The preterm gut microbiome in health and disease – metagenomic approaches for microbial diagnostics

Cristina Alcon-Giner

A thesis submitted for the degree of Doctor of Philosophy

University of East Anglia
Quadram Institute Bioscience

Norwich, United Kingdom
November 2019

Contents

LIST OF FIGURES

LIST OF TABLES

Abstract

Premature infants, born before 37 weeks of gestation, represent an important patient group at risk of developing numerous diseases such as necrotising enterocolitis and bacterial sepsis. This risk is correlated with changes in the preterm gut microbiome, which is influenced by multiple post-natal factors including gut immaturity, C-section delivery, exposure to antibiotics and difficulties in establishing breastfeeding. To reduce the risk of disease development in premature infants and reduce colonisation of bacterial pathogens, oral supplementation with beneficial bacteria including *Lactobacillus* and *Bifidobacterium* is used.

Before widespread uptake of this intervention, there is a pressing need to understand the impact of post-natal factors and how probiotic supplementation may modulate the preterm microbiome. In this thesis, I comprehensively examined the bacterial colonisation patterns of the preterm microbiome in health and disease using next-generation sequencing approaches. I also evaluated if probiotic supplementation can modify the gut microbiome in premature infants.

Short- and long-read metagenomics sequencing was used, complemented with culturing and phenotypic testing. A 16S rRNA microbiome profiling pipeline was optimised to characterise faecal samples from premature infants with and without probiotic supplementation. The methods developed provided the foundation for a large-scale clinical study (BAMBI) which sought to explore the impact of probiotic supplementation on the preterm gut microbiome in 233 infants. A subset of these faecal samples (96 samples) were examined using shotgun metagenomics to study the gut bacterial reservoir of antibiotic resistance genes (the 'resistome') after antibiotic treatment and to elucidate whether probiotic supplementation impacts the prevalence of AMR genes. Finally, MinION Nanopore sequencing was used to rapidly profile faecal samples from critically ill premature infants.

Overall, this multidisciplinary work provides novel insights into the preterm gut microbiome in health and disease, emphasises the protective role of probiotic supplementation when administered to premature infants, and evaluates whether rapid sequencing approaches can be applied for prompt microbial diagnostics.

Declaration

The work presented in this thesis is the result of my own work during my PhD. Each results chapter includes a statement detailing author's contributions to acknowledge and accredit work done in collaboration. I have also specified the work done in collaboration in the method section of each chapter.

Cristina Alcon Giner
November 2019
Norwich, UK

Acknowledgements

I would like to thank my primary supervisor Dr Lindsay Hall (QIB) for the extraordinary support and mentorship I have had throughout my PhD; this has been pivotal to the successful PhD experience I have had. Huge thanks Lindsay for the opportunity to be in your team, your supportive and enthusiastic mentorship, your creative ideas, and your constant dedication to proofreading my work. I would also like to thank my secondary supervisor Dr John Wain for his research ideas and input at the PhD review meetings, which have been key to rapidly advance my research work.

Many thanks to all present and former colleagues in the Hall lab for enriching daily my PhD journey. I have wonderful memories of shared dinners, lab team building days, and outreach and public engagement activities. Mel, Zoe, Ian, Cho-Zin, Raymond, Shab, Suparna, Magda, Lukas, Jenni R and Matthew – thanks a lot for sharing your knowledge and research experience with me, I have learnt a lot from each of you. Holly, Jennifer, Lisa, Sarah, Shannah, Iliana and Charlotte – it has been lovely to share day-to-day experiences in the lab. All of you have made the Hall lab a fantastic environment where working, learning and fun happens every day.

I would also like to thanks my collaborators who have contributed tremendously to my research and without who's input I would not have been able to complete my work; Dr Paul Clarke, Mrs Karen Few, Mrs Hayley London, Mrs Kate Lloyd (Norfolk and Norwich University Hospital); Dr Gustav Belteki (Rosie Hospital); Prof. J. Simon Kroll, Dr Alex Shaw, Dr Kathleen Sim (ICL); Dr Jonathan Swann, Dr Fahmina Fardus (Imperial College); Dr Richard Leggett, Dr Darren Heavens (Earlham Institute); Dr Matt D. Clark (Natural History Museum); Dr Will Rowe (University of Birmingham); Dr Linda Harvey and Dr Mark Wilkinson (Biorepository).

I would also like to recognise the time the following individuals have given to me at some point throughout my PhD, either via specific training or informal chats: Dr Kevin Hughes (Mars); Dr Aimee Parker, D. Dimitrios Latousakis, Dr Antonella Mechini, Dr Melinda Mayer (QIB); Mr Gary Wortley (JIC); Kate Conway and Mark Meehan (Graduate Studies Office).

Finally, I would also like to acknowledge the funding bodies that supported my research work: the BBSRC Doctoral Training Partnership (DTP) programme, and the Wellcome Trust New Investigator Award awarded to Dr Lindsay Hall.

General introduction

1.1 Gut microbiome: importance and advances in the field

The microbes that reside in the human gastrointestinal (GI) tract comprise a vast ecosystem of organisms including bacteria, fungi, viruses and protozoans, estimated to weigh ~2 kg[1, 2]. The gut microbiome is often described as a "virtual organ"[3] comprised of ~10^{14} microbial cells that work in concert with the host and are able to promote health, but may sometimes cause disease. The gut microbiome constitutes a complex and diverse web of microorganisms which require an array of different and bespoke techniques for their study.

In the last decade, the study of the human gut microbiome has become a rapidly moving field due to advances in new 'omic' technologies (in particular next-generation sequencing and metabolite profiling). These approaches allow the identification and functional characterisation of complex microbial communities without the need to cultivate individual bacteria[4]. Microbiome research is providing new insights into the associations between bacterial composition and health or disease. Indeed, large-scale studies have been completed or are underway that are investigating how beneficial bacterial populations may be associated with human wellbeing by facilitating nutrient intake[5], supporting the host's immune system[6], or providing antimicrobial protection[7] against other bacterial pathogens. These studies can be complemented by others exploring how microbiome disturbances caused by exposure to antibiotics or different diets, are linked to incidence of chronic diseases including; ulcerative colitis[8], obesity[9], autoimmune conditions[10] and infection with multi-drug resistant microbes[11].

There are significant new developments underway to harness the therapeutic properties of the gut microbiome. Microbiota therapies (MT) aim to alter the bacterial community gut composition of the individual, whilst at the same time

providing improvements in health outcomes[12]. Many of these therapies are currently at an early stage of development, but their future has an exciting potential of influencing both health and disease. Examples of these innovations include untargeted approaches such as faecal microbiota transplantation (FMT) or more targeted ones using oral supplementation with beneficial bacteria, commonly known as "probiotics"[13]. FMT has become very effective for treating recurrent *Clostridioides difficile* infections; the success of this therapy has been shown to be dependent on the microbial diversity of the donor, which should be able to provide colonisation resistance against overgrowth of *C. difficile* (treatment is now available on the UK NHS)[14]. The use of empiric probiotic supplementation in adults, on the other hand, appears to be more challenging when crediting to health claims. This is likely due to two main reasons: (i) the vast diversity of "healthy" microbiomes among study-individuals making colonisation of supplemented strains difficult and (ii) important differences between 'beneficial' traits of probiotic strains[15]. Further characterisation of underlying mechanisms of action of these microbiota therapies is needed in conjunction with large-scale clinical studies to determine health outcomes. This will help integrate these therapies into mainstream clinical practice and, importantly, will reduce variability in patient response.

The intestinal gut microbiome in early life and its role in maintaining health

Early life represents a key developmental window when the foundations for life-long health are laid down. Crucially, colonisation of the gut with beneficial bacterial pioneers (*e.g*. *Bifidobacterium* and *Lactobacillus*), contributes to optimal health and immune programming in the newborn[16]. The newborn gut is believed to be sterile[17], and then become rapidly colonised during the first hours and days of life.

A naturally-born and breastfed infant is considered to be the 'gold standard' for a healthy infant gut microbiome. Immediately after birth, the intestine of the newborn infant is colonised by bacteria residing in the mother's birth canal, the gut, and the environment, a mixed inoculum of facultative anaerobic bacteria (*Enterobacter*, *Enterococcus*, *Staphylococcus* and *Streptococcus*)[18] able to grow in an environment low in oxygen. As oxygen diminishes within the gut other obligate anaerobic bacteria thrive (*e.g. Bifidobacterium*, *Bacteroides* and *Clostridium*)[19]. These microbial pioneers probably represent the most important bacterial inoculum for the infant gut colonization. Oral feeding (by breast milk or formula) is key to stimulating the growth of these bacterial colonisers during the first four month of life. Then with weaning to solid foods (four to six months of age) the infant gut microbiome becomes exposed to additional stimuli and by about eighteen months to three years the infant gut bacterial ecosystem is established representing the infant's microbiome signature for life[20, 21].

Among these initial colonisers, members of the Bifidobacteriaceae family are important commensals which typically dominate the breastfed infant gut, reaching in some cases almost full dominance at 99%[22]. This is due, in part, to their ability to digest certain dietary components of breastmilk and produce acetate and bacteriocins which directly inhibit the growth of other bacteria[23]. Breastmilk represents a continued source of complex sugars called human milk oligosaccharides (HMOs) which act as a specialised nutrient source for the first bacterial communities such as Bifidobacteriaceae. Studies have demonstrated that there is a close association between gut bifidobacterial species and their ability to degrade the HMOs present in breastmilk[24]. This close relationship between gut bacteria and nutrient source in exclusively breastfed infants has been shown to provide beneficial health outcomes for these infants; for example, clinical studies demonstrated protective effects against diarrhoea and less incidence of long-term diabetes and obesity when compared to formula fed infants[25].

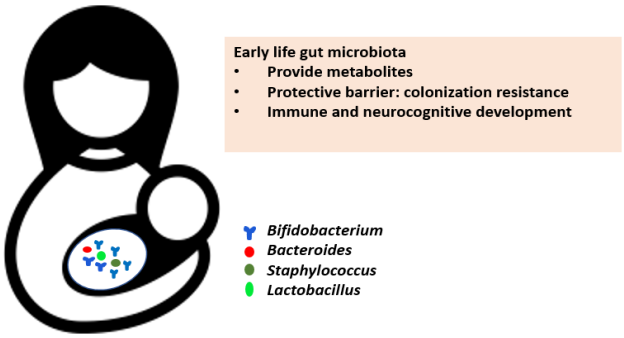


Figure 1 Microbiome in a healthy infant

The early life gut microbiome plays three main roles during infancy; (i) break-down of dietary components which results in synthesis of vitamins and the production of metabolites such as short chain fatty acids (SCFAs), (ii) prevention of pathogen colonisation through a mechanism called colonisation resistance[26], and (iii) contribution to the maturation of the infant immune and neurocognitive development[27] (Figure 1).

Vitamin synthesis and metabolite production

One of the ways the gut microbiome has been associated with the development of the infant is by synthesising important vitamins during the postnatal period. Intestinal bacteria can provide vitamin K which is necessary to synthesise certain coagulation factors in the liver, or group B vitamins such as folate which is required for cell division[28]. These vitamins in adults would be acquired by a diverse diet, but in small infants receiving a milk-based diet they can remain at low levels. Bacterial fermentation of carbohydrate products in the gut can result in production of SCFAs (e.g. butyrate, propionate and acetate) which can be used as energy by intestinal cells, and also transported from the intestinal lumen into the bloodstream where they are taken up by organs and act as substrates or signalling molecules. Butyrate for example is a SCFA which has been studied mostly for its interaction with colonocytes, improving mucosal and barrier integrity[29]. Acetate is absorbed in the liver where it is used as both an energy source and for cholesterol synthesis. Propionate has been shown to act as a precursor of the gluconeogenesis pathway in the liver[30].

Colonisation resistance

The commensal early life gut microbiome forms a stable bacterial ecosystem which can provide protection against the colonisation of potential bacterial pathogens. This phenomenon is known as "colonisation resistance" and has an important role for not only preventing potential infections, but in fighting against them as well[31]. There are several ways in which commensal bacteria can confer colonisation resistance including; (i) killing other bacteria by producing small peptides called bacteriocins which inhibit the growth of other bacteria, or by (ii) competition for nutrients and specialising in utilising a unique "nutritional niche" which can help gut colonisation. In early life, *Bifidobacterium* taxa plays an important role in conferring colonisation resistance as described later.

Maturation of infant immune system

The maturation of the infant immune system[32] develops at the same time as the infant gut microbial colonisation. Bifidobacteriaceae taxa have been associated with a role in programming the infant immune system[33]. Alterations in this process, are considered to be potential determinants of health outcomes later in life. Recent studies revealed reduced gut *Bifidobacterium* levels in 3-month-old infants were associated with higher incidence of atopy at 2 years of age, or asthma at 4 years of age[33]. *In vitro* studies have already demonstrated evidence of cross-talk among this taxa and immune cells. *Bifidobacterium* strains have been shown to induce cytokine production by peripheral blood mononuclear cells (PBMCs)[34, 35]. These studies concluded that cytokine production was strain-dependent, and careful selection should be taking into consideration when using them for therapeutic use.

Experimental work using murine models has shown that secondary bacterial metabolites produced by commensal bacteria (i.e. butyrate, propionate and acetate) enhance T-cell differentiation which can have effects on decreasing tissue inflammation and enhancing immunity[36]. Furthermore, some commensal members of the gut microbiome (i.e. *Lactobacillus, Bifidobacterium, Bacteroides*, and

Clostridium) have been shown to influence the production of non-inflammatory IgA, which contributes to overall gut homestasis[37].

As highlighted in this section, this natural process of microbial colonisation of the infant gut is believed to be instrumental in influencing the composition of the adult gut microbiome. Alterations in this process, are considered to be potential determinants of health outcomes later in life. Some researchers refer to this process as a "window of opportunity", whereby the timing of colonisation by these early life bacterial pioneers is crucial. Importantly, the process of microbial colonisation is known to be dependent on several intrinsic and extrinsic factors including; gestational age (*i.e.* time when the infant is born), mode of delivery (*i.e.* vaginal vs C-section delivery), type of feeding (*i.e.* breastmilk vs formula), and environmental exposures (*i.e.* antibiotic treatment).

1.2 Premature infants and its unique conditions which shape their gut microbiome

Every year approximately 15 million infants are born premature, *i.e.* before 37 weeks of gestation, accounting for 1 in 10 live births[38]. Premature infants have an underdeveloped immune system and are administered numerous courses of antibiotics to prevent bacterial infections. Compared to term infants, premature infants are exposed to many more pre- and post-natal factors which disrupt the natural establishment of the early life gut microbiome. Premature infants spend long periods in the neonatal intensive care unit (NICU) and they received multiple clinical procedures (intubation, mechanical ventilation and vascular access) which increases the incidence of acquiring catheter-related bloodstream infections[39]. These factors contribute to an altered profile of the gut microbial ecosystem with an increase in relative abundance of pathobionts (i.e. *Staphylococcus, Enterobacter*, *Klebsiella* or *Escherichia*) which are gut resident microbes with pathogenic potential, and decrease abundance of commensal bacteria (i.e.

Bifidobacterium). Clinical systematic reviews have linked an increased risk of infectious diseases during the NICU stay, and later life health problems such as asthma or eczema in premature infants[40, 41].

Elucidating the influence of post-natal factors involved in the developing preterm gut microbiome (Figure 2) is one of the main aims of this thesis. The following introductory pages summarise the latest research findings examining the impact of mode of delivery, gestational age, antibiotics, and diet on the preterm gut microbiome. The two most common diseases associated with disturbances of the preterm gut microbiome during their residency in NICU are necrotising enterocolitis (NEC)[42] and bacterial sepsis[43], which are discussed in more detail below. Samples from premature infants suffering from these diseases were used in this research work (Chapter 4). Microbiota therapies are highlighted, in particular those that use oral supplementation with early life bacterial members that favour the establishment of beneficial communities in the at-risk preterm gut microbiome ecosystem. Finally, as this research work was conducted using sequencing and metabolomic approaches there is a section summarising the advantages, as well as limitations, for microbial profiling and diagnostics.

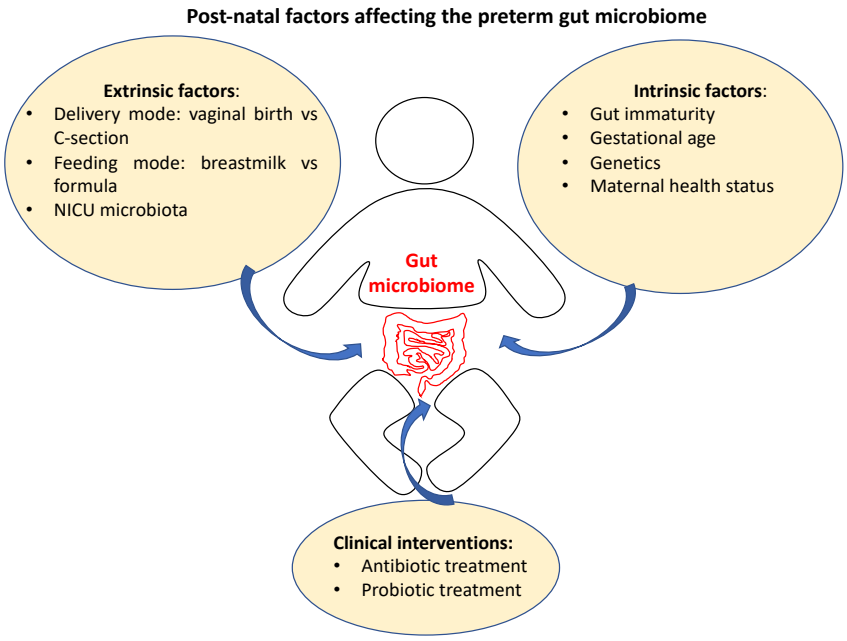


Figure 2 Post-natal factors affecting the preterm gut microbiome

Post-natal factors that have been shown to affect the establishment of the preterm gut microbiome.

1.2.1 Delivery mode and the gut microbiome of premature infants

Premature infants are often born via C-section[44], which prevents exposure to beneficial vaginal (and gut) microbes (e.g. *Bifidobacterium* and *Lactobacillus*) involved in the stimulation of the infants' immune system[45].

To date, most research studies on premature infants agree that their gut is mainly colonised by pathobionts such as *Staphylococcus*, *Escherichia*, *Klebsiella*, *Enterococcus* and *Enterobacter*[46, 47], irrespective of their mode of delivery (vaginal vs. C-section). Interestingly, a recent study performed in premature infants associated a higher relative abundance of *Bifidobacterium* in those born vaginally, exclusively breastfed and not exposed to antibiotics[48], which resemble the post-natal conditions of a full-term infant.

The findings of these studies suggest that delivery mode may not have such an influential role on the preterm gut microbiome when compared to term infants.

1.2.2 Gestational age and the effect on the preterm microbiome development

Gestational age has been described as one of the main factors influencing the gut microbiome development in early life. Korpella and colleagues studied the effect of gestational age on the preterm gut microbiome throughout the first 60 days of life[49]. They described the preterm microbiome as being typically dominated by four main genera; *Staphylococcus*, *Enterococcus*, *Enterobacter* and *Bifidobacterium,* representing >50% relative abundance in a given sample. They found that infants often switched from one pattern of microbiome to another as postmenstrual age increased. The microbiome of the premature infant progressed from *Staphylococcus*-*Enterococcus*-dominated composition (at 25- and 30-weeks postmenstrual age) to *Enterobacter (*peaking at 35 weeks postmenstrual age) and finally towards *Bifidobacterium*-dominated microbiome, which began to develop gradually after 30 weeks postmenstrual age. This study concluded that the microbiome development in the extremely low birth weight premature group (<28 gestation weeks) appeared to lag behind the moderate premature group (≥28 gestation weeks).

Maturity, indicated by postmenstrual age, is a major determinant of the preterm microbiome development, and a key factor which influences the ability of *Bifidobacterium* to reach dominance. More research investigating new interventions to improve the microbial colonisation in extremely premature infants (<28 gestation weeks) is needed.

1.2.3 Antibiotic treatment and the preterm gut microbiome

Antibiotics are the most commonly prescribed drugs in premature infants. The exposure rate of premature infants to antibiotics is in the range of 75-84%[50], and higher still for very low birth weight (VLBW) infants (<1.5 Kg). Even though antibiotic therapy is necessary when an infant suffers from a bacterial infection, in many cases antibiotics are prescribed at a very early stage before the infant

presents any sign of infection or before there is any evidence from microbiological analysis[51].

The type of antibiotic treatment given to premature infants differs widely. Most NICUs use benzylpenicillin and gentamicin as the first-choice antibiotics for empirical treatment. This combination of antibiotics should protect the premature infant from the majority of Gram-positive and Gram-negative pathogens. If the premature infant presents no signs of improvement, after a short period of time, other antibiotics are used such as cefotaxime (when there is suspicion of Gram-negative infection), vancomycin (targets Gram-positive bacteria) and metronidazole (if there is suspicion of necrotising enterocolitis). Table 1 summarises the variety of antibiotics used in the NICUs which have contributed to this research work.

Table 1 Antibiotics used in NICU, activity spectrum and mode of action

Antibiotic	Activity spectrum	Mode of action	Antibiotic Class
Benzyl penicillin	Narrow	Binds to the enzymes (transpeptidases) responsible for assembling the principal components of the cell wall (peptidoglycans). As a result, the cell wall formation process is affected and the bacterium dies. Targets most Gram-positive and a few Gram-negative bacteria.	β-lactam
Gentamicin	Broad	Binds to the 30S subunit of the bacterial ribosome, disrupting protein synthesis. Targets mostly Gram-negative bacteria and Staphylococcus as Gram-positive.	Aminoglycosides
Cefotaxime	Broad	Binds to the enzymes (transpeptidases) responsible for assembly the principal components of the cell wall (peptidoglycans). As a result, the cell wall formation process is affected, and the bacterium quickly dies. Targets numerous Gram–positive and Gram-negative bacteria.	β-lactam
Flucloxacillin	Narrow	Binds to the enzymes (transpeptidases) responsible for assembling the principal components of the cell wall (peptidoglycans). As a result, the cell wall formation process is affected, and the bacterium quickly dies. Targets most Gram-positive bacteria.	β-lactam
Vancomycin	Narrow	Inhibits the synthesis of peptidoglycan. In addition, vancomycin alters cell-membrane permeability and RNA synthesis. Targets Gram-positive bacteria.	Glycopeptide
Metronidazole	Narrow	Inhibits nucleic acid synthesis thus disrupting the DNA of bacterial cells. It works in a reduced form, which take place in anaerobic bacteria. Targets Gram-negative anaerobic and Gram-positive anaerobic bacteria.	Nitroimidazole

During the first weeks of life most antibiotics given to premature infants are administered intravenously. As the infant matures and starts feeding, oral administration is used. In general, oral administration presents a moderate absorption and bioavailability when compared to intravenous administration[52]. The majority of the infants recruited in this research work, received antibiotics intravenously due to their poor feeding skills.

Premature infants often receive antibiotic empirically, as early as the first day of life. This coincides with the initial stages of microbial gut colonisation. Disruption of this initial colonisation caused by antibiotic therapy has been associated with higher probability of premature infants developing eczema[53] and asthma[54]. Long antibiotic treatment, particularly broad-spectrum antibiotics, applies a selective pressure on the gut microbial community that diminishes colonisation resistance and favours the growth of antibiotic-resistant bacteria.

Antibiotics do not discriminate between commensal or pathogenic bacteria, the immediate effect after treatment is an overall decrease in gut microbiome diversity[55]. The gut microbiome establishes a complex network of co-dependence[56] where bacteria are producing and exchanging secondary metabolites. Therefore, the effect of antibiotics on the gut microbiome needs to be evaluated considering the gut bacteria populations as a whole entity, and factors such as mode of action of the antibiotic (broad or narrow spectrum), length of antibiotic treatment, and route of administration (oral or intravenous) will differentially impact the gut microbiome community.

Antibiotics might prevent susceptible pathogens from colonising the gut, but at the same time they might benefit the colonisation of bacterial strains that are antimicrobial resistant (AMR). Recent studies in adults have shown that the human gut microbiome acts as a huge reservoir of antibiotic resistance genes[57] referred to as the resistome[58]. How stable the gut resistome is throughout life, or how easily this reservoir of antibiotic resistance genes could be mobilised through horizontal gene transfer to dangerous pathogens are important questions that warrant further research[59].

Within the context of the resistome, more work has been done in adults than in premature infants. In 2014 Lu and colleagues studied the gut resistome by analysing the faeces of 124 volunteers from different age groups and showed that

the resistome is accumulative with age[60]. More recently, Gibson and colleagues published the first work describing the preterm gut resistome by analysing faecal samples taken during the first three months of life[55]. They correlated an increase of AMR genes with antibiotic treatment, but interestingly they also found an increase of AMR genes with no known activity against the antibiotic(s) administered. However, it is difficult to make specific conclusions associated with specific antibiotic treatments as the control group used in this study only received a short antibiotic treatment, and the time points of the study are not equally distributed amongst the babies.

Although antibiotic-induced bacterial alterations in the premature gut microbiome are only starting to be understood, there is a clear agreement that antibiotic treatment enhances the growth of multidrug resistant bacteria within gut-associated communities. There are still important gaps in our understanding of how this disruption can: (i) enhance or reduce the gut resistome as the infant ages, (ii) predispose premature infants to certain diseases/infections, or (iii) be modified if oral probiotic supplementation is used. The work presented in Chapter 3 contributes to our understanding of some of these questions.

1.2.4 Effect of diet on the preterm gut microbiome

Appropriate neonatal feeding is essential for enhancing growth, immune development and health in premature infants. Diet is considered to be one of the most influential factors at determining the composition and diversity of the neonatal gut microbiome. During the first week of life, premature infants often receive a mixture of parental nutrition (intravenously), and enteral feeding by mouth. As soon as it is possible, parental nutrition is discontinued, and infants are only fed by mouth with maternal expressed breast milk, or donor breastmilk or formula milk as alternative options in that order of priority.

Expressed breast milk (EBM) represents the preferred option for feeding premature infants due to its unique nutritional content (*i.e.* human milk oligosaccharides, lactoferrin, and beneficial bacteria) and elevated content of immunological factors (*i.e.* immunoglobulins, cytokines and growth factors), which provide protection to the infant against a range of diseases[61]. Breastmilk favours the growth of specific bacterial groups such as *Bifidobacterium* and have been shown to decrease the diversity of the gut microbial community[62].

In the absence of breastmilk, premature infants are fed with donor breast milk if available. Donor breast milk is normally fortified with protein, calcium, phosphorus, and vitamin D; premature infants are not able to take large amounts of milk, and so the fortification helps assure that they achieve adequate growth. However, this milk is often pasteurised which leads to a reduction of many beneficial bioactive components such as IgA, lactoferrin, post-natal or water-soluble vitamins[63].

If breastmilk or donor breastmilk are not available, premature infants receive formula milk. A recent study by Quigley and colleagues[64] on 1070 premature infants highlighted that infants fed with formula milk presented twice as high a risk of developing necrotising enterocolitis (described in more detail in the following section) than donor breastmilk-fed infants, despite the faster rates of growth observed. A greater variety of premature gut microbiomes from formula-fed infants has been observed, probably explained by the susceptibility of premature infants to other post-natal factors such as gestational age or antibiotic treatment[65].

1.3 Diseases associated with gut microbiome alterations in the premature infant

As described in the sections above, premature infants are predisposed to have altered bacterial colonisation patterns, which has been linked to an increased risk of developing necrotising enterocolitis (NEC) and sepsis.

NEC constitutes one of the most devastating diseases for premature infants. It is estimated that ~ 12% of premature infants <1500 g will develop NEC, and 30% of those will die[66]. NEC is a multifactorial disease, and its pathogenesis is not fully understood. It is believed that the pathological NEC process is characterised by a lack of beneficial commensal microbes in the gut, overall reduced bacterial diversity and colonisation resistance; all of which allows overgrowth of pathogenic bacteria[67]. NEC commences by an uncontrolled intestinal inflammation, induced by a presence in the gut of pathogenic bacteria linked to gut immaturity, which can lead to tissue necrosis, gut perforation and, if not controlled, sepsis[68]. *Klebsiella pneumoniae* is one the most frequent pathogenic bacteria associated with NEC-pathogenesis in premature infants[69]; in animal model studies this genus favours Paneth cell depletion[70]. Pathogenic *Shigella,* found when the infant received extended antibiotic courses, was also shown to induce NEC in premature infants[67]. Interestingly, some commensal bacterial members such as *Clostridium perfringens*, *Clostridium butyricum* and *Escherichia coli* have also been isolated during NEC outbreaks from stools of premature infants[71, 72].

Neonatal sepsis (*i.e.* bacterial blood stream infections), are the leading cause of morbidity and mortality in premature infants, accounting for 6-10% in premature infants born at 34–37 of gestational age[73]. Neonatal sepsis is classified into two groups based on the time of presentation: early-onset sepsis (EOS, refers to sepsis in premature infants presented during the first 72 hours of life) and late-onset sepsis (LOS, refer to a sepsis episode after 72 hours). Amongst the Gram positive bacterial pathogens, Coagulase Negative *Staphylococcus aureus* (CoNS) represents

48% of all bacterial infections, followed by *S. aureus*, *Enterococcus species* and Group B *Streptococcus*[74]. Amongst the Gram negative bacteria, *Klebsiella spp.* stands out [75], while *Candida albicans* is reported to be isolated in the majority of fungal infections[76].

Considering the high incidence and mortality of these diseases in the premature infant, it is crucial to be able to rapidly diagnose them. To date, the gold standard used for diagnostics in reference laboratories are culture-proven techniques. This approach requires a minimum of 24-36 hours to obtain first results, and on occasions it is not always easy to culture the pathogenic bacteria. In this thesis the real-time MinION nanopore sequencer was used to rapidly profile faecal samples from premature infants suffering from NEC and sepsis, in order to diagnose the causative bacterial pathogen and characterise its antibiotic resistance profile (Chapter 4).

1.3.1 Microbiota therapy to modulate the preterm gut microbiome

In the last twenty years, many randomised clinical trials and systematic reviews have demonstrated that routine oral supplementation of premature infants with early life commensal bacteria (*i.e. Bifidobacterium* or *Lactobacillus*) reduces the incidence of NEC and sepsis[77, 78]. It is interesting to highlight that these studies agree that the efficacy of the oral supplementation is specific to the bacterial strain used (commonly named as probiotic). Recent work is starting to relate their beneficial effects to the infant's feeding regime[78]; a meta-analysis review on twenty-five probiotic trials on premature infants confirmed that the beneficial health outcomes provided by the probiotics, were more elevated in exclusively breastfed infants, and in those where the supplementation contained more than one bacterial species.

While large scale clinical studies have demonstrated the potential of probiotics to reduce NEC and the incidence of sepsis, there is a requirement to accompany these studies with in-depth longitudinal profiling to determine the impact of this type of supplementation on the overall gut microbiome composition. Sequencing studies will allow to answer important questions; (i) did the supplemented strain colonise the preterm gut? (ii) what was the impact of the supplementation on the remaining gut bacterial populations? and (iii) for how long the supplemented strain colonised the gut? This data will be crucial to complement clinical trials if robust guidelines are to be introduced for implementing probiotic supplementation as routine clinical care for premature infants. The work presented in this thesis (Chapter 2) aims to examine the effects of Infloran, a widely used probiotic supplementation, in the preterm gut microbiome and uses sequencing and metabolomic approaches to evaluate this.

1.4 Use of metagenomics techniques for characterizing preterm gut microbiome disturbances

Metagenomics approaches have been instrumental in obtaining a better understanding of the microbial diversity present in a sample. This technology utilises directly the genetic material present in an environmental sample (*e.g.* faecal sample) without the need for culturing. Ongoing studies indicate that metagenomic approaches will transform clinical microbiology; which is also linked to the rapid drop in the cost of sequencing and turnaround time now making these technologies viable options for diagnosis in reference laboratories.

Several metagenomics approaches can be used to analyse an environmental sample including: (i) 16S rRNA gene sequencing and shotgun metagenomics (with potential application in microbial diagnostics), and (ii) functional metagenomics. Each technique has advantages and limitations, and it is important to understand how they work in order to use them efficiently.

1.4.1 16S rRNA gene sequencing

The 16S rRNA gene is the gold standard taxonomic marker for identifying bacteria in a metagenomics sample. Sequencing of 16S rRNA PCR amplicons has become a well-established and cost-effective method for profiling the bacteria community profile present in a metagenomics sample. Due to its low cost, it is normally used in large-scale studies where shotgun sequencing is not affordable.

Previous studies examining the gut microbiome using 16S rRNA gene sequencing have highlighted the influence of the DNA extraction method in the representative profile of the bacterial community obtained[79]. The 16S rRNA gene encodes nine hypervariable regions (*i.e.* V1 to V9) known to offer specificity at genus level between most bacterial groups. These hypervariable regions are surrounded by highly conserved regions in most bacterial populations, making the 16S rRNA gene an ideal taxonomic marker for this methodology. The main disadvantage of 16S rRNA gene sequencing is that it is a PCR based approach which may introduce biases such as failing to amplify a fraction of the microbial community due to differences in PCR efficiency caused by primer annealing or hairpin formation in the DNA template or primers. In addition, there are several studies which have observed differences in the 16S rRNA microbial profile depending on the hypervariable region of the 16S rRNA gene amplified[80]. Therefore, it is essential to determine the optimum region which can provide the most representative taxonomic profile for each study cohort. Chapter 1 of this thesis optimises a 16S rRNA profiling protocol from DNA extraction to bioinformatics analysis to allow robust and reproducible analysis of premature infant's samples.

1.4.2 Shotgun metagenomics

When quantifying the taxonomic diversity present in a faecal sample or evaluating the diversity of the different AMR genes present with their predicted functions, shotgun metagenomics is currently the fastest and most informative approach.

Most authors like to summarise the possibilities shotgun metagenomics data offers into two main questions: *who is there* and *what are they capable of doing?*[81]. In shotgun metagenomics, microbial DNA is extracted from the faecal sample and sheered into smaller DNA fragments, which is then end-repaired using adapters, and subsequently sequenced, typically on a sequencing-by-synthesis Illumina platform. Data from shotgun metagenomics allows quantification of the taxonomic diversity present in an environmental sample using multiple single-copy marker genes (*e.g.* rRNA genes or protein coding genes). Focusing on single-copy gene families may provide a more accurate taxonomic profile than considering methods using gene families which differ in copy number across genomes[82]. Furthermore, data from this analysis can be used to predict biological function, which is extremely useful to obtain an overview of potential metabolic pathways. Functional predictions are performed by selecting the protein coding sequences from the metagenomic reads and comparing these to protein coding sequences in a database. This analysis can help to predict a profile describing the potential biological functions present in the metagenome community.

Currently, shotgun metagenomics serves as an excellent technology to study the diversity of AMR genes (the resistome). This technology can also give an indication of whether the AMR genes detected are linked to mobile genetic elements. Associating AMR genes to mobile genetic elements is fundamental to evaluate the potential each AMR gene has to be transferred to other bacteria, *i.e.* 'horizontal gene transfer'. A recent study published by Clemente and colleges in 2015 compared the human gut resistome from an uncontacted Amerindian tribe (with no previous contact to modern medicine) and industrialised societies using shotgun metagenomics[83]. This study revealed that the Amerindian gut resistome, surprisingly, is not that different from those of an industrialised nation. However, the presence of AMR genes linked to mobile genetic elements was found to be higher in industrialised populations than in the Amerindian tribe. This work lays

the background for understanding the role that antibiotics play in promoting the mobilisation of AMR genes between bacteria[84].

Despite the advantages of shotgun metagenomics to predict biological function, this approach also has its limitations. Beyond the basic technical limitations (*e.g.* higher DNA yields are required, which may prove difficult in limited samples, such as in preterm faecal samples), and the computational issues surrounding 'big' data, which are constantly being improved, the main limitation is that this approach can only identify genes that have been previously identified. Therefore, if the main aim is to find a novel gene, then functional metagenomics are often used.

1.4.3 Functional metagenomics

Functional metagenomics involves isolating DNA from microbial communities to study the functions of the encoded proteins. This functional-based approach allows the discovery of novel enzymes whose functions would not be predicted based on DNA sequence alone. Functional metagenomics demands more time in the laboratory than shotgun metagenomics. Briefly, bacterial DNA from a metagenomic sample is extracted, sheared, cloned into a vector and transformed into a host such as *E. coli.* These clones can be later tested to gain understanding of their functional capabilities.

The work by Gibson and colleagues on the preterm gut resistome used-functional metagenomics from bacterial DNA isolated from preterm faecal samples[55]. Results from the functional metagenomics study show that the sequences of the functional AMR genes are very similar to each other (95.8% amino acid similarity), but the proteins they encoded were extremely rare (24.8% amino acid identity) and most of them were not known from current AMR databases.

1.5 Use of whole genome sequencing for characterising clinical isolates

Whole genome sequencing (WGS) on clinical isolates offers the best possibility to study the genomic evolution of a particular bacteria. This technology permits addressing specific questions such as how pathogenic bacteria acquire AMR genes during antibiotic treatment, and/or how multidrug resistance bacteria are transmitted between individuals[85]. WGS sequencing allows tracking disease outbreaks by comparison of bacterial strains among infected individuals. Other approaches used to answer these questions, such as metagenomics approaches are often limited by the difficulty of assembling complex metagenome read data.

Many authors have used WGS on multidrug resistance isolates from the same host before, during and after an antibiotic treatment with the aim of studying the evolution of resistance in a bacterial lineage. Point mutations accumulated in their genomes throughout the antibiotic course, have been associated with evolution of resistance. As an example, Mwangi and colleagues studied the genomes of *S. aureus* strains isolated from the same patient during two months of extensive antibiotic treatment with vancomycin [86]. They compared the genomes of susceptible vancomycin strains isolated at the beginning of the antibiotic course, with non-susceptible vancomycin strains isolated at the end of the course. They found that 35 point mutations had accumulated in the strain isolated at the end of the treatment. These mutations were hypothesised by the authors to have been selected by the pressure imposed by the antibiotic treatment. Another interesting finding from this work is that some of these point mutations generated resistance to other antibiotics not used in the study, indicating that these point mutations were pleiotropic in nature.

WGS can also be used to reconstruct infection events between individuals. If the genomes of several bacterial isolates are sequenced from different infected hosts, it is possible to perform an ancestry analysis with their genomes and evaluate which

individual is the focus of the infection, which is key for epidemiological analysis of outbreak events[87].

1.6 Metabolomics studies allow to explore the metabolic capacity of the intestinal microbiota.

Metabolomics studies on faecal samples provide a functional readout of the microbiome and reports on the metabolic profile among the host, diet and gut microbiome[88]. The metabolome is described as a collection of endogenous molecules including amino acids, organic acids, sugars, fatty acids, lipids, small peptides, and vitamins which all provide a "snapshot" of the biological processes taking place in the sample[89]. This technique complements sequencing-based approaches, by offering information into metabolites that mediate microbe–microbe and microbe–host interactions. Presently the main application of metabolomics approaches in the clinic include: (iii) evaluate relationships between gut microbiome and host metabolism, (ii) elucidate functional alterations in the metabolite patterns of health and disease, and (iii) find specific metabolites which can be treated as diseases markers.

A recent study on the faecal metabolome on 786 twins and examining at 1,116 metabolites revealed the gut metabolome was largely explained by the composition of the gut microbiome, and host genetics did not seem to be that influencial[90]. In infancy, diet have been shown to have a strong association with the faecal metabolite profile. Human breastmilk contains elevated amounts of unique carbohydrates known as human milk oligosaccharides (HMOs), which act as selective nutrients for certain groups of microbiota populations (*e.g.* *Bifidobacterium*) in the production of short-chain fatty acids (SCFAs). These molecules have been shown to be major players in the maintenance of gut and immune homeostasis, and are found in high levels in the proximal colon. Changes in the proportions of SCFAs, related to a breastmilk or formula diet, have been

shown to influence programming effects on inflammatory-mediated diseases, obesity or allergies[91, 92]. Altogether, metabolomics investigations linked to microbiome and nutritional studies can help obtain a further insight into microbiota-host interactions.

1.8 Overarching hypotheses

My thesis work aims to study the preterm gut microbiome in heath and disease using next-gene sequencing techniques. My entire research work lays on two main hypothesis:

- Early administration of antibiotics to premature infants can lead to disruption of gut microbiome colonisation and also contribute to increase the reservoir of antimicrobial resistance genes.
- Probiotic supplementation can contribute to re-establishment of the commensal gut microbiome after antibiotic treatment and therefore reduce the reservoir of antimicrobial resistance genes.

1.8.1 Study-case hypotheses

This thesis is divided into four main chapters:

Chapter 1: This is a methods chapter, where I used 16S rRNA gene profiling to accurately sequence the preterm gut microbiome. I hypothesise the 16S rRNA gene sequencing pipeline will be influenced by each of these stages: (i) bacterial DNA extraction method (ii) 16S rRNA hypervariable region amplified and primer choice, and (iii) bioinformatics pipeline used for data analysis.

Chapter 2: This Chapter looks at the effects of probiotic supplementation in the preterm gut microbiome. I hypothesise probiotic supplementation using *Bifidobacterium bifidum* and *Lactobacillus acidophilus* can modify the gut

microbiome and metabolome of premature infants and reduce the presence of potential pathogenic bacteria.

Chapter 3: This chapter focuses on antibiotic treatment and its impact on preterm gut microbiome. I hypothesise administration of empiric antibiotic treatment to premature infants using benzylpenicillin and gentamicin can favour the growth of multidrug resistance bacteria, and enhance the reservoir of antimicrobial resistance genes.

Chapter 4: This chapter aims to sequence rapidly the preterm gut microbiome using MinION nanopore technology. I hypothesise the MinION nanopore sequencer will be able to profile rapidly faecal samples from premature infants to obtain bacterial metataxonomic profiles and characterise their antibiotic resistance profiles.

Publications arising from this PhD

Optimisation of 16S rRNA gut microbiota profiling of extremely low birth weight infants
Alcon-Giner C, Caim S, Mitra S, Ketskemety J, Wegmann U, Wain J, Belteki G, Clarke P, Hall LJ.
BMC Genomics. 2017 Nov 2;18(1):841.
- Thesis chapter 1.

Microbiota supplementation with *Bifidobacterium* and *Lactobacillus* modifies the preterm infant gut microbiota and metabolome
Alcon-Giner C #, Dalby MJ#, Caim S, Ketskemety J, Shaw A, Sim K, Lawson M, Kiu R, Leclaire C, Chalklen L, Kujawska M, Mitra S, Fardus-Reid F, Belteki G, McColl K, R. Swan J, Kroll S, Clarke P, Hall LJ.
bioRxiv 698092.
Submitted to Cell Reports Medicine.
- Thesis chapter 2. # equal contributions

Rapid profiling of the preterm infant gut microbiota using nanopore sequencing aids pathogen diagnostics
Legget RJ#, Alcon-Giner C#, Heavens D, Caim S, C. Brook T, Kujawska M, Martin S, Hoyles L, Clarke P, Hal LJ, Clark MD.
bioRxiv 180406.
October 2019 accepted in Nature Microbiology.
- Thesis chapter 4. # equal contributions

Draft genome sequences of *Citrobacter freundii* and *Citrobacter murliniae* strains isolated from the feces of preterm infants
Chen Y, Brook TC, Alcon-Giner C, Clarke P, Hall LJ, Hoyles L.
Microbiology Resource Announcements. 2019 Aug 15;8(33).
- Contribution: Obtained and made faecal samples available.

Draft Genome Sequence of *Raoultella ornithinolytica P079F* W, isolated from the feces of a preterm infant
Chen Y, Brook TC, Alcon-Giner C, Clarke P, Hall LJ, Hoyles L.
Microbiology Resource Announcements. 2019 Aug 15;8(33).
- Contribution: Obtained and made faecal samples available.

Streaming histogram sketching for rapid microbiome analytics
Rowe WP, Carrieri AP, Alcon-Giner C, Caim S, Shaw A, Sim K, Kroll JS, Hall LJ, Pyzer-Knapp EO, Winn MD.
Microbiome. 2019 Mar 16;7(1):40.
- Contribution: Obtained samples, extracted DNA, and prepared Illumina libraries.

Preterm infant-associated *Clostridium tertium*, *Clostridium cadaveris*, and *Clostridium paraputrificum* strains: Genomic and evolutionary insights
Kiu R, Caim S, Alcon-Giner C, Belteki G, Clarke P, Pickard D, Dougan G, Hall LJ.
Genome Biology Evolution. 2017 Oct 1;9(10):2707-2714.
- Contribution: Obtained samples, extracted DNA, and prepared Illumina libraries.

Chapter 1

Title: Optimisation of 16S rRNA sequencing pipeline for profiling faeces from extremely low birth weight infants

Abstract

The gut microbiome of premature infants, particularly extremely low birth weight (ELBW) infants is altered due to a variety of factors such as birth mode and antibiotic use. Thus, new microbiota therapies (*e.g.* probiotic supplementation) are becoming increasingly popular for manipulating the ELBW infant gut microbiome to improve infant health. Microbiome profiling, via metataxonomic 16S rRNA sequence profiling, represents an important tool for understanding the outcome of any probiotic supplementation. However, it is critical to optimise these types of studies for the cohort being characterised, in this case ELBW infants. In this work, a 16S rRNA profiling protocol was optimised, to allow robust and reproducible analysis of ELBW infant faecal samples, with or without probiotic supplementation. Three different DNA extraction methods were compared, followed by comparison of three hypervariable regions primer sets (V1 + V2 + V3), (V4 + V5) and (V6 + V7 + V8). Sequencing results were analysed using two bioinformatics approaches; Operational Taxonomic Unit and Paired End. Results from this work demonstrated that appropriate primer selection when using 16S rRNA gene profiling is essential and 16S rRNA gene region (V4+V5) should be avoided for analysing faeces samples from premature infants.

A graphical abstract summarising the different stages of this study is shown in Figure 3.

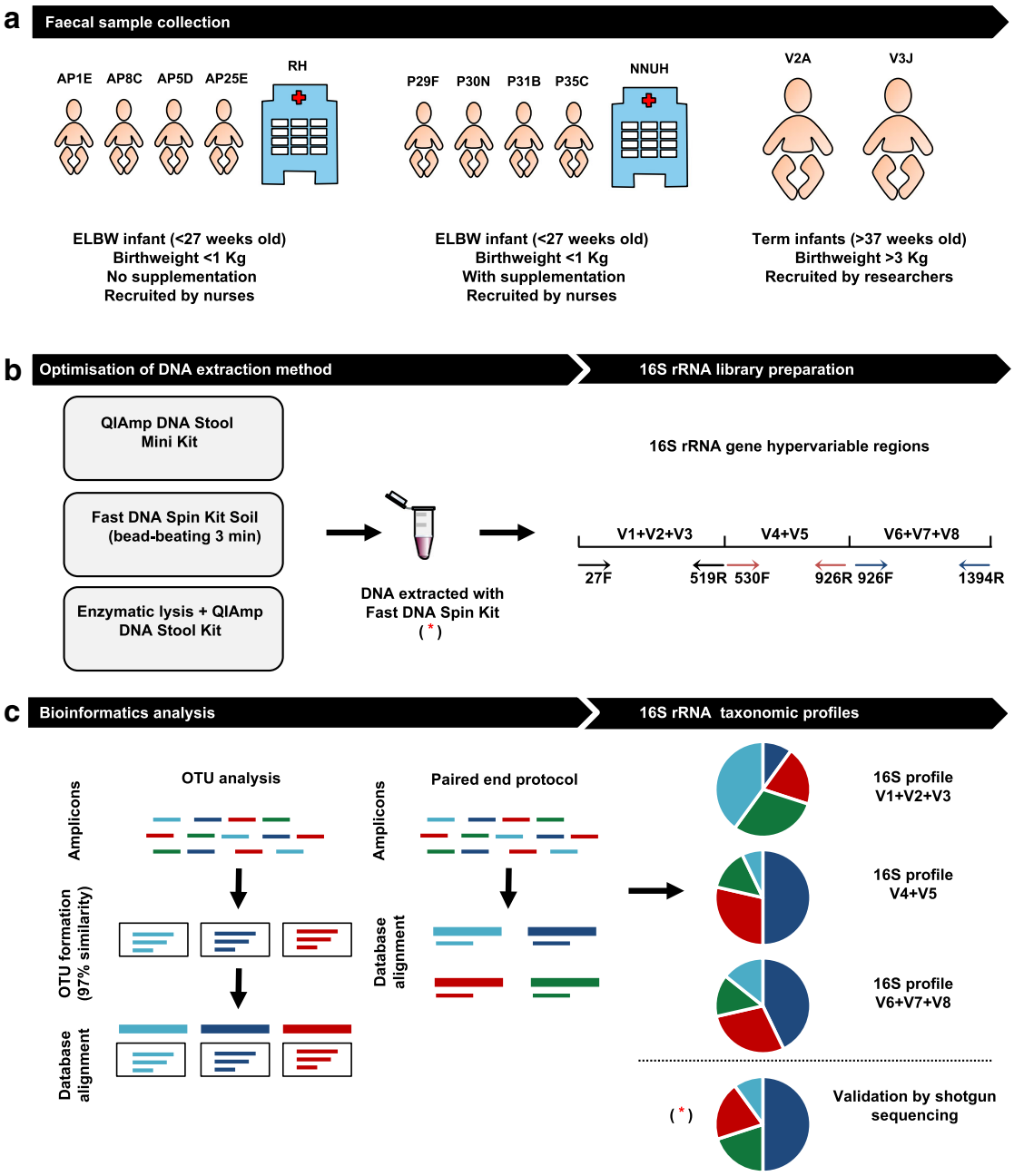


Figure 3 Study pipeline

a Recruitment of ELBW infants (<1000g) with no supplementation (AP1E, AP8C, AP5D and AP25D) and ELBW infants with supplementation (P29F, P30N, P31N, P35C) by nurses at the Rosie Hospital (RH) and the NNUH respectively. Term babies (V3J, V2A) were recruited by researchers. b Optimisation of the bacterial DNA extraction protocol from ELBW infant faeces by testing three different DNA extraction methods (QIAmp DNA Stool Mini Kit, Fast DNA Spin Kit Soil and enzymatic lysis + QIAmp DNA Stool Kit). Bacterial DNA from the study samples was extracted using the Fast DNA Spin Kit Soil and used to prepare three different 16S rRNA gene sequencing libraries. Each library was prepared using a specific pair of primers which target different hypervariable regions (prefixed by a V) of the bacterial 16S rRNA gene: (i) primers 27F-519R target (V1+V2+V3), (ii) primers 530F-926R target (V4+V5) and (iii) primers 926F-1394R target (V6+V7+V8). c A preliminary bioinformatics analysis was performed on two samples using two different bioinformatics pipelines: OTU analysis and the PE protocol. Both bioinformatics approaches were used to compare the different 16S rRNA gene sequencing profiles obtained for the different hypervariable regions tested (V1+V2+V3, V4+V5, and V6+V7+V8). (*) Validation of the 16S rRNA sequencing results was performed on three samples (AP8C, P29F and V3J) by shotgun sequencing.

Authors' contributions

This work was published in 2017 in BMC genomics[93], where I am the first author. I performed the DNA extractions, the Paired End protocol for bioinformatic analysis (supervised by Dr Suparna Mitra), the primer study, and prepared all the figures. Mr Shabhonam Caim carried out QIIME bioinformatic pipeline. Specific details of all authors' contributions can be found within the Methods section.

Introduction

Preterm birth occurs in 1:10 live births globally[94]. Premature infants, and in particular extremely low birth infants (ELBW), are born very immature, and consequently may have an underdeveloped gut and immune system[95]. Furthermore, they are often exposed to external factors which profoundly impact early life gut microbiome colonisation including; infections related to maternal health, Caesarean (C-) section delivery and long exposure to antibiotic treatments[55]. This is relevant as the microbiome plays a key role in immune programming[96], pathogen resistance[97] and neurocognitive development[98]. In comparison with full term infants, a distinct gut microbiome is found in premature infants, which is further altered in ELBW infants, and is characterised by lack of the genus *Bifidobacterium*, and overabundance of Enterobacteriaceae[99]. Importantly, these alterations in the gut microbiome can predispose premature infants to life threatening diseases such as necrotising enterocolitis (NEC)[100], which is distinguished clinically by inflammation of the intestinal tissue, and in severe cases may lead to intestinal wall perforation. Furthermore, as deterioration of the infant can occur rapidly (between 6-8 hours), there is an urgent requirement to optimally characterise microbiome profiles in patients at-risk (such as premature infants), particularly for studying the influence that novel microbiota therapies (*e.g.* probiotic supplementation) could offer for ecosystem restructuring and health outcomes. 16S rRNA gene metataxonomic profiling of faecal samples represents a cost-efficient method to gain insights into the bacterial components of the gut microbiome, and additionally

allows characterisation of cohorts where sample access and quantities are compromised (*e.g.* ELBW infants). Of the limited sequencing studies performed so far on ELBW infants, profiling clearly shows a lower abundance of Bifidobacteriaceae and *Lactobacillus,* which are commonly found in the gut microbiome of term infants and a higher abundance of Enterobacteriaceae, *Enterococci* and *Staphylococci*[101]. Importantly, Bifidobacteriaceae has been found in many studies as a dominant member of the full-term infant microbiome (particularly in vaginally delivered breast-fed infants) and has been associated with improved host wellbeing[102, 103]. Therefore, probiotic supplementation (or microbiota therapy) represents an attractive approach for manipulating the ELBW gut microbiome in order to improve health outcomes.

Previous studies examining the gut microbiome using 16S rRNA gene sequencing have highlighted the influence of DNA extraction method in the representative profile of the bacterial community obtained[79]. Also, the hypervariable region (V1 to V9) of the 16S rRNA gene targeted influences the ability to distinguish between different bacterial taxa, and only near-complete 16S rRNA gene sequences give accurate measures of taxonomic diversity[104]. Currently, the complete sequence of the 16S rRNA gene (~1400 bp) is outside the read length of short-read high-throughput sequencing technologies (*i.e.* Illumina platforms). Therefore, it is essential to determine the optimum region which can provide the most representative taxonomic profile for the study cohort.

In this study an optimised protocol for profiling the ELBW infant gut microbiome using 16S rRNA gene sequencing is presented. The study analysed faecal samples from ELBW premature infants (<1000 g; with/without probiotic supplementation) and samples from term infants as controls. Samples from ELBW infants receiving probiotic supplementation comprised the 'spiked' samples with known species of *Bifidobacterium* and *Lactobacillus*. A bacterial DNA extraction method was optimised for these samples, after comparing three different methods, and

generated amplicons to three different hypervariable regions of the 16S rRNA gene (V1+V2+V3), (V4+V5) and (V6+V7+V8) followed by Illumina sequencing were compared, with reads analysed using two different bioinformatics pipelines (OTU versus Paired End protocol). Finally, to further validate the sequencing results, shotgun sequencing was used on a subset of the tested samples. Results from this study demonstrate that inclusion of an extended bead-beating step was essential when extracting DNA from faecal samples, and that sequencing regions (V1+V2+V3) or (V6+V7+V8) of the 16S rRNA gene provided the most representative bacterial profile of the ELBW infant gut microbiome.

Hypothesis and aims

Hypothesis: The bacterial community 16S rRNA gene profile will be influenced by each stage of the 16S rRNA gene pipeline: (i) bacterial DNA extraction method (ii) 16S rRNA hypervariable region amplified and primer choice, and (iii) bioinformatics pipeline used for data analysis.

This study addressed 3 aims:

a) Evaluation of three different bacterial DNA extraction methods, to determine the optimal extraction method for profiling DNA from ELBW infants' faeces.

b) Determination of which regions of the 16S rRNA gene (V1+V2+V3, V4+V5, and V6+V7+V8) were most accurate at representing the ELBW infant gut microbiome. Complementing this analysis with an *in silico* primer aligning study to evaluate alignment of the primer pairs used, among bacterial members commonly found in the gut microbiome of ELBW.

c) Comparison of two popular bioinformatics pipelines (OTU clustering analysis and paired end protocol (PE) to assess whether the same

biological conclusions regarding ELBW microbiome composition, could be reached using different bioinformatics pipelines.

Methods

Subject recruitment and faecal sample collection

This study was approved by the University of East Anglia (UEA) Faculty of Medical and Health Sciences Ethics Committee, and sample collection was in accordance with protocols laid out by the National Research Ethics Service (NRES) approved UEA Biorepository (Licence no: 11208). Infants admitted to the Neonatal Intensive Care Units (NICUs) of the Norfolk and Norwich University Hospital (NNUH, Norwich, UK) and the Rosie Hospital (Cambridge, UK) were recruited by doctors or nurses with informed and written consent obtained from parents. Both NICUs had similar protocols for feeding and the prescription of antibiotics and antifungal drugs, with the main exception being probiotic use; the Rosie Hospital does not use probiotics, the NNUH routinely prescribed all ELBW infants an oral probiotic treatment containing *Bifidobacterium bifidum* and *Lactobacillus acidophilus* (*i.e.* Infloran®, Desma Healthcare, Switzerland) in a twice daily dose of 1 x 10^9 of each species, given from birth until 34 weeks old. A total of eight ELBW infants were recruited, four received probiotic supplementation and four did not receive any supplementation. All recruited ELBW infants were <27 week's gestation and weighed ≤1000 g at birth. Infants born vaginally and breast-fed were specifically selected, with the aim of normalising for other external factors which can influence gut colonisation of *Bifidobacterium and Lactobacillus*. A control group of two term babies were also recruited by the research team following the same protocol. Faecal samples were collected from nappies into a sterile stool container and stored at 4 °C. DNA was extracted within 4 hours of collection. Subject details are included in the table below.

Table 2 Subject details and metadata

Sample	Hospital	Birth mode	Term at week (weeks)	Birth weight (g)	Probiotic treatment (Infloran®)	Formula or breastfed	Sample collection[§] (Days after DOB)
AP1E	RH[†]	Vaginal	25	830	No	Breastfed	16
AP5D	RH	Vaginal	25	800	No	Breastfed	12
AP25E	RH	Vaginal	25	786	No	Breastfed	18
AP8C	RH	Vaginal	23	576	No	Breastfed	21
AP10B	RH	Vaginal	26	710	No	Breastfed	30
P31B	NNUH[‡]	Vaginal	23	605	Yes	Breastfed	16
P29F[*]	NNUH	Vaginal	26	1000	Yes	Breastfed	12
P30N[*]	NNUH	Vaginal	26	960	Yes	Breastfed	15
P35C	NNUH	Vaginal	23	565	Yes	Breastfed	16
P66F	NNUH	Vaginal	26	670	Yes	Breastfed	20
V3J	NNUH	Vaginal	40	3500	No	Breastfed	58
V2A	NNUH	Vaginal	40	3320	No	Breastfed	60
V3ZC	NNUH	Vaginal	40	3500	No	Breastfed[#]	365

[*] Baby P29F and P30N were twins.
[†] Rosie Hospital
[‡] Norfolk and Norwich University Hospital
[§] Faecal samples were collected in a stool container and stored at 4^0 C. DNA was extracted within 4 hours of collection.
[#] Exclusively breastfed baby until six month old.
DOB: date of birth

Sample processing and DNA extraction

Optimisation of tree bacterial DNA extraction methods was performed on faecal samples from two ELBW infants (with/without supplementation) and one term infant sample. Three different DNA extraction methods were used: (i) FastDNA Spin Kit for Soil (MP Biomedical, Santa Ana) following the manufacturer's instructions and extending the bead-beating step to 3 minutes (ii) QIAmp DNA Stool Mini Kit (Qiagen) following the manufacturer's instructions, and (iii) QIAmp DNA Stool Mini Kit (Qiagen) including an initial enzymatic lysis step of 1 hour at 37^0C (enzymatic mix: 50 mM Tris-HCl, pH 8.0, 10 mM $MgSO_4$, 5 mg/mL lysozyme and 50 U/mL mutanolysin). The DNA recovered from these samples was assessed using a Qubit® 2.0 fluorometer (Invitrogen). I performed this procedure.

