant Skin Swab	Week 3	Month 4
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-W1	BSS-5-W3	Maternal Skin Swab	Infant Skin Swab	Week 1	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-M4	BSS-5-W3	Maternal Skin Swab	Infant Skin Swab	Month 4	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-T3	BSS-5-W3	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Week 3
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-W1	BSS-5-B	Maternal Skin Swab	Infant Skin Swab	Week 1	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-M4	BSS-5-B	Maternal Skin Swab	Infant Skin Swab	Month 4	Birth
Staphylococcus epidermidis	*0.03*	t__SGB7865	5	Not Shared	MSS-5-T3	BSS-5-B	Maternal Skin Swab	Infant Skin Swab	Trimester 3	Birth
Streptococcus thermophilus	*0.05*	t__SGB8002	9	Not Shared	MST-9-W1	BST-9-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Streptococcus thermophilus	*0.05*	t__SGB8002	9	Not Shared	MST-9-T3	BST-9-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Streptococcus thermophilus	*0.05*	t__SGB8002	13	Not Shared	MST-13-W1	BST-13-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Streptococcus thermophilus	*0.05*	t__SGB8002	13	Not Shared	MST-13-W3	BST-13-M12	Maternal Stool	Infant Stool	Week 3	Month 12
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-W1	BST-6-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-B	BST-6-M4	Maternal Stool	Infant Stool	Birth	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-W3	BST-6-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-M4	BST-6-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-T3	BST-6-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-T2	BST-6-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	1	Not Shared	MST-1-T3	BST-1-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	1	Not Shared	MST-1-W3	BST-1-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	1	Not Shared	MST-1-T2	BST-1-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	14	Not Shared	MST-14-T2	BST-14-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	14	Not Shared	MST-14-T2	BST-14-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	14	Not Shared	MST-14-T2	BST-14-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	14	Not Shared	MST-14-T2	BST-14-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Streptococcus salivarius	*0.05*	t__SGB8007_group	8	Not Shared	MST-8-T2	BST-8-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	8	Not Shared	MST-8-T3	BST-8-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	8	Not Shared	MST-8-W1	BST-8-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-B	BST-5-M12	Maternal Stool	Infant Stool	Birth	Month 12

Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-W1	BST-5-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-T3	BST-5-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-T2	BST-5-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-B	BST-5-W1	Maternal Stool	Infant Stool	Birth	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-W1	BST-5-W1	Maternal Stool	Infant Stool	Week 1	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-T3	BST-5-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-T2	BST-5-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-B	BST-5-W3	Maternal Stool	Infant Stool	Birth	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-W1	BST-5-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-T3	BST-5-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-T2	BST-5-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-B	BST-5-M8	Maternal Stool	Infant Stool	Birth	Month 8
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-W1	BST-5-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-T3	BST-5-M8	Maternal Stool	Infant Stool	Trimester 3	Month 8
Streptococcus salivarius	*0.05*	t__SGB8007_group	5	Shared	MST-5-T2	BST-5-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Streptococcus salivarius	*0.05*	t__SGB8007_group	12	Shared	MST-12-W1	BST-12-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	12	Not Shared	MST-12-B	BST-12-M4	Maternal Stool	Infant Stool	Birth	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	12	Shared	MST-12-T2	BST-12-M4	Maternal Stool	Infant Stool	Trimester 2	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	12	Shared	MST-12-T3	BST-12-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Shared	MSS-11-M4	BST-11-W3	Maternal Skin Swab	Infant Stool	Month 4	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Not Shared	MST-11-W3	BST-11-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Not Shared	MST-11-M4	BST-11-W3	Maternal Stool	Infant Stool	Month 4	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Not Shared	MST-11-T3	BST-11-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Not Shared	MST-11-B	BST-11-W3	Maternal Stool	Infant Stool	Birth	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Shared	MSS-11-M4	BST-11-M4	Maternal Skin Swab	Infant Stool	Month 4	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Not Shared	MST-11-W3	BST-11-M4	Maternal Stool	Infant Stool	Week 3	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Not Shared	MST-11-M4	BST-11-M4	Maternal Stool	Infant Stool	Month 4	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Not Shared	MST-11-T3	BST-11-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Not Shared	MST-11-B	BST-11-M4	Maternal Stool	Infant Stool	Birth	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	4	Not Shared	MST-4-B	BST-4-W1	Maternal Stool	Infant Stool	Birth	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	4	Not Shared	MST-4-W3	BST-4-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	4	Not Shared	MST-4-T2	BST-4-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	4	Not Shared	MST-4-B	BST-4-W3	Maternal Stool	Infant Stool	Birth	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	4	Not Shared	MST-4-W3	BST-4-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	4	Not Shared	MST-4-T2	BST-4-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	9	Shared	MSS-9-M4	BST-9-W3	Maternal Skin Swab	Infant Stool	Month 4	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	9	Shared	MSS-9-M4	BSS-9-M4	Maternal Skin Swab	Infant Skin Swab	Month 4	Month 4
Streptococcus salivarius	*0.05*	t__SGB8007_group	9	Shared	MSS-9-M4	BST-9-W1	Maternal Skin Swab	Infant Stool	Month 4	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	12	Not Shared	MST-12-W1	BST-12-M12	Maternal Stool	Infant Stool	Week 1	Month 12
Streptococcus salivarius	*0.05*	t__SGB8007_group	12	Not Shared	MST-12-B	BST-12-M12	Maternal Stool	Infant Stool	Birth	Month 12
Streptococcus salivarius	*0.05*	t__SGB8007_group	12	Not Shared	MST-12-T2	BST-12-M12	Maternal Stool	Infant Stool	Trimester 2	Month 12
Streptococcus salivarius	*0.05*	t__SGB8007_group	12	Not Shared	MST-12-T3	BST-12-M12	Maternal Stool	Infant Stool	Trimester 3	Month 12
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Shared	MST-6-W1	BST-6-W1	Maternal Stool	Infant Stool	Week 1	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Shared	MST-6-B	BST-6-W1	Maternal Stool	Infant Stool	Birth	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Shared	MST-6-W3	BST-6-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Shared	MST-6-M4	BST-6-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-T3	BST-6-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-T2	BST-6-W1	Maternal Stool	Infant Stool	Trimester 2	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Shared	MST-6-W1	BST-6-W3	Maternal Stool	Infant Stool	Week 1	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Shared	MST-6-B	BST-6-W3	Maternal Stool	Infant Stool	Birth	Week 3

Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Shared	MST-6-W3	BST-6-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Shared	MST-6-M4	BST-6-W3	Maternal Stool	Infant Stool	Month 4	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-T3	BST-6-W3	Maternal Stool	Infant Stool	Trimester 3	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	6	Not Shared	MST-6-T2	BST-6-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Not Shared	MSS-11-M4	BST-11-W1	Maternal Skin Swab	Infant Stool	Month 4	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Shared	MST-11-W3	BST-11-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Shared	MST-11-M4	BST-11-W1	Maternal Stool	Infant Stool	Month 4	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Shared	MST-11-T3	BST-11-W1	Maternal Stool	Infant Stool	Trimester 3	Week 1
Streptococcus salivarius	*0.05*	t__SGB8007_group	11	Shared	MST-11-B	BST-11-W1	Maternal Stool	Infant Stool	Birth	Week 1
Streptococcus australis	*0.03*	t__SGB8059_group	4	Not Shared	MST-4-B	BST-4-M4	Maternal Stool	Infant Stool	Birth	Month 4
Streptococcus lactarius	*0.03*	t__SGB8065	1	Shared	MSS-1-M4	BST-1-M12	Maternal Skin Swab	Infant Stool	Month 4	Month 12
Streptococcus lactarius	*0.03*	t__SGB8065	1	Shared	MSS-1-M4	BST-1-M4	Maternal Skin Swab	Infant Stool	Month 4	Month 4
Streptococcus lactarius	*0.03*	t__SGB8065	1	Shared	MSS-1-M4	BSS-1-M4	Maternal Skin Swab	Infant Skin Swab	Month 4	Month 4
Streptococcus lactarius	*0.03*	t__SGB8065	1	Shared	MSS-1-M4	BSS-1-M8	Maternal Skin Swab	Infant Skin Swab	Month 4	Month 8
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	5	Shared	MST-5-T2	BST-5-W3	Maternal Stool	Infant Stool	Trimester 2	Week 3
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	5	Shared	MST-5-B	BST-5-W3	Maternal Stool	Infant Stool	Birth	Week 3
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	5	Shared	MST-5-T2	BST-5-M8	Maternal Stool	Infant Stool	Trimester 2	Month 8
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	5	Shared	MST-5-B	BST-5-M8	Maternal Stool	Infant Stool	Birth	Month 8
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	4	Shared	MST-4-W3	BST-4-W3	Maternal Stool	Infant Stool	Week 3	Week 3
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	4	Shared	MST-4-B	BST-4-W3	Maternal Stool	Infant Stool	Birth	Week 3
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	4	Shared	MST-4-W3	BST-4-W1	Maternal Stool	Infant Stool	Week 3	Week 1
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	4	Shared	MST-4-B	BST-4-W1	Maternal Stool	Infant Stool	Birth	Week 1
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	9	Shared	MSS-9-M4	BST-9-B	Maternal Skin Swab	Infant Stool	Month 4	Birth
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	13	Shared	MST-13-B	BST-13-M4	Maternal Stool	Infant Stool	Birth	Month 4
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	13	Not Shared	MST-13-T3	BST-13-M4	Maternal Stool	Infant Stool	Trimester 3	Month 4
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	13	Shared	MST-13-B	BSS-13-M4	Maternal Stool	Infant Skin Swab	Birth	Month 4
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	13	Not Shared	MST-13-T3	BSS-13-M4	Maternal Stool	Infant Skin Swab	Trimester 3	Month 4
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	11	Shared	MST-11-W1	BST-11-W1	Maternal Stool	Infant Stool	Week 1	Week 1
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	11	Shared	MST-11-W1	BST-11-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	6	Not Shared	MST-6-W1	BST-6-M8	Maternal Stool	Infant Stool	Week 1	Month 8
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	6	Not Shared	MST-6-M4	BST-6-M8	Maternal Stool	Infant Stool	Month 4	Month 8
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	6	Not Shared	MST-6-W1	BST-6-M4	Maternal Stool	Infant Stool	Week 1	Month 4
Streptococcus parasanguinis	*0.03607692*	t__SGB8071	6	Not Shared	MST-6-M4	BST-6-M4	Maternal Stool	Infant Stool	Month 4	Month 4

PUBLICATIONS

Serghiou IR, Webber MA, Hall LJ. An update on the current understanding of the infant skin microbiome and research challenges. Current Opinion Microbiology. 2023 Oct; 75:102364. Epub 2023 Aug 14. PMID: 37586254.

Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Microbiology

An update on the current understanding of the infant skin microbiome and research challenges

Iliana R Serghiou[1,2,*], Mark A Webber[1,3,†] and Lindsay J Hall[1,2,3,4,‡]

Multiple factors contribute to establishment of skin microbial communities in early life, with perturbations in these ecosystems impacting health. This review provides an update on methods used to profile the skin microbiome and how this is helping enhance our understanding of infant skin microbial communities, including factors that influence composition and disease risk. We also provide insights into new interventional studies and treatments in this area. However, it is apparent that there are still research bottlenecks that include overreliance on high-income countries for skin microbiome 'surveys', many studies still focus solely on the bacterial microbiota, and also technical issues related to the lower microbial biomass of skin sites. These points link to pertinent open-research questions, such as how the whole infant skin microbiome interacts and how microbial-associated functions shape infant skin health and immunity.

Addresses
[1] Quadram Institute Bioscience, Norwich Research Park, Norwich, Norfolk NR4 7UQ, UK
[2] School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk NR4 7TJ, UK
[3] Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, Norfolk NR4 7TJ, UK
[4] Intestinal Microbiome, School of Life Sciences, ZIEL – Institute for Food & Health, Technical University of Munich, Freising, Germany

Corresponding authors: Serghiou, Iliana R (iliana.serghiou2@quadram.ac.uk), Webber, Mark A (mark.webber@quadram.ac.uk), Hall, Lindsay J (lindsay.hall@quadram.ac.uk)
* ORCID: 0000-0002-0699-5892
† ORCID: 0000-0001-8938-5709
‡ ORCID: 0000-0001-9169-7592

Current Opinion in Microbiology 2023, **75**:102364

This review comes from a themed issue on **Microbiota**

Edited by **Christopher Stewart** and **Maria Carmen Collado**

Available online xxxx

https://doi.org/10.1016/j.mib.2023.102364

1369-5274/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Where the infant skin microbiome field currently stands

The human skin microbiome is organised into microbial communities at different body sites [1]. These communities include bacteria, archaea, viruses and fungi, and participate in the host's skin physiological functions and immunity, and overall skin health [1–3]. The infant skin bacteriome has been studied in various projects in the last 20 years and has been found to typically consist of members of the phyla *Bacillota*, *Actinomycetota*, *Pseudomonadota* and *Bacteroidota* [4]. As well as the bacteriome, the infant mycobiome has been found to contain *Malassezia*, *Candida*, *Cladosporium*, *Alternaria*, *Rhodotorula*, *Fusarium*, *Cryptococcus*, *Exophiala*, *Aspergillus and Nigrospora* [5–8]. Directly after birth, coagulase-negative staphylococci species (CoNS) (*Staphylococcus epidermidis* and *Staphylococcus capitis*) dominate the neonatal skin [9,10] along with *Lactobacillus iners*; not long after, these microbes decrease in abundance, which allows gradual colonisation by *Acinetobacter*, *Streptococcus*, *Veillonella*, *Bifidobacterium* and *Enhydrobacter* species [4]. Preterm neonates have been shown to have a more varied skin microbiome and exhibit more *Staphylococcus*, *Raoultella*, *Klebsiella*, *Serratia*, *Streptococcus*, *Enterococcus* and *Enterobacter* [10]. *Streptococcus* species particularly appear to be more abundant in infancy, and decrease with age [11–13]. Perturbations in these communities, particularly early in life, have been associated with dermal conditions, including staphylococcal-scalded skin syndrome (SSSS), atopic dermatitis (AD), diaper dermatitis (DD) and tinea capitis (TC) [1,4,5,7,14–17].

Most recent research has adopted higher-resolution methods to defining the skin microbiome. This has largely confirmed previous studies in terms of species present (with AD the most-studied skin disease), although there has been more of a focus on detailed interactions between factors that affect the infant skin microbiome, mainly in high-income countries (HICs) [1–3,5,15,18,19]. Delivery mode, antibiotic exposure, breastfeeding and neonatal skin structure and function are factors that have been shown to influence initial skin community establishment [2,4,20]. Longer term, the environment, skincare cosmetics, diet, age and gender also impact the composition of the skin microbiome [2,3]. The skin microbiome community make-up appears to be largely defined by age, and only transiently

influenced by factors such as drug use and cleaning and moisturising practices [4]. Recently, Zioutis et al. [21] found that community assembly was driven primarily by ecological drift, suggesting occurrence of stochastic processes. Other recent popular research topics include skin disease microbiome profiling [4,5,14–17], investigation of understudied microbial communities [4,5,22], combined studies using metagenomics, metabolomics and culturomics [5,23–25] and probiotic interventions [19,26–28].

Given the recent increase in knowledge, the aim of this review was to provide an update on the current infant skin microbiome research. We identified 33 research papers and eight reviews published since 2021, which we focus on in this paper (Table S1).

Factors affecting the infant skin microbiome

There are many factors that affect the infant skin microbiome, and more recent studies have focussed on how some of these influence skin microbiome establishment, function and disease [4–6,9,10,18,20–24,29–32].

Delivery mode influences skin microbial colonisation, which in turn affects skin barrier function [4,5]. A recent investigation tested the impact of vaginal delivery (VD) in water on the infant skin (bacterial) microbiome [4]. The authors identified a higher abundance of *Streptococcus* on the skin of infants born by elective caesarean section (CS) and VD in water, compared with normal VD [4]. Rapin et al. [4] further distinguished between elective and emergency CS, and while they found no significant difference in bacterial skin diversity, they did observe higher *Lactobacillus* abundance in emergency compared with elective CS (likely due to initial labour-associated exposure to the vaginal microbiota) [4]. Dry skin was also observed in all infants born through elective CS, which was associated with higher *Streptococcus* abundance [4]. Mycobiome-focused studies have also shown *Candida albicans* and *Rhodotorula* species in higher abundance on the skin of VD infants, while *Aspergillus* was in higher abundance in CS infants [5,6]. The pregnancy and birthing process (whether vaginally or CS delivery), can encounter complications, which may lead to the use of antibiotics. Antibiotic exposure in infants has been shown to reduce skin microbial diversity, although this appears to be transient followed by recovery, and then increasing diversity seen over time, despite exposure [9,30]. More recently, it appears that antibiotic use in pregnancy leads to colonisation of the infant skin with genitourinary tract-associated pathogens (*Ureaplasma* and *Gardnerella*) [23].

To prevent CS-associated conditions (e.g. AD), interventions have been trialled, such as vaginal seeding [20,29]. This exposes CS-delivered infants to the maternal (vaginal) microbiome that is 'bypassed' during CS delivery, although it does not account for important vertical transfer of maternal gut microbes [20,29]. However, there is concern regarding infection transmission from the mothers' vaginal microbiome to the neonate, for example, transfer of group-B *Streptococcus*, but screening steps have been put in place [20,29]. Currently, there are five vaginal seeding randomised placebo-controlled trials in infants investigating how this process correlates with immunity and health outcomes in infants (Table S2). To date, one clinical trial (ID: NCT03567707) does not appear to have published any data, however, the other four clinical trials (ID: NCT03298334, ID: NCT03809390, ID: ChiCTR2000031326 and ID: ACTRN12618000339257) have conducted associated studies, some showing that vaginal seeding can partially 'restore' microbiota profiles in CS babies and normalise neurodevelopment [20,29,33–35]. However, two associated studies demonstrated that vaginal seeding had no significant impacts on gut microbiota composition, growth, anthropometry or allergy risks during the first few months to two years of life [36,37].

Other maternal factors, such as the mothers' breastmilk and food allergies, have been associated with infant skin microbiome changes [4,18]. Mother food allergies have been associated with higher levels of *Prevotella* and *Alloprevotella* on infant skin [4], and breastfeeding has been associated with transfer of transient microbes (*Prevotellaceae*) to the infant skin, which was limited to the typical period of breastfeeding [4]. These transient microbes could have important implications in community establishment and health, however, further investigation into this is required. There have also been differing opinions as to whether breastmilk has an effect on the infant skin microbiome composition, allergic disease and AD in infants [18,21], however, further studies are needed in this area.

The maternal and infant 'environment' also appears to influence skin community composition. For example, the presence of dogs during pregnancy is associated with higher representations of *Betaproteobacteriales* and *Ralstonia* on the infant skin [4]. Longer-term hospital stays, including in neonatal intensive care units, predispose for carriage of potential pathogens on the skin of preterm infants, and show differences between countries. For example, *Mycobacterium tuberculosis* and *Mycobacteroides abscessus* were found on the skin of neonates in a hospital in India [31], *Streptococcus* was identified on neonates in a hospital in Sweden [4,31], and bacterial colonisation with Gram negatives *Klebsiella pneumoniae* and *Serratia marcescens* appeared to be common in neonates born in South African hospitals [32]. Another study identified similarities between the preterm skin microbiome and the environment of the infant's incubator/

isolator and air swabs, although they concluded that the skin microbiome swabs clustered separately from incubator/isolator and control swabs [10].

Cleansing practices are common in hospitals around the world, and many frequently use chlorhexidine gluconate to cleanse the skin of neonates, which has been shown to be effective for skin antisepsis and reducing bacterial colonisation [22,32]. These antisepsis products can be harsh on the neonatal skin, therefore emollient application may be necessary to improve skin condition, which has previously been shown to increase Gram-positive *Staphylococcus aureus* colonisation [32]. Capone et al. [24] claims to be the first study to evaluate the impact of commercialised baby skincare products on the infant skin microbiome; they showed that routine emollient use in infants increased their skin microbiome richness, ceramide and free fatty acid levels, which plays an important role in skin barrier integrity. In contrast, other studies have shown no associations between emollient use and skin bacterial composition [4]. This demonstrates the need for more in-depth investigations into the effects of emollient use on infant skin composition and health.

Other less-studied factors impacting the infant skin microbiome include genetics and gender. Genetics, specifically filaggrin (gene) functional deficiency, has shown no association with the skin microbiome, despite it being responsible for impaired skin barrier function and a major risk factor for AD [4]. There have also been few studies comparing the impact of gender on the infant skin microbiome, and have shown contrasting results: some identifying no significant effect [6], whereas others have shown a significant difference in the first week of birth, specifically in the groin area, but the reasons were unclear [30].

Infant skin disease microbiome profiling

SSSS, AD, DD and TC are common neonatal skin diseases [1,4,5,7,14–17], however, the majority of recent studies have focused on AD and DD. Skin microbiome profiling of infant AD has shown altered microbiome community composition, with *S. aureus* identified as the main bacterial driver, whilst others have shown the reverse [1,15,16], or no specific bacterial taxa associated with AD [4]. There appears to be a mixed consensus on the main causative agent of AD, which needs to be clarified through further investigation. Furthermore, both lesional and non-lesional skin of AD-afflicted infants have shown a slower temporal development of the skin microbiome than healthy infants [16]. DD has been studied in relation to the bacteriome and mycobiome, with *Candida* and *S. aureus* directly correlating with disease severity [5,17]. TC is caused by pathogenic fungi *Microsporum canis*, which has been correlated with an increase in abundance of *Cutibacterium* and *Corynebacterium*, and reduced *Streptococcus* [7].

Understudied infant skin microbiome communities

The infant skin bacteriome is better studied than the mycobiome, virome and archaeome, allowing bacteriome culture collections to be developed, such as the Skin Microbial Genome Collection (SMGC) and Skin Bacteria Culture Collection (SBCC) [38]. These collections identified a novel skin-abundant bacterial genus *Candidatus pellibacterium* in adults and infants. More recently, research has provided more insight into the infant skin fungal communities [5,39]. *C. albicans* has been shown more abundant in areas with skin-to-skin contact, whereas *Cladosporium* and *Alternaria* were more abundant in less-occluded areas [5]. Saheb Kashaf et al. [38] also recently identified three novel *Malassezia* species (*Malassezia auris*, *Malassezia palmae* and *Malassezia rara*) in adults and infants.

There is a lack of infant skin virome research, and current work is focused on disease outcomes in neonates, such as the herpes simplex virus 2 [39]. Viruses identified in adults and infants have been classified to the order *Caudovirales*, and families *Siphoviridae* and *Myoviridae* [38]. The toe web appears to be a site that is abundant in viruses, and novel jumbo phages have been identified here [38]. Archaeome research and its role in the skin microbiome is completely lacking across the lifespan; *Thaumarchaeota* is currently the main archaeal genus identified on human skin, with no specific genera relating to infant skin [3].

Research on early-life skin microbial community assembly using metagenome-assembled genomes (MAGs) is currently very limited, however, an Early-Life Skin Genomes catalogue has recently been developed for infants aged 2–3 months and 12 months [40]. This catalogue comprises of 9194 bacterial genomes from 1029 species, 206 fungal genomes from 13 species and 39 viral sequences [40]. Top genera identified from bacteria, fungi and viruses were *Streptococcus*, *Corynebacterium*, *Neisseria*, *Bifidobacterium*, *Prevotella*, *Malassezia*, *torque teno virus* and *gammapapillomavirus* [40].

Metagenomic, metabolomic and culturomic integrative studies

In recent years, more integrative approaches are being used to study the infant skin microbiome and its function, using combinations of metabolomics, culturomics and metagenomics [5,23–25,38,41,42]. This has allowed a coexistence of three distinct metabolite–microbe clusters to be identified: the first was dominated by *Cutibacterium* linked to hydrophobic elements of the skin barrier, the second associated the *Staphylococcus* genus

with amino acids relevant to water holding capacity and pH regulation of the skin surface and the third was characterised by *Streptococcus* associated with amino acid and sugar abundance [41]. Furthermore, Saheb Kashaf et al. [38] used an integrative approach of culturomics and metagenomics to identify members and functions of the human skin microbiome, and led to the development of the SMGC and SBCC from infant and adult skin. These integrative studies will have implications for identifying potential strains for probiotic interventions to treat skin disease and maintain skin health.

Probiotic interventions

Probiotic interventions, mostly using *Lactobacillus* and *Bifidobacterium* spp., and *Lacticaseibacillus rhamnosus* GG, have shown promising results in treating AD in infants [26–28]. After probiotic treatment with *L. rhamnosus* GG, higher levels of *Prevotella*, *Veillonella* and *Ralstonia* and lower levels of *Stenotrophomonas* and *Microbacterium* were identified. Certain skin and gut microbiota (*Acinetobacter* spp., *Haemophilus haemolyticus*, *Bacteroides ovatus*, *Schaalia odontolytica* and *Actinomyces graevenitzi*) have also been identified as potential candidates for probiotic treatment of AD and food allergies [19]. There are multiple clinical trials for skin probiotic interventions in infants currently being conducted, or that were completed within the last two years (Table S3); which will assess the potential of probiotics for treatment of skin disease.

Technical research limitations

Although there continues to be an increase in infant skin microbiome research, several technical and study design limitations may be impacting the associations and conclusions drawn. For example, most longitudinal and RCT studies have focused on multiple timepoints closely clustered temporally, rather than taking a longer-term outlook into infancy [1,2,4,5,9,10,14–17,21,23,24,27,28,30,31,41]. Other infant skin microbiome studies have been limited due to either low microbial taxonomic resolution and/or lack of microbial functional information, given the frequent use of targeted gene region sequencing methods [43]. The way skin samples are collected further impacts the resolution and robustness of infant skin microbiome studies; while research has identified that the deeper layers of skin tissue hold stable viable bacterial populations compared with the skin surface [44], the Manos [15] review revealed multiple studies that used skin scrapings (that can access these deeper tissues), resulted in reduced DNA quantity and quality that was not adequate for 16S rRNA gene sequencing, and resulted in elevated host DNA contamination [15]. Furthermore, skin scrapings are invasive to infants causing neonatal skin trauma. Also, although groups have used commercialised DNA extraction methods (some with slight alterations) to study the infant skin microbiome, these are not optimised for skin. The skin is a low-biomass niche [45] and has microbial components, such as the mycobiome, that are particularly difficult to lyse, which affect DNA quantity and quality, thus sequencing outcomes. For example, Teufel et al. [17] identified an issue of internal-transcribed spacer sequencing (used to profile fungal communities) not yielding enough reads for analysis. Furthermore, even with rigorous controls in place, low-biomass niches are also at a high risk of contamination [25,42]. Thus, new methods and approaches optimised for study of the skin microbiome are required to address these technical limitations and improve infant skin microbiome research.

Unmet research questions

There is a pressing need for more longitudinal (following infants from birth into infancy) and higher-resolution infant skin microbiome studies. Shotgun metagenomic sequencing will provide higher resolution over 16S rRNA gene sequencing, allowing species and strain-level profiling as well as functional information to be obtained [43]. However, to date, there are very few skin microbiome studies that utilise shotgun metagenomics [5,9,13,30,46]. A consideration for these types of studies is the ability to extract enough high-quality DNA, but ongoing studies in this area are addressing this need [13]. Profiling using shotgun metagenomics will enable researchers to robustly study the infant skin microbiome, including the mycobiome, virome and archaeome, which are significantly understudied in infants when compared with the bacteriome.

Research on how the whole microbiome (viruses, fungi, bacteria and archaea) interacts and functions as a community, and how this is associated with skin health or disease, is also lacking [5]. Moreover, although there has been a research focus on infant skin microbiome in relation to certain skin diseases [4,5,7,14–18,21], as of yet, no one has been able to distinguish between microbiome-driven and microbiome-independent cases. This is a crucial missing research gap that needs to be addressed for improved treatment interventions [4]. Clinical studies have used culturomic and viromic methods to diagnose skin disease in independent cases and examine certain factors such as the effects of cleansing and moisturising products on skin bacterial colonisation [14,22,32,39]. However, to improve the robustness of these studies, multiple techniques such as culturomics, viromics and metabolomics with metagenomics, should be more widely employed, as currently these are few and far-between [7,25,38,41]. These multi-omics studies would help to identify perturbation-associated microbiota changes and fully understand their functional role in infant skin health and immunity, before and after birth, which is lacking. This would also accelerate development of improved maintenance, preventative and treatment options that will improve standard of care of full-term and preterm neonates, therefore optimising

their long-term health. Interventions include oral and topical probiotic use, development of live biotherapeutics products and improved skin transplantation and wound dressing techniques, some of which are already underway [14,26].

Currently, the vast majority of infant skin microbiome studies have only focused on a few HICs, and skin health maintenance, prevention and treatment interventions would be significantly benefitted by diversifying the cohorts that are sampled to a wider geographical and socio-economic reach. This would also help establish a baseline microbiome composition for a healthy infant skin microbiome, the influence of factors modulating the skin microbiota and how skin conditions are linked to microbial compositional changes. Certain practices, such as kangaroo mother care (KMC) and breastfeeding, are known to improve neonatal health outcomes, and their practice is advised globally by the World Health Organisation's updated 2022 recommendations [47]. Shah, Govindarajan, Rangaiah et al. [31] revealed that the presence of *M. tuberculosis* and *M. abscessus* on Indian neonates diminished with KMC, demonstrating its importance in neonatal health outcomes. Other practices worldwide may also influence seeding of the infant skin microbiota and should be explored further to understand the effects that practices have on birth outcomes and neonatal microbiomes, which could have lasting health implications [48].

Conclusions

Infant skin microbiome research over the last two years has focused on the bacteriome and investigated a more detailed interaction of certain factors that affect it. Infant skin exposed to these factors and afflicted with skin disease (mainly AD) has been assessed and profiled (mainly in HICs), which has led to initial investigations of probiotic skin health interventions. There are clear focused research bottlenecks, which have left open questions on how the whole infant skin microbiome interacts and functions in infant skin health and immunity, before and after birth. These could be addressed by profiling cohorts across the globe, and overcoming technical limitations, which include lack of integrative approaches used, overuse of targeted gene sequencing and unoptimised skin collection and DNA extraction methods used. This would improve neonatal skin health interventions and optimise their long-term health.

CRediT authorship contribution statement

Iliana R Serghiou: Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Project administration. **Lindsay J Hall:** Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Project administration. **Mark A Webber:** Conceptualization, Methodology, Investigation, Writing – original draft, Writing – review & editing, Project administration.

Data availability

No data were used for the research described in the article.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

I.R.S. is funded by a Biotechnology and Biological Sciences Research Council (BBSRC) CTP studentship with Unilever (BB/T508974/1). L.J.H. is supported by Wellcome Trust Investigator Awards 100974/C/13/Z and 220876/Z/20/Z; and a BBSRC Institute Strategic Programme, Gut Microbes and Health BB/R012490/1, and its constituent projects BBS/E/F/000PR10353 and BBS/E/F/000PR10356. M.A.W. is supported by project Grant (BB/T014644/1) from the Biotechnology and Biological Sciences Research Council and BBSRC Institute Strategic Programmes Microbes in the Food Chain BB/R012504/1 and its constituent project BBS/E/F/000PR10349.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.mib.2023.102364.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest

1. Townsend EC, Kalan LR: **The dynamic balance of the skin microbiome across the lifespan**. *Biochem Soc Trans* 2023, **51**:71-86.

2. Dwyer LR, Scharschmidt TC: **Early life host-microbe interactions in skin**. *Cell Host Microbe* 2022, **30**:684-695.

3. Skowron K, *et al.*: **Human skin microbiome: impact of intrinsic and extrinsic factors on skin microbiota**. *Microorganisms* 2021, **9**:543.

4. • Rapin A, *et al.*: **The skin microbiome in the first year of life and its association with atopic dermatitis**. *Allergy* 2023, **7**:1949-1963.
Early-life microbial colonization of the skin may modulate the immune system and impact the development of AD and allergic diseases later in life. The authors show that skin microbiome composition during infancy is defined by age, transiently influenced by delivery mode as well as environmental, parental factors and breastfeeding. They also identify an association between the microbiome, skin barrier integrity and the onset of AD. This is the largest longitudinal study assessing the skin microbiome in the first year of life and the first to investigate the impact of delivery mode while considering VD in water and distinguishing between elective and emergency CS.

5. Nguyen UT, Kalan LR: **Forgotten fungi: the importance of the skin mycobiome**. *Curr Opin Microbiol* 2022, **70**:102235.

6. Wang YR, *et al.*: **Infant mode of delivery shapes the skin mycobiome of prepubescent children**. *Microbiol Spectr* 2022, **10**:e0226722.

7. Tao R, *et al.*: **Altered skin fungal and bacterial community compositions in tinea capitis**. *Mycoses* 2022, **65**:834-840.

8. Shah PA, *et al.*: **Exploring the skin mycobiome in very preterm babies during the early neonatal period in a Neonatal Intensive Care Unit of India**. *Trop Doct* 2022, **52**:362-364.

9. • Hellmann KT, *et al.*: **Viability-resolved metagenomics reveals antagonistic colonization dynamics of Staphylococcus epidermidis strains on preterm infant skin**. *mSphere* 2021, **6**:e0053821.
The skin is a habitat for microbes that commonly infect preterm infants, but the use of sequencing for fine-scale study of the infant skin microbial communities has been limited by technical barriers. Using a photo-reactive dye that eliminates DNA from nonviable microbes, and shotgun metagenomics to sequence the remaining DNA, the authors resolve skin species and strains from preterm skins swabs. CoNS species, *Staphylococcus epidermidis* and *Staphylococcus capitis*, dominated the neonatal skin. Two strains of *Staphylococcus epidermidis* showed an antagonistic relationship on the skin. The authors show the feasibility of gaining detailed information about strain colonization dynamics from this difficult-to-sequence body site of preterm infants, which might be used to guide novel approaches to prevent infections.