16S rRNA gene library preparation

Fast DNA Spin Kit extracted DNA was used for preparing 16S rRNA Illumina MiSeq sequencing libraries. DNA concentration was normalised to 5 ng/mL using a Qubit® 2.0 fluorometer. Three hypervariable regions of the 16S rRNA gene (V1+V2+V3 (primers 27F-519R), V4+V5 (primers 530F-926R), and V6+V7+V8 (primers 926F-1394R)) were amplified using the HotStarTaq Plus Master Mix Kit (Qiagen, USA). Details of the primer sequences used for amplification can be found in Table 3.

Table 3 Primers used in 16S rRNA sequencing library

Primer name	Shorthand	16S rRNA gene hypervariable region	Primer sequence
27Fmod	27F	V1+V2+V3	AGRGTTTGATCMTGGCTCAG
ill519Rmod	519R	V1+V2+V3	GTNTTACNGCGGCKGCTG
530F	530F	V4+V5	GTGCCAGCMGCNGCGG
bac926R	926R	V4+V5	CCGTCAATTYYTTTRAGTTT
926F	926F	V6+V7+V8	AAACTYAAAKGAATTGACGG
bac1394R	1394R	V6+V7+V8	ACGGGCGGTGTGTRC

Each DNA sample was amplified using a primer pair tagged individually with a unique barcode. PCR amplification conditions were: 1 cycle of 94 °C for 3 minutes, followed by 25 cycles of 94 °C for 45 s, 55 °C for 15 s and 72 °C for 30 s. Amplicons were pooled in equal proportions and purified using Ampure XP beads (Agencourt). The purified product was used to prepare the Illumina DNA library. Libraries were sequenced on the Illumina MiSeq platform using a read length up to 2x300 bp. I prepared the samples for this analysis, and libraries were made and sequenced by company MrDNA.

Whole genome shotgun metagenomics library preparation

Genomic DNA (approximately 500 ng) from two ELBW infants' samples (with/without supplementation) and one term infant sample was fragmented to an

average size of 250 bp and subjected to DNA library creation using established Illumina paired end protocols. Adapter-ligated libraries were amplified and indexed via PCR. A portion of each library was used to create an equimolar pool and enriched libraries were subjected to 100 base paired end sequencing (HiSeq 2000 V3; Illumina). I prepared the samples for this analysis and samples were sequenced at Sanger Institute.

Bioinformatics analysis

16S rRNA gene sequencing analysis

Two bioinformatics pipelines were used to analyse the 16S rRNA gene sequencing data: OTU clustering analysis and paired end protocol (PE). OTU clustering analysis was performed using the QIIME bioinformatics pipeline[105]. First, read pairs were assembled using PEAR[106], a highly accurate pair-end read merger. Second, sequences were quality filtered using QIIME's split_libraries_fastq.py and chimeras were identified and removed using identify_chimeric_seqs.py and filter_fasta.py respectively. Following, OTU picking step was run using pick_open_reference_otus.py (percent_subsample parameter set at 0.1) and QIIME SILVA_128 as our reference database. OTUs were formed by clustering to 97% similarity, and a representative sequence was picked for each OTU aligned using PyNAST [107] and taxonomy was assigned using uclust[108]. Filtering prior to build the tree that was done by removing the positions with gaps and specified as 0 in the lanemask. FastTree is used to create a tree file for the represented sequences. Final taxonomic output was saved as a biom file.

An in-house PE protocol was used following the quality control of the raw paired reads using FASTX-Toolkit[109] (with a minimum quality threshold of 33 for at least 50% of the bases in each read sequence. Reads that passed the threshold were aligned against the SILVA database (version: SILVA_128_SSURef_tax_silva)[110] and BLASTN (ncbi-blast-2.2.25+; Max e-value 10e-3)[111]. The BLAST files

obtained were imported into MEGAN6[112] to create MEGAN-own files ("rma6" files) using the following parameters: 100 as maximum number of matches per reads, and "Min Score = 50" and "Top Percent = 10". All output files (rma6) of paired read sequences were then normalised and compared using MEGAN6. I performed the Paired End Protocol supervised by Dr Suparna Mitra, and Mr Shabhonam Caim performed the QIIME pipeline.

Whole genome shotgun gene sequencing

Whole genome paired sequences from samples AP8C (an ELBW infant without supplementation), P29F (an ELBW infant who received supplementation) and V3J (term infant) were obtained from an Illumina HiSeq 2000 V3 sequencer. The first 10 bases were trimmed using FASTX-Toolkit[113]. Subsequently, trimmed sequences were aligned against the NCBI non-redundant database (version 04/2016)[114] using DIAMOND[115]. All output files of paired read sequences were then imported and analysed using the PE protocol of MEGAN with non-default settings.

Functional profiles were performed on the same samples using the KEGG pathway database. Mapping files used for this analysis were obtained from MEGAN's website. This analysis was done by Dr Suparna Mitra.

Sequencing reads statistics

Read counts at different stages of the bioinformatics analysis are provided in Appendix 1. To compare study samples, sequences were normalised using values from the sample with the lowest number of reads. In other cases, read counts were displayed in percentage of number of reads.

Principal Coordinate Analysis plot was performed using Bray-Curtis distances on the 16S rRNA bacterial community profiles using MEGAN. The Shannon diversity

index was obtained by exporting genus level profile (normalised) from all 30 samples in MEGAN and plotting them in Excel. I performed this analysis.

Primer annealing study

Amplicon sequences from the most common bacterial taxa found in sample P29F (ELBW infant with supplementation) were extracted using MEGAN [112]. Full length sequences of the respective 16S rRNA genes were obtained from Genbank after identified the respective database entries using BLASTN [116]. Primer annotation of the 16S rRNA sequences was performed using Genedoc 2.7 [117]. I performed this analysis with the help of Dr Udo Wegmann.

Validation of primers 530F-926R: PCR and melting curves qPCR

PCR

DNA extracted from *B. bifidum* (isolated from the probiotic supplement) and seven different *Bifidobacterium* strains (from NCIMB strain collection, Aberdeen, Scotland), was amplified by PCR using primers 530F-926R. Table 4 provides the details of the NCIMB collection strains used in this study.

Table 4 *Bifidobacterium* strains used for validating primers 530F-926R using PCR

***Bifidobacterium* strains**	**NCIMB collection number**	**Isolated from**
B. longum	8809	Nursing stools
B. bifidum	13922	Not described
B. catenatum	702239	Human faeces
B. angulatum	702236	Human faeces
B. adolescentis	702204	Adult intestine
B. breve	8807	Infant intestine
B. infantis	702255	Infant intestine

A faecal metagenomic sample and a *Lactobacillus acidophilus* strain (isolated directly from the probiotic supplement) were used as positive controls. Amplicon samples were run on 1% agarose gel for 30 minutes at 100 V. DNA was visualized under UV light after staining with ethidium bromide. I performed this analysis.

qPCR

Melting curves of PCR amplicons obtained from the probiotic strains (*Bifidobacterium bifidum* and *Lactobacillus acidophilus*) and two bacterial isolates from an ELBW infant with supplementation (*Enterococcus faecium* and *Streptococcus infantarius*) were performed using a LightCycler 480 (Roche Molecular Diagnostics). Conditions for the melting curves were: 95 °C for 5 s, 65 °C for 1 min and a final stage at 97 °C continuous. As an additional experiment, a melting curve from an amplicon obtained from a mixed DNA sample (containing 5 ng DNA from of each of the above bacterial species) was run. Conditions used for this melting curve were the same as the ones described previously. I performed this procedure.

Results

Effect of DNA extraction method in sample preparation.

DNA extraction is the first critical step in sample preparation for sequencing studies[118]. A preliminary study was performed with two faecal samples from two ELBW infants, and one term infant as a control. Two different DNA extraction kits (Fast DNA Spin kit and QIAamp) were tested with two different conditions with the aim of optimising the best extraction method for profiling DNA from ELBW infants' faeces. Average DNA concentrations in elution from each method of 21.4 - 1.97 ng μL^{-1}, and 0.016 - <0.0005 ng μL^{-1} were obtained with the Fast DNA Spin Kit and QIAamp, respectively (Table 5). The DNA concentration was significantly

higher with the Fast DNA Spin kit than the QIAamp (p <0.05), and elevated amounts of DNA were obtained with the inclusion of a longer bead-beating step.

Table 5 DNA yield from different DNA extraction methods

Sample	Extraction Method	Qubit (ng/µl)
ELBW infant no probiotics (AP10B)	Fast DNA Spin Kit (3 min bead-beating)	2.25
	Fast DNA Spin Kit (30 sec bead-beating)	1.97
	QIAamp DNA stool kit	<0.0005
	Enzymatic lysis and QIAamp DNA stool kit	0.0146
ELBW infant with probiotics (P66F)	Fast DNA Spin Kit (3 min bead-beating)	13.8
	Fast DNA Spin Kit (30 sec bead-beating)	7.38
	QIAamp DNA stool kit	<0.0005
	Enzymatic lysis and QIAamp DNA stool kit	0.0156
Term baby(V3ZC)	Fast DNA Spin Kit (3 min bead-beating)	21.4
	Fast DNA Spin Kit (30 sec bead-beating)	7.7
	QIAamp DNA stool kit	0.0164
	Enzymatic lysis and QIAamp DNA stool kit	0.77

Assessing coverage of 16S rRNA sequencing data

Coverage of sequencing data was assessed by performing rarefaction curves, which correlates numbers of reads sequenced with number of genus found in the sample. As a rule, when the rarefaction curve *plateaus* the majority of bacterial genus present in the sample are detected. This study was performed as there were 5x and 10x differences in the number of reads obtained from regions (V1+V2+V3 and V6+V7+V8) compared to region (V4+V5). After discussions with the sequencing company, no specific reasons (*e.g.* library preparation or MiSeq settings) could explain the higher number of reads obtained when using region (V4+V5).

The coverage analysis using rarefaction curves indicated that at 25,000 reads the majority of bacterial populations were sequenced (Figure 4), with all samples presenting >25,000 reads, which subsequently enabled robust comparison and normalisation of the data.

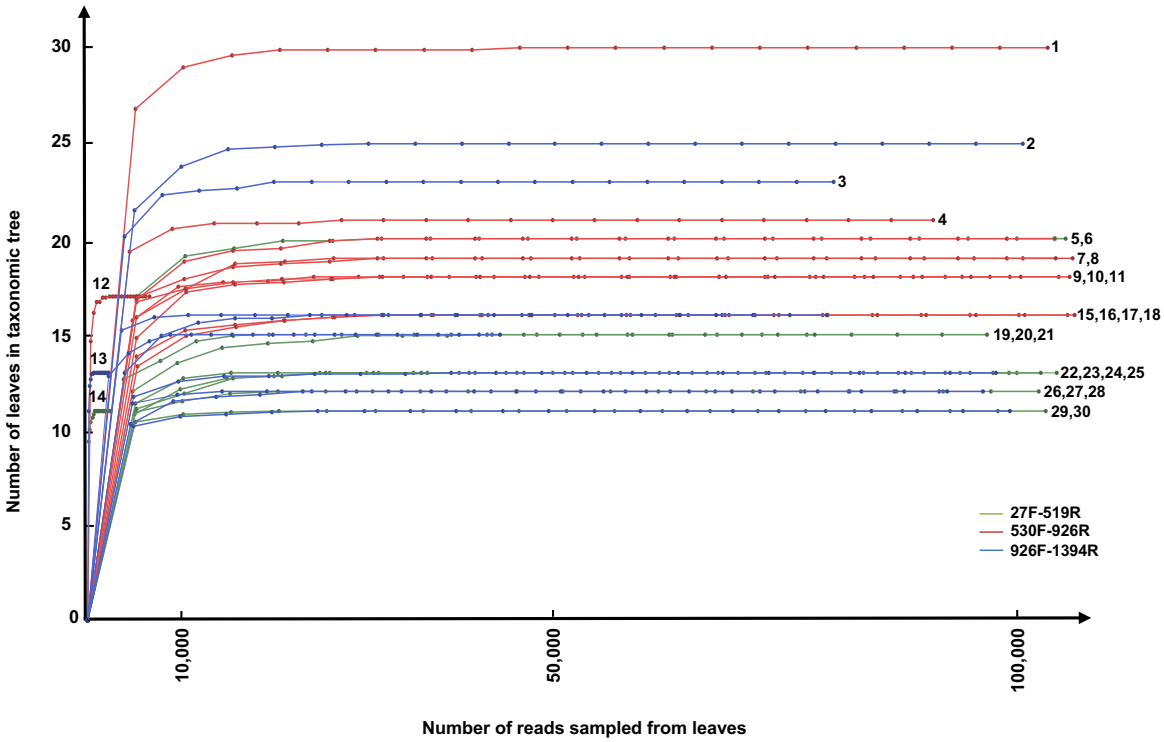


Figure 4 Rarefaction curves 16S rRNA gene sequencing data

Rarefaction curves representing number of species (leaves) detected at genus level versus number of reads sampled. Three different 16S rRNA gene sequencing data were used for this study: (i) green curves represent sequencing data from 16S rRNA library prepared using primers 27F-519R, (ii) red curves represent sequencing data from 16S rRNA library prepared using primers 530F-926R and (iii) blue curves represent sequencing data from 16S rRNA library prepared using primers 926F-1394R. Rarefaction curves are labelled with numbers to differentiate among the samples used in the study: 1 (V2A.530F), 2 (V2A.926F), 3 (V3J.926F), 4 (AP8C.530F), 5 (V2AJ.27F), 6 (AP5D.530F), 7 (P35C.530F), 8 (V3J.530F), 9 (P29F.530F), 10 (P31B.530F), 11 (P30N.530F), 12 (AP25E.530F), 13 (AP25E.926F), 14 (AP25E. 27F), 15 (AP8C.926F), 16 (P35C.926F), 17 (V3J.27F), 18 (P31B.926F), 19 (AP1E.530F), 20 (P31B.27F), 21 (P29F.27F), 22 (P30N.27F), 23 (AP5D.926F), 24 (AP5D.27F), 25 (P29F.926F), 26 (AP1E.926F) ,27 (AP1E.27F), 28 (AP8C.27F), 29 (P30N.926F), 30 (P35C.27F). Numbers 12, 13 and 14 correspond to sample AP25E where majority of sequenced reads assigned at family level (i.e. Enterobacteriaceae).

Optimisation of 16S rRNA bioinformatics pipeline (OTU pipeline versus paired end protocol)

To evaluate and determine the optimal 16S rRNA bioinformatics pipeline a preliminary study was performed on one ELBW infant (AP1E) and one term infant sample (V3J) using two different bioinformatics approaches (OTU and paired end protocol). A reference-based OTU clustering analysis (QIIME) was used; this approach organises the raw reads within OTUs of 97% similarity, and then compares against public databases. In contrast, the paired end protocol (PE) aligns raw reads to the 16S rRNA gene databases directly after quality control. At genus level both methods tested showed similar taxonomic profiles for the majority of the bacterial populations (*e.g. Bacteroides* and *Staphylococcus* for sample AP1E (Figure 5a)*, and Bifidobacterium* and *Streptococcus* for sample V3J (Figure 5b).

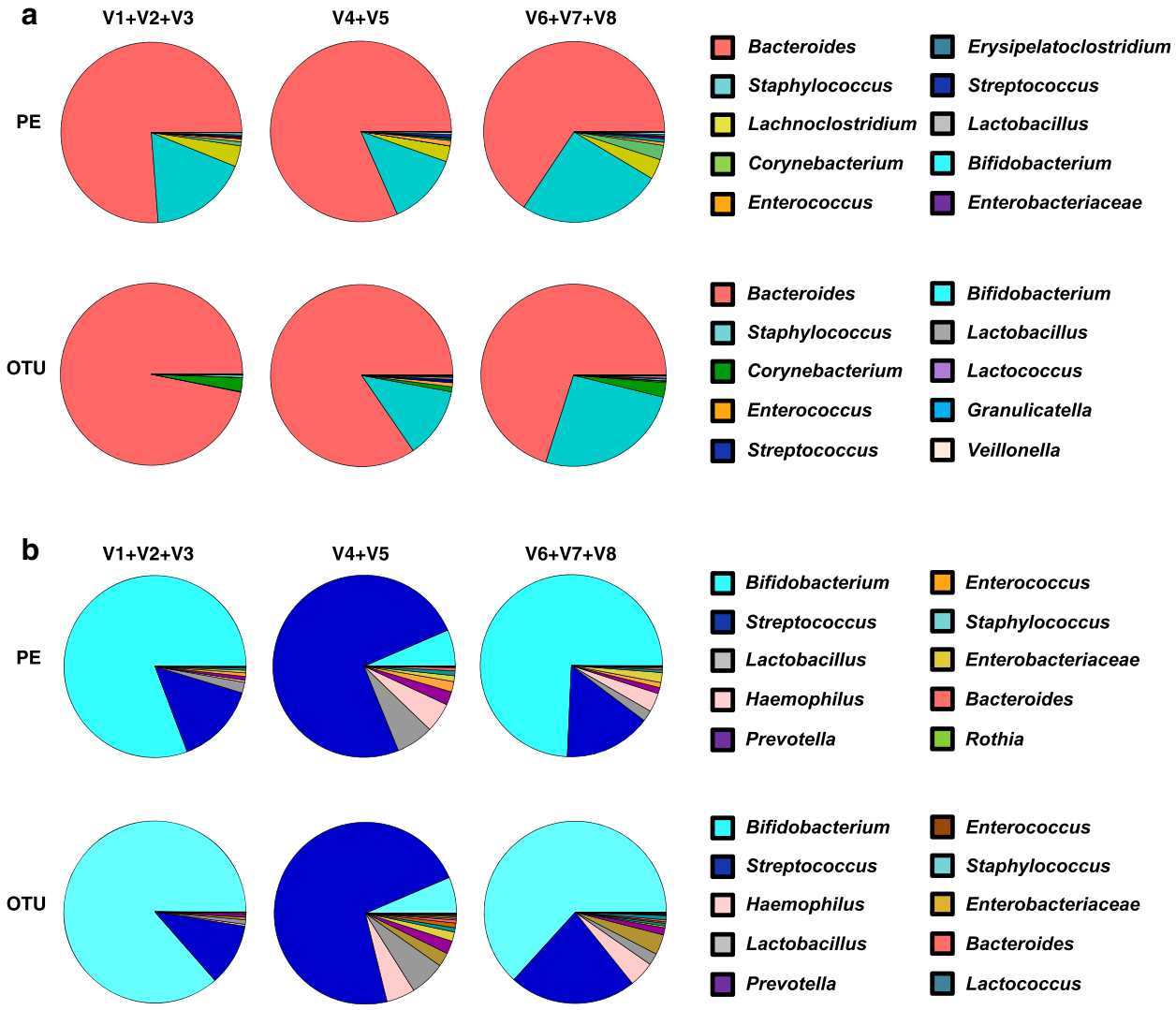
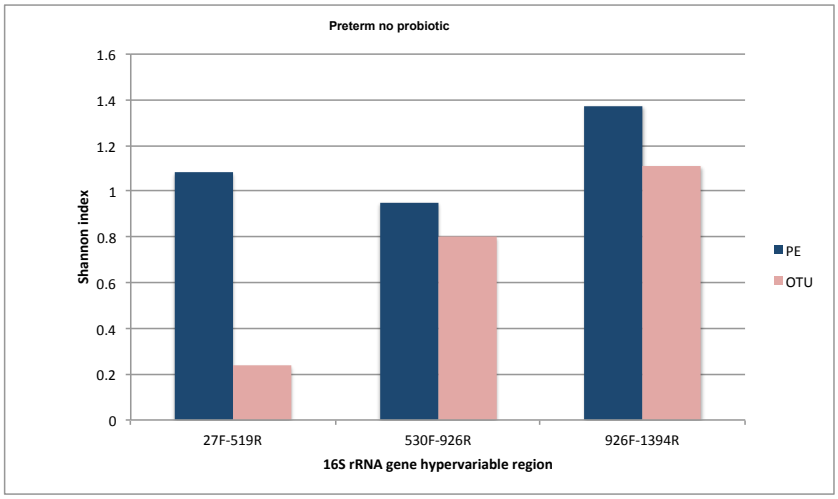


Figure 5 Comparison of bioinformatics analyses (OTU versus PE protocol)

Preliminary study comparing two different bioinformatics approaches: OTU clustering performed using QIIME and PE protocol. Both bioinformatics approaches used the same database (SILVA version 128). a Taxonomic profiles obtained using PE protocol and OTU clustering for sample AP1E (ELBW infant no supplementation). b Taxonomic profiles obtained using PE protocol and OTU clustering for sample V3J (term infant sample). Three different 16S rRNA gene libraries were prepared for each sample, (i) (V1+V2+V3, primers 27F-519R), (ii) (V4+V5, primers 530F-926R) and (iii) (V6+V7+V8, primers 926F-1394R).

To complement this analysis, the Shannon Diversity Index (Figure 6) for OTU and PE approaches was calculated, with results between both pipelines comparable, except for region (V1+V2+V3, 27F-519R) where the OTU approach presented the lowest value.

a



b

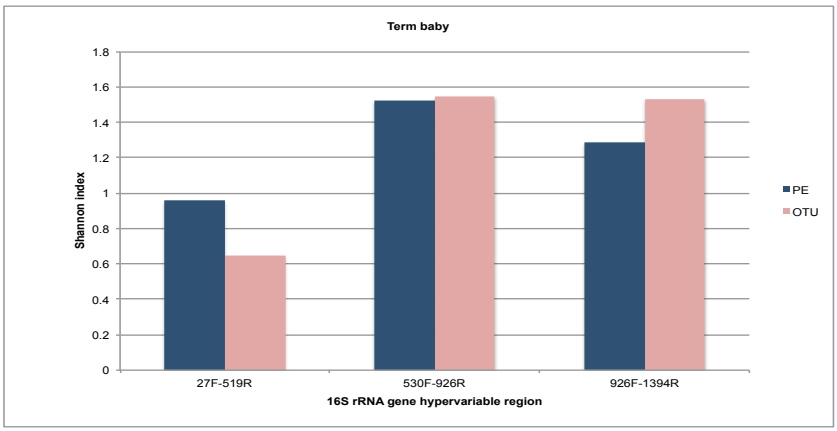


Figure 6 Shannon diversity index on 16S rRNA gene sequencing data analysed using OTU and PE protocol

Shannon diversity index was calculated using 16S rRNA bacterial community profiles for sample AP1E (ELBW infant without probiotic supplementation) and sample V3J (term infant). a Shannon diversity indexes comparison of three different 16S rRNA libraries (27F-519R (region V1+V2+V3), 530F-926R (region V4+V5) and 926F-1394R (region V6+V7+V8)) using OTU and PE protocol pipelines for sample AP1E (premature infant no supplementation). b Shannon diversity indexes comparison of three different 16S rRNA libraries (27F-519R (region V1+V2+V3), 530F-926R (region V4+V5) and 926F-1394R (region V6+V7+V8)) using OTU and PE protocol pipelines for sample V3J (term infant).

This preliminary study served to highlight that different bioinformatics protocols provided similar bacterial profiles for high relative abundance taxa, however there were some differences, which is particularly relevant to bacterial populations present in low abundance (*e.g. Lachnoclostridium* and *Corynebacterium* for sample AP1E (Figure 5a)*, and Prevotella* and *Enterococcus* for sample V3J (Figure 5b).

Impact of 16S rRNA gene hypervariable region amplified on taxonomic assignments.

Targeting different hypervariable regions of the 16S rRNA gene is known to be an important factor influencing the bacterial genus profiles obtained[119]. To determine which hypervariable region is optimal at profiling the ELBW infant gut microbiome, three 16S rRNA gene sequencing libraries were prepared amplifying different regions of the 16S rRNA gene (V1+V2+V3, V4+V5, and V6+V7+V8).

Results from the taxonomic assignments showed that the most abundant bacterial populations found in ELBW samples (*e.g. Enterococcus, Staphylococcus,* Enterobacteriaceae and *Streptococcus*), were similar between regions (V1+V2+V3, V4+V5, and V6+V7+V8) (Figure 7). These data indicate that the three hypervariable regions similarly target these bacterial taxa, which agrees with previous studies[120].

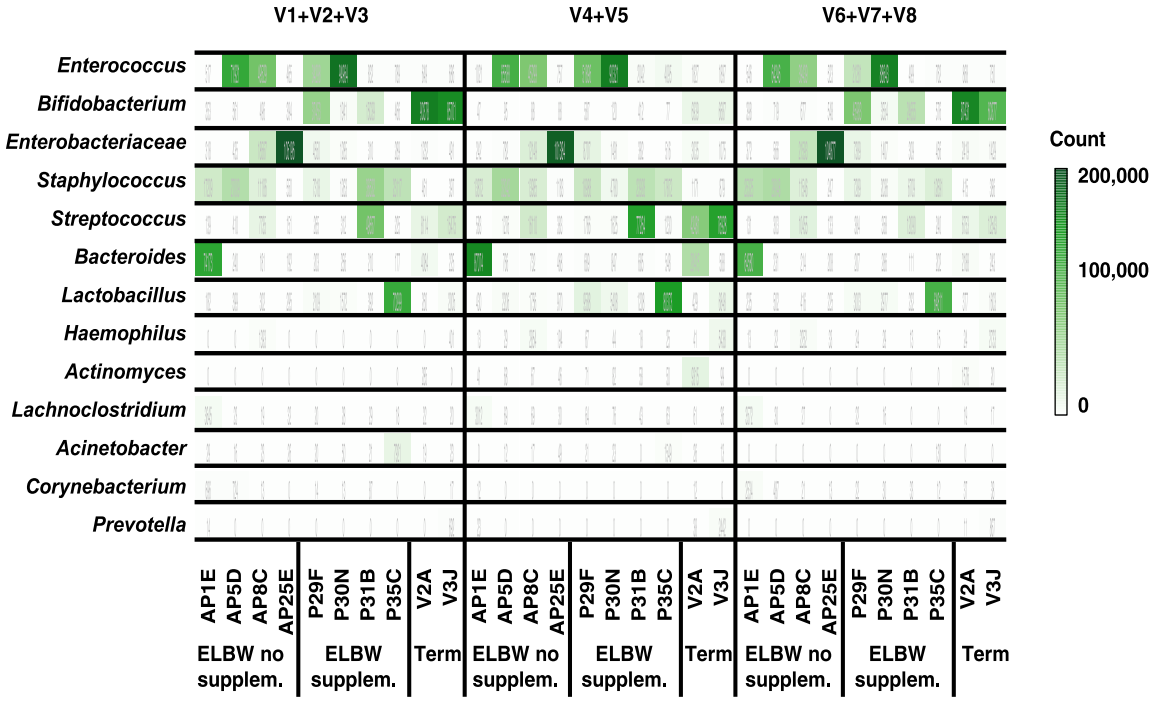


Figure 7 Comparison of taxonomic assignments among the 16S rRNA gene hypervariable regions tested using PE protocol approach

Heat map displaying number of reads assigned to the most common bacterial taxa found in the study samples. Top panel row divides the figure in the different regions of the 16S rRNA gene analysed, namely: V1+V2+V3 (primers 27F-519R), V4+V5 (primers 530F-926R) and V6+V7+V8 (primers 926F-1394R). The vertical axis of the panel indicates 13 most common bacterial taxa found. The horizontal axis labels the different samples used in the study: premature infants without supplementation (AP1E, AP5D, AP8C, AP25C), premature infants with supplementation (P29F, P30N, P31B, P35C), and term baby samples (V2A, V3J). The intensity of the green colour highlights the abundance of the number of reads found. Probiotic supplementation has been abbreviated to supplem.

A novel feature of this study is that it includes faecal samples obtained from ELBW infants supplemented with probiotics. These samples can be considered as 'spiked samples' containing known bacterial taxa (*Lactobacillus* and *Bifidobacterium*), which enabled more thorough comparison of profiles obtained when the different hypervariable region primer sets.

Results for the taxonomic assignments assigned to *Bifidobacterium* indicated prominent differences between the different regions. Profiles obtained from regions (V1+V2+V3 and V6+V7+V8) highlighted *Bifidobacterium* in two out of the four samples (P29F and P30N) analysed from ELBW infants who had received probiotic supplementation (Figure 7). In contrast, analysis of region (V4+V5) did not indicate *Bifidobacterium* in any of the four samples (P29F, P30N, P31B and P35C) tested from ELBW infants who received supplementation (*i.e*. 'spiked' samples). Interestingly, one of the remaining two supplemented ELBW infants recently finished a five-day course of vancomycin, which could explain the underrepresentation of *Bifidobacterium* in this sample. Notably, when comparing the results from region (V4+V5) in samples which are known to contain a high proportion of *Bifidobacterium i.e.* full term infants, a 92% decrease in the number of reads assigned to *Bifidobacterium* was observed when compared to the other regions tested (V1+V2+V3 and V6+V7+V8), which is in agreement with the results found in ELBW infants receiving supplementation.

In the case of *Lactobacillus*, the three hypervariable regions (V1+V2+V3, V4+V5, and V6+V7+V8) were able to detect the presence of this taxon in three (P29F, P30N and P35C) out of the four samples from ELBW infants who received supplementation, and in one term baby sample (V3J). Region (V4+V5) presented the highest number of reads obtained for this taxon, and amplicons amplifying region (V4+V5) revealed 3 and 4 times higher number of reads for *Lactobacillus* when compared to the other regions (V1+V2+V3 and V6+V7+V8). These data

indicate that the (V4+V5) region may over represent this bacterial population, which is discussed in more detail when comparing to the shotgun analysis.

For additional bioinformatics comparison, the same analysis as described above but using the QIIME bioinformatics pipeline was performed. Similar findings were found when using the QIIME pipeline; region V4+V5 showed overrepresentation of *Lactobacillus* and underrepresentation of *Bifidobacterium* (Figure 8).

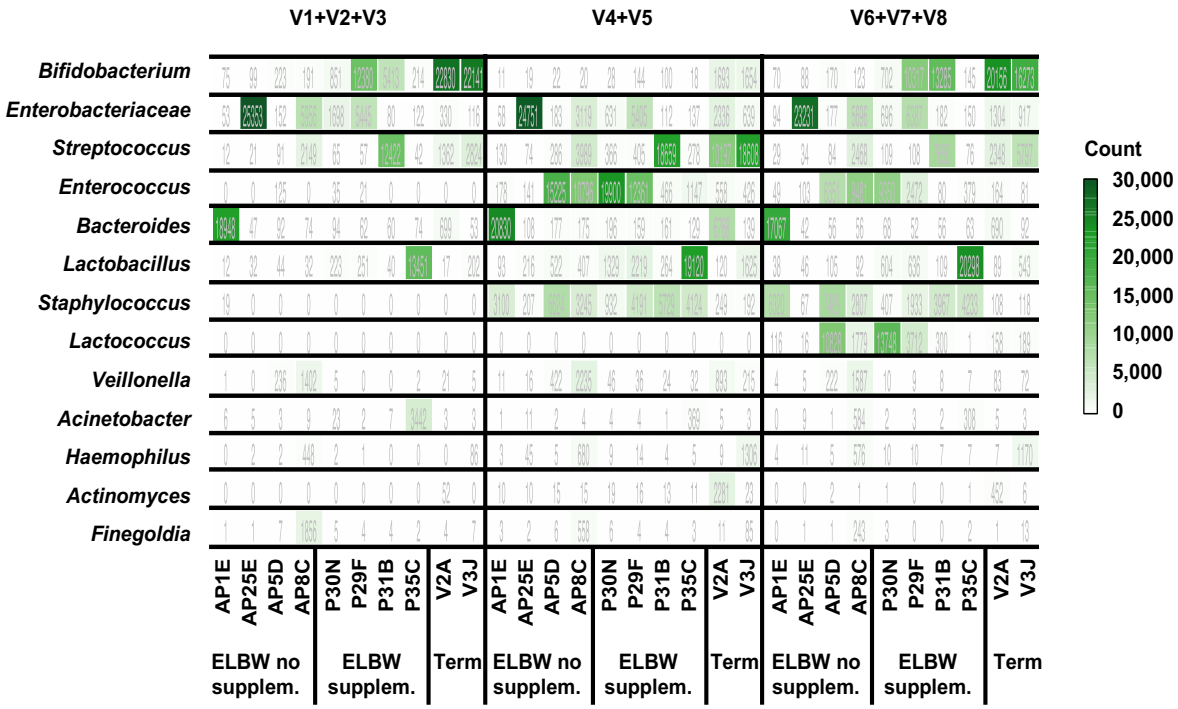


Figure 8 Comparison of taxonomic assignments among the 16S rRNA gene hypervariable regions tested using QIIME approach

Heat map displaying number of reads assigned to the most common bacterial taxa found in the study samples using QIIME bioinformatics pipeline. Top panel row divides the figure in the different regions of the 16S rRNA gene analysed, namely: V1+V2+V3 (primers 27F-519R), V4+V5 (primers 530F-926R and V6+V7+V8 (primers 926F-1394R). The vertical axis of the panel indicates a selection of the 13 most common bacterial taxa found. The horizontal axis labels the different samples used in the study: premature infants without supplementation (AP1E, AP5D, AP8C, AP25C), premature infants with supplementation (P29F, P30N, P31B, P35C), and term baby samples (V2A, V3J). The intensity of the green colour highlights the abundance of the number of reads found. Probiotic supplementation has been abbreviated to supplem.

These data indicated that primer choice, or hypervariable region of the 16S rRNA gene amplified ply a stronger role in overall profiles obtained, rather than biases introduced from different bioinformatic analysis.

Next, a Principal Coordinate Analysis (PCoA) based on 16S rRNA community profiles of the hypervariable regions tested (Figure 9) was performed to visualise the similarities and differences of all the bacterial communities assigned for each sample. Results highlighted that the distribution of samples amplified using region (V4+V5) was distinct from samples amplified using region (V1+V2+V3 and V6+V7+V8). These differences were more accentuated in faecal samples which contain *Bifidobacterium* such as P31B and P29F (from ELBW infants with supplementation) and V3J and V2A (from term infants).

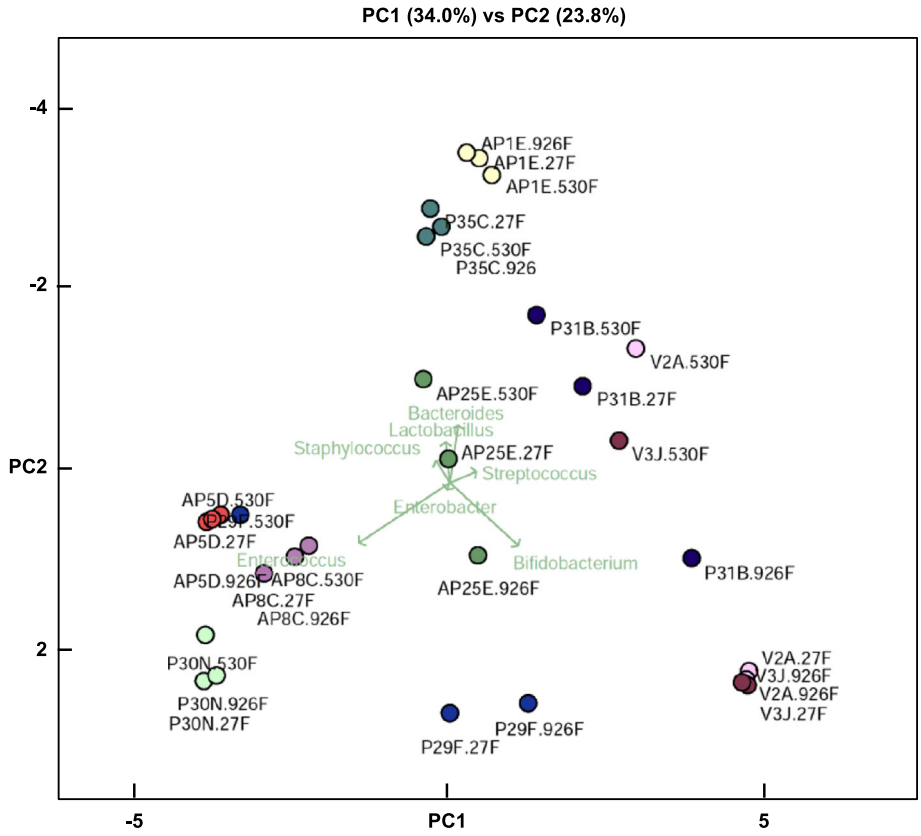


Figure 9 Principal Coordinate Analysis (PCoA) based on 16S rRNA community profiles analysed using PE protocol of the hypervariable regions tested

PCoA was performed based on the taxonomic assignments obtained from the 16S rRNA gene sequencing libraries analysed. Samples used for this plot were classified in main three groups: (i) premature infants without supplementation (AP1E, AP5D, AP8C, AP25C), (ii) premature infants with supplementation (P29F, P30N, P31B, P35C), and (iii) term baby samples (V2A, V3J). Samples names are coded highlighting the 16S rRNA gene library they belong. Sample names ending in (.27F) belong to 16S rRNA gene library prepared using primers 27F-519R (target region V1+V2+V3), sample names ending in (.530F) belong to 16S rRNA gene library prepared using primers 530F-926R (region V4+V5), and sample names ending in (.926F) belong to 16S rRNA gene library amplified using primers 926F-1394R (region V6+V7+V8). PCoA plot indicates that distribution of samples targeting (V4+V5) region was distinct from samples targeting (V1+V2+V3) and (V6+V7+V8).

When performing the same analysis using the QIIME pipeline results were comparable (Figure 10), again suggesting bioinformatics analysis does not play a major role in explaining these differences.

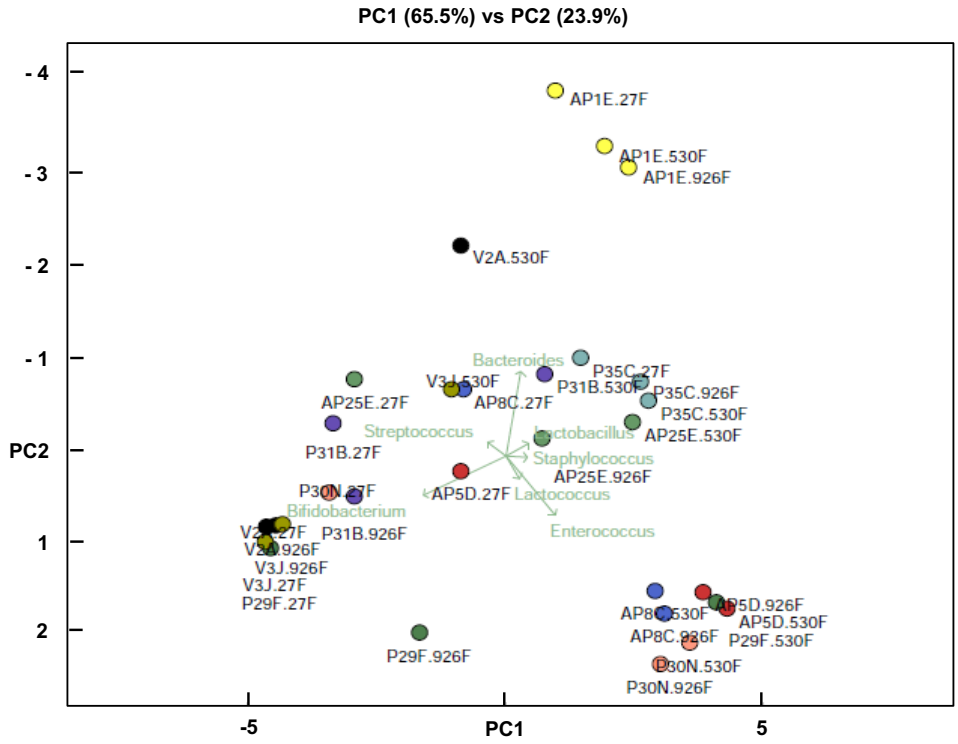


Figure 10 Principal Coordinate Analysis (PCoA) based on 16S rRNA community profiles analysed using QIIME of the hypervariable regions tested

PCoA was performed based on the taxonomic assignments obtained from the 16S rRNA gene sequencing libraries analysed. Samples used for this plot were classified in main three groups: (i) premature infants without supplementation (AP1E, AP5D, AP8C, AP25C), (ii) premature infants with supplementation (P29F, P30N, P31B, P35C), and (iii) term baby samples (V2A, V3J). Samples names are coded highlighting the 16S rRNA gene library they belong. Sample names ending in (.27F) belong to 16S rRNA gene library prepared using primers 27F-519R (target region V1+V2+V3), sample names ending in (.530F) belong to 16S rRNA gene library prepared using primers 530F-926R (region V4+V5), and sample names ending in (.926F) belong to 16S rRNA gene library amplified using primers 926F-1394R (region V6+V7+V8). PCoA plot indicates that distribution of samples targeting (V4+V5) region was distinct from samples targeting (V1+V2+V3) and (V6+V7+V8).

Finally, the Shannon diversity index on all samples was calculated, which is an indicator of sample diversity, and reflects how many different taxa are present, considering how evenly they are distributed. Results from this study indicated that the different hypervariable regions provided different indices (Figure 11); Shannon

indexes were higher in samples containing high amounts of *Bifidobacterium* such as V2A and V3J, which agrees with results obtained by PCoA analysis.

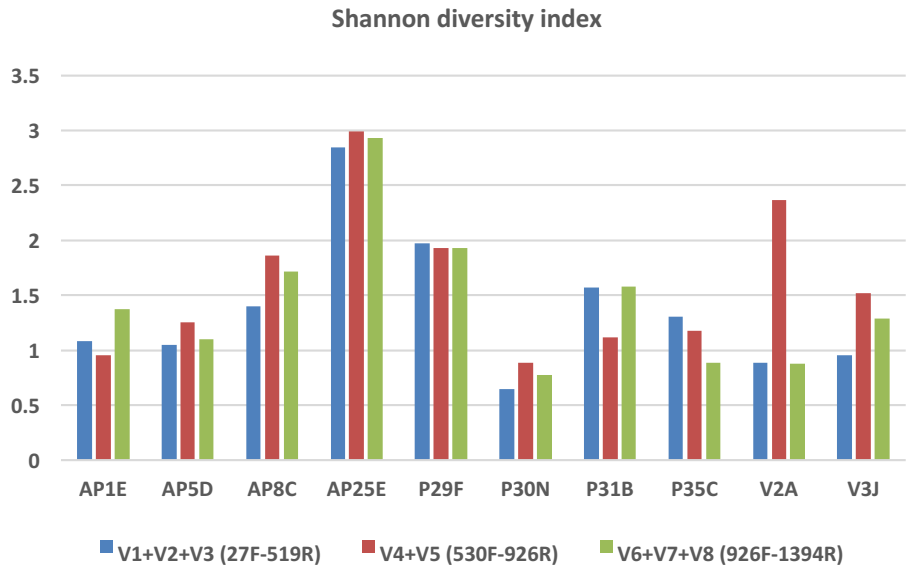


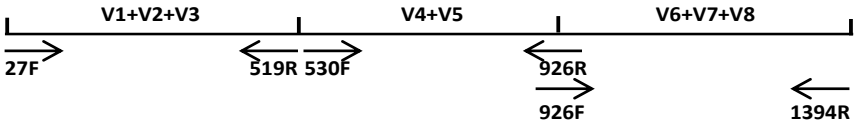
Figure 11 Shannon diversity index calculation of 16S rRNA gene sequencing data

Shannon diversity index was calculated using 16S rRNA bacterial community profiles for the three different 16S rRNA libraries tested in this study (27F-519R (region V1+V2+V3), 530F-926R (region V4+V5) and 926F-1394R (region V6+V7+V8)). Sequencing data was analysed using the PE protocol.

Validation of 16S rRNA gene primers used against *Bifidobacterium*

A primer aligning study was performed to evaluate the alignment of the primer pairs used to construct the 16S rRNA libraries to *Bifidobacterium bifidum* (species in the probiotic supplementation) and other bacterial members commonly found in ELBW infants. Surprisingly, this study did not reveal any obvious annealing disadvantage (mismatch) for any particular pair of primers against the bacterial taxa tested (Figure 12).

a



b

	22 41
16S rRNA *B. bifidum* CP010412	AGGGTTCGATTCTGGCTCAG
16S rRNA *S. epidermis* NR_074995	AGAGTTTGATCCTGGCTCAG
16S rRNA *E. cloacae* CP012165	AGAGTTTGATCATGGCTCAG
16S rRNA *E. faecalis* CP014949	AGAGTTTGATCCTGGCTCAG
16S rRNA *B. infantis* M58738.1	AGAGTTTGATCCTGGCTCAG
16S rRNA *B. longum* ATCC 156697	AGGGTTCGATTCTGGCTCAG
27Fmod primer	AGRGTTTGATCMTGGCTCAG

	536 553
16S rRNA *B. bifidum* CP010412	CAGCAGCCGCGGTAATAC
16S rRNA *S. epidermis* NR_074995	CAGCAGCCGCGGTAATAC
16S rRNA *E. cloacae* CP012165	CAGCAGCCGCGGTAATAC
16S rRNA *E. faecalis* CP014949	CAGCAGCCGCGGTAATAC
16S rRNA *B. infantis* M58738.1	CAGCNGCCGCGGTAATAC
16S rRNA *B. longum* ATCC 156697	CAGCAGCCGCGGTAATAC
519R primer	CAGCMGCCGCNGTAANAC

	540 555
16S rRNA *B. bifidum* CP010412	GTGCCAGCAGCCGCGG
16S rRNA *S. Epidermis* NR_074995	GTGCCAGCAGCCGCGG
16S rRNA *E. cloacae* CP012165	GTGCCAGCAGCAGCGG
16S rRNA *E. faecalis* CP014949	GTGCCAGCAGCCGCGG
16S rRNA *B. infantis* M58738.1	GTGCCAGCNGCCGCGG
16S rRNA *B. longum* ATCC 156697	GTGCCAGCAGCCGCGG
530F primer	GTGCCAGCMGCNGCGG

	930 949
16S rRNA *B. bifidum* CP010412	AAACTCAAAGAAATTGACGG
16S rRNA *S. Epidermis* NR_074995	AAACTCAAAGGAATTGACGG
16S rRNA *E. cloacae* CP012165	AAACTCAAATGAATTGACGG
16S rRNA *E. faecalis* CP014949	AAACTCAAAGGAATTGACGG
16S rRNA *B. infantis* M58738.1	AAACTCAAAGAAATTGACGG
16S rRNA *B. longum* ATCC 156697	AAACTCAAAGAAATTGACGG
bac926R primer	AAACTYAAARRAATTGACGG

	930 949
16S rRNA *B. bifidum* CP010412	AAACTCAAAGAAATTGACG
16S rRNA *S. Epidermis* NR_074995	AAACTCAAAGGAATTGACG
16S rRNA *E. cloacae* CP012165	AAACTCAAATGAATTGACG
16S rRNA *E. faecalis* CP014949	AAACTCAAAGGAATTGACG
16S rRNA *B. infantis* M58738.1	AAACTCAAAGAAATTGACG
16S rRNA *B. longum* ATCC 156697	AAACTCAAAGAAATTGACG
bac926F primer	AAACTYAAAKGAATTGACG

	1418 1432
16S rRNA *B. bifidum* CP010412	GTACACACCGCCCGT
16S rRNA *S. Epidermis* NR_074995	GTACACACCGCCCGT
16S rRNA *E. cloacae* CP012165	GTACACACCGCCCGT
16S rRNA *E. faecalis* CP014949	GTACACACCGCCCGT
16S rRNA *B. infantis* M58738.1	GTACACACNGCCCGT
16S rRNA *B. longum* ATCC 156697	GTACACACCGCCCGT
bac1394R primer	GYACACACCGCCCGT

Figure 12 Primer alignment study of the most common bacterial taxa found in ELBW (P29F)

a Representation of primers used in this study along the 16S bacterial rRNA gene. b Primer alignment study using 16S rRNA gene from *Bifidobacterium bifidum* CP 010412 (isolated from Infloran) and the most common bacterial taxa found in an ELBW infant (P29F) with supplementation (*Staphylococcus epidermis* NR_074995, *Enterobacter cloacae* CP012165 and *Enterococcus faecalis* CP014949). We also included two strains of *Bifidobacterium* as control samples (*B. infantis* M58738.1 and *B. longum* ATCC 156697) All sequences are represented in 5´-3´orientation using UPAC nucleotide code, where Y = C or T, R = A or G, K = G or T, M = A or C.

To complement this *in silico* analysis, the primer pair 530F and 926R (region (V4+V5)) was used to amplified genomic DNA isolated from seven different strains of bifidobacteria including the probiotic strain *B. bifidum* (Figure 13). Results from this experiment confirmed that this primer pair did not encounter any annealing problem when working with DNA isolated from pure strains.

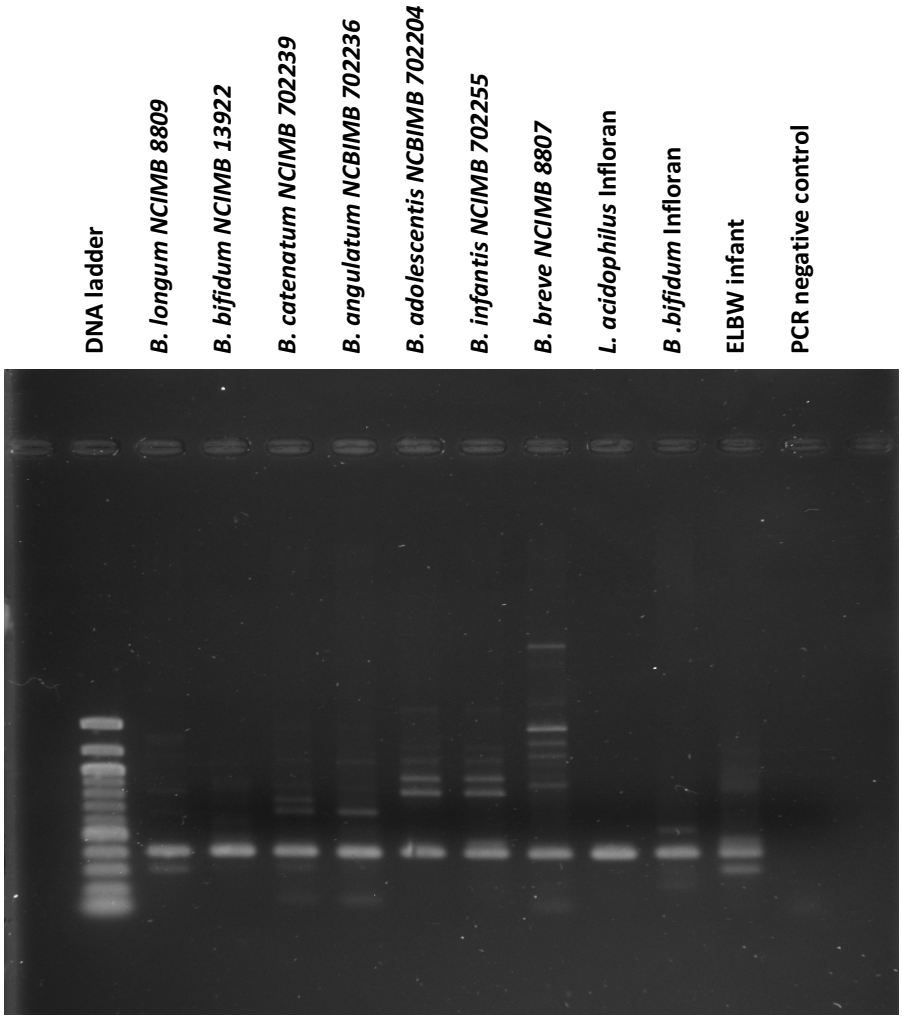
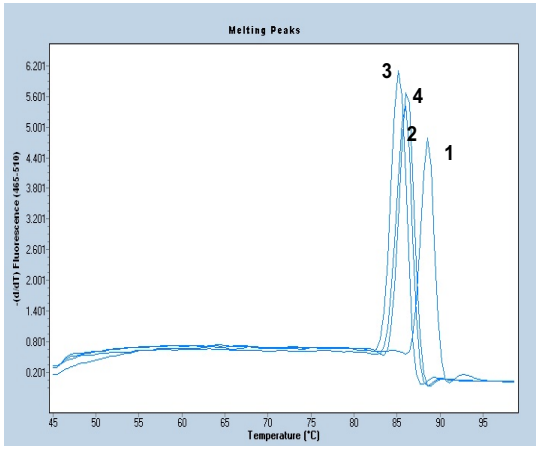


Figure 13 PCR amplification using primers 530F-926R on 8 *Bifidobacterium* strains

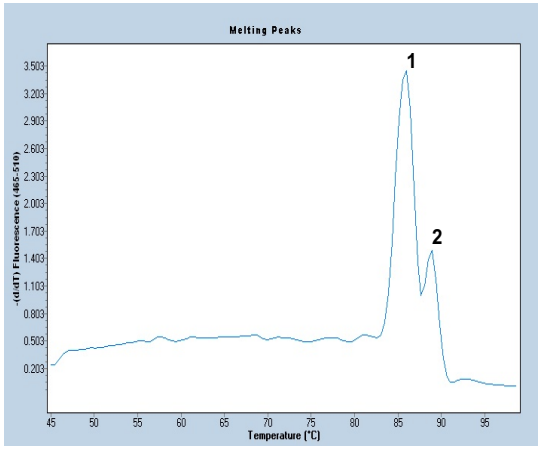
PCR amplification targeting the bacterial 16S rRNA gene using primers 530F-926R on Bifidobacterium collection strains. DNA extracted from *B. bifidum* (isolated from the commercial probiotic supplementation) and seven different *Bifidobacterium* NCIMB collection strains, was amplified using primers 530F-926R. Positive controls for this study were a faecal metagenomic sample and a *L. acidophilus* strain (isolated from the probiotic supplementation). Pure water was used as negative control. Amplicon samples were run on 1% agarose gel for 30 minutes at 100 V. DNA was visualised under UV light after staining with ethidium bromide. All tested samples gave a PCR product.

Subsequently, the melting curves of the amplicons generated using primers 530F-926R were calculated, to check the GC content of region (V4+V5). Templates selected for this PCR were the strains used in the probiotic supplementation *(B. bifidum* and *L. acidophilus*), and two other species which were overrepresented by region (V4+V5), *Enterococcus faecium* and *Streptococcus infantarius.* As expected, *B. bifidum* contained the highest GC content, in comparison with the other strains (Figure 14a). When running the same PCR, this time using a mixed DNA sample as template DNA (in order to simulate the conditions of a metagenomic sample) of the four bacterial strains used previously (*B. bifidum, L. acidophilus, E. faecium and S. infantarius*), primers 530F-926R preferentially annealed to the other bacterial genomic DNA, showing peak 1 in Figure 14b, when compared to peak 2 which corresponded with the melting temperature of *B. bifidum*.

a Amplicon melting curves from single bacterial strains



b Amplicon melting curve from a mixed DNA sample



1	***Bifidobacterium bifidum*** **(Infloran isolate)**	**Tm (°C)= 88.60**
2	***Lactobacillus acidophilus*** **(Infloran isolate)**	**Tm (°C)= 85.90**
3	***Streptococcus infantarius*** **(preterm isolate)**	**Tm (°C)= 85.24**
4	***Enterococcus faecium*** **(preterm isolate)**	**Tm (°C)= 86.20**

Peak 1 Tm (°C)= 85.89
Peak 2 Tm (°C)= 88.93

Figure 14 Melting curves of PCR amplicons from probiotic strains and bacterial isolates

a Melting curves of PCR amplicons from probiotic strains (*Bifidobacterium bifidum* and *Lactobacillus acidophilus)* and bacterial isolates (*Enterococcus faecium* and *Streptococcus infantarius*). Primers used to generate these amplicons were 530F-926R targeting region (V4+V5). *Bifidobacterium bifidum* displayed the highest melting temperature. b Melting curve of PCR amplicon obtained from a mixed DNA sample (5ng of *Bifidobacterium bifidum*, 5 ng of *Lactobacillus acidophilus*, 5 ng of *Enterococcus faecium* and 5 ng of *Streptococcus infantarius*). Primers used to generate these amplicons were 530F-926R targeting region (V4+V5). Peak name (2) presents a melting temperature (Tm) similar to the melting temperature (Tm) obtained for *B. bifidum*.

Validation of 16S rRNA gene sequencing data using shotgun metagenomics analysis

To validate the 16S rRNA gene sequencing data, whole genome shotgun sequencing was performed on two ELBW infants (one receiving supplementation and another one without it), and one term baby sample as control. Whole genome shotgun sequencing allows the entire DNA content to be sequenced with less biases, therefore it can be considered as 'gold-standard'. Results confirmed the presence of the predominant bacterial populations detected using 16S rRNA gene sequencing, namely *Bifidobacterium, Enterococcus, Staphylococcus, Enterobacter* and *Streptococcus* (Figure 15).

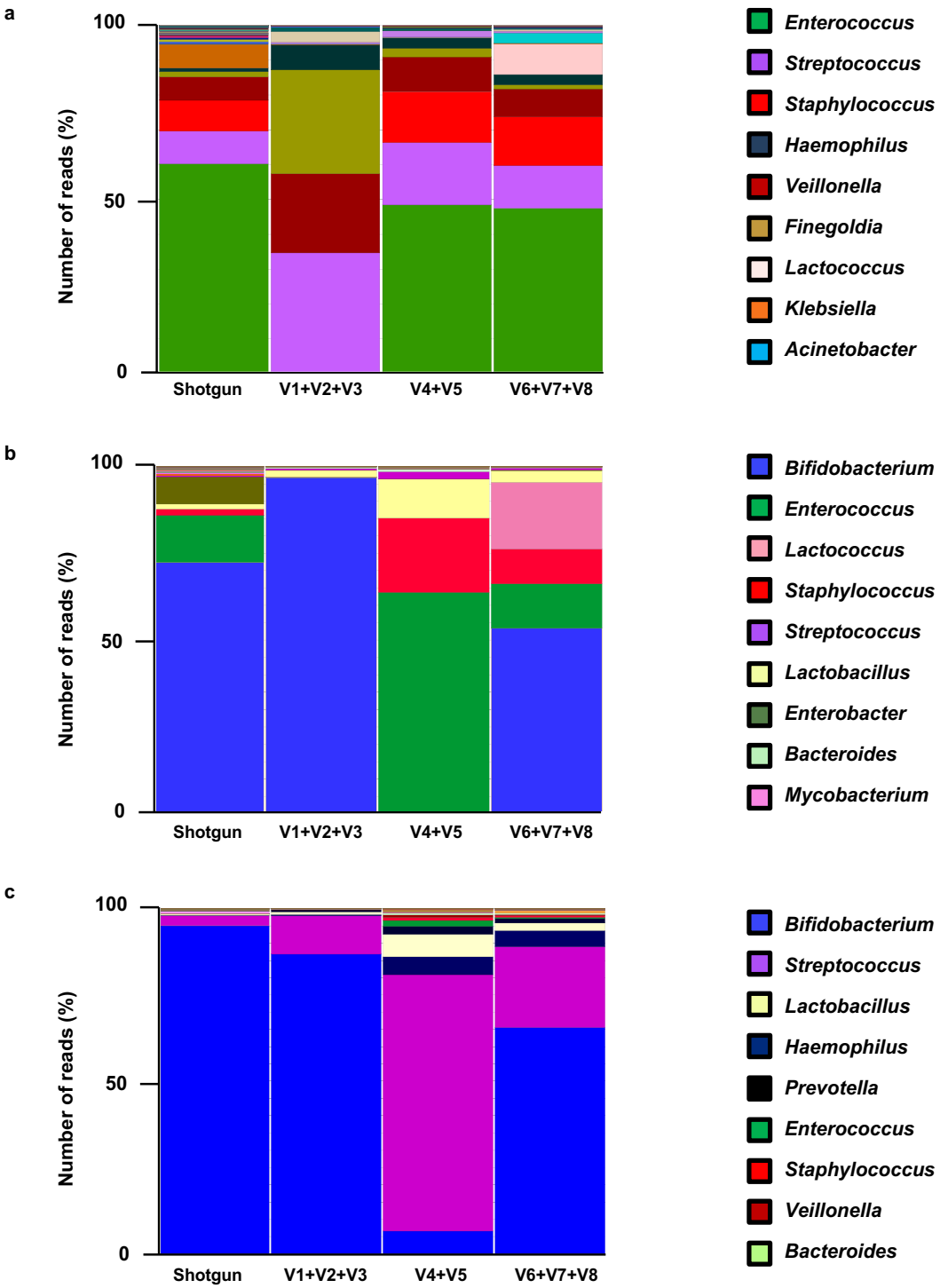


Figure 15 Bacterial community profiles determined by shotgun and 16S rRNA gene sequencing data

Comparison of bacterial profiles analysed by shotgun and 16S rRNA gene sequencing data. Relative abundance of the bacterial taxa was represented in percentages of number of reads. Bar colours represent different genus taxa, and bar lengths signify the relative abundance of each taxon. 16S rRNA bacterial profiles are named according to the different 16S rRNA hypervariable region amplified: (i) (V1+V2+V3, primers 27F-519R), (ii) (V4+V5, primers 530F-926R) and (iii) (V6+V7+V8, primers 926F-1394R). **a** Bacterial community profiles determined by shotgun and 16S rRNA gene sequencing from an ELBW infant (sample AP8C) with no supplementation. **b** Bacterial community profiles determined by shotgun and 16S rRNA gene sequencing from an ELBW infant (sample P29F) with supplementation. **c** Bacterial community profiles determined by shotgun and 16S rRNA gene sequencing from a term baby (sample V3J).