10. Schoch JJ, *et al.*: **Skin microbiome sampling in the preterm neonate**. *Pediatr Dermatol* 2023, **40**:129-131.

11. Capone KA, *et al.*: **Diversity of the human skin microbiome early in life**. *J Investig Dermatol* 2011, **131**:2026-2032.

12. Zhu T, *et al.*: **Age and mothers: potent influences of children's skin microbiota**. *J Investig Dermatol* 2019, **139**:2497-2505.e6.

13. • Serghiou IR, *et al.*: **An efficient method for high molecular weight bacterial DNA extraction suitable for shotgun metagenomics from skin swabs**. *Microb Genom* 2023, **9**, https://doi.org/10.1099/mgen.0.001058.
The relatively low bacterial biomass of skin, in comparison to other areas such as the gut microbiome, makes obtaining sufficient DNA for shotgun metagenomic sequencing challenging. Shotgun metagenomic sequencing is needed for taxonomic resolution and functional information in order to study community compositions and functional capabilities of the skin microbiome. The authors optimised a DNA extraction method that produces enough DNA from skin swabs, suitable for shotgun metagenomics, and demonstrated it can be used to effectively characterise the skin microbiota. This method will allow future studies to identify taxonomic and functional changes in the skin microbiota, which is needed to develop interventions to improve and maintain skin health.

14. McKeown KE, Baker RB: **Staphylococcal scalded skin syndrome, identification, and wound care: a case report series**. *Adv Neonatal Care* 2022, **22**:325-332.

15. Manos J: **The human microbiome in disease and pathology**. *APMIS* 2022, **130**:690-705.

16. Olesen CM, *et al.*: **Altered maturation of the skin microbiome in infants with atopic dermatitis**. *Acta Derm Venereol* 2022, **102**:adv00817.

17. Teufel A, *et al.*: **Characterization of the microbiome in the infant diapered area: insights from healthy and damaged skin**. *Exp Dermatol* 2021, **30**:1409-1417.

18. Gołębiewski M, *et al.*: **Mother's milk microbiome shaping fecal and skin microbiota in infants with food allergy and atopic dermatitis: a pilot analysis**. *Nutrients* (10) 2021, **13**:3600.

19. Łoś-Rycharska E, *et al.*: **A combined analysis of gut and skin microbiota in infants with food allergy and atopic dermatitis: a pilot study**. *Nutrients* (5) 2021, **13**:1682.

20. Hourigan SK, Dominguez-Bello MG, Mueller NT: **Can maternal-child microbial seeding interventions improve the health of infants delivered by cesarean section?** *Cell Host Microbe* 2022, **30**:607-611.

21. Zioutis C, *et al.*: **Ecological processes shaping microbiomes of extremely low birthweight infants**. *Front Microbiol* 2022, **13**:812136.

22. Sharma A, *et al.*: **Aqueous chlorhexidine 1% versus 2% for neonatal skin antisepsis: a randomised non-inferiority trial**. *Arch Dis Child — Fetal Neonatal Ed* 2021, **106**:643-648.

23. Orchanian SB, *et al.*: **Multiomic analyses of nascent preterm infant microbiomes differentiation suggest opportunities for targeted intervention**. *Adv Biol* 2022, **6**:e2101313.

24. • Capone KA, *et al.*: **A randomized clinical study on the effects of emollient use on the developing infant skin microbiome and metabolome**. *Exp Dermatol* 2023, **32**:75-77.
There is limited information available on the effects of skin care products on the skin microbiome of healthy individuals. The authors are the first to study the impact of baby skin care products on the infant skin microbiome. The routine use of an appropriately formulated emollient in infants increases their skin microbiome richness, ceramide levels and free fatty acids that play an important role in skin barrier integrity.

25. Mishra A, *et al.*: **Microbial exposure during early human development primes fetal immune cells**. *Cell* 2021, **184**:3394-3409.e20.

26. Andrade PDSMAd, *et al.*: **Efficacy of probiotics in children and adolescents with atopic dermatitis: a randomized, double-blind, placebo-controlled study**. *Front Nutr* 2022, **8**:833666.

27. Pan H, Su J: **Association of probiotics with atopic dermatitis among infant: a meta-analysis of randomized controlled trials**. *Oxid Med Cell Longev* 2022, **2022**:5080190.

28. Carucci L, *et al.*: **Therapeutic effects elicited by the probiotic *Lacticaseibacillus rhamnosus* GG in children with atopic dermatitis. The results of the ProPAD trial**. *Pediatr Allergy Immunol* 2022, **33**:e13836.

29. • Song SJ, *et al.*: **Naturalization of the microbiota developmental trajectory of cesarean-born neonates after vaginal seeding**. *Med* 2021, **2**:951-964.e5.
Cesarean section-born babies show altered microbiota development in relation to babies born vaginally, which can have immunological implications. The authors present the first statistically powered longitudinal study to determine the effect of restoring exposure to maternal vaginal fluids after cesarean section birth. Microbiota trajectories of exposed cesarean section-born babies aligned more closely with that of vaginally born babies. The majority of amplicon sequence variants from maternal vaginal microbiomes on the day of birth were shared with other maternal sites, in contrast to non-pregnant women from the Human Microbiome Projec t study. The results of this observational study prompt urgent randomized clinical trials to test whether microbial restoration reduces the increased disease risk associated with cesarean section birth and the underlying mechanisms. It also provides evidence of the plur-ipotential nature of maternal vaginal fluids to provide pioneer bacterial colonizers for the newborn body sites.

30. Xu Y, *et al.*: **Antibiotic exposure prevents acquisition of beneficial metabolic functions in the preterm infant gut microbiome**. *Microbiome* 2022, **10**:103.

31. • Shah PA, *et al.*: **Mycobacterium species on the cutaneous microbiome of very preterm neonates**. *J Trop Pedia* (2) 2022, **68**:fmac020.
Disruption to the neonatal skin microbiome could lead to pathological disease. There is a need to understand the neonatal skin microbiome of very preterm neonates admitted to the neonatal intensive care unit at a tertiary healthcare setting before and after kangaroo mother care, using next-generation sequencing, which the authors investigated. The authors showed a relative abundance of *Mycobacterium tuberculosis* and *Mycobacteroides abscesses* on the skin microbiome of very preterm neonates before and after kangaroo mother care. The presence of these bacilli as commensals or as potential pathogens is alarming due to the risk of early exposure and incidence of tuberculosis from birth. These authors identify their findings to be the first established in such a setting.

32. Dramowski A, *et al.*: **Impact of 1% chlorhexidine gluconate bathing and emollient application on bacterial pathogen colonization dynamics in hospitalized preterm neonates — a pilot clinical trial**. *EClinicalMedicine* 2021, **37**:100946.

33. Zhou L, *et al.*: **Effects of vaginal microbiota transfer on the neurodevelopment and microbiome of cesarean-born infants: a blinded randomized controlled trial**. *Cell Host Microbe* 2023, **31**:1232-1247.e5.

34. Dominguez-Bello MG, *et al.*: **Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer**. *Nat Med* 2016, **22**:250-253.

35. Mueller NT, *et al.*: **Maternal bacterial engraftment in multiple body sites of cesarean section born neonates after vaginal seeding — a Randomized Controlled Trial**. *mBio* 2023, **14**:e00491-23.

36. Liu Y, *et al.*: **Effects of vaginal seeding on gut microbiota, body mass index, and allergy risks in infants born through cesarean delivery: a randomized clinical trial**. *Am J Obstet Gynecol MFM* 2023, **5**:100793.

37. Wilson BC, *et al.*: **Oral administration of maternal vaginal microbes at birth to restore gut microbiome development in infants born by caesarean section: a pilot randomised placebo-controlled trial**. *eBioMedicine* 2021, **69**:103443.

38. • Saheb Kashaf S, *et al.*: **Integrating cultivation and metagenomics for a multi-kingdom view of skin microbiome diversity and functions**. *Nat Microbiol* 2022, **7**:169-179.
Previous metagenomic investigations into the role of the skin microbiome in health or disease have found that much of the sequenced data do not match reference genomes, making it difficult to interpret metagenomic datasets. The authors combined bacterial cultivation and metagenomic sequencing to assemble the SMGC, which comprises 622 prokaryotic species derived from 7 535 MAGs and 251 isolate genomes. This collection includes 174 newly identified bacterial species and 12 newly identified bacterial genera, including the abundant genus '*Candidatus pellibacterium*', which has been newly associated with the skin. The SMGC increased the characterized set of known skin bacteria by 26% and recovered 12 eukaryotic species, and assembled thousands of viral sequences, including newly identified clades of jumbo phages. This collection provides a comprehensive view of skin microbiome diversity.

39. Hayes CK, *et al.*: **Herpes simplex virus-2 variation contributes to neurovirulence during neonatal infection**. *J Infect Dis* 2022, **226**:1499-1509.

40. Shen Z, *et al.*: **A genome catalog of the early-life human skin microbiome**. *bioRxiv* 2023, https://doi.org/10.1101/2023.05.22.541509

41. • Roux PF, Oddos T, Stamatas G: **Deciphering the role of skin surface microbiome in skin health: an integrative multiomics approach reveals three distinct metabolite–microbe clusters**. *J Investig Dermatol* 2022, **142**:469-479.e5.
There is still much to learn about the role the microbiome plays in skin physiology. The authors employed an integrative methodology for microbiome and metabolome analysis of skin surface samples collected from healthy infants aged 3–6-months. Their integrative analyses delineated the coexistence of three distinct metabolite–microbe clusters.

42. Kennedy KM, *et al.*: **Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies**. *Nature* 2023, **613**:639-649.

43. Sfriso R, *et al.*: **Revealing the secret life of skin — with the microbiome you never walk alone**. *Int J Cosmet Sci* 2020, **42**:116-126.

44. Acosta EM, *et al.*: **Bacterial DNA on the Skin Surface Overrepresents the Viable Skin Microbiome**. eLife Sciences Publications, Ltd; 2023.

45. Bjerre RD, *et al.*: **Effects of sampling strategy and DNA extraction on human skin microbiome investigations**. *Sci Rep* 2019, **9**:17287.

46. Fleming E, *et al.*: **Cultivation of common bacterial species and strains from human skin, oral, and gut microbiota**. *BMC Microbiol* 2021, **21**:278.

47. WHO: **WHO Recommendations for Care of the Preterm or Low Birth Weight Infant**. World Health Organization; 2022.

48. Nsereko E, *et al.*: **Consideration of cultural practices when characterizing the vaginal microbiota among African and African American women**. *Biol Res Nurs* 2021, **23**:91-99.

Serghiou IR, Baker D, Evans R, Dalby MJ, Kiu R, Trampari E, Phillips S, Watt R, Atkinson T, Murphy B, Hall1 LJ, and Webber MA. An efficient method for high molecular weight bacterial DNA extraction suitable for shotgun metagenomics from skin swabs. Microbial Genomics. 2023 July; 9(7).

MICROBIAL GENOMICS

METHOD

Serghiou *et al.*, *Microbial Genomics* 2023;9:001058
DOI 10.1099/mgen.0.001058

MICROBIOLOGY SOCIETY
OPEN DATA OPEN ACCESS

An efficient method for high molecular weight bacterial DNA extraction suitable for shotgun metagenomics from skin swabs

Iliana R. Serghiou[1,2,*], Dave Baker[1], Rhiannon Evans[1], Matthew J. Dalby[1], Raymond Kiu[1], Eleftheria Trampari[1], Sarah Phillips[1], Rachel Watt[1], Thomas Atkinson[1], Barry Murphy[3], Lindsay J. Hall[1,2,4,*] and Mark A. Webber[1,4,*]

Abstract

The human skin microbiome represents a variety of complex microbial ecosystems that play a key role in host health. Molecular methods to study these communities have been developed but have been largely limited to low-throughput quantification and short amplicon-based sequencing, providing limited functional information about the communities present. Shotgun metagenomic sequencing has emerged as a preferred method for microbiome studies as it provides more comprehensive information about the species/strains present in a niche and the genes they encode. However, the relatively low bacterial biomass of skin, in comparison to other areas such as the gut microbiome, makes obtaining sufficient DNA for shotgun metagenomic sequencing challenging. Here we describe an optimised high-throughput method for extraction of high molecular weight DNA suitable for shotgun metagenomic sequencing. We validated the performance of the extraction method, and analysis pipeline on skin swabs collected from both adults and babies. The pipeline effectively characterised the bacterial skin microbiota with a cost and throughput suitable for larger longitudinal sets of samples. Application of this method will allow greater insights into community compositions and functional capabilities of the skin microbiome.

DATA SUMMARY

All sequence data and codes can be accessed at:

NCBI Bio Project ID: PRJNA937622

DOI: https://github.com/quadram-institute-bioscience/coronahit_guppy

DOI: https://github.com/ilianaserghiou/Serghiou-et-al.-2023-Codes

INTRODUCTION

The skin microbiome is a complex ecosystem organised into distinct microbial communities present at different body sites [1, 2]. These microbial ecosystems participate in the host's skin physiological functions and immunity [3, 4]. Perturbations in these communities can negatively impact skin health, particularly early in life [5]. Studying how the skin microbiota forms and changes over time is therefore important to understand how interventions that alter the microbiota affect skin health.

Previous skin microbiome studies have commonly used traditional 16S rRNA gene amplicon sequencing (metataxonomics) to taxonomically classify these complex communities [6]. Whilst relatively cheap, 16S rRNA gene amplicon sequencing provides

Received 27 February 2023; Accepted 04 June 2023; Published 10 July 2023
Author affiliations: [1]Quadram Institute Bioscience, Norwich Research Park, Norwich, Norfolk, NR4 7UQ, UK; [2]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, Norfolk, NR4 7TJ, UK; [3]Unilever R&D Port Sunlight, Bebington, CH63 3JW, UK; [4]Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, Norfolk, NR4 7TJ, UK.
***Correspondence:** Iliana R. Serghiou, iliana.serghiou2@quadram.ac.uk; Lindsay J. Hall, lindsay.hall@quadram.ac.uk; Mark A. Webber, mark.webber@quadram.ac.uk
Keywords: Long and short read sequencing; Microbial abundance; Skin microbiome.
Abbreviations: BR, broad range; CF, consent forms; HMW, high molecular weight; HS, high sensitivity; LB, lysogeny broth; LMW, low molecular weight; MAGs, metagenome assembled genomes; ONT, Oxford Nanopore Technology; PBS, phosphate buffered saline; PEARL, pregnancy and early life study; PIS, participant information sheets; QIB, Quadram Institute Bioscience; SMS, shotgun metagenomic sequencing.
Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files.Four supplementary files are available with the online version of this article.
001058 © 2023 The Authors

This is an open-access article distributed under the terms of the Creative Commons Attribution License. This article was made open access via a Publish and Read agreement between the Microbiology Society and the corresponding author's institution.

Impact Statement

Determining the functional capabilities of microbial communities within different human microbiomes is important, to understand their impacts on health. Extraction of sufficient DNA is challenging, especially from low biomass samples, such as skin swabs, suitable for shotgun metagenomics, which is needed for taxonomic resolution and functional information. Here we describe an optimised DNA extraction method that produces enough DNA from skin swabs, suitable for shotgun metagenomics, and demonstrate it can be used to effectively characterise the skin microbiota. This method will allow future studies to identify taxonomic and functional changes in the skin microbiota, which is needed to develop interventions to improve and maintain skin health.

limited taxonomic information on bacteria and archaea, and some variations in this method can identify species level, but cannot provide information about strain variations or functional capacities. Alternatively, the use of Shotgun Metagenomic Sequencing (SMS) for taxonomic classification follows sequencing of all genetic material and is not limited to targeted regions [7–9]. This reduces bias from selective amplification efficiency and can provide taxonomic information at species/strain level as well as being able to provide information about functional capacities present in the microbiome and individual species [6, 9, 10]. SMS can be performed using multiple technologies, including the Illumina, Oxford Nanopore (ONT) and PacBio Single Molecule Real-Time (SMRT) platforms [11, 12]. In contrast to the Illumina technology, the ONT and PacBio SMRT technologies produce long sequence reads. Data produced with these platforms will usually reconstruct more complete genomes than from short reads and facilitates the generation of high-quality Metagenome Assembled Genomes (MAGs) [12], which can be used for higher taxonomic resolution and functional information [10, 13].

The relatively low bacterial biomass of skin complicates the extraction of sufficient DNA quantities for SMS [14, 15]. This is particularly true for longer read technologies where more input material is needed [16]. There are a limited number of commercialised kit protocols available that can produce high molecular weight (HMW) DNA from skin in sufficient quantities for SMS, although none have been specifically optimised within a protocol that includes sample collection, lysis, DNA extraction, then informatic analysis from skin microbiome samples. To address this need we describe here an optimised high-throughput automated DNA extraction method, for recovery of HMW microbial DNA from skin swabs. This was validated using skin swabs from adult volunteers and babies enrolled in the Pregnancy and Early Life (PEARL) study [17]. The method results in DNA with yield and molecular weight suitable for SMS.

METHODS

DNA extraction method development

To optimise extraction of microbial DNA from skin swabs, different diluents and lysis procedures were evaluated for effectiveness before DNA was extracted using a Promega Maxwell RSC 48 Instrument and RSC Blood DNA Kit (see File S1, available in the online version of this article for protocol). This instrument and kit were chosen as they produce HMW DNA [18, 19], with a higher binding capacity and cleaner eluate than traditional silica-based DNA purification systems [20–23]. The platform is also suitable for high-throughput automated genomic DNA isolation capable of processing from 48 samples in 40 min [22] making this system compatible with larger sample sets. The focus of this work was to develop an efficient, optimised, and validated lysis method for recovering DNA from skin microbiota. Although we developed this using the Maxwell platform, the method should be applicable to any other automated and column-based kits on the market, our goal was not however to benchmark extraction kits, which was beyond the scope of producing this protocol.

To obtain enough DNA from skin swabs, suitable for SMS, we optimised the RSC protocol by testing different variables including the initial diluent and various lysis procedures. After dilution and lysis, samples were heated, following the RSC Blood DNA Kit protocol, and loaded to the Maxwell instrument for the automated extraction (Fig. 1).

Testing initial diluents: measuring extracted bacterial DNA quantity and cellular viability

To allow for a protocol where a swab could be processed allowing both DNA extraction and, in parallel, culture of organisms, it was desirable to remove material from the swab into a diluent. To determine if diluents impacted bacterial viability and ability to extract DNA, 1 x Phosphate Buffered Saline (PBS) and Ultrapure Milli-Q water, for collecting skin bacteria, were compared by measuring extracted bacterial DNA quantity recovered from swabs inoculated with bacteria. Fourty-four sterile charcoal cotton swabs (M40-A2, Technical Service Consultants Ltd.) were used to collect a single colony from an agar plate inoculated with *Staphylococcus aureus* NCTC 8532 to act as a target for DNA extraction. These 'spiked' swab heads were snapped into 1.5 ml Eppendorf tubes containing 1 ml of either 1 x PBS or Ultrapure Milli-Q water. These were then extracted following the Promega Maxwell RSC 48 Instrument and RSC Blood DNA Kit protocol in File S1, with the following modification: the swabs were vortexed

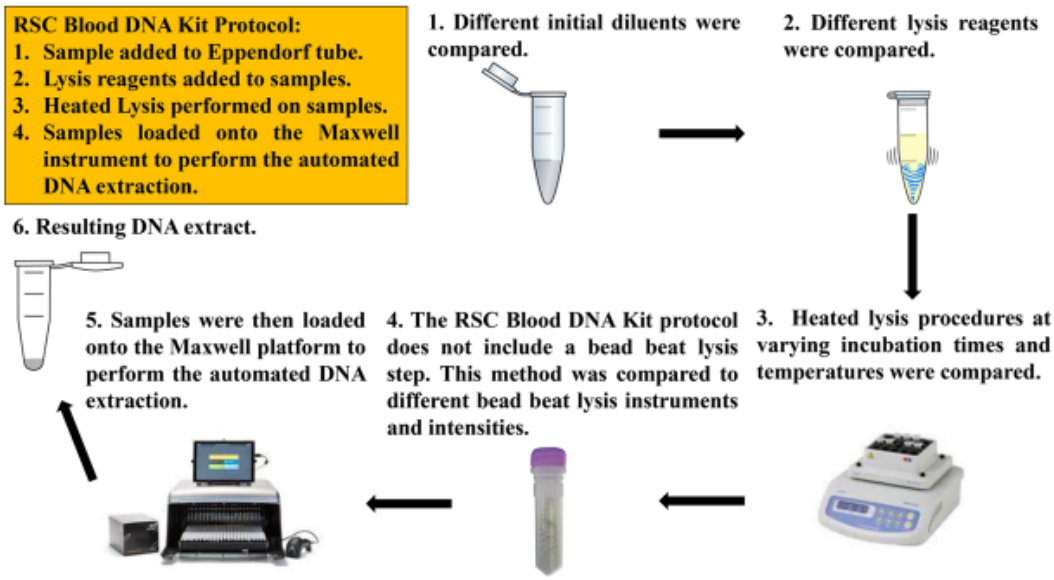


Fig. 1. The RSC Blood DNA Kit protocol (yellow box) and alterations to test different initial diluents and lysis procedures.

at full speed for 2 min and then centrifuged at 14000 ***g*** for 15 min to pellet the cells before the supernatant was removed, and cells were resuspended in 300 µl of 1 x PBS or Ultrapure Milli-Q water. Steps 4 and 6–8 of the RSC protocol were then followed. A bead beating step was then performed using the MP Biomedicals FastPrep instrument for 3 min at speed setting 6.0. The samples were centrifuged again at 14000 ***g*** for 15 min to pellet the cells before sample supernatants were loaded onto the Maxwell instrument and the extraction started following steps 9–21 of the RSC protocol.

The effectiveness of 1 x PBS and Ultrapure Milli-Q water, as initial diluents for collecting skin bacteria, were further compared by measuring bacterial cell viability through the recovery of bacteria from liquid cultures. Cell viability is an important factor as we wanted an initial dilution step, which maintained bacterial viability and was therefore compatible with both culture of bacteria from samples and efficient DNA extraction. Overnight liquid cultures (10 ml) were grown from isolates of three species (*S. aureus* NCTC 8532, *Pseudomonas aeruginosa* PA14 and *Escherichia coli* EC18PR-0166–1, a food isolate of ST10), with three replicates for each. For each replicate, 1 ml was transferred into a 15 ml falcon tube and pelleted by centrifugation at 14000 ***g*** for 15 min. Samples were then resuspended in 200 µl of LB, 1 x PBS or Ultrapure Milli-Q water and left for 1 h at ambient temperature. Serial dilutions of the resuspended samples were made and plated onto drug-free agar and incubated, which were then used to count viable numbers of cells in each sample. A total of nine independent samples were tested for each species in each diluent.

Testing lysis methods: six extraction method procedures

Six lysis methods were compared to identify the best method for high yields of high molecular weight DNA from both Gram-negative and Gram-positive bacteria. Each method varied factors from common lysis methods used in commercial kits for research – heat, chemical, enzymatic, and mechanical [24, 25]. Table 1 lists the differences between the six methods. Methods were tested using both overnight liquid cultures and sterile swab heads inoculated with harvested bacteria from overnight plate cultures.

Duplicate 10 ml overnight liquid cultures were grown for each species (*S. aureus*, *P. aeruginosa* and *E. coli*), from each, 300 µl was added into two 1.5 ml Eppendorf tubes resulting in six tubes, which were tested for method 1 and 2. A further 400 µl of each liquid culture was added into four tubes resulting in six tubes tested for each remaining method. All samples were then extracted following the Promega Maxwell RSC 48 Instrument and RSC Blood DNA Kit protocol (detailed in File S1) with changes to the lysis procedure for each of the six methods tested. All Eppendorf tubes were then vortexed at full speed for 2 min and centrifuged at 14000 ***g*** for 15 min to pellet the cells; the supernatants were removed, and pellets resuspended in 300 µl (methods 1 or 2) or 400 µl (methods 3–6) of 1 x PBS.

For method 1 and 2 samples, 30 µl of Proteinase K and 300 µl of Lysis Buffer (both reagents included in the RSC Blood DNA Kit) were added to the 300 µl sample suspensions. These were then incubated in a heating block at 56 °C for 20 min. For methods 3 and 4 samples, 3 µl of Epicentre Ready-Lyse lysozyme (diluted to 250 U μl^{-1} in Tris-EDTA Buffer [Sigma-Aldrich]) was added to

Table 1. Comparison of extraction methods

	Method 1	*Method 2*	*Method 3*	*Method 4*	*Method 5*	*Method 6*
Heated lysis Step	Yes	Yes	Yes	Yes	Yes	Yes
Time	20 mins	20 mins	18 h	18 h	18 h	18 h
Temperature	56 °C	56 °C	37 °C	37 °C	37 °C	37 °C
Reagents	Proteinase K, lysis buffer	Proteinase K, lysis buffer	Epicentre ready-lyse lysozyme	Epicentre ready-lyse lysozyme	Thermo Fischer lysozyme	Thermo Fischer lysozyme
Agitation	No	No	300 r.p.m.	300 r.p.m.	300 r.p.m.	300 r.p.m.
Bead beat step	Yes	Yes	Yes	Yes	Yes	Yes
Instrument	FastPrep	Tissue Lyser	FastPrep	Tissue Lyser	FastPrep	Tissue Lyser
Settings	3 mins at 6.0 FastPrep	3 mins at 20 Hz	3 mins at 6.0 FastPrep	3 mins at 20 Hz	3 mins at 6.0 FastPrep	3 mins at 20 Hz
Heated offboard lysis step	No	No	Yes	Yes	Yes	Yes
Temperature	N/A	N/A	68 °C	68 °C	68 °C	68 °C
Time	N/A	N/A	15 mins	15 mins	15 mins	15 mins
Reagents	N/A	N/A	Proteinase K, buffer ATL, carrier RNA, buffer ACL	Proteinase K, buffer ATL, carrier RNA, buffer ACL	Proteinase K, buffer ATL, carrier RNA, buffer ACL	Proteinase K, buffer ATL, carrier RNA, buffer ACL
Agitation	N/A	N/A	300 r.p.m.	300 r.p.m.	300 r.p.m.	300 r.p.m.

the 400 µl sample suspensions. For methods 5 and 6 samples, 3 µl of Thermo Fischer lysozyme (diluted to 250 U μl^{-1} in Tris-EDTA Buffer [Sigma-Aldrich]) was added to the 400 µl sample suspensions. Samples from methods 3–6 were then incubated with agitation at 300 r.p.m., 37 °C for 18 h. A bead beating step was performed on all samples. Method 1, 3 and 5 samples used the MP Biomedicals FastPrep instrument for 3 min at speed setting 6.0 and method 2, 4 and 6 samples used the Qiagen Tissue Lyser instrument for 3 min at 20 Hz to compare the impact of a less intense bead beating step. An off-board lysis was performed on method 3–6 samples, which included addition of 40 µl proteinase K, 165 µl Buffer ATL, 120 µl Carrier RNA (lyophilised Carrier RNA was resuscitated with Buffer AVE to make a 1 µg μl^{-1} solution), and 315 µl Buffer ACL (all off-board lysis reagents from Qiagen) into the 400 µl sample suspensions. These samples were then incubated at 68 °C for 15 min. Samples from all methods were centrifuged at 14000 ***g*** for 15 min to pellet cells and the supernatants were loaded onto the Maxwell instrument and the extraction started following steps 9–21 of the initial RSC protocol.

After evaluation of the performance of the different methods from cultured cells, method six performed the best (see Results) and was chosen for validation using swab samples. For validation, sterile charcoal cotton swabs (M40-A2, Technical Service Consultants Ltd.) were spiked with one colony from overnight plate cultures of each of the three species and eight independent swabs were processed per species. Swab heads were snapped off into 1.5 ml Eppendorf tubes containing 1 ml of 1 x PBS and samples were vortexed for 2 min before being centrifuged at 14000 ***g*** for 15 min to pellet the cells. The supernatants were removed, and the pellets were resuspended with 400 µl 1 x PBS. The method six procedure was then followed as described above.

Validation of DNA extraction method using volunteer and PEARL study skin swabs

The optimised DNA extraction method was tested on skin swabs from adults and babies to validate the selected method ability to obtain appropriate bacterial DNA for SMS and confirm data was suitable for analysing the taxonomic profiles of bacterial communities present on skin. Samples were cultured in parallel to DNA sequencing; this allowed us to identify organisms which should be represented in the SMS data whilst also enabling the creation of a skin microbiota culture collection for future functional work with strains of interest. Swabs were cultured aerobically and anaerobically on Columbia blood agar plates as in previous studies [26]. For each swab, cells grown on the aerobic and anaerobic plates were harvested into one glycerol stock, a sample of which was then used for DNA extraction and SMS to compare to results direct from swabs.

Study design for adult volunteer and PEARL study baby skin swab collection

The Norwich Research Park Biorepository recruited and consented 12 adult volunteers between the age of 23–65. There was no contact between the researcher and participants to ensure anonymity. Eligible volunteer participants had no current skin conditions or had been prescribed antibiotics over the last 3 months. The volunteer participants were provided with Participant Information Sheets (PIS) and were consented with Consent Forms (CF) and provided samples using a self-swabbing protocol

under observation and following instruction from Biorepository staff (File S2). The volunteers collected two swabs (using charcoal cotton swabs M40-A2, Technical Service Consultants Ltd.), one from the right arm and one from the left arm, to produce 24 samples in total. Samples were stored in a 4 °C fridge and anonymised with a unique barcode before being collected and tested on the same day swabbing was performed. In addition to the adult volunteers, swabs (using charcoal cotton swabs M40-A2, Technical Service Consultants Ltd.) from the skin of ten babies collected at 4 months of age, as part of the PEARL study, were also included (see Phillips *et al.*, [17] for study design and inclusion criteria, and Table S1 for baby participant metadata); the PEARL Study selected the use of charcoal cotton swabs due to their suitability for sampling multiple body sites (i.e. vagina, gut and skin) and long-term preservation of microbes in the −80 degrees Celsius freezer. Charcoal cotton swabs were therefore selected for use in all optimisation steps in this study to avoid a bias from the swabs used within our study and the PEARL study.

Volunteer and baby skin swab processing and finalised DNA extraction procedure

The skin swabs were processed as described above with the optimised method, a cell-free, diluent-only sample was included as a negative control on each extraction run and an established commercial mock community (the ATCC skin microbiome whole cell mix) was included as a positive control [27]. Dilutions of the positive control microbiome mix were also prepared to validate extraction efficiency and identify a cut-off point of starting material needed for SMS. For full details on the sample processing, DNA extraction protocol and the ATCC positive control protocol, see File S3.

DNA quantification and quality assessment

A High Sensitivity (HS) assay using the Qubit 2.0 fluorometer instrument and HS Qubit Invitrogen kit, was used to quantify all samples. If a concentration was out of range, i.e. too high, the Broad Range (BR) Qubit assay was used instead, using the Qubit 2.0 fluorometer instrument and BR Qubit Invitrogen kit. Tapestation assays were used to determine DNA molecular weight. A D5000 or HS D5000 Tapesation assay were used with an Agilent 2200 instrument and Agilent D5000 or HS D5000 kits.

Shotgun metagenomic sequencing using Illumina and Oxford Nanopore Technology

Preparation of libraries for SMS for both Illumina (Illumina DNA Prep Kit: 20018704) and ONT (Illumina DNA Prep: 20018704, Tagmentation: 20060059) platforms included DNA normalisation, tagmentation, PCR barcoding, quantification, pooling, and quality control. Samples were then loaded onto the Illumina NextSeq500 Instrument using a Mid-output 300 cycle kit (Illumina Catalogue FC-404–2003) or the MinION flow cell ONT instrument (R9.4.1). The QIB Bioinformatics team converted the Illumina raw data to eight FASTQ files for each sample, and the ONT raw data was converted into FASTQ files using the customised guppy method. All FASTQ files were then run through Fastp (V.0.19.5+galaxy1) [28], which is a quality-filtering tool for FASTQ files that removes adaptors. For full details on the SMS protocol for Illumina and ONT, view File S4.

Generating taxonomic profiles

All SMS data was automatically deposited in a local instance of IRIDA (irida-19.09.2) [29] and uploaded to the QIB Galaxy platform (V.19.05) [30]. Here, data was cleaned by removing adaptors and trimming reads, and filtered for quality using Fastp (V.0.20.0) (-q 20) [28], before reads mapping against a human reference database (human_20200311) were removed using Kraken2 (V.2.1.1+galaxy0) [31]. Remaining reads were then analysed to obtain microbiota taxonomic profiles using Kraken2 (V.2.1.1+galaxy0) [31] and Bracken (V.2.2) [32].

Metagenome-assembled genomes (MAG) extraction

Using the trimmed and filtered reads, host-associated (human) sequences were removed via Kneaddata (V.0.10.0) (The Huttenhower Lab [33]) with human genome database (GRCh38.p13) to generate clean fastq reads. Shotgun metagenome raw reads were co-assembled with MEGAHIT (V.1.2.9) [34] prior to extraction of MAGs. The MetaWRAP (V.1.3.2) pipeline [35] was used to extract MAGs based upon metagenome assemblies generated and metagenome clean reads via binning software 'metaBAT' (V.2.12.1) [36], 'MAXBIN2' (V.2.2.6) [37] and 'CONCOCT' (V.1.1.0) [38] using the sub-module 'binning'. MAGs were then refined using sub-module 'bin_refinement' to select the high-quality bins from each sample with completeness >80% and contamination <10% using CheckM (V.1.1.3) [39]. All MAGs were taxonomically ranked using gtdb-tk (V.1.5.1) [40] via module gtdbtk classify_wf.