The additional coverage that shotgun provides at species level (Figure 16) enabled confirmation of *B. bifidum* (present in the probiotic supplementation) in sample P29F (ELBW infant receiving supplementation). *Lactobacillus acidophilus* (present in the supplementation) was not detected in the sample tested (P29F).

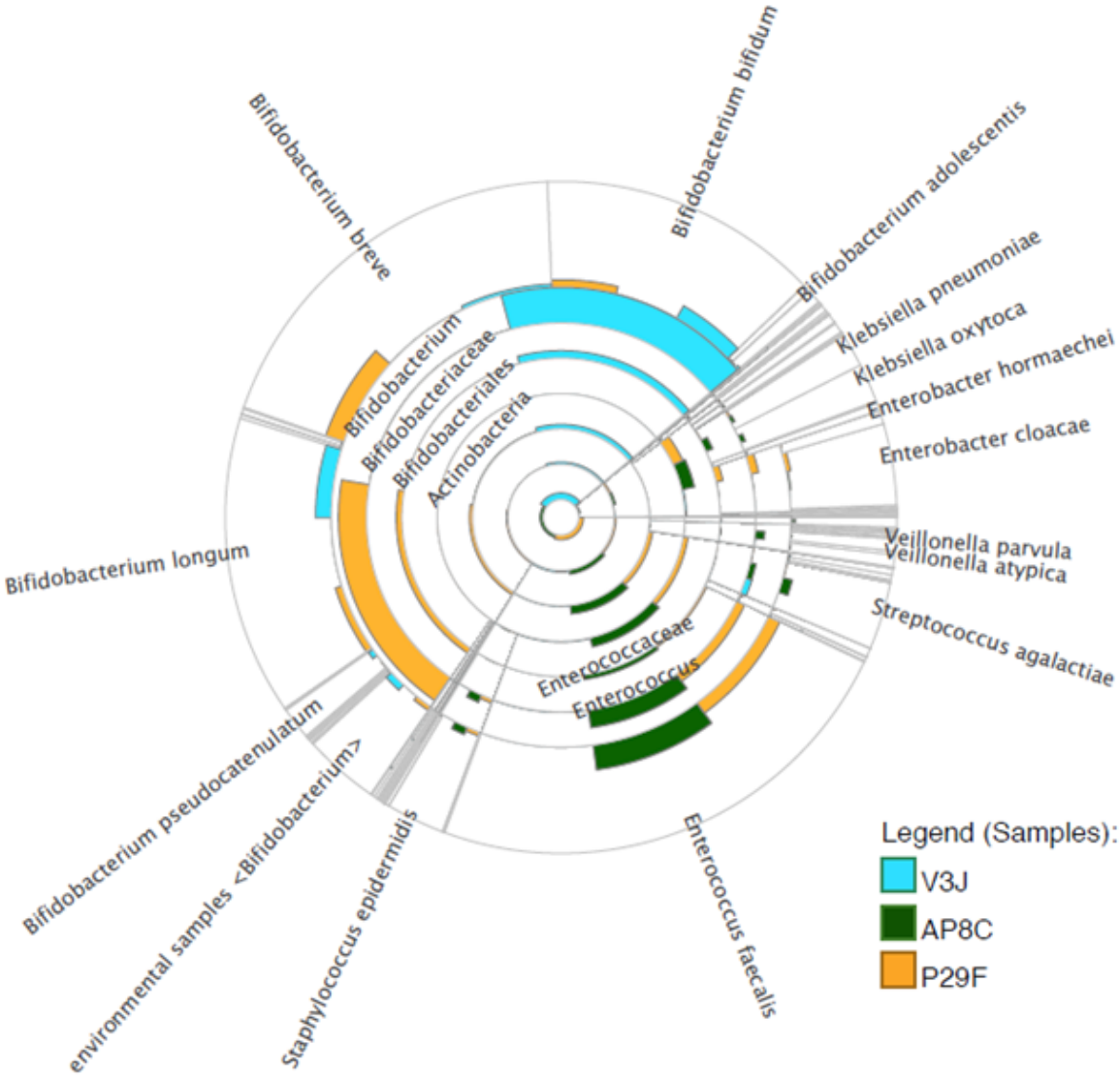


Figure 16 Shotgun taxonomic profiles from two ELBW infants with/without supplementation and a term infant

Radial taxonomic tree displaying shotgun community profiles from faecal samples of an ELBW infant with no supplementation (AP8C, represented in green) an ELBW infant with supplementation (P29F, represented in yellow) and a term baby (V3J, represented in blue). Relative abundance was indicated according to the length of the coloured bars in the figure. The centre of the radial tree indicates phylum level, and the subsequent concentric layers of the radial tree indicate class, order, family, and genus and species level. Term baby (V3J) and ELBW infant with supplementation (P29F) samples presented a higher abundance of *Bifidobacterium* when compared to an ELBW infant with no supplementation (AP8C).

Complementing the shotgun analysis, functional analysis (Figure 17) was performed on the same sample, which indicated increased representation of glycan

metabolism pathways. Some *Bifidobacterium* strains such as *B. bifidum* have been previously shown to metabolise breast milk-derived human milk oligosaccharides, which are known to contain these glycan structures [121].

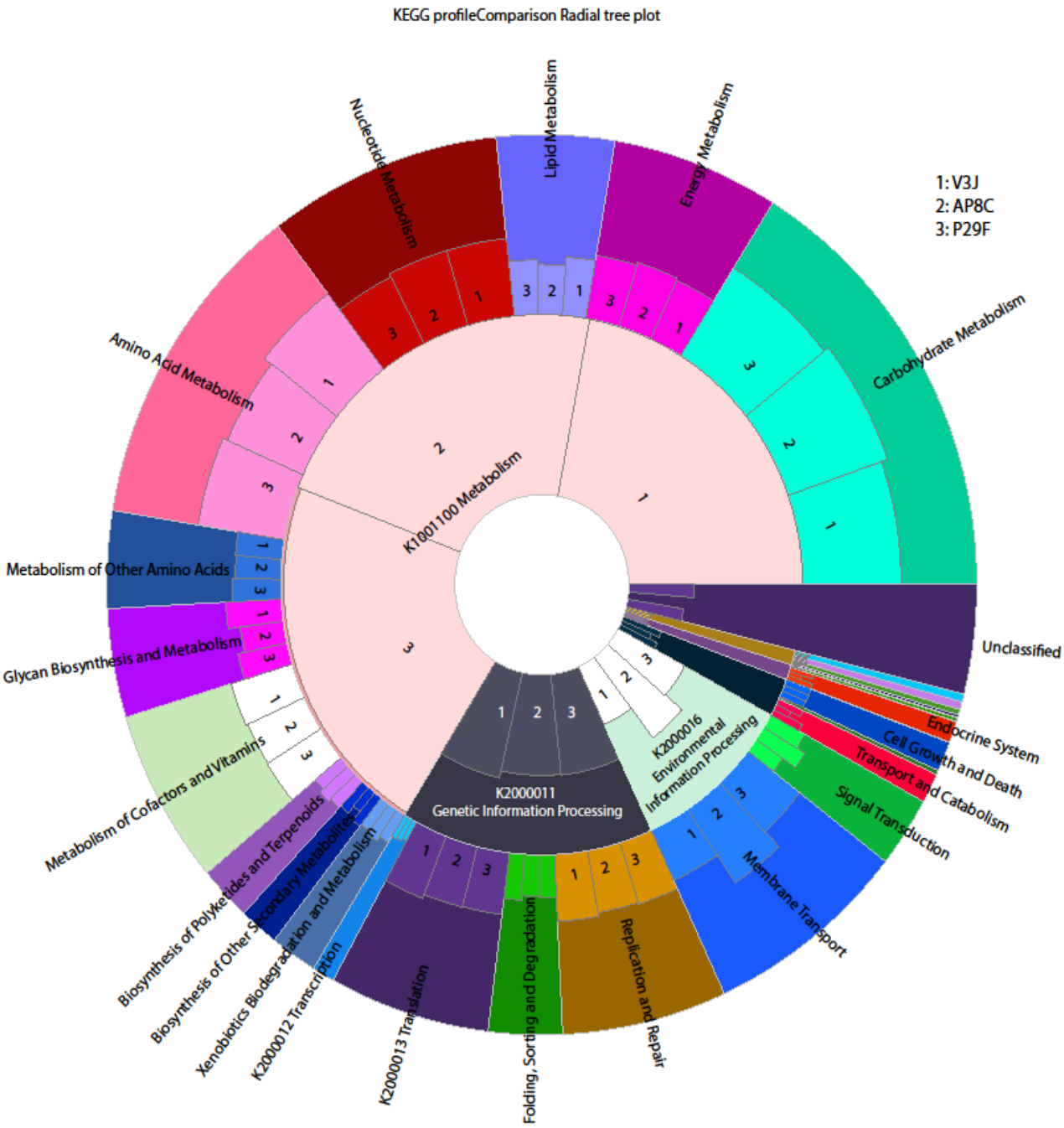


Figure 17 Shotgun functional profiles from two ELBW infants with /without supplementation and a term infant

Radial tree displaying shotgun functional profiles from an ELBW infant with no supplementation (sample AP8C, represented by the number 2 in the figure) an ELBW infant with supplementation (sample P29F, represented by the number 3 in the figure) and a term baby sample (sample V3J, represented by the number 1 in the figure). Functional analysis was performed using the KEGG pathway analysis.

Results from the whole genome shotgun sequencing analysis confirmed previous 16S rRNA metataxonomic analysis, in that (V4+V5) region failed to adequately discriminate gut bacterial populations in ELBW infants. This region overrepresented *Streptococcus, Enterococcus* and *Lactobacillus* populations, and underrepresented *Bifidobacterium* in comparison to the other regions tested. The percentages of the number of reads obtained for one ELBW infant no supplementation, one ELBW infant with supplementation, and one term baby are indicated in Table 6.

Table 6 Percentage of number of reads obtained for shotgun and 16S rRNA gene sequencing

Bacteria taxa genus	**AP8C Shotgun**	**AP8C 27F-519R**	**AP8C 530F-926R**	**AP8C 926F-1394R**	**P29F Shotgun**	**P29F 27F-519R**	**P29F 530F-926R**	**P29F 926F-1394R**	**V3J Shotgun**	**V3J 27F-519R**	**V3J 530F-926R**	**V3J 926F-1394R**
Bifidobacterium	0.02	0.70	0.11	1.03	72.39	37.33	0.60	49.56	95.01	81.00	6.46	75.31
Enterococcus	61.75	69.54	55.17	59.87	13.60	34.19	53.35	22.94	0.02	0.65	1.83	0.94
Enterobacter	0.26	0.37	0.66	0.74	7.87	18.29	13.43	15.68	0.00	0.06	0.13	0.12
Staphylococcus	8.98	15.70	17.03	17.53	1.72	7.29	19.52	7.86	0.01	0.37	0.85	0.48
Lactobacillus	0.09	0.42	2.14	0.63	1.45	2.10	9.86	3.24	0.36	1.89	6.61	2.00
Streptococcus	9.20	10.86	19.62	15.87	0.32	0.26	1.81	0.41	2.95	14.63	74.26	15.65
Escherichia	0.30	0.02	0.05	0.00	0.28	0.25	0.06	0.00	0.10	0.04	0.06	0.11
Actinomyces	0.00	0.00	0.08	0.00	0.02	0.00	0.07	0.00	0.06	0.00	0.10	0.02
Acinetobacter	0.02	0.03	0.02	0.00	0.00	0.02	0.02	0.00	0.00	0.02	0.01	0.00
Bacteroides	0.00	0.25	0.89	0.33	0.00	0.20	0.71	0.22	0.00	0.21	0.57	0.30
Granulicatella	0.00	0.00	0.19	0.00	0.00	0.00	0.14	0.00	0.00	0.00	0.01	0.00
Haemophilus	0.99	1.96	3.48	3.42	0.00	0.00	0.07	0.03	0.09	0.38	5.25	3.41
Lachnoclostridium	0.00	0.03	0.08	0.04	0.00	0.02	0.07	0.02	0.00	0.02	0.04	0.02

To visualise if there were strong patterns between the 16S rRNA bacterial profiles of the different hypervariable regions tested (V1+V2+V3, V4+V5 and V6+V7+V8), and the shotgun sequencing data (used as gold standard), a Principal Coordinate Analysis (Figure 18) was performed using the PE protocol and QIIME pipelines. This investigation confirmed that region (V4+V5) presented the most diverse distribution among samples when *Bifidobacterium* was a resident member

of the gut microbiome (*e.g.* differences were greater in sample P29F belonging to an ELBW with probiotic supplementation and sample V3J from a term infant sample).

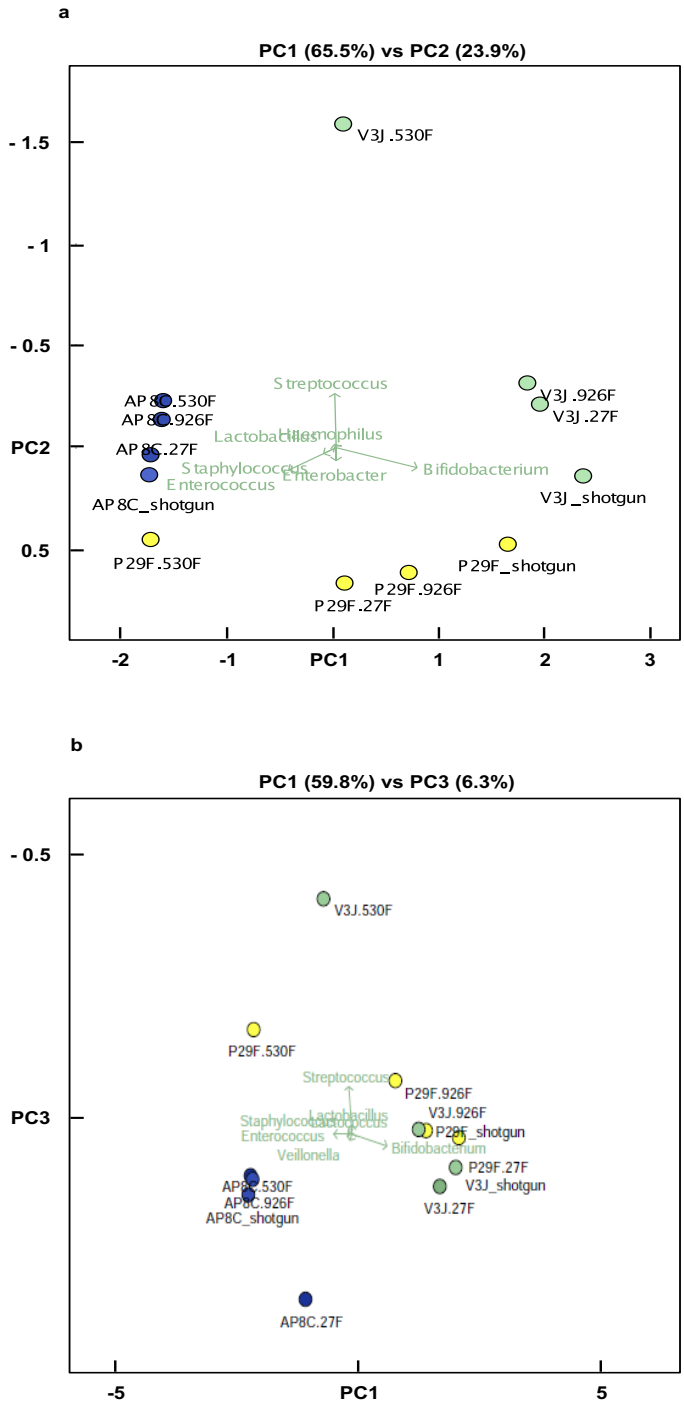


Figure 18 PCoA plots based on 16S rRNA gene sequencing and shotgun data

a PCoA based on 16S rRNA gene sequencing data analysed using PE protocol. **b** PCoA based on 16S rRNA gene sequencing data analysed using QIIME. Blue circles represent sequencing data of premature infants without supplementation (AP8C), yellow circles data of premature infants with supplementation (P29F), and green circles data of term baby (V3J). Each sample was analysed using three different 16S rRNA gene libraries (.27F, targets region (V1+V2+V3), .530F targets region (V4+V5), and .926F targets region (V6+V7+V8)). Samples ended with (_shotgun) represents shotgun data used as 'gold standard' in this study.

Discussion

Metataxonomic microbiome profiling using 16S rRNA gene sequencing is a cost-effective amplicon sequencing method ideal for use when profiling large samples numbers (*i.e*. clinical trials). There are several studies examining the gut microbiome using 16S rRNA gene sequencing in adults and term infants that have highlighted that the DNA extraction method can significantly impact the representative profile of the bacterial community obtained[79]. Furthermore, the 16S rRNA hypervariable region (V1 to V9) targeted can influence the ability to distinguish between different bacterial taxa[104] as well as the bioinformatics pipeline chosen for the data analysis. When profiling samples from at-risk cohorts (such as ELBW infants), it is essential to optimise and standardise sample preparation, sequencing methods, and bioinformatic tools among clinical studies, particularly for studying the influence of microbiota therapies (*e.g.* probiotic supplementation) on community profiles and health outcomes.

In this study, an optimal protocol for 16S rRNA profiling of premature infants' samples is described standardising the following specific steps: (i) DNA extraction method, (ii) primer choice and hypervariable region of the 16S rRNA gene, and (iii) bioinformatics pipeline. The first step in this pipeline evaluated different DNA extraction methods, which revealed the Fast DNA Spin Kit for Soil, including a bead-beating step, was the best method for extracting high quantity and quality DNA from ELBW infants (Table 5). Bead-beating has been previously been shown to facilitate disruption of cell membrane components such as high molecular weight capsules, and these results extend these findings to ELBW infant samples. Furthermore, extending the bead-beating time to 3 min lead to higher DNA yields from all samples, particularly for *Bifidobacterium* supplemented ELBW and *Bifidobacterium*-rich term infants. This highlights that samples expected to have high *Bifidobacterium* levels (genus known to express exopolysaccharide capsules)[122] are optimally processed using an extended bead-beating DNA extraction protocol.

One of the critical stages when comparing microbiome studies is the influence that bioinformatics pipelines can have on the taxonomic assignments obtained when bacterial population are at lower levels. Faecal samples from ELBW infants generally exhibit low bacterial diversity, and therefore excellent sequencing coverage is obtained, and this may indicate why both bioinformatics pipelines used (OTU vs QIIME) showed similar trends (Figure 5), however this may be somewhat different if a more complex (*e.g.* adult) sample was compared. Interestingly, for low relative abundance bacterial populations such as *Acinetobacter* or *Haemophilus,* the results indicated small differences between the pipelines. This may be explained by the fact that the OTU approach accepts all assignments of taxa even when only one OTU is assigned (which may result in false positives), whereas the paired end protocol discards low-confidence taxa (assignments <25 reads). Therefore, it is important to be aware of these differences in bioinformatics pipelines when studying low abundance bacterial populations.

Characterisation of which hypervariable region of the 16S rRNA gene was most suitable for representing the gut microbiome of ELBW infants, was also determined in this study (Figures 7-8). Three hypervariable regions of the 16S rRNA gene (V1+V2+V3), (V4+V5) and (V6+V7+V8) were evaluated, and results revealed an underrepresentation of *Bifidobacterium* when amplifying region (V4+V5). This is in agreement with previous studies that also highlighted problems with amplifying (V4+V5) region of the 16S rRNA gene for *Bifidobacterium*-rich samples when using faeces samples from adults and infants [123,124].

To probe why these striking differences were observed just for this 16S rRNA region, primer alignment studies (Figure 13) surprisingly did not reveal any specific mismatched with primers amplifying region (V4+V5, 530F-926R) for eight different *Bifidobacterium* strains, which suggests the DNA template used in the PCRs may be the issue. Indeed, several studies have described that templates containing a high GC content are more difficult to amplify than non-GC-rich

templates [125, 126] and in the context of a metagenomic sample where different genomes are competing against the same pair of primers this factor could play a role. Notably, the genus *Bifidobacterium* contains a high GC genome content (56-67%), and *B. bifidum* (present in the supplementation) and contains a higher GC content in region (V4+V5) than other strains commonly present in the ELBW infant gut microbiome (Figure 14a). This may lead into an underrepresentation of *Bifidobacterium* when it is present in a metagenomic sample, and other studies using the same region (V4+V5), but different primers, have also encountered an underrepresentation of *Bifidobacterium*[123]. It is also interesting to highlight that primer 926R presented the lowest GC content among the primers used in this study and does not have a strong GC clamp at its 3'end, which could as well interfere with the binding to genomes with high GC content. Among the most common bacteria found in premature infants *Bifidobacterium* (GC ~ 60%), *Lactobacillus* (GC ~ 60%) and *Corynebacterium (GC~53%)* would more affected this by this issue.

Finally, the 16S rRNA sequencing data was further benchmarked to shotgun sequencing by analysing a subset of the samples we used for 16S rRNA analysis (AP8C, P29F and V3J, Figure 15). Notably, shotgun metagenomics introduces less PCR biases and artefacts, but is significantly more expensive to scale up and requires additional computing power for downstream analysis, which in large-scale *in vivo* and clinical studies are important factors to consider. From a sample collection stand-point, shotgun also requires a higher yield of bacterial DNA (500 ng is the recommended amount of DNA compared to 25 ng required for 16S rRNA gene sequencing) which can be challenging to obtain from case-specific ELBW infants (*e.g.* at an early time point of the study where most of infants are administered antibiotics). Results from shotgun metagenomics analysis validated an underrepresentation of *Bifidobacterium* in region (V4+V5) and over-representation of *Lactobacillus* (Table 6).

Conclusions

This study highlights the importance of selecting the optimal DNA extraction method when using 16S rRNA gene sequencing to analyse metagenomic samples; *i.e.* include a bead-beating step to lyse capsulated bacteria such a bifidobacteria, and is now considered gold standard by many research teams [127]. We additionally underlined the influence that bioinformatics pipelines may have at detecting bacterial population present in low numbers.

Appropriate primer selection when using 16S rRNA gene profiling is essential for analysing gut-associated metagenomic samples. Region (V4+V5) should be avoided in metagenomics studies where the genus *Bifidobacterium,* either resident or supplemented, is evident. More specifically, it was demonstrated differences in GC content of the (V4+V5) region of the 16S rRNA gene between *Bifidobacterium* and other low GC content bacterial populations present in the ELBW infant gut microbiome (*e.g. Streptococcus* and *Enterococcus*), significantly biases profiling in mixed bacterial communities.

Future work

16S rRNA gene sequencing is a more cost-efficient method than shotgun metagenomics and can be incredibly useful in large scale projects (*e.g.* clinical trials) with hundreds or thousands of samples. This metataxonomic profiling provides the added advantage that it can sequence samples with very low bacterial content, such as in ELBW infants, due to the PCR step. This optimised pipeline represents a good choice and has been used in large clinical trials such as the BAMBI study (Chapter 2).

To complement this study, it would be interesting to analyse in more depth the bacterial populations present at lower levels, to potentially trace bacterial pathogens at initial stages of infection. This could initially be done by inoculating

different amounts of a ´known´ pathogen into the faecal sample, to estimate the sensitivity of the method. Once the sensitivity is established, a more practical outcome could be added to the study, by using clinical samples from ELBW infants profiled showing early signs of infection.

An important factor which has not been covered in the present study is that bacterial species differ in their copy number of the 16S rRNA gene[128]. This can have a confounding effect on the ´relative abundance´ of the microbial community profile obtained. Notwithstanding, it was decided not to correct for copy number of the 16S RNA gene as the main aim of this study was to assure most of the bacterial communities comprising the preterm microbiome were detected. However, if we were to perform a study focusing on bacterial populations present at low abundance, correction for 16S rRNA gene copy number would be important, and publicly available bioinformatics tools can be used to correct for copy number variation (*e.g.* Copyrighter[129] or rrNDB[130]).

In the near future the use of 16S rRNA gene sequencing for microbial identification will likely be replaced by shotgun genome sequencing. Sequencing the whole genomes of the microbial communities provides a more complete picture (and it is not restricted to bacterial members, but also includes fungi, archaea and viruses) and has great potential to be applied in the clinical field (*e.g.* to identify antibiotic resistance profiles or virulence traits). Throughout my PhD, I have been fortunate to explore these options, using shotgun metagenomics in combination with: (i) the most widely used sequencing technology (Illumina, Chapter 3) and (ii) the fastest sequencing technology (MinION sequencing, Chapter 4).

Chapter 2

Title: *Bifidobacterium* and *Lactobacillus* supplementation modifies the microbiome and metabolome of premature infants residing in Neonatal Intensive Care Units

Abstract

Supplementation with early life bacterial members or 'probiotics' is becoming increasingly popular in neonatal intensive care units (NICUs), as a way of manipulating the gut microbiome of premature infants. The findings of numerous clinical studies support that this practice reduces the incidence of sepsis and necrotising enterocolitis (NEC) in premature infants[131, 132]. However, very few of these studies have explored how this supplementation modifies the early life microbiome, and none have done this for large patient cohorts using a combination of microbiome profiling and other characterisation approaches e.g. metabolomics and whole genome sequencing of bacterial strains. Thus, further studies are required if standardised guidelines are to be introduced for implementing this as routine clinical care for premature infants.

The work presented here includes a large longitudinal study from two different cohorts of premature infants; 101 orally supplemented with *Bifidobacterium* and *Lactobacillus* (Bif/Lacto group) and 133 non-supplemented (control group). This work sought to determine the impact of this supplementation on the wider gut microbiome via 16S rRNA metataxonomic profiling, and additionally the impact on the metabolome using paired faecal samples and untargeted NMR metabolomics analysis. Microbiome profiling on 591 samples indicated higher relative abundances of beneficial *Bifidobacterium,* and lower relative abundances of potentially pathogenic *Klebsiella, Staphylococcus* and *Escherichia* in the Bif/Lacto group, when compared to non-supplemented control premature infants. A subset of these samples (n=157) analysed with NMR revealed elevated levels of lactate and

acetate in the Bif/Lacto group which correlated to a lower faecal pH, whilst the control group presented high level of human milk oligosaccharides (3-Fucosyllactose, 2- Fucosyllactose). In addition, whole genome sequencing on the *Bifidobacterium* supplemented strain confirmed preterm gut colonisation, and ability to utilise human milk oligosaccharides from breastmilk. Phenotypic antibiotic testing suggested the *Bifidobacterium* strain used in this study was susceptible to most commonly prescribed NICU antibiotics.

This study demonstrates that probiotic supplementation can modify the preterm microbiome and the gastrointestinal environment to more closely resemble the gut microbiomes profiles found in full-term infants[133, 134]**.** A graphical abstract summarising the different stages of this study is shown in Figure 19.

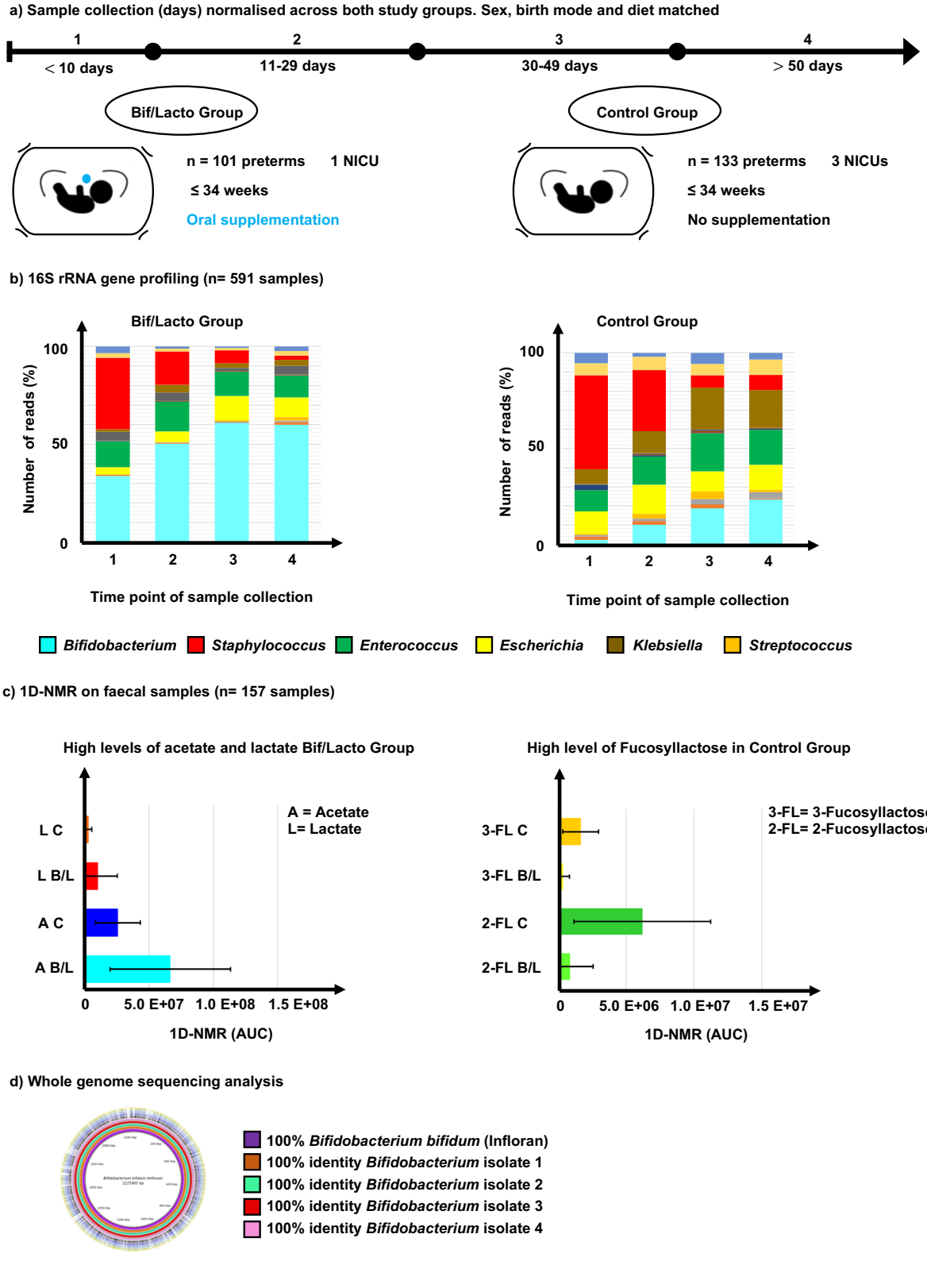


Figure 19 Graphical abstract of the study

a, Study outline and sample collection times. The study comprised two groups: (i) Bif/Lacto Group (received oral supplementation containing *Bifidobacterium* and *Lactobacillus*, and (ii) control group which was not administered any probiotic supplementation **b,** Taxonomic profiles showing percentage of number of reads of the most common bacterial taxa at each study time point (1 = <10 days from birth, 2 = 11-29 days, 3 = 30-49 days and 4 = >50 days) in the control group and the Bif/Lacto group **c,** Summary of main faecal metabolites derived from ^{1}H-NMR spectra in the Bif/Lacto group and control group **d,** Comparison of whole genome sequencing analysis of *Bifidobacterium bifidum* from Infloran and other *Bifidobacterium bifidum* isolates.

Authors' contributions

This large body of work is available as a preprint, and has been submitted to Cell Reports Medicine where I am first author. I was involved from the very beginning of this large multi-centre cohort study (recruiting 233 at-risk premature infants), contributing to overall study design, contact point for clinical leads, sample preparation for genomic and metabolomic analysis, data analysis, figure preparation, and drafting of the manuscript. This chapter details the parts of the study I have specifically led. Specific details of authors' contributions are highlighted in the method section of this chapter.

Introduction

Microbial colonisation of the gut during the early life developmental window plays an instrumental role in the maturation of the immune system, nutrient acquisition and pathogen exclusion (Chapter 1). Immediately after birth, and during the first days of life, the intestine of the newborn infant is colonised by bacteria residing in the mother's birth canal and the environment, a mixed inoculum of facultative anaerobic bacteria (*Enterobacter*, *Enterococcus*, *Staphylococcus* and *Streptococcus*)[18] able to grow under the presence of low amounts of oxygen. As oxygen diminishes within the gut environment this then allows obligate anaerobic bacteria to thrive (*e.g. Bifidobacterium*, *Bacteroides* and *Clostridium*)[19]. These microbial pioneers are key players in shaping the rapidly changing ecosystem until a climax or 'adult-like' microbiome is established at 2-3 years of age[20, 21].

Among these pioneers, *Bifidobacterium* comprise the most abundant group (representing up to 70-90% based on relative abundance of this genus in breast-fed infants), and importantly species and strains of this genus have been shown to interact with immune cells[135], which promotes immune development and maturation.

In the case of premature infants (born before 37 weeks' gestation), this natural process of bacterial gut colonisation is often disrupted due to a variety of factors including; C-section delivery, prolonged hospitalisation, and prescription of multiple courses of antibiotics. Due to these microbial-altering factors, premature infants have a particularly perturbed early life gut microbiome, with sequencing studies indicating low relative abundance of commensals (*Bifidobacterium* and *Lactobacillus*), while the relative abundace of potential pathogenic bacteria is increased (*Enterobacter, Enterococcus, Escherichia and Klebsiella*). Crucially, these microbiome disturbances in premature infants increase the risk of developing necrotising enterocolitis (NEC, with a 5-12% incidence in premature infants with birthweight lower than 1500 gr[136]) and late onset sepsis[43] (LOS, 15-25%). Furthermore, these initial gut microbiome disruptions, also increase the risk of developing allergies, asthma, or eczema during childhood and as an adult[40, 41]. Therefore, modulating the preterm gut microbiome to increase the presence of beneficial commensal bacteria could overcome these disruptive effects.

Oral supplementation of *Bifidobacterium* and *Lactobacillus* from birth in premature infants is becoming an alternative and cost-effective gut microbiota therapy while the infant resides in the Neonatal Intensive Care Unit (NICU). There are several clinical studies which support that oral supplementation of commensal bacterial such as *Bifidobacterium* or *Lactobacillus* exert a protective effect against NEC[137, 138] despite the heterogeneity of results and the need for larger longitudinal studies[139]. Notably, the largest ever neonatal probiotic trial (PiPS study, n=1310 infants recruited[140]) found no evidence of benefit for prevention of NEC and LOS in premature infants when using supplementation with *Bifidobacterium breve BBG1*. However, shortly after the publication, various researchers highlighted several inconsistencies[141] in this study: (i) the probiotic dose given to the premature infants was lower (100 million CFU) than the one used in the pilot study which indicated beneficial effects (1 billion CFU), (ii) the PiPS study reported a cross-colonisation of 49% of placebo samples at all trial sites by 36 weeks' gestation, and

(iii) some authors question whether the size of the study cohort was sufficiently large to draw conclusions. These points serve to highlight there is currently not clear standardised guidelines to administer 'probiotics' for premature infants. Furthermore, most of the published clinical studies using microbiota therapy have not explored how this supplementation modifies the early life preterm microbiome, which is required if robust guidelines are to be introduced for implementing this as routine clinical care for premature infants.

The work presented here analysed two different cohorts of premature infants, one group supplemented with *Bifidobacterium* and *Lactobacillus* strains (Bif/Lacto group), and a control non-supplemented group (control group). Infants in the Bif/Lacto group were routinely prescribed an oral supplementation containing 10^9 colony forming units (CFU) of *Bifidobacterium bifidum* and 10^9 CFU of *Lactobacillus acidophilus* (Infloran®, Desma Healthcare). A recent clinical audit of the NICU administering this supplementation found a more than 50% reduction in NEC rates and late-onset sepsis when comparing 5-year epochs before and after introducing this probiotic supplementation[142]. Cohorts were matched by gestational age (< 34 weeks gestation), sex, birth mode, time points of sample collection, and diet across four different NICUs. 16S rRNA gene profiling was used to obtain an overview of their gut microbiomes (n = 591). Complementing this analysis and to evaluate the colonisation of the supplemented *Bifidobacterium* strain, genomes of different *Bifidobacterium* strains isolated from stool samples were compared to the supplemented strain using whole genome sequencing analysis. Finally, paired metabolomic analysis on a subset of stool samples from both cohorts (n=157) evaluated the effects of *Bifidobacterium* and *Lactobacillus* supplementation on the metabolite profile from both study groups.

Hypothesis and aims

This study aims to investigate the outcome of *Bifidobacterium* and *Lactobacillus* supplementation on the microbiome and metabolome of premature infants residing in NICUs.

Hypothesis: Microbial supplementation with the early life members *Bifidobacterium* and *Lactobacillus* facilitates colonisation of these strains in the preterm gut modifying the wider microbiome and metabolome.

The study involved three aims/sub-sections:

a) A metataxonomic analysis using 16S rRNA gene profiling (n=591 samples) aimed to (i) characterise the development of the preterm microbiome for both study groups (supplemented and non-supplemented), and (ii) evaluate the interaction of environmental factors such as antibiotic treatment, delivery mode, or diet.

b) Whole genome sequence analysis of ten *Bifidobacterium* isolates from the supplemented group, aimed to investigate the colonisation of the supplemented *Bifidobacterium* strain.

c) 1D-NMR metabolomic analysis (n=157 samples) aimed to (i) investigate potential metabolites significantly enhanced in any of the study cohorts, (ii) relate them alongside their bacterial community profiling and (iii) elucidate potential relationships of immune and metabolic function in relation to health outcomes.

Methods

Study design

This study consisted of two distinct patient groups; routine oral supplementation of *Bifidobacterium* and *Lactobacillus* (*i.e*. Bif/Lacto group n=101 infants, 40 ≤ 1000g and 61 > 1000g), and a control non-supplemented group (*i.e*. control group n= 142, 63 ≤ 1000g and 79 > 1000g infants). Infants in the Bif/Lacto group were routinely prescribed an oral supplementation containing 10^9 colony forming units (CFU) of *Bifidobacterium bifidum* and 10^9 CFU of *Lactobacillus acidophilus* (Infloran®, Desma Healthcare). This supplementation was given twice daily as soon as enteral feeds commenced (usually on day 1 postnatal), until 34 weeks old for infants with a birthweight > 1000 g, and until discharge in ELBW infants (< 1000g). Each capsule of Infloran (250 mg) was dissolved in 2 ml of expressed breastmilk and/or sterile water and given to the infant via a nasogastric tube or by mouth.

Inclusion criteria for the infants recruited were gestational age ≤ 34 weeks, and remanence in the same NICU for duration of the study. Infants diagnosed with NEC at the time of consent or with severe congenital abnormalities were excluded from the study. Premature infants enrolled belonged to four different Neonatal Intensive Care Units (NICUs) across England; infants from the Bif/Lacto Group came from Norfolk and Norwich University Hospital (NNUH, Norwich); infants from the control Group came from The Rosie Hospital (Cambridge), Queen Charlotte's and Chelsea Hospital (London) and St Mary's Hospital (London). All these NICUs utilised comparable antibiotic and antifungal policies. Cohorts were matched by gestational age (< 34 weeks gestation), sex, birth mode, time points of sample collection and diet across four different NICUs.

Ethical approval for the study

Faecal collection from infants from NNUH and The Rosie Hospital was approved by the Faculty of Medical and Health Sciences Ethics Committee at the University of East Anglia (UEA) and followed the protocols laid out by the UEA Biorepository (Licence no: 11208). Faecal collection for Queen Charlotte's and Chelsea Hospital and St Mary's Hospital was approved by West London Research Ethics Committee (REC) under the REC approval reference number 10/H0711/39. In all cases, doctors and nurses recruited infants after parents gave written consent.

Time points of sample collection for this study included <10 days, 10-29 days, 30-49 days, >50 days. Clinical data collected on the premature infants included gestational age, delivery mode (C-section vs. vaginal), antibiotic courses received, and dietary information (see Appendix 2 for details of every infant recruited in this study).

DNA extraction of stool samples from premature infants

FastDNA Spin Kit for Soil (MP Biomedical, Santa Ana) was used to extract the bacterial DNA from the faeces samples following the protocol recommended by the manufacturer but extending the bead-beating step to 3 minutes. The DNA recovered was assessed using a Qubit® 2.0 fluorometer (Invitrogen).

This study involved the extraction of 593 faecal samples. This job was carried out by myself, and the technicians working at the Hall lab (Jennifer Ketskemety and Lisa Chalken).

16S rRNA gene sequencing: library preparation and bioinformatics analysis

The 16S rRNA region (V1-V2) was amplified to determine the bacterial community composition on the premature infant stool samples. Primers used for library construction are detailed in table 7 and PCR conditions were: 1 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 RNA gene libraries was performed on the Illumina MiSeq platform with 300 bp paired end reads.

Sequencing reads were filtered through quality control using trim galore (0.4.3) keeping a minimum quality threshold of phred 33 and minimum read length of 60 of bases. Reads that passed the threshold were aligned against SILVA database (version: SILVA_132_SSURef_tax_silva) using BLASTN (ncbi-blast-2.2.25+; Max e-value 10e-3) separately for both pairs. After performing the BLASTN alignment, all output files were imported and annotated using the paired-end protocol of MEGAN on default LCA parameters.

Table 7 Primer sequences for amplifying V1+V2 region of 16S rRNA gene using MiSeq Illumina

9 forward (FW) primers 12 reverse (RV) primers	Total Primer sequences
V1FW_SD501	AATGATACGGCGACCACCGAGATCTACACAAGCAGCATATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD502	AATGATACGGCGACCACCGAGATCTACACACGCGTGATATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD503	AATGATACGGCGACCACCGAGATCTACACCGATCTACTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD504	AATGATACGGCGACCACCGAGATCTACACTGCGTCACTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD505	AATGATACGGCGACCACCGAGATCTACACGTCTAGTGTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD506	AATGATACGGCGACCACCGAGATCTACACCTAGTATGTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD507	AATGATACGGCGACCACCGAGATCTACACGATAGCGTTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD508	AATGATACGGCGACCACCGAGATCTACACTCTACACTTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SA501	AATGATACGGCGACCACCGAGATCTACACATCGTACGTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V2RV_SD701	CAAGCAGAAGACGGCATACGAGATACCTAGTAAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD702	CAAGCAGAAGACGGCATACGAGATACGTACGTAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD703	CAAGCAGAAGACGGCATACGAGATATATCGCGAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD704	CAAGCAGAAGACGGCATACGAGATCACGATAGAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD705	CAAGCAGAAGACGGCATACGAGATCGTATCGCAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD706	CAAGCAGAAGACGGCATACGAGATCTGCGACTAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD707	CAAGCAGAAGACGGCATACGAGATGCTGTAACAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD708	CAAGCAGAAGACGGCATACGAGATGGACGTTAAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD709	CAAGCAGAAGACGGCATACGAGATGGTCGTAGAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD710	CAAGCAGAAGACGGCATACGAGATTAAGTCTCAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD711	CAAGCAGAAGACGGCATACGAGATTACACAGTAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD712	CAAGCAGAAGACGGCATACGAGATTTGACGCAAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT

16S library preparation was performed by myself, Jennifer Ketskemety and Lisa Chalken, as well as management of sequencing files with the sequencing provider. Initial QC and 16S rRNA gene bioinformatics analysis was led by Mr Shabhonam Caim (Hall lab bioinformatician), and I supported and carried out further analysis.

Data analysis

Statistical analyses and diversity calculations were completed using GraphPad (version7). Wilcoxon non-parametric test was used to look for statistical significance among external factors influencing *Bifidobacterium*. Non-parametric two tailed student t-test was used to estimate differences in pH values. I performed this analysis, except for non-metric multidimensional scaling (NMS) plots (Figure 21), which was performed by Dr Matthew Dalby.

Genomic DNA Extraction from bacterial isolates

Isolation of strains present in the supplementation (*Bifidobacterium bifidum* and *Lactobacillus acidophilus)* as well as other potential pathogenic strains from the infants' samples such as *Staphylococcus, Escherichia,* and *Klebsiella* was performed. *Bifidobacterium* strains were isolated using MRS (Difco) agar with 50 mg/L of mupirocin, *Baird-Parker* agar (Oxoid) for *Staphylococcus*, and MacConkey agar for *Escherichia* and *Klebsiella*.

DNA extraction of the isolates was prepared using an overnight pure culture in Brain Heart Infusion Broth (BHI). Bacterial pellets were resuspended in 2 ml 25% sucrose in 10 mM Tris and 1 mM EDTA at pH 8. Cells were subsequently lysed adding 50 µl 100 mg/ml lysozyme (Roche) and incubating the mixture at 37 °C for 1 h. Following, 100 µl 20 mg/ml Proteinase K (Roche), 30 µl 10 mg/ml RNase A (Roche), 400 µl 0.5 M EDTA (pH 8.0) and 250 µl 10% Sarkosyl NL30 (Fisher) was added into the lysed bacterial suspension, incubated 1 h on ice and left overnight at 50 °C. Next day protocol comprised washes of phenol-chloroform-isoamyl alcohol (PCIA, Sigma) using 15 ml gel-lock tubes (Qiagen). E Buffer (10mM Tris pH 8 (Fisher Scientific, UK) was added to the sample to a final volume of 5 ml, mixed with 5 ml of PCIA (Sigma) and centrifuged for 15 min at 1792g. The CIA step was repeated three times, after which the final aqueous phase was transferred into a sterile Corning ™ 50 ml centrifuge tube, and 2.5 volumes of ethanol (VWR

Chemicals, USA) were added, incubated for 15 min at -20 °C, and centrifuged 10 min at 1792g and 4 °C. Finally, the pellet was washed twice with 10 ml of 70% ethanol and centrifuged at 1792g for 10 min, dried overnight, and re-suspended in 300 µl of E Buffer. I performed the isolations of the bacterial strains and the DNA extraction.

Whole genome sequencing analysis: library preparation and bioinformatics analysis

DNA of pure cultures was subjected to multiplexed sequencing using standard Illumina library protocols followed by sequencing via HiSeq 2500 platform with 125 bp paired end reads. Genome assemblies using pipeline described by Page et al [143]. All the assembled contigs were passed through prokka (1.12).

Library preparation of the bacterial isolates was done at the Wellcome Sanger Institute (Cambridge) and bioinformatic analysis was carried out by Mr Shabhonam Caim.

Determination of Minimal Inhibitory Concentration (MIC) of Infloran strains.

Broth microdilution method[144] was used to test the Minimal Inhibitory Concentration of the probiotic strains (*Bifidobacterium bifidum* and *Lactobacillus acidophilus*) against the routinely used antibiotics at NICUs (benzylpenicillin, gentamicin, vancomycin, metronidazole, meropenem and cefotaxime). Serial two fold dilutions of the antibiotics in MRS medium (Difco) were prepared and added to 10 µl from a fresh overnight culture of *Bifidobacterium bifidum* and *Lactobacillus acidophilus*. Microplates were incubated for 24 h at 37 °C under anaerobic conditions. Cell density was monitored using a plate reader (BMG Labtech, UK) at 595 nm. MICs were determined as the lowest concentration of antibiotic inhibiting any bacterial growth. Test were done in triplicate. I performed the MIC tests.

Metabolomic profiling using ^{1}H- nuclear magnetic resonance spectroscopy (NMR)

A subset of 157 paired faecal samples (75 from Bif/Lacto group, and 81 from control group) were analysed by standard one-dimensional (1D) ^{1}H-NMR spectroscopy. The selection of these samples was performed randomly. 1D-NMR samples were normalised to 50 mg faecal content, added 700 µl of phosphate buffer (pH 7.4; 100% D_2O) containing 1 mmol/L of 3-trimethylsilyl-1-[2,2,3,3-$^{2}H_4$] propionate (TSP), plus 1 mm diameter of zirconium beads, (BioSpec Products). Samples were bead beated using Precellys bead beater (Bertin) for 2 cycles of 40s at 6,500Hz speed[145], centrifuged at 14,000 x g for 10 min and supernatant was transferred to NMR tubes. 1D-NMR spectra were acquired for each sample using a nuclear overhauser effect pulse sequence for water suppression as described by Beckonert and colleagues[146]. Spectra was imported into Matlab 9.4 (R2018a), manually corrected by removing signals corresponding to TSP and water and normalised using probabilistic quotient method. Data analysis was performed using orthogonal projection to latent structures discriminant analysis (OPLS-DA) and custom in-house scripts. Metabolites were assigned manually using Chenomx 7.0.

I selected the samples for metabolomics analysis, prepared for metabolomics analysis and analysed them under the supervision Dr Jonathan Swann and Dr Fahmina Fardus at Imperial College (London).

pH measurement of the faecal samples

Faecal pH was measured on a subset of samples used in the metabolomics analysis (39 samples from the Bif/Lacto Group and 39 samples from the control Group). Samples selected for this pH analysis were used previously in the metabolomic analysis. Samples were weighed to 50 mg of faecal material, added 1ml of sterile water and used a glass electrode pH meter to measure the pH (Martini Mi151). I performed the pH measurements of the samples.

Kinetic growth curves of *Staphylococcus, Klebsiella* and *Escherichia* gut bacterial isolates using supernatant of *Bifidobacterium bifidum* from Infloran.

The test was done using a 96 well plate including: 180 µl of *B. bifidum* supernatant, 2 µl of an overnight culture of *Staphylococcus*, *Klebsiella* or *E. coli* and 10 µl of a 10x BHI broth (pH adjusted to 6.3) to adjust for nutrient consumption. The plates were incubated in aerobic conditions at 37 °C and analysed every 15 min over a period of 24 h. The optical density (OD) at 595 nm was determined for each well using a plate reader (Tecan, Infinite 50) and samples were tested in triplicate. Controls were grown alongside each sample: (i) control for microbiological media BH (Brain Heart) broth and (ii) control of pH consisting of BH (Brain Heart) broth with pH adjusted to the supernatant of *B. bifidum* (6.3). I performed the kinetic growth curves of this study.

Results

Study design

In this study, a total of 591 stool samples from 224 premature infants from four Neonatal Intensive Care Units in UK (Norwich, Cambridge and London) were analysed. Two groups were included in this study; Bif/Lacto group (101 infants, routine oral Infloran supplementation), and the control group (133 infants, non-supplemented infants). All infants recruited (n=234) were ≤ 34 weeks of gestational age, with 103 of them classified as ELBW <1000 g. Probiotic supplementation was given twice daily until 34 weeks old for infants with a birthweight >1000 g, and until discharge in ELBW infants (<1000 g). On average ELBW infants received a minimum of 20 days of probiotic supplementation and infants weighing >1000 g received an average of 14 days supplementation. The study excluded premature infants diagnosed with NEC at the time of sample collection or with severe congenital abnormalities. The study groups also comprised predominantly of premature infants who received breast milk or donor breastmilk (78% in Bif/Lacto Group and 77% control Group), mixed (breastmilk,

formula or donor breast milk, 19% Bif/Lacto Group and 20% control Group), and exclusively formula fed (4% Bif/Lacto Group, and 2% control Group). Stool samples were collected at four time points: t=1(<10 days), t=2 (11-29 days), t=3 (20-49 days) and t=4 (>50 days). Figure 20 shows a summary of the study design highlighting number of infants recruited in each selected criterion.

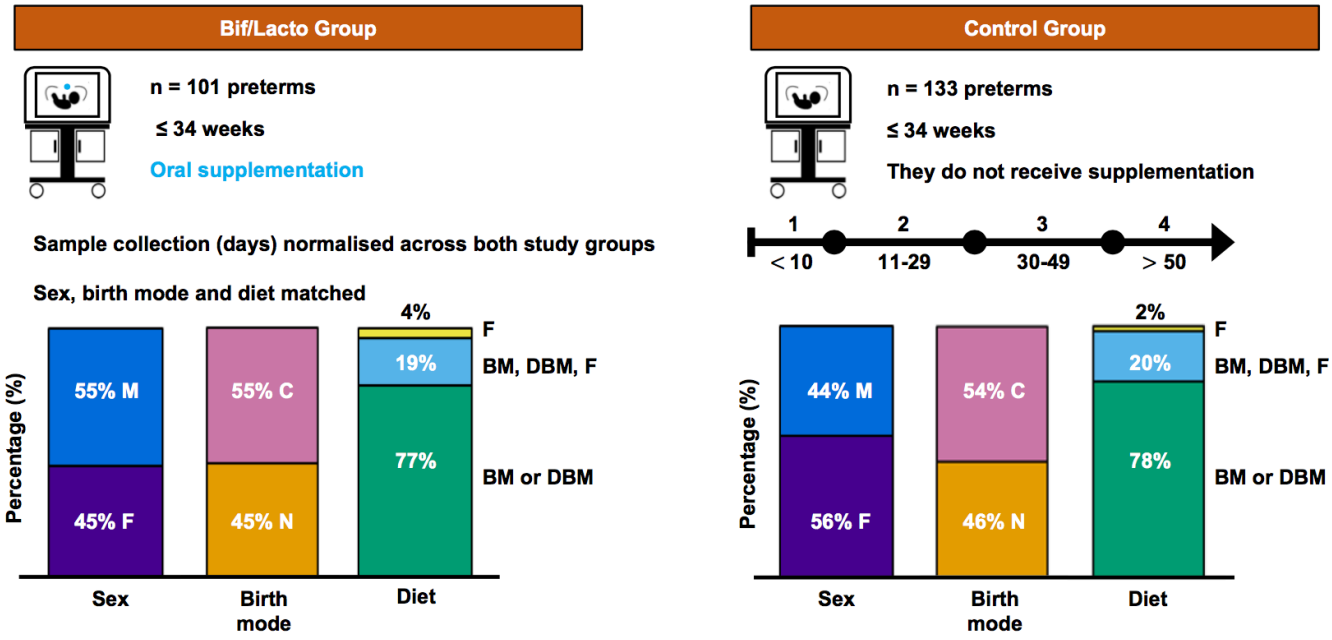


Figure 20 Study design

Figure highlighting number of premature infants recruited in each study cohort, times of stool collection and percentages of infants recruited detailing sex (M= male or F= female), birth mode (C= C-section or N= Natural birth) and diet (BM= breastmilk, DBM= donor breast milk, F= formula).

Oral Bif/Lacto supplementation influences bacterial genus abundance and bacterial diversity

To determine if supplementation with *Bifidobacterium* and *Lactobacillus* could alter the preterm gut microbiome a 16S rRNA gene bacterial profiling was performed on 101 infants who received routine oral supplementation, and 133 infants non-supplemented as control group. Sequencing files with < 25,000 reads were discarded, due to my previous work indicating that the optimal threshold in reads counts for an accurate representation of bacterial populations in the preterm gut microbiome[147].

Initial, clustering of premature infants samples using non-metric multidimensional scaling (NMDS) plots indicated a clear separation in gut microbiome profiles when comparing Bif/Lacto group and the control group. NMDS plots can be used when data does not have a normal distribution; the closer the points in the 2D representation the more similar their microbial communities. *Bifidobacterium* was the most prevalent genus in the Bif/Lacto group, while *Staphylococcus*, *Escherichia*, and *Klebsiella* were the most abundant in the control group (Figure 21). Contrary to expectations, only a minority of premature infants in the Bif/Lacto group had detectable relative abundances of *Lactobacillus*, suggesting this taxon did not efficiently colonise the premature gut.

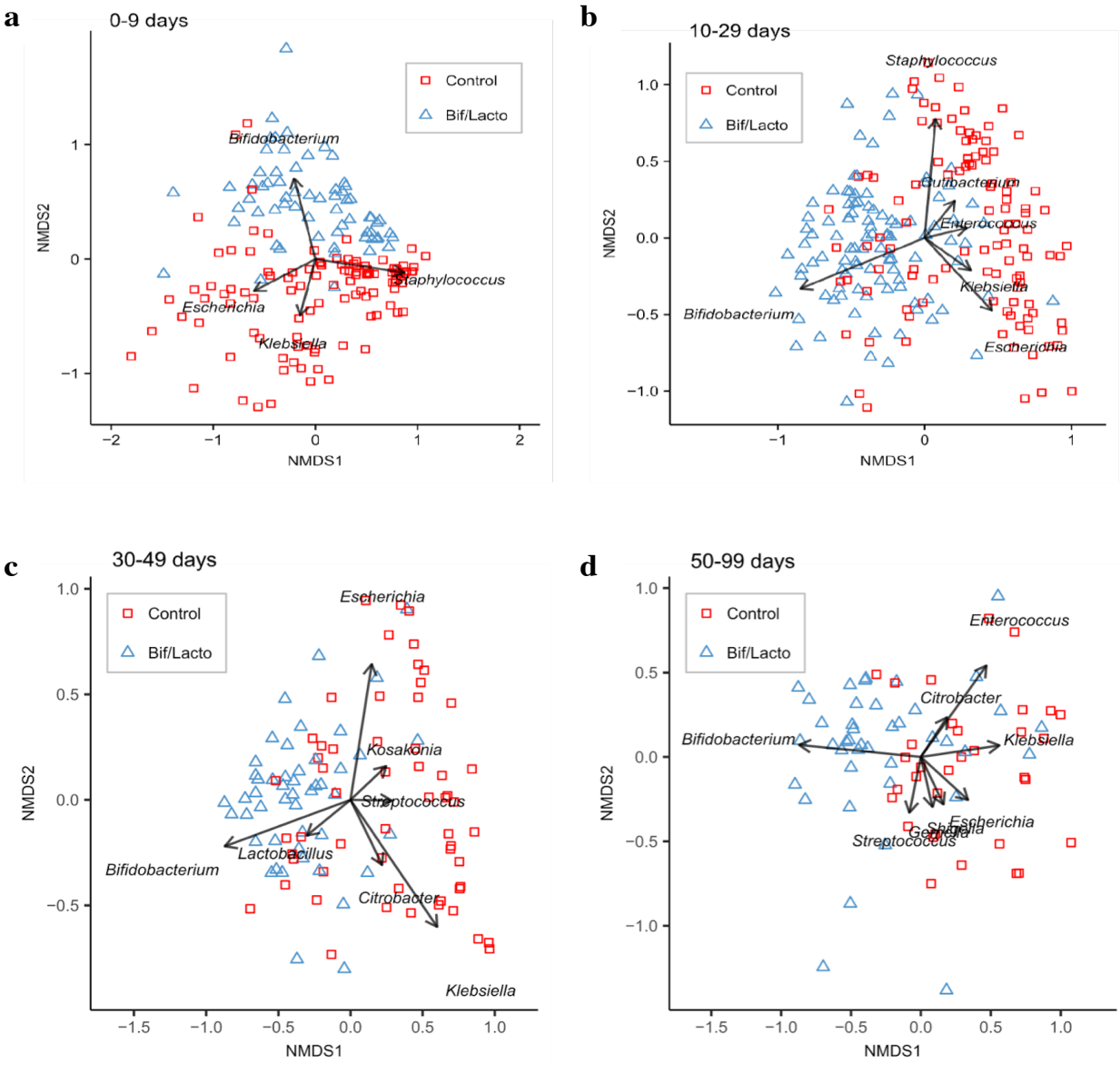


Figure 21 NMDS plots from Bif/Lacto group and control group differentiating time points of sample collection

NMDS plots showing premature faecal samples clustered using a Bray-Curtis similarity matrix at (a) <10 days, (b) 11-29 days (c) 30-49 days and (d) at >50 days of age. Arrows indicate bacterial genus driving the separation points on the NMDS plots.

Bar charts displaying the top twelve bacterial genera for all samples across the study period indicated that the premature faecal microbiome was typically dominated by a maximum of four bacterial genus; *Bifidobacterium*, *Staphylococcus*, *Escherichia* and *Klebsiella* (Figure 22). *Bifidobacterium* comprised > 30% of the overall relative abundances in the total microbiome for the Bif/Lacto group at all time points (Figure 22a), with control group infants having < 20% (Figure 22b). *Staphylococcus* was initially relatively abundant in both groups,

but decreased across later time points, which may correlate with initial inoculation of skin associated bacteria. Surprisingly, *Lactobacillus* was only detected in a minority of infants, but with a higher relative abundance in Bif/Lacto infants compared to the Control group at all time points. Importantly, the Bif/Lacto group presented lower abundance of potential pathobiont bacteria (*Klebsiella*, *Escherichia and Enterobacter)* when compared to the control group. Overall these data, indicate that oral supplementation from birth may modulate the preterm microbiome, including 'displacing' other potentially pathogenic and commonly-associated premature resident taxa.

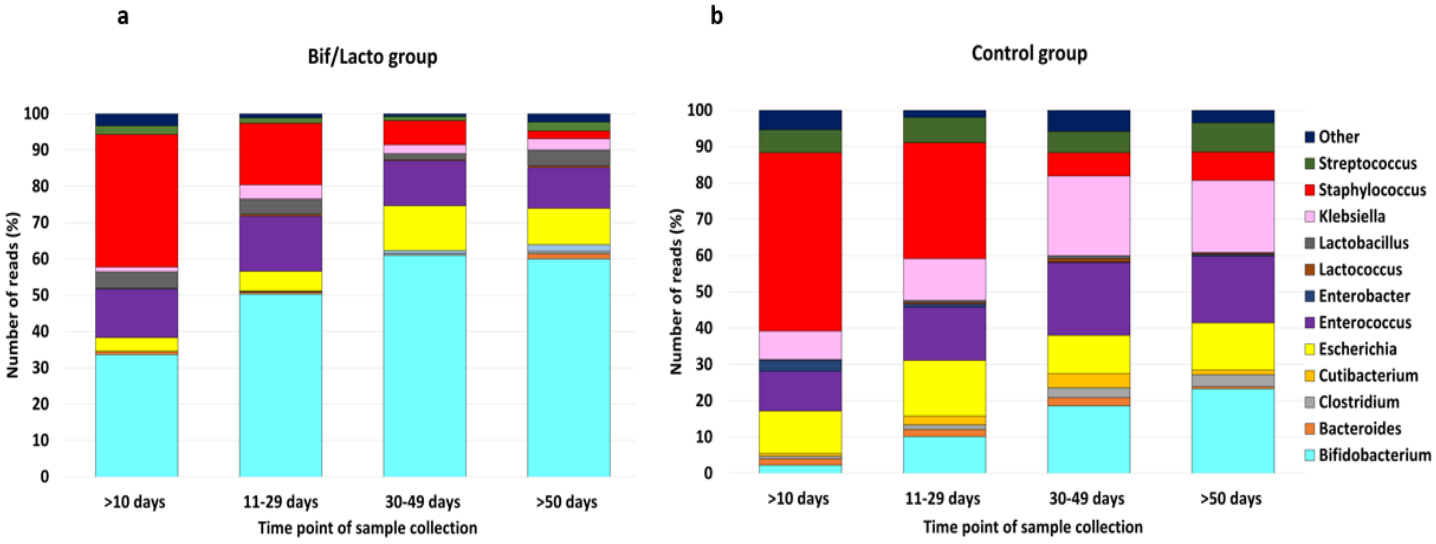
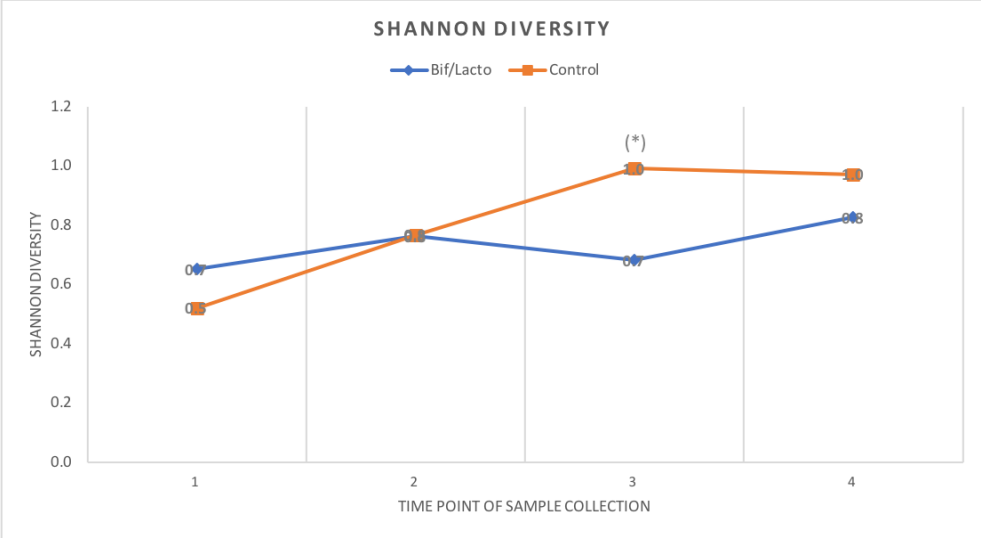


Figure 22 Genus abundance between Bif/Lacto group and control group

Mean proportional abundance of the most common bacterial genus at each time point (<10 days, 11-29 days, 30-49 days and >50 days of age) for Bif/Lacto group (a), and control group (b).

When plotting average Shannon diversity indexes (Figure 23a), the Bif/Lacto group presented a lower index than the control group, which may be due to the dominance of *Bifidobacterium* in these samples. Interestingly, the number of bacterial genus detected in the Bif/Lacto group was smaller during the first three time points of the study (up to 50 days) when compared to the control group, however at time point 4 (>50 days) both cohorts converged (Figure 23b).

a)



b)

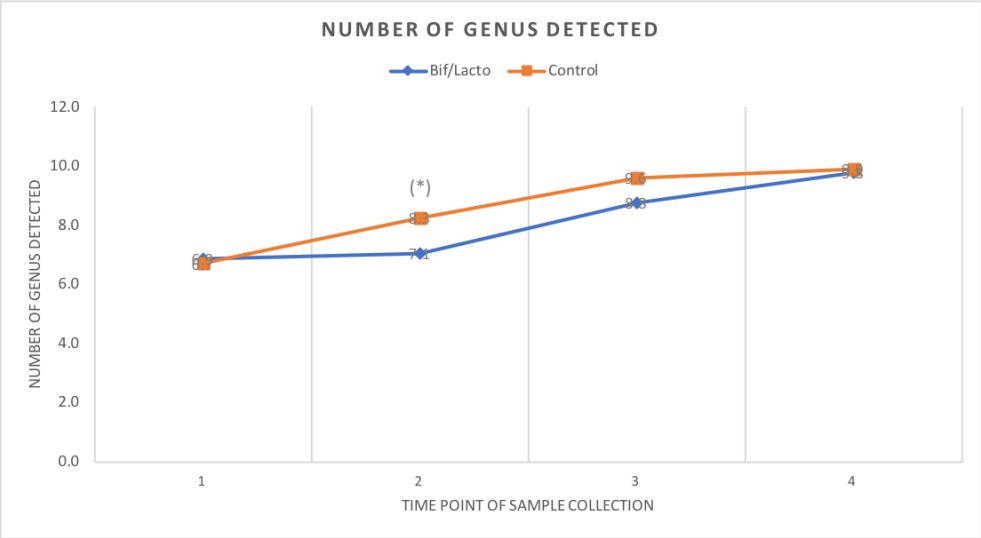


Figure 23 Shannon diversity index and bacterial genus detected among the study cohorts (Bif/Lacto and control group)

a, Shannon diversity index increased more rapidly in the control group than in the Bif/Lacto. b, Number of bacterial genus detected from the start of the study to the end. Asterisks represent $p < 0.05$.

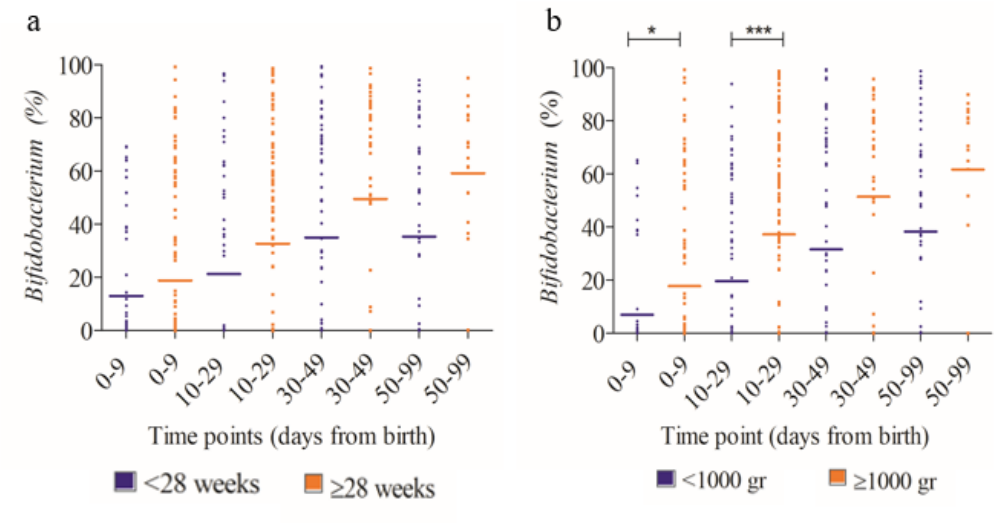
External factors including birth weight and antibiotics negatively influenced *Bifidobacterium* abundance in recruited infants

As shown above *Bifidobacterium* was the dominant taxa that differentiated between the two premature infant groups (Bif/Lacto and control). The influence that other external factors, such as gestational age, birth weight, antibiotics, delivery mode or diet, was also analysed.

A recent study by Korpella and colleagues highlighted that gestational age was the main determinant impacting the preterm microbiome[148], with gut maturity correlating with *Bifidobacterium* abundance. To evaluate the influence of gestational age in this study, all infants studied were grouped into very low gestational age (<28 weeks) or low gestational age (≥28 weeks). Results indicate a tendency for lower relative abundance of *Bifidobacterium* in infants with very low gestational age (<28 weeks), however the differences shown among the two gestational age groups where not significant (Figure 24a). Using birth weight as a defining factor revealed that premature infants with extremely low birth weight (<1000 g) presented significantly lower relative abundance of *Bifidobacterium* (Figure 24b) up until day 29. As shown in Figure 24c there is a positive correlation with gestational age and birthweight.

Gestational age

Birth weight



Gestational age/Birth weight

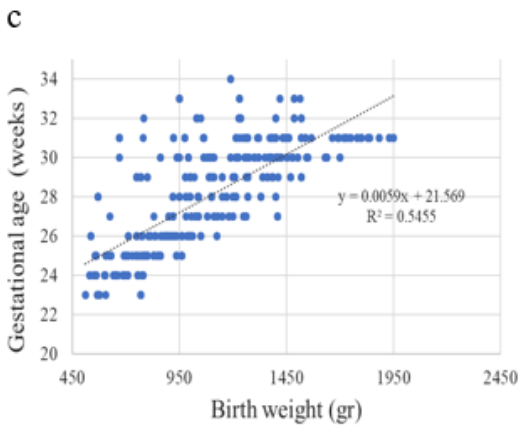


Figure 24 Effect of gestational age and birth weight on *Bifidobacterium* relative abundance

a, *Bifidobacterium* abundance in premature infants with very low gestational age (<28 weeks) and low gestational age (≥28 weeks). **b**, *Bifidobacterium* abundance between very low birth weight (<1000 g) and low birth weight (>1000 g) infants. **c**, Gestational age correlated with birth weight (gr). Asterisks represent p values: *P < 0.05, ***P < 0.001.

Antibiotics have been shown to have a strong, but temporary effect, on the preterm microbiome[49]. Correlation of microbiome profiles from all infants recruited with antibiotic usage, revealed significant differences in *Bifidobacterium* abundance when comparing premature infants receiving antibiotics at the time of sample collection, with those who did not receive any treatment (Figure 25).

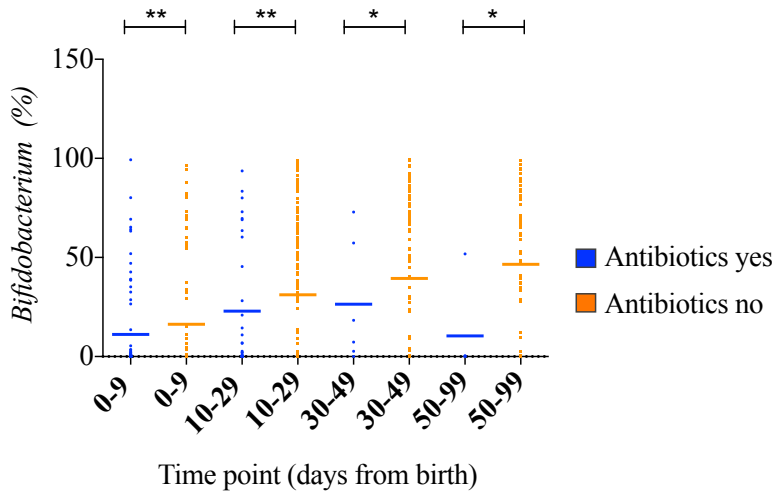


Figure 25 Effect of antibiotics on *Bifidobacterium* relative abundance

Bifidobacterium abundance in infants receiving antibiotics at the time of sample collection (Antibiotic yes) and infants who did not receive antibiotic treatment (Antibiotic no). Asterisks represent *p* values: *P < 0.05, **P < 0.01.

Delivery mode (vaginal birth or C-section) is another external factor which determines whether the gut is primarily colonised by maternal vaginal and faecal microbiome, or by skin microbes[149]. This study evaluated the effect that delivery mode had on *Bifidobacterium* relative abundance, and surprisingly no significant differences where observed when comparing natural (n=133) versus C-section (n=101) birth (Figure 26), both in the Bif/Lacto group and control group. These data indicate that premature infants born by vaginal birth (45% of the infants recruited in this study) may not get the colonisation effects observed in their full-term counterparts. In addition, ~80% of premature infants receive antibiotics during the first week of life[150], which may significantly alter the initial microbial inoculation process when an infant pass through the birth canal.

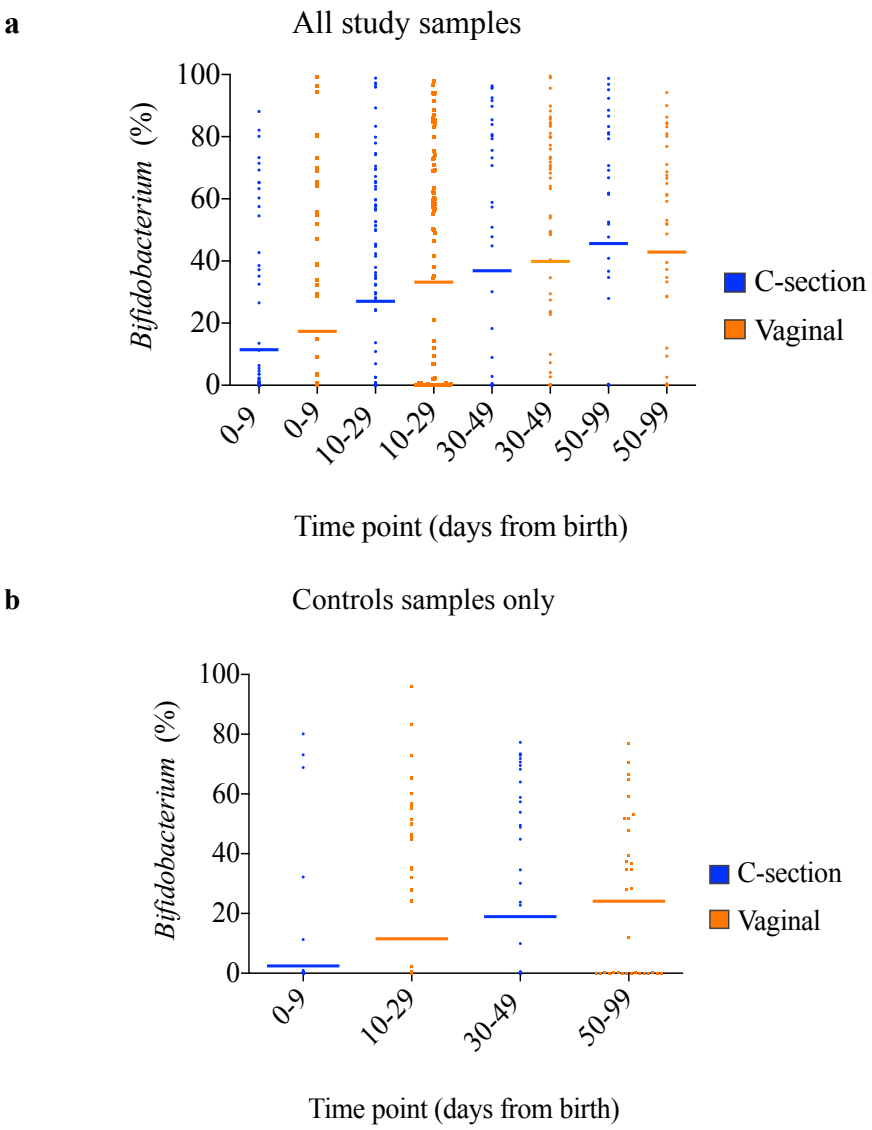


Figure 26 Effect of delivery mode on *Bifidobacterium* abundance

a, *Bifidobacterium* abundance in all study samples differentiating infants born by C-section (blue) and vaginal birth (orange) **b**, *Bifidobacterium* abundance in control infants highlighting infants born by C-section (blue) and vaginal birth (orange). Data was grouped according to time points of sample collection.

It is widely accepted that *Bifidobacterium* represents an important commensal in the infant gut, and that healthy breastfed infants are predominantly colonised by this taxa, which is linked to an ability to digest human milk oligosccharides[148]. Contrary to expectation, formula fed infants presented higher relative abundance of *Bifidobacterium* (Figure 27). However, only a very small proportion of infants were exclusively formula fed in this study (18 infants out of 234) making robust statistical analysis difficult. It is interesting to note that the formula given to these exclusively formula fed infants contained prebiotics (*i.e.* FOS), which has been shown to enhance the growth of beneficial bacteria such as *Bifidobacterium*.

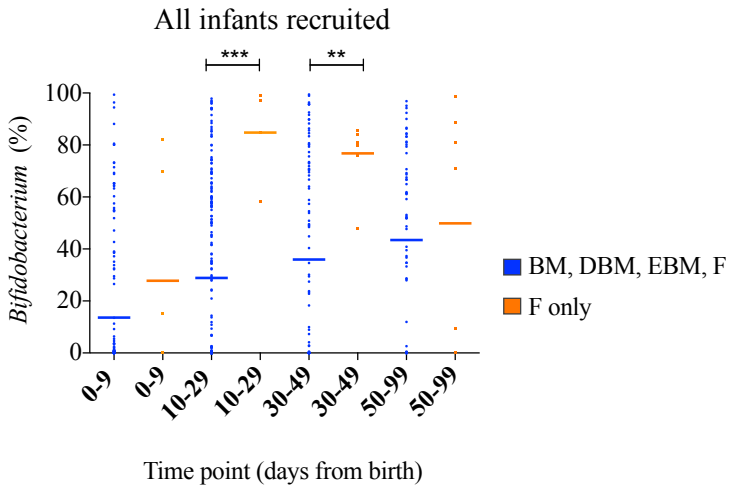


Figure 27 Effect of diet on *Bifidobacterium* abundance

Bifidobacterium abundance in infants receiving breastmilk (blue) or formula only (orange). BM = breastmilk, DBM = Donor breastmilk, EBM = Expressed breastmilk and F = formula. Asterisks represent *p* values: *P < 0.05, **P < 0.01 ***P < 0.001.

Whole genome sequence analysis demonstrates gut colonisation of supplemented *Bifidobacterium strain.*

Previous probiotic studies have indicated only transient or short-term colonisation of supplemented strains[151, 152]. Thus, next it was determined if the *B. bifidum* strain present in the supplementation (*i.e.* Infloran), was able to colonise the gut of the supplemented Bif/Lacto group. WGS analysis was performed, comparing Infloran *B. bifidum* (Infloran) to ten other *Bifidobacterium* isolates extracted from Bif/Lacto infants (five of them received supplementation at the time of sample collection and two had stopped supplementation). Phylogenetic core genome analysis showed more than 99.9% similarity among all the *B. bifidum* isolates (Figure 28), with pangenome analysis also indicating high sequence similarities (Figure 29). Notably, the study indicated four premature infants (P19, P8, P35, and P15) were found to harbour the same Infloran *B. bifidum* isolate. Interestingly, samples belonging to premature infant P8 were collected at 41 and 50 days after supplementation had finished, indicating longer-term colonisation of this strain.

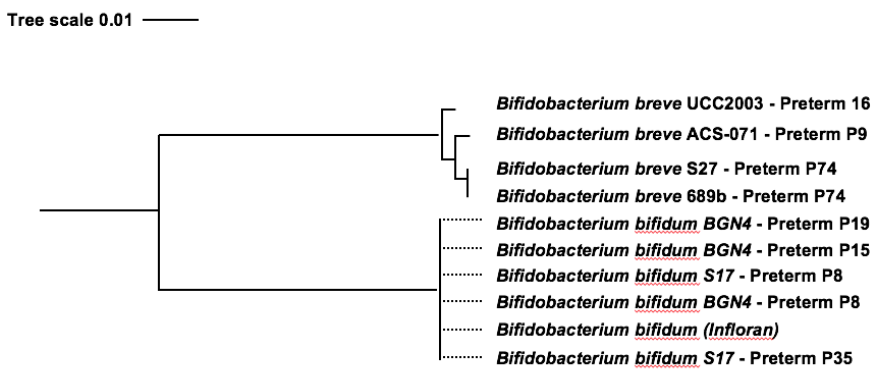


Figure 28 Core genome tree from *Bifidobacterium bifidum* present in the oral supplementation and other Bifidobacterium species isolated from premature infants

Comparison of core genomes from *Bifidobacterium bifidum* present in the oral supplementation, and ten other Bifidobacterium isolates from the study premature infants. Roary core gene alignment output was used to create a maximum likelihood (ML) phylogenetic tree.

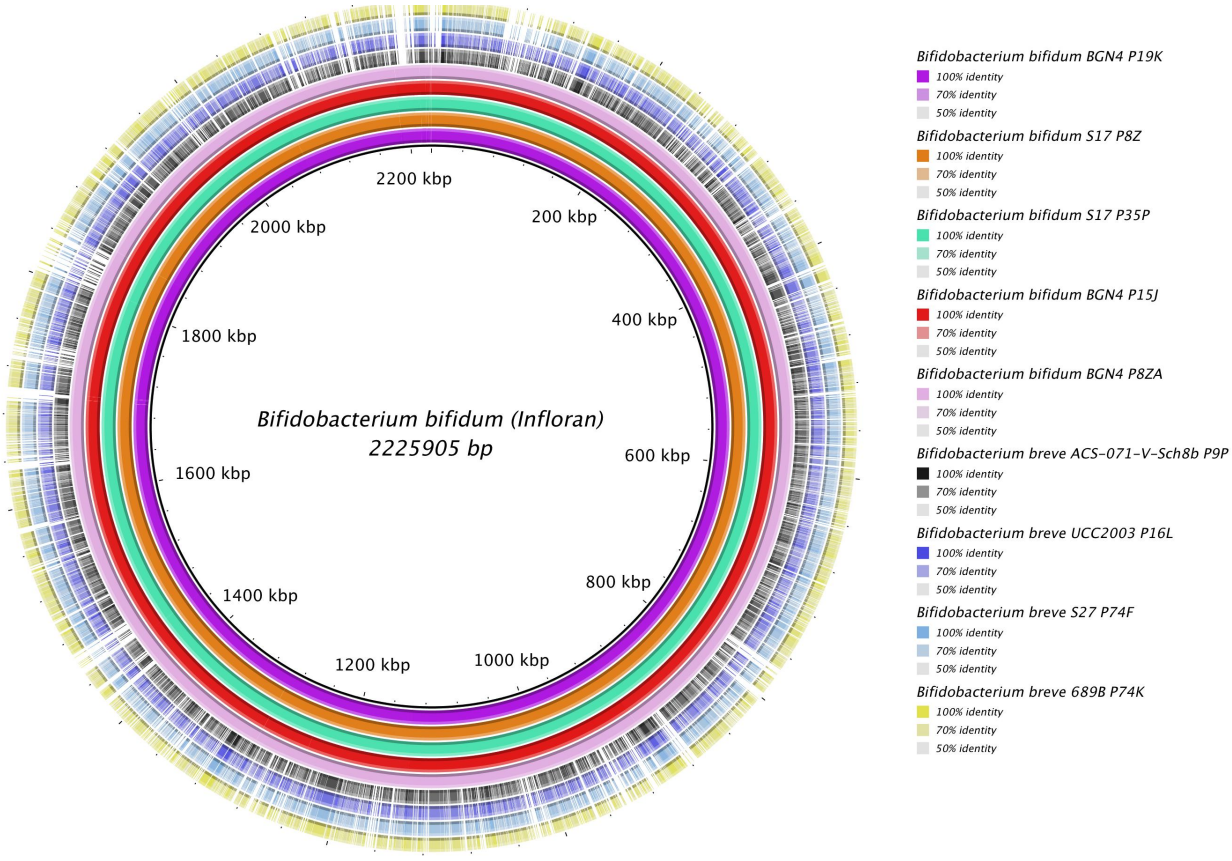


Figure 29 Genome diagrams from *Bifidobacterium bifidum* (Infloran) and other *Bifidobacterium*

BRIG software was used to compare whole genomes *of Bifidobacterium bifidum* present in the supplementation and other *Bifidobacterium isolates*. Similarity was calculated using BLAST.

Minimal Inhibitory Concentration (MIC) testing of *Bifidobacterium bifidum* (Infloran) demonstrated susceptibility to the most common antibiotics used in NICUs.

16S rRNA gene sequencing data indicated that *Bifidobacterium* abundance was significantly affected by antibiotic treatment (Figure 25). To probe this finding, a phenotypic antibiotic test was performed using Minimal Inhibitory Concentration (MIC) testing. Results showed *B. bifidum* (Infloran) presented a lower (MIC) breakpoint value than those put forward by the European Committee on Antimicrobial Susceptibility Testing (EUCAST[153]) to all antibiotics tested (*i.e.* benzylpenicillin, gentamicin and meropenem), which suggests this strain is susceptible to these commonly prescribed NICU antibiotics (Table 8).

Table 8 Minimal Inhibitory (MIC) concentrations for *Bifidobacterium bifidum* (Infloran)

Antibiotic	*Bifidobacterium bifidum (Infloran)*	
	MIC (mg/L)	**Eucast value (mg/L)**
Benzylpenicillin	0.11	0.12 (ampicillin)
Gentamicin	39	64
Meropenem	0.095	ND

Probiotic supplementation drives differences in faecal metabolomic profiles

Metataxonomic profiling of both study groups (Bif/Lacto group indicated a clear separation in gut microbiome profiles. Thus, next it was determined whether these differences may also link to metabolome profiles of both cohorts using a subset of randomly selected faecal samples (n= 157) and untargeted NMR metabolomics analysis.

Principal component analysis (PCA) of all the metabolites found in samples, showed clear group clustering, differentiating sample cohorts (Bif/Lacto group and control group, see Figure 30).

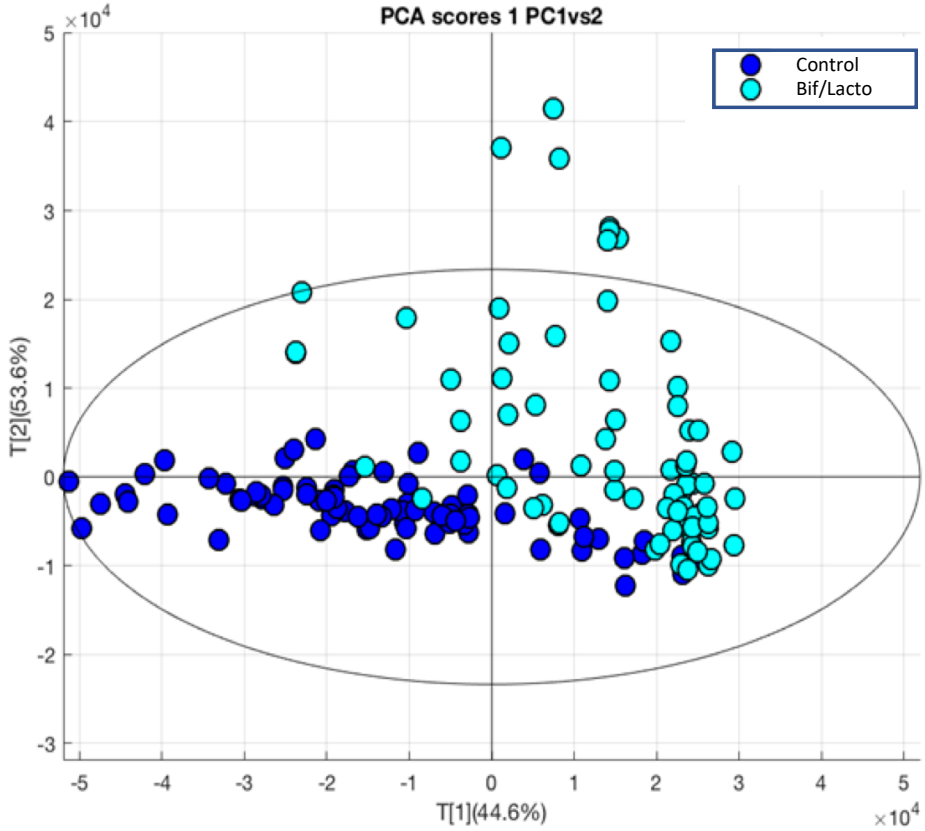


Figure 30 Principal component analysis (PCA) plot of metabolite profile from Bif/Lacto group and control group

PCA plot representing faecal metabolite profiles from Bif/Lacto group (light blue) and control group (dark blue).

To further investigate whether these differences were maintained throughout the study, orthogonal signal corrected partial least squares discriminant analysis[18] (OPLS-DA) on the spectrum data was performed. The predicted performance of the OPLS-DA models (Q^2Y) and the *p*-values obtained (all less than 0.01, Table 9), suggested metabolite differences observed continued throughout the study period. *Bifidobacterium* abundance was the main driver in the model that influenced these differences with a $Q^2Y = 0.44$ and *p* value <0.01. It is interesting to note that the OLPS-DA model showed significant differences for *Staphylococcus* abundance in the Bif/Lacto group only, and for time point of sample collection in the control group.

Table 9 Q^2Yvalues obtained from OPLS-DA Models

Y	*All samples (Q^2Y)*	*Bif/Lacto Group (Q^2Y)*	*Control Group (Q^2Y)*	*p value*
***Bifidobacterium* abundance**	**0.44**	**0.29**	**0.20**	0.01
Time points of sample collection	-0.01	-0.27	**0.24**	0.01
Birth weight	0.08	-0.17	-0.01	NA
Gestational age	0.06	-0.13	-0.125	NA
***Staphylococcus* abundance**	**0.10**	**0.20**	0.06	0.01
***Klebsiella* abundance**	0.002	-0.11	-0.13	NA
***Escherichia* abundance**	-0.07	-0.002	-0.05	NA
***Enterococcus* abundance**	-0.007	-0.15	-0.08	NA
***Streptococcus* abundance**	-0.11	-0.13	-0.11	NA

Left column indicates the different matrices used (Y). Models were evaluated using all samples or differentiating among cohorts (Bif/Lacto and control). NA = not applicable

Next, a loading coefficient plot was generated to display the covariance between the Y-response matrix (Bif/Lacto or control Group) and the signal intensity of the metabolites in the NMR data; this allows detection of the metabolites responsible for differentiation between the study groups. Colours projected onto the coefficient plot indicated the correlation coefficient (R^2) between each metabolite and the Y-response variable (Figure 31), with red indicating strong significance and blue indicating weak significance.

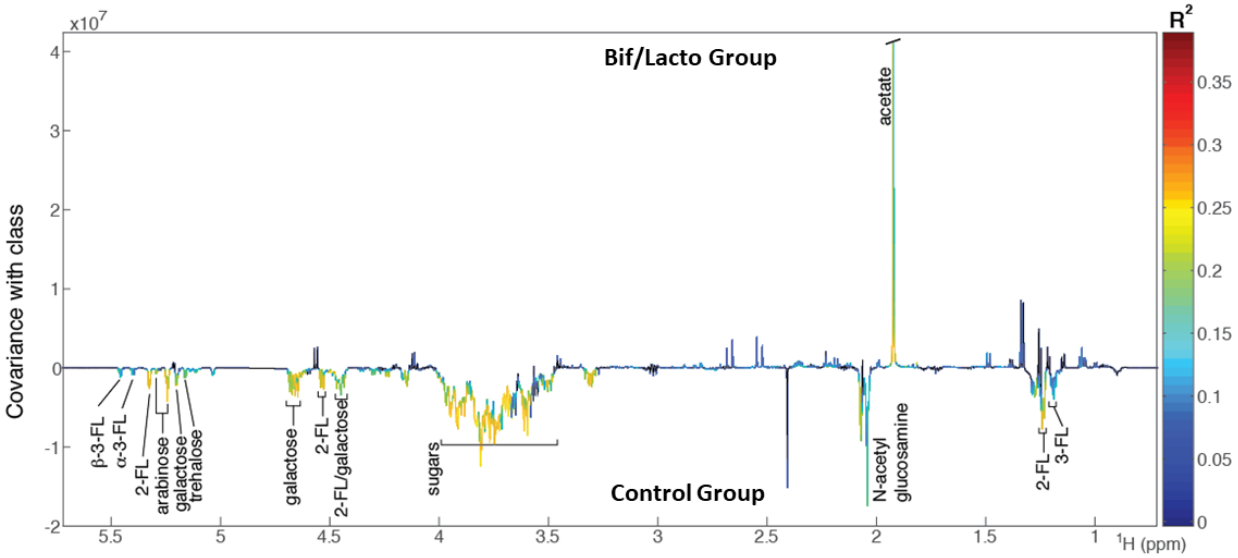


Figure 31 OPLS-DA loading plot on *Bifidobacterium* abundance

OPLS-DA loading plot differentiating sample group. Peaks pointing upwards correspond to metabolites enhanced within the Bif/Lacto group, and peaks pointing downwards were elevated in the control group.

Statistical analysis highlighted six metabolites which were significantly different throughout the study period: (i) acetate and lactate were enhanced in faecal samples from Bif/Lacto Group (Figures 32a-d), (ii) while sugars 3-Fucosyllactose (Figures 32e-f), 2- Fucosyllactose (Figures 32e-f), arabinose (Figures 32g-h) and trehalose (Figures 32g-h) were found elevated in faeces from the control group.

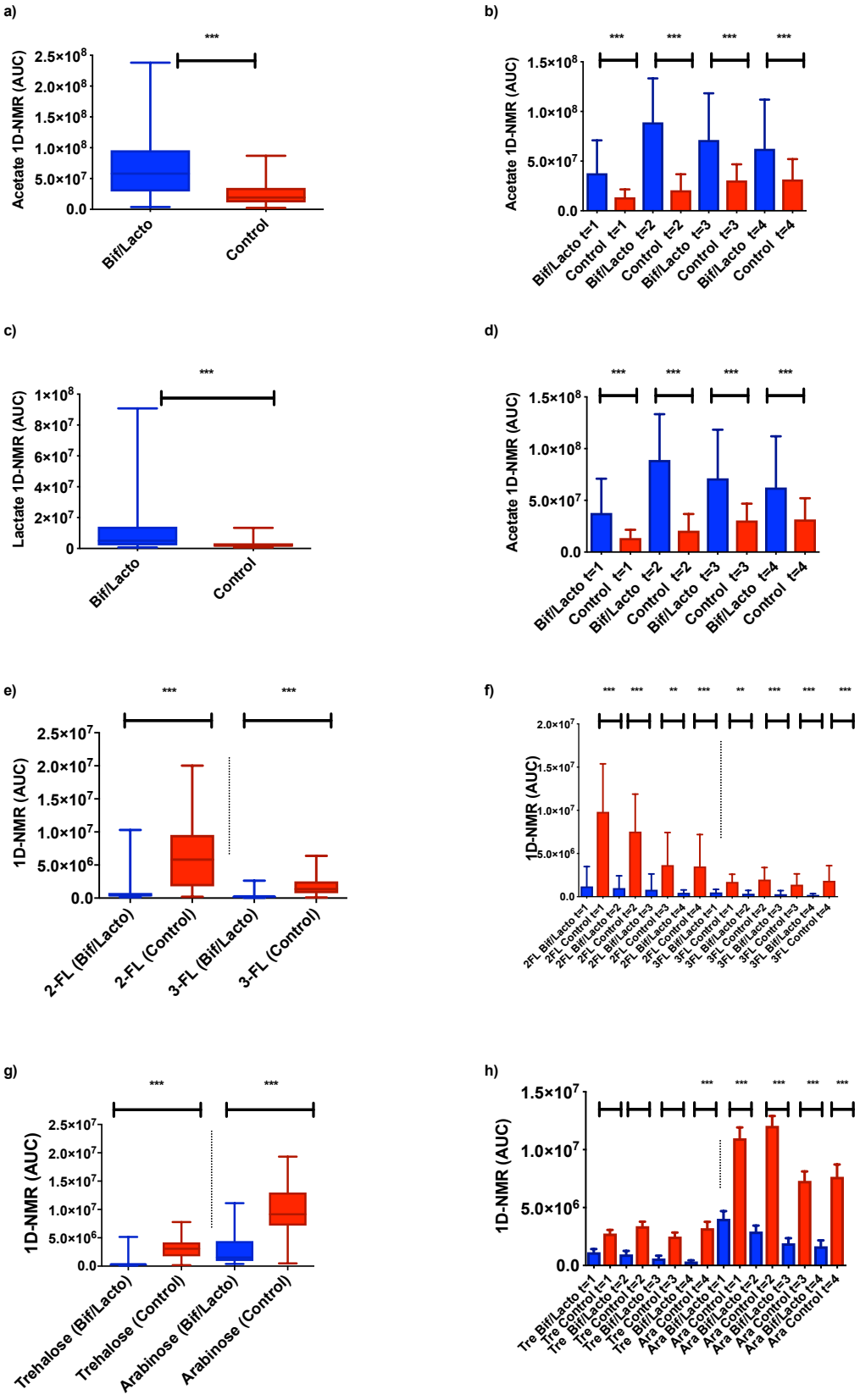


Figure 32 Summary of main 1D-NMR metabolites found in faeces from Bif/Lacto and control group

a, Relative Acetate AUC levels **b**, Relative acetate AUC levels differentiating time point of sample collection **c,** Relative lactate AUC levels **d,** Relative lactate AUC levels differentiating time point of sample collection **e,** Relative trehalose and arabinose AUC levels **f,** Relative trehalose and arabinose AUC levels differentiating time point of sample collection **g,** Relative 2-fucosyllactose and 3-fucosyllactose AUC levels **h**, Relative 2-fucosyllactose and 3-fucosyllactose AUC levels differentiating time point of sample collection. Time point of sample collection t=1 (>10days), t=2 (11-29 days), t=3 (30-49 days) and t=4 (>50days). Asterisks represent *p* values: $^{**}p < 0.01$, $^{***}p < 0.001$.

Acetate and lactate are microbial-derived short chain fatty acids (SCFAs), and fermentation products from *Bifidobacterium,* as well as other beneficial members of the gut microbiome such as *Lactobacillus*[154]. These SCFAs, and acetate in particular, have been shown to promote defence functions in host epithelial cells[155]. Interestingly, *Bifidobacterium* abundance appears to be correlated to acetate levels; samples with high level of *Bifidobacterium* abundance presented higher levels of acetate (Figure 33).

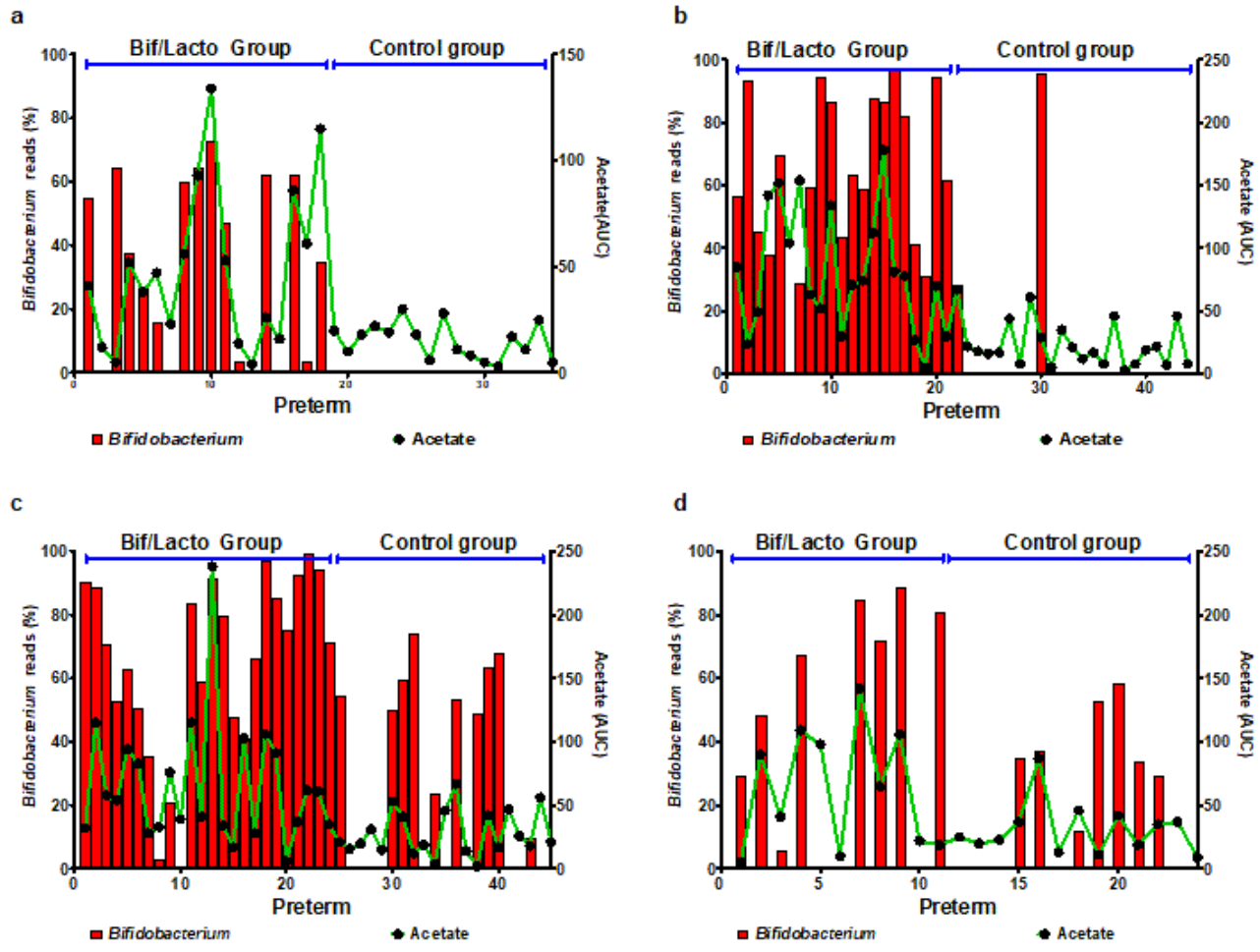


Figure 33 *Bifidobacterium* reads, and acetate levels found in the Bif/Lacto Group and control group

a, *Bifidobacterium* reads and acetate levels found in the Bif/Lacto Group and control group at time point of collection 1= >10 days **b**, *Bifidobacterium* reads and acetate levels found in the Bif/Lacto Group and control Group at time point of collection 2= 11-29 days **c**, *Bifidobacterium* reads and acetate levels found in the Bif/Lacto Group and control Group at time point of collection 3= 30-49 days **d**, *Bifidobacterium* reads and acetate levels found in the Bif/Lacto Group and control Group at time point of collection 4= >50 days. *Bifidobacterium* reads were displayed in left Y axis, and acetate levels were displayed in right Y axis.

2- Fucosyllactose and 3- Fucosyllactose are oligosaccharides naturally occurring in human breast milk, commonly known as HMOs. It is interesting to note that infants belonging to the control group had higher levels in faeces of these HMOs, than the Bif/Lacto Group (Figure 32e-f). Recent scientific studies have shown that *B. bifidum* plays an important role in degrading HMOs[156], and the products produced from this metabolism, such as SCFAs, mediate cross-feeding amongst other *Bifidobacterium* species and strains and also other microbiome members. In this study infants belonging to the Bif/Lacto group presented high abundance of *Bifidobacterium bifidum* (supplemented strain), which could have contributed to the degradation of these HMOs and facilitated colonisation of other *Bifidobacterium* taxa.

pH measurements from faeces from Bif/Lacto group were more acidic than control group, and were correlated with higher levels of acetate and lactate

A recent study in breast-fed infants associated elevated *Bifidobacterium* abundance in the gut with a decrease of faecal pH[157]. To determine if there were differences in faecal pH of the Bif /Lacto group vs. the control group, a subset of 74 samples used in the metabolomic analysis were measured for pH. Faeces from infants in the Bif/Lacto group presented a significantly lower pH (5.79 ± 0.80, Figure 34a) than those in the control group (6.85 ± 0.58). Furthermore, this reduced pH in the Bif/Lacto group was maintained throughout the study (from birth to 50 days of life, Figure 34b), and correlated to higher levels of acetate (Figure 34c) and lactate (Figure 34d).

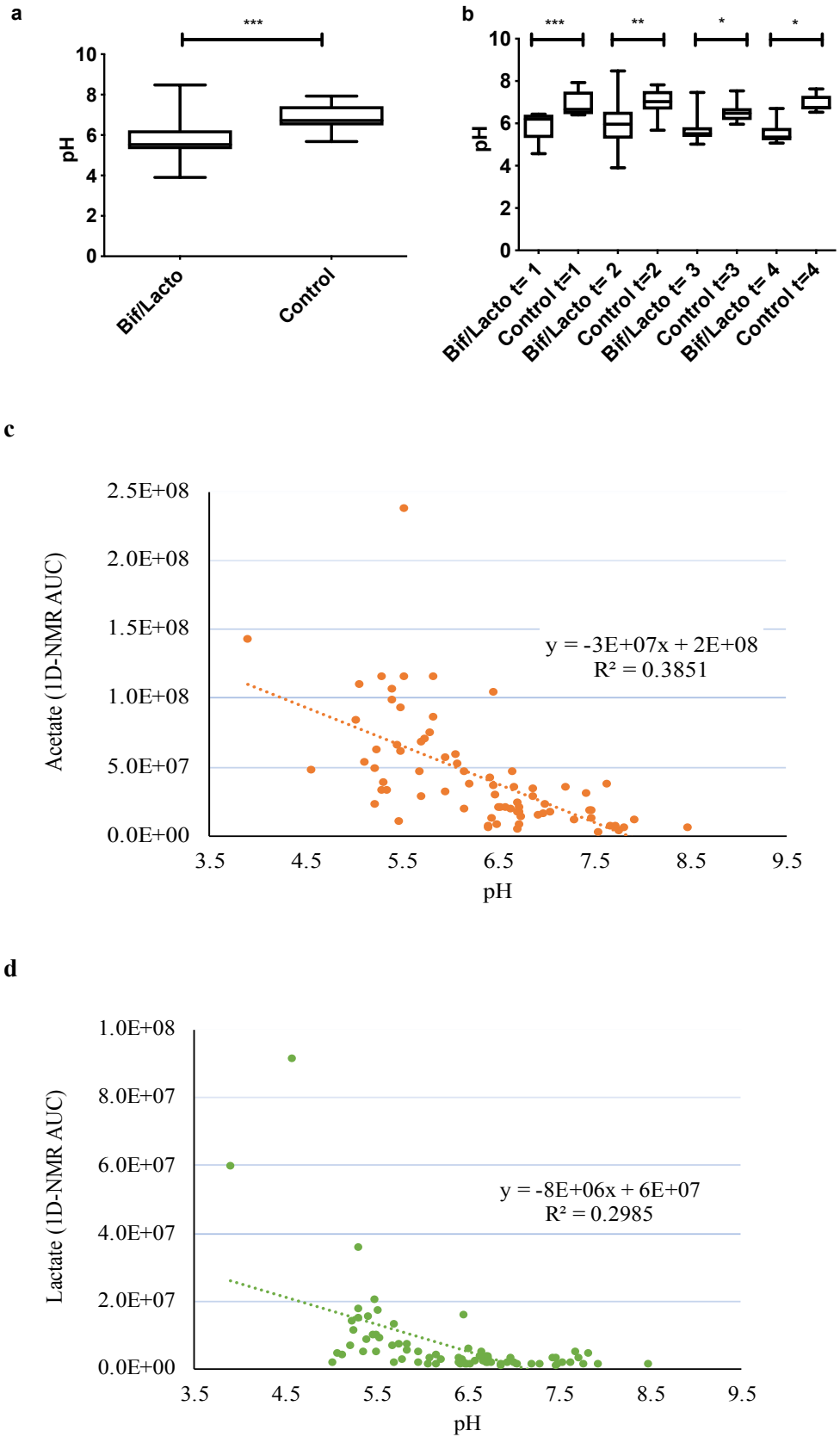


Figure 34 pH faeces measurements from Bif/Lacto Group and control Group

a, pH measurements from faeces samples belonging to Bif/Lacto Group and Control Group. **b**, pH measurements from Bif/Lacto Group and control Group detailing time points of sample collection t=1(<10 days), t=2 (11-29 days), t=3 (20-49 days) and t=4 (>50 days) **c**, Correlation between acetate and pH **d**, Correlation between lactate and pH. Asterisks represent *p* values: *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Bifidobacterium bifidum (Infloran) supernatant inhibited growth of *Staphylococcus haemolyticus* isolated from premature infants' samples

Previous research studies have indicated that *Bifidobacterium* may directly inhibit opportunistic pathogens such as *Staphylococcus epidermis*[158] or *Escherichia coli*[159]. To test the potential production of antimicrobials from *B. bifidum* (Infloran), kinetic growth curves were performed on several pathobionts isolates; *Staphylococcus*, *Klebsiella* and *Escherichia*. Supernatant from an overnight culture of *B. bifidum* (Infloran) including an extra dose of nutrients (see methods for details), was used to grow isolates, and optical density at 595 nm (OD_{595}) was monitored during 24 hours in anaerobic conditions. Interestingly, there was a delay in growth for the *S. haemolyticus* isolate during the first 8 hours when compared to the control (Figure 35a) and suggests presence of a *B. bifidum*-derived antimicrobial. In the case of the *Klebsiella pneumoniae* and *Escherichia coli* isolates, there was no effect on growth when culturing the isolates with the supernatant of *B. bifidum* or when using media with lower pH (to mimic the reduced pH observed in the supplemented premature infants (Figures 35c-d).

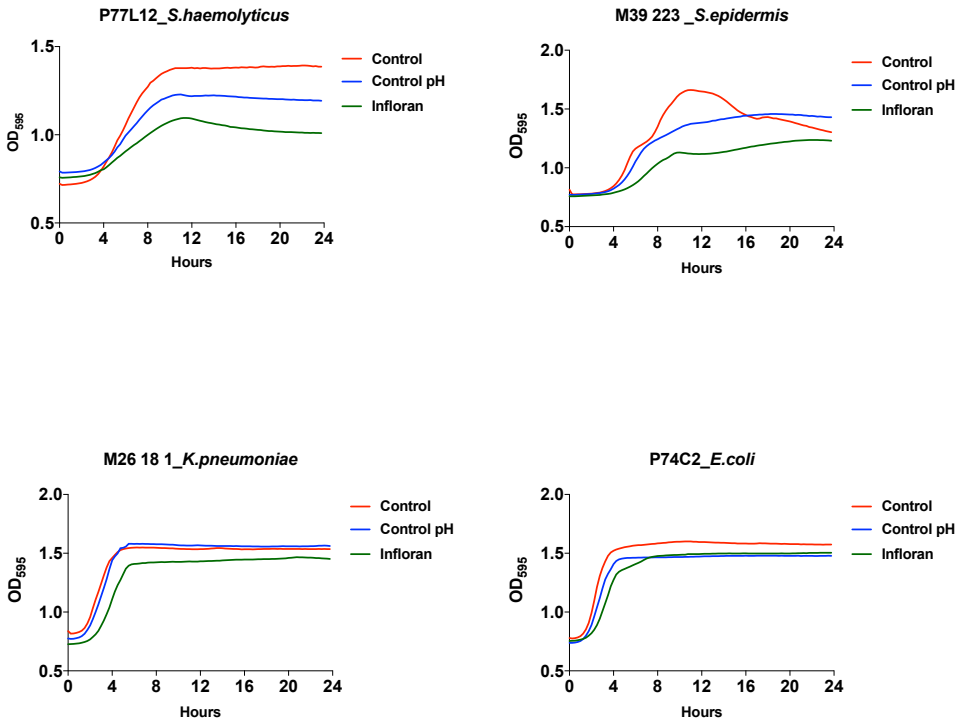


Figure 35 Kinetic growth curves of *Staphylococcus, Klebsiella* and *Escherichia* isolates using supernatant of *Bifidobacterium bifidum* from Infloran

Growth curves of *Staphylococcus* (**a**, **b**)*, Klebsiella* (**c**) and *Escherichia* (**d**) bacterial isolates. Each graph comprises three different growth conditions: (i) ´control´ (red) the bacterial isolate is grown with a rich microbiological media (ii) ´control pH´ the bacterial isolate is grown with a rich microbiological media with pH adjusted to Infloran supernatant and (iii) ´Infloran´ isolate is grown using supernatant of *Bifidobacterium bifidum* (Infloran).

To validate the findings obtained with *Staphylococcus haemolyticus*, the same growth curves were monitored using classical microbial counts (Figure 36). Unexpectedly, the results showed no differences between the growth observed when the *Staphylococcus haemolyticus* isolate was grown using supernatant of *B. bifidum* or with control samples, which may be due to the turbity of bacterial suspension at ~ 8 hours affecting the OD_{595} measurement (Figure 35a).

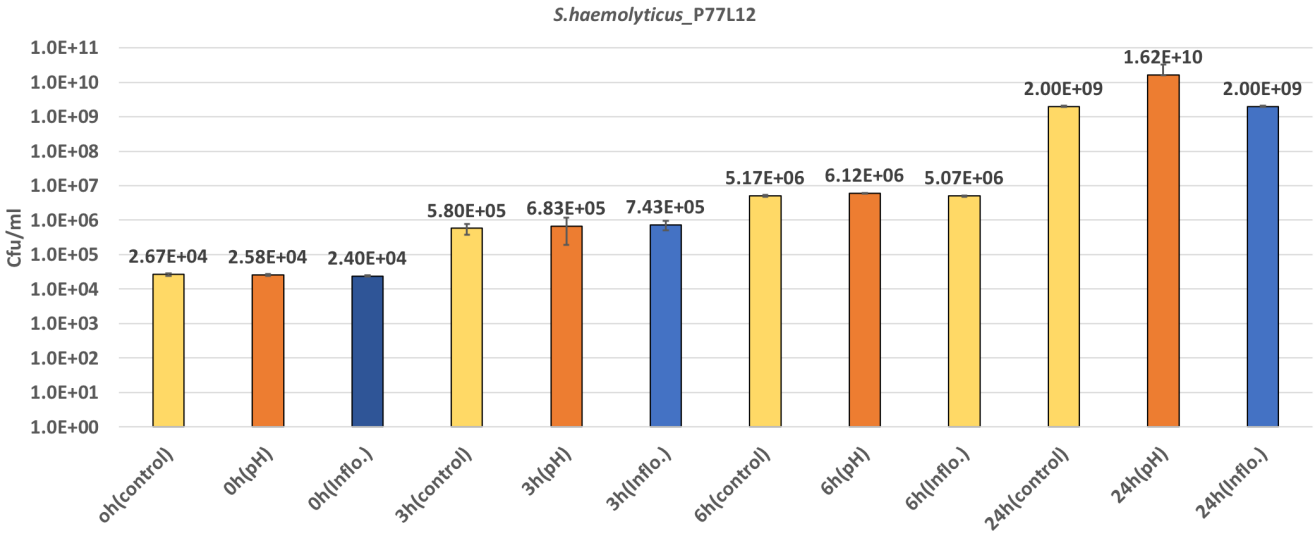


Figure 36 Growth curves of *S. haemolyticus* isolate on supernatant of *Bifidobacterium bifidum* (Infloran)

Growth curves of *S. haemolyticus* isolate using supernatant of BH (yellow), BH with pH adjusted (orange) and *Bifidobacterium bifidum* supernatant (blue).

Discussion

Premature infants are exposed to a wide variety of clinical and environmental factors such as multiple antibiotic treatments, reduced exposure to maternal microbiome, reduced breastfeeding, and prolonged stays in NICU. This can negatively affect the early stages of bacterial gut colonisation[49], and consequently impact on the long-term health outcomes of these infants. Oral supplementation with members of early life microbiome such as *Bifidobacterium* or *Lactobacillus,* has become an alternative for accelerating the establishment of a healthier gut microbiome in premature infants. Numerous clinical studies support that bacterial supplementation significantly reduces the risk of NEC and the incidence of nosocomial infections[160-163]. Notably, a recent clinical audit in the same NICU where the oral supplementation was administered (i.e. Norfolk and Norwich University Hospital), indicated a >50% reduction in NEC rates and late-onset sepsis when comparing 5-year epochs before and after introducing probiotic supplementation, with no episodes of probiotic 'sepsis' indicated[164]. However, before widespread uptake in NICUs across different countries, there is a pressing need to complement these findings with larger-scale meta-studies and state-of-the-

art sequencing and metabolomic techniques, to better understand the complex dynamics of bacterial communities residing in the preterm gut microbiome.