Data visualisation

R (V.4.1.2) [41] and the package ggplot2 [42] were used to plot taxonomic profiles and alluvial and box plots. GraphPad Prism (V.5.04) [43] was used to generate scatter plots.

Statistical analysis

Statistical analysis was performed using unpaired t-tests in GraphPad Prism (V.5.04) [43]. A significance level of 0.05 was used to identify results likely to be different.

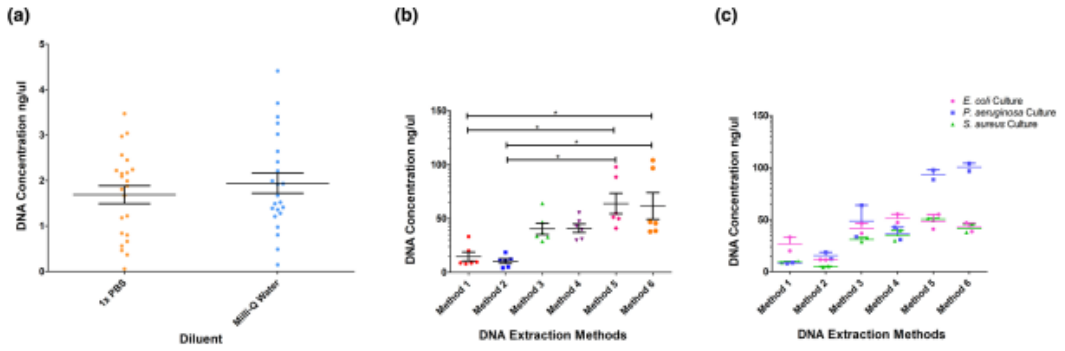


Fig. 2. Results of the variables tested. (a): Total DNA yield (ng µl^{-1}) from spiked swabs (n=44; 22 per diluent) processed in 1xPBS and Milli-Q water during a DNA extraction. (b): Total DNA yield (ng µl^{-1}) obtained from each DNA extraction method (n=36; six cultures per method). (c): Total DNA yield (ng µl^{-1}) per species for each method (n=36; six cultures per method). Horizontal bars on each plot show averages, vertical bars show the standard error of the mean (SEM) and lines with an asterisk (*) indicate significant (P<0.05) differences.

RESULTS

Optimisation of DNA extraction method

Impact of initial diluents on extracted bacterial DNA quantity and cell viability

There was no significant difference between amounts of bacterial DNA extracted from the 44 sterile charcoal cotton swabs spiked with *S. aureus* and processed in either PBS or water (Fig. 2a). Recovery of *S. aureus*, *P aeruginosa* and *E. coli*, also showed no significant differences in viable numbers recovered after suspension in either diluent (P>0.05; Fig. S1). As there was no significant difference in both DNA extraction and bacterial recovery between PBS and water, future experiments used PBS.

Testing lysis methods: six extraction method procedures

DNA extracted from liquid cultures of *S. aureus*, *P. aeruginosa* and *E. coli* using the six methods (Table 1), showed that methods 5 and 6 yielded the most DNA, (40.9–97.7 ng µl^{-1} and 37–104 ng µl^{-1} respectively), and there was a significant difference in DNA concentrations between methods 5 and 6 and methods 1 and 2 (Table S2; Fig. 2b, c). Methods 5 and 6 yielded higher DNA concentrations for *P. aeruginosa* (Fig. 2c) compared to other species, but there was no statistically significant difference in extraction efficiency between each bacterial species. DNA extraction methods 2, 4 and 6 produced higher molecular weights than the others, ranging from 20232 to 31786 bp (Table 2). Together, these results demonstrated that method six produced the most DNA of highest molecular weight. This method was also the most cost effective due to the cheaper lysozyme used and was chosen for further validation. This method included overnight lysis with lysozyme, a further heated offboard lysis step and a bead beating lysis using the Qiagen Tissue Lyser.

DNA extractions from sterile charcoal cotton swabs spiked with independent cultures were successful, with DNA concentrations averaged at 22.1 ng µl^{-1}.

DNA extraction method validation using swabs from volunteers or babies

DNA concentrations from adult and baby skin swabs, that were extracted using method 6, ranged from <0.50 (no detected DNA) – 10.5 ng µl^{-1} (Table S3) with DNA successfully extracted from all the baby samples, but only 15/24 adult volunteer

Table 2. Average molecular weight (bp) of DNA extracted

Sample	Method 1	Method 2	Method 3	Method 4	Method 5	Method 6
E. coli	0	25321	1735	21693	1909	24219
E. coli	0	25697	2057	20232	1877	22730
P. aeruginosa	0	23771	1763	23163	1678	29459
P. aeruginosa	0	23826	1631	23742	1538	31786
S. aureus	0	25085	1831	23954	1668	22118
S. aureus	0	22011	1762	24084	1711	22904

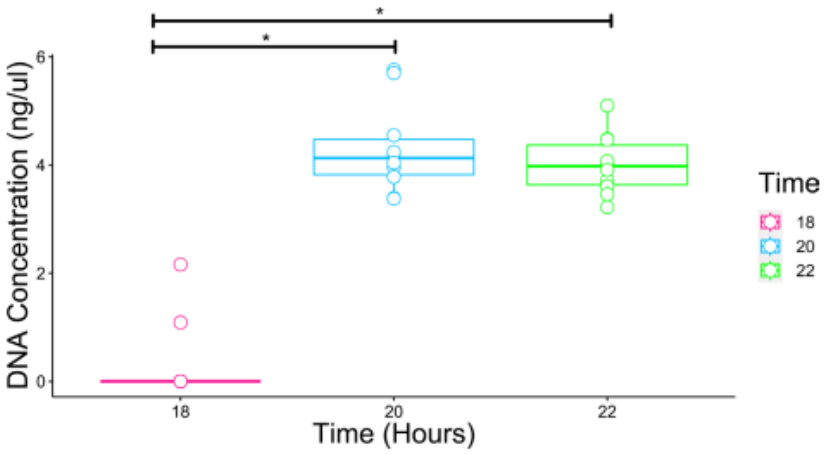


Fig. 3. Comparison of DNA yield (ng µl^{-1}) from swabs incubated for different periods (n=30; 10 swabs per incubation time). The box plots show the average DNA concentrations (ng µl^{-1}) for each incubation time. Horizontal bars on each plot show averages, vertical bars show the standard error of the mean (SEM) and lines with an asterisk (*) indicate significant (P<0.05) differences.

samples. Cultured plates recovered bacteria from all adult skin swabs, although recovery of cultures from the baby samples was only successful for 4/10 swabs. Concentrations of DNA extracted from cultured bacteria averaged at 79.9 ng µl^{-1}.

As some swabs did not yield DNA using method 6, we compared DNA yield from the extracted swabs after different overnight lysis incubation times. Samples were randomly incubated for either 18, 20 or 22 h (Fig. 3). A significant difference between 18 and 20 h and 18 and 22 h (P<0.05) was observed, but no significant difference between 20 and 22 h (P>0.05). Samples that did not yield detectable amounts of DNA were those incubated for 18 h therefore, future samples were incubated between 20–22 h to obtain higher DNA yield.

DNA extracted from skin microbiota is suitable for shotgun metagenomics with short and long read platforms

After removal of human reads, microbial taxonomic profiles were generated using both Illumina and ONT sequence data using Kraken2 and Bracken (Figs 4–6) swabs and cultures.

The positive controls displayed the expected microbiota from the ATCC skin microbiome whole cell mix – *Acinetobacter johnsonii*, *Corynebacterium striatum*, *Micrococcus luteus*, *Cutibacterium acnes*, *Staphylococcus epidermidis*, *Streptococcus mitis*. There was also a clear reduction in reads from these samples across the dilution series. This demonstrated that the DNA extraction method was able to effectively extract DNA from the diverse range of species present in the ATCC skin microbiome mix. Some background contamination was detected in negative controls, although the number of reads was always much lower than in samples and most DNA fragments identified in the negative controls mapped against organisms not seen in the test samples.

Both Illumina and ONT data indicated a typical skin microbiota from both adult and baby skin swabs and generated enough reads allowing downstream taxonomic analysis at the species level. The adult swabs identified bacteria, viruses, and phages, whereas the baby swabs only displayed bacterial diversity. Baby skin swabs contained more *Streptococcus* and fewer *Staphylococcus* species when compared to adult skin swabs. The baby skin swabs also indicated the presence of *Bifidobacterium longum*, *Bifidobacterium breve* and *Bifidobacterium bifidum*, which are not typical skin residents, but common residents of the infant gut, which likely demonstrates transient skin contamination on the babies [44, 45]. Importantly, data was generated for skin swabs that had very low DNA concentrations. Illumina and ONT platforms identified very similar microbiota profiles for both skin swabs and cultures, with comparable percentage total counts of the most abundant species (those representing more than 0.5% of each sample) (Fig. S2). Analysis of the taxonomic profile from cultured samples exhibited less microbial diversity than the skin swabs as expected but confirmed the presence of species identified in the SMS. As in the SMS data, adult cultures exhibited more *Staphylococcus* species than *Streptococcus*.

Once a successful DNA extraction method was established, the depth of sequence data required to provide optimal phylogenetic resolution and to construct MAGs were both assessed. This was done by comparing outcomes using 5Gbp per sample and subsamples thereof down to 1Gbp of data. For species identification a rarefaction curve was produced, which showed more species identified as more data was used; though statistical analysis showed there was not a significant difference in species recovery between 2.5 and 5Gbp of data (Fig. S3A). Recovery of MAGS was also higher from samples where 5Gbp of data were used than 1Gbp, although this difference was not found to be statistically significant (Fig. S3B). We were able to isolate MAGs predicted to be from diverse genera showing the method can isolate and

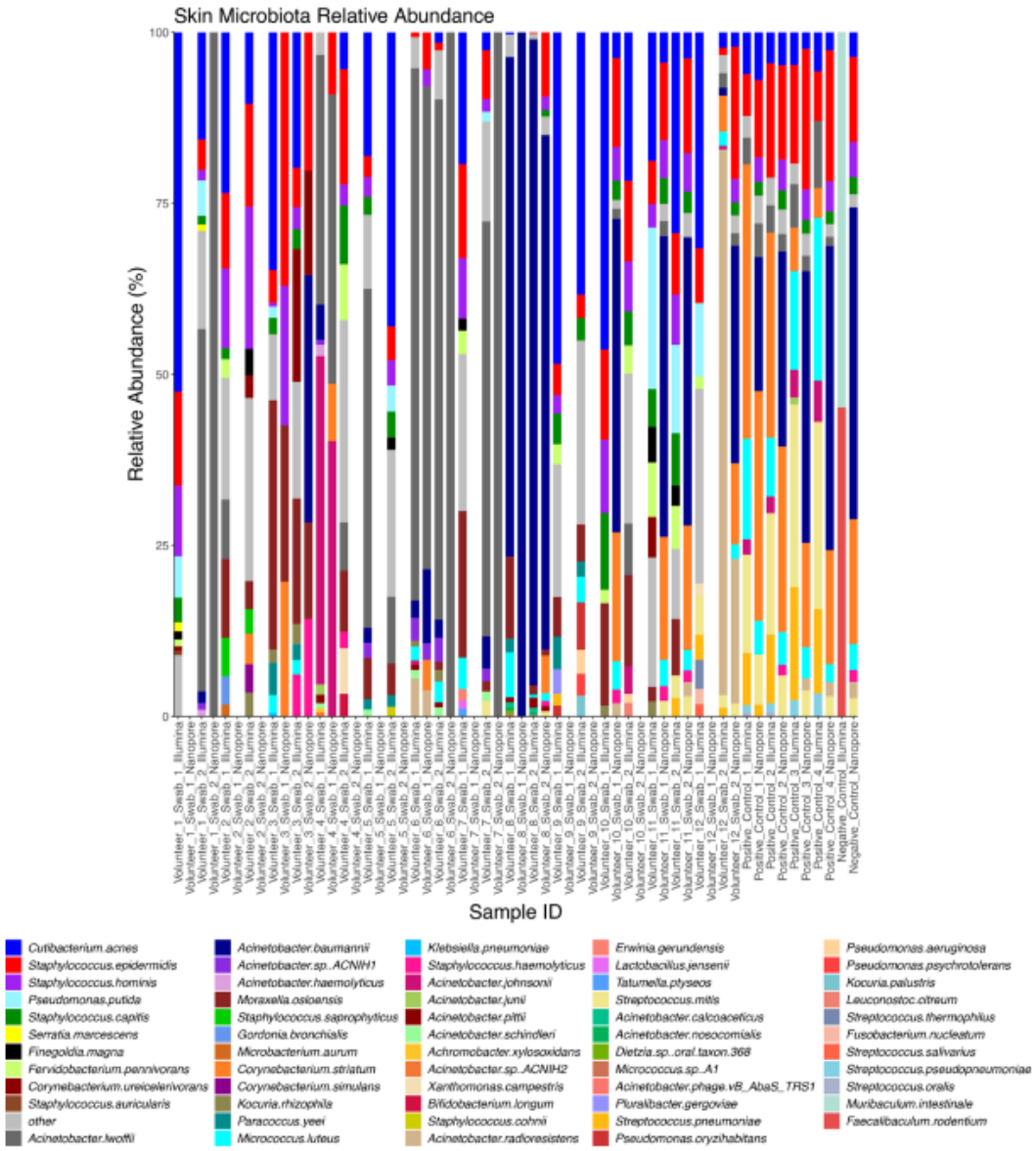


Fig. 4. Taxonomic profiles of skin swab microbiota from 12 adult volunteers (two swabs collected from both forearms from each volunteer) generated using Illumina and Nanopore data. Profiles show the relative abundance (percent) of the 10 most abundant species that occur within each sample.

assemble DNA from different parts of the phylogeny of the skin microbiota (Table S4). Based on this analysis, 5Gbp of data appears to be adequate for phylogenetic analysis of the skin microbiota using this method, whilst also providing enough data for generation of MAGs and therefore allowing detailed study of the functional capacity of the skin microbiome to be studied based on these results.

DISCUSSION

We aimed to develop an efficient protocol for DNA extraction suitable for use from both skin swabs and cultured bacterial cells. Initial testing showed both water and PBS were suitable diluents to maintain viability and for DNA extraction in agreement with previous studies [46–48] and PBS was then used throughout. Comparison of a variety of lysis procedures identified the



Fig. 5. Taxonomic profiles of skin swab culture microbiota from 12 adult volunteers (two swabs collected from both forearms from each volunteer) generated using Illumina and Nanopore data. Profiles show the relative abundance (percent) of the 10 most abundant species that occur within each sample.

effectiveness of a combined approach using both overnight heated enzymatic (lysozyme) and mechanical (bead beat) lysis methods to result in sufficient DNA yield of a high molecular weight from both Gram-positive and Gram-negative bacteria. Previous work has indicated that the type of enzyme and mechanical intensity is also important for lysis of different bacterial species [49–51]; however, our combined use of a mechanical and enzymatic lysis approach resulted in an unbiased extraction of Gram-positive and Gram-negative bacteria, which was validated by the production of expected profiles from the positive control mock community [52]. We used the Maxwell DNA isolation platform which was able to reliably produce DNA of good quality

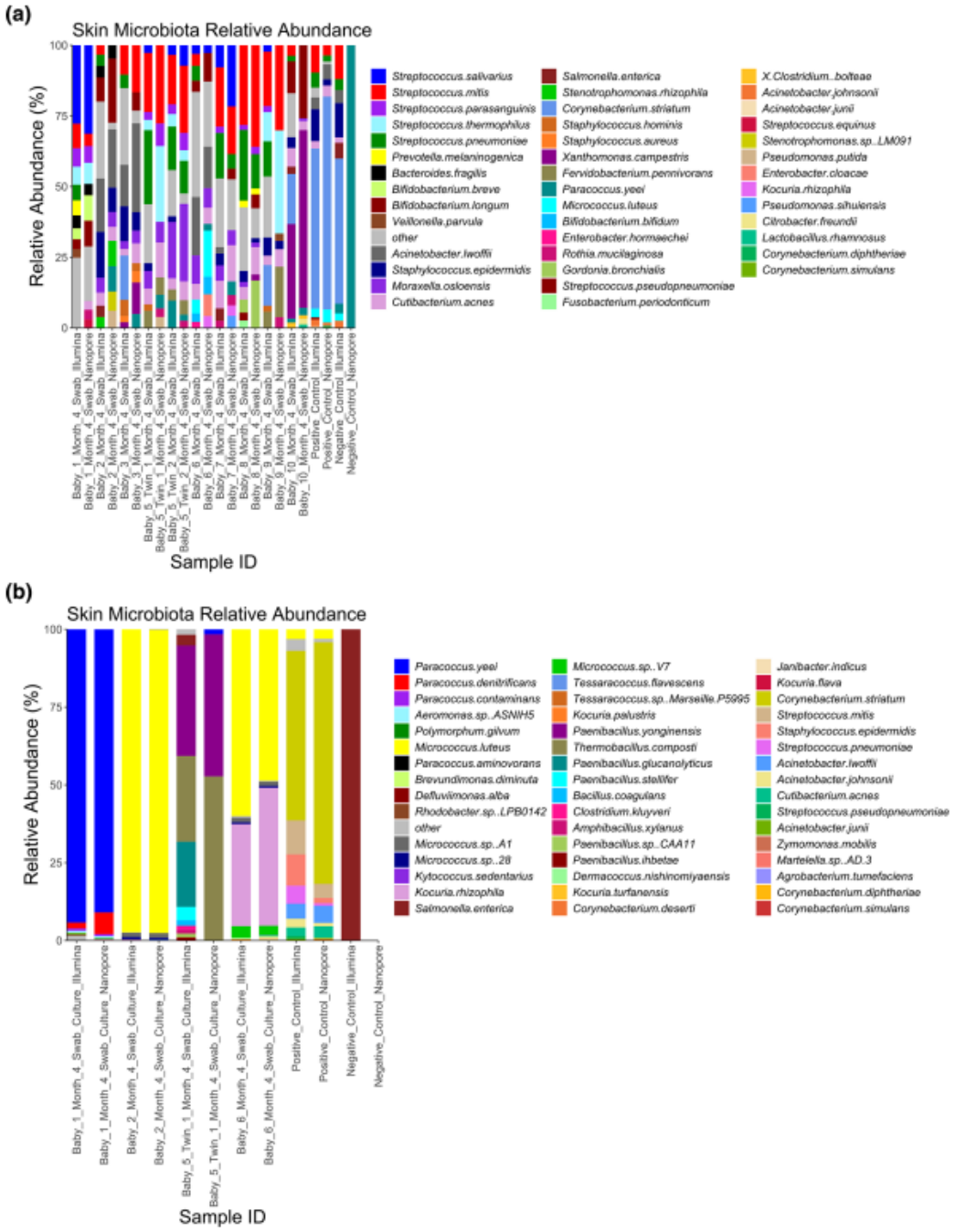


Fig. 6. Taxonomic profiles of the skin swab and culture microbiota from ten PEARL babies (one swab collected off one forearm from each baby) generated using Illumina and Nanopore data. Profiles show the relative abundance (percent) of the 10 most abundant species that occur within each sample. (a): Illumina and Nanopore skin swab data. (b): Illumina and Nanopore skin culture data.

and molecular weight suitable for SMS. Other platforms are available and likely to work well with this method, we did not seek to benchmark extraction kits, but this could be done in future to refine this method further.

Given the low biomass of skin microbiota, some of the adult skin swabs produced very low/absent DNA concentrations and paired cultures also indicated low bacterial burden. Individual variations when swabbing (pressure, direction, frequency) can affect the yield of DNA and viable bacteria, and are difficult to control [9] and may be responsible for this variation. A swabbing method was used as it is commonly used to collect skin microbiome samples [9] and was already used by our local PEARL study to collect samples due to its non-invasive nature, which is suitable for neonates, who have an underdeveloped skin structure [53, 54]. We also found a difference in sensitivity between platforms for samples with low amounts of DNA, some adult swabs did not produce data using the ONT platform although these same samples generated bacterial cultures. As the ONT platform requires more input DNA to generate data than Illumina platforms [16], the inability to generate data for some samples was not surprising as skin swabs can be low biomass [14, 15]. However, increasing the overnight incubation time did improve DNA yield, and the Illumina sequencing resulted in generated data for all samples.

Most samples did generate data from both Illumina and ONT platforms which presented similar microbiota profiles from skin swabs and cultures. Typical adult skin microbiota (Phyla; *Pseudomonadota*, *Actinomycetota*, and *Bacillota*) [2, 55, 56] and infant skin microbiota (Phyla; *Bacillota*, *Actinomycetota*, *Pseudomonadota*, and *Bacteroidota*) [57] were detected. We focused on bacterial species identified, but the protocol did identify other skin microbiota (viruses, phages and fungi), although only from adult volunteers [55]. Other researchers can use this protocol as a starting point to be adapted if these organisms are their focus. Baby profiles only contained bacteria, and demonstrated less microbial diversity than adults, which has been shown in previous studies [58]. Baby skin did exhibit more *Streptococcus* species than adult skin, which agrees with previous work demonstrating a predominance of *Streptococcus* species in early age, which decreases with age [57, 58]. Interestingly, sequencing of swabs from infant skin identified *Bifidobacterium* species, which are not typical skin residents, but rather maternal and infant gut residents and they can also be found in breast milk [44, 45]. Given the paired cultures did not result in any *Bifidobacterium* isolates, this is likely to indicate transient transfer to the babies' skin through breast feeding. The babies with available metadata that showed *Bifidobacterium* presence on the skin were all breast fed at some point between birth and month 4.

Whilst skin is a relatively low biomass environment, we did not need to include any methods to mechanically deplete human DNA or selectively enrich microbial DNA before SMS [59], which have been needed in some other studies on low biomass samples. These enrichment approaches do not reliably target all species [59], can skew the resulting genomic profiles [60] and depletion can result in some loss of bacteria [59], thus further steps are required for downstream analysis. In our described method, we generated enough data, and depleted human DNA computationally, therefore precluding the need for any additional steps that may introduce biases and skew skin microbiota profiles.

Both Illumina and ONT sequence data allowed identification of all ATCC positive control species, with a clear reduction in read number across the dilution series. These results further demonstrate the effectiveness of the extraction method and utility of both sequencing platforms. Inclusion of a commercially available mixed community positive control, with a known cell concentration, is important for standardising the extraction process, and serial diluting the positive control can determine the limit of detection [61]. This is also helpful when comparing different sequencing runs and sample sets, allowing more robust comparisons to be made. Although, we tried to define a limit of detection for DNA concentration and read number required for effective SMS, we had several swab samples that did not obtain a DNA concentration reading, but usable reads were produced for taxonomic profiling. Therefore, no obvious cut-off for a limit of detection was determined, and indeed there is also no 'defined' limit identified in the literature for low biomass samples, such as skin swabs.

We did identify some background contamination in the negative controls, contamination commonly occurs in metagenomic studies, especially those with low biomass samples [62]. Several studies have identified contamination sources occurring from neighbouring samples and the 'kitome' [62, 63]. Contamination within a dataset can be identified and removed using bioinformatic techniques [64, 65] although low biomass samples have a higher risk of true microbial microbiota members being removed [66]. Given the background contaminants in the controls were at a very low level and mostly represented species not seen in the test samples, we did not remove them as they had a negligible impact on the profiles produced.

We determined that the generation of 5Gbp of Illumina data from a skin swab was suitable for microbial species profiling and produced a limited number of MAGs. MAGs are important for in-depth functional information [13] and indicate genome quality [39, 67, 68], and they can be used to identify novel taxa and allow further comparison with whole genome sequence data from isolates. Our method is compatible with both Illumina and ONT platforms, and combining a higher sequencing depth with ONT, the data has potential to improve the number and quality of MAGs to be recovered [13, 69, 70].

CONCLUSION

An optimised medium-throughput DNA extraction, SMS, and analysis approach can effectively characterise the skin microbiota from adults and babies. This method can be applied for in-depth analysis of cohort studies allowing identification of taxonomic

and functional changes of mothers and infants over time and should allow comparison to other body sites (e.g. the gut). Robust microbiota profiling, particularly in less well studied niches such as the skin, is important for the development of methods to alter microbiome compositions for health.

Funding information
This research was supported in part by the NBI Computing infrastructure for Science (CiS) group through the provision of a High-Performance Computing (HPC) Cluster. I.R.S. is funded by a Biotechnology and Biological Sciences Research Council (BBSRC) CTP studentship with Unilever (BB/T508974/1). L.J.H. is supported by Wellcome Trust Investigator Awards 100974/C/13/Z and 220876/Z/20/Z; and a BBSRC Institute Strategic Programme, Gut Microbes and Health BB/R012490/1, and its constituent projects BBS/E/F/000PR10353 and BBS/E/F/000PR10356. M.A.W. is supported by project grant (BB/T014644/1) from the Biotechnology and Biological Sciences Research Council and BBSRC Institute Strategic Programmes Microbes in the Food Chain BB/R012504/1 and its constituent project BBS/E/F/000PR10349.

Author contributions
Conceptualisation, I.R.S., B.M., M.A.W., L.J.H.; Data curation, I.R.S., B.M., M.A.W., L.J.H., M.D., R.K., D.B., R.E., S.P., R.W. and T.A.; Formal analysis, I.R.S., B.M., M.A.W., L.J.H., M.D. and R.K.; Funding acquisition, B.M., M.A.W. and L.J.H.; Investigation, I.R.S., B.M., M.A.W., L.J.H., D.B., R.E., S.P., R.W. and T.A.; Methodology, I.R.S., B.M., M.W., L.J.H., D.B. and R.E.; Project administration, I.R.S., B.M., M.A.W., L.J.H., S.P., R.W. and T.A.; Resources, I.R.S., B.M., M.A.W., L.J.H., D.B., R.E., M.D., R.K., S.P., R.W., T.A.; Supervision, I.R.S., B.M., M.A.W. and L.J.H.; Validation, I.R.S., D.B., R.E., M.D. and R.K.; Visualisation, I.R.S., M.D. and R.K.; Writing – original draft, I.R.S., B.M., M.A.W., L.J.H., D.B., R.E., S.P., R.W., T.A., M.D., and R.K.; Writing – review and editing, I.R.S., E.T., B.M., M.A.W. and L.J.H.

Conflicts of interest
The authors declare no competing interests.

Ethical statement
Ethical approval was obtained for the adult volunteer recruitment, skin swab sampling and processing from the University of East Anglia (UEA) Faculty of Medicine and Health Sciences (FHM); The recruitment, sampling and processing was performed under the Norfolk and Norwich University Hospital (NNUH) Biorepository ethics - FMH ethical approval reference: 2020/21–065. The PEARL study was approved by The Quadram Institute Biosciences (QIB) Human Research Governance Committee (HRGC), local Research Ethics Committee (REC), and Health Research Authority (HRA). This study was conducted in accordance with the principles of the Declaration of Helsinki. The proposed research was conducted in accordance with the conditions and principles of the International Conference on Harmonisation Good Clinical Practice (ICH GCP), and in compliance with the UK national law. The research meets the requirements of the EU General Data Protection Regulation (GDPR), UK Data Protection Act 2018 and relevant sponsor's policies - IRAS number: 241880.

References

1. **National Academies of Sciences, Engineering and Medicine (NASEM)**. *Environmental Chemicals, the Human Microbiome, and Health Risk: A Research Strategy*. Washington, DC, USA: National Academies Press; 2018.
2. **Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, *et al.*** Bacterial community variation in human body habitats across space and time. *Science* 2009;326:1694–1697.
3. **Cho I, Blaser MJ**. The human microbiome: at the interface of health and disease. *Nat Rev Genet* 2012;13:260–270.
4. **Human Microbiome Project Consortium**. Structure, function and diversity of the healthy human microbiome. *Nature* 2012;486:207–214.
5. **Kong HH**. Skin microbiome: genomics-based insights into the diversity and role of skin microbes. *Trends Mol Med* 2011;17:320–328.
6. **Jo JH, Kennedy EA, Kong HH**. Research techniques made simple: bacterial 16S ribosomal RNA gene sequencing in cutaneous research. *J Invest Dermatol* 2016;136:e23–e27.
7. **Allaband C, McDonald D, Vázquez-Baeza Y, Minich JJ, Tripathi A, *et al.*** Microbiome 101: studying, analyzing, and interpreting gut microbiome data for clinicians. *Clin Gastroenterol Hepatol* 2019;17:218–230.
8. **Kuczynski J, Lauber CL, Walters WA, Parfrey LW, Clemente JC, *et al.*** Experimental and analytical tools for studying the human microbiome. *Nat Rev Genet* 2012;13:47–58.
9. **Sfriso R, Egert M, Gempeler M, Voegeli R, Campiche R**. Revealing the secret life of skin - with the microbiome you never walk alone. *Int J Cosmet Sci* 2020;42:116–126.
10. **Liu Y-X, Qin Y, Chen T, Lu M, Qian X, *et al.*** A practical guide to amplicon and metagenomic analysis of microbiome data. *Protein Cell* 2021;12:315–330.
11. **Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, *et al.*** Opportunities and challenges in long-read sequencing data analysis. *Genome Biol* 2020;21:30.
12. **Pearman WS, Freed NE, Silander OK**. Testing the advantages and disadvantages of short- and long- read eukaryotic metagenomics using simulated reads. *BMC Bioinformatics* 2020;21:220.
13. **Singleton CM, Petriglieri F, Kristensen JM, Kirkegaard RH, Michaelsen TY, *et al.*** Connecting structure to function with the recovery of over 1000 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing. *Nat Commun* 2021;12:2009.
14. **Bjerre RD, Hugerth LW, Boulund F, Seifert M, Johansen JD, *et al.*** Effects of sampling strategy and DNA extraction on human skin microbiome investigations. *Sci Rep* 2019;9:17287.
15. **de Goffau MC, Lager S, Salter SJ, Wagner J, Kronbichler A, *et al.*** Recognizing the reagent microbiome. *Nat Microbiol* 2018;3:851–853.
16. **Wang Y, Zhao Y, Bollas A, Wang Y, Au KF**. Nanopore sequencing technology, bioinformatics and applications. *Nat Biotechnol* 2021;39:1348–1365.
17. **Phillips S, Watt R, Atkinson T, Savva GM, Hayhoe A, *et al.*** The Pregnancy and EARly Life study (PEARL) - a longitudinal study to understand how gut microbes contribute to maintaining health during pregnancy and early life. *BMC Pediatr* 2021;21:357.
18. **Bey BS, Fichot EB, Dayama G, Decho AW, Norman RS**. Extraction of high molecular weight DNA from microbial mats. *Biotechniques* 2010;49:631–640.
19. **MandrekarPV, Ma Z, Krueger S, Cowan C**. High-Concentration (>100ng/µl) Genomic DNA From Whole Blood Using the Maxwell® 16 Low Elution Volume Instrument. American Medical Association, Manual of Style, 10th edition. Promega Corporation Web site; 2010. https://www.promega.co.uk/resources/pubhub/high-concentration-genomic-dna-from-whole-blood-using-the-maxwell-16-low-elution-volume-instrument [accessed 1 November 2020].
20. **Dunbar J, Gallegos-Graves LV, Gans J, Morse SA, Pillai S, *et al.*** Evaluation of DNA extraction methods to detect bacterial targets in aerosol samples. *J Microbiol Methods* 2018;153:48–53.
21. **Moeller JR, Moehn NR, Waller DM, Givnish TJ**. Paramagnetic cellulose DNA isolation improves DNA yield and quality among diverse plant taxa. *Appl Plant Sci* 2014;2:apps.1400048.