This study represents one the largest reported longitudinal studies of premature infants (n = 591 stool samples), and one of the very few where study groups (Bif/Lacto, control) were matched by gestational age (< 34 weeks gestation), sex, birth mode, time points of sample collection and diet across four different NICUs. Due to the large-scale nature of the study, 16S rRNA sequencing and analysis was used, rather than the significantly more expensive shotgun metagenomics approach. Notably, *Bifidobacterium* represented the most abundant taxa in the Bif/Lacto group, whilst *Staphylococcus*, *Escherichia*, and *Klebsiella* were the most prevalent in the control group. This analysis strongly suggested that supplementation with *Bifidobacterium* and *Lactobacillus* from birth in premature infants, augmented the abundance of *Bifidobacterium* and reduced the incidence of pathobionts which are common denizens of hospital environments. To date, most of the research studies analysing the preterm microbiome in infants not receiving probiotic supplementation, indicated low abundance of the favourable genus *Bifidobacterium*[47, 49, 165]. *Bifidobacterium* represents an important commensal bacteria in early life able to synthesize compounds which can influence host immunity (*e.g. Bifidobacterium breve* has been shown to produce the beneficial metabolite linoleic acid which has potential roles in immune modulation, anticarcinogenic, and antiobesity activities[166]), or promote the production of regulatory T cells ($CD4^+$ $CD25^+$ $Foxp3^+$) involved in fortifying the intestinal gut barrier[167]. Contrastingly, *Lactobacillus acidophilus* (Infloran) did not appear to colonise the lower part of the intestinal tract in premature infants. This latter finding agrees with a small research study done in 2016 using the same supplementation (Infloran) and seven infants[168]. Previous research studies have shown *Lactobacillus* is able to survive in acidic environments with 0.3% bile salts[169], and a study using a single bioreactor to simulate the passage from stomach to intestines showed that *Lactobacillus gasseri* K7 presented a high survival rate in

the stomach-intestine passage [170]. There is the possibility that *Lactobacillus acidophilus* (Infloran) preferentially colonised the premature small intestine, but only a biopsy sample could be used to screen this part of the intestine, which was outwith the ethic boundaries of this study.

As the *Bifidobacterium* genus was shown to be the predominant member in the supplemented group (Bif/Lacto) exerting a potential protective role against colonisation of pathobionts, the study determined which external factors were affecting the abundance of this taxa. Low birthweight and antibiotic treatment clearly impacted negatively on *Bifidobacterium* abundance (Figures 24 and 25), which is in agreement with previous studies[49, 171]. The study also validated antibiotic susceptibility of *Bifidobacterium bifidum* (Infloran) *in vitro* (Table 8), supporting the genomic findings. Surprisingly, when investigating the influence of delivery mode, premature infants born by vaginal birth did not seem to obtain the beneficial commensal bacteria observed in term infants (Figure 26), probably due to the high incidence of prophylactic antibiotic prescription given to premature infants during the first week of life[172].

It has been well documented that not all microbial supplementations have equivalent efficacy at colonising the premature gut[151]. Therefore, this study takes the advantage of whole-genome sequencing and its broad sequencing coverage to assess the colonisation of the supplemented *Bifidobacterium* strain. Results strongly suggested that *Bifidobacterium bifidum* from Infloran colonised the gut of premature infants in the Bif/Lacto group. The fact that the supplementation was given in early days (from first enteral feed) at a high dose (twice daily) and continued for a prolonged period, may have provided the optimal conditions for *Bifidobacterium bifidum* (Infloran) to establish and colonise the premature gut. In addition, it is interesting to note that the majority of the infants in the study were fed with breastmilk, and this particular species of *Bifidobacterium* has been shown

to be an avid utiliser of HMOs[173] contributing to the establishment of other pioneering species of the gut.

Metabolomic analysis on a subset of samples (n = 157) was performed to determine if the changes observed in the gut microbiome were reflected at the functional level, and to elucidate potential interactions involving host and bacteria. Faecal metabolomes from the Bif/Lacto group were distinct from the control group and indicated elevated levels of the short chain fatty acids (SCFAs) acetate and lactate. Notably, a recently study associated the production of SCFAs with a reduction on the faecal pH[157], which correlated with our findings (Figure 32c). In early childhood SCFAs have been shown to play a key role in enhancing innate immunity[174] and increasing the maturation of the enteric nervous system[175]. Moreover, a mouse study associated acetate production by *Bifidobacterium* to gut barrier function, preventing the infection from enterohaemorrhagic *Escherichia coli* O157:H7[176]. In contrast, the metabolome of premature infants from the control group were shown to have high levels of HMOs (2-fucosyllactose and 3-fucosyllactose), probably due to the low levels of HMO bacterial utilisers in their gut.

In vitro analysis measuring the potential production of bacteriocins using the supernatant of *Bifidobacterium bifidum* (Infloran) against other potential pathobionts (*Staphylococcus, Escherichia* and *Klebsiella*) did not reveal any impact on their bacterial growth (Figure 35). Previous research studies indicated the production of bacteriocins in *Bifidobacterium* is generally associated with late logarithmic phase or early stationary phase of growth, but considerable variation exists among species tested, microbiological media utilised and pH conditions used[177]. Further analysis optimising this test could be done, or alternatively, a more complex, but perhaps more informative system, could be trialled, using the model colon with a bacterial mock community representing the preterm microbiome.

Limitations of this study include that this was an observational study and was not planned as a randomised clinical study. It was not possible to make associations of the microbiome and metabolome profiles found with relevant clinical data of the premature infants studied. Looking at the effect of different antibiotic regimes or antibiotic dosing on the premature infant gut microbiome was also out of the scope of this work. Alongside the main microbiological findings of this study, this work provides context for further clinical trials focusing on future intervention studies in this at-risk infant population.

Conclusions

Infloran supplementation modified the microbiome and metabolome of premature infants residing at NICU. Antibiotic treatment and low birth weight are relevant factors influencing disturbances in the preterm microbiome, whilst delivery mode did not have a significant effect. Infloran supplementation exerted an important functional effect enhancing the abundance of short chain fatty acids in the supplemented group.

The large-scale longitudinal study presented here, contributes to the growing knowledge of the preterm infant microbiome, and emphasizes that probiotic supplementation plays a crucial role in exerting protective functional effects on the preterm gut microbial communities.

Future work

This work has offered me the possibility of working closely with doctors and nurses. One of the main challenges of this study has been to obtain the clinical data. As the Bif/Lacto supplementation (Infloran) was so successful at colonising the preterm gut, I would continue this work by addressing in-depth clinical questions where there are still disagreements in probiotic supplementation studies.

Questions such as (i) is Infloran involved in helping premature infants gain weight[178] during their stay at NICUs?, or (ii) does Infloran contribute to reduce NICU stay for premature infants? need answers in this field.

Another interesting area to complement this work would be to evaluate the effects of probiotic supplementation on intestinal mucosa integrity. Commensal bacteria, SCFAs, and antimicrobial peptides are known to be key players at promoting health in the intestinal mucosa by facilitating the assembly of tight junctions, renewal of intestinal cells, and enhance production of mucin[67], a mucus gel coat which forms a protective barrier. Research studies show commensal bacteria are involved in promoting IgA secretion from plasma cells within the gut lumen, defending the mucosa from invasion by pathogens and reducing proinflammatory signals[179]. It would be interesting to measure IgA as well as proinflammatory signals on a subset of faeces samples analysed in this study, to assess whether the Infloran supplementation is enhancing an effect on the mucosal integrity.

On the genomic side, future research on microbial supplementation studies needs to include the utilization of shotgun metagenome profiling, which is becoming more affordable with the decrease in sequencing cost. This approach will provide more information of species and strain level information, as well as identify important bacterial functional pathways such as genes involved in antimicrobial resistance mechanisms.

Chapter 3

Title: Effect of antibiotic treatment on the preterm infant gut microbiome and resistome

Abstract

Premature infants, particularly very low birth weight (VLBW, <1.5 kg), often receive prophylactic antibiotic treatments from birth, to prevent early onset infections. Notably, administration of antibiotics also disrupts the resident gut microbiome, and may create an important reservoir of resistant strains, and of transferable resistance genes, the so called 'resistome', which may correlate with the increasing incidence of antimicrobial resistance (AMR). This study evaluated the impact of antibiotic regimes on the preterm microbiome, and the 'resistome'.

Shotgun metagenomics was performed on longitudinal faecal samples isolated from VLBW infants (n=34) with/without an antibiotic course of benzylpenicillin and gentamicin (short and long courses). Bioinformatics analysis were used to characterise the taxonomic diversity, and the frequency of antibiotic resistance genes. The study cohorts were differentiated between VLBW infants who received probiotic supplementation (probiotic cohort) and VLBW infants who did not receive supplementation (no probiotic cohort), to elucidate whether supplementation contributes to re-establishment of the commensal gut microbiome after antibiotic treatment.

Results indicate that antibiotics and the NICU environment significantly alter the preterm microbiome, with increased representation of potentially 'pathogenic' species such as *Klebsiella pneumoniae, Escherichia coli and Staphylococcus spp.* Probiotic-supplemented VLBW infants presented a higher relative abundance of

Bifidobacterium throughout the study period when compared to non-supplemented infants. The total reservoir of AMR genes within the two study cohorts did not significantly vary among the AMR gene categories detected, nor their AMR gene abundance. These findings may have implications for preliminary guidance for recommendations of the use of antibiotics in VLBW infants residing within neonatal intensive care units.

A graphical abstract summarising the different stages of this study is shown in Figure 37.

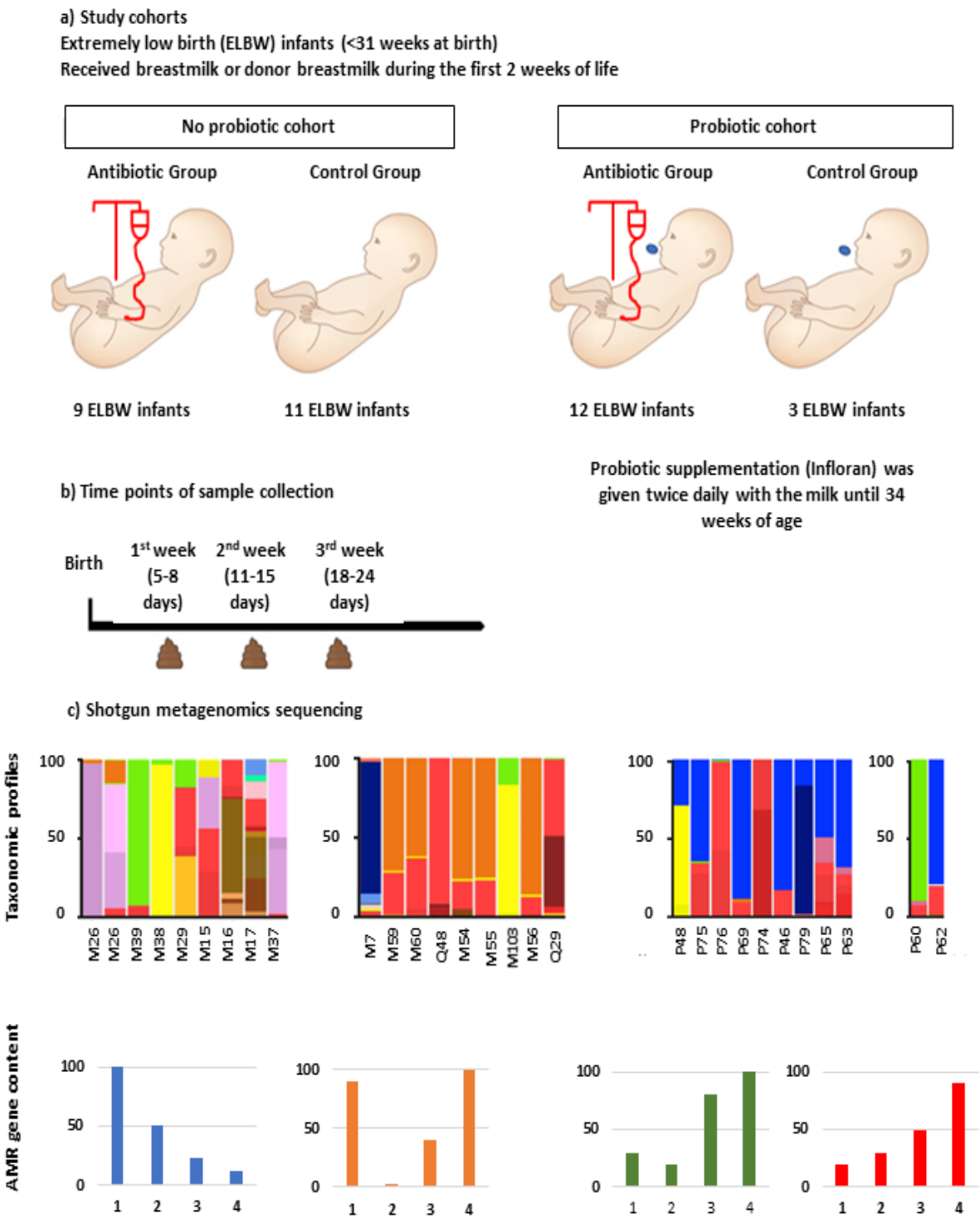


Figure 37 Study pipeline

Premature infants recruited in this study were very low birth (VLBW) infants (<1.5 Kg) who received breastmilk or donor breastmilk during the first two weeks of life. This study includes two different cohorts of VLBW infants: no probiotic supplemented and probiotic supplemented. Samples were collected by nurses at 1st week (ranging from 5-8 days), 2nd week (ranging from 11-15 days) and 3rd week (ranging from 18-24 days). Shotgun metagenomic sequencing was performed to determine bacterial taxonomic profiles and AMR gene content.

Author´s contributions:

I contributed to the overall study design, recruited the samples, prepared them for genomic analysis, performed data analysis, and designed the final figures. Mr Shabhonam Caim ran the bioinformatic pipelines related to taxonomic, AMR gene detection and functional analysis. Dr Will Rowe designed a bespoke bioinformatic pipeline to assign AMR genes to specific taxa. Specific details of all authors' contributions can be found within the Methods section.

Introduction

The immature immune system of premature infants increases their risk of developing neonatal infections *e.g.* bacteraemia/sepsis. Sepsis is a life-threatening condition triggered by the body's immune system in response to a bacterial infection, which can damage vital organs and cause sudden death if not treated at an early stage[180]. Estimates indicate that 23% of neonatal deaths worldwide are caused by infections, and approximately half of these occur during the first week of life[181, 182]. Mortality rates associated with an early-onset sepsis episode increase with prematurity[183]. Very low birth weight (VLBW, <1.5 kg) infants represent the most vulnerable cohort of premature infants with elevated risk of developing early-onset infections (EOIs)[184]. EOIs typically occur in the first 3 days of life and are normally caused by endogenous bacterial pathogens inhabiting the mother's genitourinary tract (*e.g.* Group B *Streptococcus*, *Escherichia coli*, and *Staphylococcus*)[183], which can be vertically transmitted to the infant before or during birth[185, 186].

To protect premature infants from acquiring an EOI, doctors often prescribe antibiotics empirically during the first days of life. The World Health Organisation guidelines recommends a combination of amoxicillin (a common penicillin) and gentamicin as a preventative measure for EOIs[187]. The duration of the antibiotic course is very subjective and is based on the clinician's opinion, antepartum factors

and prematurity of the baby, and often varies from two to seven days. Notably, administration of antibiotics disrupts the resident preterm gut microbiome[188]. Therefore, it is important to understand how antibiotic usage impacts the developing preterm gut microbiome. Antibiotic treatment may have consequences for preterm metabolic and immune development[189], and may also enhance the presence of antibiotic resistance genes (ARGs) in the gut, defined as the 'resistome'[55].

Some Neonatal Intensive Care Units (NICUs) use probiotic supplementation to modify the preterm gut microbiome and reduce the incidence of a devastating bowel disease called necrotising enterocolitis (NEC)[77, 190, 191]. Meta-analysis from several clinical trials have shown that probiotic supplementation using *Bifidobacterium* reduce the incidence of NEC in premature infants[190, 192, 193]. To date, there is a paucity of data available to evaluate the interplay between this supplementation and antibiotic therapy on the developing preterm gut microbiome and the preterm gut resistome. To date there is only one published study on premature infants analysing these interactions[194]. This study demonstrated a higher *Bifidobacterium* abundance in the probiotic group when compared to non-supplemented infants. However the resistome was not significantly different when comparing probiotic supplemented very low birth weight infants (>28 weeks gestation), with more mature non-supplemented infants (28-31 weeks gestation) at 4 weeks and 4 months. Thus, to further understand whether probiotic supplementation contributes to reduced AMR gene carriage, further more controlled studies are required *e.g.* with infants belonging to probiotic supplemented and non-supplemented groups, who also have the same gestational age and antibiotic regimes.

In this study, shotgun metagenomics on faecal samples isolated from 34 VLBW infants was used to interrogate the evolution of the preterm gut microbiome and resistome throughout the first 3 weeks of life. The study includes infants with and

without antibiotic treatment to characterise their metataxonomic diversity, and frequency of antibiotic resistance genes. Furthermore, it includes two study cohorts (probiotic and non-probiotic supplemented) to evaluate if supplementation can re-establish the commensal gut microbiome and reduce the preterm resistome after antibiotic treatment. This is the first preterm study where cohorts and control groups were matched by gestational age (< 33 weeks gestation), antibiotic treatment (*i.e*. benzylpenicillin and gentamicin), diet (mostly breastfeed), birth mode (predominantly born by C-section), and time points of sample collection. Genomic findings were validated using phenotypic testing (*i.e*. Minimal Inhibitory Concentration).

Results from this study indicate that antibiotics and the NICU environment significantly altered the preterm microbiome, and increase representation of potentially multidrug 'pathogenic' species such as *Klebsiella pneumoniae, Escherichia coli, Enterococcus faecalis and Enterobacter cloacae*. The probiotic supplementation used in this study may contribute to reduced abundance of potential bacterial pathogens. Surprisingly, the AMR gene content from both study cohorts (non supplemented and probiotic supplemented) was comparable.

Hypothesis and aims

Hypothesis: Early life administration of antibiotics can lead to disruption of the preterm gut microbiome and also enhance the reservoir of antimicrobial resistance genes (the 'resistome'). Probiotic supplementation can contribute to re-establishment of the commensal gut microbiome after antibiotic treatment and reduce the reservoir of antimicrobial resistance genes.

The study involved two aims/sub-sections:

a) Evaluation of the taxonomic profiles of VLBW infants treated with and without antibiotic treatment to determine the impact of the antibiotic

stewardship on the bacterial populations inhabiting the preterm gut microbiome. As this study includes antibiotic-naive samples, the effect of the NICU environment can also be investigated.

b) Evaluation of the VLBW preterm gut reservoir of antimicrobial resistance genes. This study evaluates whether antibiotic treatment increases the total AMR gene content in the preterm gut microbiome, and determines whether probiotic treated infants have reduced AMR gene content.

Methods

Subject recruitment

Premature infants recruited in this study resided in three Neonatal Intensive Care Units (NICUs): Norfolk and Norwich University Hospital (NNUH, Norwich, UK), Queen Charlotte's and Chelsea Hospital (London, UK) and St Mary's Hospital (London, UK). Sample collection for NNUH was approved by the Faculty of Medical and Health Sciences Ethics Committee at University of East Anglia (UEA) and followed the protocols laid out by the UEA Biorepository (Licence no: 11208). Sample collection for Queen Charlotte's and Chelsea Hospital and St Mary's Hospital was approved by West London Research Ethics Committee (REC) under the REC approval reference number 10/H0711/39. Doctors and nurses recruited infants after parents gave written consent.

Sample collection

All NICUs included in this study presented similar protocols for feeding, prescription of antibiotics and antifungal drugs. The main exception was that the NNUH routinely prescribed all VLBW an oral probiotic supplementation (Infloran®, Desma Healthcare, Switzerland) twice daily, whereas St Mary's Hospital and Queen Charlotte's and Chelsea Hospital did not use probiotic

supplementation. This supplementation contained *Bifidobacterium bifidum* and *Lactobacillus acidophilus* in a dose of 1 x 10^9 cfu of each species, and was given from birth until 34 weeks old.

A total of 34 VLBW infants under 31 weeks gestation were recruited for this study, and were either treated with only benzylpenicillin and gentamicin (antibiotic treated) or no antibiotic treatment (control group) during the first days of life. We specifically selected infants who were given breastmilk or donor breastmilk during the first 2 weeks of life, with the aim of normalising feeding regimes between the two study cohorts. Faeces were collected at time points 1st, 2nd and 3rd week of their NICU stay. Details of the VLBW infants recruited for this study can be found in Appendix 2. I performed the selection of the samples included in this study.

DNA extraction

DNA extraction was performed using the FastDNA Spin Kit for Soil (MP Biomedical, Santa Ana) following the manufacturer's instructions, but extending the bead-beating step to 3 minutes, and eluting the DNA with 55 °C sterile water. The DNA recovered from these samples was assessed using a Qubit® 2.0 fluorometer (Invitrogen). I performed the DNA extractions.

Shotgun metagenomics library preparation and sequencing

Samples containing 500 ng of genomic DNA were placed into a Covaris plate with glass wells and DNA was sheared into fragments of approximately 450 bp. The sheared DNA was purified and concentrated using an SPRI-cleanup kit (Beckman,USA). Library construction entailed an end repair, A-tailing and adapter ligation steps. Adapter ligated samples were amplified and indexed by PCR using established Illumina paired end protocols. A portion of each library was used to create an equimolar pool and enriched libraries were subjected to 125 base paired

end sequencing on a HiSeq 2500 V4. Library preparation and sequencing were performed at the Sanger Institute, Cambridge, UK.

Taxonomic and functional profile analyses

Sequencing files were quality assessed with the FASTX-Toolkit. Subsequent taxonomic analysis was performed using MetaPhlAn v2.0 (http://huttenhower.sph.harvard.edu/metaphlan2) and depicting of the paired read sequences was performed using R. Principal Coordinate Analysis (PCoA) was performed using Bray-Curtis distances on shotgun taxonomic profiles in MEGAN version 6.0. Functional annotation was performed using EggNOG mapper (version 1.0.3) based on EggNOG orthology[195] data. Mr Shabhonam Caim ran the bioinformatic analysis and I finalised the figures using MEGAN.

Identifying antimicrobial resistance genes

Presence/absence of AMR genes within the samples was performed using two different bioinformatics approaches: (i) linear approach where metagenomic contigs produced by MEGAHIT[196] were aligned to the CARD database version 2.0.1 (https://card.mcmaster.ca/download) using a filtering criteria of e-value $1e^{-10}$ and 90% identity, and (ii) a non-linear approach which combines variation graph representation of gene sets with a Locality Sensitive Hashing (LSH) indexing scheme. This latter approach was performed in collaboration with Dr Will Rowe who is the developer of this pipeline[197].

Isolation and characterisation of *Bifidobacterium*, *Enterococcus*, *Staphylococcus*, *Klebsiella* and *Escherichia* strains from VLBW infant faeces

In order to validate the results obtained from the genomic analysis, I performed bacterial isolation, targeting the most abundant bacterial taxa present in samples. 50-25 mg of faecal sample was homogenised in 5 mL of phosphate buffer saline

(PBS) by vortexing. Homogenates were serially diluted to 10^{-4} in PBS buffer and aliquots of 100 ml were spread on different selected medium: MacConkey (Oxoid), MRS (Difco^{tm}) with 50 mg/L mupirocin (Oxoid), Baird-Parker (Oxoid) and Slanetz and Bartley medium (Oxoid). Agar plates were incubated aerobically (MacConkey, Baird-Parker and Slanetz and Bartley agar) and anaerobically (MRS agar) at 37 °C over three days. Five colonies from each agar plate were streaked three consecutive times onto new nutrient agar plates to assure purity. The DNA was extracted using the FastDNA Spin Kit for Soil (MP Biomedical, Santa Ana). The 16S rRNA gene was PCR amplified with the primers: fD1 (5'- AGA GTT TGA TCC TGG CTC AG - 3'), fD2 (5'- AGA GTT TGA TCA TGG CTC AG - 3') and rP1 (5' - ACG GTT ACC TTG TTA CGA CTT - 3')[198]. The PCR conditions were: 1 cycle of 94 °C for 5 min, followed by 35 cycles of 94 °C for 1 min, 43 °C for 1 min, and 72 °C for 2 min followed by a final strand extension at 72 °C for 7 min. All amplicons were sequenced using an automated Sanger sequencing service (Eurofins Genomics, Luxembourg), and sequences were used to identify the species isolated using Basic Local Alignment Search Tool (BLAST). I performed this procedure.

Determination of Minimal Inhibitory Concentration (MIC) for bacterial isolates

Calculation of the Minimal Inhibitory Concentration (MIC) of bacterial isolates against commonly used antibiotics in NICUs was performed using the broth microdilution method[199]. Five *Klebsiella pneumoniae* isolates from baby M26 at time points 7, 14 and 21 days, were tested against benzylpenicillin and meropenem. Serial two-fold dilutions of these two antibiotics were added to sterile nutrient broth, and 10 µl from a fresh overnight culture of the isolates was added in each well. Microplates were incubated for 24 h at 37 °C under aerobic conditions. Cell density was monitored using a plate reader (BMG Labtech, UK) at 595 nm. MICs were determined as the lowest concentration of antibiotic inhibiting any bacterial growth. I performed this procedure.

DNA extraction from bacterial isolates for whole-genome analysis

Overnight bacterial isolates were centrifuged, re-suspended in 30 ml of PBS (Sigma-Aldrich, UK) and centrifuged again. The pellet was re-suspended in 2 ml of 25% sucrose (Fisher Scientific) in TE buffer (10 mM Tris and 1 mM EDTA at pH 8.0) and 50 µl of Roche Lysozyme (Roche Molecular Systems, UK) at 100 mg/ml in 0.25 M Tris pH 8.0. The mixture was incubated at 37 °C for 1 h, and 100 µl of Proteinase K at 20 mg/ml (Roche Molecular Systems), 30 µl of RNase A at 10 mg/ml (Roche Molecular Systems), 400 µl of 0.5 M EDTA pH 8.0 (VWR Chemicals), and 250 µl of freshly prepared 10% Sarkosyl NL30 (Sigma-Aldrich) were added. The mixture was incubated on ice for 2 h and transferred to a water bath at 50 °C overnight. Next, Elution Buffer (10 mM Tris pH 8.0 (Fisher Scientific) was added to the sample to a final volume of 5 ml, mixed with 5 ml Phenol:Chloroform:Isoamyl Alcohol (PCIA) (25:24:1) (Sigma-Aldrich) in a Qiagen MaXtract High Density tube (Qiagen, Germany) and centrifuged for 15 min at 1792g. The aqueous phase was then transferred into a new Qiagen MaXtract High Density tube, made up with Elution Buffer to the volume of 5 ml, mixed with 5 ml of PCIA, and centrifuged for 10 min at 1792g. This procedure was repeated, with 5 min centrifugation time. Next, the aqueous phase was transferred into a Qiagen MaXtract High Density tube, made up to 5 ml with Elution Buffer, mixed with 5 ml of Chloroform:Isoamyl Alcohol (CIA) (24:1) (Sigma-Aldrich, UK), and centrifuged for 5 min at 1792g. The CIA step was then repeated once more, after which the final aqueous phase was transferred into a sterile Corning ^{TM} 50 ml centrifuge tube, and 2.5 volumes of ethanol (Ethanol absolute AnalaR NORMAPUR®, VWR Chemicals, USA) were added. The sample was incubated for 15 min at 20 °C, and centrifuged for 10 min at 1792g at 4 °C. Finally, the pellet was washed with 10 ml of 70% ethanol and centrifuged at 1792g for 10 min twice, dried overnight, and re-suspended in 300 µl of Elution Buffer. I performed this procedure.

Results

The preterm gut microbiome in non-probiotic supplemented VLBW infants is different from probiotic supplemented VLBW infants

Shotgun metagenomics sequencing was used to capture the bacterial community profiles between the two study cohorts (non-probiotic supplemented and probiotic supplemented). Each cohort included VLBW infants treated with antibiotics (benzylpenicillin and gentamicin) from 3 to 8 days of treatment, and VLBW infants who did not receive any antibiotic treatment (considered as 'control group').

I initially performed Bray Curtis distances and Principal Coordinate Analysis to determine the overall taxonomic profiles of the two study cohorts (non-probiotic and probiotic supplemented). PCoA visualisation indicated that each cohort clustered into two separate groups (Figure 38), suggesting the taxonomic profiles from both cohorts are clearly distinct.

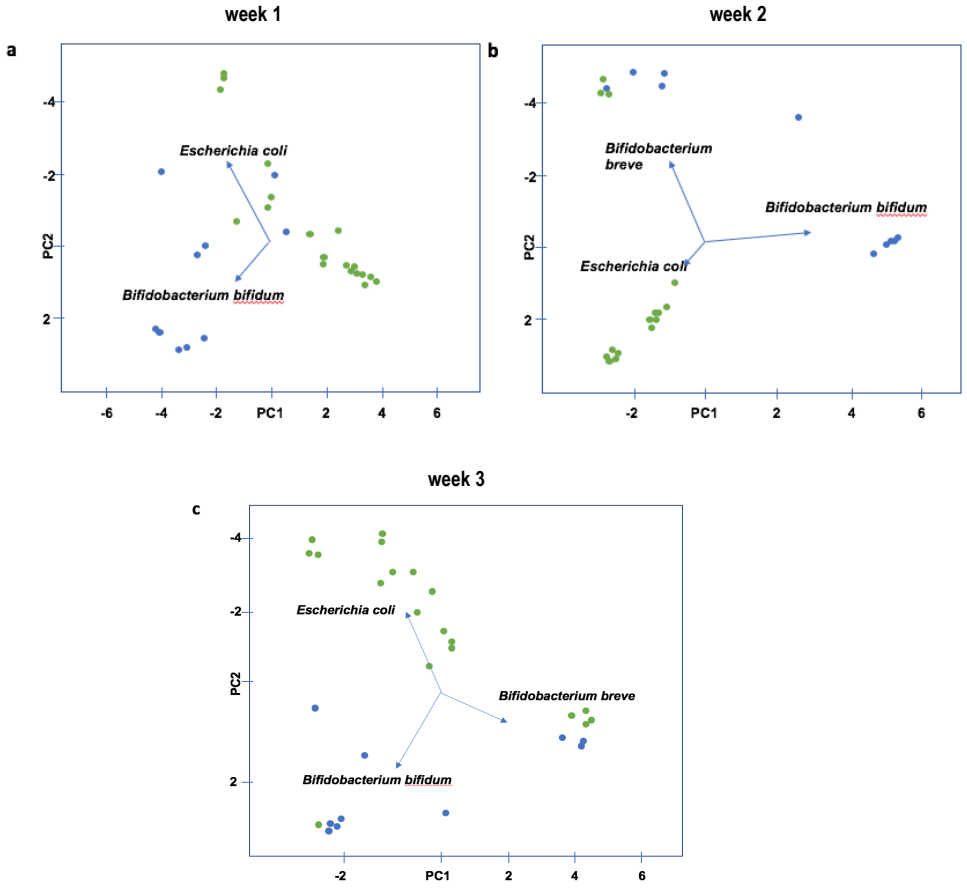


Figure 38 Principal Coordinate Analysis (PCoA) from samples tested in the study

PCoA plots were performed based on the Bray-Curtis distances calculated from the relative abundance of the taxonomic profiles. Samples were grouped according to the time point of collection: a) PCoA plot from samples analysed at 1st week of the study b) PCoA plot from samples analysed at 2nd week and c) PCoA plot from samples analysed at 3rd week. Samples from non-probiotic supplemented infants were highlighted in green, and samples from probiotic supplemented infants were marked in blue.

The main differences between the two study cohorts were driven by the genus *Bifidobacterium,* which was consistently enhanced in the probiotic cohort throughout the study period, and the genera *Escherichia* and *Klebsiella* which were increased in the non-probiotic cohort at the 2nd and 3rd week of the study (Figure 39).

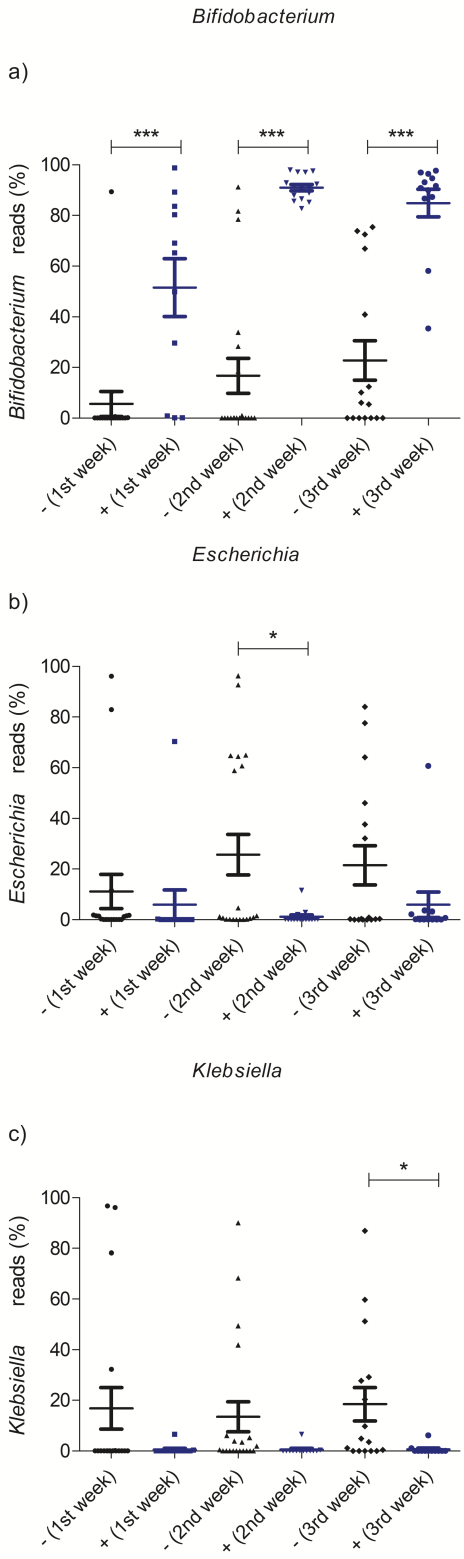


Figure 39 *Bifidobacterium*, *Escherichia* and *Klebsiella* abundance throughout the study period

Relative abundance of *Bifidobacterium* (a), *Escherichia* (b) and *Klebsiella* (c) at the sample time points of the study: 1st week, 2nd week, and 3rd week. Samples were grouped into VLBW infants receiving probiotic treatment (highlighted with +) and VLBW infants non-supplemented (highlighted with -).

Antibiotic treatment and the NICU environment impact profiles of multidrug resistance bacteria within the preterm gut microbiome

More in-depth visualisation of taxonomic profiles from the non-probiotic cohort, indicated that the most abundant bacterial species found were *Klebsiella pneumoniae*, *Escherichia coli*, *Enterococcus faecalis*, *Staphylococcus epidermis* and *Enterobacter cloacae* (Figure 40). These bacteria have all previously been described as prevalent members of the gut microbiome of VLBW infants[120, 200, 201]. Interestingly, premature infants who received the longest antibiotic treatment (infants M26, M36 and M39), presented the highest levels of *K. pneumoniae* (90%-96%), *Klebsiella oxytoca* (47%-49%) and *Enterococcus faecalis* (92%-33%) during the first two weeks of the study (Figure 40a and 40b). Taxonomic profile similarities at the 2nd and 3rd week of the study, in both the antibiotic and the non-antibiotic group (Figure 40b and 40c), suggested the NICU environment may also play a role (in tandem with antibiotic treatment) at colonising the premature infant gut microbiome with pathobionts.

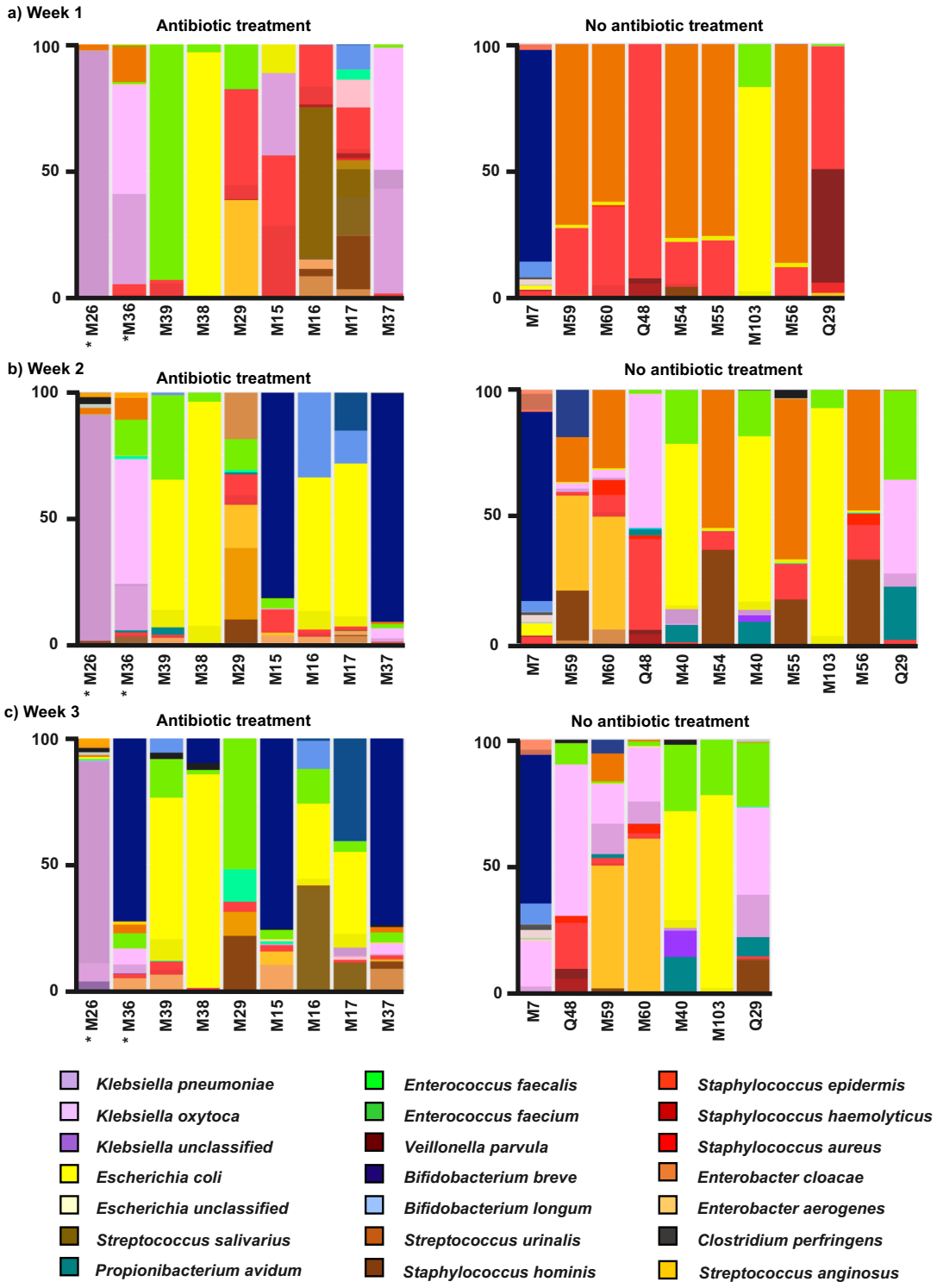


Figure 40 Comparison of bacterial profiles analysed for non-supplemented probiotic cohort (antibiotic vs non-antibiotic treatment)

Relative abundance of each taxon was represented in percentage value, and bar colours represent different species found. Time points of study were classified as a) 1st week of study, b) 2nd week of study and c) 3rd week of study. Asterisks represent infants receiving a long antibiotic treatment.

Microbial profiles for the probiotic supplemented cohort indicated key differences from the non-probiotic cohort. Overall, there was a predominance of *Bifidobacterium bifidum* throughout the study period; average abundance of this bacteria taxa per infant was 44% at 1st week of study, 56% at 2nd week, and 43% at 3rd week (Figure 41). Notably, *B. bifidum* was present in the daily probiotic supplementation these infants received. Interestingly, colonisation of other *Bifidobacterium* species at the 2nd and 3rd week of the study was also observed, such as *Bifidobacterium breve* and *Bifidobacterium longum* (Figures 41b, and 41c). Different *Bifidobacterium* spp. and strains are known to colonise infants, and to potentially cross-feed within the bifidobacterial community[202]. I thus hypothesise that initial colonisation of *B. bifidum* at the 1st week of life, may promote establishment of subsequent bifidobacterial colonisers.

It is important to highlight that other bacterial taxa such as *K. pneumoniae*, *E. coli*, *E. faecalis*, and *E. cloacae,* which were found in high abundance in the non-probiotic cohort, were all present at low levels (e.g.: <10% for *Escherichia coli* and <6% for *Klebsiella pneumoniae*), with the only exception of three infants in this cohort (P48, P60 and P42, Figures 41a and 41c).

The probiotic cohort was also affected by the antibiotic treatment with increased levels of *Staphylococcus* and *Escherichia* (Figure 41a infants P48, P75, P76, P74, P65 and P63) and reduced levels of *Bifidobacterium* during the 1st week of the study. Despite this, the cohort had a predominance of *Bifidobacterium* throughout the study period, which suggested this genus was able to displace other potential pathogenic bacteria present in the NICU environment.

Overall, these data suggest that empiric antibiotic treatment favours the growth of potential pathogenic bacteria (*e.g. Staphylococcus, Escherichia* and *Klebsiella*) in both probiotic and non-probiotic treated cohorts. The NICU environment alongside the antibiotic treatment influenced the preterm gut microbiome in the non-probiotic

cohort, whilst the probiotic supplementation may provide colonisation resistance against pathogenic bacteria.

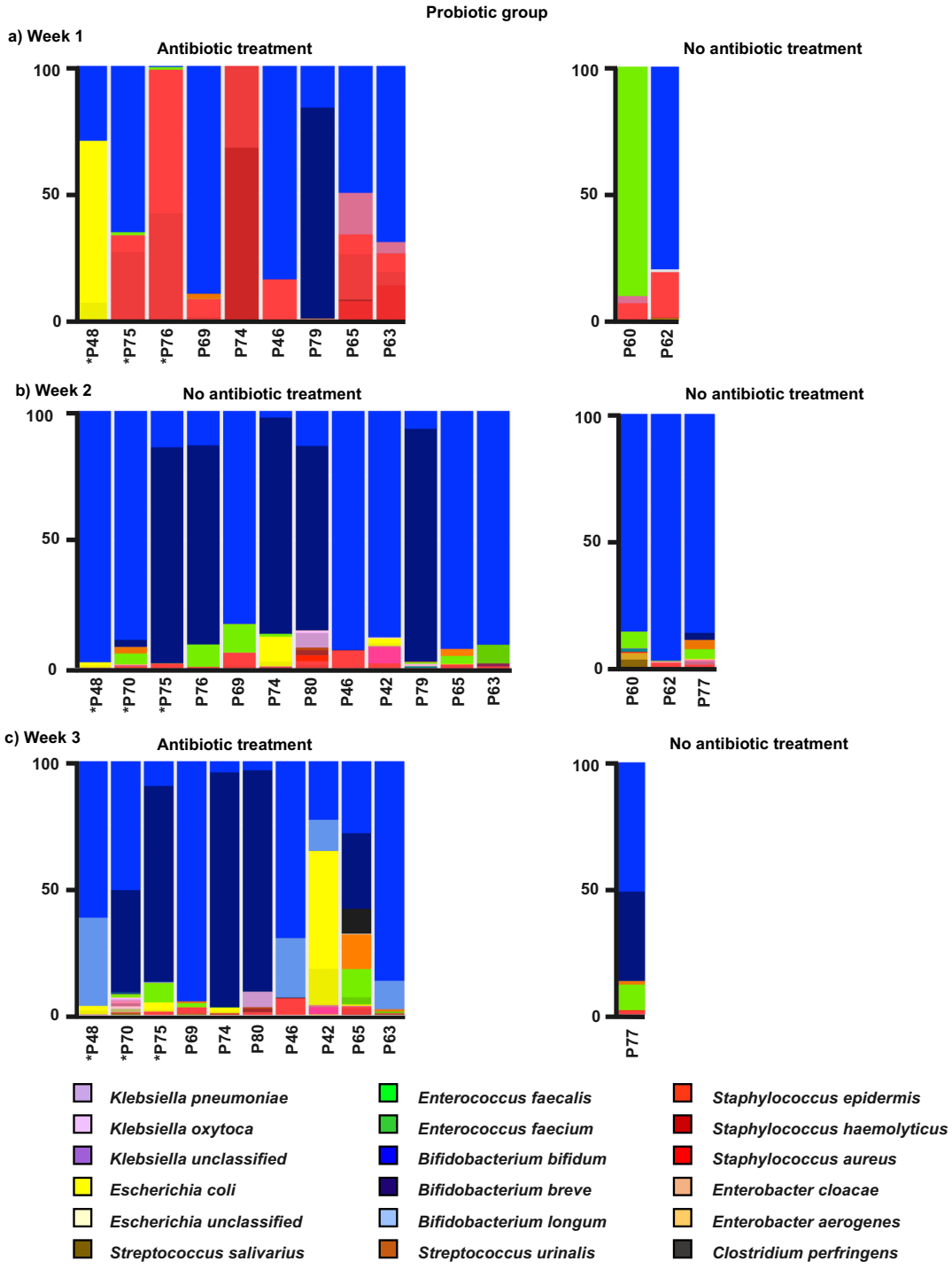


Figure 41 Comparison of bacterial profiles analysed for supplemented probiotic cohort (antibiotic treated vs non-antibiotic)

Relative abundance of each taxon was represented in percentage value, and bar colours represent different species found. Time points of study were classified as a) 1st week of study, b) 2nd week of study and c) 3rd week of study. Asterisks represent infants receiving a long antibiotic treatment.

Functional bacterial categories from the non-probiotic supplemented VLBW infants and probiotic supplemented VLBW infants were comparable

Functional pathways were calculated using EggNOG which allows determination of the functional annotations from sequencing data. EggNOG uses clusters of orthologous groups (OGs), in this case homologous sequences that started diverging from the same speciation event, which are later on functionally annotated using phylogenetic methods[203]. When comparing the functional categories from the non-probiotic and probiotic cohort, I did not observe significant differences (Figure 42), which was in contrast to the observed differences in taxonomic profiles, shown in the previous section (Figures 40 and 41). Four categories predominated among both study cohorts: "Energy production and conversion (C)", "Amino-acid transport (E)", "Carbohydrate transport (G)" and "Inorganic ion transport (P)" (Figure 43). It is interesting to highlight that there was a big proportion of functional categories with "unknown function (S)".

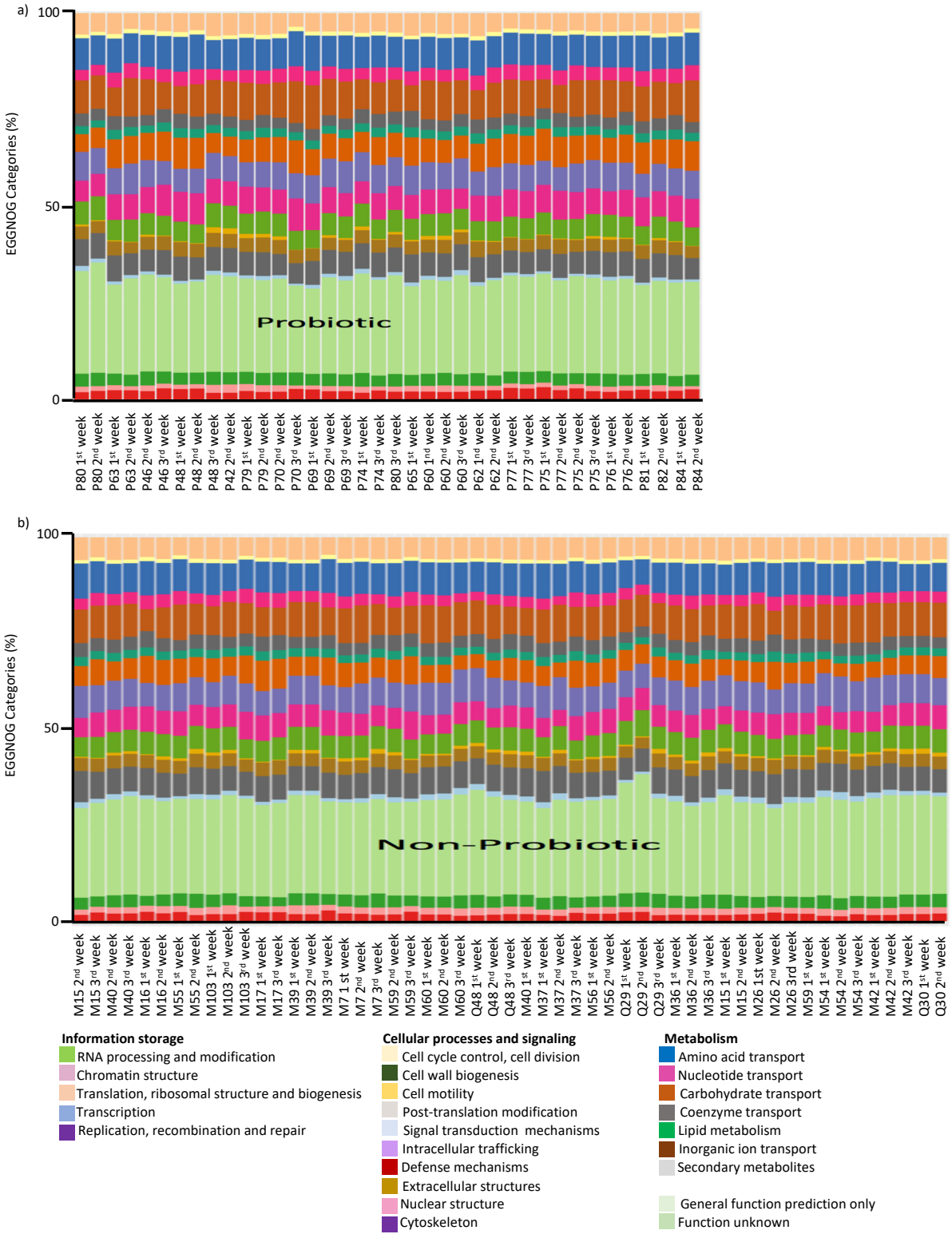


Figure 42 EggNOG functional category analysis

Results from EggNOG functional category analysis for a) study samples belonging to probiotic cohort, and b) study samples belong to non-probiotic cohort. All identified proteins were classified into molecular families species in the figure legend.

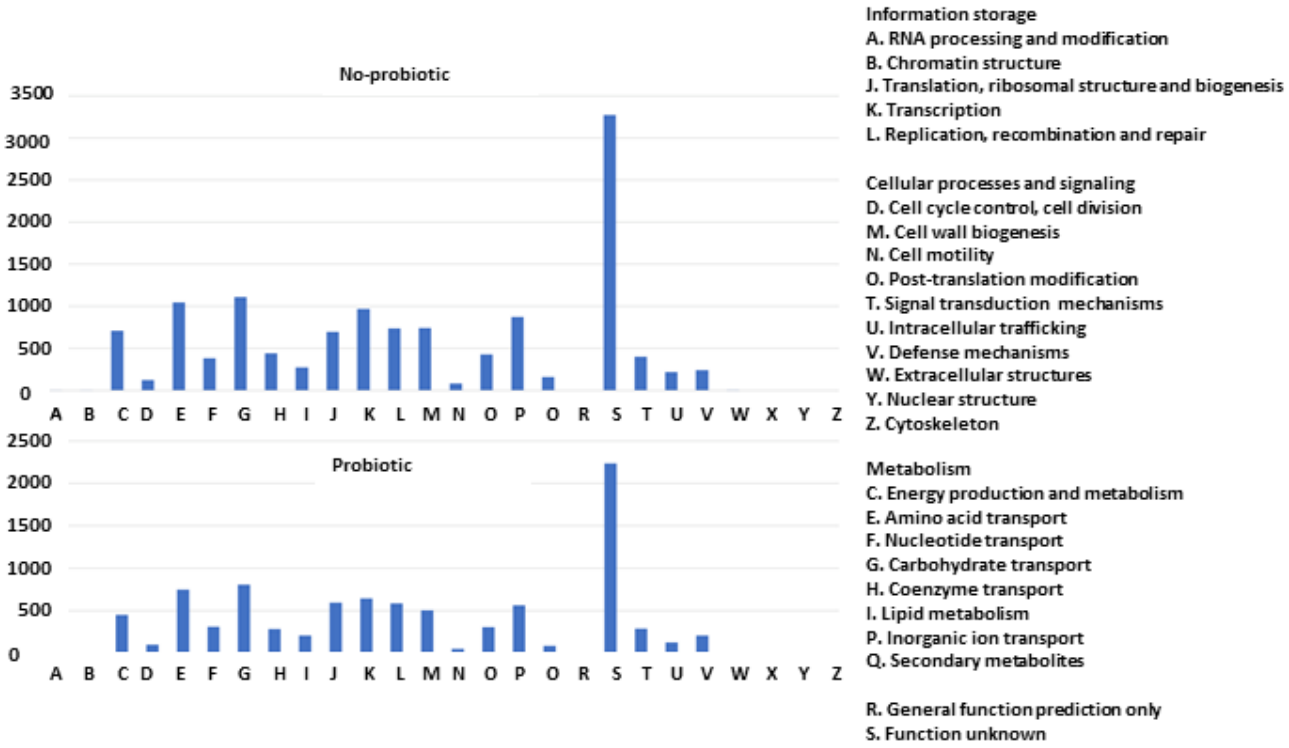


Figure 43 Average of EggNOG functional categories for non-probiotic and probiotic cohorts

The letters on the x-axes represent the different EggNOG categories. The legend on the right-hand side explains the details of the functional categories.

Prevalence of antimicrobial resistance genes in the non-probiotic and probiotic cohort was similar

The presence of antimicrobial resistance genes within the two study cohorts (probiotic and non-probiotic) was calculated by initially aligning the metagenomic contigs to the Comprehensive Antibiotic Resistance Database (CARD). This approach provided an overall picture of the antimicrobial resistance gene profiles from the probiotic and non-probiotic cohort. In total, 229 different AMR genes were detected for the non-probiotic cohort and probiotic cohort. The average number of AMR genes identified per infant in the non-probiotic cohort was 27 (ranging from 0-70) and in the probiotic cohort was 17 (ranging from 1-56), see Table 10. These values indicate there is variability within the subject cohorts, which has been shown previously[200, 204].

Table 10 AMR genes detected in the study cohorts. Samples are classified according to study cohorts (probiotic and non- probiotic). Cells are colour coded using a colour scale for different cell values.

			Infants recruited								
			M39	M26	M37	M38	M15	M16	M17	M29	M36
No probiotic	Antibiotic treatment	1st week	0	9	13	61	2	17	16	19	10
		2nd week	51	10	7	62	11	59	58	18	16
		3rd week	62	15	22	57	11	61	69	14	19
			M7	M59	M60	Q48	M40	M55	M103	M56	Q29
	No antibiotic	1st week	51	12	12	13	ND	7	60	11	16
		2nd week	70	8	19	12	67	10	1	16	11
		3rd week	9	12	17	25	67	ND	61	ND	7

			Infants recruited									
			P69	P74	P65	P75	P76	P63	P46	P48	P79	P80
Probiotic	Antibiotic treatment	1st week	16	8	19	19	12	11	12	42	3	ND
		2nd week	16	47	13	13	13	9	9	23	2	25
		3rd week	18	30	56	56	ND	1	9	13	ND	12
			P60	P62	P77							
	No antibiotic	1st week	14	12	ND							
		2nd week	14	4	14							
		3rd week	ND	ND	17							

Interestingly, the total AMR gene content, when comparing samples treated with and without antibiotics within each cohort (probiotic and non-probiotic), was not significantly different (Table 10). However, an overall comparison of the cohorts indicated that the total AMR content trended towards higher in the non-probiotic cohort than the probiotic cohort, although *p-values* were not significant (1st week *p-value* =1, 2nd week *p-value*=0.398 and 3rd week *p-value*= 0.230). A higher number of infant samples would be needed to obtain a statistical significance.

When AMR genes were grouped according to mechanism of action, these aligned to 22 different antibiotic classes. Table 11 summarises the number of AMR genes (normalised per number of infants tested) for each antibiotic class. This table highlights study cohorts (probiotic and non-probiotic supplemented) as well as time points of the study (1st week, 2nd week, and 3rd week). The most prevalent antibiotic classes included antibiotic efflux pumps (commonly found in many bacterial

taxa[205]), aminoglycoside resistance genes (confer resistance to gentamicin), β-lactamases (confer resistance to benzylpenicillin) and bacitracin resistance (a common antibiotic used for topical applications). Contrary to expectations, there was not a higher abundance of aminoglycoside and β-lactamase genes during the first week when infants were administered antibiotics. Notably, the preterm gut microbiome appears to be colonised with a higher number of bacterial species/strains at the end of the study, (*i.e.* week 2 to 3) than at the earlier stages (*i.e.* week 1). Thus, simple qualitative analysis of presence or abundance of AMR genes may be of limited use to fully assess these differences.

Table 11 Antimicrobial resistance genes detected for probiotic and non-probiotic cohort. Numbers represent the average of AMR genes detected using CARD database divided by the number of premature infants included in each group. AMR genes are grouped according to mechanism of action. Cells are colour coded using a colour scale for different cell values.

	Probiotic group			No probiotic group		
	1st week	2nd week	3rd week	1st week	2nd week	3rd week
Aminoglycoside	1.6	0.8	0.1	1.2	1.4	2.1
Efflux pump	5.4	8.4	9.6	7.4	14.2	14.4
Isoleucyl-tRNA synthetase (ileS)	1.0	1.0	1.0	0.0	0.0	0.0
Bacitracin	1.6	2.8	3.9	1.7	6.4	6.4
beta-lactamase	0.4	1.2	1.2	1.6	2.0	2.9
Fosfomycin	0.1	0.1	0.0	0.4	0.3	0.7
Fusidic acid	0.8	0.2	0.2	0.4	0.2	1.0
Lincosamide	0.1	0.1	0.1	0.1	0.2	0.0
Macrolide	1.3	0.8	0.6	0.6	0.9	0.3
Methicillin	0.8	0.5	0.4	0.4	0.2	0.3
Rifamycin	0.3	0.8	0.0	0.3	1.3	1.4
Streptothricin acetyltransferase	0.1	0.0	0.0	0.1	0.2	0.0
Sulfonamide	0.3	1.2	1.6	0.7	2.7	2.2
Tetracycline	0.3	0.8	0.8	0.3	0.9	1.2
Trimethoprim	0.8	0.8	0.6	0.7	1.2	1.9
Bacterial porin	0.0	0.0	0.1	0.1	0.0	0.1
Defensin	0.0	0.0	0.1	0.0	0.0	0.0
Streptogramin vat transferase	0.0	0.0	0.1	0.0	0.0	0.0
Porin to beta-lactams	0.0	0.0	0.1	0.0	0.0	0.0
Quinolone	0.0	0.0	0.1	0.0	0.0	0.0
Vancomycin	0.0	0.0	0.0	0.2	0.1	0.0

Assigning antimicrobial resistance genes to specific taxa using the GROOT pipeline

Although some differences in the total AMR gene abundance was detected within the study cohorts, it is clinically important to determine whether a pathogen or a commensal bacteria carries specific AMR genes. A new bioinformatic tool called GROOT developed by Dr Will Rowe was used to assign antimicrobial resistance genes to specific taxa. This approach combined variation graph representation of gene sets with a Locality Sensitive Hashing (LSH) indexing scheme[206]. The bioinformatic work performed in this part of the study was carried out by Dr Will Rowe.