22. Promega Maxwell® RSC Blood DNA Kit; (n.d.). https://www.promega.com/-/media/files/resources/protocols/technical-manuals/101/maxwell-rsc-blood-dna-kit-protocol.pdf?la=en [accessed 1 November 2020].
23. **Sui H-Y, Weil AA, Nuwagira E, Qadri F, Ryan ET, *et al.*** Impact of DNA extraction method on variation in human and built environment microbial community and functional profiles assessed by shotgun metagenomics sequencing. *Front Microbiol* 2020;11:953.
24. **Gill C, van de Wijgert J, Blow F, Darby AC, Larsen PE.** Evaluation of lysis methods for the extraction of bacterial DNA for analysis of the vaginal microbiota. *PLoS One* 2016;11:e0163148.
25. **Martzy R, Bica-Schröder K, Pálvölgyi ÁM, Kolm C, Jakwerth S, *et al.*** Simple lysis of bacterial cells for DNA-based diagnostics using hydrophilic ionic liquids. *Sci Rep* 2019;9:13994.
26. **Ogai K, Nagase S, Mukai K, Iuchi T, Mori Y, *et al.*** A comparison of techniques for collecting skin microbiome samples: swabbing versus tape-stripping. *Front Microbiol* 2018;9:2812.
27. ATCC Skin Microbiome Whole Cell Mix; (n.d.). https://www.atcc.org/products/msa-2005 [accessed 1 December 2022].
28. **Chen S, Zhou Y, Chen Y, Gu J.** fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 2018;34:i884–i890.
29. **Matthews TC, Bristow FR, Griffiths EJ, Petkau A, Adam J, *et al.*** The Integrated Rapid Infectious Disease Analysis (IRIDA) platform. *bioRxiv* 2018. DOI: 10.1101/381830.
30. **Afgan E, Baker D, Batut B, van den Beek M, Bouvier D, *et al.*** The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res* 2018;46:W537–W544.
31. **Wood DE, Lu J, Langmead B.** Improved metagenomic analysis with Kraken 2. *Genome Biol* 2019;20:257.
32. **Lu J, Breitwieser FP, Thielen P, Salzberg SL.** Bracken: estimating species abundance in metagenomics data. *PeerJ Comput Sci* 2017;3:e104.
33. The Huttenhower Lab. Kneaddata; (n.d.). https://huttenhower.sph.harvard.edu/kneaddata/ [accessed 1 January 2023].
34. **Li D, Liu C-M, Luo R, Sadakane K, Lam T-W.** MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph . *Bioinformatics* 2015;31:1674–1676.
35. **Uritskiy GV, DiRuggiero J, Taylor J.** MetaWRAP-a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 2018;6:158.
36. **Kang DD, Froula J, Egan R, Wang Z.** MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 2015;3:e1165.
37. **Wu Y-W, Simmons BA, Singer SW.** MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 2016;32:605–607.
38. **Alneberg J, Bjarnason BS, Bruijn I, Schirmer M, Quick J, *et al.*** CONCOCT: clustering contigs on coverage and composition. *arXiv* 2013:1312.4038.
39. **Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW.** CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–1055.
40. **Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH, Hancock J.** GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 2019;36:1925–1927.
41. **RStudio Team.** RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA; 2021. http://www.rstudio.com/
42. **Wickham H.** ggplot2. In: *Elegant Graphics for Data Analysis*, 1st Edition. New York, NY: Springer-Verlag, 2009.
43. **GraphPad Software.** GraphPad Prism 5.04. San Diego California; 2010. https://www.graphpad.com/company [accessed 1 January 2023].
44. **Toscano M, De Grandi R, Grossi E, Drago L.** Role of the human breast milk-associated microbiota on the newborns' immune system: a mini review. *Front Microbiol* 2017;8:2100.
45. **Yan W, Luo B, Zhang X, Ni Y, Tian F.** Association and occurrence of bifidobacterial phylotypes between breast milk and fecal microbiomes in mother–linfant dyads during the first 2 years of life. *Front Microbiol* 2021;12:669442.
46. **Banning N, Toze S, Mee BJ.** Escherichia coli survival in groundwater and effluent measured using a combination of propidium iodide and the green fluorescent protein. *J Appl Microbiol* 2002;93:69–76.
47. **Downey AS, Da Silva SM, Olson ND, Filliben JJ, Morrow JB.** Impact of processing method on recovery of bacteria from wipes used in biological surface sampling. *Appl Environ Microbiol* 2012;78:5872–5881.
48. **Liao C-H, Shollenberger LM.** Survivability and long-term preservation of bacteria in water and in phosphate-buffered saline*. *Lett Appl Microbiol* 2003;37:45–50.
49. **Albertsen M, Karst SM, Ziegler AS, Kirkegaard RH, Nielsen PH.** Back to basics--the influence of DNA extraction and primer choice on phylogenetic analysis of activated sludge communities. *PLoS One* 2015;10:e0132783.
50. **Schindler CA, Schuhardt VT.** Lysostaphin: a new bacteriolytic agent for the Staphylococcus. *Proc Natl Acad Sci* 1964;51:414–421.
51. **Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ, *et al.*** Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PLoS One* 2012;7:e33865.
52. **Maghini DG, Moss EL, Vance SE, Bhatt AS.** Improved high-molecular-weight DNA extraction, nanopore sequencing and metagenomic assembly from the human gut microbiome. *Nat Protoc* 2021;16:458–471.
53. **Chiou YB, Blume-Peytavi U.** Stratum corneum maturation. A review of neonatal skin function. *Skin Pharmacol Physiol* 2004;17:57–66.
54. **Narendran V, Visscher MO, Abril I, Hendrix SW, Hoath SB.** Biomarkers of epidermal innate immunity in premature and full-term infants. *Pediatr Res* 2010;67:382–386.
55. **Byrd AL, Belkaid Y, Segre JA.** The human skin microbiome. *Nat Rev Microbiol* 2018;16:143–155.
56. **Grice EA, Kong HH, Conlan S, Deming CB, Davis J, *et al.*** Topographical and temporal diversity of the human skin microbiome. *Science* 2009;324:1190–1192.
57. **Capone KA, Dowd SE, Stamatas GN, Nikolovski J.** Diversity of the human skin microbiome early in life. *J Invest Dermatol* 2011;131:2026–2032.
58. **Zhu T, Liu X, Kong F-Q, Duan Y-Y, Yee AL, *et al.*** Age and mothers: potent influences of children's skin microbiota. *J Invest Dermatol* 2019;139:2497–2505.
59. **Marquet M, Zöllkau J, Pastuschek J, Viehweger A, Schleußner E, *et al.*** Evaluation of microbiome enrichment and host DNA depletion in human vaginal samples using Oxford Nanopore's adaptive sequencing. *Sci Rep* 2022;12:4000.
60. **Hammond M, Homa F, Andersson-Svahn H, Ettema TJG, Joensson HN.** Picodroplet partitioned whole genome amplification of low biomass samples preserves genomic diversity for metagenomic analysis. *Microbiome* 2016;4:52.
61. **Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, *et al.*** Contamination in low microbial biomass microbiome studies: issues and recommendations. *Trends Microbiol* 2019;27:105–117.
62. **Lou YC, Hoff J, Olm MR, West-Roberts J, Diamond S, *et al.*** Using strain-resolved analysis to identify contamination in metagenomics data. *BioRxiv* 2022. DOI: 10.1101/2022.01.16.476537.
63. **Olomu IN, Pena-Cortes LC, Long RA, Vyas A, Krichevskiy O, *et al.*** Elimination of "kitome" and "splashome" contamination results in lack of detection of a unique placental microbiome. *BMC Microbiol* 2020;20:157.
64. **Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ.** Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 2018;6:226.

65. **Zhou Q**, **Su X**, **Ning K**. Assessment of quality control approaches for metagenomic data analysis. *Sci Rep* 2014;4:6957.
66. **Díaz S**, **Escobar JS**, **Avila FW**, **Zhu L**. Identification and removal of potential contaminants in 16S rRNA gene sequence data sets from low-microbial-biomass samples: an example from mosquito tissues. *ASM* 2021;6:e0050621.
67. **Bowers RM**, **Kyrpides NC**, **Stepanauskas R**, **Harmon-Smith M**, **Doud D**, *et al.* Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat Biotechnol* 2017;35:725–731.
68. **Sczyrba A**, **Hofmann P**, **Belmann P**, **Koslicki D**, **Janssen S**, *et al.* Critical assessment of metagenome interpretation-a benchmark of metagenomics software. *Nat Methods* 2017;14:1063–1071.
69. **DeMaere MZ**, **Darling AE**. bin3C: exploiting Hi-C sequencing data to accurately resolve metagenome-assembled genomes. *Genome Biol* 2019;20:46.
70. **Gweon HS**, **Shaw LP**, **Swann J**, **De Maio N**, **AbuOun M**, *et al.* The impact of sequencing depth on the inferred taxonomic composition and AMR gene content of metagenomic samples. *Environ Microbiome* 2019;14:7.

Five reasons to publish your next article with a Microbiology Society journal

1. When you submit to our journals, you are supporting Society activities for your community.
2. Experience a fair, transparent process and critical, constructive review.
3. If you are at a Publish and Read institution, you'll enjoy the benefits of Open Access across our journal portfolio.
4. Author feedback says our Editors are 'thorough and fair' and 'patient and caring'.
5. Increase your reach and impact and share your research more widely.

Find out more and submit your article at microbiologyresearch.org.

Dalby MJ, Kiu R, Serghiou IR, Miyazaki A, Acford-Palmer H, Tung R, Cain S, Phillips S, Kujawska M, Matsui M, Iwamoto A, Taking B, Cox SE, Hall LJ. Gut microbiota and faecal cytokine profiles of rural Cambodian infants and links to early life nutrition and diarrhoeal episodes. npj Biofilms Microbiomes 10, 85 (2024).

npj | biofilms and microbiomes

Article

Published in partnership with Nanyang Technological University

https://doi.org/10.1038/s41522-024-00562-0

Faecal microbiota and cytokine profiles of rural Cambodian infants linked to diet and diarrhoeal episodes

Check for updates

Matthew J. Dalby[1,2], Raymond Kiu[1,2], Iliana R. Serghiou[2], Asuka Miyazaki[3], Holly Acford-Palmer[2], Rathavy Tung[4], Shabhonam Caim[2], Sarah Phillips[2], Magdalena Kujawska[1,2,5], Mitsuaki Matsui[3], Azusa Iwamoto[6], Bunsreng Taking[7], Sharon E. Cox[3,8,9] & Lindsay J. Hall[1,2,5,10]

The gut microbiota of infants in low- to middle-income countries is underrepresented in microbiome research. This study explored the faecal microbiota composition and faecal cytokine profiles in a cohort of infants in a rural province of Cambodia and investigated the impact of sample storage conditions and infant environment on microbiota composition. Faecal samples collected at three time points from 32 infants were analysed for microbiota composition using 16S rRNA amplicon sequencing and concentrations of faecal cytokines. Faecal bacterial isolates were subjected to whole genome sequencing and genomic analysis. We compared the effects of two sample collection methods due to the challenges of faecal sample collection in a rural location. Storage of faecal samples in a DNA preservation solution preserved *Bacteroides* abundance. Microbiota analysis of preserved samples showed that *Bifidobacterium* was the most abundant genus with *Bifidobacterium longum* the most abundant species, with higher abundance in breast-fed infants. Most infants had detectable pathogenic taxa, with *Shigella* and *Klebsiella* more abundant in infants with recent diarrhoeal illness. Neither antibiotics nor infant growth were associated with gut microbiota composition. Genomic analysis of isolates showed gene clusters encoding the ability to digest human milk oligosaccharides in *B. longum* and *B. breve* isolates. Antibiotic-resistant genes were present in both potentially pathogenic species and in *Bifidobacterium*. Faecal concentrations of Interlukin-1alpha and vascular endothelial growth factor were higher in breast-fed infants. This study provides insights into an underrepresented population of rural Cambodian infants, showing pathogen exposure and breastfeeding impact gut microbiota composition and faecal immune profiles.

Research on the gut microbiota has primarily focussed on populations in high-income countries (e.g. Europe and North America)[1], and the problems and limitations this bias is introducing to microbiome research are gaining increasing attention[2]. Although more recent studies have increased the representation of populations in low- and low-middle-income countries (LMICs), infant studies lag behind those working on adult cohorts[3]. This is important for defining what a 'healthy' human gut microbiota[4] is across individuals in LMICs[3], and because optimal infant development is particularly important for infants from LMICs due to increased risk of (mal) nutrition and infection. While a small proportion of microbiota samples studied globally so far have come from South Asia, including India and Bangladesh[1], research into the gut microbiota of people living in other countries in the region remains very limited. Moreover, due to the logistics of sample collection and storage in rural areas, research in LMICs has often

[1]Microbes, Infection and Microbiomes, College of Medicine and Health, University of Birmingham, Birmingham, B15 2TT, UK. [2]Food, Microbiome & Health, Quadram Institute Bioscience, Norwich Research Park, Norwich, NR4 7UQ, UK. [3]School of Tropical Medicine & Global Health, Nagasaki University, Nagasaki, Japan. [4]National Maternal and Child Health Centre, Ministry of Health, Phnom Penh, Cambodia. [5]Intestinal Microbiome, School of Life Sciences, ZIEL – Institute for Food & Health, Technical University of Munich, Freising, 80333, Germany. [6]Bureau of International Health Cooperation, National Centre for Global Health and Medicine, Tokyo, Japan. [7]Kampong Cham Provincial Health Department, Ministry of Health, Kampong Cham, Cambodia. [8]London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK. [9]Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan. [10]Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK. e-mail: l.hall.3@bham.ac.uk

focused on individuals in urban areas, neglecting key infant populations. Cambodia, in particular, has a scarcity of microbiota studies, and given that this country has such a high prevalence of infants classified as stunted/wasted and has amongst the highest infant mortality rates in Southeast Asia, with infections accounting for many of these deaths[5], new studies targeting Cambodian infant populations are needed. Previous research on the gut-associated microbiota of people in Cambodia has been limited with studies focusing on particular bacterial pathogens, such as antimicrobial resistance in *Enterobacterales* in faecal samples from babies and children[6], the epidemiology of antibiotic resistance in *Escherichia coli* and *Klebsiella pneumoniae* in children[7], and effects of iron supplementation in Cambodian women[8]. However, an analysis of the composition of the gut microbiota as a whole in infants from Cambodia has not previously been published.

While the nutritional status of infants is improving in Cambodia, children under the age of 5 in rural areas still face challenges, with 22% suffering linear growth faltering or 'stunting'[9]. This is also true for infants (<1 year) in rural Cambodia, who suffer from similar levels of growth stunting at 19.2%[9], with a high proportion also lacking access to appropriate hygiene facilities. This is coupled with a high proportion reportedly experiencing symptoms of diarrhoea, fever or cough, and frequent use of unregulated antibiotic treatment to treat these childhood illnesses[10]. Malnutrition has previously been linked to differences in the gut microbiota of children from several different low-income countries[11,12], although research into these associations is lacking for children from Cambodia.

The establishment of the gut microbiota during infancy is an important aspect of healthy infant development with roles in immune, metabolic, endocrine, and other host developmental pathways[13]. Vaginally delivered, breast-fed infants are characterised by their very high abundance of the bacterial genus *Bifidobacterium*. This keystone microbiota member orchestrates wider microbiome structuring and modulates host physiology through the breakdown of complex dietary components such as human milk oligosaccharides (HMOs) in breast milk and plant-based carbohydrates, leading to the production of key metabolites that facilitate immune system development and maturation and cognitive responses[14]. Previous research indicates that LMIC infant microbiomes are characterised by *Bifidobacterium* species and strains, although these are often different genetic variants when compared to those infants from high-income country settings[3,15].

Infants born into resource-poor environments can have difficulties in establishing a healthy gut microbial ecosystem (Acosta et al, 2014)[16] due to multiple factors such as birth and breast-feeding practices, antibiotic exposure, maternal microbiota, diet, socio-economic status, host genetics, agricultural dependence, urbanisation, access to clean water and hygiene practices[3,13]. Infant gut microbiomes of those in LMICs are reportedly higher in *Segatella* (previously *Prevotella*) species, associated with both positive and negative health outcomes, and may correlate with higher plant-based, fibre-rich diets[3]. The phylum Pseudomonadota (previously Proteobacteria), which includes a variety of pathogenic and opportunistic pathogens, is highly prevalent in microbiomes of those in LMICs[3]. These factors may negatively affect the composition of the early infant gut microbiota[13] and may drive these observed microbiome differences across populations[3].

In this study, we carried out an exploratory analysis of the gut microbiota in faecal samples collected from 32 infants that are part of the Nutrition for Health of Aka-chan (baby in Japanese) and Mamas' cohort (NHAM) based in Cambodia, which included background data on infant diet, health, and living conditions. This provides an opportunity to explore the composition of the gut microbiota in rural Cambodian infants that has not previously been investigated. The samples from this cohort are unusual due to the rural location in which these infants live, with poor access to hygiene facilities, high prevalence of pathogen exposure, and frequent antibiotic usage. We analysed the effects of faecal sample storage on the composition of the microbiota of samples collected under suboptimal conditions. Then described the infant faecal microbiota composition and explored its associations with environmental factors, such as infant diet and growth. Individual faecal bacteria were cultured and isolated, and their genomes were sequenced to identify the presence of antibiotic resistance and HMO genes. Finally, the concentrations of faecal cytokines were measured to identify any associations with the gut microbiota. We aimed to explore the composition of the gut microbiota of these infants and its associations with their environment.

Results

The longitudinal infant faecal samples analysed in this study were sourced from 32 infants representing a subset of infants that are part of the NHAM birth cohort based in the rural Kampong Cham province in southeast Cambodia, and ~60 km from the provincial capital Kampong Cham[9,10] (Table 1). Samples were collected from infants living in a riverside community stretching along ~15.5 km of riverbank of the Mekong River in the communes of Khbop Ta Nguon, Preah Andoung, and Peam Koh Sna[9]. The infants were living in a rural area, with most families being small-scale farmers.

Sample storage conditions influence gut microbiota profiles

Sample collection and storage are challenging in areas without the facilities to immediately freeze faecal samples after collection. To help ensure better preservation of DNA in faecal samples, half of each faecal sample was immediately placed into a DNA preservation solution (DNAShield), with the other half remaining untreated. This later enabled the comparison of microbiota composition under different storage conditions. The faecal samples originated in a rural location with ~2 days between the sample being produced and final laboratory storage. Stool samples were collected by mothers or other caregivers and then kept in a relatively cooler place in the house or in cooler bags with ice until the field worker visited to collect them. After collection, the untreated half of each sample was transported in ice-boxes and then initially frozen at −20 °C, and then later stored at −80 °C. Half of each sample stored in DNAShield was transported at ambient temperature and then stored refrigerated at 4 °C. Three samples from 32

Table 1 | Infant baseline characteristics

Variable	Infants
n	32
Age (days) (mean (SD))	214.9 (37.3)
Sex = Female (%)	15 (46.9)
Breast feeding = No (%)	9 (28.1)
Complementary feeding = No (%)	17 (53.1)
Feeding bottle use = Yes (%)	21 (65.6)
Low birth weight (<2.5 kg) (%)	
Normal	29 (90.6)
Unknown	2 (6.2)
Low	1 (3.1)
Growth = Stunted (%)	3 (9.4)
Length-for-age *z*-score (mean (SD))	−0.9 (0.8)
Underweight *z*-score (mean (SD))	−0.6 (0.8)
Weight-for-length (mean (SD))	0.0 (1.0)
Hand washing score (mean (SD))	5.2 (1.6)
Improved toilet facility = Yes (%)	22 (68.8)
Household toilet facility = Septic tank (%)	22 (68.8)
Number in household (mean (SD))	5.8 (2.2)
Antibiotic use (previous 7 days) = No (%)	28 (87.5)
Any recent illness (previous 7 days) = No (%)	11 (34.4)
Diarrhoeal illness (previous 7 days) = No (%)	26 (81.2)
Fever (previous 7 days) = No (%)	17 (53.1)
Cough (previous 7 days) = No (%)	20 (62.5)

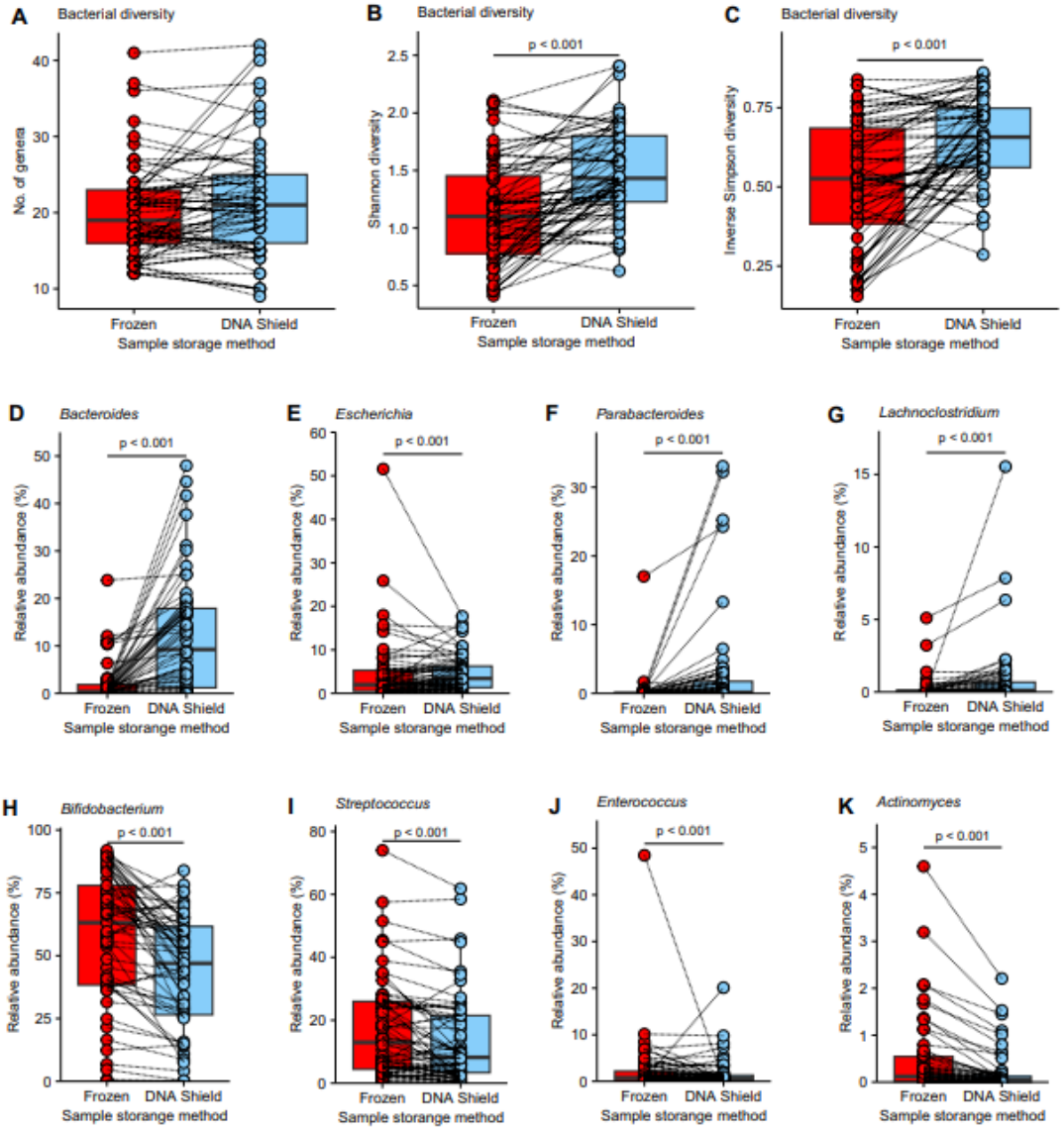


Fig. 1 | Differences in genus abundance between paired frozen samples and DNAShield stored samples. A The number of genera detected in each sample. **B** The Shannon diversity index of each sample. **C** The Inverse Simpson diversity index of each sample. Relative abundance between paired samples after different storage conditions of **D** *Bacteroides*, **E** *Escherichia*, **F** *Parabacteroides*, **G** *Lachnoclostridium*, **H** *Bifidobacterium*, **I** *Streptococcus*, **J** *Enterococcus* and **K** *Actinomyces*. Box plots show median and interquartile range with circles showing individual samples. Dotted lines link together paired Frozen and DNAShield samples from the same infant sample. Statistical significance between groups was tested using a Wilcoxon signed rank test with correction for multiple testing amongst bacteria genera using the false discovery rate method.

infants were collected (a total of 96 samples), with the two halves of each sample split between the two storage conditions. Both frozen and DNAShield paired infant gut microbiota samples underwent 16S rRNA amplicon sequencing to analyse their bacterial composition.

Out of the 96 paired samples, a total of 94 were successfully sequenced when stored in DNAShield, while only 70 were successfully sequenced from frozen samples (Supplementary Data 1). There were 69 pairs of samples with successful 16S rRNA amplicon sequencing from both the frozen and DNAShield portion of the same faecal sample. The resulting bacterial genus composition was compared between these 69 pairs of samples (Fig. 1). While the total number of bacterial genera detected was similar between frozen and DNAShield samples (Fig. 1A), microbiota diversity calculated using the Shannon index and Inverse Simpson index was higher in DNAShield samples than paired frozen samples (Fig. 1B, C). *Bacteroides, Escherichia, Parabacteroides, and Lachnoclostridium* were the four bacterial genera with the largest median difference in relative abundance between storage conditions that were higher in DNAShield samples compared to frozen samples (Fig. 1D–G; Supplementary Data 1). *Bacteroides* showed the

largest difference in favour of DNAShield storage, with a median relative abundance of 0.4% in frozen samples and 8.6% in DNAShield samples (Fig. 1D; Supplementary Data 1). The relative abundance of *Bifidobacterium*, *Streptococcus*, *Enterococcus*, and *Actinomyces* were the genera with the largest difference and significantly higher abundance in favour of frozen samples (Fig. 1H–K; Supplementary Data 1). *Bifidobacterium* comprised a median relative abundance of 63% in frozen samples and 46% in DNAShield samples (Fig. 1H; Supplementary Data 1). Further analysis of bacterial microbiota composition in this study was performed using samples stored in DNAShield due to the larger number of samples with successful sequencing results and the higher abundance of individual bacterial genera surviving in DNAShield stored samples.

These results indicate that sample storage conditions can selectively alter the abundance of different genera in the analysis of bacterial composition with a significantly higher abundance of *Bacteroides* and *Parabacteroides* remaining when samples were stored in DNA preservation solution. This shows that DNA preservation solution can preserve genus abundance in final composition analysis even when samples are initially transported at ambient temperatures and then stored refrigerated and not frozen.

Bacterial genus composition of the infant gut microbiota is influenced by breastfeeding and diarrhoeal illness

The gut microbiota of infants was dominated by *Bifidobacterium*, which comprised a median relative abundance of 48.7% (Fig. 2A). In addition, other genera, including *Bacteroides, Blautia, Enterococcus, Erysipelatoclostridium, Escherichia, Lactobacillus, Megamonas, Parabacteroides*, and *Streptococcus*, comprised the top 10 most abundant genera, and formed the majority of the gut microbiota in these infants (Fig. 2A).

An exploratory analysis was carried out using metadata available for the infants from which samples were collected. This included cross-sectional data collected at the start of the study before the collection of the faecal sample at time point 1 and additional information collected at each sample collection visit (Supplementary Data 2). The exploratory analyses used non-metric multidimensional scaling (NMDS) and sample beta-diversity to cluster the samples by the similarity of their overall microbiota composition at time point 1. Associations between the infant metadata and the beta-diversity of the gut microbiota were analysed using Envfit, and significant associations were illustrated on the NMDS plot (Fig. 2B). Only breast-feeding status was significantly associated with microbiota beta-diversity with clear separation in clusters between infants breast-fed and those infants not breast-fed (Fig. 2B).

To identify associations between the relative abundance of individual genera and infant metadata, an analysis using the MaAsLin2 package was carried out that included the samples from all three-time points, with the time point variable included as a random effect (Fig. 2C; Supplementary Data 3). When background metadata variables were analysed individually as a fixed effect, both breastfeeding status and recent diarrhoeal illness within the previous seven days showed significant associations with genus abundance, which was maintained when both variables were included together as a multivariable analysis (Fig. 2C). Genera with a differential abundance that was significantly associated with breast-feeding and/or diarrhoeal illness are shown as a heat map with positive and negative symbols indicating the direction of association (Fig. 2C). The genera positively associated with breast-feeding were *Bifidobacterium* and *Streptococcus* (Fig. 2C and D), whilst those with the strongest negative associations were *Klebsiella* and *Clostridium* (Fig. 2C and E). Breast-feeding status remained unchanged for individual infants across the three time points. At time point 1 the mean age of breast-fed infants was 207 days and non-breast fed infants 214 days. Including age as a variable in the analysis did not alter the associations with breastfeeding. Other measures of infant diet, the use of a feeding bottle or complementary feeding, were not associated with genus abundance. MaAsLin2 analysis also showed that recent diarrhoeal illness within the previous 7 days was associated with a lower relative abundance of *Bacteroides* (Fig. 2F) and a higher relative abundance of *Clostridium*, *Klebsiella* and *Shigella* (Fig. 2G, H). *Shigella* was detected in 78% of samples, *Klebsiella* in 82% and *Clostridium* in 65% indicating that these infants have a high prevalence of these potentially pathogenic genera in their gut microbiota (Supplementary Data 4). Breast-feeding was the only variable associated with differences in microbial diversity with the total number of the bacterial genus (Fig. 2I) and the Shannon diversity was lower in breast-fed infants than non-breast-fed infants (Fig. 2J).

Only three infants included in the analysis were classified as having stunted growth with a height-for-age *z*-score of less than 2 and this was not associated with microbiota composition. Tested as continuous measures, length-for-age *z*-score, underweight *z*-score, and weight-for-length *z*-score were also not associated with microbiota composition. However, this number of infants was too small for robust testing, preventing any strong conclusions being drawn. Measures of household hygiene facilities and the total number of people in the household were not associated with microbiota composition. Of particular note was the lack of an association between genus composition and recent antibiotic use within the previous 7 days (Supplementary Data 3).

These results indicate that breast-feeding was the most influential factor shaping the gut microbiota composition in favour of increased *Bifidobacterium* abundance. These results also show that recent diarrhoeal illness was associated with higher relative abundance of the potentially pathogenic genera *Clostridium*, *Klebsiella* and *Shigella*.

Bacterial species composition of the infant gut microbiota is influenced by breast-feeding status

The same analysis carried out on genus-level composition was carried out on the data at the species level with the limitation that species classifications using 16S rRNA amplicon data are less robust than genus (Supplementary Data 5). The gut microbiota at the species level was dominated by a single *Bifidobacterium* species classified as *Bifidobacterium longum* (Fig. 3A). Other species within the top 10 most abundant species at Time Point 1 included three other species of *Bifidobacterium* identified as *B. breve*, *B. bifidum*, and *B. dentium*, while *Escherichia coli*, *Enterococcus faecium*, *Klebsiella pneumoniae*, *Bacteroides fragilis* and an unclassified *Streptococcus* (Fig. 3A).

Repeating previous exploratory MaAsLin2 analysis of the species level 16S rRNA amplicon sequencing data identified the same association of breast-feeding status on microbiota composition with a stronger effect with species level data (Fig. 3C; Supplementary Data 6). Species with a relative abundance that was statistically significantly associated with breast-feeding and recent diarrhoeal illness are shown as a heat map with positive and negative symbols indicating the direction of association (Fig. 2C). *B. longum* showed the strongest positive association with breast-feeding, with consistently higher median relative abundance in breast-fed infants compared to those not breast-fed (Fig. 3D) while *Klebsiella pneumoniae* showed the strongest negative association with breast-feeding (Fig. 3E). The species *Shigella flexneri* (Fig. 3F) and *Clostridium perfringens* (Fig. 3G) were higher in samples taken after recent diarrhoeal illness. Breast-feeding was the only variable associated with differences in microbial diversity at the species level of classification. The total number of bacterial species (Fig. 3H) and the Shannon diversity was lower in breast-fed infants than in non-breast-fed infants (Fig. 3I). *S. flexneri* was detected in 49% of samples and *C. perfringens* detected in 24% of samples indicating that these infants have a high prevalence of these potentially pathogenic genera in their gut microbiota (Supplementary Data 7).

As described at the genus level, classification as stunted with a height-for-age *z*-score of <2 was not associated with gut microbiota species composition. Similarly, length-for-age *z*-score, underweight *z*-score, and weight-for-length z-score as continuous measures were also not associated with species composition. Hygiene measures of household toilet facilities and the total number of people in the household were again not associated with species composition. Also noted was the lack of an association between the gut microbiota composition and recent antibiotic use (Supplementary Data 6).

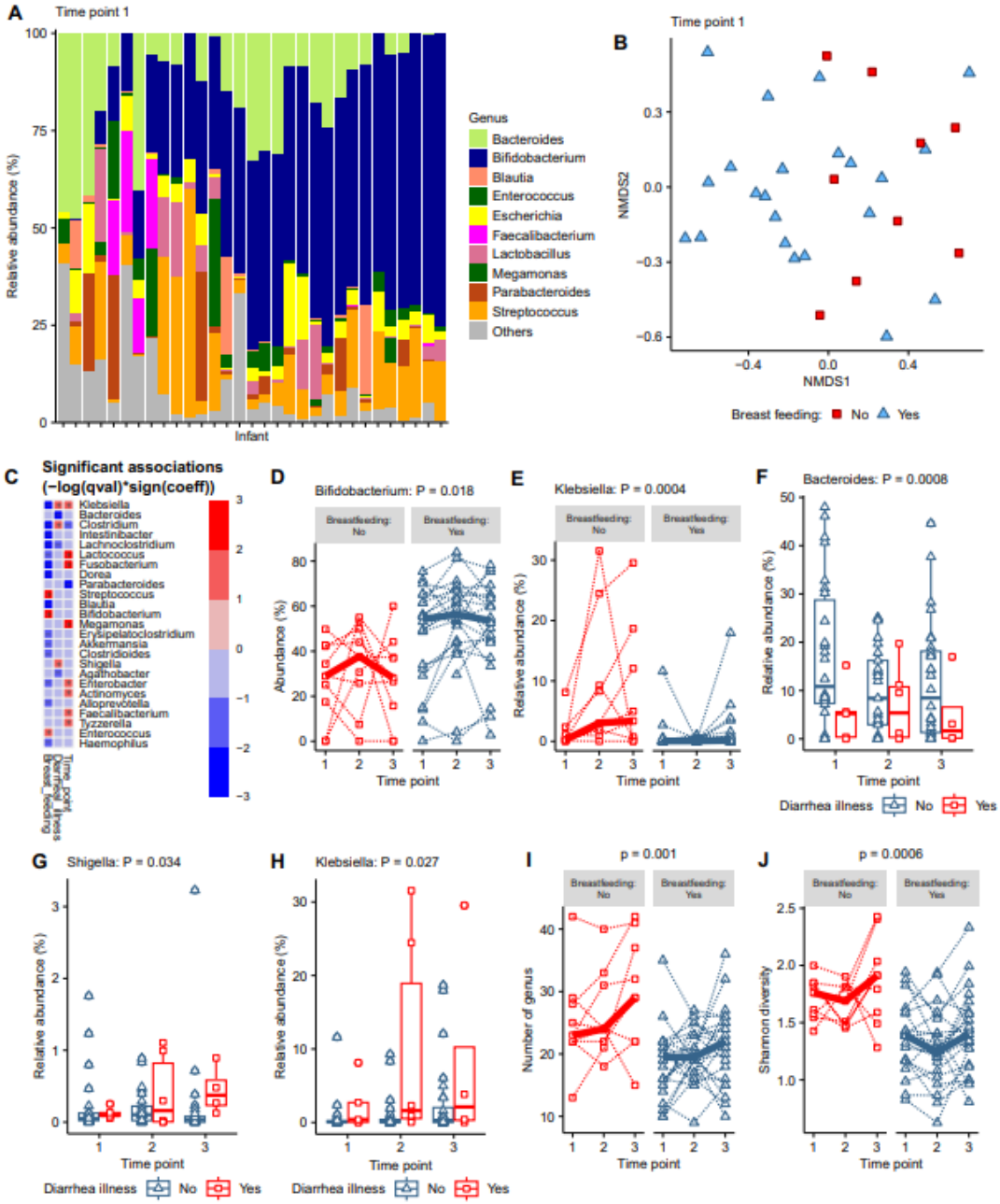


Fig. 2 | Genus composition and associations of genus composition with breast-feeding status and diarrhoeal illness. A The 10 most abundant bacterial genera across all samples are ordered on the *x*-axis by *Bifidobacterium* abundance in each infant. **B** Non-metric multidimensional scaling (NMDS) plot representing beta-diversity of infant genus microbiota at time point 1 clustered by similarity and coloured by breastfeeding status with arrows showing top 10 bacterial genera diving separation of samples. **C** Heat map showing significant associations determined by MaAsLin2 analysis between genus and breastfeeding, diarrhoeal illness and time point. **D** The relative abundance of the genus *Bifidobacterium* by breastfeeding status. **E** The relative abundance of the genus *Klebsiella* by breastfeeding status. **F** The relative abundance of the genus *Bacteroides* by diarrhoeal illness. **G** The relative abundance of the genus *Shigella* by diarrhoeal illness within the previous 7 days. **H** The relative abundance of the genus *Klebsiella* by diarrhoeal illness within the previous 7 days. **I** The number of genera by breast-feeding status. **J** The Shannon diversity index by breast-feeding status; Points on plots show individual infant data and dotted lines on spaghetti plots link individual infants across time points with the bold line showing the group median.