Tables 12 and 13 summarise the results obtained when running the metagenomic shotgun data from this study using the GROOT pipeline. Overall, potentially pathogenic bacteria found in the study samples (*e.g. Staphylococcus spp, Enterococcus faecalis, Escherichia coli*, *Klebsiella pneumoniae,* and *Streptococcus anginosus*) were the main carriers of the AMR genes detected, and commensal taxa like *Bifidobacterium* did not carry AMR genes with the only exception being infant P62 (Table 12). Interestingly, all the potential pathogenic bacteria detected harbored resistance genes towards the empiric antibiotic treatment used in this study (*i.e*. β-lactamases and aminoglycosides). These data suggest the empiric antibiotic treatment used in this study may not protect VLBW infants from potential pathogens, and it may actually favor growth of those strains that are also multidrug resistance.

Table 12 AMR genes detected for probiotic cohort using GROOT

Infant	Time point	GENE detected using GROOT	AMR Group	ASSOCIATED TAXA			
				top genus	% classified	top species	% classified
P74	1st wek	aac(6')-aph(2'')_1_M13771	Aminoglycoside	*Staphylococcus*	84	haemolyticus	69
		aph(3')-III_1_M26832	Aminoglycoside	*Staphylococcus*	83	haemolyticus	53
		blaZ_36_AJ400722	β-lactamase	*Staphylococcus*	84	haemolyticus	58
		mecA_10_AB512767	β-lactamase	*Staphylococcus*	84	haemolyticus	53
		mecA_15_AB505628	β-lactamase	*Staphylococcus*	82	haemolyticus	53
		mph(C)_2_AF167161	Macrolide antibiotic	*Staphylococcus*	84	haemolyticus	58
P60	1st week	aac(6')-aph(2'')_1_M13771	Aminoglycoside	*Enterococcus*	32	faecalis	31
		aph(3')-III_1_M26832	Aminoglycoside	*Enterococcus*	32	faecalis	31
		erm(B)_1_JN899585	tetracycline resistance	*Enterococcus*	41	faecalis	39
		erm(B)_12_U18931	tetracycline resistance	*Enterococcus*	41	faecalis	40
		mecA_10_AB512767	β-lactamase	*Staphylococcus*	53	epidermis	24
		mecA_15_AB505628	β-lactamase	*Staphylococcus*	55	epidermis	24
P62	1st week	fosB_4_CP000029	fosfomycin thiol transferase	*Staphylococcus*	52	epidermis	30
		mecA_10_AB512767	β-lactamase	*Staphylococcus*	63	epidermis	18
		tet(W)_4_AJ427422	tetracycline resistance	*Bifidobacterium*	99	bifidum	92
P62	2nd week	erm(C)_13_M13761	Macrolide antibiotic	-	-	-	-
		tet(M)_7_FN433596	tetracycline resistance	-	-	-	-
P77	1st week	aac(6')-aph(2'')_1_M13771	Aminoglycoside	-	-	-	-
		blaZ_35_AJ302698	β-lactamase	-	-	-	-
P77	2nd week	blaZ_36_AJ400722	β-lactamase	*Staphylococcus*	100	-	-
		tet(K)_4_U38428	tetracycline resistance	*Staphylococcus*	83	-	-
		tet(K)_5_J01764	tetracycline resistance	*Staphylococcus*	100	-	-
P75	1st	aac_6__-aph_2____1_M13	Aminoglycoside	*Staphylococcus*	79	haemolyticus	74
		blaZ_36_AJ400722	β-lactamase	*Staphylococcus*	75	haemolyticus	31
		erm_C__3_M17990	tetracycline resistance	*Staphylococcus*	20	-	-
		erm_C__13_M13761	tetracycline resistance	*Staphylococcus*	25	-	-
		fusB_1_AM292600	fusidic acid	*Staphylococcus*	100	epidermis	40
		mecA_10_AB512767	β-lactamase	*Staphylococcus*	71	haemolyticus	23
		mecA_15_AB505628	β-lactamase	*Staphylococcus*	71	haemolyticus	23
		mph_C__2_AF167161	tetracycline resistance	*Staphylococcus*	75	haemolyticus	35
		tet_K__4_U38428	tetracycline resistance	*Staphylococcus*	100	-	-
		tet(K)_5_J01764	tetracycline resistance	*Staphylococcus*	100	-	-
P75	2nd week	mph(C)_2_AF167161	Macrolide antibiotic	*Staphylococcus*	100	-	-
		mecA_15_AB505628	β-lactamase	-	-	-	-
		mecA_10_AB512767	β-lactamase	-	-	-	-

Table 13 AMR genes detected for non-probiotic cohort using GROOT

Infant	Time point	GENE	AMR Group	ASSOCIATED TAXA			
				top genus	% classified	top species	% classified
M26	1st week	blaSHV-11	β-lactamase	Klebsiella	67	pneumoniae	40
M26	2nd week	blaSHV-11	β-lactamase	Klebsiella	77	pneumoniae	59
M26	3rd week	blaSHV-11	β-lactamase	Klebsiella	78	pneumoniae	66
		blaSHV-40	β-lactamase	Klebsiella	78	pneumoniae	65
		blaSHV-79	β-lactamase	Klebsiella	77	pneumoniae	65
M38	1st week	aadA1	Aminoglycoside	Escherichia	97	coli	54
		mphA	Macrolide antibiotic	Escherichia	95	coli	56
		sul2	Sulfonamide antibiotic	Escherichia	100	coli	67
		dfrA1	Trimethoprim resistance	Escherichia	97	coli	58
		blaTEM-95	β-lactamase	Escherichia	95	coli	56
		tetB	Tetracycline resistance	Escherichia	90	coli	53
M38	2nd week	aadA1	Aminoglycoside	Escherichia	97	coli	54
		blaTEM-95	β-lactamase	Escherichia	95	coli	56
		drfA1	Trimethoprim resistance	Escherichia	97	coli	58
		mphA	Macrolide antibiotic	Escherichia	95	coli	57
		sul2	Trimethoprim resistance	Escherichia	100	coli	75
		tetB	Tetracycline resistance	Escherichia	90	coli	54
M39	1st week	aac6-aph2	Tetracycline resistance	Staphylococcus	70	haemolyticus	64
		ermC	Tetracycline resistance	Staphylococcus	50	haemolyticus	36
		mecA	β-lactamase	Staphylococcus	66	haemolyticus	57
		mphC	Macrolide antibiotic	Staphylococcus	76	haemolyticus	72
		tetM	Tetracycline resistance	Enterococcus	98	faecalis	85
M7	1st week	mecA	β-lactamase				
		tetW	Tetracycline resistance	Bifidobacterium	100	breve	35
M7	2nd week	blaOXY-1	β-lactamase	Klebsiella	94	LTGPAF-6F	14
		tetW	Tetracycline resistance	Bifidobacterium	95	breve	34
M59	1st week	aac6-aph2	Tetracycline resistance	Staphylococcus	42	epidermis	18
		blaZ-36	β-lactamase	Staphylococcus	97	epidermis	18
		fusB	fusidic acid				
M59	2nd week	aac6-aph2	Aminoglycoside				
		blaOXY-6	beta lactamase	Klebsiella	33	pneumoniae	11
		ermA	Lincosamide resistance	Streptococcus	88	anginosus	58
		tetO	Tetracycline resistance	Streptococcus	86	anginosus	38
		blaTEM-95	β-lactamase	Escherichia	62	coli	51
		lnuB	Lincosamide resistance	Enterococcus	98	faecalis	49
		str1	Streptomycin resistance	Enterococcus	98	faecalis	49
		tetB	Tetracycline resistance	Escherichia	62	coli	57
		dfrG	Trimethoprim resistance	Enterococcus	99	faecalis	99
		cat5	chloramphenicol resistance	Enterococcus	98	faecalis	49
		ermB	Macrolide antibiotic	Enterococcus	98	faecalis	50
M103	2nd week	ermB	Macrolide antibiotic	Enterococcus	98	faecalis	51
		str1	Streptomycin resistance	Enterococcus	98	faecalis	49
		aph3-III	Aminoglycoside	Enterococcus	98	faecalis	50
		ant6-Ia	Aminoglycoside	Enterococcus	98	faecalis	51
		dfrG	Trimethoprim resistance	Enterococcus	98	faecalis	96
		tetB	Tetracycline resistance	Enterococcus	62	faecalis	57
		blaTEM-95	β-lactamase	Escherichia	62	coli	52

Validation of the GROOT pipeline

To test the GROOT pipeline, an infant heavily colonised with *Klebsiella* was selected (infant M26); the gut microbiome of this infant was heavily populated with *K. pneumoniae* at 7 and 14 days (Figure 40a-b), while at 18 days (Figure 40c), this infant presented with three different populations of *Klebsiella* (*K. pneumoniae,*

K. oxytoca and *K. unclassified*). When running GROOT analysis on this infant, week 1 and week 2 associated *K. pneumoniae* encoded the blaSHV-11 allele, and at week 3, alleles blaSHV-11 blaSHV-40 and blaSHV-79 were present (Table 13). The difference between these three alleles is only five nucleotides, and notably blaSHV-40 has been shown to confer extended spectrum β-lactamase activity, whereas blaSHV-11 and blaSHV-79 have not[207].

To validate these results, five isolates of *K. pneumoniae* were isolated from these samples, and their antibiotic resistance phenotypes studied. The Minimal Inhibitory concentrations (MICs) against a common first line β-lactam antibiotic used in NICUs (benzylpenicillin), and meropenem, which is used in NICUs on suspicion of infection with Enterobacteriaceae with extended spectrum β-lactamase activity[208] were obtained (Table 14). The MIC values for the five *K. pneumoniae* isolates against benzylpenicillin (30,000 mg/L) indicated this antibiotic is inefficient in killing, which was expected as all these isolates carried the blaSHV-11 allele. MIC values using meropenem were also calculated and results were all above the epidemiological cut off set up by Eucast (http://www.eucast.org/), which suggested all isolates may display extended β-lactamase activity.

Table 14 MICs of *Klebsiella pneumoniae* isolates

	Isolate #1 (7 days) mg/L	**Isolate #2 (7 days) mg/L**	**Isolate #3 (18 days) mg/L**	**Isolate #4 (18 days) mg/L**	**Isolate #5 (18 days) mg/L**
Benzylpenicillin	**30,000**	**30,000**	**30,000**	**30,000**	**30,000**
Meropenem	**0.601**	**0.601**	**0.601**	**0.601**	**0.305**

Isolation and validation of AMR strains detected by the GROOT pipeline

To further validate the GROOT pipeline it is necessary to phenotypically test other potential AMR species/strains detected by GROOT in the study samples. *Enterococcus*, *Enterobacter*, *Escherichia*, and *Staphylococcus* are commonly found in the gut microbiome of VLBW infants. A number of these species such as *Staphylococcus aureus*, *Klebsiella pneumoniae*, and *Enterococcus faecium* are classified as 'ESKAPE pathogens', which have become a significant burden in hospitals, due to multi-drug resistance[209].

In preparation for this analysis I performed bacterial isolations from a representative subset of samples included in this study. Table 15 presents a summary of all the bacterial isolates I obtained and its corresponding identification using sequencing of their full 16S rRNA gene. Further work characterising the genomes and AMR gene carriage of these isolates, using whole genome sequencing approaches, is required to verify the assignment of AMR genes using GROOT pipeline and metagenomic shotgun data (Tables 12 and 13).

Table 15 Bacterial isolates from a subset of VLBW infants included in this study and characterisation using their 16S rRNA gene sequence

Baby ID	Time point of study	Isolate #	Blast result (16S RNA gene sequence)
P74C	6	1	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
	6	2	*Escherichia fergusonii* strain ATCC 35469 16S ribosomal RNA, complete sequence
	6	3	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
	6	4	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
	6	5	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
	6	6	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
	6	7	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
P74G	23	1	*Enterococcus faecalis strain* NBRC 100480 16S ribosomal RNA gene, partial sequence
	23	3	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	23	4	*Shigella sonnei* strain CECT 4887 16S ribosomal RNA gene, partial sequence
	23	5	*Escherichia fergusonii* strain ATCC 35469 16S ribosomal RNA gene, partial sequence
	23	6	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
	23	7	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
	23	1 Bif	*Bifidobacterium breve* strain DSM 20213 16S ribosomal RNA gene, partial sequence
	23	8	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
	23	9	*Staphylococcus epidermidis* strain 21.1 16S ribosomal RNA gene, partial sequence
P75B	5	1	*Bifidobacterium breve* strain DSM 20213 16S ribosomal RNA gene, partial sequence
	5	7	*Staphylococcus capitis* strain ATCC 27840 16S ribosomal RNA gene, partial sequence
	5	9	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
P75F	24	5	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	24	2	*Escherichia fergusonii* strain ATCC 35469 16S ribosomal RNA gene, partial sequence
	24	6	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	24	7	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	24	8	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	24	2	*Bifidobacterium bifidum* strain PRI 1 chromosome, complete genome
	20	4	*Enterobacter xiangfangensis* strain 10-17 16S ribosomal RNA gene, partial sequence
P77L	20	12	*Staphylococcus haemolyticus* strain JCM 2416 16S ribosomal RNA gene, partial sequence
	20	2	*Enterococcus faecalis* strain NBRC 100480 16S ribosomal RNA gene, partial sequence
	20	5	*Enterococcus faecalis* strain NBRC 100480 16S ribosomal RNA gene, partial sequence
	20	2	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
	20	1	*Enterobacter cloacae* strain ATCC 13047 16S ribosomal RNA gene, partial sequence
	20	7	*Enterobacter cloacae* strain ATCC 13047 16S ribosomal RNA gene, partial sequence
P77L	20	3	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
P62J	18	1	*Bifidobacterium bifidum* strain 1579 16S ribosomal RNA gene, partial sequence
	18	2	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
	18	3 Bif	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
	18	3	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	18	4	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
	18	A	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	18	6	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	18	5	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	18	B	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
P62C	7	A	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	7	2	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence

	7	3	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
	7	C	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	7	D	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	7	B	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
P60L	17	1 Bif	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
	17	2	*Enterobacter xiangfangensis* strain 10-17 16S ribosomal RNA gene, partial sequence
	17	1	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	17	4	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	17	5	*Staphylococcus aureus* strain S33 R 16S ribosomal RNA, complete sequence
	17	6	*Staphylococcus aureus* strain S33 R 16S ribosomal RNA, complete sequence
	17	7	*Staphylococcus aureus* strain S33 R 16S ribosomal RNA, complete sequence
	17	8	*Enterococcus faecalis* strain NBRC 100480 16S ribosomal RNA gene, partial sequence
	17	9	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
P60F	5	1	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
	5	2	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
	5	3	*Bifidobacterium longum subsp. suillum* strain Su 851 16S ribosomal RNA, partial sequence
M26 7	7	1	*Klebsiella pneumoniae* strain QLR-3 16S ribosomal RNA gene, partial sequence
	7	2	*Klebsiella pneumoniae* strain QLR-3 16S ribosomal RNA gene, partial sequence
M26 18	18	1	*Klebsiella pneumoniae* strain QLR-3 16S ribosomal RNA gene, partial sequence
	18	2	*Klebsiella pneumoniae* strain B-3-4 16S ribosomal RNA gene, partial sequence
	18	3	*Klebsiella pneumoniae* strain B-3-4 16S ribosomal RNA gene, partial sequence
	18	4	Klebsiella pneumoniae strain B-3-4 16S ribosomal RNA gene, partial sequence
M7 8	8	1	*Enterococcus sp.* H184 16S ribosomal RNA gene, partial sequence
	8	2	*Escherichia coli strain* noha905 16S ribosomal RNA gene, partial sequence
	8	2 bif	*Bifidobacterium breve* strain DSM 20213 16S ribosomal RNA gene, partial sequence
	8	4	*Enterococcus faecalis* strain NBRC 100480 16S ribosomal RNA gene, partial sequence
	8	3	*Enterococcus sp.* H185 16S ribosomal RNA gene, partial sequence
M7 21	21	1	*Klebsiella michiganensis* strain W14 16S ribosomal RNA gene, partial sequence
	21	2	*Klebsiella sp.* SI-AL-1B 16S ribosomal RNA gene, partial sequence
	21	3	*Klebsiella michiganensis* strain W14 16S ribosomal RNA gene, partial sequence
	21	7	*Bifidobacterium breve* strain DSM 20213 16S ribosomal RNA gene, partial sequence
M16 18	18	1	*Klebsiella pneumoniae* strain DSM 30104 16S ribosomal RNA gene, partial sequence
	18	2	*Klebsiella pneumoniae* strain DSM 30104 16S ribosomal RNA gene, partial sequence
	18	3	*Klebsiella pneumoniae* strain DSM 30104 16S ribosomal RNA gene, partial sequence
P48C	8	1	*Bifidobacterium animalis subsp. lactis* strain YIT 4121 16S ribosomal RNA gene, partial sequence
	8	2	*Bifidobacterium bifidum* strain NBRC 100015 16S ribosomal RNA gene, partial sequence
	8	3	*Bifidobacterium bifidum* strain PRI 1 chromosome, complete genome
M59 22	22	1	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	22	1	*Bifidobacterium animalis subsp. lactis* strain YIT 4121 16S ribosomal RNA gene, partial sequence
	22	4	*Bifidobacterium bifidum* strain PRI 1 chromosome, complete genome
	22	2	*Bifidobacterium animalis subsp. lactis* strain YIT 4121 16S ribosomal RNA gene, partial sequence
	22	8	*Klebsiella oxytoca* strain JCM1665 16S ribosomal RNA gene, partial sequence
	22	5	*Enterobacter cloacae* strain ATCC 13047 16S ribosomal RNA gene, partial sequence
	22	4	*Enterococcus faecalis* strain NBRC 100480 16S ribosomal RNA gene, partial sequence
	22	7	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	22	3	*Enterococcus saigonensis* strain VE80 16S ribosomal RNA, partial sequence
	22	2	*Enterococcus faecalis* strain NBRC 100480 16S ribosomal RNA gene, partial sequence
M38 20	22	2	*Bifidobacterium breve* strain DRBB29 chromosome, complete genome
M39 22	22	2	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence
	22	3	*Staphylococcus epidermidis* strain NBRC 100911 16S ribosomal RNA gene, partial sequence

Discussion

Premature infants, and in particular VLBW infants (<1.5 kg) are often administered empiric antibiotic treatment during the first week of life because they are extremely vulnerable to bacterial infections. Approximately 78%-87% of VLBW infants receive antibiotics within the first three days of life[210], however, antibiotic treatment may affect the early stages of the developing host-microbiome ecoystem[49], and consequently impact on long-term health outcomes. Some Neonatal Intensive Care Units (NICUs) have started to administer probiotic supplementation to modify the preterm gut microbiome, as studies have shown this reduces the incidence of bacterial sepsis and necrotising enterocolitis[138, 168]. To date, very few studies have evaluated the short-term effects of antibiotics on the VLBW infant gut microbiome, side by side with the effect of probiotic supplementation on the overall antimicrobial resistance carriage.

This study represents one of the few longitudinal studies including VLBW infants (34 recruited and 95 stool samples), where antibiotic regimes were matched from two different study cohorts (with and without probiotic supplementation). All the infants in this study were born before the 33st week of gestation, weighed less than <1.5 kg, were mostly breastfed and born by C-section. One of the novelties of this study is that it comprises antibiotic-naïve premature infants who did not receive any antibiotic treatment (n=16 infants). Moreover, this is the first clinical premature study to also include this number of VLBW infants receiving no antibiotic treatment. Previous published studies have included fewer premature infants with no exposure to antibiotics (only two (1)[204], and three (2)[211]) and crucially antibiotic regimes were not matched. Because of the difficulties in obtaining antibiotic-naïve VLBW infants, previously published research studies have used healthy term infants or late-premature infants (born at 36 weeks

gestation)[212] whose gut maturation and NICU residency is very different to that of an VLBW infant.

To understand the short-term effects associated with empiric antibiotic treatment and whether this contributes to enhance gut antimicrobial resistance carriage, whole-metagenome shotgun sequencing was performed on all study samples. By week 1 of the study, there was an increase of *Escherichia coli* and *Staphylococcus spp* in both study cohorts and a proliferation of *Klebsiella spp* and *Enterococcus faecalis* in the non-probiotic cohort. Notably, antibiotic treatment is known to increase Enterobacteriaceae [212] (*e.g. Escherichia* and *Klebsiella*) abundance and enhance the presence of multidrug-resistance bacteria[213]. As such, the results from this study concords with previous published studies, where Enterobacteriaceae and Enterococcaceae populations have been described as predominant populations in the preterm microbiome during the first month of life. *Klebsiella spp, Escherichia coli, Enterococcus faecalis* and *Staphylococcus spp* are common sepsis pathogens in VLBW infants[214-216] and their presence in the preterm gut could predispose the at-risk premature infant to a bacterial infection.

Interestingly, the results from the microbial functional analysis (Figure 43) did not show great differences among study individuals nor between study cohorts. This finding suggest that the main microbial functions remain conserved across hosts, despite the taxonomic perturbation observed by the antibiotic treatment. Enteric bacterial pathogens are known to impact three major physiological functions of the intestinal epithelium via various specific virulence factors: (i) disruption of the tight junction barrier (*e.g.* secreting proteases), (ii) dysregulation of intestinal ion transporters, and (iii) activation of the inflammatory response in the gastrointestinal mucosa[217, 218]. I hypothesize that greater variation in microbial functional pathways (and identification of virulence factors) will be found if the functional pathways were studied at a deeper level.

This study showed that alongside antibiotic treatment influencing the preterm gut microbiome, the NICU environment also plays an important role, especially in premature infants who did not receive probiotic supplementation. It is well known that hospital environments represent a reservoir of potential pathogens, in particular those who possess multidrug resistance to antibiotics. *Staphylococcus aureus*, Enterobacteriaceae (*e.g. Escherichia coli*, *Klebsiella pneumoniae*), *Acinetobacter baumannii* and *Pseudomonas aeruginosa* are some of the most common hospital-related pathogens characterised by having potential multidrug resistance mechanisms[219, 220]. All non-supplemented premature infants presented elevated levels of pathobionts (*e.g. Klebsiella spp, Escherichia coli* and *Staphylococcus spp*) in their gut until the very end of the study period (week 3) regardless of whether they received antibiotic treatment or not. This observation supports the notion that colonisation of the preterm gut by pathobionts is not only related to antibiotic treatment; the NICU environment also influences the preterm microbiome. In contrast, in the case of probiotic supplemented infants there was a predominance of *Bifidobacterium spp.* (Figure 41) in infants treated with and without antibiotic. Similar results were found in a preterm research study[194] using the same probiotic supplementation; which showed a 64% relative abundance of *Bifidobacterium* from week 1 of the study. These results suggest that daily probiotic supplementation may induce colonisation resistance against hospital-related pathogens and may facilitate a quicker recovery of the preterm gut microbiome after antibiotic treatment. Colonisation resistance of *Bifidobacterium* against common intestinal pathogens has been reported previously in *in-vitro* studies using different mechanisms: (i) synthesising antimicrobial products which impaired adhesion of *Clostridioides difficile* to enterocytes [221]or (ii) producing acetate which improved the integrity of the epithelial barrier and reduced translocation of *E. coli* toxins[176].

Surprisingly, the total AMR content between infants treated with antibiotics and non-treated infants (probiotic and no probiotic) was not significantly different

(Table 10). However, this does correspond with the taxonomic profiles within the study cohorts, which are also similar (Figure 40 and Figure 41). This suggests that the gut antibiotic resistome may not only be established by antibiotic treatment; the colonisation of pathobionts from the NICU environment may also influence this outcome. Contrary to expectation, infants treated with empiric antibiotic treatment (penicillin and gentamicin) did not have a higher number of specific AMR genes against this treatment if week 1 is compared to week 2 and 3 (Table 11), which supports this latter statement.

When comparing the AMR content of probiotic supplemented infants and non-supplemented, the non-supplemented cohort had a higher number of AMR genes (Table 10) although the *p-values* did not show significance. A higher number of samples would be required to evaluate robustly whether the non-probiotic infants, containing higher levels of pathobionts in their gut, contained a higher AMR gene content. A similar study involving 66 premature infants was not able to find differences among the total AMR content when benchmarking probiotic and non-probiotic premature infants[194], which agrees with our findings.

Even though the bioinformatics approach using the CARD database provided insight into the overall AMR gene content, a specific bioinformatic tool which could associate specific AMR genes to specific taxa was needed. One of the most relevant questions clinically is to determine whether a pathogen or a commensal bacterium is the carrier of specific AMR genes. GROOT was used in collaboration with Dr Will Rowe, to assign antimicrobial resistance genes to specific taxa. This approach successfully managed to associate most of the AMR genes detected to potential pathogenic bacteria (*i.e*. *Staphylococcus spp, Enterococcus faecalis, Escherichia coli*, *Klebsiella pneumoniae,* and *Streptococcus anginosus*). All these potential pathogens harbored resistance genes towards the empiric antibiotic treatment used in this study (*i.e*. β-lactamases and aminoglycosides), while none of the commensal bacteria were assigned AMR genes with the exemption of only one

infant. One of the main concerns regarding the use of prolonged empirical antibiotic therapy is that it perturbs the colonisation of the beneficial bacteria[222]. This data showed that the empiric antibiotic treatment used in this study was facilitating the growth of these potential pathogens and preventing the growth of commensal bacteria which normally do not encode AMR genes.

Finally, the use of the GROOT pipeline highlighted specific information regarding the antimicrobial activity of the AMR genes detected. GROOT was able to associate an extended spectrum β-lactamase activity (blaSHV-40) to pathogenic *K. pneumoniae* from shotgun metagenomic data. This finding was validated phenotypically (Table 14). Further work will be needed to determine the applicability of this bioinformatic pipeline against other potential AMR species/strains (*e.g. Staphylococcus, Enterobacter, Enterococcus or Escherichia*) commonly found in the gut microbiome of VLBW infants. Whole-genome sequence of 96 bacterial isolates (Table 15) was prepared in order to examine whether the AMR genes detected using GROOT and metagenome data were also found in the genomes of these isolates. Phenotypic validation is also planned after this genomic analysis.

Conclusions

The taxonomic analyses indicated clear differences between study cohorts (non-probiotic and probiotic supplemented). Antibiotic treatment had a transient effect on the preterm gut microbiome, especially for VLBW infants not receiving probiotic supplementation. Probiotic supplemented VLBW infants had a higher relative abundance of *Bifidobacterium* throughout the study period, which may induce colonisation resistance against hospital-related pathogens.

The total AMR content within the two study cohorts did not significantly influence the AMR gene categories detected, nor their abundance. The use of the GROOT pipeline allowed most of the AMR genes detected to be associated with the potential pathogenic bacteria detected (*i.e. Staphylococcus spp, Enterococcus faecalis, and Escherichia coli),* while commensal bacteria did not carry AMR genes. This suggest empiric antibiotic treatment was not effective against these pathogens and may only be favoring the colonisation of these bacteria in the vulnerable preterm gut.

More personalised antibiotic regimes should be used to treat VLBW infants, and antibiotic stewardship after a microbiological diagnosis should be given consideration.

Further work

Future plans for this study include continuation of analysis for another subset of VLBW infants heavily populated with other potential pathogenic bacteria such as *Enterococcus*, *Enterobacter*, *Escherichia*, and *Staphylococcus,* and to determine if the GROOT pipeline can robustly be applied to in-depth AMR analysis of shotgun metagenomic datasets. To further validate this data, whole genome sequencing of the bacterial isolates has been planned to verify whether the AMR genes found in the shotgun data using GROOT are also encoded with the genomes.

I believe that to be able to administer more personalised antibiotic treatments will require faster microbial diagnosis methods. This idea forms the basis for the study described in the following Chapter, where MinION Nanopore technology was used to rapidly profile the preterm microbiome and resistome of premature infants suffering from sepsis.

To complement this study, I would include metabolomic analysis of the study samples. Alterations in the gut microbial populations may change the intra-community metabolic interaction, and influence host metabolic, hormonal and immune homeostasis. Administration of antibiotic therapy in young mice has been associated with substantial increases in SCFAs (*e.g.* acetate, propionate and butyrate in the caecal contents), which, when delivered in increased quantities through the blood circulation to the liver, enhanced the production of fat[223]. The timeframe of this study is relatively short, but as the taxonomic profiles from the two study cohorts were very different, I hypothesise their metabolomic profiles would be as well.

Chapter 4

Title: Rapid diagnostics of the preterm gut microbiome using MinION Nanopore technology for monitoring microbiota intervention strategies, and antibiotic resistance profiles

Abstract

MinION nanopore technology is an exciting new sequencing platform that offers the possibility of long reads that can be analysed in real time. These features, plus the compact and portable nature of this platform makes the MinION an attractive sequencing technology that can be applied to the field of microbiome monitoring and rapid clinical diagnostics. The work presented here demonstrates how MinION technology can be used to rapidly profile faecal samples from premature infants; to monitor colonisation of probiotic strains, and to profile samples obtained from infants suffering from sepsis or necrotising enterocolitis (NEC). Initially, the MinION technology was validated using a mock microbial community, and relevant clinical samples from the same infant were analysed at different time points, with benchmarking against Illumina technology. Next steps involved demonstration of how MinION technology can be used in clinical settings; utilising MinION technology to diagnose bacterial pathogens and antimicrobial resistant (AMR) profiles of premature infants suffering from NEC. Finally, a 'real-time' run was performed which involved timing all stages, from sample preparation, sequencing, to downstream real time analysis, which culminated in a <5h determination of the pathogenic bacteria and corresponding AMR profiles. Bacterial isolation of the bacterial pathogen and phenotypic antibiotic testing using Minimal Inhibitory Concentration (MIC) was performed to validate the MinION results.

Figure 44 represents a graphical abstract summarising the different stages of this work

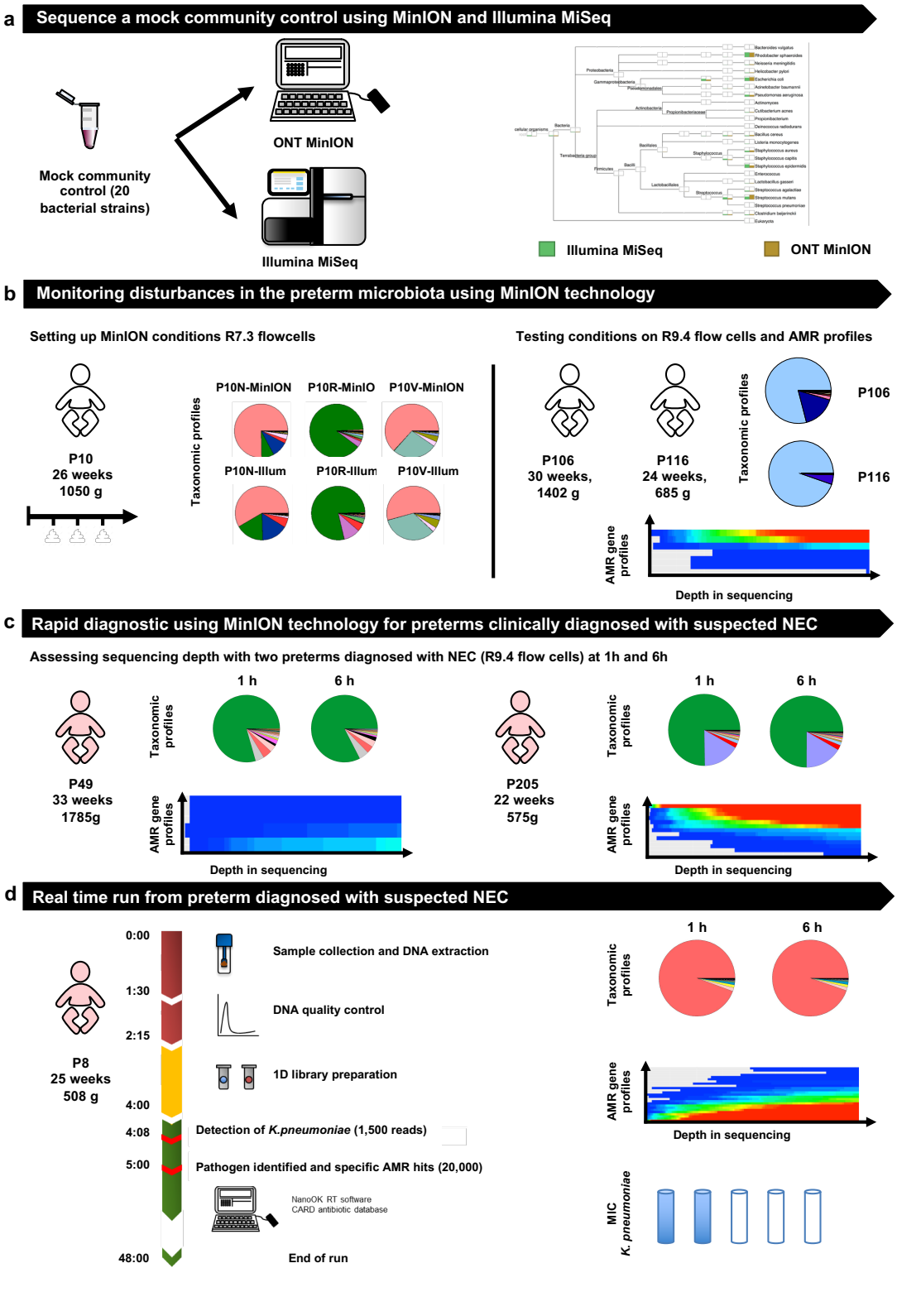


Figure 44 Graphical abstract summarising study pipeline

This study aims to use MinION technology to profile faecal samples from healthy and critically ill premature infants. The work had several stages: a) Sequencing a microbial mock community using MinION and Illumina, b) Setting up conditions for MinION technology and testing flowcells R7.3 and R9.4 c) Diagnosing taxonomic and AMR profiles from premature infants with suspected NEC d) Performing a 'real time' run with a NEC infant where all the stages of the pipeline were timed. Bacterial isolation of the bacterial pathogen and phenotypic antibiotic testing using Minimal Inhibitory Concentration (MIC) was performed to validate the MinION findings

Authors' contribution to this work

The completion of this large and multi-disciplinary study has been in close collaboration with Dr Richard Leggett at Earlham Institute. I was involved in the design of the study and took the lead on selecting samples for the study, extracting DNA, analysing and graphing the data. Dr Richard Leggett developed the bespoke NanoOK RT MinION pipeline and carried out bioinformatics analysis. This work has been accepted in Nature Microbiology and I am joint first author with Dr Richard Leggett. Specific details of authors' contributions can be found in the method section of this chapter.

Introduction

High throughput next generation DNA/RNA sequencing (NGS) platforms have revolutionised our ability to drive forward genomics research. Within the context of studying complex microbial communities, NGS has opened culture-independent and rapid profiling of samples in both environmental and clinical settings using both short-read sequencing (*e.g.* Illumina) and long-read (*e.g.* Pacific Biosciences (PacBio), and Oxford Nanopore Technologies).

Current NGS platforms are expensive, and often occupy a large footprint, which makes their usage challenging to apply in the field or clinic. Notably, Oxford Nanopore Technologies released the MinION device in 2014, a pocket-sized Nanopore sequencing platform capable of producing sequencing runs similar to Illumina and PacBio. This portable and low-cost technology has the advantage that it can be used at the point of care, and it offers real-time sequencing[224] therefore results are available in a matter of hours.

One at-risk patient cohort who can benefit from this technology are premature infants. Premature infants are at high risk of acquiring bacterial infections (*i.e.* sepsis) and life threatening diseases (*i.e.* necrotising enterocolitis (NEC)) due to

their immature immune system[225], underdeveloped gut physiology[226] and long residency in the hospital environment (see Introduction Chapter). Furthermore, these infants often receive early and extended courses of antibiotic treatment during their stay at the neonatal intensive care unit (NICU), which significantly alters their gut microbiome and promotes the growth of multidrug-resistant bacteria. In recent years, some NICUs have introduced probiotic supplementation using common early life gut colonisers (*e.g. Bifidobacterium* and *Lactobacillus*) as microbiota therapies[195, 227] to prevent premature infants acquiring gut bacterial infections. Therefore, these at-risk premature infants represent an important cohort for developing and optimising advances in microbial detection platforms. The fast turnaround time and long-read genomic data provided by MinION technology, could contribute to early patient diagnosis by identifying bacterial taxonomic profiles and antibiotic resistance genes in a timely fashion.

This work demonstrates the 'real-world' utility of the MinION platform in the analysis of whole genome shotgun of metagenomic samples (*i.e*. faeces) from at risk premature infants. This study involved the development of a bespoke software (*i.e*. NanoOK RT) able to provide bacterial metataxonomic profiles and associate its antimicrobial resistance genes. Firstly, a 20 species human microbiome mock community was used to prove nanopore metagenomic data can be classified reliably and rapidly. Secondly using DNA from 3 different time points of one patient spanning 2 clinical interventions, the study captured the complete diversity of the immature gut microbiome and the dynamics it undergoes over time. Sequencing of these samples was paired with Illumina platform, and demonstrated identical taxonomic profiles and no significant bias in the nanopore sequencing. Thirdly, the MinION pipeline developed here was tested profiling samples from two healthy premature infants and two critically ill premature infants diagnosed with NEC. Finally, a 'real-time' run was performed which involved timing all stages of the pipeline and culminated in a less than 5h determination of the pathogenic bacteria and corresponding AMR profiles. Bacterial isolation of the

bacterial pathogen and phenotypic antibiotic testing were used to validate the findings. Results demonstrated MinION sequencers offer the ability to progress from clinical samples to tailored patient antimicrobial treatment in a few hours.

Hypothesis and aims

Hypothesis: The MinION nanopore sequencer will be able to profile rapidly faecal samples from premature infants to obtain bacterial metataxonomic profiles and characterise their antibiotic resistance profiles.

The study involved four aims/sub-sections:

i) Sequencing a bacterial mock community to evaluate the accuracy of MinION technology for metagenomic sample profiling and benchmarking results with Illumina sequencing.

ii) Performing a longitudinal study using samples from the same patient, to test whether MinION technology can monitor microbial disturbances (*e.g.* antibiotic treatment, or probiotic supplementation) in the preterm gut microbiome and resistome.

iii) Using a bespoke software (NanoOK RT pipeline developed by Dr Richard Leggett) to rapidly profile faeces from critically ill premature infants, to test if turnaround time is comparable to other rapid molecular tests used in the clinic.

iv) Whole genome sequencing and phenotypic testing of one of the bacterial pathogens detected in this study, to validate the metataxonomic and resistome profiles obtained using MinION technology and the NanoOK RT pipeline.

Methods

Genomic DNA from mock community

Genomic DNA from a microbial mock community was used (HM-277D, BEI Resources, Manassas, VA). This mock community contained a mixture of 20 bacterial strains. Details of the strains present in the mock community along with their concentrations are indicated in the table below.

Table 16 Bacterial strains present in the microbial mock community (HM-277D, BEI Resources)

Organism	NCBI Reference Sequence	Number of operons	Concentration
Acinetobacter baumannii, strain 5377	NC_009085	100,000	82 pg/ul
Actinomyces odontolyticus, strain 1A.21	NZ_AAYI02000000	10,000	10 pg/ul
Bacillus cereus, strain NRS 248	NC_003909	1,000,000	450 pg/ul
Bacteroides vulgatus, strain NCTC 11154	NC_009614	10,000	7.6 pg/ul
Clostridium beijerinckii, strain NCIMB 8052	NC_009617	1,000,000	440 pg/ul
Deinococcus radiodurans, strain R1 (smooth)	NC_001263	10,000	10 pg/ul
Enterococcus faecalis, strain OG1RF	NC_017316	10,000,000	7 pg/ul
Escherichia coli, strain K12, substrain MG1655	NC_000913	10,000,000	6.8 ng/ul
Helicobacter pylori, strain 26695	NC_000915	100,000	86 pg/ul
Lactobacillus gasseri, strain 63 AM	NC_008530	100,000	32 pg/ul
Listeria monocytogenes, strain EGDe	NC_003210	100,000	50 pg/ul
Neisseria meningitides, strain MC58	NC_003112	100,000	58 pg/ul
Propionibacterium acnes, strain KPA171202	NC_006085	100,000	88 pg/ul
Pseudomonas aeruginosa, strain PAO1-LAC	NC_002516	1,000,000	1.6 ng/ul
Rhodobacter sphaeroides, strain ATH 2.4.1	NC_007493	10,000,000	14 ng/ul
Staphylococcus aureus, strain TCH1516	NC_010079	1,000,000	590 pg/ul
Staphylococcus epidermidis, FDA strain PCI 1200	NC_004461	10,000,000	5.1 ng/ul
Streptococcus agalactiae, strain 2603 V/R	NC_004116	1,000,000	32 pg/ul
Streptococcus mutans, strain UA159	NC_004350	10,000,000	4.1 ng/ul
Streptococcus pneumoniae, strain TIGR4	NC_003028	10,000	5.5 pg/ul

Illumina sequencing of mock community

DNA was sheared (600ng) in a 60 µl volume on a Covaris S2 (Covaris, Massachusetts, USA) for 1 cycle of 40 s with a duty cycle of 5%, cycles per burst of 200 and intensity of 3. Illumina paired end libraries were constructed with inserts spanning from 600 bp to >1000 bp. Sheared DNA was then end-repaired using the NEB End Repair Module (NEB, Hitchin, UK), size selected with a 0.58 x Hi Prep bead clean-up (GC Biotech, The Netherlands) and followed by A tailing using the NEB A tailing module (NEB) and ligation of adapters using the NEB Blunt/TA Ligase Master Mix (NEB). Three 1x bead clean-ups were then undertaken to remove all traces of adapter dimers. DNA was then assessed by running an Agilent BioAnalyser High Sensitivity chip and quantified using the Kappa qPCR Illumina quantification kit. Finally, based on the qPCR results 9 pM DNA was sequenced on an Illumina MiSeq with 300 bp paired reads. This work was done in collaboration with Dr Darren Heavens at Earlham Institute.

MinION sequencing of mock community

A total of 1 µg of DNA was fragmented in a 46 µl volume in a Covaris G-tube (Covaris, Massachusetts, USA) at 6,000 rpm in an Eppendorf centrifuge 5417. MinION 2D libraries were prepared targeting inserts bigger than 8 kbp. Sheared DNA was then subjected to a repair step using the NEB FFPE repair mix (NEB, Hitchin, UK) and purified with a 1x Hi Prep bead clean-up (GC Biotech, Alphen aan den Rijn, The Netherlands). The repaired DNA was then end-repaired and A-tailed using the NEBNext Ultra II End Repair and A-Tailing Module (NEB), purified with a 1x Hi Prep bead clean-up and then the AMX and HPA MinION adapters ligated using the NEB Blunt/TA Ligase Master Mix (NEB). An HP tether was then added and incubated for 10 min at room temperature followed by 10 min room temperature incubation with an equal volume of pre-washed MyOne C1 beads (Thermo Fisher, Cambridge, UK). The library bound beads were washed

twice with bead binding buffer (ONT) before the final library eluted via a 10 min incubation at 37 ºC in the presence of the MinION Elution Buffer. The final library was then mixed with running buffer, fuel mix and nuclease free water and loaded onto an R7.3 flowcell following the manufacturer's instructions. Sequencing data was collected for 48 h. Dr Darren Heavens performed the laboratory analysis, Dr Richard Leggett performed the bioinformatic analysis and I prepared the final figure.

Clinical samples

Ethical approval and sample collection

Subject recruitment for this study was approved by The Ethics Committee of the Faculty of Medical and Health Sciences in the University of East Anglia (Norwich, UK). Faeces collection protocol was in accordance with the Norwich Research Park (NRP) Biorepository Human Tissue Act 2004 (HTA, license number 11208). Recruitment of infants admitted to the Neonatal Intensive Care Unit (NICU) of the Norfolk and Norwich University Hospital (NNUH, Norwich, UK) was done by doctors or nurses with informed and written consent obtained from parents. All infants recruited in this study received oral probiotic supplementation containing *Bifidobacterium bifidum* and *Lactobacillus acidophilus* (*i.e*. Infloran®, Desma Healthcare, Switzerland) strains with a daily dose of 2×10^9 of each species. Collection of faecal samples was carried out by researchers and stored at -80 °C prior to DNA extraction.

Clinical details of all the premature infants recruited for this study are detailed in the diagrams below. Figure 45 contains clinical details for the healthy premature infants recruited and Figure 46 information about premature infants diagnosed with suspected NEC.

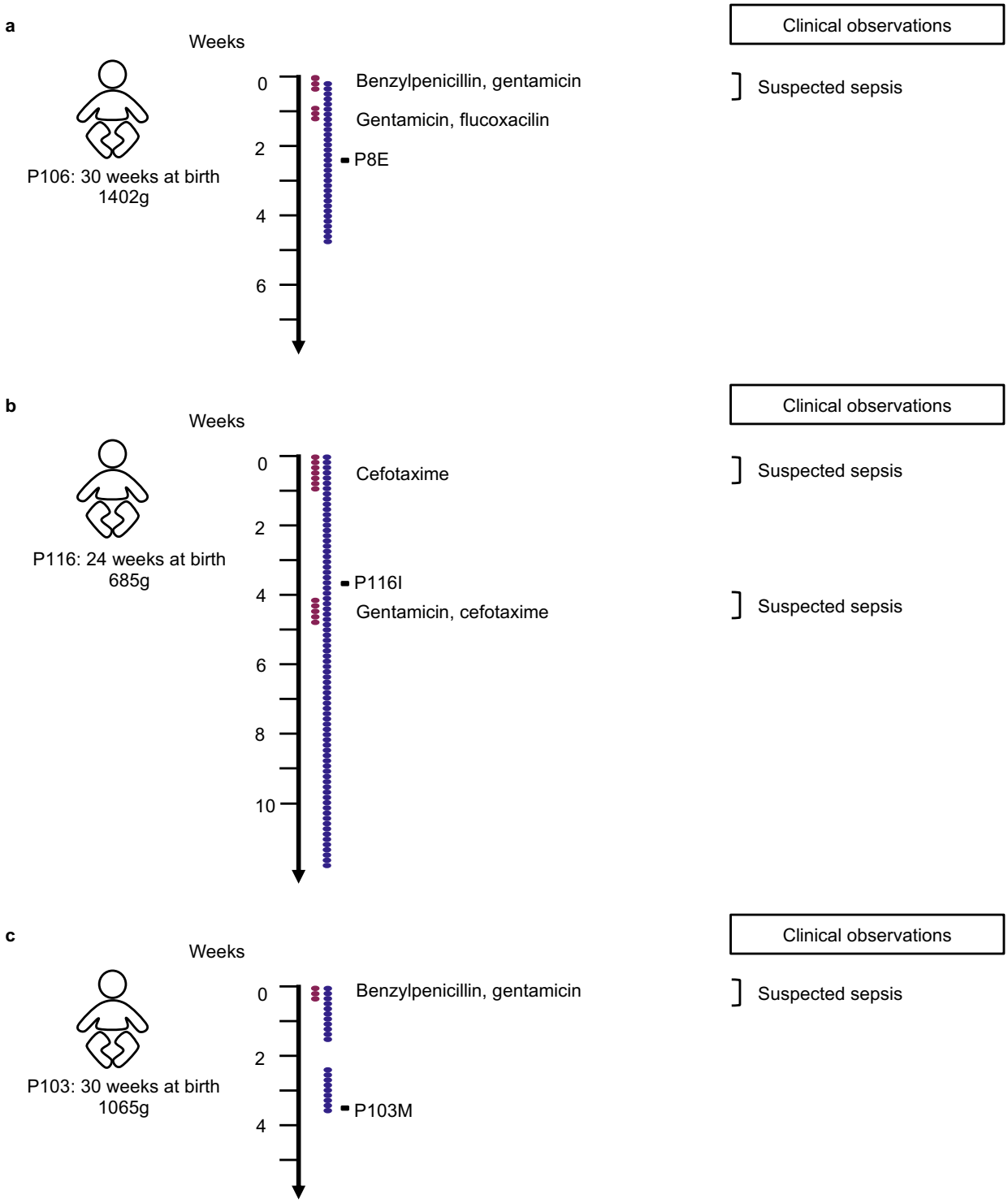


Figure 45 Timeline diagrams for healthy premature infants (P106, P116, P103)

Timeline diagrams indicating time points of faecal sample collection, duration of antibiotic, probiotic treatment, and relevant clinical observations. The timeline diagrams are divided in weeks and dots represent days within the scale. Blue dots indicate days of probiotic treatment, red dots antibiotic treatment, and black squares time points for sample collection. (a) timeline diagram for premature infant P106, (b) timeline diagram for premature infant P116 and (c) timeline diagram for premature infant P103.

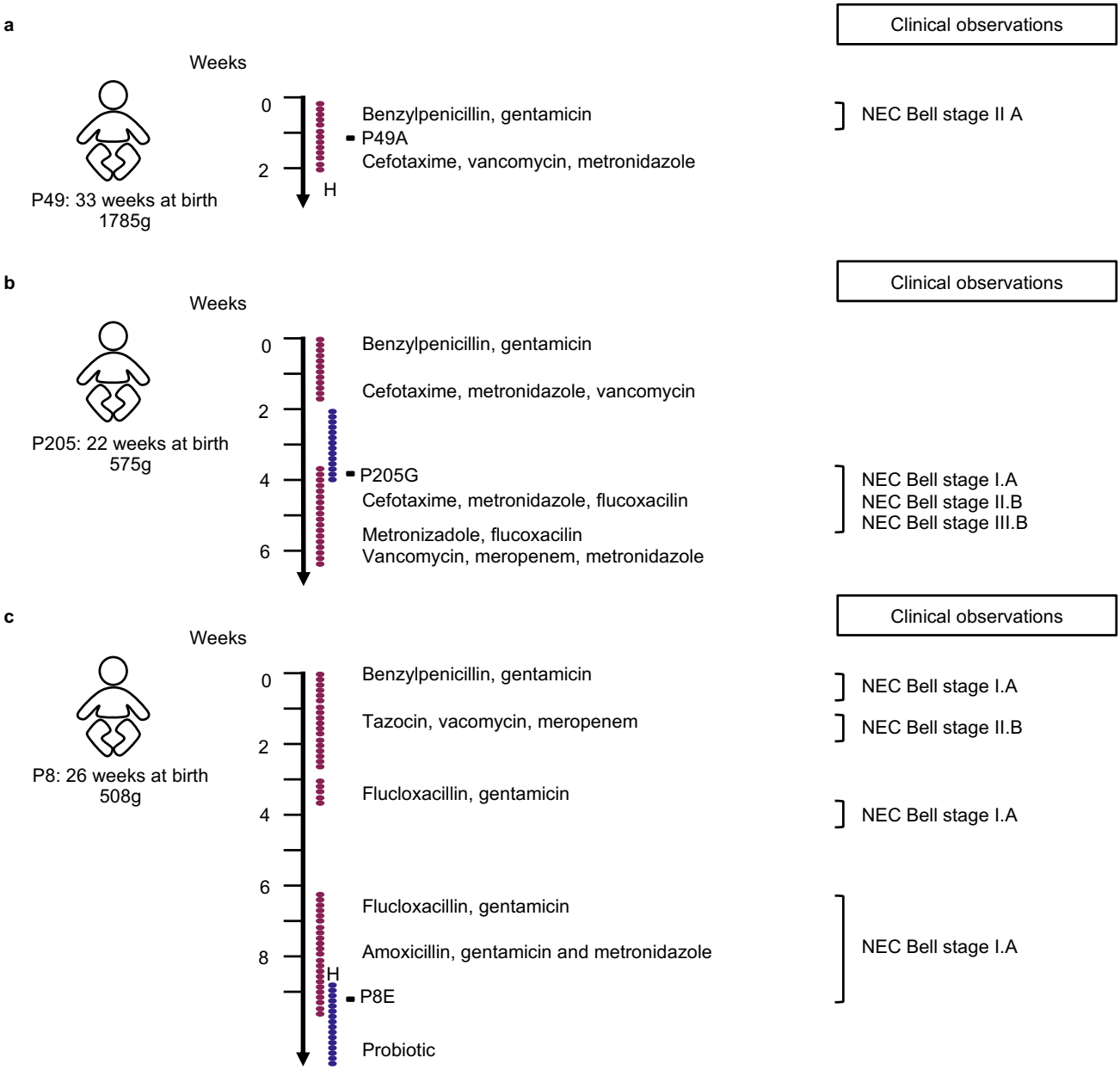


Figure 46 Timeline diagrams for premature infants diagnosed with NEC (P49, P205 and P8)

Timeline diagrams indicating time points of faecal sample collection, duration of antibiotic, probiotic treatment, and relevant clinical observations. The timeline diagrams are divided in weeks and dots represent days within the scale. Blue dots indicate days of probiotic treatment, red dots antibiotic treatment, black squares time points for sample collection, letter H transfer of the premature infant to another hospital. Clinical observations highlight Bell stages of necrotising enterocolitis (Gregory, DeForge et al. 2011) commonly used by clinicians to assign the severity of this disease. (a) timeline diagram for premature infant P49, (b) timeline diagram for premature infant P205 and (c) timeline diagram for premature infant P8.

DNA extraction of faecal samples

Bacterial DNA was extracted from 100-150 mg of faecal material using the FastDNA Spin Kit for Soil (MP Biomedical, Santa Ana) following the manufacturer's instructions but extending the bead-beating step to 1 min and eluting the DNA with 55 °C DES. The purity and concentration of the DNA was assessed using a NanoDrop 2000c Spectrophotometer and Qubit® 2.0 fluorometer. DNA concentrations higher than 25 ng/µl were considered acceptable to proceed further analysis. I performed this lab work and analysis.

MinION shotgun library preparation and flowcells

MinION 1D and MinION 2D libraries were prepared to assess if the clinical samples used in this study could be equally analysed. Both libraries differ in their preparation times. The preparation time for MinION 1D libraries is 30 min. MinION 1D library were constructed incubating 200 ng of DNA with 2.5 µl FRM for 1 min at 30 °C, then 1 min at 75 °C followed by the addition of 1 µl RAD and 0.2 µl NEB Blunt/TA Ligase Master Mix (NEB), and room temperature incubation for 5 min. The final library DNA was mixed with running buffer containing fuel mix, library loading beads and nuclease free water.

MinION 2D libraries, had an estimated preparation time of 2 h. This library was prepared following the same steps outlined for the mock community library preparation (for details see above). The final library DNA was eluted with same reagents specified above for MinION 1D libraries.

Different MinION cells were used along the study. The MinION technology evolves very fast, and throughout the duration of this study R7.3 flowcell was used with an estimated error rate of 15%, and moved to R9.4 with an error rate less than 7%[228].

The table below details samples, MinION libraries and flowcells types used in this study.

Table 17 Summary of samples used in this study detailing flowcell, sequencing kit and flow cell used

Run	Sample	Flow cell	Seq'ng kit	Library type
1	Mock	R7.3 (MAP103)	MAP006	2D
2	P10N	R7.3 (MAP103)	MAP006	2D
3	P10R (1)	R7.3 (MAP103)	MAP006	2D
4	P10R (2)	R7.3 (MAP103)	MAP006	2D
5	P10V (1)	R7.3 (MAP103)	MAP006	2D
6	P10V (2)	R7.3 (MAP103)	MAP006	2D
7	P8	R9.4 (MIN106)	LSK208	2D
8	P8	R9.4 (MIN106)	RAD002	1D Rapid
9	P8	R9.5 (MIN107)	LSK108	1D Ligation
10	P49A	R9.5 (MIN107)	LSK208	1D Ligation
11	P205G	R9.5 (MIN107)	LSK208	1D Ligation
12	P106I	R9.5 (MIN107)	LSK208	1D Ligation
13	P116I	R9.5 (MIN107)	LSK208	1D Ligation
14	P103M	R9.5 (MIN107)	RAD004	1D Rapid

Mr Darren Heavens performed the MinION libraries and Dr Richard Leggett performed the bioinformatic analysis. I prepared the final figures of this analysis.

Illumina HiSeq 2500 shotgun library preparation

Illumina libraries for samples (P10N, P10R and P10V) were prepared using TruSeq Nano DNA Library Prep Kit following manufacturer's instructions and sequenced using the HiSeq Illumina 2500 machine with 150 bp paired end reads. The Illumina library for P8 was prepared following the same protocol used for the mock community (see above) and run at 9 pM on an Illumina MiSeq with a 2 x 300 bp read metric. This work was done in collaboration with Mr Darren Heavens who

performed the laboratory analysis, Dr Richard Leggett performed the bioinformatic analysis and I prepared the final figures of this analysis.

16S rRNA gene library preparation and bioinformatics analysis

Libraries were constructed using bacterial DNA from samples P10N and P10V normalised to 5 ng ml-1. V4 region of the 16S rRNA gene was amplified using the following primers, 5' AAT GAT ACG GCG ACC GAG ATC TAC A and, 5' CAA GCA GAA GAC GGC ATA CGA GAT AAC T. PCR conditions used for this amplification were: 1 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 using a 96 well Thermal Cycler PCR machine. Sequencing of the 16S RNA gene libraries was performed on the Illumina MiSeq platform with 250 bp paired end reads.

Quality control was subjected to raw reads obtained using FASTX-Toolkit keeping a minimum quality threshold of 33 for at least 50% of the bases. Reads that passed the threshold were aligned against SILVA database (version: SILVA_119_SSURef_tax_silva) using BLASTN (ncbi-blast-2.2.25+; Max e-value 10e-3) separately for both pairs. After performing the BLASTN alignment, all output files were imported and annotated using the paired-end protocol of MEGAN. This work was performed at Sanger Institute in Cambridge, I performed the bioinformatic analysis and prepared the final figure.

Time series study for infant P10

Illumina and MinION sequencing raw reads for samples P10N, P10V and P10R from infant P10 were analysed using the following bioinformatics pipelines. For the Illumina samples, we removed PCR duplicates (remove_pcr_duplicates.pl, script from **https://github.com/richardmleggett/scripts**), ran Trimmomatic[61] to remove adaptors and applied a sliding window quality filter (size 4, mean quality greater than or equal to 15). A random set of 1,000,000 reads was subsampled

(subsample.pl script, **https://github.com/richardmleggett/scripts**) to represent the yield. These reads were used as the input to a blastn search of NCBI's nt database.

For the Nanopore sequencing, the analysis with the reads classified as 'pass' reads was followed (defined as 2D reads with a mean Q value >9) and performed no further pre-processing before running blastn. Taxonomic analysis was done using MEGAN6[30]. The bioinformatics pipeline was run by Dr Richard Leggett and I prepared the final figures.

'Real-time' study for premature infants using MinION Nanopore and NanoOK RT

Both 1D and 2D Nanopore libraries were prepared using the SQK-RAD002 Rapid Sequencing Kit 1D and SQK-LSK208 Ligation Sequencing Kit 2D, respectively, and each library was sequenced on a R9.4 flowcell. Local basecalling through MinKNOW software was used to collect signal data.

To enable real-time analysis of the MinION data, an improved version of NanoOK software[229] was used. This new software NanoOK RT monitored a specified directory for FAST5 files as they were created. For each new file, a FASTA file was extracted automatically. FASTA files were grouped into batches of 500 to improve practicability, and each batch was BLAST searched against the NCBI nt database (downloaded February 2017) and the CARD database[230] (v1.1.1, downloaded October 2016) of antibiotic resistance genes. NanoOK RT has also the advantage that it can write out command files for MEGAN, which allows detailed analysis of community composition, either as the run proceeds, or after completion. NanoOK RT is available as an extension to NanoOK, selectable as a run-time option, from https://github.com/TGAC/NanoOK. The development of this tool was done by Richard Leggett and further information can be found in the manuscript.

Complementing this bioinformatic tool, NanoOK Reporter, was also developed for this project. This tool provided a graphical user interface to monitor the run, view summaries of species, and antibiotic resistance genes identified. NanoOK Reporter is available from https://github.com/richardmleggett/NanoOKReporter and allows the user to browse through data in real time as batches are processed, or after all of the results were in using their timestamps to indicate when a result was first obtained. The summary data can be exported as plain text files, and these were subsequently used for later analysis (below). Dr Richard Leggett was the developer of this tool, I prepared the figures for my thesis.

Generation of resistance heat maps

CARD results were obtained from NanoOK Reporter and used the option save summary data as a plain text file. This saved a text file for the analysis at each time point (batches of timestamped 500 reads) summarising the counts of resistance genes identified up to that point. We took the latest time point file (chunk 459, available at https://github.com/richardmleggett/bambi) and extracted a list of the ARO (Antibiotic Resistance Ontology). Each unique ARO was manually assigned to an antibiotic group. We subsequently wrote a script (gather_heatmap_data.pl, same GitHub repository) to take the summary files, together with this mapping and to generate a file (BAMBI_P8_2D_Local_070317_hits.txt) summarising hits per group at each time point. An R script (plot_card_heatmap.R, same GitHub repository) took this file and rendered the heat map. Dr Richard Leggett developed this tool and I assigned the AMR genes detected to its corresponding antibiotic group. I prepared the final figures for my thesis.

'Walking out' study from resistance genes to identify the encompassing bacteria

A shell script was written to go through all the CARD BLAST hits and wrote each CARD hit and the corresponding nucleotide hits for the same read

(walk_out_preprocess.sh, available at https://github.com/richardmleggett/bambi). A second script (walk_out.pl) took the output from the first script and parsed it, read-by-read. If there was a hit in nucleotides that began at least 50 bases before the start of the CARD hit, or at least 50 bases after the end of the CARD hit, then this species was taken as the encompassing species. The script also recorded count of the number of times each species was seen. The person who performed this analysis was Dr Richard Leggett. I was involved in preparing the final figure.

Isolation and biochemical characterisation of the P8 *Klebsiella pneumoniae* strain

Faecal sample P8 was homogenised in 1 mL TBT buffer (100 mM Tris/HCl, pH 8.0; 100 mM NaCl; 10 mM $MgCl_2 \bullet 6H_2O$) at 1500 rpm for 1 h. Homogenates were serially diluted to 10^{-4} in TBT buffer and aliquots of 50µl were spread on MacConkey (Oxoid) agar plates and incubated aerobically at 37 °C overnight.

Lactose-positive colonies (pink colour) were re-streaked on MacConkey agar three times to purify. Biochemical characterisation was performed using API 20E tests (Biomerieux) according to manufacturer's instructions. This work was performed in collaboration with Mr Tom Brook.

Determination of Minimal Inhibitory Concentration (MIC) for P8 *Klebsiella pneumoniae* isolate

Broth microdilution method[231] was used to calculate the Minimal Inhibitory Concentration (MIC) when antibiotic was added to the medium. MICs were determined as the lowest concentration of antibiotic inhibiting any bacterial growth.

Samples used for this study were from a *Klebsiella pneumoniae* isolate from baby P8 (sample P8E). Serial two-fold dilutions of the most common antibiotics used at NICU (benzylpenicillin, gentamicin, vancomycin, metronidazole, meropenem and

cefotaxime) were added to sterile nutrient broth. A fresh overnight culture was used as bacterial inoculum (10 µl). Microplates were incubated for 24 h at 37 °C under aerobic conditions. Cell density was monitored using a plate reader (BMG Labtech, UK) at 595 nm. I performed this experiment and analysis.

DNA extraction from P8 *Klebsiella pneumoniae* isolate for WGS analysis

10 ml of an overnight culture of the *Klebsiella pneumoniae* isolate was centrifuged at 1792g for 10 min, re-suspended in 30 ml of Phosphate Buffered Saline (PBS) (Sigma-Aldrich, UK) and centrifuged again. The pellet was then re-suspended in 2 ml of 25% sucrose (Fisher Scientific, USA) in TE (10 mM Tris (Fisher Scientific, USA) and 1 mM EDTA at pH 8.0 (VWR Chemicals, USA); 50 µl of Roche Lysozyme (Roche Molecular Systems, UK) at 100 mg/ml in 0.25 M Tris pH 8.0 was added. The mixture was incubated at 37 ^{o}C for 1 h, and 100 µl of Proteinase K at 20 mg/ml (Roche Molecular Systems, UK), 30 µl of RNase A at 10 mg/ml (Roche Molecular Systems, UK), 400 µl of 0.5 M EDTA pH 8.0 (VWR Chemicals, USA), and 250 µl of freshly prepared 10% Sarkosyl NL30 (Sigma-Aldrich, UK) were added. The mixture was then incubated on ice for 2 h and subsequently transferred to a water bath at 50 ^{o}C overnight. Next, E Buffer (10 mM Tris pH8.0 (Fisher Scientific, UK)) was added to the sample to a final volume of 5 ml, mixed with 5 ml Phenol:Chloroform:Isoamyl Alcohol (PCIA) (25:24:1) (Sigma-Aldrich, UK) in a Qiagen MaXtract High Density tube (Qiagen, DE) and centrifuged for 15 min at 1792g. The aqueous phase was transferred into a new Qiagen MaXtract High Density tube, made up with E Buffer to the volume of 5 ml if necessary, mixed with 5 ml of PCIA, and centrifuged for 10 min at 1792g. This procedure was repeated, with 5 min centrifugation time. Next, the aqueous phase was transferred into a Qiagen MaXtract High Density tube, made up to 5 ml with E Buffer if necessary, mixed with 5 ml of Chloroform:Isoamyl Alcohol (CIA) (24:1) (Sigma-Aldrich, UK), and centrifuged for 5 min at 1792g. The CIA step was repeated once more, after which the final aqueous phase was transferred into a sterile Corning ^{TM}

50ml centrifuge tube, and 2.5 volumes of ethanol (Ethanol absolute AnalaR NORMAPUR®, VWR Chemicals, USA) were added. The sample was incubated for 15 min at -20 ºC, then centrifuged for 10 min at 1792g and 4 ºC. Finally, the DNA pellet was washed with 10 ml of 70% ethanol and centrifuged at 1792g for 10 min twice, dried overnight, and re-suspended in 300 µl of E Buffer. This analysis was done by Mr Tom Brook.

Whole genome sequencing library preparation and sequencing of P8 *Klebsiella pneumoniae* isolate

DNA samples containing 500 ng genomic DNA were analysed. DNA was sheared into fragments of 400-600 bp using a Covaris plate with glass wells and AFA fibres. SPRI clean-up kit (Beckman, USA) was used to remove smaller sized fragments and concentrate the sheared DNA samples. Whole genome library construction performed by a liquid handling robot comprised end repair, A-tailing and adapter ligation reactions. Adapter ligated samples were subsequently amplified using the following PCR conditions: 5 min 95 °C, 10 cycles of (30 sec 98 °C, 30 sec 65 °C, 1 min 72 °C) and 10 min at 72 °C. LabChip GX was then used to size and assess the quality of the libraries and determine pooling volumes for each library using Beckman Coulter Biomek NXp (span-8). Final pools were finally loaded on the HiSeq 2500 sequencers. Sequencing was done at Sanger.

AMR gene characterisation of P8 *Klebsiella pneumoniae* isolate

Presence or absence testing of AMR genes was performed on one *Klebsiella pneumoniae* isolate assembled using Velvet. Contigs were aligned using BLAST (v2.2.30) against the CARD database under double filtering criteria of expected e-value $1e^{-10}$ and 90% identity. Customised in-house scripts were used to generate a data matrix that was then used to construct a heat map using the R heatmap2 package. This part of the work was performed by Mr Shabhonam Caim. I grouped

the AMR genes detected and prepared the final figure for my thesis and manuscript.

Results

Sequencing of a microbial mock community using MinION and Illumina technology

To evaluate the accuracy of MinION technology at sequencing metagenomics samples, a bacterial mock community mixture (HM-276D, BEI Resources, Manassas, VA) of known composition and abundance was sequenced. MinION results were validated with Illumina technology, using it as a research 'standard'.

Initial yield and length metric results from the MinION sequencing data showed that one flowcell (R7.3) produced 148,441 total reads, with 71,675 reads passing default quality filter, with a mean size of 3,047 bp, and longest read size of 40,561 bp. Table 18.

Table 18 Nanopore flow cell version and yield for mock community

Run	Sample	Flow cell	Seq'ng kit	Library type	Total no. of reads	No. of pass reads	Mean length of pass reads (bp)	Pass read N50 (bp)	Longest pass read (bp)
1	Mock	R7.3 (MAP103)	MAP006	2D	148,441	71,675	3,047	5,497	40,561

When comparing the taxonomic assignments obtained for each bacterial species using MinION and Illumina technology, results showed broadly similar abundance levels across both platforms (Figure 47). In some cases, a greater proportion of Nanopore reads were assigned at species level rather than genus or family, while in other cases, a greater proportion of Illumina reads were able to be assigned to species level. This can be explained by the nature of the longer reads, in some cases the longer length of Nanopore reads will provide better specificity, however

in other cases, these longer reads may also contain errors which may lessen specificity.

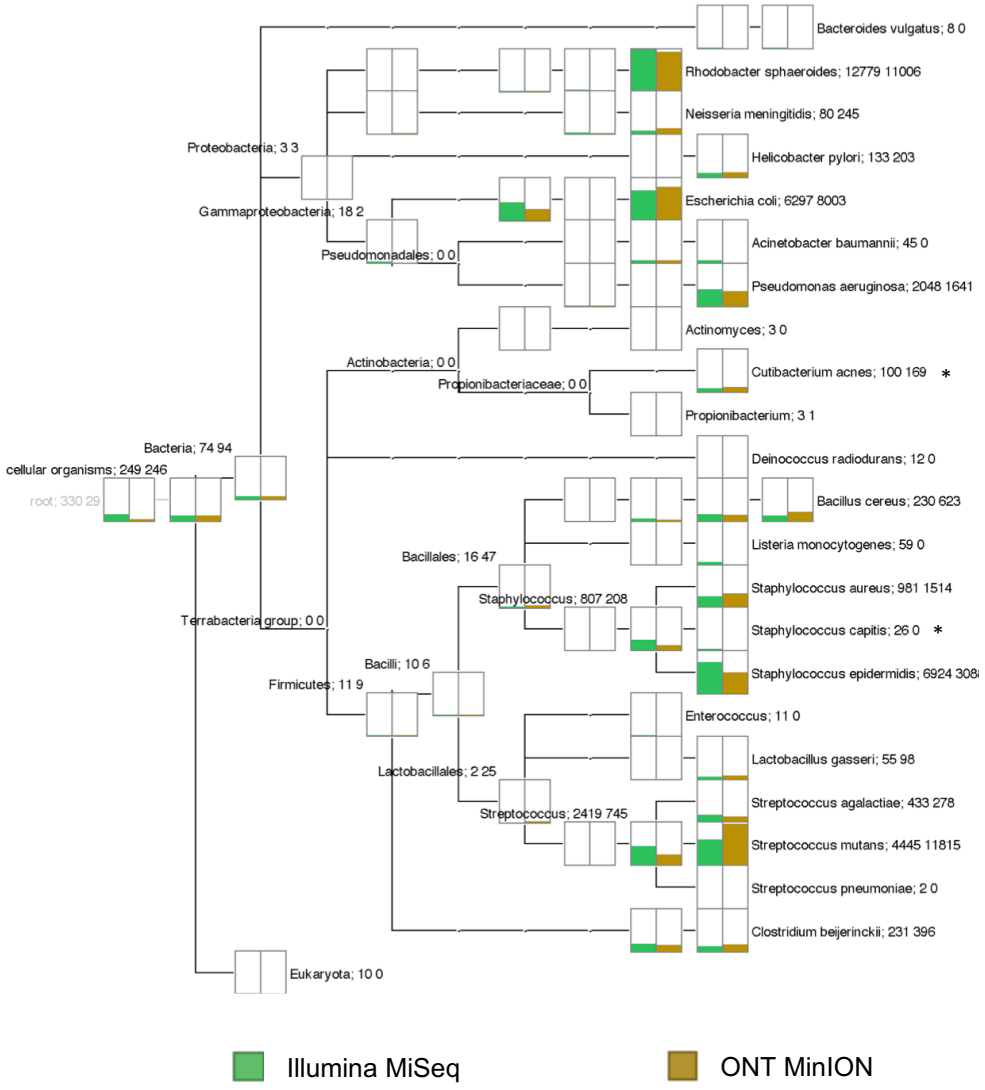


Figure 47 Sequencing of microbial mock community (HM-277D, BEI Resources) using Illumina MiSeq and Oxford Nanopore Technologies MinION sequencing

MEGAN taxonomy tree representing species assigned from the mock community control sequenced by Illumina MiSeq (green) and ONT MinION (brown). The height of the bars represents the number of reads assigned for each species taxa. (*) represents species assigned by Megan but not specified as members of the community.

Monitoring microbial disturbances in the preterm gut microbiome using MinION

The utility of the MinION platform in a clinical setting was determined next. Three different faecal samples (from days 13, 28 and 64 after birth) from one premature infant (P10) were used to monitor the diversity of the immature gut microbiome,

and study responses to clinical interventions. Results were benchmarked to 'gold standard' Illumina sequencing technology, both 16S rRNA amplicon and metagenomic sequencing. To standardise results from these different sequencing technologies, the DNA extraction protocol was kept consistent.

When comparing the taxonomic assignments obtained using MinION vs. Illumina it was observed that the results were comparable for the majority of the bacterial genera present in infant P10. Results from the most abundant bacteria detected *Klebsiella, Enterobacter, Veillonella, Staphylococcus* and *Bifidobacterium* were similar among the samples tested (Figure 48b). Interestingly, sample P10N (collected when the infant was receiving probiotic supplementation) confirmed colonisation of *Bifidobacterium bifidum,* and confirmed the ability of MinION to detect a known species in a complex microbial community from a patient sample. Subsequent samples (P10R and P10V) were collected after probiotic supplementation had stopped, and after additional courses of antibiotics. Furthermore, MinION analysis also detected presence of *Enterobacter cloacae* (in sample P10R), a well-known late onset sepsis (LOS) pathogen in premature infants[232], which correlated with the clinical diagnosis of suspected sepsis at the time of sample collection (Figure 48a).

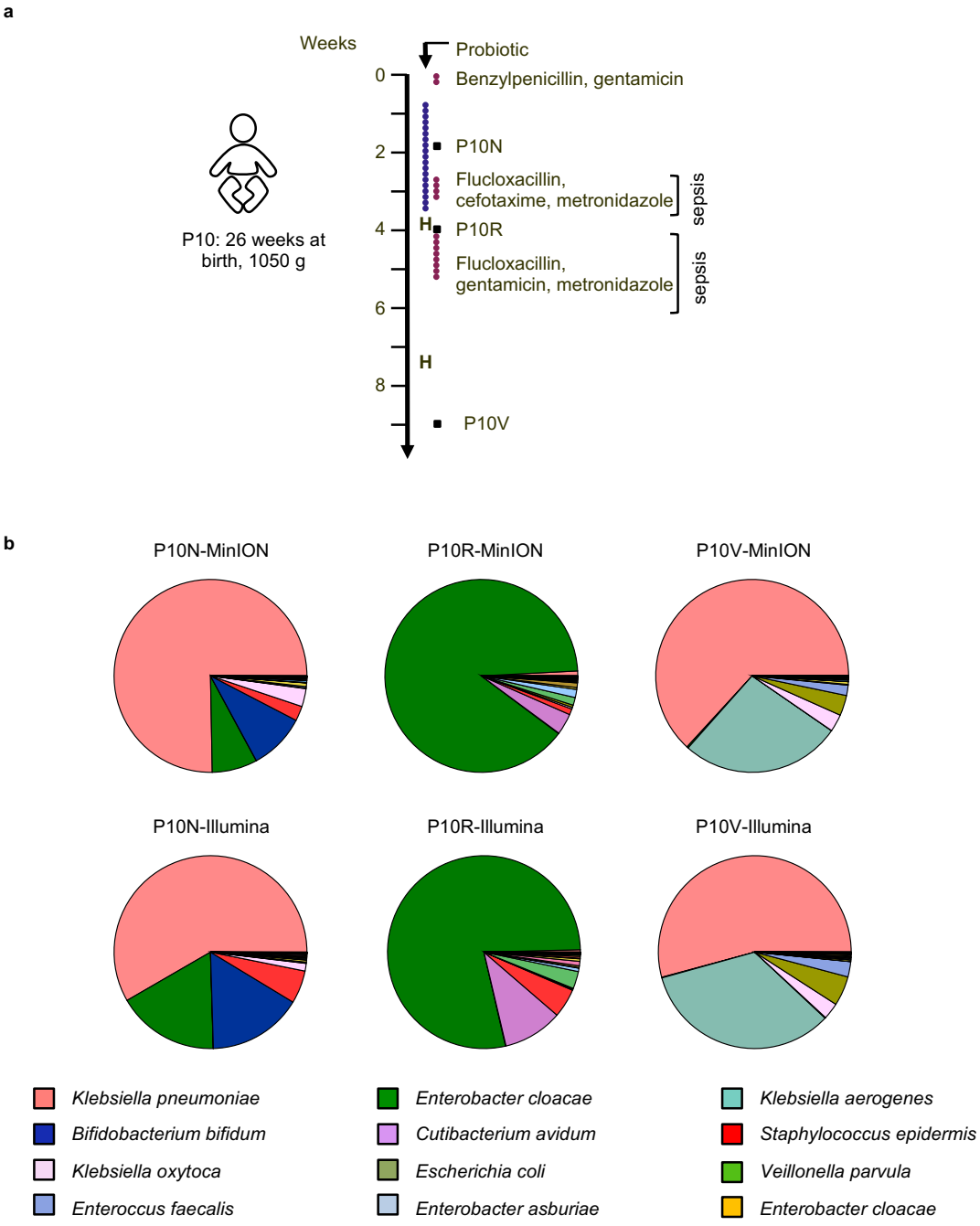


Figure 48 Longitudinal study on premature infant P10 using MinION and Illumina sequencing

(a) Timeline diagram of premature infant P10 indicating times of faecal sample collection (P10N, P10R and P10V), duration of antibiotic and probiotic treatment, and relevant clinical observations. The timeline diagram is divided into weeks and dots represent days within the scale. Blue dots represent days of probiotic treatment, red dots antibiotic treatment, black squares time points for sample collection and letter H transfer of the premature infant to another hospital.

(b) Sequencing data from ONT MinION and Illumina HiSeq 2500 from premature infant P10. Pie charts represents taxonomic profiles at different time points P10N, P10R and P10V, as assigned by MEGAN. The top row corresponds to results obtained using MinION sequencing and bottom row displays results using Illumina HiSeq. The figure legend shows all species with abundance ≥ 0.01% in at least one sample.

When comparing Illumina results to 16S rRNA gene data (Figure 49) similar taxonomic profiles were obtained. It is important to highlight that in some cases, the short 16S rRNA gene reads failed to differentiate some bacteria taxa even at genus level *e.g.* members of the family Enterobacteriaceae, which comprises commensal gut bacteria as well as opportunistic pathogens, and whose full-length 16S rRNA genes are often indistinguishable from one another.

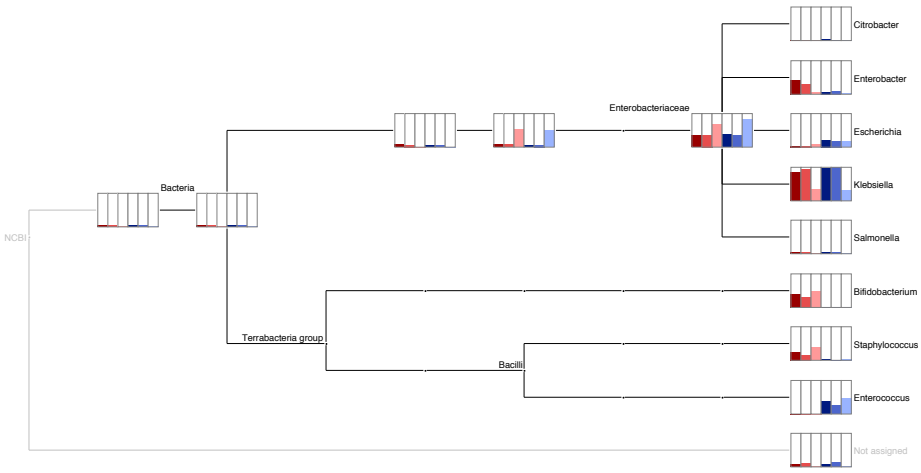


Figure 49 Megan taxonomic tree comparing assignments obtained by Illumina HiSeq 2500 WGS, ONT MinION, and Illumina 16S rRNA gene sequencing

Megan taxonomic tree showing bacteria taxa identified and their corresponding abundances. The height of the bars represents the number of reads assigned for each species taxa. Samples highlighted in red belong to sample P10N sequenced using Illumina HiSeq 2500 WGS, ONT MinION, and Illumina 16S rRNA gene sequencing. Samples highlighted in blue belong to sample P10V sequenced using Illumina HiSeq 2500 WGS, ONT MinION, and Illumina 16S rRNA gene sequencing.

To assess whether sequencing depth covered the total bacterial diversity present in the samples tested, rarefaction curves were performed and compared to Illumina technology (Figure 50). Results from this comparison showed that for both sequencing technologies most species were detected at approximately 25,000 reads, highlighted by the fact that the rarefaction curves reached saturation at this point. This result indicated that the level of coverage of MinION and Illumina HiSeq sequencing was very similar when using samples from premature infants,

which are known to have a less diverse gut microbiome when compared to adult samples[47].

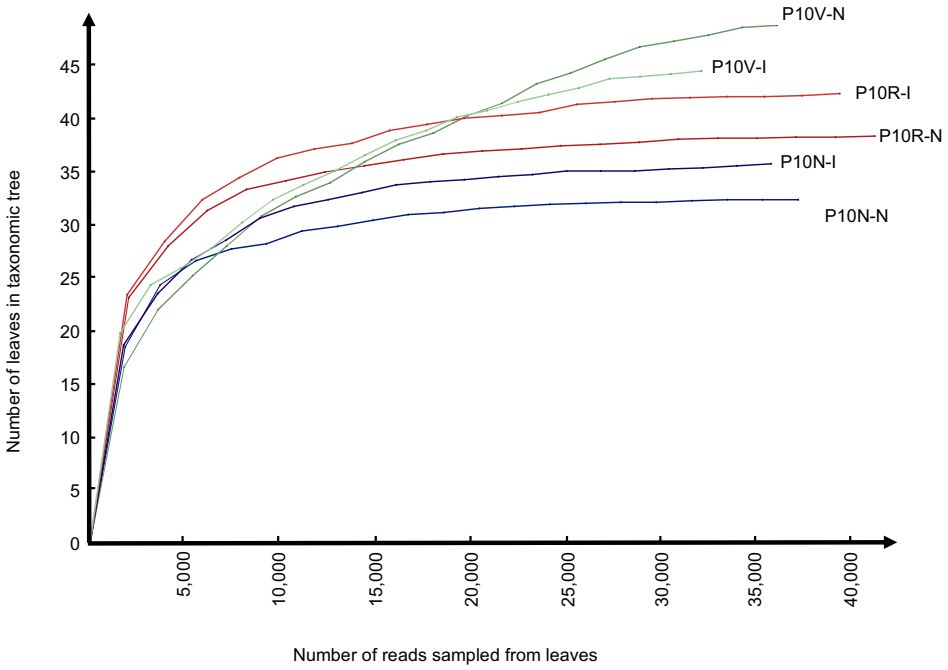


Figure 50 Rarefaction curves comparing MinION and Illumina HiSeq 2500 for premature infant P10

Rarefaction curves representing number of species (leaves) detected in the taxonomic tree vs number of reads sampled. Three samples from premature infant P10 were analysed: samples (P10N-N, P10R-N and P10V-N) were sequenced with MinION technology, while samples (P10N-I, P10R-I and P10V-I) were sequenced with Illumina technology.

Finally, a characterisation of the AMR profile in premature infant P10 was performed using the CARD database (a popular bioinformatic database use to trace antibiotic resistance genes) with comparison between MinION and Illumina sequencing data (Figure 51). Overall four groups of AMR genes were detected in high abundance including efflux pumps, β-lactams, aminoglycosides and fluoroquinolones. Elevated expression of efflux pumps has previously been observed in clinical multidrug resistant isolates, including *e.g.* AcrAB-TolC[233], which was detected in the three samples analysed. β-lactamase and aminoglycoside genes confer resistance to the antibiotics prescribed to preterm infant P10 (benzylpenicillin and gentamicin), while fluoroquinolone resistance genes correlate with the heavy use of fluoroquinolone antibiotics in hospitals[234] . Furthermore, the detection of AMR genes specific for certain species such as FosA2 present in

Enterobacter cloacae (for sample P10R), illustrates the ability of the MinION technology to not only detect the pool of AMR genes present in the samples tested, but also determine the species taxa-specific AMR genes.

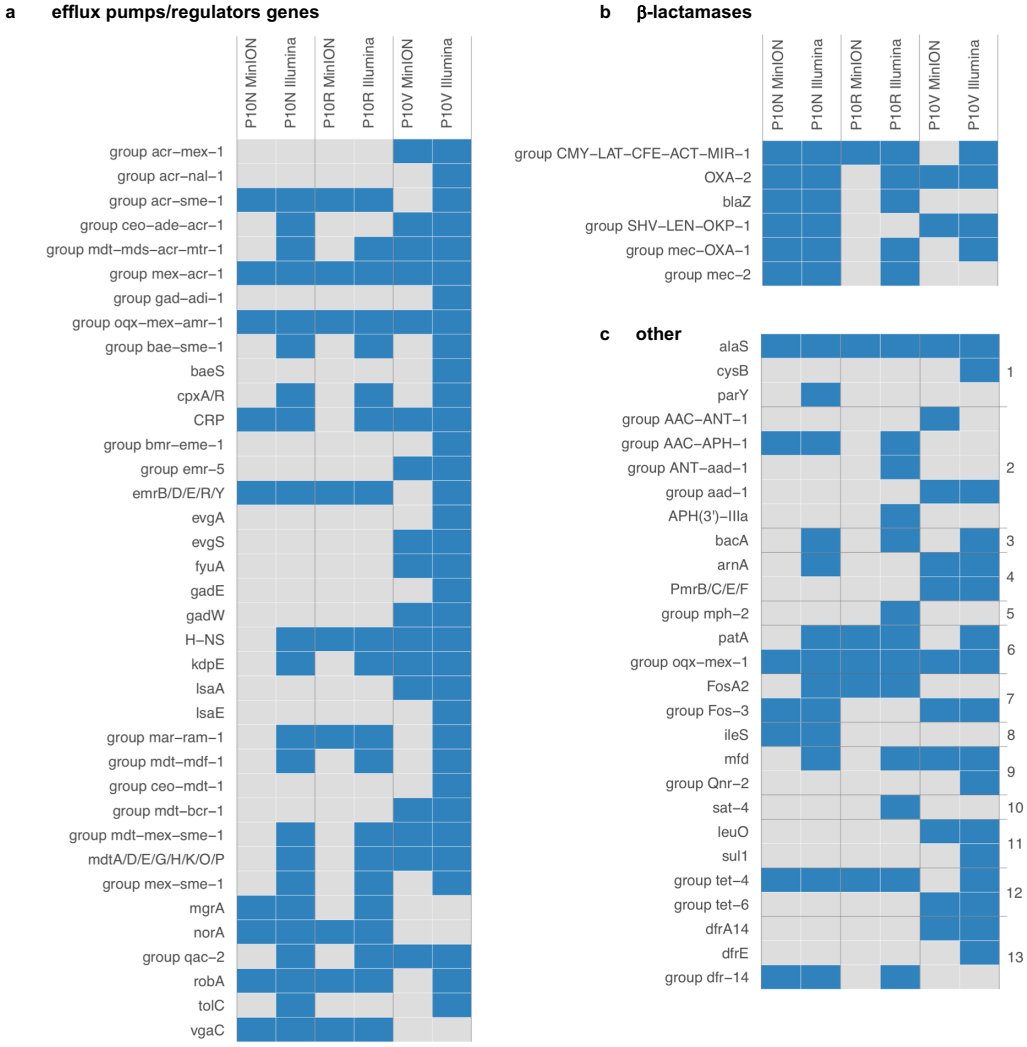


Figure 51 Heat map displaying presence or absence of AMR hits found in premature infant P10 using MinION and Illumina HiSeq 2500

Three samples of preterm P10 were analysed and sequenced using MinION (P10N-N, P10R-N and P10V-N), and Illumina technology (P10N-I, P10R-I and P10V-I). (**a**) Heat map representing efflux pumps inhibitors or regulators genes found in samples P10N, P10R and P10V. (**b**) Heat map highlighting b-lactamases. (**c**) Heat map showing (1) aminocoumarin resistance genes (2) aminoglycosides resistance (3) bacitracin resistance (4) colistin resistance (5) erythromycin resistance (6) fluoroquinolone resistance (7) fosfomycin resistance (8) mucopirocin resistance (9) quinolone resistance (10) streptothricin resistance (11) sulphonamide resistance (12) tetracycline resistance (13) trimethoprim resistance. AMR genes were grouped according to sequence similarity. AMR genes were grouped according to sequence similarity.

Using MinION technology to profile samples from healthy premature infants receiving probiotic supplementation

To validate the optimised MinION pipeline, two samples from healthy premature infants P106 and P116 (clinical details see Figure 45) were sequenced. The taxonomic profiles of these infants indicated that a dominant bifidobacterial gut microbiome profile correlated with improved health (Figure 52). Interestingly, MinION was able to detect *B. bifidum* which was present in the probiotic supplementation. Furthermore their 'resistome' was markedly reduced in comparison to infant P10.

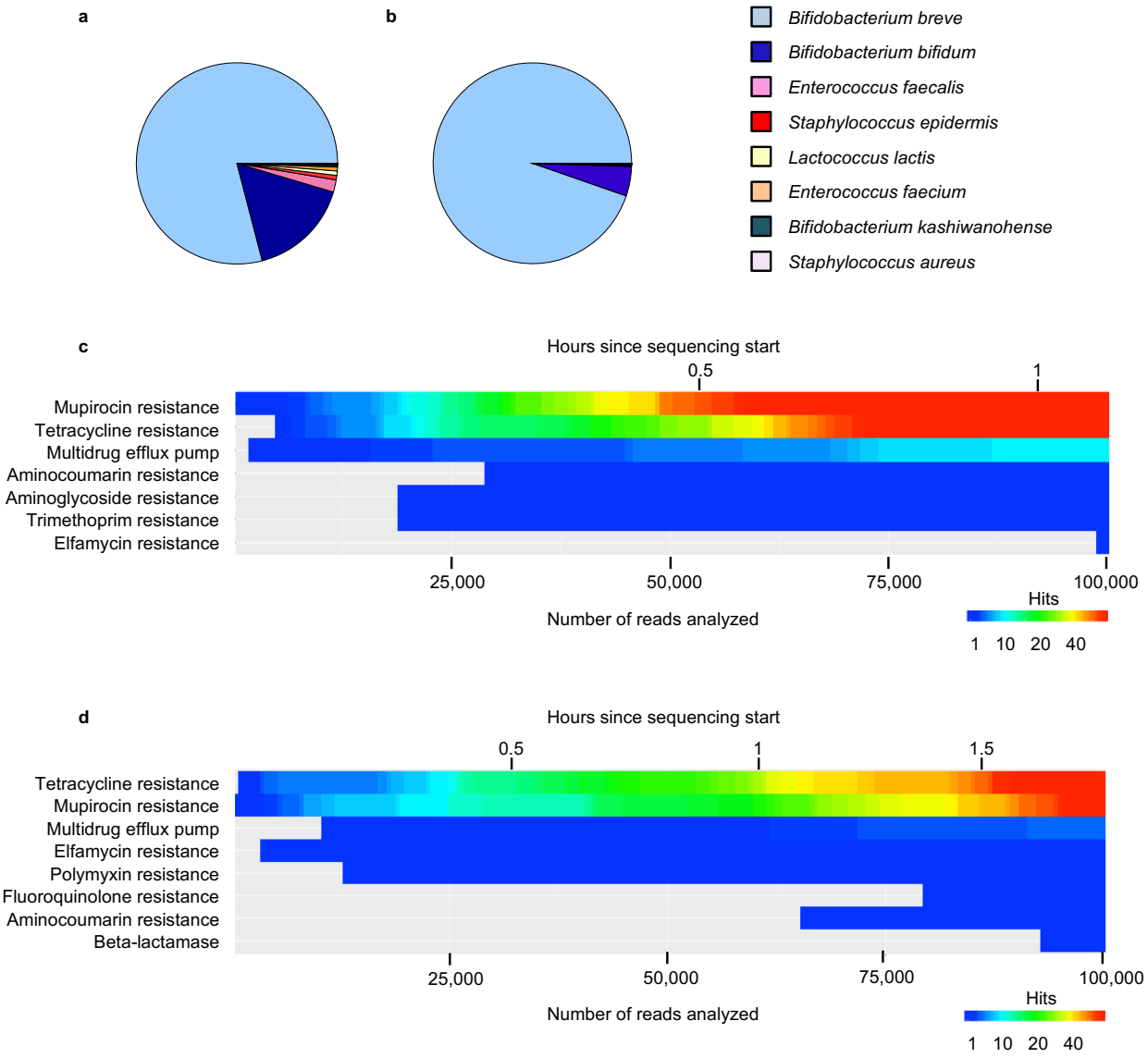


Figure 52 Rapid diagnostic of healthy premature infants P106 and P116 receiving probiotic supplementation

(a), (b) Taxonomic profiles obtained using MinION technology for premature infants P106 and P116, respectively. Figure legend comprises the 8 most abundant taxonomic taxa obtained.

(c), (d) Heat maps displaying number of CARD database hits detected among the most common groups of antibiotic resistance genes found in premature infants P106 and P116. Top and lower panel indicate the hours since sequencing start, and the number of reads analysed, respectively within this timeframe.

New bioinformatics tools utilise MinION specific features for improved rapid characterisation of gut-associated pathogenic bacteria and resistance profiles

New bioinformatic tools (NanoOK RT and NanoOK Reporter) were developed by Dr Richard Leggett to add real-time functionality to the optimised MinION pipeline. These tools were able to provide a report with the most prevalent AMR

genes, and perform a ‘walkout’ analysis from the AMR genes into flanking DNA of the host bacteria containing these genes.

Using these bespoke bioinformatics tools, three samples from three critically ill infants suffering from NEC (P49, P205 and P8) were profiled (clinical details of these premature infants can be found in the Figure 46). For this study, the latest flowcell version available R9.4 was used. Taxonomic results show that infants P49 and P205 samples both contained high proportions of *E. cloacae* (Figure 53). Analysis of the ‘resistome’ of these infants highlighted a significant number of AMR genes, particularly in P205 (*i.e*. aminoglycoside resistance and β-lactamases), which were detected within minutes of sequencing start (Figure 53b-d). Although these babies had *E. cloacae* dominated gut microbiome profiles, and a significant community ‘resistome’, they also harboured other potentially pathogenic bacteria (*e.g. Klebsiella pneumoniae*), highlighting the clinical importance of determining which bacteria are harbouring AMR genes for downstream treatment options.

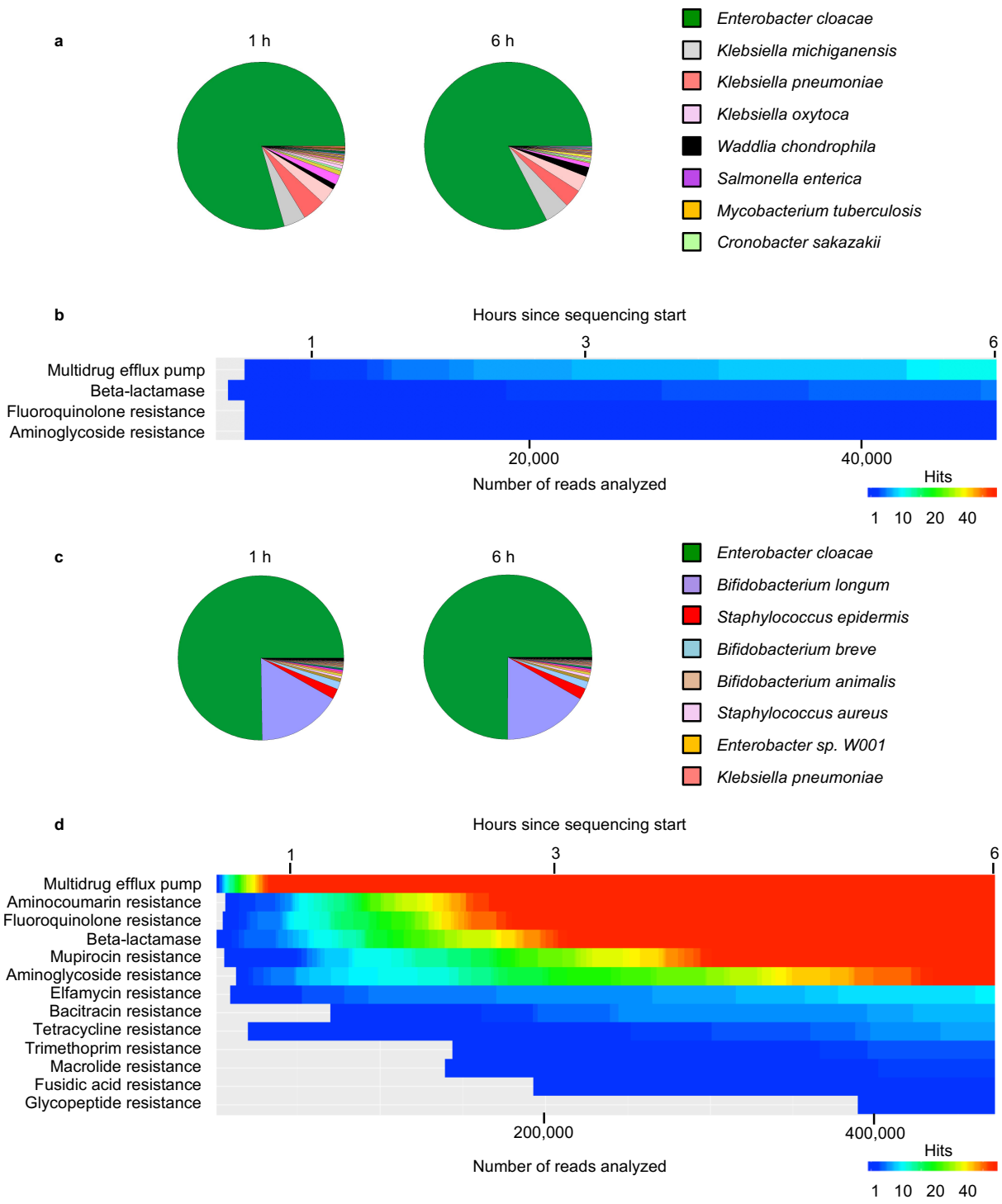


Figure 53 Rapid diagnostic using MinION technology for premature infants clinically diagnosed with suspected NEC (P49 and P205)

(a), (c) Taxonomic profiles comparing results obtained at 1h and 6h since sequencing started. Results for premature infant P49 are highlighted in a and for premature infant P205 are highlighted in c. Figure legends comprise the 8 most abundant taxonomic taxa obtained.

(b), (d) Heat maps displaying number of CARD database hits detected among the most common groups of antibiotic resistance genes found in premature infant P49 and P205, respectively. Top and lower panel indicate the hours since sequencing started and the number of reads analysed.

As MinION reads are typically longer than Illumina reads, a 'walkout' approach was used to extract additional information by examining flanking sequences either side of each AMR hit, and searching the NCBI nt database for hits that were independent (defined as ≥ 50 bp) from the AMR sequence. Results of this analysis showed that the vast majority of AMR genes of infant P205, mapped back to *E. cloacae* (88%, Figure 54a), which was also the dominant species taxonomically, with the remaining resistance genes (6%) associated with *B. longum* (*i.e*. mupirocin resistance). Contrastingly, although infant P49 had very similar levels of *E. cloacae* compared to infant P205, only 60% of AMR hits were associated with *E. cloacae. Klebsiella* represented a small proportion taxonomically in P49 (~13%), however *Klebsiella* species (*K. pneumoniae*, *K. michiganensis* and *K. oxytoca*) appeared to encode a range of AMR genes (*e.g.* OXA-2 (β-lactamases), CRP and mexB (efflux pumps) and patA and mfd (fluoroquinolone resistance)), making up >30% of total AMR genes present in this infant sample (Figure 55b and Appendix 4). These data highlight that MinION sequencing data coupled with the NanoOK Reporter analysis software is able to map specific AMR determinants to specific pathogenic bacteria and could facilitate targeted antibiotic treatment.

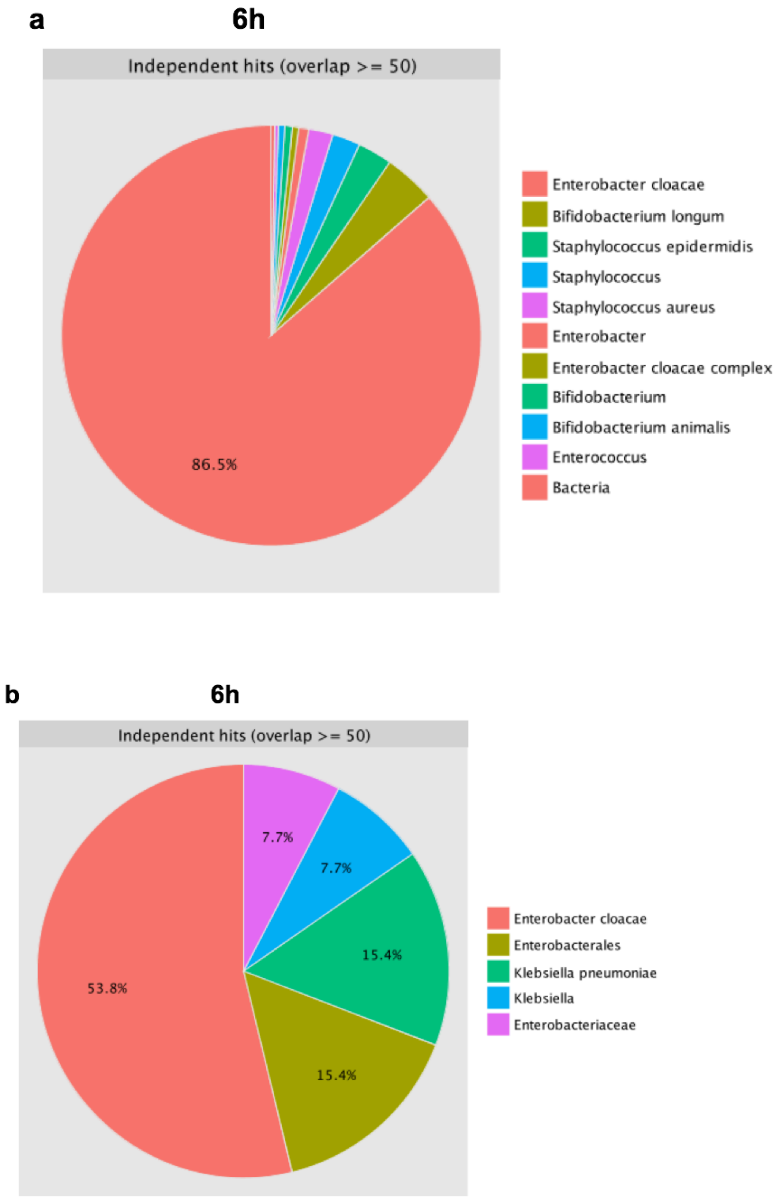


Figure 54 Walkout study for premature infants P205 and P49 reported by NanoOK RT software

(a) Results from independent hits (defined as ⩾ 50 bp overlap from the AMR sequence) at 6 hours of sequencing for premature infant P205.

(b) Results from independent hits (defined as ⩾ 50 bp overlap from the AMR sequence) at 6 hours of sequencing for premature infant P49.

Next, a 'real-time' run (from sample preparation to data analysis) using MinION R9.4 flow cells and 1D libraries was performed. The faecal sample chosen for this analysis came from a premature infant (P8) clinically diagnosed with suspected NEC (clinical details in Figure 46c, methods section). This infant was exposed,

before sample collection, to 43 days of non-concurrent antibiotic treatments (*i.e.* benzylpenicillin, gentamicin, meropenem, tazocin, vancomycin, flucloxacillin, metronidazole and amoxicillin). Current rapid clinical microbiology tests can take between 36 and 48 h if strain susceptibility to antibiotics is determined. All stages of this run were timed including sample preparation (90 min), DNA quality control (45 min), 1D MinION library preparation and loading onto the MinION flow cell (1 h and 45 min), and sequencing-and-data analysis (8 min for first specific AMR hit at 1,500 reads, and 5 h in total to obtain saturation point at 20,000 reads, Figure 55).

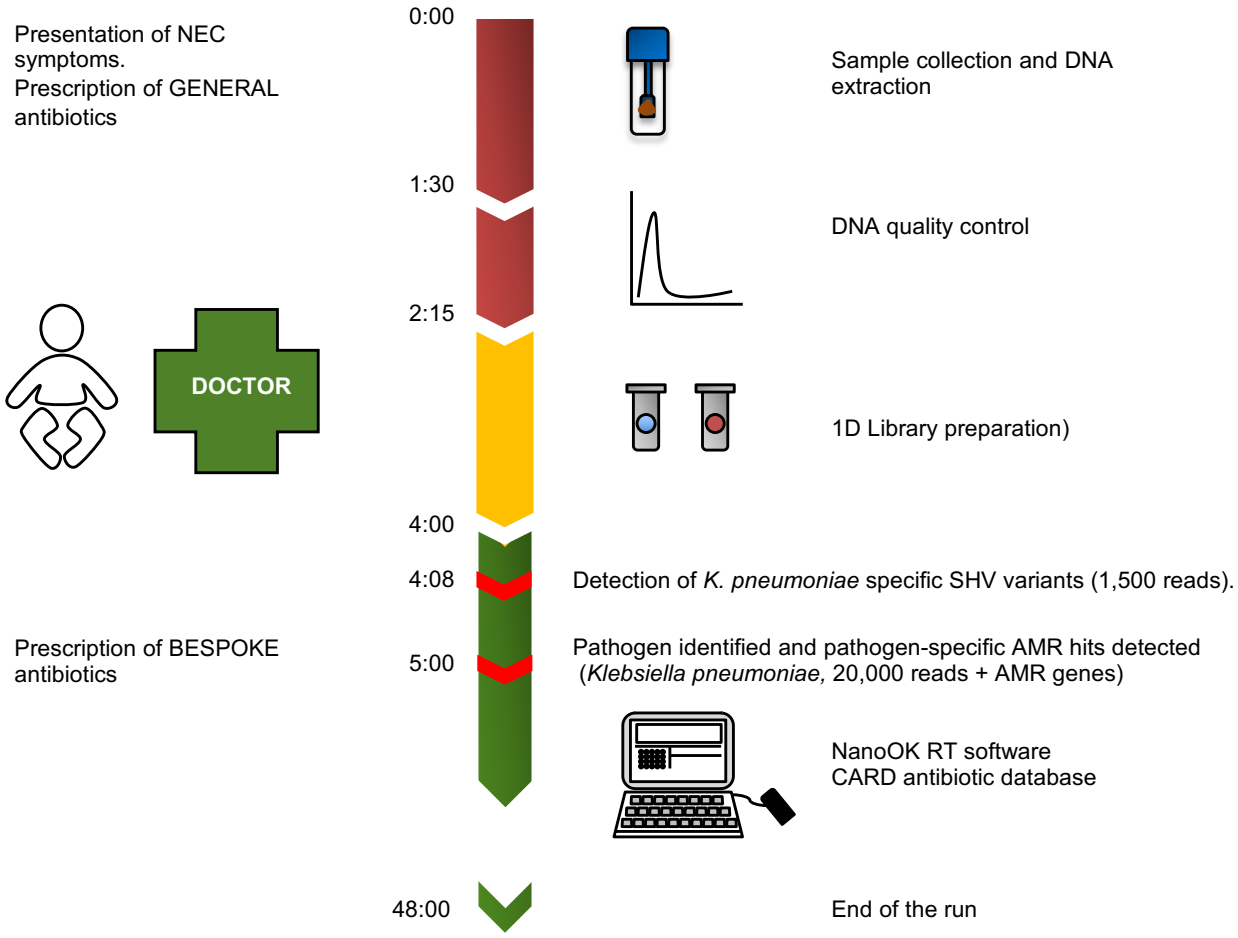


Figure 55 Timeframe diagram for 'real time' run performed for rapid diagnostic of premature infant (P8) suffering from NEC

Step 1 (red, 2h 25min): Sample collection, DNA extraction and quality control. Step 2 (yellow, 1h 45min): 1D library preparation incorporating bead clean up and DNA repair. Step 3 (green): data analysis using local base calling and NanoOK RT. Pathogen detection (*Klebsiella pneumoniae*) and *K. pneumoniae* specific AMR genes were first detected at 4 hours and 8 minutes (1,500 reads analysed). Left side of the panel indicates clinical symptoms and general guidelines for antibiotic prescription.

This run generated a total of 1.37 million reads in a full 48 hour run. By 1 h after sequencing start (5 h total), the pipeline had analysed 20,000 reads and *K. pneumoniae* accounted for around 70% of reads. These reads were much longer (N50 3,479 bp) than the previous 2D runs from samples P49 and P205 (Table 19), meaning that at the 1h time point NanoOK RT and NanoOK Reporter had analysed over 3x more sequence data in this new 1D run.

Table 19 MinION runs of premature infants suffering from NEC

Sample	Flow cell	Seq'ng kit	Library type	Total no. of reads	Mean length of pass reads (bp)	Pass read N50 (bp)	Longest pass read (bp)
P8	R9.5 (MIN107)	LSK108	1D Ligation	1,369,544	1,838	3,479	897,734
P49A	R9.5 (MIN107)	LSK208	1D Ligation	84,527	1,046	1,338	34,975
P205G	R9.5 (MIN107)	LSK208	1D Ligation	2,745,619	966	1,102	11,619

The first 500 reads immediately indicated a dominance of *K. pneumoniae* (a potential causative organism that has been associated with NEC pathogenesis in premature infants[235]), as well as *Proteus mirabilis*. To further verify enough of the bacterial diversity was sequenced at 1 h time point, a comparison with the taxonomic profile at 6 h was performed; time-point chosen due to clinical relevance to NEC deterioration (101,000 reads, 10 h total time). This comparison verified that there were no significant qualitative differences between the two taxonomic profiles (Figure 56).

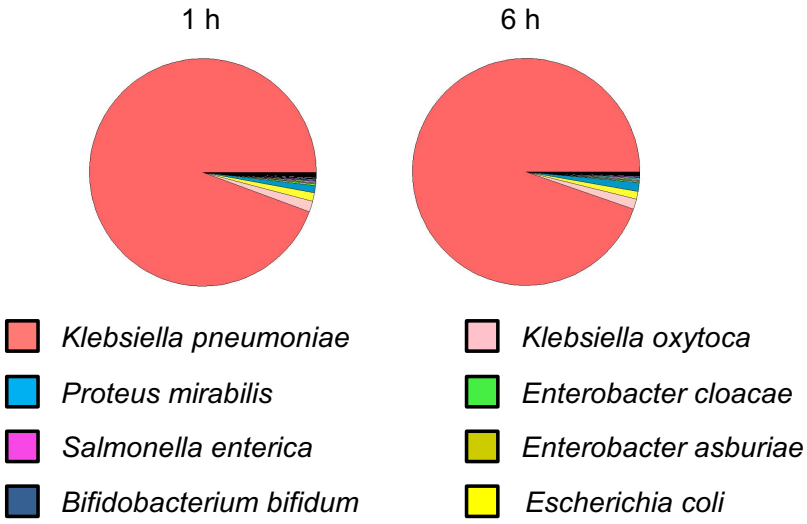


Figure 56 Rapid diagnostic of premature infant P8 clinically diagnosed with suspected NEC

Taxonomic profiles obtained using MinION technology at 1h, and 6h since sequencing started. Figure legend comprises the 12 most abundant taxonomic taxa obtained.

As highlighted previously, it is clinically important to detect AMR genes in metagenomic samples from premature infants to guide appropriate antibiotic prescription. In this 'real-time' run it was determined how rapidly we could map AMR genes to the CARD database. Figure 57 shows the huge number of AMR gene classes detected throughout the run, including polymyxin, aminoglycoside, tetracycline, quinolone resistance, β-lactamases and efflux pumps, all of which were detected in as little as 1 h after sequencing start. *K. pneumoniae*-specific SHV variants[236] were detected as early as 6 min (at 1,500 reads, 4 h 8 min total time), whilst other lower abundance AMR genes in the sample, such as those conferring trimethoprim, sulphonamide and streptothricin resistance, were not detected until 3-4 h post sequencing (7-8 h total).

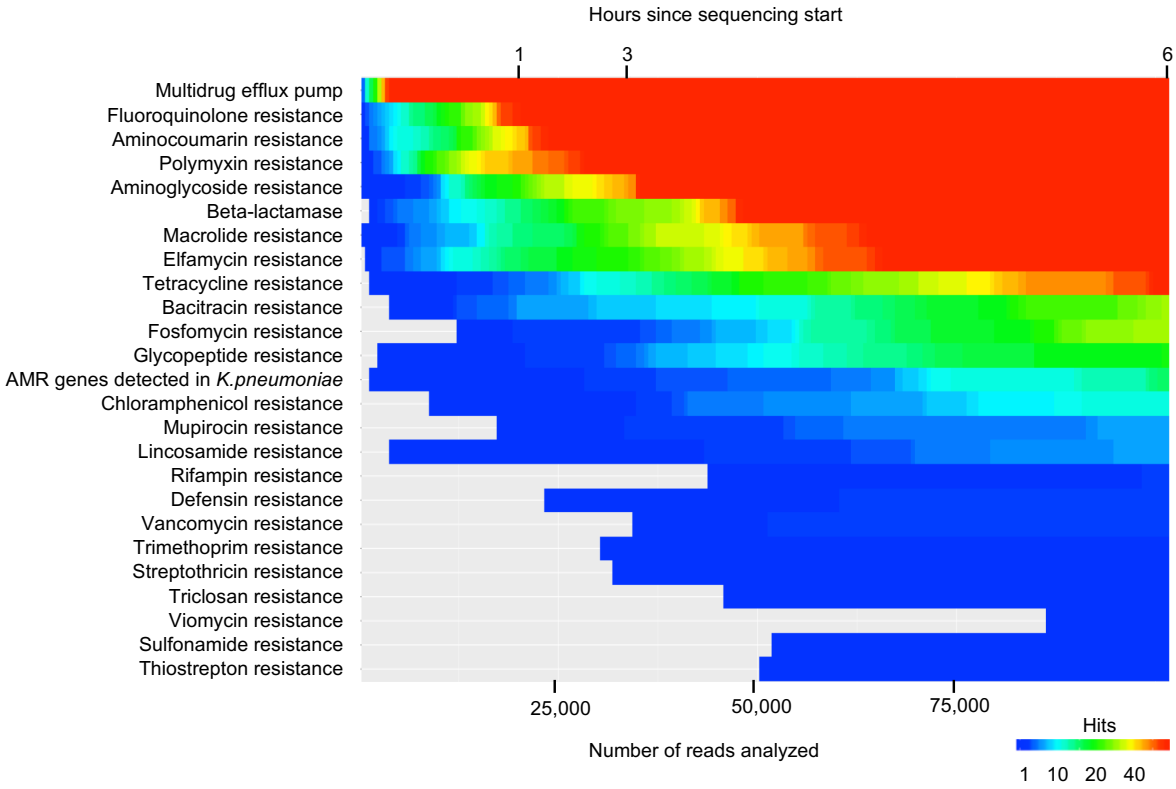


Figure 57 Rapid diagnostic of AMR genes detected for premature infant P8 clinically diagnosed with suspected NEC

Heat map displaying number of CARD database hits detected among the most common groups of antibiotic resistance genes found in premature infant P8. Top and lower panel indicate the hours since sequencing started and the number of reads analysed, respectively within this timeframe.

Finally, when using NanoOK Reporter to perform AMR 'walkout' analysis on this infant sample, the majority (~90%) of AMR genes within the whole metagenomic sample mapped to *K. pneumoniae* (Figure 58, Appendix 4), including multidrug exporters such as *acrB* or *oqxA*, conferring resistance to tetracycline, chloramphenicol, and fluoroquinolones, *vanSC* (resistance to vancomycin), *tet 41* (resistance to tetracycline) and *dfrA20* (resistance to trimethoprim). There were also specific AMR gene cassettes to *P. mirabilis* including OXA-63, which can confer cephalosporin resistance, and *tet34* resistance to tetracycline.

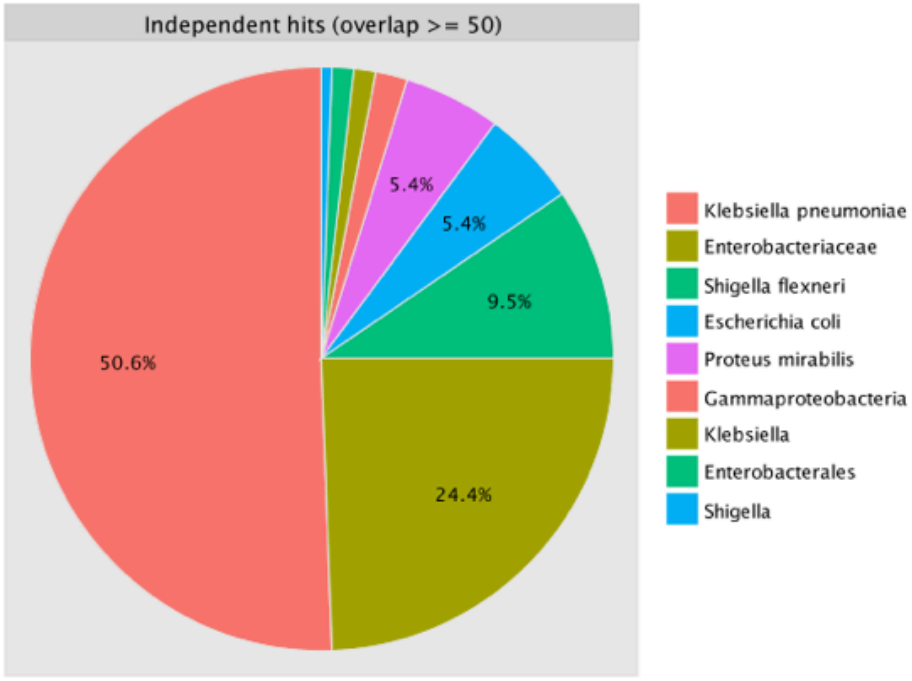


Figure 58 Walkout study of premature infant P8 reported by NanoOK RT software

Results from independent hits (defined as ≥ 50 bp overlap from the AMR sequence) at 6 hours of sequencing.