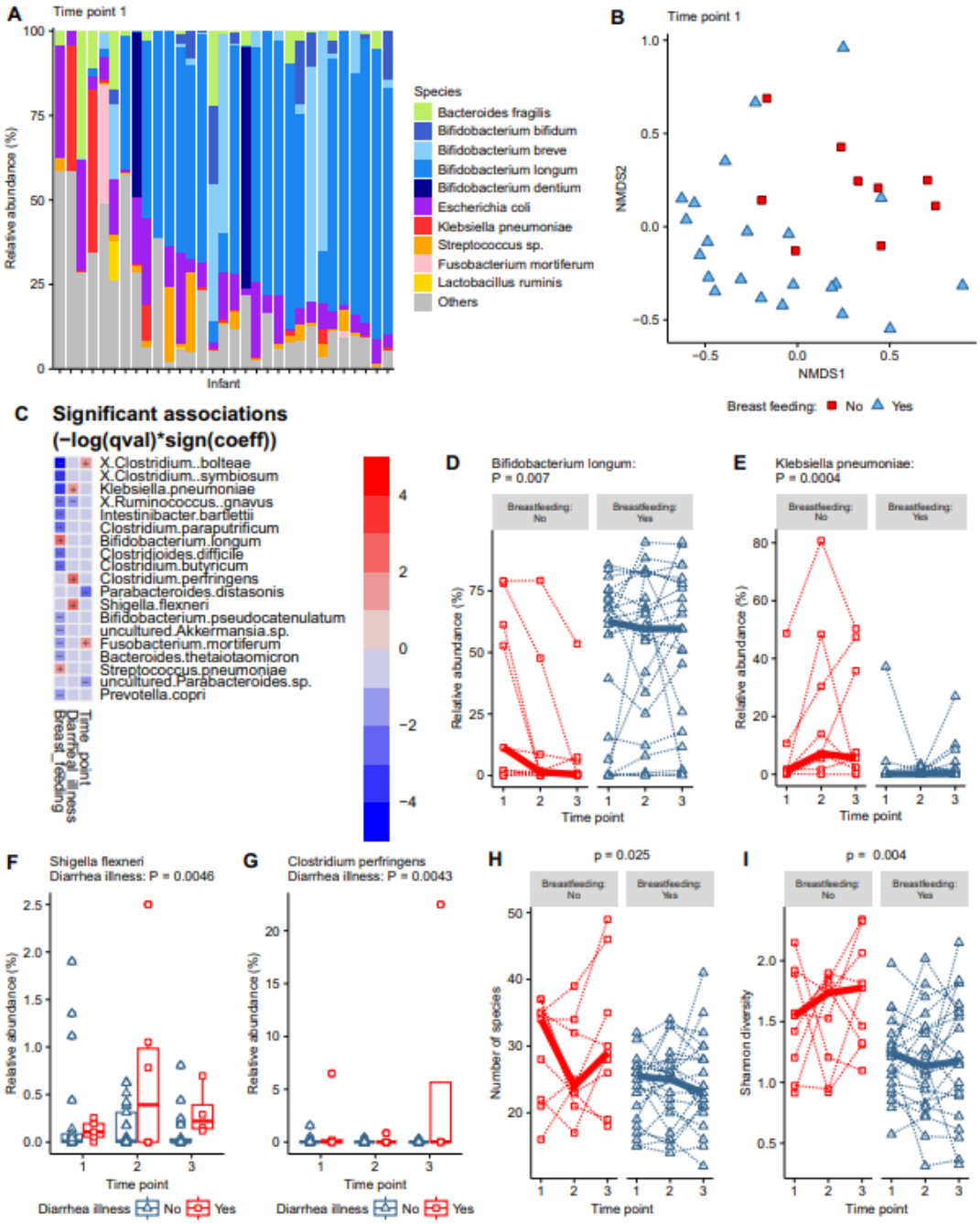


Fig. 3 | Species composition and association of species composition with breastfeeding and diarrhoeal illness. A The top 10 most abundant bacterial species across all samples at time point 1 ordered by total *Bifidobacterium* species abundance. **B** Non-metric multidimensional scaling (NMDS) plot representing beta-diversity of infant species microbiota at time point 1 clustered by similarity and coloured by breastfeeding status with arrows showing top five bacterial genus diving separation of samples. **C** Heat map showing significant associations determined by MaAsLin2 analysis between species and breast feeding, recent diarrhoeal illness and time point. **D** The relative abundance of the *Bifidobacterium longum* by breastfeeding status. **E** The relative abundance of *Klebsiella pneumoniae* by breast-feeding status. **F** The relative abundance of *Shigella flexneri* by recent diarrhoeal illness within the previous 7 days. **G** The relative abundance of *Clostridium perfringens* by recent diarrhoeal illness within the previous 7 days. **H** The number of species by breast-feeding status. **I** The Shannon diversity index by breast-feeding status, Points on plots show individual infant data and dotted lines on spaghetti plots link individual infants across time points with the bold line showing the group median.

These species-level results support the finding that the breast-feeding status was a major influence on the gut microbiota composition and associated with the abundance of *B. longum*. Recent diarrhoeal illness was associated with higher abundance bacterial species including *S. flexneri* and *C. perfringens* that are known to be potential causes of diarrhoea.

Infants harbour bacterial strains that encode diverse human milk oligosaccharide and antimicrobial resistance genes

The frozen faecal samples that were not stored in DNA preservation solution were used to culture and isolated individual colonies for whole genome sequencing to gain further insight into the species of bacteria in the gut microbiota of these infants. Genomic analysis was carried out on the sequenced isolates to identify the species (and strains) and for the presence of virulence traits, antimicrobial resistance genes and HMO gene clusters. From the bacterial isolates genome sequenced, 2 were identified as *Bacteroides*, 6 as *Bifidobacterium*, 1 as *Ruminococcus*, 3 as *Enterococcus*, 21 as *Clostridium*, 3 as *Clostridioides*, and 2 as *Paraclostridium* (Fig. 4A).

Two *C. perfringens* strains were isolated and genome-sequenced. Both contained five toxin-producing genes and six colonisation factor genes that relate to the ability to cause infection and disease in the gut (Fig. 4B). This confirms the presence of *C. perfringens* that were identified using 16S rRNA amplicon sequencing and the presence of toxin and colonisation factor genes supports a potential role in its association with cases of infant diarrhoea.

Bifidobacterium isolates comprised three species, including *B. pseudocatenulatum*, *B. breve*, and *B. longum*. Comparing the four genome-sequenced bacterial isolates identified as *B. longum* using average nucleotide identity with a 98% cut-off, three were identified as *B. longum* subsp. *longum* and one as *B. longum* subsp. *suis* (Supplementary Data 8; Supplementary Fig. 1). Both *B. breve* and *B. longum* isolates contained complete HMO gene clusters, with *B. pseudocatenulatum* containing partial HMO gene clusters (Fig. 4C). The inclusion of *B. pseudocatenulatum* genomes isolated from infants in neighbouring Vietnam showed the same pattern of partial HMO gene clusters the Cambodian *B. pseudocatenulatum* isolate (Fig. 4C). This indicates the ability of the *B. breve* and *B. longum* present to digest the oligosaccharides present in human breast milk, with only *B. longum* showing a higher relative abundance in breast-fed infants compared to non-breast-fed (Fig. 3C, D).

Antibiotic use was not associated with compositional differences in the infant gut microbiota as shown using an NMDS plot showing the samples from all time points combined (Fig. 5A). Additionally, the relative abundance of *Bifidobacterium* did not show any difference in samples from infants receiving antibiotics (Fig. 5B). Genomic analysis of whole genome sequenced isolates showed the presence of antimicrobial resistance genes in the majority of strains isolated (Fig. 5C). Additional infant isolates of B. *pseudocatenulatum* originating from Vietnam and *E. faecium*, *Enterococcus faecalis*, and *C. perfringens* isolates originating from the United Kingdom some similar patterns of antibiotic resistance genes (Fig. 5C). Of particular note was the presence of the antimicrobial resistance genes *erm*(X), *tet*(O), *tet*(O/23/O), *tet*(Q), and *tet*(W) in the genomes of isolated *Bifidobacterium* species (Fig. 5C), which provide resistance to erythromycin (*erm*) and tetracycline (*tet*) respectively. Amongst this group of infants, the most commonly received antibiotic was amoxycillin, with individual instances of cefixime, and cefadroxile also received (Supplementary Data 2).

Faecal cytokines profiles are associated with breastfeeding status

The early-life gut microbiota is known to play a key role in shaping gut and mucosal immune responses. A multi-plex assay, including a panel of 30 cytokines, was used to determine concentrations in the infant faecal samples (Supplementary Data 9).

A linear mixed model incorporating infant metadata variables as a fixed effect and the infant as a random effect was used as an exploratory analysis of associations between each cytokine and metadata variables in all samples across the three time points (Supplementary Data 10). The most robust and consistent association was identified between two cytokines and breast-feeding (Fig. 6A–B). The cytokines Interleukin-1-alpha (IL-1α) (Fig. 6A) and Vascular endothelial growth factor (VEGF) (Fig. 6B) were at significantly higher concentrations in breast-fed infants. Both cytokines showed substantial variation in concentration across time points within samples from the same infant. IL-1α (Fig. 6A) and VEGF (Fig. 6B) were also the two cytokines with the highest concentrations in the faecal samples across all infants.

To identify whether differences in major faecal cytokines by breast-feeding status were linked to differences seen in the microbiota composition a correlation analysis was carried out between faecal cytokine concentrations and the relative abundance of the 5 most abundant genus (Fig. 6C) and species (Supplementary Fig. 2) in the microbiota. Weakly consistent positive associations were seen between VEGF and *Lactobacillus* (Fig. 6D) and *Streptococcus* (Fig. 6E) across the three time points, significant at two time points after correction for multiple testing. There were no consistent correlations between the concentration of other cytokines and the relative abundance of individual bacterial genus or species, including *Bifidobacterium*.

Discussion

This study represents the first time that the infant gut microbiota composition of a cohort of rural Cambodian infants has been analysed in detail. We show that a shortage of samples under rural conditions can radically reduce the abundance of certain genera within the samples. Infants had a gut microbiota dominated by the genus *Bifidobacterium*, shaped to a large degree by breastfeeding, which also impacts faecal cytokine profiles. Previous gut microbiota research in Cambodia has been limited to an investigation into antimicrobial resistance genes in Enterobacterales[6], the epidemiology of antibiotic resistance in *E. coli* and *K. pneumoniae*[7], and the effects of iron supplementation on individual genus abundance in women[8]. This present study goes further and examines in-depth genus and selected species and strain composition in Cambodian infants.

Our study results represent an unusual example of different sample storage conditions, as these samples were stored without immediate freezing. Overall, the effect of storage conditions on the composition of the microbiota was small, but selectively affected individual genera with a higher relative abundance of *Bacteroides* in particular was seen in samples stored in DNA preservation solution and a lower abundance of common species, including *Bifidobacterium* due to the proportional nature of the microbiota data. *Bacteroides* represents an important genus in the infant gut microbiota, transferred from the mother during delivery and has been found to be higher in vaginally delivered infants[17], with some *Bacteroides* species having the ability to consume components of breast milk[18]. The relatively high proportion of *Bacteroides* in these infants and their almost disappearance from samples without DNA preservation solution shows the importance of a storage method that preserves this significant component of the infant microbiota for analysis.

Previous research has investigated the effects of storage conditions on microbiota analysis and reported alterations in *Bacteroides* abundance, although this has focused on storage conditions while frozen or the number of freeze–thaw cycles[19,20]. Room temperature storage of faecal samples for up to 72 h has been reported to reduce genus, including *Bacteroides*, in adult faecal samples over time[21]. While defrosting frozen samples for 3 h before analysis reduced *Bacteroides* abundance by half[22]. The duration of cool but unfrozen temperatures that samples in the present study were exposed to appear to have resulted in the particular low of *Bacteroides*. Other novel sample storage methods have been used for samples collected in difficult conditions. One study comparing rainforest hunter-gatherers and agriculturalist Bantu farmers in the Central African Republic used RNAlater or ethanol to preserve faecal samples for microbiota and metabolomic analysis, respectively[23]. Another unusually difficult situation collecting faecal samples from wild mountain gorillas and chimpanzees in Uganda found similar DNA preservation resulting from faecal storage in either RNAlater solution or dried using silica gel beads, although no microbiota analysis was carried

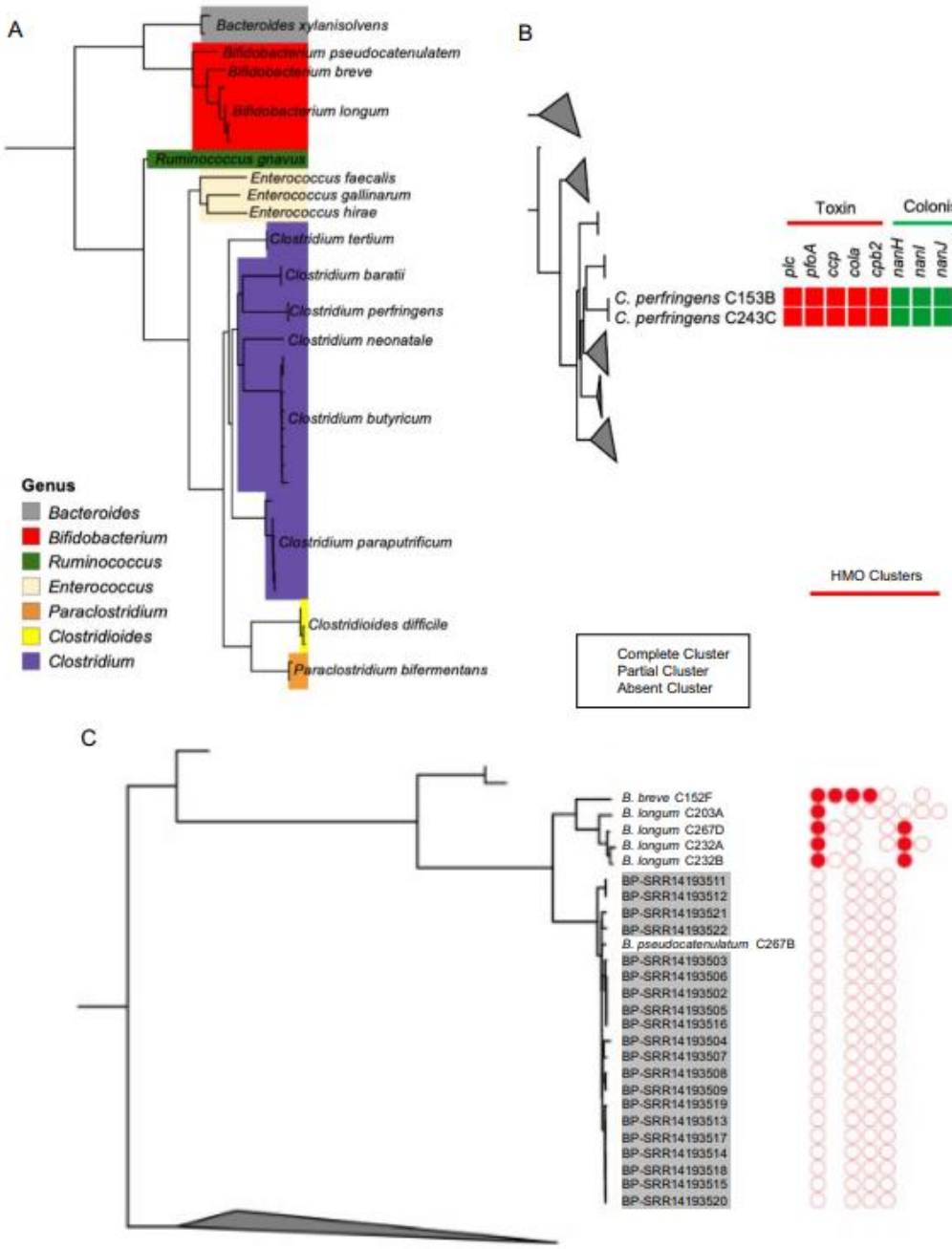


Fig. 4 | Genomic analysis of isolates for species determination, virulence genes, and human milk oligosaccharide (HMO) gene clusters. A Tree showing species and relatedness of whole genome-sequenced isolates from Cambodian infants. **B** Presence of toxin genes and colonisation factors in *Clostridium perfringens* isolates. **C** Presence of HMO gene clusters in *Bifidobacterium* isolates. Grey-shaded sample names indicate samples included for comparison from previous studies.

out in this case[24]. However, the unpreserved samples in the present study were also important for culturing and isolating live bacteria for genomic sequencing, highlighting the importance of having samples stored without preservation when possible. There is an opportunity for future research to directly compare the effects of a greater range of storage conditions on downstream microbiota analysis to determine optimal methods for the collection of samples in suboptimal conditions.

The overall composition of the gut microbiota in these Cambodian infants was predominantly comprised of the genus *Bifidobacterium*. Comparative research indicates that a high abundance of *Bifidobacterium* is

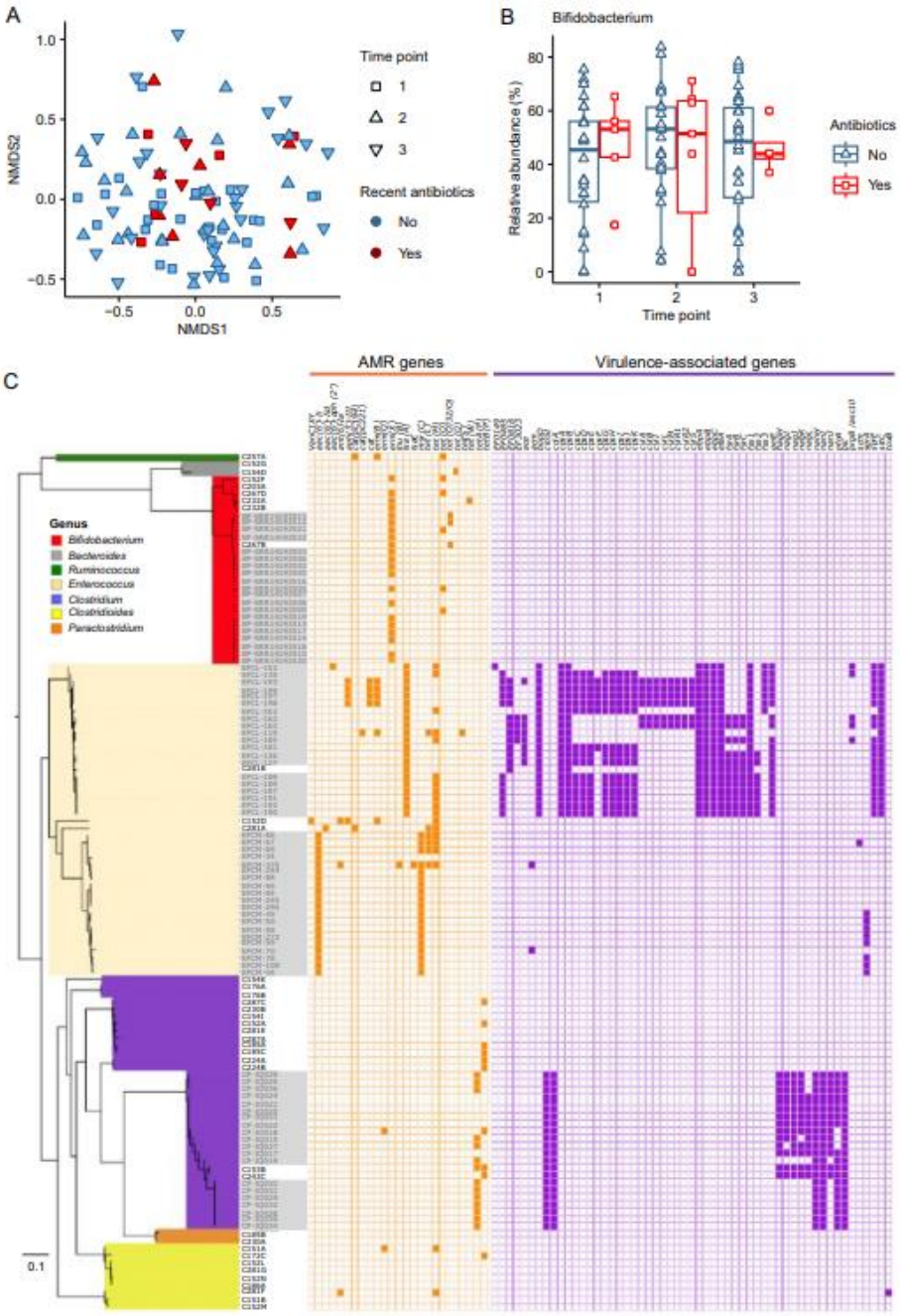


Fig. 5 | Compositional differences and *Bifidobacterium* abundance by recent antibiotic use and genomic analysis of isolates for antimicrobial resistance genes. **A** NMDS plot showing clustered faecal samples from all three time points coloured by recent antibiotic use. **B** *Bifidobacterium* genus abundance by antibiotic use. **C** Antimicrobial resistance genes and virulence-associated genes identified in the isolate genomes, plus public genomes. Grey-shaded sample names indicate samples included for comparison from previous studies.

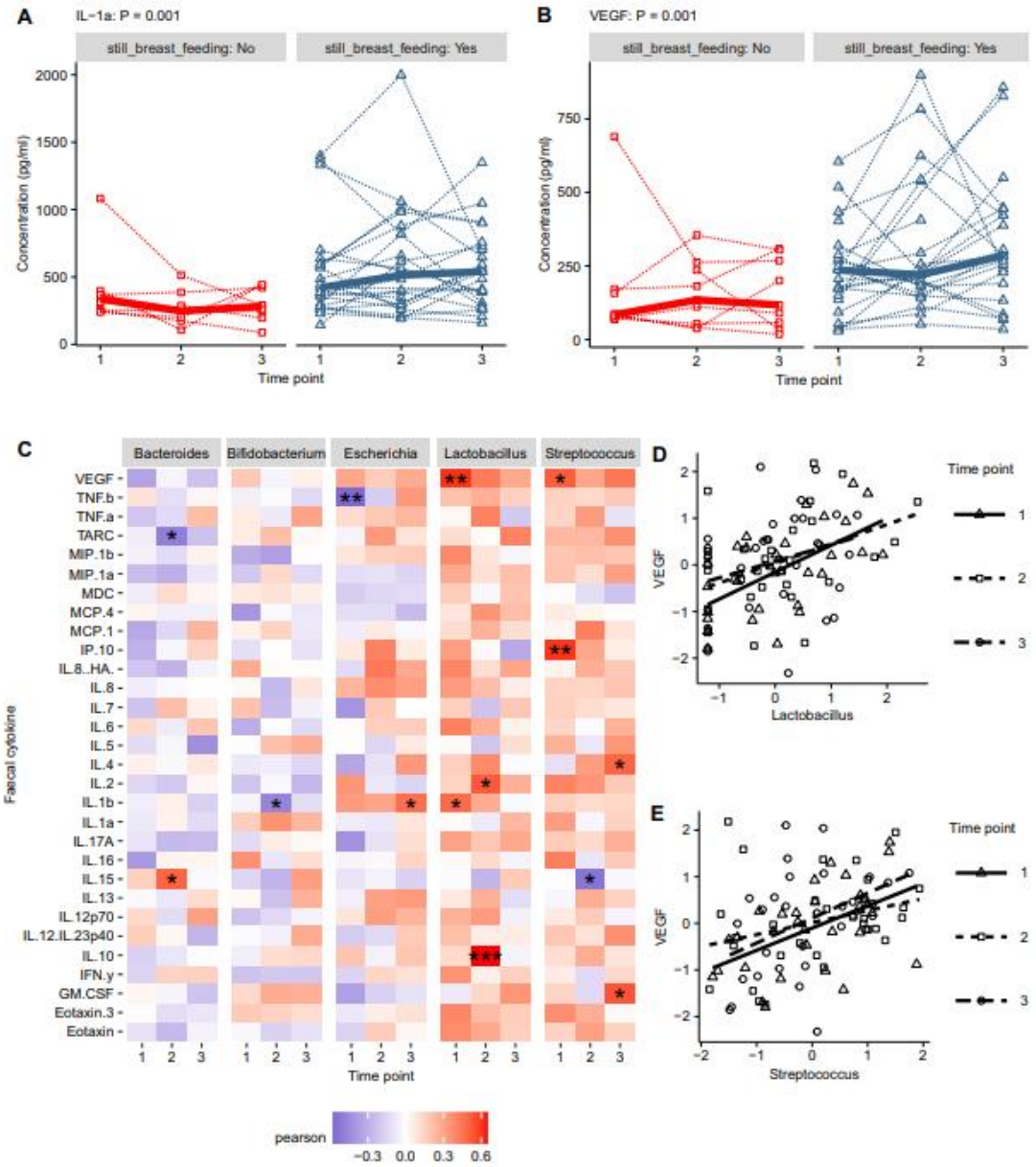


Fig. 6 | Infant faecal cytokines concentration and associations with breast feeding status. **A** Individual infant IL-1α concentration by breastfeeding status across time points. **B** Individual infant VEGF concentration by breastfeeding status across time points. **C** Pearson correlations between faecal cytokine concentrations and five highest abundance genus with variables in both datasets individually normalised using the bestNormalize package. **D** Scatter plot of VEGF and *Lactobacillus* correlations labelled for each time point. **E** Scatter plot of VEGF vs. *Streptococcus* correlations labelled for each time point.

typical among infants in rural areas of less industrialised countries[25] whereas it is generally lower in more westernised countries, such as the United States[26], and also reported to be low in younger infants from China[27]. Other abundantly present genera in Cambodian infants were similar to those observed in other regional populations. For instance, infants from Bangladesh of a similar age exhibited the presence of *Lactobacillus*, *Streptococcus*, *Escherichia*, and *Klebsiella*[28]. One genus notably absent from this cohort of Cambodian infants is *Prevotella* (now reclassified under new genus names including *Segatella*), which has been observed in Bangladeshi infants at a similar age[28], and found to be a dominant member of the gut microbiota in rural West Africa, yet rare in Italian children[29]. More recent studies have confirmed the high prevalence of *Segatella* in children from low-income countries such as The Gambia, Kenya, and Mali[30], and specifically in Gambian infants from an early age similar to the Cambodian infants in the current study[31]. While *Segatella* abundance was lower in Bangladeshi children, it was still detectable in children of a similar age to this cohort of

Cambodian infants[30]. In contrast, Cambodian infants exhibited the presence of *Bacteroides*, a genus typically associated with the gut microbiota of European and US inhabitants[29]. The dietary differences that may account for these discrepancies cannot be determined from this study alone, underscoring the need for further investigation into underrepresented populations. Whether the presence of *Bacteroides* and the absence of *Segatella* persist into adulthood remains unknown, as the composition of the adult gut microbiota in this region has yet to be explored.

Bifidobacterium was present at lower abundance in non-breast-fed infants and not breast-feeding was associated with a higher abundance of potentially pathogenic bacterial genera like *Klebsiella* and *Clostridium*. Although the number of infants not breast-fed was too low to draw strong conclusions, previous research has indicated that increased *Bifidobacterium* in the preterm infant gut can help reduce the proportions of potential pathogens[32]. This is linked to the ability of *Bifidobacterium* to metabolise HMOs in breast milk and reduce gut pH to make it a more inhospitable environment for the overgrowth of opportunistic pathogens. *B. longum* was the dominant species in this cohort of infants, and *B. longum* was the only *Bifidobacterium* species with higher abundance in breast-fed infants. The relatively low abundance of *B. pseudocatenulatum* in these infants may be due to their lack of complete HMO gene clusters. The matching pattern of partial HMO gene clusters *B. pseudocatenulatum* from Cambodia and Vietnam suggests a similar dietary niche for the species in infants in the region. The four *B. longum* isolates genome-sequenced all contained one or two HMO gene clusters indicating that they were potentially capable of metabolising the sugars present in breast milk. Interestingly, three of the *B. longum* isolates were confirmed to be *B. longum* subsp. *longum*, which do not typically dominate the gut microbiota of infants in high-income countries nor other LMICs with high rates of breastfeeding. It was recently reported that *B. longum* subsp. *infantis* was the dominant *Bifidobacterium* in 1–2-month-old infants from the Gambia and Bangladesh with *B. longum* subsp. *longum* only forming a minor proportion of the microbiota[15]. The isolation of *B. longum* subsp. *longum* from Cambodian samples may be due to sampling at an older age of around 7 months, with recent research from Bangladesh indicating that a transition occurs from *B. longum* subsp. *infantis* to *B. longum* subsp. *longum* with increasing infant age[28]. One *B. longum* isolate was identified as most closely matching *B. longum* subsp. *suis*, a subspecies originally isolated from pigs[33] and found to be a dominant member of the piglet gut microbiota[34]. *Bifidobacterium* isolates corresponding to *B. longum* subsp. *suis* were recently identified in Bangladeshi infants[28], which suggests that this may be an under-recognised member of the infant gut microbiota. From only a single isolate we cannot determine the proportion of *B. longum* subsp. *suis* within Cambodian samples, and this could represent the transient colonisation of this bacterium from close proximity to pigs or that this is a resident of the gut microbiota in this region. The presence of an HMO gene cluster capable of metabolising breast milk sugars indicates that this *B. longum* subsp. *suis* isolate may be capable of metabolising breast milk HMOs. The dominant abundance of the species *B. longum* in breast-fed infants underline the central effect of diet on the gut microbiota composition of these infants, overwhelming any other environmental factor in this small cohort.

Previous research has identified associations between restricted infant growth and stunting due to malnutrition and differences in the infant gut microbiota[12,35]. We did not find any associations between the gut microbiota and infant's growth. This was potentially due to the small cohort of infants involved, with only three out of the thirty infants included showing stunted growth and the relatively young age of the infants starting at an average of 7 months of age. Previous research has looked at older children[35] or older infants, with a higher proportion of the infants suffering from stunting[12]. This indicates that in this small cohort of infants it is the presence or absence of breast milk having the largest effect, potentially overwhelming the ability to detect other environmental influences.

In Cambodia and many other LMICs, infants can be exposed to pathogens through the provision of contaminated drinking water in the first months of life (Poirot et al., 2018). Whilst we show that there are high proportions of beneficial bacteria in most samples, pathogenic taxa, including *Shigella*, *Escherichia*, and *Clostridium* genus were also detected in most infants. Although *S. flexneri* and *C. perfringens* abundance in the gut was associated with recent diarrhoea these bacterial species were also detectable in a large minority of infant samples indicating environmental exposure to these organisms and potential asymptomatic carriage.

Antibiotics can significantly alter the profile of early gut colonisation, resulting in reduced diversity[32,36] and susceptibility to other infections[3]. Notably, in Cambodia, there is widespread, unregulated and inappropriate antibiotic use, and these selection pressures may further facilitate the dissemination of AMR genes[7]. An estimated one-third of the Cambodian population is <18 years old, and this group may be a significant reservoir for the spread of antimicrobial-resistant organisms[7]. Indeed, we observed multiple antibiotic-resistance genes in potentially pathogenic species, including those associated with diarrhoeal episodes. The pattern of antimicrobial resistance genes in *C. perfringens*, *E. faecium*, and *E. faecalis*, showed some similarities between Cambodian infant isolates and United Kingdom infant isolates and between *B. pseudocatenulatum* from Cambodian and Vietnamese infant isolates. This suggests geographically distant infant-derived bacterial genomes can share similar antibiotic-resistance genes. As these samples were from healthy infants, this potentially indicates a pattern of normal resistance in commensals in the gut. We also observed the presence of antimicrobial resistance genes in *Bifidobacterium*. Although this may seem unexpected, as *Bifidobacterium* is very rarely reported to be antibiotic resistant, a recent report from the neighbouring country of Vietnam indicates that *Bifidobacterium* isolated from healthy infants harbour antibiotic resistance genes[37]. We also found that Cambodian *Bifidobacterium* isolates contained resistance genes to the antibiotics erythromycin and tetracycline. This may link to the fact that this cohort of infants has previously been reported to be exposed to frequent and unregulated use of antibiotics[30], thus potentially driving AMR dynamics in beneficial members of the infant gut microbiota. However, although amoxicillin was the most commonly used antibiotic in this cohort, we did not observe any beta-lactam resistance genes in the isolates. Previous studies have indicated that while it appears, there is high and unregulated use of antibiotics within the Cambodian population—they may be ineffective due to inappropriate storage, thus reducing their effectiveness. However, the widespread presence of resistance genes in other isolates may point towards active selection through exposure to antibiotics. The apparent maintenance of *Bifidobacterium* abundance despite antibiotic treatment may point to the presence of other novel resistance genes in these *Bifidobacterium*, which, if identified, may have future potential for combining effective probiotic therapy with antibiotic treatment.

Consistent differences in the faecal concentrations of the cytokines IL-1α and VEGF point to the potential effects of breastfeeding on the infant's gut mucosal immune system. VEGF has been reported to be secreted by the mammary glands, and breast milk contains high concentrations of free VEGF that can bind to receptors on the surface of cells in the infant intestine[38–40]. Reduced VEGF in preterm infants is suspected to be involved in necrotising enterocolitis, which is associated with the overgrowth of opportunistic pathogens, including *C. perfringens*[41,42]. While the role of VEGF originating from breast milk in term infants remains unexplored, these growth factors could exert growth-promoting and protective effects on the infant gut[38]. While the role of IL-1α in the infant's gut has not been explored, it has also been reported to be present in breast milk[43]; however, whether this has any effects on the infant's gut is currently unknown. Although we are unable to demonstrate a breast milk origin from these results, the presence of higher concentrations of these three faecal cytokines in breast-fed infants does suggest that they may be derived from breast milk, with a potential to influence infant development that warrants further investigation.

While there was little consistent association across time points between the gut microbiota at the genus level and faecal cytokines, a weak positive correlation was noted between *Lactobacillus* and *Streptococcus* and the faecal cytokine VEGF. As *Streptococcus* abundance was higher in breast-fed

infants, this correlation is potentially linked to breastfeeding. As both microbiota composition and faecal cytokines are variable with time, it may not be possible to identify links between factors in a limited sample size of infants. Interestingly, despite being the genus most differentiated by breastfeeding status, there are no consistent correlations between *Bifidobacterium* IL-1α or VEGF. Indeed, many species and strains from this genus are known to positively influence the gut and systemic immune system—including *B. longum*, which is proposed to be in part mediated by metabolites produced as a result of HMO degradation, e.g. acetate[14,44]. However, further studies, including a larger number of infants, would need to be performed to explore microbial-immune interactions in more detail. In addition, the use of the frozen samples for faecal cytokine analysis may be limited as it is unknown how much storage conditions may impact the degradation of the cytokines tested.

The initial sample storage was pragmatic, as the samples were not collected with the intention of determining the effects of storage conditions on microbiota composition. A study specifically designed to investigate this question would include more detailed documentation of individual sample collection methods and compare various storage conditions. The available samples for microbiota analysis covered a short duration, with infants aged between an average of 201 to 235 days. These samples included infants who were either still breastfeeding or had stopped but did not cover periods before and after weaning. This allowed for a cross-sectional analysis of microbiota composition but limited the ability to assess how the infant microbiota developed over time. Future research could address this gap by collecting longitudinal samples from birth through childhood, providing a more comprehensive investigation of the maturation of the infant microbiota. Conducting such a study with infants in Cambodia would offer valuable insights and present an important opportunity for further research.

We show that this cohort of infants from rural Cambodia has a *Bifidobacterium*-dominated gut microbiota (including atypical species) that is primarily shaped by breastfeeding. This highlights the influence of breastfeeding in supporting a beneficial infant microbiota in rural conditions. As *Bacteroides* represents a common and important genus in the human gut the effects of sample storage are an important consideration when immediate freezing of samples is not possible. The lack of impact of antibiotic use may indicate either widespread antibiotic resistance or the use of ineffective antibiotics. The finding that faecal cytokines are shaped by breast-feeding and could potentially originate from breast milk is a finding with unknown but potentially interesting implications. This study adds significantly to the available data on the composition of infant gut microbiota communities in a previously neglected and difficult-to-sample geographical area of the world. Moreover, given that Cambodian profiles are somewhat different from other infant cohorts, including LMICs, studies such as these show the importance of looking at infants from underrepresented areas for defining what can constitute a normal infant microbiota.

Methods

Study design and ethics

Samples were collected from a sub-group of infants enroled in the NHAM cohort based in villages along the Mekong River in rural Kampong Cham province in Cambodia. This registered all infants born between April 2016 and March 2019 with the objective of investigating the causes of stunting. Full ethical approvals, including National Ethical Committee for Health Research in the Ministry of Health, Cambodia (016) and the Ethical Committee of the National Centre for Global Health and Medicine in Japan (NCGM-G-001870-01), and methods for the collection of detailed metadata for infants included in this sub-study are as previously described[9,10]. Weight and length were measured to the nearest 0.1 kg and 1 mm by trained field workers at each home visit using a weighing scale (Seca 877, Seca, Hamburg, Germany) and a length board (Seca 417, Seca).

Sample collection

Infant faecal samples were analysed from a subset of 32 infants out of 47 enroled in a detailed 3-month longitudinal follow-up study[9,10] collected between March and July 2017. Mothers/primary care-givers and their infants were visited by study field workers with three samples collected and analysed from each infant for a total of 96 individual samples. The mean age of infants was 201 days (Standard deviation: 32) at Time Point 1, 216 days (Standard deviation: 30) at Time Point 2, and 236 days (Standard deviation: 30) at Time Point 3. The trained field worker gave instructions on how to collect and equipped the caregivers with the material for the child's stool collection prior to routine home visits. Once stool sample collection by mothers or caregivers was completed, the sample was kept in a relatively cooler place in the house or in cooler bags with ice until the field worker visited to collect them. The majority of the samples were collected by the field worker within 10 h of the time of child defecation reported by mothers (maximum: 47 h and minimum: 15 min among all samples). Each sample collected was divided into two and half stored in DNAShield and stored unaltered at ambient temperatures until delivered to laboratory. The other half of the sample was stored in iceboxes with refrigerant and transported to a −20 °C freezer on the same day, and once delivered was stored at −80 °C. The first sample was collected after an initial visit where cross-sectional metadata was collected. Subsequent samples were collected at approximately two-week intervals, along with further individual visit metadata. For recent diarrhoeal illness parent or caregiver was asked at the initial visit whether the child infant experienced diarrhoea within the past 7 days and since the previous visit at each subsequent sample collection. Underweight *z*-scores (weight-for-age), weight-for-length z-scores, length-for-age *z*-score were calculated previously[9]. A handwashing score was generated as a composite score combining the frequency and consistency of handwashing with soap. Improved toilet facility was defined as having access to a toilet flushing to a septic tank or pit latrine compared to unimproved having only access to a pit latrine or no toilet facility. Household toilet facility was defined as having a toilet within the home flushing to a septic tank.

DNA extraction of preterm stool samples

FastDNA Spin Kit for Soil (MP) was used to extract DNA from preterm faeces following manufacturer instructions, with extended 3 min bead-beating. DNA concentration and quality were quantified using a Qubit® 2.0 fluorometer (Invitrogen).

16S rRNA amplicon sequencing of faecal samples

16S rRNA region (V1–V2) primers were used for library construction as described previously[32]. This set of primers allowed the amplification of one 16S rRNA gene sequencing library containing 96 different samples. PCR conditions used were: cycle of 94 °C 3 min and 25 cycles of 94 °C for 45 s, 55 °C for 15 s and 72 °C for 30 s. Sequencing of the 16S rRNA gene libraries was performed using the Illumina MiSeq platform with 300 bp paired-end reads.

16S rRNA amplicon data analysis

The raw sequence reads obtained was subjected to quality filtering using trim_galore v0.4.3. The resulting filtered sequence reads were then mapped against the 16S rRNA SILVA database (SILVA_132_SSURef_tax_silva) using BLASTn v2.2.25+ with a maximum e-value threshold of 10e-3. This mapping process was performed separately for both pairs of sequences[45]. After completing the BLASTn alignments, the output files were annotated using the paired-end protocol of MEGAN6, employing the default Lowest Common Ancestor (LCA) parameters[46,47].

Further analysis was conducted using R version 4.2.2 (2022-10-31 ucrt) and RStudio (2023.06.0+421). To determine the number of reads required to adequately represent the microbiota in each sample, rarefaction curves were generated using the vegan package in R. For the analysis at the genus level, the sample with the smallest number of reads contained 37,306 classified reads, which was deemed an acceptable minimum. For the species-level analysis, the sample with the second lowest number of reads contained 9317 reads, which was considered an acceptable minimum. The 16S rRNA gene sequence data were then normalised by subsampling using the rarefy function in the phyloseq package, resulting in an even depth of 37,306 reads

for the genus data and 9317 reads for the species data. One sample was excluded from the species data analysis as it only contained 3965 reads.

Anaerobic bacterial culturing and isolation from infant faecal samples

Anaerobic bacterial culturing and isolation were performed on 17 Cambodian infant faecal samples (C151, C152, C153, C154, C172, C176, C185, C186, C203, C224, C230, C232, C243, C257, C267, C281, and C287) following[48–50]. Sample preparation was performed under anaerobic conditions. YCFA media and PBS were pre-reduced in an anaerobic cabinet at 37 °C for a minimum of 12 h before culturing. Two samples, weighing 100 mg each were prepared from each faecal sample. The first sample was homogenised in reduced PBS, serially diluted, and plated directly onto the YCFA agar. The second sample was ethanol treated using 70% ethanol under ambient aerobic conditions for 1 h to kill vegetative cells; the sample was then washed three times with reduced PBS to remove the ethanol, homogenised in the reduced PBS, serially diluted and plated onto YCFA + sodium taurocholate agar to stimulate spore germination. All agar plates were placed in the anaerobic cabinet at 37 °C for 48 h. Single colonies were picked and streaked onto pre-reduced YCFA media to isolate and purify the colonies; YCFA media was pre-reduced in the anaerobic cabinet at 37 °C for a minimum of 12 h before isolating colonies. Colonies from the non-ethanol-treated samples were streaked onto the reduced YCFA agar plates, and colonies from the ethanol-treated samples were streaked onto the reduced YCFA + Sodium taurocholate agar plates. Plates were then left to incubate in the anaerobic cabinet at 37 °C for 48–72 h. This was carried out a second time to make sure the isolates were pure. Falcon tubes containing reduced YCFA broth were inoculated with a single colony from the YCFA and YCFA + sodium taurocholate agar plates and left to incubate in the anaerobic cabinet at 37 °C for 48hs, with bacterial stocks prepared in 1.5 ml cryovials with 30% glycerol YCFA broth, and stored at -80 °C. Additionally, two infant faecal samples C203 and C232 were processed for targeted *Bifidobacterium* isolation. Briefly, serially diluted aliquots of faecal homogenates (10^{-1}–10^{-6} in PBS) were plated onto MRS agar supplemented with 50 mg/l mupirocin and 0.5 g/l L-cysteine HCL and incubated in an anaerobic cabinet at 37 °C for 48 h. Colonies with desired morphology were randomly selected and re-streaked on MRS agar to purity. Pure cultures were prepared in 1.5 ml cryovials with 30% glycerol-reinforced clostridial medium (RCM) and stored at −80 °C.

Whole genome sequencing (WGS)

DNA extractions were performed using the MPbio Fast DNA Spin Kit for soil and bead beating (but with an extended 3 min bead beating time). DNA quantification was carried out using the Qubit 2.0 fluorometer. DNA from pure bacterial cultures was sequenced at Wellcome Trust Sanger Institute using 96-plex Illumina HiSeq 2500 platform to generate 125 bp paired-end reads as described previously[51].

Genomic analysis

Paired-end raw sequence reads (fastq) were quality-filtered (-q 20) via fastp v0.20.0[52] prior to de novo genome assembly using SPAdes v3.14.1[53] at default parameters. Next, contigs smaller than 500 bp were discarded in each genome assembly. All genome assemblies were subjected to contamination check via CheckM v1.1.3[54], contaminated (contamination > 5%) and/or incomplete (completeness < 90%) genome assemblies were excluded from further analyses ($n = 5$). All genomes were identified taxonomically using gtdb-tk v1.5.1[55], using ANI cut-off 95% against type strains of each species.

These high-quality genome assemblies ($n = 71$) were then dereplicated at ANI 99.9% (cut-off for strain level) using dRep v3.2.2[56] prior to generation of a distance tree using Mashtree v1.2.0[57,58] on default parameters. Additional children/infant-associated reference bacterial genomes ($n = 80$) from recent studies (retrieved from NCBI) were added for comparison purposes —*Enterococcus faecium*, *Enterococcus faecalis*, *Bifidobacterium pseudocatenulatum* and *C. perfringens* (Supp X)[37,41,58]. Genome reads (reference genomes) downloaded from SRA (Sequence Read Archive) were subjected to quality filtering (-q 20) by fastp v0.20.0 followed by genome assembly via unicycler v0.50.0 prior to further gene screening analysis. Mash-distance tree was annotated in iTOL v6.5.8[59].

These genomes were also subjected to sequence search analysis via ABRicate v1.0.1[60] based on ResFinder v4.0 database[61] (--minid=95 and --mincov=90) to determine antimicrobial resistance (AMR) profiles, and VFDB database[62] (--minid=90 and –mincov=90) to determine virulence profiles. Furthermore, HMO clusters were detected using an in-house HMO sequence database via ARBricate at –minid=70[63]. In addition, *C. perfringens* virulence profiles were predicted via database TOXIper v1.1 using ABRicate at –minid=90[64].

For *Bifidobacterium longum* subspecies analysis, *B. longum* strains were compared with four subspecies type strains, namely, subspecies *infantis*, subspecies *longum*, subspecies *suillum* and subspecies *suis*, to generate a core gene alignment via Panaroo v1.2.8[65], resulting in 1212 core genes, which core-gene alignment constructed by MAFFT[66] were SNP-extracted (via SNP-sites v2.3.3)[67] and subsequently used to generate a phylogenetic tree via IQ-TREE v2.0.5[68] with ultra-fast bootstrap replicates -B 1000 and automatic evolutionary model selection option -m TEST (best fit model was determined to be GTR+F+ASC+G4). Subspecies identification was performed using fastANI v1.3[69] at cut-off 98%.

Faecal cytokine measurement

Naïve (i.e. samples not stored in DNASheild, which would interfere with cytokine analysis) faecal samples were homogenised with PBS using a FastPrep® Bead Beater (4.0 m/s, 3 min), centrifuged (14,000 rpm, 15 min) and 25 µl of supernatant was used for the assay. Samples were analysed using MULTI-SPOT™ plates, MESO Quickplex SQ120 and discovery workbench software according to the manufacturer's protocol. Pre-coated immunoassays V-PLEX Proinflammatory Panel 1 (human) and V-PLEX Cytokine Panel 1 (human) were used to detect a set of 20 different cytokines: IFNγ, IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, IL-12p70, IL-13, TNFα, GMCSF, IL-1α, IL-5, 273 IL-7, IL-12p40, IL-15, IL-16, IL-17A, TNFβ, and VEGF-A.

Statistical analysis

Statistical analysis was conducted using R version 4.2.2 (2022-10-31 ucrt) and RStudio (2023.06.0+421). To examine differences in the relative abundance of genera among the 69 infant faecal samples, which had 16S rRNA sequence data obtained from both samples stored in DNAShield solution and Frozen, a Wilcoxon rank sum test was performed in R. This test was chosen due to the non-parametric nature of the majority of the genus abundance data. To account for multiple testing, a Benjamini and Hochberg false discovery rate (FDR) method was applied, and a p-value of <0.05 was considered statistically significant for adjusted p-values.

The overall composition of the 16S rRNA gene sequence data was analysed using NMDS (Non-metric multidimensional scaling). The NMDS plots were generated in R Studio using the vegan package version 2.6-4, employing a Bray-Curtis dissimilarity calculation. To determine the statistical significance of differences in microbial community structure based on infant metadata variables, permutational MANOVA was conducted using the Adonis function from the vegan R package version 2.6-4.

Microbial diversity at both the genus and species taxa levels was calculated using the vegan package version 2.6-4. The number of genera or species was determined for each sample, and the Shannon diversity index was calculated accordingly. Linear mixed-effect models, utilising the "lmer" function in R, were employed to test the associations between the relative abundance of genera or species and metadata variables. The number of genera or species, as well as the Shannon diversity index, were assessed against each individual metadata variable, with Time Point included as a fixed effect and Patient ID included as a random effect in the model.

To test for a differential abundance of genera and species in relation to available background metadata variables for the infant samples, the Maaslin2 package version 1.12.0 (Microbiome Multivariable Associations with Linear Models)[70] was utilised. The Maaslin2 analysis was performed with

default parameters, including Benjamini-Hochberg FDR multiple testing correction and a significance threshold of 0.25. No minimum prevalence or abundance cut-offs were applied, and the data input was previously normalised. Each metadata variable was individually tested, incorporating Time Point as a Fixed Effect and Patient ID as a random effect grouping variable in a mixed-effect model. Positive associations were identified individually with Breast Feeding and Diarrhoeal Illness and these two variables where tested together in the same model as Fixed Effects with Time Point and with Patient ID included as a Random Effect in the final analysis.

Associations between faecal cytokine concentrations and metadata variables were tested using the MaAsLin2 package version 1.12.0[70], differences in concentration against background metadata variables available for the infant samples. For faecal cytokine analysis, MaAsLin2 was run with default parameters, including multiple testing correction using Benjamini–Hochberg FDR with a significance threshold of 0.25, except no minimum prevalence or abundance cut-offs were applied, normalised methods used were "CLR" and the transformation "LOG". Each metadata variable was tested individually with time point as a fixed effect and with patient ID included as a random effect grouping variable to run a mixed effect model.

Correlations between faecal cytokine concentration and the relative abundance of genus and species were tested using a Pearson correlation with correction for multiple testing using Benjamini and Hochberg FDR correcting between both cytokines and time points. Both cytokine concentration and genus/species abundance were first normalised using the bestNormalize package version 9.1.1 in R.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Bacterial pure genome assemblies are now deposited at NCBI GenBank under project accession PRJNA958097, while 16S rRNA amplicon sequencing raw reads are deposited in NCBI SRA under project PRJNA958098.

Code availability

Bacterial pure genome assemblies are now deposited at NCBI GenBank under project accession PRJNA958097, while 16S rRNA amplicon sequencing raw reads are deposited in NCBI SRA under project PRJNA958098.

Received: 19 January 2024; Accepted: 28 August 2024;
Published online: 14 September 2024

References

1. Abdill, R. J., Adamowicz, E. M. & Blekhman, R. Public human microbiome data are dominated by highly developed countries. *PLoS Biol.* **20**, e3001536 (2022).
2. Browne, H. P. et al. Boosting microbiome science worldwide could save millions of children's lives. *Nature* **625**, 237–240 (2024).
3. Porras, A. M. & Brito, I. L. The internationalization of human microbiome research. *Curr. Opin. Microbiol.* **50**, 50–55 (2019).
4. Lackey, K. A. et al. What's normal? Microbiomes in human milk and infant feces are related to each other but vary geographically: the INSPIRE study. *Front. Nutr.* **6**, 45 (2019).
5. Cambodia Demographic and Health Survey. *Cambodia Demographic and Health Survey* https://dhsprogram.com/publications/publication-PR136-Preliminary-Reports-Key-Indicators-Reports.cfm (2022).
6. Auguet, O. T. et al. Population-level faecal metagenomic profiling as a tool to predict antimicrobial resistance in Enterobacterales isolates causing invasive infections: An exploratory study across Cambodia, Kenya, and the UK. *EClinicalMedicine* **36**, 100910 (2021).
7. van Aartsen, J. J. et al. Epidemiology of paediatric gastrointestinal colonisation by extended spectrum cephalosporin-resistant *Escherichia coli* and *Klebsiella pneumoniae* isolates in north-west Cambodia. *BMC Microbiol.* **19**, 59 (2019).
8. Finlayson-Trick, E. et al. The effect of oral iron supplementation on gut microbial composition: a secondary analysis of a double-blind, randomized controlled trial among Cambodian women of reproductive age. *Microbiol. Spectr.* **11**, e0527322 (2023).
9. Miyazaki, A. et al. Determinants of growth measurements in rural Cambodian infants: a cross-sectional study. *Int Health* **13**, 49–56 (2021).
10. Miyazaki, A. et al. Frequent unregulated use of antibiotics in rural Cambodian infants. *Trans. R. Soc. Trop. Med. Hyg.* **114**, 401–407 (2020).
11. Robertson, R. C. The Gut Microbiome in Child Malnutrition. in *Nestlé Nutrition Institute Workshop Series* Vol. 93 (eds. Michaelsen, K. F., Neufeld, L. M. & Prentice, A. M.) 133–144 (S. Karger AG, 2020).
12. Surono, I. S., Widiyanti, D., Kusumo, P. D. & Venema, K. Gut microbiota profile of Indonesian stunted children and children with normal nutritional status. *PLoS ONE* **16**, e0245399 (2021).
13. Robertson, R. C., Manges, A. R., Finlay, B. B. & Prendergast, A. J. The human microbiome and child growth—first 1000 days and beyond. *Trends Microbiol.* **27**, 131–147 (2019).
14. Henrick, B. M. et al. Bifidobacteria-mediated immune system imprinting early in life. *Cell* **184**, 3884–3898.e11 (2021).
15. Taft, D. H. et al. Bifidobacterium species colonization in infancy: a global cross-sectional comparison by population history of breastfeeding. *Nutrients* **14**, 1423 (2022).
16. MAL-ED Network Investigators. The MAL-ED Study: A Multinational and Multidisciplinary Approach to Understand the Relationship Between Enteric Pathogens Malnutrition Gut Physiology Physical Growth Cognitive Development and Immune Responses in Infants and Children Up to 2 Years of Age in Resource-Poor Environments. *Clin. Infect. Dis.* **59**, S193–S206 (2014).
17. Matharu, D. et al. Bacteroides abundance drives birth mode dependent infant gut microbiota developmental trajectories. *Front. Microbiol.* **13**, 953475 (2022).
18. Kijner, S., Cher, A. & Yassour, M. The infant gut commensal *Bacteroides dorei* presents a generalized transcriptional response to various human milk oligosaccharides. *Front. Cell. Infect. Microbiol.* **12**, 854122 (2022).
19. Pérez-Burillo, S. et al. Effect of freezing on gut microbiota composition and functionality for in vitro fermentation experiments. *Nutrients* **13**, 2207 (2021).
20. Poulsen, C. S., Kaas, R. S., Aarestrup, F. M. & Pamp, S. J. Standard sample storage conditions have an impact on inferred microbiome composition and antimicrobial resistance patterns. *Microbiol. Spectr.* **9**, e01387–21 (2021).
21. Roesch, L. F. W. et al. Influence of fecal sample storage on bacterial community diversity. *Open Microbiol. J.* **3**, 40–46 (2009).
22. Cardona, S. et al. Storage conditions of intestinal microbiota matter in metagenomic analysis. *BMC Microbiol.* **12**, 158 (2012).
23. Gomez, A. et al. Gut microbiome of coexisting BaAka pygmies and Bantu reflects gradients of traditional subsistence patterns. *Cell Rep.* **14**, 2142–2153 (2016).
24. Nsubuga, A. M. et al. Factors affecting the amount of genomic DNA extracted from ape faeces and the identification of an improved sample storage method. *Mol. Ecol.* **13**, 2089–2094 (2004).
25. Olm, M. R. et al. Robust variation in infant gut microbiome assembly across a spectrum of lifestyles. *Science* **376**, 1220–1223 (2022).
26. Casaburi, G. et al. Metagenomic insights of the infant microbiome community structure and function across multiple sites in the United States. *Sci. Rep.* **11**, 1472 (2021).
27. Kuang, Y.-S. et al. Composition of gut microbiota in infants in China and global comparison. *Sci. Rep.* **6**, 36666 (2016).

28. Vatanen, T. et al. A distinct clade of *Bifidobacterium longum* in the gut of Bangladeshi children thrives during weaning. *Cell* **185**, 4280–4297.e12 (2022).
29. De Filippo, C. et al. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proc. Natl Acad. Sci. USA* **107**, 14691–14696 (2010).
30. Pop, M. et al. Diarrhea in young children from low-income countries leads to large-scale alterations in intestinal microbiota composition. *Genome Biol.* **15**, R76 (2014).
31. de Goffau, M. C. et al. Gut microbiomes from Gambian infants reveal the development of a non-industrialized *Prevotella*-based trophic network. *Nat. Microbiol.* **7**, 132–144 (2022).
32. Alcon-Giner, C. et al. Microbiota supplementation with *Bifidobacterium* and *Lactobacillus* modifies the preterm infant gut microbiota and metabolome: an observational study. *Cell Rep. Med.* **1**, 100077 (2020).
33. Matteuzzi, D., Crociani, F., Zani, O. & Trovatelli, L. D. *Bifidobacterium suis* n. sp.: a new species of the genus *Bifidobacterium* isolated from pig faces. *Z. Allg. Mikrobiol.* **11**, 387–395 (1971).
34. Zani, G., Biavati, B., Crociani, F. & Matteuzzi, D. Bifidobacteria from the faeces of piglets. *J. Appl. Bacteriol.* **37**, 537–547 (1974).
35. Monira, S. et al. Gut microbiota of healthy and malnourished children in Bangladesh. *Front. Microbiol.* **2**, 228 (2011).
36. Luchen, C. C. et al. Impact of antibiotics on gut microbiome composition and resistome in the first years of life in low- to middle-income countries: A systematic review. *PLoS Med* **20**, e1004235 (2023).
37. Chung The, H. et al. Exploring the genomic diversity and antimicrobial susceptibility of *Bifidobacterium pseudocatenulatum* in a Vietnamese population. *Microbiol. Spectr.* **9**, e00526–21 (2021).
38. Kobata, R. et al. High levels of growth factors in human breast milk. *Early Hum. Dev.* **84**, 67–69 (2008).
39. Siafakas, C. G., Anatolitou, F., Fusunyan, R. D., Walker, W. A. & Sanderson, I. R. Vascular endothelial growth factor (VEGF) is present in human breast milk and its receptor is present on intestinal epithelial cells. *Pediatr. Res.* **45**, 652–657 (1999).
40. Vuorela, P., Andersson, S., Carpén, O., Ylikorkala, O. & Halmesmäki, E. Unbound vascular endothelial growth factor and its receptors in breast, human milk, and newborn intestine. *Am. J. Clin. Nutr.* **72**, 1196–1201 (2000).
41. Kiu, R. et al. Particular genomic and virulence traits associated with preterm infant-derived toxigenic *Clostridium perfringens* strains. *Nat. Microbiol.* **8**, 1160–1175 (2023).
42. Sabnis, A. et al. Intestinal vascular endothelial growth factor is decreased in necrotizing enterocolitis. *Neonatology* **107**, 191–198 (2015).
43. Zanardo, V. et al. Effect of maternal smoking on breast milk interleukin-1α, β-endorphin, and leptin concentrations. *Environ. Health Perspect.* **113**, 1410–1413 (2005).
44. Laursen, M. F. et al. Bifidobacterium species associated with breastfeeding produce aromatic lactic acids in the infant gut. *Nat. Microbiol.* **6**, 1367–1382 (2021).
45. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
46. Huson, D. H. et al. MEGAN community edition—interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput. Biol.* **12**, e1004957 (2016).
47. Huson, D. H. & Mitra, S. Introduction to the analysis of environmental sequences: metagenomics with MEGAN. In *Evolutionary Genomics: Statistical and Computational Methods* Vol. 2 (ed. Anisimova, M.) 415–429 (Humana Press, Totowa, NJ, 2012).
48. Browne, H. P. *Development of Anaerobic Culturing Methods to Study the Evolution And Transmission of the Human Intestinal Microbiota*. (London School of Hygiene & Tropical Medicine, 2018).
49. Browne, H. P. et al. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* **533**, 543–546 (2016).
50. Forster, S. C. et al. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat. Biotechnol.* **37**, 186–192 (2019).
51. Harris, S. R. et al. Evolution of MRSA during hospital transmission and intercontinental spread. *Science* **327**, 469–474 (2010).
52. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
53. Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).
54. Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
55. Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P. & Parks, D. H. G. T. D. B.- Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**, 1925–1927 (2020).
56. Olm, M. R., Brown, C. T., Brooks, B. & Banfield, J. F. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **11**, 2864–2868 (2017).
57. Katz, L. S. et al. Mashtree: a rapid comparison of whole genome sequence files. *J. Open Source Softw.* **4** (2019).
58. Shao, Y. et al. Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* **574**, 117–121 (2019).
59. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).
60. Seemann, T. *ABRicate: Mass Screening of Contigs for Antimicrobial and Virulence Genes* (2023). https://github.com/tseemann/abricate.
61. Bortolaia, V. et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* **75**, 3491–3500 (2020).
62. Liu, B. et al. 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* **47**, D687–D692 (2019).
63. Lawson, M. A. E. et al. Breast milk-derived human milk oligosaccharides promote *Bifidobacterium* interactions within a single ecosystem. *ISME J.* **14**, 635–648 (2020).
64. Kiu, R. *TOXIper: Toxinotype assignment of Clostridium perfringens via ABRicate* (2023). https://github.com/raymondkiu/TOXIper.
65. Tonkin-Hill, G. et al. Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol.* **21**, 180 (2020).
66. Katoh, K. & Standley, D. M. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
67. Page, A. J. et al. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genom.* **2**, e000056 (2016).
68. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
69. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).
70. Mallick, H. et al. Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput. Biol.* **17**, e1009442 (2021).

Acknowledgements

We thank the families for participating in the study. L.J.H. is supported by Wellcome Trust Investigator Awards 100974/C/13/Z and 220876/Z/20/Z; the Biotechnology and Biological Sciences Research Council (BBSRC), Institute Strategic Programme Gut Microbes and Health BB/R012490/1, and its constituent projects BBS/E/F/000PR10353 and BBS/E/F/000PR10356.

Author contributions

Conceptualisation: L.J.H. and S.E.C.; Methodology: M.J.D., R.K., I.R.S., A.M., H.A.-P., R.T., S.C., S.P., M.K., M.M., A.I., and B.T.; Software: M.J.D.,

R.K., S.C., and M.K.; Validation: M.J.D., K.S., S.E.C., and L.J.H.; Formal analysis: M.J.D., R.K., A.M., S.E.C., and L.J.H.; Investigation: M.J.D., R.K., I.R.S., A.M., H.A.-P., R.T., S.C., S.P., M.K., M.M., A.I., B.T., S.E.C., and L.J.H.; Resources: M.J.D., A.M., R.T., M.M., A.I., B.T., S.E.C., and L.J.H.; Data curation: M.J.D., R.K., A.M., S.E.C.; Writing—original draft, M.J.D. and L.J.H.; Writing—review and editing, M.J.D., R.K., I.R.S., A.M., H.A.-P., R.T., S.C., S.P., M.K., M.M., A.I., B.T., S.E.C., and L.J.H.; Visualisation: M.J.D. and R.K.; Supervision: S.E.C. and L.J.H.; Project administration: M.J.D., S.E.C., and L.J.H.; Funding acquisition: S.E.C. and L.J.H.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41522-024-00562-0.

Correspondence and requests for materials should be addressed to Lindsay J. Hall.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2024

Jain V, Dalby MJ, Alexander EC, Burford C, Acford-Palmer H, Serghiou IR, Teng NMY, Kiu R, Gerasimidis K, Zafeiropoulou K, Logan M, Verma A, Davenport M, Hall LJ, Dhawan A, 2024. Association of gut microbiota and gut metabolites and adverse outcomes in biliary atresia: A longitudinal prospective study. Hepatology Communications; 8(11): e0550.

Received: 31 July 2024 | Accepted: 1 August 2024

DOI: 10.1097/HC9.0000000000000550

ORIGINAL ARTICLE

OPEN

Association of gut microbiota and gut metabolites and adverse outcomes in biliary atresia: A longitudinal prospective study

Vandana Jain[1] | Matthew J. Dalby[2,3] | Emma C. Alexander[1] | Charlotte Burford[1] | Holly Acford-Palmer[2] | Iliana R. Serghiou[2] | Nancy M.Y. Teng[2] | Raymond Kiu[2,3] | Konstantinos Gerasimidis[4] | Konstantina Zafeiropoulou[4] | Michael Logan[4] | Anita Verma[5] | Mark Davenport[6] | Lindsay J. Hall[2,3,7] | Anil Dhawan[1]

[1]Paediatric Liver, GI and Nutrition Centre and Mowatlabs, King's College Hospital, London, UK

[2]Quadram Institute Bioscience, Norwich Research Park, Norwich, UK

[3]Institute of Microbiology and Infection, College of Medical and Dental Sciences, University of Birmingham, Birmingham, UK

[4]Human Nutrition, School of Medicine, Dentistry and Nursing, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow Royal Infirmary, Glasgow, UK

[5]Department of Infection Science, Paediatric Liver, GI and Nutrition Centre and Mowatlabs, King's College Hospital, London, UK

[6]Department of Paediatric Surgery, King's College Hospital, London, UK

[7]Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, UK

Correspondence
Vandana Jain, Department of Paediatric Hepatology, King's College Hospital, Denmark Hill, London SE5 9RS, UK.
Email: vjain@nhs.net

Abstract

Background: The Kasai portoenterostomy (KPE) aims to re-establish bile flow in biliary atresia (BA); however, BA remains the commonest indication for liver transplantation in pediatrics. Gut microbiota-host interplay is increasingly associated with outcomes in chronic liver disease. This study characterized fecal microbiota and fatty acid metabolites in BA.

Methods: Fecal samples were prospectively collected in newly diagnosed BA infants (n = 55) before and after KPE. Age-matched healthy control (n = 19) and cholestatic control (n = 21) fecal samples were collected. Fecal 16S rRNA gene amplicon sequencing for gut microbiota and gas chromatography for fecal fatty acids was performed.

Results: Increased abundance of *Enterococcus* in pre-KPE BA and cholestatic control infants, compared to healthy infants, was demonstrated. At the early post-KPE time points, increased alpha diversity was revealed in BA versus healthy cohorts. A lower relative abundance of *Bifidobacterium* and increased *Enterococcus*, *Clostridium*, *Fusobacterium*, and *Pseudomonas* was seen in infants with BA. Fecal acetate was reduced, and fecal butyrate and propionate were elevated in early post-KPE BA infants. Higher post-KPE alpha diversity was associated with nonfavorable clinical outcomes (6-month jaundice and liver transplantation). A higher relative abundance of post-KPE *Streptococcus* and

Abbreviations: BA, biliary atresia; CC, cholestatic control; HC, healthy control; KPE, Kasai portoenterostomy; LT, liver transplantation; MCT, medium-chain triglyceride; NLS, native liver survival; NMDS, nonmetric multidimensional scaling; PELD, Pediatric End-Stage Liver Disease; SCFA, short-chain fatty acid; TB, total bilirubin.

Vandana Jain and Matthew J. Dalby contributed equally to the manuscript and should be considered co-first authors.

Lindsay J. Hall and Anil Dhawan also contributed equally to the manuscript.

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.hepcommjournal.com.

This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright © 2024 The Author(s). Published by Wolters Kluwer Health, Inc. on behalf of the American Association for the Study of Liver Diseases.

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

Fusobacterium and a lower relative abundance of *Dorea*, *Blautia*, and *Oscillospira* were associated with nonfavorable clinical outcomes. *Blautia* inversely correlated to liver disease severity, and *Bifidobacterium* inversely correlated to fibrosis biomarkers. *Bifidobacterium* abundance was significantly lower in infants experiencing cholangitis within 6 months after KPE.

Conclusions: Increased diversity, enrichment of pathogenic, and depletion of beneficial microbiota early post-KPE are all factors associated with nonfavorable BA outcomes. Manipulation of gut microbiota in the early postsurgical period could provide therapeutic potential.

Keywords: biliary atresia, liver transplantation, microbiota, short-chain fatty acid

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

INTRODUCTION

Biliary atresia (BA) is a disease characterized by fibro-obliteration of the biliary tree, presenting in infancy with obstructive jaundice that blocks the flow of bile from the liver to the gut.[1] The etiology of BA remains uncertain but likely involves genetic and developmental factors that result in an immune response targeting the bile ducts. The Kasai portoenterostomy (KPE) surgery relieves the bile duct obstruction and re-establishes the flow of bile by creating a bilio-enteric conduit. However, jaundice clearance is only achieved in 60% of infants, and the progression of liver disease to cirrhosis that requires liver transplantation (LT) often continues even in infants with successful jaundice clearance.[2] BA remains the commonest cause of LT in pediatrics. Post-KPE pharmacotherapy that aims to dampen fibro-inflammatory processes has not translated into improved native liver survival (NLS).

Over the last decade, the development of high-throughput genetic sequencing techniques has facilitated advanced detection and classification of the complex ecosystem of the *gut microbiota* (gut bacterial community) and their encoded genes, the *gut microbiome*. An emerging role for the gut microbiota in disease processes has been described, with gut microbiota perturbations (often referred to as dysbiosis) being described in numerous chronic disorders, including across the spectrum of liver disease in adult cohorts.[3] The term "gut-liver axis" highlights the close anatomical and functional relationship between the liver and the gut microbiota. Metabolites produced by gut microbes, such as short-chain fatty acids (SCFAs) and secondary bile acids, are involved in modulating host metabolism[4,5] and are being investigated in liver disease.

From as early as the 1980s, studies based on culturing of fecal samples have identified alterations in the gut microbiota in BA. Cholangitis, the inflammation of the biliary system, after KPE, provides the most convincing evidence of gut microbiota pathogenesis in BA, with potential pathways including the reflux of gut bacteria through the bilio-enteric conduit or bacterial translocation through the enterohepatic circulation.[6] In light of data implicating gut microbiota pathogenesis in cholangitis, as well as in adult cholestatic and cirrhotic disease,[7] sequencing techniques are identifying gut microbial disturbance in BA.[8–16]

In this prospective longitudinal study, a comprehensive analysis of the fecal microbiota and paired fecal fatty acid in BA pathogenesis and disease progression was conducted.

METHODS

Patient and control cohorts

In this prospective longitudinal single-center study, infants were recruited into 3 separate cohorts between January 2016 and December 2019: (i) BA, (ii) healthy control (HC), and (iii) cholestatic control (CC).

BA cohort

All infants with newly diagnosed BA, confirmed by intraoperative cholangiogram and/or liver biopsy pathology at King's College Hospital (KCH), were candidates for the present study. Longitudinal assessment at baseline (pre-KPE), 6 weeks (6w), 12 weeks (12w), and 24 weeks (24w) post-KPE time points was performed for all infants. Exclusion criteria included previous bowel resection, candidate for primary LT, an extrahepatic disease associated with significant morbidity/mortality, and international candidates unable to attend longitudinal study visits (see Supplemental Methods for Clinical BA Protocol, http://links.lww.com/HC9/B57).

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

HC cohort

HC infants were recruited from the local community at 1 of the following chronological age-matched time points to the BA cohort: (i) 6 weeks (pre-KPE equivalent), (ii) 3 months (6w-post-KPE equivalent), (iii) 5 months (12w-post-KPE equivalent), and (iv) 8 months (24w-post-KPE equivalent). HC infants were not the same infants followed up longitudinally. Inclusion criteria included >37 weeks gestation and exclusively breastfed from birth. Exclusion criteria included any surgical intervention or postnatal medication administration.

Non-BA CC cohort

Infants presenting to KCH for investigation of conjugated jaundice (defined as conjugated fraction >20% total bilirubin) between the ages of 4 weeks and 2 months of age were recruited to the study as CCs. Exclusion criteria included previous bowel resection and subsequent BA diagnosis.

Study data and clinical endpoints

All cohorts

Demographic, dietary, and pharmacotherapy characteristics were recorded for all 3 study cohorts (BA, HC, and CC) at baseline (pre-KPE time point and equivalent). The predominant baseline milk diet, reported by parent(s), from birth to initial feces collection, was summarized as (i) exclusively breast-fed, (ii) mixed breast/formula–fed, and (iii) formula-fed. The administration of high medium-chain triglyceride (MCT) milk before feces collection in BA/CC cohorts was also documented. For BA and CC infants, administration of ursodeoxycholic acid and/or postnatal antibiotics before initial feces collection was documented. Post-KPE and equivalent time points, demographic, dietary, and pharmacotherapy characteristics for BA and healthy infants were recorded. Dietary characteristics were documented using 3-day pre-visit dietary diaries, completed by parent(s): (i) exclusively breastfed, (ii) exclusively MCT-fed, (iii) mixed breast/MCT-fed, (iv) standard formula–fed, and (v) weaned versus non-weaned. "Weaned" refers to the process of introducing solid food.

BA cohort

Biochemical, hematological, and anthropometric data were collected at all time points (pre-KPE and 6w-, 12w-, and 24w-post-KPE). Biochemical parameters included total bilirubin (TB, μmol/L), AST (IU/L), GGT (IU/L), and albumin (g/L). Hematological parameters included international normalized ratio and platelet count ($\times 10^9$/L). Liver stiffness measurement (kPa) was measured using transient elastography (Fibroscan) (see Supplemental Methods for transient elastography details, http://links.lww.com/HC9/B57). Mathematical scores, Pediatric End-Stage Liver Disease (PELD), and AST-to-platelet ratio index were calculated at each time point (see Supplemental Methods for scoring systems, http://links.lww.com/HC9/B57). Anthropometric parameters included weight (kg) and length (cm) measurements required for PELD calculation.

Primary clinical outcome measures were (i) 6-month jaundice clearance (BA-JC, TB <20 μmol/L by 6 mo after KPE: BA-J, TB ≥20 μmol/L by 6 mo after KPE) and (ii) 2-year NLS (BA-NLS, NLS by 2 y after KPE; BA-LT, LT/listed for LT by 2 y after KPE). Favorable clinical outcomes were defined as BA-JC and BA-NLS, and nonfavorable clinical outcomes were defined as BA-J and BA-LT. The secondary clinical outcome measure was cholangitis (see Supplemental Methods for cholangitis definitions, http://links.lww.com/HC9/B57).

Fecal sample collection

Fecal samples were collected from all BA and HC infants at pre-KPE and post-KPE (6w-, 12w-, and 24w) equivalent time points. Fecal samples from CC infants were collected at the pre-KPE time point only. Samples were collected directly from the nappy using a sterile spatula and pot. All samples were kept refrigerated during transport at approximately +4°C for a maximum of 24 hours before being processed. Fecal samples were (i) stored at −80°C for DNA extraction and sequencing and (ii) mixed with an equal volume of 1 M sodium hydroxide (NaOH) and stored at −80°C, for fatty acid analysis.

DNA extraction, 16S rRNA gene amplicon sequencing of fecal samples, and sequence analysis

FastDNA Spin Kit for Soil (MP) was used to extract genomic DNA from feces following the manufacturer's instructions, with an extended 3-minute bead-beating. DNA concentration was quantified using a Qubit 2.0 fluorometer (Invitrogen) (see Supplemental Methods for 16S rRNA gene amplicon sequencing of fecal samples and sequence analysis, http://links.lww.com/HC9/B57).

Determination of fecal SCFAs

SCFAs, acetate, propionate, and butyrate, were measured using gas chromatography (Agilent 7820A; Agilent

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

Technologies) with a DB-WAX UI column on diethyl ether extracts, with nitrogen as the carrier gas.

Statistical analysis

Continuous variables comprising non-normal data were reported as median values and IQR. Two independent groups were analyzed using the Mann-Whitney *U* test. Multiple independent groups were analyzed using the Kruskal-Wallis test, with pairwise comparisons performed using the Dunn (1964) procedure and Bonferroni correction. Categorical variables were reported as frequency and percentages (%). Pearson chi-square test was used to analyze categorical variables. The Fisher exact test was used to analyze categorical variables if expected cell frequencies are < 5. A p value < 0.05 was considered statistically significant for all tests. Mann-Whitney *U*, Kruskal-Wallis, and chi-square analyses were performed using IBM SPSS version 28.

Microbial diversity

Microbial diversity, at the genus level, was calculated using the vegan package version 2.6-4. The number of genera or species was determined for each sample, and the Shannon diversity index and Inverse Simpson diversity index were calculated accordingly.

Microbiota composition

The overall composition of the 16S rRNA gene sequence data was analyzed using NMDS (nonmetric multidimensional scaling). These plots were generated in R Studio using the vegan package version 2.6-4, employing a Bray-Curtis dissimilarity calculation. To determine the statistical significance of differences in microbial community structure based on infant metadata variables, permutational MANOVA was conducted using the Adonis function from the vegan R package version 2.6-4. Arrows and genus labels on the NMDS plots indicate bacterial genera driving the separation of points on the NMDS plots. To test for the differential abundance of genera and fecal fatty acids in relation to the infant cohort group and clinical outcome variables for the infant samples, the Maaslin2 package (Microbiome Multivariable Associations with Linear Models)[17] was used. The analysis in Maaslin2 was performed using default options recommended for the software package, with an association being considered significant at a p value threshold of $p < 0.05$ and q value of < 0.25 after correction for multiple testing using the Benjamini-Hochberg false discovery rate method. A genus-level association with a false discovery rate–adjusted p value < 0.25 has commonly been considered statistically significant in previous microbiome studies. The Maaslin2 analysis was performed using Benjamini-Hochberg false discovery rate multiple testing correction, with a significance threshold of 0.25 used as the default threshold recommended for the Maaslin2 package. For microbiota genus abundance data, the analysis method "LM" was used with a minimum prevalence cutoff of 0.1, and an abundance cutoff of 0.1 was applied, with no normalization method used because the data input was previously normalized. For SCFA data, the analysis method "LM" was used with a minimum prevalence cutoff of 0 and an abundance cutoff or 0 applied, with the normalization method "TSS" and transformation method "LOGIT" applied to the data. Each infant cohort group and clinical outcome variable was tested individually as a fixed effect. For analyses involving more than 1 time point, the individual participant ID was included as a random effect.

Correlation analysis

Pearson's correlation with correction for multiple genus testing (using the Benjamini-Hochberg method) was conducted to determine (a) associations between disease severity and relative abundance of the top 10 most genera and (b) associations between SCFA concentrations and the relative abundance of the top 10 genera. Variables were normalized using the bestNormalize package (version 1.9.1). Individual correlations were plotted using ggplot2 (version 3.1.4) in R as a scatter plot with a line of best fit determined by linear regression.

Ethical statement

The experimental protocol was approved by the London-Harrow Research Ethics Committee (REC 15/LO/1966) in accordance with the ethical standards established in the 1964 Declaration of Helsinki. Informed consent was obtained from the parent/guardian before specimen collection. CC samples were stored under the Biobank Ethical Agreement (REC 18/WA/0009), and access to samples was granted by the Biobank Access Committee.

RESULTS

Characteristics of pre-KPE fecal microbiota and SCFA profiles in BA, compared to age-matched control cohorts

Fifty-five infants with newly diagnosed BA were recruited before KPE, and followed at 6w-, 12w-, and 24w after KPE. In addition, 18, 19, 18, and 17 "separate" HC infants were recruited at equivalent ages

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

to pre-KPE, 6w-, 12w-, and 24w-post-KPE time points. Twenty-one non-BA cholestatic infants were recruited at an equivalent age to the pre-KPE time point (see Supplemental Results for further details on CC, http://links.lww.com/HC9/B57). Fecal sample collection and 16S rRNA gene amplicon sequencing were successful in 33 BA, 17 HC, and 19 CC infants. Gas chromatography measurement of fecal fatty acids was successfully obtained from 34 BA, 11 HC, and 15 CC infants. Factors contributing to the missing DNA sequencing data included insufficient stool sample available or DNA quantification too low for sequencing. Factors contributing to missing fatty acid data included insufficient stool samples available or insufficient material for extraction after freeze-drying.

Table 1 shows baseline characteristics for the total infant cohort recruitment; Supplemental Tables S1 and S2, http://links.lww.com/HC9/B57, show baseline characteristics for fecal microbiota and fecal SCFA subcohorts, respectively. Demographic characteristics were similar between BA and HC infants at baseline, while there were fewer vaginal deliveries and lower birthweights among CC infants. Baseline diet was similar between BA and CC cohorts but significantly different between these "diseased" cohorts and the healthy cohort, with 33% of BA and 29% of CC infants exclusively breastfed compared to 100% of healthy infants. The majority of BA (82%) and CC (71%) infants were already receiving specialized high MCT formula milk. UCDA and postnatal antibiotic use was higher in CC versus BA infants. CC infants displayed a milder cholestatic phenotype (lower TB and GGT levels) compared to BA infants.

Comparison of pre-KPE BA fecal microbiota and SCFA profiles against age-matched control cohorts

To examine the gut microbiota composition at the pre-KPE time point, the 10 bacterial genera with the highest abundance were visualized as stacked bar plots for BA, HC, and CC infants (Figure 1A). *Streptococcus* and *Bifidobacterium* were visually dominant in all 3 groups, with *Enterococcus* and *Clostridium* being more apparent in BA and CC infants. Microbial diversity was similar between the 3 groups (Figure 1B, Supplemental Figure S1A, http://links.lww.com/HC9/B57). Clustering samples using NMDS, based on a Bray-Curtis matrix, showed distinct clustering between BA and healthy infants (PERMANOVA test: R^2 = 6.6%, p = 0.034), with no separation between BA and CC infants

TABLE 1 Baseline characteristic comparison between pre-KPE BA and age-matched HC and CC cohorts for total cohort recruitment

	Total cohort recruitment					
Baseline characteristics	**BA, n = 55**	**HC, n = 18**	**CC, n = 21**	**BA vs. HC, *p***	**BA vs. CC, *p***	**CC vs. HC, *p***
Demographics						
Age at time point (d), median (IQR)	49 (17, 127)	40 (29, 61)	51 (22,95)	0.82	0.07	0.16
Mode of birth delivery (vaginal)	31 (56)	12 (67)	5 (24)	0.58	***0.02***	***0.01***
Sex (M)	33 (60)	12 (67)	13 (62)	0.78	0.1	0.1
Prematurity (< 37 wk)	6 (11)	0 (0)	5 (24)	0.33	0.17	0.05
Birth weight (kg), median (IQR)	3.09 (2.75, 3.50)	3.31 (3.08, 3.63)	2.7 (1.79, 3.06)	0.09	***0.004***	***< 0.001***
Laboratory parameters						
TB (μmol/L), median (IQR)	130 (102, 148)	—	95 (61, 131)	—	***0.008***	—
GGT (IU/L), median (IQR)	366 (242, 724)	—	165 (88, 429)	—	***0.009***	—
Pharmacotherapy						
UDCA	33 (60.0)	0 (0)	18 (86)	—	0.05	—
Postnatal antibiotics	19 (35)	0 (0)	17 (81)	—	***< 0.001***	—
Predominant diet from birth						
Exclusively breastfed	18 (33)	18 (100)	6 (29)	***< 0.001***	0.79	***< 0.001***
Mixed breast/formula–fed	25 (46)	0 (0)	8 (38)	—	0.61	—
Formula-fed	12 (22)	0 (0)	7 (33)	—	0.38	—
MCT administration						
MCT-fed	45 (82)	0 (0)	15 (71)	—	0.36	—

Note: Data are n (%) unless stated otherwise. Statistical significance is represented as $p < 0.05$.
Bold and italic values are significant p-values < 0.05.
Abbreviations: BA, biliary atresia; CC, cholestatic control; HC, healthy control; KPE, Kasai portoenterostomy; MCT, medium-chain triglyceride; TB, total bilirubin; UDCA, ursodeoxycholic acid.

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

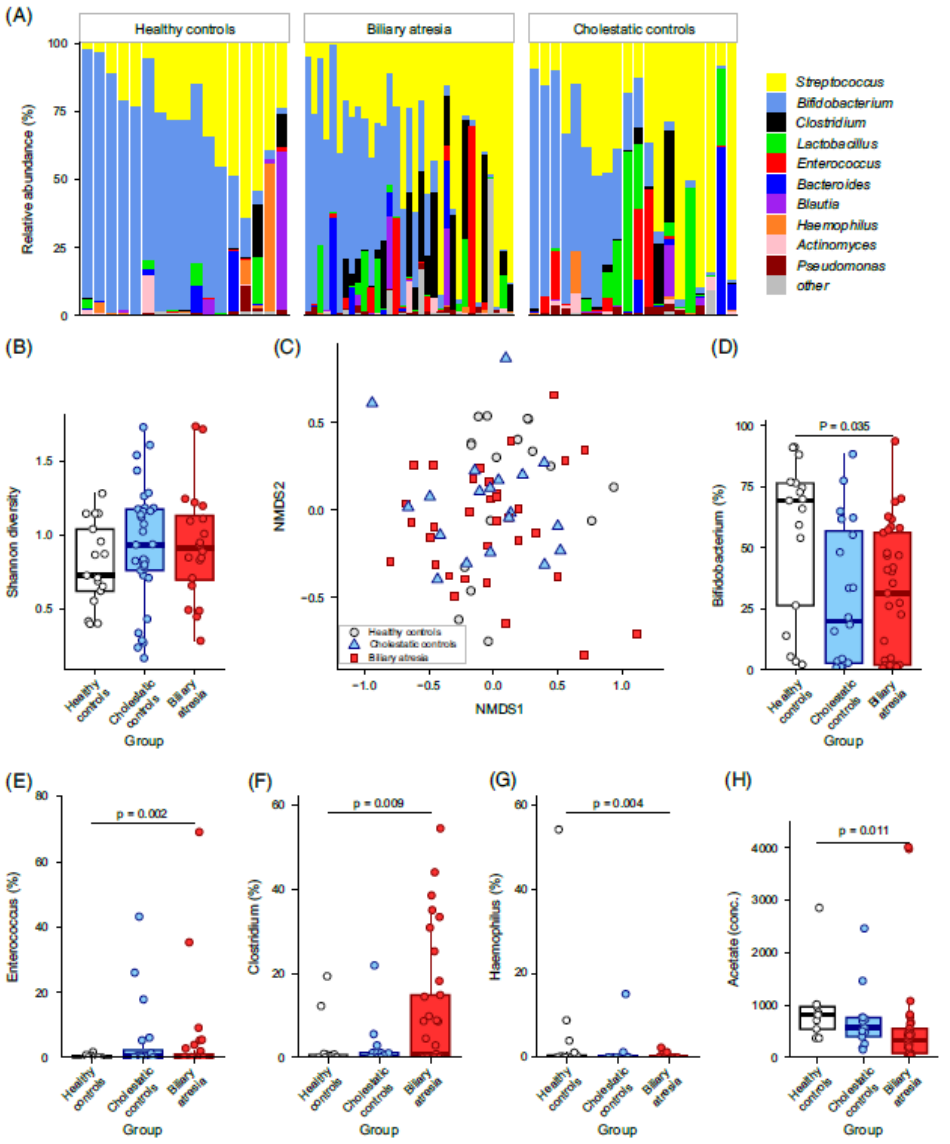


FIGURE 1 Pre-KPE composition of fecal microbiota and SCFAs in BA infants compared to healthy and cholestatic infants. (A) Pre-KPE relative abundance of the 10 most abundant bacterial genera in BA infants displayed as stacked bar plots compared to healthy and cholestatic control infants; (B) Shannon diversity of pre-KPE BA, age-matched healthy, and cholestatic infants. (C) NMDS plot showing clustering of genus level pre-KPE fecal microbiota (Stress = 0.18). PERMANOVA test: R^2 = 6.6%, p = 0.034. Pairwise analysis: HC versus BA, p = 0.066, HC versus CC: p = 0.091, BA versus CC: p = 1. Relative abundance of (D) *Bifidobacterium*, (E) *Enterococcus*, (F) *Clostridium*, and (G) *Haemophilus* between the infant groups. (H) Concentration of fecal acetate between infant groups. Statistical analysis of genus differences determined using Masslin2 with BA set as the reference group. Boxplots represent the median and interquartile range. Statistical significance was determined as $p < 0.05$. Abbreviations: BA, biliary atresia; CC, cholestatic control; HC, healthy control; KPE, Kasai portoenterostomy; NMDS, nonmetric multi-dimensional scaling; SCFA, short-chain fatty acids.

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

(Figure 1C). The differential abundance of genera between groups was determined using the MaAsLin2 package with BA as the reference group. The relative abundance of *Bifidobacterium* (Figure 1D) was lower (p = 0.035) in BA infants than in healthy infants, while the relative abundance of *Enterococcus* (p = 0.002) and *Clostridium* (p = 0.009) was higher in BA (Figures 1E, F), compared to healthy infants, but similar to CC infants. *Haemophilus* was also lower in BA infants compared to HC infants (Figure 1G).

Fecal SCFA concentration analyzed using the MaAsLin2 package, with BA as the reference group, showed lower acetate (p = 0.011) in BA at pre-KPE compared to healthy infants (Figure 1H). Butyrate concentrations also tended to be higher (p = 0.056) in BA compared to healthy infants (Supplemental Figure S1B, http://links.lww.com/HC9/B57). Fecal SCFA concentrations were similar between BA and cholestatic infants.

Longitudinal assessment of BA fecal microbiota

Differential abundance analysis of BA microbiota, with pre-KPE as the reference group, was conducted to understand their progression over time. *Enterococcus*, *Pseudomonas*, *Fusobacterium*, *Actinomyces*, and

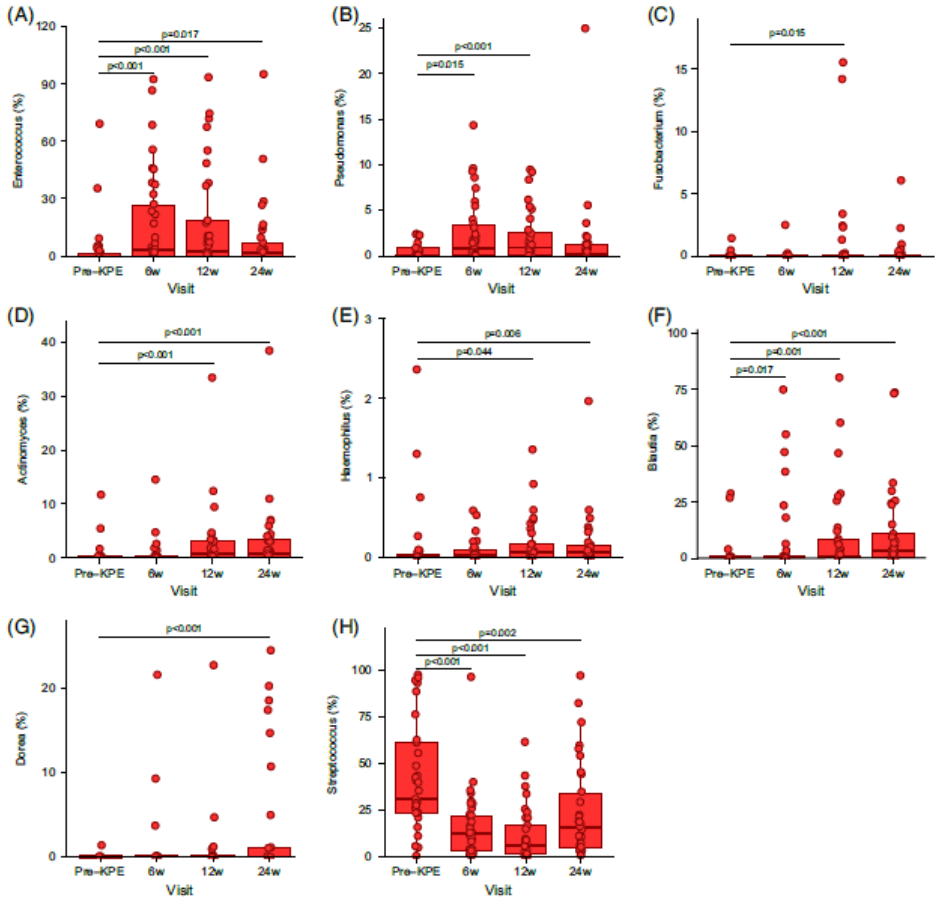


FIGURE 2 Longitudinal microbiota composition in BA infants. Longitudinal relative abundance of (A) *Enterococcus*, (B) *Pseudomonas*, (C) *Fusobacterium*, (D) *Actinomyces*, (E) *Haemophilus*, (F) *Blautia*, (G) *Dorea*, and (H) *Streptococcus* in BA infants across time points. Abbreviations: BA, biliary atresia; KPE, Kasai portoenterostomy.

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

Haemophilus abundance was higher after KPE compared to before KPE (Figures 2A–E). *Blautia* abundance progressively increased over time (Figure 2F), Dorea was higher at 24w after KPE (Figure 2G), and *Streptococcus* abundance was lower after KPE (Figure 2H).

Comparison of post-KPE fecal microbiota and SCFA profiles against age-matched HC

Fecal microbiota and fecal SCFAs were compared longitudinally at 6w-, 12w-, and 24w-post-KPE in BA and age-matched healthy infants. Table 2 and Supplemental Table S3, http://links.lww.com/HC9/B57, compare characteristics between BA and HC at all time points for fecal microbiota and fecal SCFA subcohorts, respectively. Breastfeeding was significantly lower in BA infants at all time points; pharmacotherapy administration was abundant in BA, compared to pharmacotherapy-naïve healthy infants. Weaning rates were similar between both cohorts.

Figure 3A illustrates a marked difference in the distribution of the top 10 genera between BA and HC infants at each time point. Clustering samples, using NMDS, highlighted distinct clustering between BA and healthy infants at all time points (PERMANOVA; 6w-post-KPE, $R^2 = 0.11$, $p < 0.001$; 12w-post-KPE, $R^2 = 0.17$, $p < 0.001$; 24w-post-KPE, $R^2 = 0.08$, $p = 0.002$; Supplemental Figures S2A–C, http://links.lww.com/HC9/B57). Microbiota diversity measured using Shannon and Inverse Simpson diversity was higher at 6w- and 12w-post-KPE in BA compared to healthy infants but similar between groups at 24w-post-KPE (Figure 3B, Supplemental Figure S2D, http://links.lww.com/HC9/B57). A lower abundance of *Bifidobacterium* in BA versus HC infants was demonstrated at all time points (Figure 3C), accompanied by a consistently higher relative abundance of *Enterococcus*, *Pseudomonas*, and *Clostridium* at 6w- and 12w-post-KPE, *Fusobacterium* at 12w-post-KPE (Figures 3D–F, Supplemental Figure S2E, http://links.lww.com/HC9/B57), and *Haemophilus* and *Actinomyces* (Supplemental Figures S2F, G, http://links.lww.com/HC9/B57) at 6w-post-KPE.

Fecal SCFA profiles differed significantly between BA and HC infants at each time point. Acetate concentration was significantly lower at 6w- ($p = 0.003$) and 12w-post-KPE ($p = 0.024$) in BA compared to healthy infants (Figure 3H). The concentrations of propionate and butyrate were strikingly higher in BA infants at 6w-post-KPE (Figures 3H, I). The Pearson correlation analysis was used to identify associations between individual genus abundance and fecal SCFA concentrations within the BA infant group (Supplemental Figures S3A–D, http://links.lww.com/HC9/B57). Significant microbiota-SCFA correlations were noted at 24w-post-KPE; notably, *Bifidobacterium* revealed a strong positive correlation with acetate ($r = 0.56$, $p = 0.017$) (Figure 3J).

Association between fecal microbiota and SCFA profiles and primary clinical outcome measures

To explore the clinical relevance of fecal microbiota in BA infants, fecal microbiota composition was compared between favorable and nonfavorable clinical outcome measures. Baseline characteristics (demographics, diet, and pharmacotherapy) were similar between outcome groups (Supplemental Tables S4 and S6, http://links.lww.com/HC9/B57). Dietary and pharmacotherapy characteristics for longitudinal time points are represented in Supplemental Tables S5 and S7, http://links.lww.com/HC9/B57.

Jaundice clearance

Shannon diversity and Inverse Simpson diversity were higher at 12w-post-KPE in BA-J versus BA-JC groups (Figure 4A, Supplemental Figure S4A, http://links.lww.com/HC9/B57). *Fusobacterium* abundance ($p = 0.009$) was higher at 24w-post-KPE and *Streptococcus* higher ($p = 0.003$) at 12w-post-KPE in BA-J compared to BA-JC groups (Figures 4B, C). Relative abundance of *Blautia* ($p = 0.021$), *Dorea* ($p = 0.046$), and *Oscillospira* ($p = 0.011$) was lower at 24w-post-KPE in BA-J compared to BA-JC groups (Figures 4D–F).

NLS

Shannon diversity and Inverse Simpson diversity were higher at 12w-post-KPE in BA-LT compared to BA-NLS groups (Figure 4G, Supplemental Figure S4B, http://links.lww.com/HC9/B57). *Blautia* ($p = 0.003$) and *Dorea* ($p = 0.026$) abundance was lower (Figures 4H, I), and *Oscillospira* tended to be lower ($p = 0.055$) (Figure 4J) at 24w-post-KPE in BA-LT compared to BA-NLS groups. No differences between SCFA abundance and primary clinical outcome measures were revealed.

Association between fecal microbiota and cholangitis

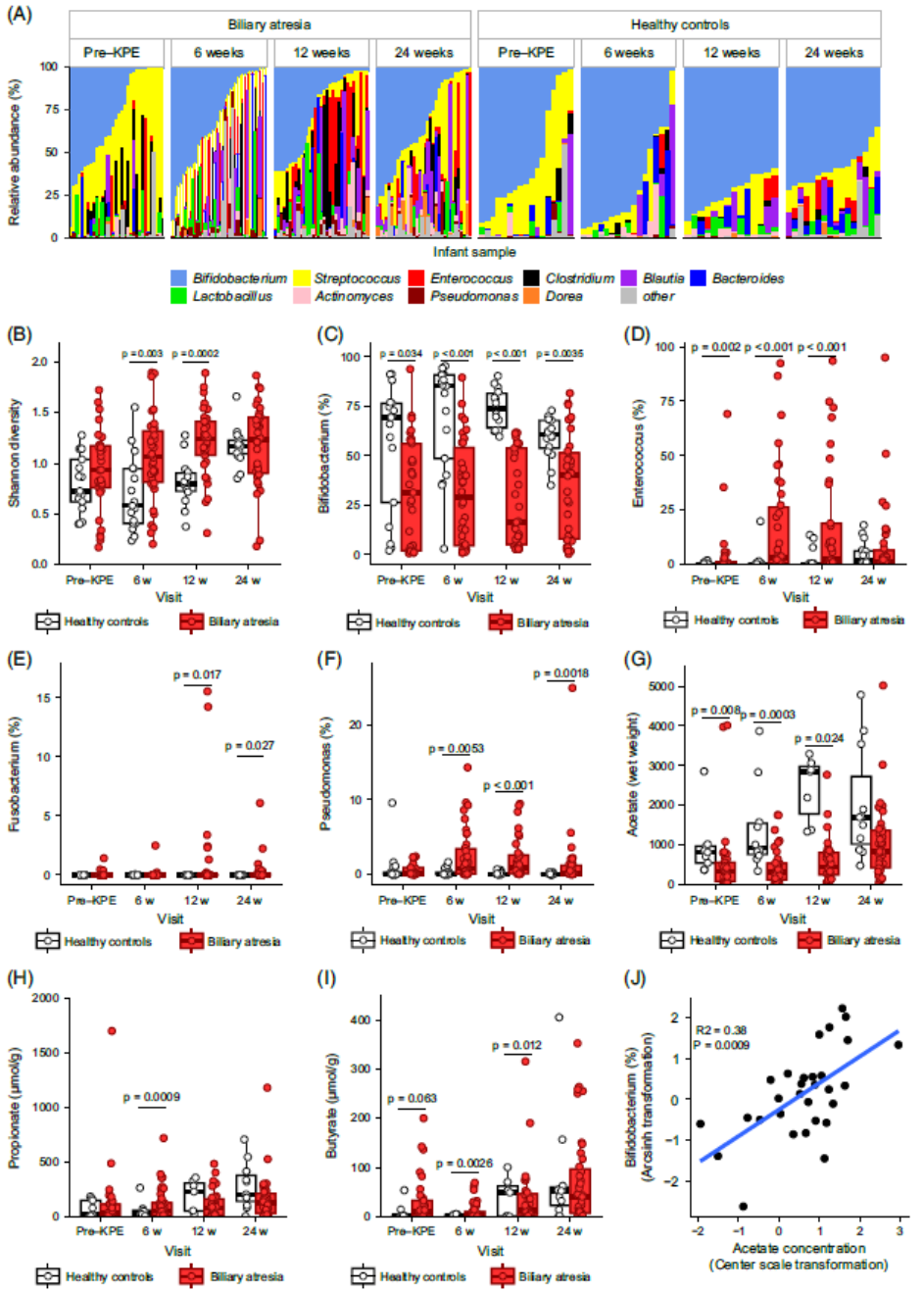
Fecal microbiota was compared between BA infants who had experienced "one or more episode(s) of cholangitis" versus "no cholangitis" between each time point. Sixteen percent (n = 6), 30% (n = 10), and 27% (n = 9) developed an episode(s) of cholangitis between "KPE to 6w-post-KPE," "6w- to 12w-post-KPE," "12w- to

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

TABLE 2 Baseline characteristic comparison between longitudinal BA and age-matched HC and CC subcohorts for fecal microbiota subcohort

	Fecal microbiota analysis								
	6w-post-KPE			12w-post-KPE			24w-post-KPE		
Confounder variable	BA, n = 38	HC, n = 17	*p*	BA, n = 33	HC, n = 14	*p*	BA, n = 33	HC, n = 16	*p*
Baseline characteristics									
Age at time point (wk)	12.4 (10.7, 15)	13.3 (13, 13.9)	0.2	19.4 (17, 22)	21.3 (20, 23)	0.1	32.9 (30, 35)	34.2 (32.8, 35.4)	0.2
Mode of birth delivery (vaginal)	23 (61)	11 (65)	0.8	18 (55)	8 (57)	0.9	20 (61)	12 (75)	0.3
Sex (M)	22 (58)	7 (41)	0.3	18 (55)	7 (50)	0.7	17 (52)	8 (50)	0.9
Prematurity (< 37 wk)	4 (11)	0 (0)	0.2	2 (6)	0 (0)	0.4	4 (12)	0 (0)	0.2
Birth weight (kg)	3.04 (2.7, 3.5)	3.4 (3, 3.5)	***0.04***	3.01 (2.8, 3.4)	3.4 (2.9, 3.6)	0.2	3 (2.7, 3.3)	3.4 (3, 3.6)	0.07
Pharmacotherapy									
UDCA	37 (97)	0 (0)	***<0.001***	32 (97)	0 (0)	***<0.001***	32 (97)	0 (0)	***<0.001***
Antibiotics between time points	38 (100)	0 (0)	***<0.001***	15 (46)	0 (0)	***0.002***	17 (52)	0 (0)	***<0.001***
Diet									
Exclusively breastfed	1 (3)	17 (100)	***<0.001***	1 (3)	14 (100)	***<0.001***	0 (0)	16 (100)	***<0.001***
Mixed breast/ MCT-fed	7 (18)	0 (0)	0.1	5 (15)	0 (0)	0.1	4 (12)	0 (0)	0.2
MCT-fed	30 (79)	0 (0)	***<0.001***	29 (88)	0 (0)	***<0.001***	24 (73)	0 (0)	***<0.001***
Extensively formula-fed	0 (0)	0 (0)	NA	3 (9)	0 (0)	0.2	7 (22)	0 (0)	0.05
Weaned	0 (0)	0 (0)	NA	6 (18)	2 (14)	0.7	29 (88)	16 (100)	0.2

Note: Data are n (%) unless stated otherwise. Statistical significance is represented as $p < 0.05$.
Abbreviations: BA, biliary atresia; CC, cholestatic control; HC, healthy control; KPE, Kasai portoenterostomy; MCT, medium-chain triglyceride; NA, assessment not appropriate; UDCA, ursodeoxycholic acid.



Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

FIGURE 3 Longitudinal fecal microbiota and SCFA composition comparison between BA and healthy infants. Stacked bar plots (A) showing the 10 most abundant genera in individual BA infants and age-matched healthy controls at pre-KPE, 6 weeks, 12 weeks, and 24 weeks after KPE. (B) Longitudinal Inverse Simpson diversity in BA and age-matched HC infants. Relative abundance comparison of (C) *Bifidobacterium*, (D) *Enterococcus*, (E) *Fusobacterium*, and (F) *Pseudomonas* between BA and healthy infants. Fecal (G) acetate, (H) propionate, and (I) butyrate concentrations in BA and age-matched HC infants. Only significant *p* values included in bar chart comparisons. (J) Significant correlation between *Bifidobacterium* abundance and acetate concentration at 24 weeks in BA infants. The blue line on the scatter plot indicates the line of best fit determined by linear regression. Abbreviations: BA, biliary atresia; HC, healthy control; KPE, Kasai portoenterostomy; SCFA, short-chain fatty acids.

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

24w-post-KPE," respectively. Differential abundance testing identified lower *Bifidobacterium* ($p = 0.012$) abundance by 24w-post-KPE in infants experiencing cholangitis in the preceding time period (Figure 4K). No difference in SCFA profile was revealed between cholangitis and no cholangitis groups. No difference in fecal microbiota/SCFA profiles in infants who subsequently developed an episode(s) of cholangitis was observed (data not shown).

Correlation between fecal microbiota and SCFA profiles and liver disease severity and biomarkers of liver fibrosis

Pearson correlation analysis was used to determine associations between the top 10 abundant genera and parameters of liver disease severity (PELD and TB) and liver fibrosis (AST-to-platelet ratio index and liver stiffness measurement) (Supplemental Figures S5A–D, http://links.lww.com/HC9/B57). Notably, *Bifidobacterium* inversely correlated with fibrosis markers (AST-to-platelet ratio index, $r^2 = 0.21$, $p = 0.01$; liver stiffness measurement, $r^2 = 0.34$, $p = 0.027$) (Figure 4LM) and *Blautia* inversely correlated with parameters of liver disease severity (TB, $r^2 = 0.17$, $p = 0.028$; PELD, $r^2 = 0.21$, $p = 0.025$) (Figures 4N, O; Supplemental Figure S5E, http://links.lww.com/HC9/B57). No correlation between SCFAs and disease severity or fibrosis was revealed.

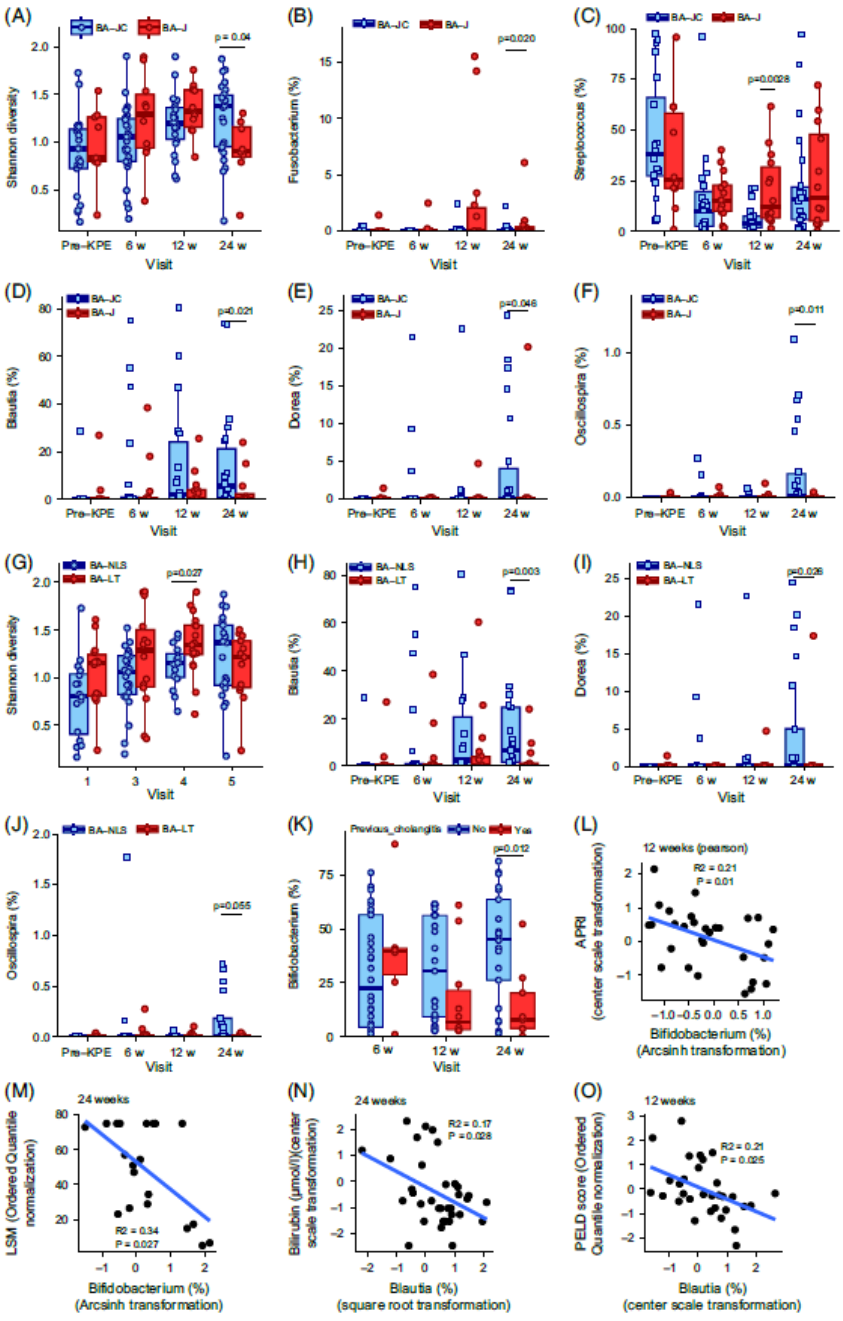
DISCUSSION

Exploring the role of the gut microbiota and its potential downstream effects in early BA could uncover a crucial modifiable pathway for early BA therapeutics. In this novel longitudinal study, pre- and post-KPE dysbiosis (increased diversity, expansion of pathobionts, and depletion of beneficial microbiota) and SCFA disturbance in BA infants were described. Importantly, dysbiosis was associated with relevant BA clinical outcomes.

A notable divergence in beta diversity, indicating the composition of the gut microbiota, was revealed when comparing the BA microbiota against the microbiota of healthy infants.[18] BA microbiota was characterized by the elevation of potentially harmful bacteria (referred to as "pathobionts"; *Enterococcus*, *Clostridium*, *Fusobacterium*, and *Pseudomonas*) and a reduction in beneficial bacteria (*Bifidobacterium*). The expansion of pathobionts has been reported in studies of adult cholestasis and cirrhosis, as well as recent BA studies, with *Bifidobacterium* depletion being increasingly recognized in BA.[8–16]

The similarity between gut microbiota profiles between pre-KPE BA and non-BA cholestatic infants suggests cholestasis as a microbiota-shaping factor. Germ-free mouse models for biliary disease have highlighted an association between the gut microbiota and cholestasis.[19,20] Gut microbiota interact with bile acids through exerting bile salt hydrolase activity, which deconjugates primary bile acids and, through 7-alpha-hydroxylation, converts primary to secondary bile acids. In turn, bile acids influence microbial composition through mechanisms such as microbial gene regulation, immune modulation, and regulation of gut barrier integrity.[5] This intricate microbiota-bile acid interplay was illustrated in an adult cirrhosis cohort, where a positive correlation between *Ruminococcaceae* (7-alpha-hydroxylating bacteria) and deoxycholic acid (secondary bile acid) was demonstrated. In BA, serum bile acid profiles have been explored as a discriminatory tool between BA and non-BA infantile cholestasis, but increasingly, microbiota studies are incorporating fecal bile acid analysis as a functional pathway.[9,10] One BA study demonstrated a positive correlation between *Enterococcus* and secondary bile acid derivatives,[10] potentially posing a functional pathway for this pathobiont in BA.

A striking feature of this study was the markedly increased alpha diversity in the gut microbiota in the early post-KPE period and its association with poor jaundice clearance and LT. Beta diversity, incorporating pathobiont expansion (*Fusobacterium* and *Streptococcus*) and depletion of beneficial microbiota "probiotics" (*Blautia*, *Dorea*, and *Oscillospira*) were also linked to these poorer outcomes. *Streptococcaceae* and *Fusobacteriaceae* have been similarly associated with cholestatic disease and disease severity in cirrhosis in adult and pediatric cohorts[7,9]; however, unraveling the extent to which liver disease is propelling dysbiosis or vice versa is challenging in the clinical setting. Human isolates of *Fusobacterium varium*[21] have been shown



Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

FIGURE 4 Fecal microbiota and SCFA associations with clinical outcomes. (A) Shannon diversity of fecal microbiota between BA-J and BA-JC groups. Relative abundance comparison of (B) *Fusobacterium*, (C) *Streptococcus*, (D) *Blautia*, (E) *Dorea*, and (F) *Oscillospira* between BA-J and BA-JC groups. (F) Shannon diversity of fecal microbiota between BA-NLS and BA-LT groups. Relative abundance comparison of (G) *Blautia*, (H) *Dorea*, (I) *Oscillospira* between BA-NLS and BA-LT groups. Correlation between (J) APRI versus *Bifidobacterium* at 12 weeks. (K) Correlation between LSM versus *Bifidobacterium* at 24 weeks. Only significant *p* values included in bar chart comparisons. (L) Correlation between Bilirubin versus *Blautia* at 12 weeks. (M) Correlation between PELD versus *Blautia* at 12 weeks. The blue line on scatter plots (I–M) indicates the line of best fit determined by linear regression. (N) *Bifidobacterium* relative abundance at 6, 12, and 24 weeks in BA infants, based on previous cholangitis. Abbreviations: BA, biliary atresia; HC, healthy control; JC, jaundice clearance; KPE, Kasai portoenterostomy; LSM, liver stiffness measurement; LT, liver transplantation; NLS, native liver survival; PELD, Pediatric End-Stage Liver Disease; SCFA, short-chain fatty acids.

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

to induce colonic mucosal erosion in mice by rectal enema, postulated as a mechanistic link in inflammatory bowel disease. Surprisingly, *Enterococcus* expansion was not associated with clinical outcomes, potentially highlighting more complex downstream interactions rather than linear genus-outcome correlation, requiring further interrogation.

The link between "probiotic" depletion and clinical outcomes is intriguing and could play a crucial role in the development of future BA therapeutics. *Blautia* inversely correlated with liver disease severity parameters. *Blautia* depletion has been described in adult cirrhosis cohorts, with its beneficial interaction with bile acid metabolism, SCFA production, and immunomodulatory potential[22] influencing host interactions. *Bifidobacterium,* a Gram-positive anerobic genus, which is dramatically depleted after KPE, although not directly associated with clinical outcomes, is inversely correlated with fibrosis markers, potentially highlighting an immune role for this genus in BA. *Bifidobacterium* is well known to confer positive health benefits to the human host, namely through its antibacterial activity,[23] immunomodulatory potential, stabilization of the intestinal barrier integrity,[24] and bile acid metabolism through its ability to secrete bile salt hydrolase.[25] Depletion of *Bifidobacterium* in infancy has been linked to numerous conditions, including necrotizing enterocolitis[26] and poorer neurodevelopmental outcomes.[27] *Bifidobacterium* play a critical role in fermenting human milk oligosaccharides (HMOs), non-digestible carbohydrates found in breastmilk; hence, a symbiotic relationship between breastmilk-*Bifidobacterium* exists.[28] Although higher microbiota diversity is often linked with a "healthy state," in early infancy, a low diverse, *Bifidobacterium*-dominant signature[28] is associated with health. Therefore, the substantial decline in breastfeeding rates following KPE likely contributes to *Bifidobacterium* depletion, enrichment of pathobionts, and higher diversity observed in BA infants.

While interactions between dietary factors, cholestasis, and the microbiota are likely to play pivotal roles in BA dysbiosis, the significant changes observed in gut microbiota diversity and composition during the early postsurgical period prompt inquiries into the potential involvement of standard postoperative pharmacotherapy. Prophylactic antibiotics are part of post-KPE protocols worldwide, although their impact on BA clinical outcomes remains uncertain. Antibiotic-associated dysbiosis has been revealed in several infant studies.[29] In our study, *Bifidobacterium* abundance was significantly reduced following cholangitis infection, raising the possibility of antibiotic-associated *Bifidobacterium* depletion.[29] Corticosteroids and choleretic administration after KPE pose uncertain clinical benefits[30] and have also been shown to influence microbial patterns.[31,32] Therefore, gaining a deeper comprehension of the interplay between BA therapies and microbiota patterns can facilitate the development of more precise mechanistic pathways and refine clinical practice.

The production of SCFAs has surfaced as a pivotal microbiota-downstream pathway. This study, to our knowledge, is the first to examine SCFAs alongside BA gut microbiota. SCFAs, including acetate, propionate, and butyrate, represent the primary end-products of the fermentation of nondigestible carbohydrates exclusively by the gut microbiota. SCFAs influence essential host interactions, such as immune regulation, gut barrier integrity maintenance, and antimicrobial properties (including lowering pH to inhibit pathogens).[33] In this study, fecal acetate concentration was significantly reduced in pre-KPE BA infants and further exacerbated after KPE. Acetate is the most abundant SCFA in healthy infants, suggesting an evolutionary contribution to infant health. *Bifidobacterium* is a dominant acetate-producing bacterium in infancy; hence, its depletion in BA potentially influences acetate depletion. Microbiota-SCFA correlation analysis did reveal a positive correlation between *Bifidobacterium* and acetate by 24w-post-KPE. The absence of a *Bifidobacterium*-acetate association at earlier time points may highlight insufficient substrate (eg, breastmilk) or overall *Bifidobacterium* abundance to exert its functional capacity.

Animal models have highlighted the immunomodulatory and antimicrobial properties of acetate, including inhibition of lipopolysaccharide-induced TNFa secretion.[34] An association between reduced stool acetate at 3 months of age[35] and subsequent allergic disease suggests a role for early acetate-mediated developmental immune priming. Hence, downstream implications of acetate depletion in BA fibro-inflammatory pathways require further interrogation. In contrast to acetate reduction, butyrate abundance

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

was higher in BA infants at early post-KPE time points. Butyrate is the preferred fuel used for colonocyte proliferation and regulates gut epithelial barrier integrity through coordinated regulation of the tight junction proteins.[36] Butyrate-producing microbiota (eg, *Ruminococcus* and *Clostridales*) are reduced in adult cirrhosis cohorts, and butyrate has been inversely correlated with portal hypertension and inflammatory mediators[37]; however, the role of butyrate in infancy has not been well-defined.

Microbiota-modulatory clinical trials are showing promise in adult liver disease. Several randomized, placebo-controlled clinical trials have revealed probiotic efficacy in various liver diseases, and clinical trials on fecal microbiota transplantation (FMT) have demonstrated safety and favorable changes in gut microbiota composition and function in alcohol-associated hepatitis and HE, with preclinical and pilot data suggesting a role for FMT in nonalcoholic liver disease and primary sclerosing cholangitis.[38] Few studies have explored the role of probiotics in BA, and only culture-dependent techniques have been used. One study demonstrated similar efficacy in cholangitis prevention between prophylactic *Lactobacillus* casei rhamnosus (probiotic) versus neomycin[39] in jaundice-free BA patients (0–3 y-old) and beneficial microbiota shifts in BA infants after LT have been demonstrated.[40] A recent position paper on the use of probiotics in preterm infants to prevent necrotizing enterocolitis supports *Bifidobacterium* prophylaxis.[41] The findings of this latest study suggest the need for additional exploration into the potential benefits of pre- and probiotic prophylactic supplementation during the early post-KPE period and in cholangitis management.

This study's strength lies in its comprehensive characterization of the fecal microbiota, coupled with paired fecal SCFA analysis in BA, using both cholestatic patients and HCs. Nonetheless, certain limitations should be acknowledged. First, while the rationale behind selecting exclusively breastfed, pharmacotherapy-naïve infants as HCs has been discussed, the divergence of dietary and pharmacotherapy characteristics between cohorts are significant confounders. Future studies could consider incorporating formula-fed, antibiotic-exposed infants alongside the "gold-standard" control arm to mitigate these confounders. Second, a non-BA CC cohort was only available at the first time point due to the challenges of obtaining CCs matched for microbiota-shaping factors at the longitudinal time points. Third, although 16S rRNA gene amplicon sequencing data provide valuable genus-level profiling in a relatively new disease area, combining it with shotgun metagenomics could offer species/strain data and potential functional signatures for future studies. Lastly, dietary data collection was limited to 3-day milk diaries, which may introduce recall bias, and extensive pharmacotherapy administration (eg, antireflux medications) was not examined due to small subcohorts.

CONCLUSIONS

The microbiota and SCFA profiles in BA represent a notable departure from healthy microbiota patterns, especially following Kasai surgery. Microbial disturbance after Kasai has been linked to clinical outcomes, highlighting the importance of scrutinizing post-Kasai protocols in mitigating their potential contributions and exploring strategies to refine these protocols, such as the incorporation of prebiotics or probiotics.

DATA AVAILABILITY STATEMENT

Data, analytical methods, and study materials are available to other researchers under specific requests. The data collected for the study, including anonymous individual participant data and the data dictionary defining each field in the set, will be made available to others for scientific purposes. The data and related documents, including analytical methods and study materials, will be available after the publication date on specific request to vjain@nhs.net, with a signed data access agreement and restriction of publication without the authors' consent.

AUTHOR CONTRIBUTIONS

Vandana Jain and Anil Dhawan conceptualized the study. Anil Dhawan supervised and mentored Vandana Jain. Vandana Jain oversaw the investigation. Vandana Jain was responsible for acquiring the resources. Vandana Jain was responsible for sample collection. Emma C. Alexander and Charlotte Burford contributed to sample collection and processing. Vandana Jain, Holly Acford-Palmer, Iliana R. Serghiou, Nancy M.Y. Teng, Konstantinos Gerasimidis, Michael Logan, and Konstantina Zafeiropoulou were involved in sample processing and analysis. Matthew J. Dalby and Raymond Kiu carried out a bioinformatic analysis. Mark Davenport contributed to acquiring resources and data interpretation. Anita Verma and Konstantinos Gerasimidis contributed to data interpretation. Matthew J. Dalby, Vandana Jain, and Charlotte Burford participated in the statistical analysis. Matthew J. Dalby was responsible for the visualizations of data. Vandana Jain and Matthew J. Dalby wrote the initial draft. Vandana Jain, Anil Dhawan, Matthew J. Dalby, and Lindsay J. Hall edited and reviewed the subsequent drafts. All authors reviewed and commented on subsequent drafts. Vandana Jain and Matthew J. Dalby accessed and verified the data.

FUNDING INFORMATION

Joint BSPGHAN/CLDF charity fund.

CONFLICTS OF INTEREST

Konstantinos Gerasimidis is on the speakers' bureau for Nutricia Danone and Nestle. Anil Dhawan consults and owns stock in BitBio and Aspect Bio. He consults for Astellas. The remaining authors have no conflicts to report.

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024

REFERENCES

1. Hartley JL, Davenport M, Kelly DA. Biliary atresia. Lancet. 2009; 374:1704–13.
2. Jain V, Burford C, Alexander EC, Sutton H, Dhawan A, Joshi D, et al. Prognostic markers at adolescence in patients requiring liver transplantation for biliary atresia in adulthood. J Hepatol. 2019;71:71–7.
3. Davis BC, Bajaj JS. The human gut microbiome in liver diseases. Semin Liver Dis. 2017;37:128–40.
4. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, et al. Host-gut microbiota metabolic interactions. Science. 2012; 336:1262–7.
5. Ridlon JM, Kang DJ, Hylemon PB, Bajaj JS. Bile acids and the gut microbiome. Curr Opin Gastroenterol. 2014;30:332–8.
6. Luo Y, Zheng S. Current concept about postoperative cholangitis in biliary atresia. World J Pediatr. 2008;4:14–9.
7. Reuter B, Bajaj JS. Microbiome: Emerging concepts in patients with chronic liver disease. Clin Liver Dis. 2020;24:493–520.
8. Tessier MEM, Cavallo L, Yeh J, Harpavat S, Hoffman KL, Petrosino JF, et al. The fecal microbiome in infants with biliary atresia associates with bile flow after Kasai portoenterostomy. J Pediatr Gastroenterol Nutr. 2020;70:789–95.
9. Wang J, Qian T, Jiang J, Yang Y, Shen Z, Huang Y, et al. Gut microbial profile in biliary atresia: A case-control study. J Gastroenterol Hepatol. 2020;35:334–42.
10. Song W, Sun L-Y, Zhu Z-J, Wei L, Qu W, Zeng Z, et al. Association of gut microbiota and metabolites with disease progression in children with biliary atresia. Front Immunol. 2021; 12:698900.
11. van Wessel D, Nomden M, Bruggink J, de Kleine R, Kurilshikov A, Verkade H, et al. Gut microbiota composition of biliary atresia patients before Kasai portoenterostomy associates with long-term outcome. J Pediatr Gastroenterol Nutr. 2021;73:485–90.
12. Jain V, Alexander EC, Burford C, Verma A, Dhawan A. Gut microbiome: A potential modifiable risk factor in biliary atresia. J Pediatr Gastroenterol Nutr. 2021;72:184–93.
13. Song W, Sun LY, Zhu ZJ. Effects of previous Kasai surgery on gut microbiota and bile acid in biliary atresia with end-stage liver disease. Front Med. 2021;8:704328.
14. Sun X, Cai Y, Dai W, Jiang W, Tang W. The difference of gut microbiome in different biliary diseases in infant before operation and the changes after operation. BMC Pediatr. 2022;22:502.
15. Zheng L, Wu Y, Gong Z, Lv Z, Xu W, Sheng Q, et al. The composition of the gut microbiota is altered in biliary atresia with cholangitis. Front Surg. 2022;9:990101.
16. Yang T, Yang S, Zhao J, Wang P, Li S, Jin Y, et al. Comprehensive analysis of gut microbiota and fecal bile acid profiles in children with biliary atresia. Front Cell Infect Microbiol. 2022;12:914247.
17. Mallick H, Rahnavard A, McIver LJ, Ma S, Zhang Y, Nguyen LH, et al. Multivariable association discovery in population-scale meta-omics studies. PLoS Comput Biol. 2021;17:e1009442.
18. World Health Organization. Breastfeeding. 2022. Accessed October 2024. https://www.who.int/health-topics/breastfeeding#tab=tab_2
19. Tabibian JH, O'hara SP, Trussoni CE, Tietz PS, Splinter PL, Mounajjed T, et al. Absence of the intestinal microbiota exacerbates hepatobiliary disease in a murine model of primary sclerosing cholangitis. Hepatology. 2016;63:185–96.
20. Schrumpf E, Kummen M, Valestrand L, Greiner TU, Holm K, Arulampalam V, et al. The gut microbiota contributes to a mouse model of spontaneous bile duct inflammation. J Hepatol. 2017; 66:382–9.
21. Ohkusa T, Okayasu I, Ogihara T, Morita K, Ogawa M, Sato N. Induction of experimental ulcerative colitis by Fusobacterium varium isolated from colonic mucosa of patients with ulcerative colitis. Gut. 2003;52:79–83.
22. Liu X, Mao B, Gu J, Wu J, Cui S, Wang G, et al. Blautia—A new functional genus with potential probiotic properties? Gut Microbes. 2021;13:1875796.
23. Inturri R, Stivala A, Fumeri PM, Blandino G. Growth and adhesion to HT-29 cells inhibition of Gram-negatives by Bifidobacterium longum BB536 e Lactobacillus rhamnosus HN001 alone and in combination. Eur Rev Med Pharmacol Sci. 2016;20:4943–9.
24. Lim HJ, Shin HS. Antimicrobial and immunomodulatory effects of Bifidobacterium strains: A review. J Microbiol Biotechnol. 2020; 30:1793–800.
25. Long SL, Gahan CGM, Joyce SA. Interactions between gut bacteria and bile in health and disease. Mol Aspect Med. 2017; 56:54–65.
26. Thanert R, Keen EC, Dantas G, Warner BB, Tarr PI. Necrotizing enterocolitis and the microbiome: Current status and future directions. J Infect Dis. 2021;223(12 suppl 2):s257–S263.
27. Beghetti I, Barone M, Turroni S, Biagi E, Sansavini A, Brigidi P, et al. Early-life gut microbiota and neurodevelopment in preterm infants: Any role for Bifidobacterium? Eur J Pediatr. 2022;181: 1773–7.
28. Ma J, Li Z, Zhang W, Zhang C, Zhang Y, Mei H, et al. Comparison of gut microbiota in exclusively breast-fed and formula-fed babies: A study of 91 term infants. Sci Rep. 2020;10: 15792.
29. Tanaka S, Kobayashi T, Songjinda P, Tateyama A, Tsubouchi M, Kiyohara C, et al. Influence of antibiotic exposure in the early postnatal period on the development of intestinal microbiota. FEMS Immunol Med Microbiol. 2009;56:80–7.
30. Burns J, Davenport M. Adjuvant treatments for biliary atresia. Transl Pediatr. 2020;9:253–65.
31. Chen W, Wei Y, Xiong A, Li Y, Guan H, Wang Q, et al. Comprehensive analysis of serum and fecal bile acid profiles and interaction with gut microbiota in primary biliary cholangitis. Clin Rev Allergy Immunol. 2020;58:25–38.
32. Hart L, Farbod Y, Szamosi JC, Yamamoto M, Britz-McKibbin P, Halgren C, et al. Effect of exclusive enteral nutrition and corticosteroid induction therapy on the gut microbiota of pediatric patients with inflammatory bowel disease. Nutrients. 2020;12:1691.
33. Tan J, McKenzie C, Potamitis M, Thorburn AN, Mackay CR, Macia L. The role of short-chain fatty acids in health and disease. Adv Immunol. 2014;121:91–119.
34. Masui R, Sasaki M, Funaki Y, Ogasawara N, Mizuno M, Iida A, et al. G protein-coupled receptor 43 moderates gut inflammation through cytokine regulation from mononuclear cells. Inflamm Bowel Dis. 2013;19:2848–56.
35. Arrieta M-C, Stiemsma LT, Dimitriu PA, Thorson L, Russell S, Yurist-Doutsch S, et al. Early infancy microbial and metabolic alterations affect risk of childhood asthma. Sci Transl Med. 2015; 7:307ra152.
36. Liu B, Qian J, Wang Q, Wang F, Ma Z, Qiao Y. Butyrate protects rat liver against total hepatic ischemia reperfusion injury with bowel congestion. PLoS One. 2014;9:e106184.
37. Juanola O, Ferrusquía-Acosta J, García-Villalba R, Zapater P, Magaz M, Marín A, et al. Circulating levels of butyrate are inversely related to portal hypertension, endotoxemia, and systemic inflammation in patients with cirrhosis. FASEB J. 2019;33:11595–605.
38. Bajaj JS, Ng SC, Schnabl B. Promises of microbiome-based therapies. J Hepatol. 2022;76:1379–91.
39. Lien T-H, Bu L-N, Wu J-F, Chen HL, Chen AC, Lai MW, et al. Use of Lactobacillus casei rhamnosus to prevent cholangitis in biliary atresia after Kasai operation. J Pediatr Gastroenterol Nutr. 2015;60:654–8.
40. Waldner B, Aldrian D, Zöggeler T, Oberacher H, Oberhuber R, Schneeberger S, et al. The influence of liver transplantation on

the interplay between gut microbiome and bile acid homeostasis in children with biliary atresia. Hepatol Commun. 2023;7:e0151.
41. van den Akker CHP, van Goudoever JB, Shamir R, Domellöf M, Embleton ND, Hojsak I, et al. Probiotics and preterm infants: A position paper by the European Society for Paediatric Gastroenterology Hepatology and Nutrition Committee on Nutrition and the European Society for Paediatric Gastroenterology Hepatology and Nutrition Working Group for Probiotics and Prebiotics. J Pediatr Gastroenterol Nutr. 2020;70:664–80.

How to cite this article: Jain V, Dalby MJ, Alexander EC, Burford C, Acford-Palmer H, Serghiou IR, et al. Association of gut microbiota and gut metabolites and adverse outcomes in biliary atresia: A longitudinal prospective study. Hepatol Commun. 2024;8:e0550. https://doi.org/10.1097/HC9.0000000000000550

Downloaded from http://journals.lww.com/hepcomm by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC1y0abggQZXdtwnfKZBYtws= on 10/29/2024