Genomic characterisation of *Klebsiella pneumoniae* isolate from premature infant P8

To validate the metagenomic genotypic data using NanoOK RT tool and its 'walkout' analysis, *K. pneumoniae* was isolated from patient P8, with whole genome sequencing (WGS, using both Illumina and MinION) performed. The same AMR pipeline using CARD database was used to compared the results. Many of the AMR genes/groups detected in the 'walk-out' analysis from the metagenomic sample P8 correlated with both the Illumina and MinION whole genome sequencing data of the isolate (Figure 59). A significant proportion (~60%) of the resistance genes/groups in the metagenomics 'walk-out' and isolate WGS correlated with efflux pumps (e.g. group *mdt-mds-acr-mtr*, group *mdt-mex-sme*, group *mex-acr*, and group *oqx-mex-amr*), whilst other 'hits' correlated to known *K. pneumoniae* AMR genes/groups including b-lactamases genes (e.g. group *SHV-LEN-OKP*), or Fosfomycin (group *Fos-3*).

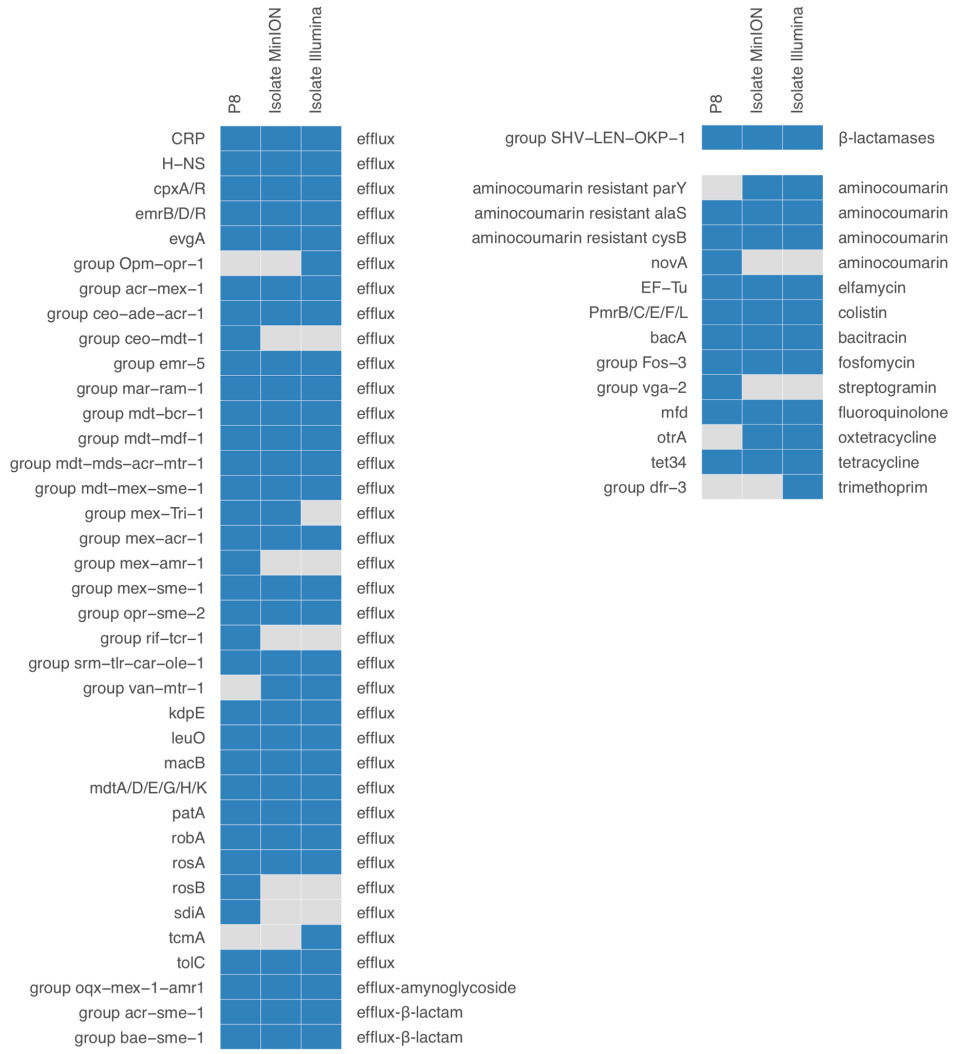


Figure 59 AMR genes associated to *Klebsiella pneumoniae* from a metagenomic sample compared to those found in a *Klebsiella pneumoniae* isolate from the same sample

Heat maps displaying AMR genes associated to *Klebsiella pneumoniae* from faecal sample P8 and AMR genes detected from a *Klebsiella pneumoniae* isolate from the same sample. Sequencing of the metagenomic sample was performed for only 6h using MinION and NanoOK RT tool (highlighted as 'P8''). Sequencing and assembly of the *Klebsiella pneumoniae* isolate was performed using both MinION ('Isolate MinION') and Illumina HiSeq ('isolate Illumina''). AMR genes were grouped according to sequence similarity.

Overall these results indicate MinION metagenomic sequencing using NanoOK Reporter and 'walkout' analysis is faster and provides robust clinically relevant AMR data that may help guide antibiotic treatment.

Phenotypic characterisation of *Klebsiella pneumoniae* isolate from premature infant P8

Antibiotic resistance phenotype of the same *K. pneumoniae* isolate was tested against the seven most commonly used antibiotics in NICUs (Table 20).

Interestingly, the *K. pneumoniae* isolate was found to have a higher minimum inhibitory concentration (MIC) breakpoint value for those antibiotics that were prescribed to the premature infant P8 *i.e.* benzylpenicillin, amoxicillin, metronidazole, gentamicin and vancomycin. In contrast, the only MIC breakpoint value lower than those put forward by the European Committee on Antimicrobial Susceptibility Testing (EUCAST[153]) was for cefotaxime, an antibiotic not prescribed to the infant at sample collection point. Notably, these data also correlated with the AMR data generated by NanoOK reporter and the 'walkout analysis', with the only exception of metronidazole resistance, which was only detected after WGS of the isolate (gene *msba***,** Appendix 4).

Table 20 Broth microdilution test for *Klebsiella pneumoniae* isolate from baby P8 (sample P8E) and Eucast values

Antibiotic	**MIC (mg/L)**	**Eucast (mg/L)**
Gentamicin	3.12	2
Benzylpenicillin	780	No data
Amoxicillin	3,900	>512
Metronidazole	1,250	No data
Vancomycin	1,562	No data
Meropenem	6,25	0.125
Cefotaxime	0.19	0.25

Discussion

This work demonstrates that MinION technology can be used at clinical point of care for profiling microbial communities and corresponding AMR determinants, which may facilitate antibiotic treatment guidance, ultimately improving patient outcomes. This technology is smaller, faster and cheaper than any other sequencing technologies currently available (*e.g.* Illumina platforms). Using a combination of improved Nanopore sequencing chemistries, and bespoke Nanopore analysis packages (NanoOK RT and NanoOK Reporter, developed by Dr Richard Leggett), the MinION platform can successfully profile known metagenomes, and clinical samples from critically ill premature infants. Importantly, MinION sequencing data using the new R9.4 flow cells were comparable in discriminatory power to the conventional Illumina sequencing platform and provided clinically relevant information within just 5 h from sample receipt.

Initially, it was important to benchmark the MinION against a known metagenomic sample (mock community), to determine the usability of this new technology. Results indicated that the MinION is suitable for detection of a wide range of microbes, which is in agreement with a previous study also using R7.3 flow cells[237]. When testing longitudinal samples from a premature infant residing in NICU and sequenced the same samples using Illumina as a gold standard, results showed analyses were comparable for both sequencing platforms, detecting the probiotic strain (*i.e. B. bifidum*) during the supplementation period, and the pathogenic strain *E. cloacae* during sepsis episodes (Figure 48b). The investigation of AMR genes detected by MinION and Illumina, revealed that both sequencing platforms generated reads mapping to genes with similar antibiotic resistance mechanisms (Figure 51), and only 4 genes (*mphC*, *fusB*, *sat-4 and vanRG*) with unique resistance mechanisms out of all 146 AMR genes were detected exclusively by Illumina. This result may be correlated to the lower MinION read count, and so could be mitigated by ongoing improvements in MinION technology. Notably, we observed the presence of AMR genes that corresponded to prescribed antibiotics;

β-lactamase and aminoglycoside genes (conferring resistance to benzylpenicillin and gentamicin respectively), while fluoroquinolone resistance genes did not correlate to any prescribed antibiotics, and so may relate to AMR gene transfer of strains from other sources. Thus, AMR profiling using MinION technology may be able to guide clinical treatment decisions at an earlier stage of patient care.

For critically ill patients, rapid actionable information is crucial for improving patient outcomes. Thus, it was demonstrated that the entire pipeline of sample preparation, library construction, sequencing and analysis could be carried out rapidly and in 'real-time'. Importantly, the most recent flow cells (R9.4) were used for this study, which have an improved error rate (~5-7% for 1D) and yield. Specifically, for this study new bioinformatic tools (NanoOK RT and NanoOK Reporter) were developed by Richard Leggett to improve functionality. Two ill infants were initially diagnosed with suspected NEC, P49 and P205. Taxonomic results showed both samples presented high levels of *E. cloacae,* and a significant resistome, which may correlate with the extensive course of antibiotics the infants received (Figure 53). Importantly, additional analysis using the new NanoOK Reporter functionality, allowed determination of what specific taxa harboured these AMR genes, and efflux pumps, including *E. cloacae*, which was the dominant species present (Figure 54). Interestingly, although *Klebsiella spp.* represented a more minor component of the P49 microbial community they appeared to harbour > 30% of the total AMR genes present. Thus, in patient P49 a poorly chosen antibiotic treatment could target *Klebsiella* and miss the predominant *E. cloacae* or could target only *E. cloacae* leading to an increase in *Klebsiella* pathogenic species, whereas the best treatment would target both sets of pathogenic species. Performing a 'walkout' study rather than *de novo* metagenomic assembly is considerably less compute intensive. These data indicated that relevant AMR genes detected in a metagenomic sample and further mapped to known pathogenic species may facilitate tailored antibiotic treatment strategies for critically ill patients.

Next, a real-life diagnostic approach was executed by performing a real-time run. The sample for this experiment was from an extremely ill premature infant P8 (born after 26 weeks' gestation with a birthweight of only 508 g), who had received multiple courses of antibiotics since birth (46 days of antibiotic treatment out of 63 days of life at sample collection) and presented suspicion of NEC at the time of sample collection.

Initial attempts to real-time sequence infant P8 were unsuccessful, due to problems related to quality of flow cells and base calling. A third attempt was needed to generate impressively high yields (101,000 reads) at only 6 hour of sequencing and 1.37 million reads in a full 48 hour run. Taxonomic profiling in real time revealed a *K. pneumoniae*-dominated profile, after just 1 h of sequencing (~20,000 reads), enabling to confidently 'call' this potential pathogen. This analysis was further strengthened as more sequencing at 6 h, gave almost identical microbial profiles (Figure 56). *K. pneumoniae* has been linked to NEC (and is supported by corresponding clinical observations), with overgrowth in the intestine linked to pathological inflammatory cascades, facilitated by a 'leaky' epithelial barrier[238]. It should be noted that the single species domination in this sample facilitates early detection at lower read depth and low-level pathogen abundance would require deeper sequencing. Profiling of additional more complex samples from NEC diagnosed infants (*i.e*. P49 and P205) also indicates distinct and differential microbiome profiles (when compared to P8) also at 1 h post sequencing start (Figure 53), highlighting how rapid diagnosis of pathogen overgrowth is possible using R9.4 Nanopore flow cells.

Whilst detection of individual pathogens is important, a critical additional requisite is identification of AMR profiles so that tailored antibiotic treatment can be used. Real-time analysis of MinION data using NanoOK RT highlighted the presence of a significant metagenomic 'resistome', including presence of colistin resistance, a last resort antibiotic, by the detection of genes *arnA*[239], *PmrB*[240] and *PmrC*. It was

noted the greater the sequencing depth the greater the number of AMR genes detected, although importantly there was a significant number of AMR genes detected as rapidly as 1 h after sequencing start, including β-lactamases, quinolone, aminoglycoside, and tetracycline resistance genes (Figure 57).

To further benchmark the MinION pipeline developed, the pathogenic *K. pneumoniae* strain was isolated, sequenced and tested phenotypically (MIC test). Notably, WGS (both Illumina and MinION) indicated agreement with resistance genes determined by the MinION metagenomic run, and the 'walkout' approach (Figure 59). When subjecting strains to MIC testing (current gold standard for profiling AMR), phenotypic resistance was observed to all main groups of antibiotics that had been prescribed to infant P8, Table 20). There was good association between AMR gene sequence detection and MIC testing, *i.e.* group *SHV-LEN-OKP* and β-lactam antibiotics, *aac* and *aph* genes and gentamicin, and *van* genes resistant to vancomycin which highlights that MinION could be extremely useful for rapid AMR profiling (Appendix 4). Only metronidazole resistance was identified via MIC testing, that was not present in the MinION metagenomic analysis. The corresponding resistance gene was only detected in the isolate whole genome sequencing data (gene *msba*, Appendix 4). As such it is expected that rapid analysis using MinION sequencing, would inform early and more appropriate antibiotic choices for patient care, halting the rapid deterioration observed in critically ill patients.

This work indicated that MinION metagenomic profiling can provide robust, clinically useable data in less than 5 hours. In comparison to standard sequencing platforms such as Illumina MiSeq sequencing (paired 250 bp reads) and PacBIO (>15,000 bp reads) where obtaining first sequencing results normally take ~39 hours[241] or ~7 hours[242], respectively. Profiling preterm faecal samples using MinION technology as a routine basis could offer insights into overall preterm microbiome dynamics, and could complement other rapid molecular diagnostics

tests currently used for detecting bacterial bloodstream infections such as QuickFISH technology or MALDI-TOF[243]. The QuickFISH technology uses a fluorescent peptide nucleic acid probe able to target the bacterial 16S rRNA gene and has the advantage that it is extremely fast (~20 min turnaround time) but has the disadvantage that only known bacterial pathogens can be traced. In contrast, the MALDI-TOF approach using mass spectrometry can identify any bacteria directly from blood culture in approximately 1 hour. QuickFISH and MALDI-TOF technologies have the disadvantage that they cannot perform AMR profile.

Nevertheless, this application has limitations if the sample is not treated adequately to remove human cells, and presents difficulties to detect bacterial pathogens in very low amounts[244]. Overall the fast turnaround time, high sensitivity and vast amount of data provided by the MinION Nanopore pipeline developed in this study, makes this technology very attractive to complement the current rapid molecular diagnostics tests used in the clinic.

Conclusion

This work demonstrates that MinION technology has the potential to diagnose, in 'real time', premature infants suffering from NEC, and to monitor the effectiveness of microbiota therapy (probiotic supplementation). Data obtained may allow clinicians to rapidly tailor antibiotic treatment strategies in a rapid (~6 h decision from sample receipt) and timely manner. The utility of this approach was confirmed when compared to Illumina metagenomic sequencing and isolation and characterisation of *K. pneumoniae* strain including WGS and phenotypic (*i.e*. MIC) testing. This suggests that MinION may be used in a clinical setting, potentially improving health care strategies and antibiotic stewardship for at-risk premature infants in the future.

Future work

This work has probed the applicability of MinION technology for diagnosing faecal samples from critically ill premature infants suffering from sepsis or NEC. I envisage future work where we can use this technology in a broader scale, perhaps using the PromethION technology, which allows you to profile higher number of samples (up to 48 flowcells in one run). It would be very informative to run several samples from premature infants residing in the same ward or hospital, to perform epidemiological studies to evaluate whether there are correlations among the bacterial groups at taxonomic and resistome level. Currently the MinION error rate (3% for 2D and 6% for 1D using R9.4 flowcells[245]) does not allow to distinguish between closely related alleles of AMR genes, and published work has tackled this issue by assembling the raw reads[246]. However, it is conceivable that continued improvements in Nanopore chemistry and base calling algorithms could improve raw read accuracy, to the point that intermediate assembly is not necessary prior to AMR gene identification so long read length could be then be used to give a greater insight of the location of the AMR genes (*i.e.* plasmid or genome) or discriminate against closely relates AMR gene alleles.

Final summary

Premature infants represent an at-risk patient cohort exposed to multiple potential microbiota 'damaging' post-natal factors from the first day of life; infections related to maternal health, caesarean delivery, prolonged antibiotic courses, difficulties in establishing breastfeeding, and prolonged residency in the NICUs. The gut microbiome in premature infants is characterised by overall reduced bacterial diversity, but high levels of potentially pathogenic bacteria (*e.g.* *Klebsiella*, *Escherichia* or *Staphylococcus*) and low levels of beneficial genera *Bifidobacterium*[235]. Many studies have associated a decrease in the abundance of beneficial microbes in premature infants, to a higher incidence of developing; (i) neonatal-associated bacterial infections/sepsis or necrotising enterocolitis[247, 248] (NEC), and (ii) later onset diseases in childhood, such as atopy[249] or obesity[250].

There are several ways to beneficially modify the preterm microbiome, including maximising the exposure to breastmilk, or donor breast milk if absent, or administering probiotics orally. The potential beneficial role exerted by commensal bacteria (*e.g. Bifidobacterium* and *Lactobacillus*) has been reported in several meta-analyses which have shown reductions in NEC and sepsis incidence[138, 168]. Interestingly, many of these studies highlight that only the genus *Bifidobacterium* colonises the gut long-term, and that some probiotic mixtures are not as effective as others. More information about the efficacy of probiotic strains is essential to be able to improve preterm health. In this context, the use of next generation sequencing approaches (*e.g.* 16S rRNA gene or whole-genome shotgun sequencing) can provide important insights to help guide future clinical trials at optimising probiotic therapies ahead of promoting large-scale administration.

In this study, a 16S rRNA gene sequencing pipeline was optimised to depict the complexity of the gut microbiome in premature infants with and without probiotic supplementation (Chapter 1). 16S rRNA profiling is a cost-efficient sequencing

method, that is useful in large scale projects (*e.g.* clinical trials) involving hundreds or thousands of samples. This metataxonomic profiling provides the added advantage that it can sequence samples with a very low bacterial content (*e.g.* faecal samples from premature infants), due to the PCR-mediated amplification step. This work emphasizes the importance of; (i) choosing an adequate DNA extraction method, and (ii) targeting the hypervariable region of the 16S rRNA gene which offers the most representative bacterial community profile.

The optimised 16S rRNA gene sequencing pipeline was used to assess the effectiveness of the probiotic supplementation Infloran®, a probiotic mix of *Bifidobacterium bifidum* and *Lactobacillus acidophilus,* on the preterm microbiome. This study represents the largest reported longitudinal study currently performed in premature infants (to our knowledge), and included 591 stool samples from 234 infants. This large body of work included many scientists, clinicians, and nurses, and I had the privilege to lead this work from establishing SOPs, obtaining ethical consent, engaging with health practitioners, recruiting patients, and right through to its publication. The main findings from this study (Chapter 2) were that (i) Infloran® supplementation beneficially modified the preterm microbiome by enhancing *Bifidobacterium* populations and reducing potential pathogenic bacteria, and (ii) probiotic supplementation contributed to augment the abundance of short chain fatty acids (*i.e.* acetate and lactate) in the supplemented group. The importance of this work is that it emphasises the exerting protective functional effects of microbial therapies on the preterm gut microbial communities, and contributes to provide evidence for changing clinical practice in NICUs.

In recent times, there has been an increase in our knowledge of the human gut microbiome which has allowed the characterisation of gut microbiome profiles in health and disease. This has opened the door to new preventive microbiota therapies called Live Bio-Therapeutic Products (LBPs), which are probiotics used for prevention, treatment or cure of a disease[251]. A mechanistic understanding of

how probiotic supplementation provides health benefits to premature infants is yet to be determined, such an understanding would help justify changes in clinical policies among neonatal intensive care units. An important question concerns the mechanism(s) whereby health benefits are derived from microbial supplementation. This could be elucidated by focusing on the following questions:

(i) How does *Bifidobacterium* contribute to provide colonisation resistance against other potential bacterial pathogens?

(ii) How important is the role of infant diet or prebiotics to achieve a successful colonisation of *Bifidobacterium* in the preterm gut?

Advances in the application of high-throughput shotgun metagenomic sequencing are expected to significantly contribute to answering some of these key questions. Shotgun metagenomic profiling, by virtue of sequencing the genomes of whole bacterial communities, can provide detailed information on the species/strain level, as well as identify important bacterial functional pathways such as genes involved in antimicrobial resistance. Approximately 78%-87% of premature infants receive antibiotics prophylactically within the first three days of life[210], which are administered to prevent episodes of early-sepsis. Antibiotics are well-known to influence the composition of the preterm gut microbiome and, importantly, enhance the presence of antibiotic resistance genes (ARGs), which can be detrimental when fighting bacterial infections. In the work presented here, whole-metagenome shotgun sequencing was performed on 95 stool samples from 34 premature infants (Chapter 3). The aim of this work was to evaluate the short-term effects associated with prophylactic antibiotic treatment on the preterm gut microbiome, and to determine whether this practise contributes to enhanced gut-associated antimicrobial resistance carriage. An important novelty of this study relates to the fact that several ‘sample sets´ included samples from premature infants receiving probiotic supplementation (to probe the role of this treatment regime on the carriage of AMR genes within the preterm gut microbiome), and samples from premature infants that did not receive supplementation (to explore

the impact of antibiotics alone in driving AMR over time in at-risk populations). The latter was particularly innovative as the administration of antibiotics in premature infants is so widespread, and yet this study managed to recruit 11 premature infants who did not receive antibiotic treatment. Unexpectedly, it appears that (short-term) prophylactic antibiotic treatment had only a transient effect on the preterm gut, while the NICU environment appeared to more significantly alter preterm-associated microbial communities. Another surprising finding came from comparison of the total reservoir of AMR genes between infants treated with or without antibiotics, there was no significant difference in the profile of gene categories and their abundance. Furthermore, AMR genes conferring resistance to the prophylactic antibiotic treatments were already present in all bacterial pathogens detected. These findings may have implications for preliminary guidance for recommendations of the use of prophylactic antibiotic treatment in premature infants.

The above data highlights that a more personalised antibiotic regime should be used to protect at-risk premature infants. Blood cultures are currently the "gold standard" for diagnosis of blood stream infections and characterisation of their antimicrobial susceptibility. The turnaround time of this technique is usually within 18 to 24 h[252], which highlights the necessity to use faster microbiological diagnosis techniques for at-risk patients. MinION Nanopore sequencing presents an attractive alternative to rapidly profile faecal samples from critically-ill premature infants suffering from necrotising enterocolitis or sepsis. The work performed in Chapter 4 of this thesis demonstrates that MinION technology can rapidly diagnose bacterial pathogens as well as its AMR gene content in less than 6 h from sample receipt. However, there are still a few considerations to be able to implement this approach in large-scale testing: (i) this technology should be more affordable and (ii) able to sequence a larger number of samples. PromethION from Oxford Nanopore, uses the same technology as MinION but offers the possibility of sequencing 48

samples at the same time[253]. If cost are reduced, we may soon see the use of PromethION for routine microbial diagnosis in clinical laboratories.

Overall, this multidisciplinary clinically-relevant work using high-throughput sequencing (*i.e.* 16S rRNA gene sequencing and shotgun metagenome sequencing) provides novel insights into the preterm gut microbiome in health and disease. In particular, it emphasises the protective role that probiotic supplementation plays when administered to premature infants, and the risks of using antibiotic treatment prophylactically. It also evaluates the applicability of a state-of-the-art technology such as MinION Nanopore platform for prompt microbial diagnostics. More fundamentally, this work illustrates how next-generation sequencing platforms linked to laboratory science can translate knowledge into a clinical application. Future considerations into how research institutions, pharmaceutical industry and clinicians must find ways to collaborate and integrate their knowledge is needed, to further expand and accelerate microbiome science.

Appendix 1

Number of raw read counts for shotgun and 16S rRNA gene sequencing data

Samples for Shotgun sequencing

Baby ID	Shotgun samples	Raw_read count	in MEGAN (Paired when possible)
V3J	S1_1.fastq	9,536,453	8,420,963
	S1_2.fastq	9,536,453	
AP8C	S2_1.fastq	12,382,963	10,589,359
	S2_2.fastq	12,382,963	
P29F	S3_1.fastq	10,896,924	9,622,411
	S3_2.fastq	10,896,924	

Samples for 16S rRNA sequencing using PE

Baby ID	Shotgun samples	Raw_read count	After Quality filter	in MEGAN (Paired when possible)
	V1+V2+V3			
AP1E	AP1E.27F_R1.fastq	154,509	146,552	
	AP1E.27F_R2.fastq	154,509	124,575	108,497
AP25E	AP25E.27F_R1.fastq	155,207	147,971	
	AP25E.27F_R2.fastq	155,207	122,997	270,759
AP5D	AP5D.27F_R1.fastq	101,963	97,419	
	AP5D.27F_R2.fastq	101,963	80,605	178,024
AP8C	AP8C.27F_R1.fastq	83,883	80,074	
	AP8C.27F_R2.fastq	83,883	66,301	146,371
P29F	P29F.27F_R1.fastq	100,264	95,214	
	P29F.27F_R2.fastq	100,264	75,983	171,169
P30N	P30N.27F_R1.fastq	113,827	108,139	
	P30N.27F_R2.fastq	113,827	88,764	196,897
P31B	P31B.27F_R1.fastq	104,542	100,022	
	P31B.27F_R2.fastq	104,542	83,931	183,910
P35C	P35C.27F_R1.fastq	131,266	124,227	
	P35C.27F_R2.fastq	131,266	101,482	225,702
V2A	V2A.27F_R1.fastq	104,410	99,819	
	V2A.27F_R2.fastq	104,410	80,254	180,052
V3J	V3J.27F_R1.fastq	94,344	90,100	
	V3J.27F_R2.fastq	94,344	72,869	162,932
	V4+V5			
AP1E	AP1E.530F_R1.fastq	1,079,921	1,054,856	
	AP1E.530F_R2.fastq	1,079,921	940,830	1,995,544
AP25E	AP25E.530F_R1.fastq	542,529	527,076	
	AP25E.530F_R2.fastq	542,529	444,981	967,069
AP5D	AP5D.530F_R1.fastq	754,988	737,055	
	AP5D.530F_R2.fastq	754,988	648,579	1,385,492
AP8C	AP8C.530F_R1.fastq	489,498	477,261	

	AP8C.530F_R2.fastq	489,498	415,314	892,360
P29F	P29F.530F_R1.fastq	469,124	457,067	
	P29F.530F_R2.fastq	469,124	394,279	846,302
P30N	P30N.530F_R1.fastq	576,331	564,022	
	P30N.530F_R2.fastq	576,331	493,048	1,056,787
P31B	P31B.530F_R1.fastq	423,862	413,347	
	P31B.530F_R2.fastq	423,862	363,728	776,976
P35C	P35C.530F_R1.fastq	492,809	483,177	
	P35C.530F_R2.fastq	492,809	429,850	912,978
V2A	V2A.530F_R1.fastq	679,918	662,228	
	V2A.530F_R2.fastq	679,918	574,822	1,236,324
V3J	V3J.530F_R1.fastq	418,955	408,725	
	V3J.530F_R2.fastq	418,955	356,165	763,712
	V6+V7+V8			
AP1E	AP1E.926F_R1.fastq	148,260	142,900	
	AP1E.926F_R2.fastq	148,260	117,216	260,092
AP25E	AP25E.926F_R1.fastq	139,882	135,615	
	AP25E.926F_R2.fastq	139,882	115,599	251,211
AP5D	AP5D.926F_R1.fastq	98,244	94,729	
	AP5D.926F_R2.fastq	98,244	80,025	174,754
AP8C	AP8C.926F_R1.fastq	99,080	95,770	
	AP8C.926F_R2.fastq	99,080	81,301	177,054
P29F	P29F.926F_R1.fastq	66,289	63,985	
	P29F.926F_R2.fastq	66,289	50,986	114,970
P30N	P30N.926F_R1.fastq	101,995	98,807	
	P30N.926F_R2.fastq	101,995	84,480	183,269
P31B	P31B.926F_R1.fastq	131,601	126,617	
	P31B.926F_R2.fastq	131,601	100,729	227,339
P35C	P35C.926F_R1.fastq	120,933	116,860	
	P35C.926F_R2.fastq	120,933	96,829	213,685
V2A	V2A.926F_R1.fastq	99,339	94,646	
	V2A.926F_R2.fastq	99,339	68,738	163,384
V3J	V3J.926F_R1.fastq	126,399	121,345	
	V3J.926F_R2.fastq	126,399	92,676	213,519

Samples for 16S rRNA sequencing using QIIME

Baby ID	Shotgun samples	Raw_read count	Assembly	After Quality filter	in MEGAN (Paired when possible)
	V1+V2+V3				
AP1E	AP1E.27F_R1.fastq	154,509	149,927	74,768	
	AP1E.27F_R2.fastq	154,509			73,286
AP25E	AP25E.27F_R1.fastq	155,207	149,309	66,989	
	AP25E.27F_R2.fastq	155,207			66,233
AP5D	AP5D.27F_R1.fastq	101,963	96,606	31,555	
	AP5D.27F_R2.fastq	101,963			29,589
AP8C	AP8C.27F_R1.fastq	83,883	79,707	28,356	

	AP8C.27F_R2.fastq	83,883			25,808
P29F	P29F.27F_R1.fastq	100,264	95,641	37,141	
	P29F.27F_R2.fastq	100,264			35,173
P30N	P30N.27F_R1.fastq	113,827	106,924	32,515	
	P30N.27F_R2.fastq	113,827			31,450
P31B	P31B.27F_R1.fastq	104,542	100,864	44,990	
	P31B.27F_R2.fastq	104,542			41,677
P35C	P35C.27F_R1.fastq	131,266	122,234	36,039	
	P35C.27F_R2.fastq	131,266			34,051
V2A	V2A.27F_R1.fastq	104,410	102,296	59,276	
	V2A.27F_R2.fastq	104,410			57,969
V3J	V3J.27F_R1.fastq	94,344	92,210	51,365	
	V3J.27F_R2.fastq	94,344			46,407
	V4+V5				
AP1E	AP1E.530F_R1.fastq	1,079,921	1,078,121	1,052,924	
	AP1E.530F_R2.fastq	1,079,921			1,028,760
AP25E	AP25E.530F_R1.fastq	542,529	541,235	521,658	
	AP25E.530F_R2.fastq	542,529			504,693
AP5D	AP5D.530F_R1.fastq	754,988	753,758	733,507	
	AP5D.530F_R2.fastq	754,988			715,114
AP8C	AP8C.530F_R1.fastq	489,498	488,558	474,597	
	AP8C.530F_R2.fastq	489,498			450,256
P29F	P29F.530F_R1.fastq	469,124	467,292	452,341	
	P29F.530F_R2.fastq	469,124			430,128
P30N	P30N.530F_R1.fastq	576,331	575,366	560,498	
	P30N.530F_R2.fastq	576,331			546,603
P31B	P31B.530F_R1.fastq	423,862	423,043	412,194	
	P31B.530F_R2.fastq	423,862			396,517
P35C	P35C.530F_R1.fastq	492,809	491,243	479,886	
	P35C.530F_R2.fastq	492,809			462,862
V2A	V2A.530F_R1.fastq	679,918	678,285	658,393	
	V2A.530F_R2.fastq	679,918			635,528
V3J	V3J.530F_R1.fastq	418,955	417,757	405,838	
	V3J.530F_R2.fastq	418,955			384,789
	V6+V7+V8				
AP1E	AP1E.926F_R1.fastq	148,260	146,736	99,832	
	AP1E.926F_R2.fastq	148,260			94,851
AP25E	AP25E.926F_R1.fastq	139,882	138,988	103,182	
	AP25E.926F_R2.fastq	139,882			101,554
AP5D	AP5D.926F_R1.fastq	98,244	97,503	69,336	
	AP5D.926F_R2.fastq	98,244			62,979
AP8C	AP8C.926F_R1.fastq	99,080	98,425	71,562	
	AP8C.926F_R2.fastq	99,080			59,316
P29F	P29F.926F_R1.fastq	66,289	65,658	41,768	

	P29F.926F_R2.fastq	66,289			36,853
P30N	P30N.926F_R1.fastq	101,995	101,345	74,108	
	P30N.926F_R2.fastq	101,995			69,663
P31B	P31B.926F_R1.fastq	131,601	129,856	76,795	
	P31B.926F_R2.fastq	131,601			31,552
P35C	P35C.926F_R1.fastq	120,933	119,842	82,133	
	P35C.926F_R2.fastq	120,933			61,947
V2A	V2A.926F_R1.fastq	99,339	97,790	48,016	
	V2A.926F_R2.fastq	99,339			45,181
V3J	V3J.926F_R1.fastq	126,399	124,563	70,010	
	V3J.926F_R2.fastq	126,399			49,171

Appendix 2

Metadata all infants samples

sample_ID	sex	delivery	birthweight	gestageweeks	antibiotics	diettype	daysofprobiotics	hospital	Treatment
AP10A	F	C	710	26	y	BM	NA	Addenbrookes	Control
AP10E	F	C	710	26	n	BM	NA	Addenbrookes	Control
AP11A	M	C	740	25	n	BM	NA	Addenbrookes	Control
AP12B	F	C	785	25	n	BM_F	NA	Addenbrookes	Control
AP12D	F	C	785	25	n	BM_F	NA	Addenbrookes	Control
AP12G	F	C	785	25	n	BM	NA	Addenbrookes	Control
AP12I	F	C	785	25	n	BM	NA	Addenbrookes	Control
AP14A	F	V	980	27	n	BM	NA	Addenbrookes	Control
AP16B	F	V	1190	34	y	BM	NA	Addenbrookes	Control
AP16C	F	V	1190	34	n	BM	NA	Addenbrookes	Control
AP17A	M	C	1400	30	y	BM	NA	Addenbrookes	Control
AP19A	M	V	1180	27	y	BM	NA	Addenbrookes	Control
AP19C	M	V	1180	27	n	BM	NA	Addenbrookes	Control
AP19E	M	V	1180	27	y	BM	NA	Addenbrookes	Control
AP1D	F	C	830	25	y	BM	NA	Addenbrookes	Control
AP1E	F	C	830	25	n	BM	NA	Addenbrookes	Control
AP1F	F	C	830	25	n	BM	NA	Addenbrookes	Control
AP1J	F	C	830	25	n	BM	NA	Addenbrookes	Control
AP20B	F	V	710	25	y	BM	NA	Addenbrookes	Control
AP20D	F	V	710	25	n	BM	NA	Addenbrookes	Control
AP21	F	V	790	25	n	BM	NA	Addenbrookes	Control
AP21C	F	V	790	25	n	BM	NA	Addenbrookes	Control
AP21D	F	V	790	25	y	BM	NA	Addenbrookes	Control
AP21E	F	V	790	25	n	BM	NA	Addenbrookes	Control
AP22B	F	C	605	25	n	BM	NA	Addenbrookes	Control
AP22D	F	C	605	25	y	BM	NA	Addenbrookes	Control
AP23H	F	C	700	25	n	BM	NA	Addenbrookes	Control
AP24A	F	V	785	28	NA	NA	NA	Addenbrookes	Control
AP25C	F	C	785	25	n	BM_DBM	NA	Addenbrookes	Control
AP25F	F	C	785	25	n	BM	NA	Addenbrookes	Control
AP26A	F	C	830	27	n	BM	NA	Addenbrookes	Control
AP28A	M	C	985	26	n	BM	NA	Addenbrookes	Control
AP28C	M	C	985	26	n	BM	NA	Addenbrookes	Control
AP2D	F	V	600	24	n	BM	NA	Addenbrookes	Control
AP2F	F	V	600	24	n	BM	NA	Addenbrookes	Control
AP2J	F	V	600	24	n	BM	NA	Addenbrookes	Control
AP2K	F	V	600	24	n	BM	NA	Addenbrookes	Control
AP3B	F	V	1520	29	y	BM	NA	Addenbrookes	Control
AP3C	F	V	1520	29	n	BM	NA	Addenbrookes	Control
AP3D	F	V	1520	29	n	BM_F	NA	Addenbrookes	Control
AP3E	F	V	1520	29	n	BM_F	NA	Addenbrookes	Control
AP5A	F	V	800	25	y	BM	NA	Addenbrookes	Control
AP5B	F	V	800	25	y	BM	NA	Addenbrookes	Control
AP5H	F	V	800	25	n	BM_F	NA	Addenbrookes	Control
AP5I	F	V	800	25	n	BM_F	NA	Addenbrookes	Control

AP6B	M	C	570	28	y	BM	NA	Addenbrookes	Control
AP6F	M	C	570	28	n	BM	NA	Addenbrookes	Control
AP6I	M	C	570	28	n	BM	NA	Addenbrookes	Control
AP6O	M	C	570	28	n	BM	NA	Addenbrookes	Control
AP8B	M	V	576	23	n	BM	NA	Addenbrookes	Control
AP9C	F	C	1020	29	n	BM	NA	Addenbrookes	Control
M100.1	M	V	1380	30	n	BM_DBM	NA	Imperial	Control
M100.2	M	V	1380	30	n	BM_DBM	NA	Imperial	Control
M101.1	M	C	1320	30	n	BM_DBM	NA	Imperial	Control
M101.2	M	C	1320	30	n	BM_DBM	NA	Imperial	Control
M103.1	F	C	1210	28	n	BM	NA	Imperial	Control
M103.2	F	C	1210	28	n	BM_F	NA	Imperial	Control
M111.1	F	V	1380	30	n	BM_DBM	NA	Imperial	Control
M111.2	F	V	1380	30	n	BM_DBM	NA	Imperial	Control
M114.1	F	V	1535	30	y	BM	NA	Imperial	Control
M114.2	F	V	1535	30	n	BM	NA	Imperial	Control
M115.2	F	C	1476	30	n	BM_DBM	NA	Imperial	Control
M118.1	F	V	1190	30	n	BM_DBM	NA	Imperial	Control
M118.2	F	V	1190	30	n	BM_F	NA	Imperial	Control
M121.1	F	C	1270	29	n	BM_DBM	NA	Imperial	Control
M121.2	F	C	1270	29	n	BM_DBM	NA	Imperial	Control
M121.3	F	C	1270	29	n	BM	NA	Imperial	Control
M121.4	F	C	1270	29	n	BM	NA	Imperial	Control
M121.5	F	C	1270	29	n	BM	NA	Imperial	Control
M122.1	M	V	1005	27	n	BM	NA	Imperial	Control
M122.3	M	V	1005	27	n	BM	NA	Imperial	Control
M122.6	M	V	1005	27	n	BM	NA	Imperial	Control
M123.1	M	V	1950	31	y	BM	NA	Imperial	Control
M123.2	M	V	1950	31	n	BM	NA	Imperial	Control
M124.1	F	C	1567	31	n	F	NA	Imperial	Control
M125.2	M	C	1230	31	n	F	NA	Imperial	Control
M126.1	M	C	1470	29	n	BM	NA	Imperial	Control
M126.3	M	C	1470	29	n	BM	NA	Imperial	Control
M126.4	M	C	1470	29	n	BM	NA	Imperial	Control
M127.1	F	C	1080	30	y	BM_DBM	NA	Imperial	Control
M127.2	F	C	1080	30	n	BM_DBM	NA	Imperial	Control
M128.1	M	C	1320	30	n	BM_DBM	NA	Imperial	Control
M129.1	F	C	800	29	n	BM_DBM	NA	Imperial	Control
M129.2	F	C	800	29	n	BM_F	NA	Imperial	Control
M130.1	M	C	975	31	n	BM_DBM	NA	Imperial	Control
M131.1	F	C	1640	30	n	BM_F	NA	Imperial	Control
M131.2	F	C	1640	30	n	BM_F	NA	Imperial	Control
M134.1	F	C	1371	31	n	BM_F	NA	Imperial	Control
M135.1	F	C	1750	31	n	BM_F	NA	Imperial	Control
M135.2	F	C	1750	31	n	BM_F	NA	Imperial	Control
M15.1	M	C	1370	30	n	BM	NA	Imperial	Control
M15.2	M	C	1370	30	n	BM	NA	Imperial	Control
M15.3	M	C	1370	30	n	BM	NA	Imperial	Control
M15.4	M	C	1370	30	n	BM	NA	Imperial	Control
M15.5	M	C	1370	30	n	BM	NA	Imperial	Control

M15.6	M	C	1370	30	n	BM	NA	Imperial	Control
M16.1	M	C	1340	29	n	BM	NA	Imperial	Control
M16.2	M	C	1340	29	n	BM	NA	Imperial	Control
M17.1	M	C	1200	29	n	BM	NA	Imperial	Control
M17.2	M	C	1200	29	n	BM	NA	Imperial	Control
M18.1	F	C	1100	29	n	BM	NA	Imperial	Control
M18.2	F	C	1100	29	n	BM	NA	Imperial	Control
M22.1	M	C	1260	31	n	BM	NA	Imperial	Control
M22.2	M	C	1260	31	n	BM_F	NA	Imperial	Control
M26.1	F	C	1100	28	n	BM_F	NA	Imperial	Control
M26.2	F	C	1100	28	n	BM_F	NA	Imperial	Control
M27.1	F	C	923	31	y	BM	NA	Imperial	Control
M27.2	F	C	923	31	n	BM	NA	Imperial	Control
M27.3	F	C	923	31	n	BM	NA	Imperial	Control
M29.1	F	C	1100	28	n	BM	NA	Imperial	Control
M29.2	F	C	1100	28	n	BM	NA	Imperial	Control
M36.1	F	V	1780	31	y	BM	NA	Imperial	Control
M36.2	F	V	1780	31	y	BM	NA	Imperial	Control
M36.3	F	V	1780	31	n	BM	NA	Imperial	Control
M37.1	F	V	1320	31	n	BM_F	NA	Imperial	Control
M37.2	F	V	1320	31	n	BM_F	NA	Imperial	Control
M37.3	F	V	1320	31	n	BM_F	NA	Imperial	Control
M37.4	F	V	1320	31	n	BM_F	NA	Imperial	Control
M37.5	F	V	1320	31	n	BM_F	NA	Imperial	Control
M38.1	M	C	1110	29	n	BM_DBM	NA	Imperial	Control
M38.2	M	C	1110	29	n	BM_DBM	NA	Imperial	Control
M39.1	F	V	1330	28	n	BM_DBM	NA	Imperial	Control
M39.2	F	V	1330	28	n	BM_DBM	NA	Imperial	Control
M39.3	F	V	1330	28	n	BM	NA	Imperial	Control
M39.4	F	V	1330	28	n	BM	NA	Imperial	Control
M54.1	M	C	1000	28	n	BM_DBM	NA	Imperial	Control
M54.2	M	C	1000	28	n	BM_DBM	NA	Imperial	Control
M55.1	F	C	1100	28	n	BM_DBM	NA	Imperial	Control
M55.2	F	C	1100	28	n	BM_DBM	NA	Imperial	Control
M56.1	F	C	950	28	n	BM_DBM	NA	Imperial	Control
M56.2	F	C	950	28	n	BM_DBM	NA	Imperial	Control
M58.1	M	C	980	28	n	BM_DBM	NA	Imperial	Control
M58.2	M	C	980	28	n	BM	NA	Imperial	Control
M59.1	F	C	1000	28	n	BM_DBM	NA	Imperial	Control
M59.2	F	C	1000	28	n	BM	NA	Imperial	Control
M60.1	F	C	950	28	n	BM_DBM	NA	Imperial	Control
M60.2	F	C	950	28	n	BM	NA	Imperial	Control
M62.1	M	C	1840	31	n	F	NA	Imperial	Control
M62.2	M	C	1840	31	n	BM_F	NA	Imperial	Control
M63.1	M	V	1830	31	y	BM_F	NA	Imperial	Control
M63.2	M	V	1830	31	n	BM_F	NA	Imperial	Control
M65.1	M	V	1100	28	y	BM_F	NA	Imperial	Control
M65.2	M	V	1100	28	y	BM_F	NA	Imperial	Control
M7.1	M	V	1460	31	n	BM_F	NA	Imperial	Control
M7.2	M	V	1460	31	n	BM_F	NA	Imperial	Control

M7.3	M	V	1460	31	n	BM_F	NA	Imperial	Control
M70.1	F	V	1720	31	n	BM_F	NA	Imperial	Control
M70.2	F	V	1720	31	n	BM_F	NA	Imperial	Control
M73.1	M	C	1730	31	n	BM	NA	Imperial	Control
M73.2	M	C	1730	31	n	BM	NA	Imperial	Control
M77.1	F	C	1300	30	n	BM_DBM	NA	Imperial	Control
M77.2	F	C	1300	30	n	BM_DBM	NA	Imperial	Control
M77.3	F	C	1300	30	n	BM	NA	Imperial	Control
M78.1	M	V	1360	29	y	BM_F	NA	Imperial	Control
M78.2	M	V	1360	29	n	BM_F	NA	Imperial	Control
M79.1	F	C	1260	30	n	BM_DBM	NA	Imperial	Control
M79.2	F	C	1260	30	n	BM	NA	Imperial	Control
M82.1	M	C	1100	30	y	BM_DBM	NA	Imperial	Control
M82.2	M	C	1100	30	n	BM_DBM	NA	Imperial	Control
M84.1	M	C	1630	30	y	BM_DBM	NA	Imperial	Control
M84.2	M	C	1630	30	n	BM_F	NA	Imperial	Control
M85.1	M	C	1350	30	n	BM_F	NA	Imperial	Control
M85.2	M	C	1350	30	n	BM_DBM	NA	Imperial	Control
M86.1	F	C	1090	30	n	BM	NA	Imperial	Control
M86.2	F	C	1090	30	n	BM_DBM	NA	Imperial	Control
M87.1	M	C	920	27	n	BM_DBM	NA	Imperial	Control
M87.2	M	C	1081	27	n	BM_DBM	NA	Imperial	Control
M87.3	M	C	1080	27	n	BM_DBM	NA	Imperial	Control
M88.1	M	C	920	27	n	BM_DBM	NA	Imperial	Control
M88.2	M	C	920	27	n	BM_DBM	NA	Imperial	Control
M88.3	M	C	920	27	n	BM_DBM	NA	Imperial	Control
M88.4	M	C	920	27	n	BM_DBM	NA	Imperial	Control
M89.1	F	C	1215	30	y	BM_DBM	NA	Imperial	Control
M89.2	F	C	1215	30	n	BM_DBM	NA	Imperial	Control
M90.1	F	C	1275	30	n	BM_DBM	NA	Imperial	Control
M90.2	F	C	1275	30	n	BM_DBM	NA	Imperial	Control
M91.1	F	V	1090	29	y	BM_DBM	NA	Imperial	Control
M91.2	F	V	1090	29	n	BM_DBM	NA	Imperial	Control
M91.3	F	V	1090	29	n	BM_DBM	NA	Imperial	Control
M92.1	M	V	1270	28	n	BM_DBM	NA	Imperial	Control
M92.2	M	V	1270	28	n	BM	NA	Imperial	Control
M92.3	M	V	1270	28	n	BM	NA	Imperial	Control
M93.1	M	V	978	29	y	BM_DBM	NA	Imperial	Control
M93.2	M	V	978	29	n	BM_DBM	NA	Imperial	Control
M93.3	M	V	978	29	n	BM	NA	Imperial	Control
M94.1	M	V	1040	28	y	BM_DBM	NA	Imperial	Control
M94.2	M	V	1040	28	n	BM_DBM	NA	Imperial	Control
M94.3	M	V	1040	28	n	BM_F	NA	Imperial	Control
M95.1	M	V	1410	29	y	BM_DBM	NA	Imperial	Control
M95.2	M	V	1410	29	n	BM	NA	Imperial	Control
M95.3	M	V	1410	29	n	BM	NA	Imperial	Control
M95.4	M	V	1410	29	n	BM	NA	Imperial	Control
P100A	M	C	750	28	y	BM	4	NNUH	Probiotic
P100H	M	C	750	28	n	BM	14	NNUH	Probiotic
P102F	F	V	1930	31	y	BM	7	NNUH	Probiotic

P102J	F	V	1930	31	n	BM	11	NNUH	Probiotic
P103A	F	C	1065	30	n	BM	7	NNUH	Probiotic
P103G	F	C	1065	30	n	BM	11	NNUH	Probiotic
P103O	F	C	1065	30	n	BM	20	NNUH	Probiotic
P104E	M	C	1200	27	y	BM	10	NNUH	Probiotic
P104K	M	C	1200	27	n	BM_F	39	NNUH	Probiotic
P104L	M	C	1200	27	n	BM	39	NNUH	Probiotic
P105E	M	V	1410	30	n	BM	9	NNUH	Probiotic
P105I	M	V	1410	30	n	BM	14	NNUH	Probiotic
P105L	M	V	1410	30	y	BM	24	NNUH	Probiotic
P106C	F	V	1402	30	n	BM	4	NNUH	Probiotic
P106H	F	V	1402	30	y	BM	10	NNUH	Probiotic
P106M	F	V	1402	30	n	BM	31	NNUH	Probiotic
P107B	F	C	1515	33	n	BM	0	NNUH	Probiotic
P108C	F	V	549	24	y	BM	17	NNUH	Probiotic
P108K	F	V	549	24	n	BM	42	NNUH	Probiotic
P108M	F	V	549	24	n	BM	63	NNUH	Probiotic
P109D	F	C	1315	31	n	BM_DBM	17	NNUH	Probiotic
P10N	M	V	1050	26	n	BM	9	NNUH	Probiotic
P10T	M	V	1050	26	NA	BM	18	NNUH	Probiotic
P10V	M	V	1050	26	NA	NA	18	NNUH	Probiotic
P110B	F	V	567	23	n	BM	14	NNUH	Probiotic
P110P	F	V	567	23	n	BM	44	NNUH	Probiotic
P110Q	F	V	567	23	y	BM	55	NNUH	Probiotic
P111C	F	V	602	24	y	BM	11	NNUH	Probiotic
P111D	F	V	602	24	y	BM	12	NNUH	Probiotic
P111G	F	V	602	24	n	BM	37	NNUH	Probiotic
P111H	F	V	602	24	n	BM	39	NNUH	Probiotic
P111O	F	V	602	24	n	BM	60	NNUH	Probiotic
P112C	M	C	1151	29	y	BM	6	NNUH	Probiotic
P112H	M	C	1151	29	n	BM	14	NNUH	Probiotic
P112N	M	C	1151	29	n	BM	35	NNUH	Probiotic
P113A	F	V	1385	29	n	F	17	NNUH	Probiotic
P113F	F	V	1385	29	n	F	36	NNUH	Probiotic
P114B	F	V	1260	29	n	F	18	NNUH	Probiotic
P114H	F	V	1260	29	n	F	42	NNUH	Probiotic
P115C	F	C	554	25	y	BM	14	NNUH	Probiotic
P115M	F	C	554	25	n	DBM	41	NNUH	Probiotic
P116F	M	V	685	24	n	BM	11	NNUH	Probiotic
P117B	F	V	930	26	y	BM	11	NNUH	Probiotic
P118E	M	C	1485	33	y	BM	10	NNUH	Probiotic
P118G	M	C	1485	33	n	BM	13	NNUH	Probiotic
P119B	F	V	1420	33	n	BM	6	NNUH	Probiotic
P119D	F	V	1420	33	n	NA	6	NNUH	Probiotic
P11F	M	C	1640	30	y	BM	18	NNUH	Probiotic
P120B	M	C	624	27	n	BM	9	NNUH	Probiotic
P120K	M	C	624	27	y	BM	34	NNUH	Probiotic
P12A	M	V	1000	26	y	BM	2	NNUH	Probiotic
P12F	M	V	1000	26	n	BM	10	NNUH	Probiotic
P12K	M	V	1000	26	n	BM	31	NNUH	Probiotic

P12M	M	V	1000	26	n	BM	31	NNUH	Probiotic
P13A	F	C	700	25	y	BM	2	NNUH	Probiotic
P14C	F	V	1400	32	y	BM_F	4	NNUH	Probiotic
P14F	F	V	1400	32	n	BM_F	10	NNUH	Probiotic
P15G	F	V	1520	32	n	BM_F	6	NNUH	Probiotic
P15I	F	V	1520	32	n	BM_F	9	NNUH	Probiotic
P16D	F	V	707	24	y	BM	6	NNUH	Probiotic
P16J	F	V	707	24	y	BM	13	NNUH	Probiotic
P16S	F	V	707	24	n	BM	46	NNUH	Probiotic
P16V	F	V	707	24	n	BM	46	NNUH	Probiotic
P17A	F	V	774	24	y	BM	4	NNUH	Probiotic
P17B	F	V	774	24	y	BM	6	NNUH	Probiotic
P18D	F	V	750	26	n	BM_DBM	17	NNUH	Probiotic
P18H	F	V	750	26	n	BM_DBM	18	NNUH	Probiotic
P18I	F	V	750	26	n	BM_DBM	18	NNUH	Probiotic
P18K	F	V	750	26	n	BM_DBM	18	NNUH	Probiotic
P19K	M	V	780	24	y	BM	9	NNUH	Probiotic
P19S	M	V	780	24	n	NA	9	NNUH	Probiotic
P20B	M	V	1000	29	n	BM	4	NNUH	Probiotic
P20H	M	V	1000	29	n	BM	10	NNUH	Probiotic
P20O	M	V	1000	29	n	BM	39	NNUH	Probiotic
P20Q	M	V	1000	29	n	BM	39	NNUH	Probiotic
P21A	F	V	1209	29	n	BM	5	NNUH	Probiotic
P21D	F	V	1209	29	n	BM	7	NNUH	Probiotic
P21L	F	V	1209	29	n	BM	14	NNUH	Probiotic
P21U	F	V	1209	29	n	BM	44	NNUH	Probiotic
P22	M	C	1443	31	y	BM	5	NNUH	Probiotic
P22A	M	C	1443	31	n	BM	6	NNUH	Probiotic
P23C	M	C	975	31	n	BM	6	NNUH	Probiotic
P23D	M	C	975	31	n	BM	6	NNUH	Probiotic
P24B	F	C	1215	31	y	BM	5	NNUH	Probiotic
P24G	F	C	1215	31	n	BM	11	NNUH	Probiotic
P25O	F	V	860	25	y	BM_DBM	12	NNUH	Probiotic
P25S	F	V	860	25	n	NA	58	NNUH	Probiotic
P27F	M	C	1229	32	y	BM	7	NNUH	Probiotic
P27J	M	C	1229	32	y	BM	12	NNUH	Probiotic
P28A	M	V	780	25	y	BM_DBM	0	NNUH	Probiotic
P28G	M	V	780	25	y	BM_DBM	6	NNUH	Probiotic
P28R	M	V	780	25	n	BM_DBM	20	NNUH	Probiotic
P29C	M	V	935	26	y	BM	1	NNUH	Probiotic
P29G	M	V	935	26	y	BM	6	NNUH	Probiotic
P29L	M	V	935	26	n	BM	39	NNUH	Probiotic
P31Y	F	V	605	23	n	BM	85	NNUH	Probiotic
P32C	F	C	775	26	y	BM	3	NNUH	Probiotic
P32H	F	C	775	26	y	BM	10	NNUH	Probiotic
P32O	F	C	775	26	n	BM	53	NNUH	Probiotic
P33	M	C	1170	28	y	BM	4	NNUH	Probiotic
P33E	M	C	1170	28	n	BM	13	NNUH	Probiotic
P33J	M	C	1170	28	n	BM	39	NNUH	Probiotic
P35E	M	V	565	23	y	BM	14	NNUH	Probiotic

P35M	M	V	565	23	y	BM	39	NNUH	Probiotic
P35P	M	V	565	23	n	BM	52	NNUH	Probiotic
P37C	M	C	1050	32	y	BM	11	NNUH	Probiotic
P38E	F	C	950	33	n	BM	11	NNUH	Probiotic
P38K.1	F	C	950	33	n	NA	48	NNUH	Probiotic
P38K.2	F	C	950	33	n	NA	48	NNUH	Probiotic
P39C	M	C	1129	27	n	BM	13	NNUH	Probiotic
P39D	M	C	1129	27	n	BM	13	NNUH	Probiotic
P39K	M	C	1129	27	n	NA	53	NNUH	Probiotic
P40H	M	C	1039	27	y	BM	12	NNUH	Probiotic
P40S	M	C	1039	27	n	BM_F	42	NNUH	Probiotic
P40V	M	C	1039	27	y	F	53	NNUH	Probiotic
P41	M	C	795	25	y	BM	2	NNUH	Probiotic
P41D	M	C	795	25	n	BM	8	NNUH	Probiotic
P42H	F	V	1384	29	n	DBM	20	NNUH	Probiotic
P42I	F	V	1384	29	n	DBM	31	NNUH	Probiotic
P45C	M	V	1372	30	n	F	6	NNUH	Probiotic
P45D	M	V	1372	30	n	F	7	NNUH	Probiotic
P45H	M	V	1372	30	n	F	12	NNUH	Probiotic
P46	M	V	1700	30	n	BM_F	6	NNUH	Probiotic
P46E	M	V	1700	30	n	BM_F	11	NNUH	Probiotic
P47F	F	C	1485	32	y	BM_F	4	NNUH	Probiotic
P47M	F	C	1485	32	n	BM_F	9	NNUH	Probiotic
P48C	M	V	1425	30	y	BM_F	6	NNUH	Probiotic
P48K	M	V	1425	30	n	F	9	NNUH	Probiotic
P48P	M	V	1425	30	n	BM_F	22	NNUH	Probiotic
P49D	M	V	1785	33	y	F	0	NNUH	Probiotic
P50F	F	V	935	25	n	BM	10	NNUH	Probiotic
P50H	F	V	935	25	n	BM	15	NNUH	Probiotic
P50L	F	V	935	25	n	BM	38	NNUH	Probiotic
P51E	F	V	1434	31	y	BM	1	NNUH	Probiotic
P51J	F	V	1434	31	n	BM	9	NNUH	Probiotic
P51K	F	V	1434	31	n	BM	11	NNUH	Probiotic
P51O	F	V	1434	31	n	BM	20	NNUH	Probiotic
P51P	F	V	1434	31	n	BM_F	NA	NNUH	Probiotic
P52F	M	V	900	27	y	BM	7	NNUH	Probiotic
P52H	M	V	900	27	n	NA	39	NNUH	Probiotic
P53D	M	V	1250	27	y	BM_DBM	3	NNUH	Probiotic
P53G	M	V	1250	27	n	BM_DBM	13	NNUH	Probiotic
P53K	M	V	1250	27	y	BM_DBM	19	NNUH	Probiotic
P53L	M	V	1250	27	n	BM_F	40	NNUH	Probiotic
P53O	M	V	1250	27	n	NA	40	NNUH	Probiotic
P55B	F	C	749	29	y	BM	8	NNUH	Probiotic
P55D	F	C	749	29	n	BM	12	NNUH	Probiotic
P55E	F	C	749	29	n	BM	13	NNUH	Probiotic
P55L	F	C	749	29	n	BM	39	NNUH	Probiotic
P55N	F	C	749	29	n	BM	44	NNUH	Probiotic
P56A	M	C	1110	27	y	BM	4	NNUH	Probiotic
P56D	M	C	1110	27	n	BM	10	NNUH	Probiotic
P56H	M	C	1110	27	n	BM	41	NNUH	Probiotic

P57E	F	C	900	26	n	BM	6	NNUH	Probiotic
P57M	F	C	900	26	n	BM	13	NNUH	Probiotic
P57O	F	C	900	26	n	BM	19	NNUH	Probiotic
P58H	F	C	1010	26	y	BM	3	NNUH	Probiotic
P58K	F	C	1010	26	n	BM	9	NNUH	Probiotic
P59E	F	C	784	32	n	BM	31	NNUH	Probiotic
P60F	F	C	1230	33	n	BM	3	NNUH	Probiotic
P60I	F	C	1230	33	n	BM	7	NNUH	Probiotic
P60K	F	C	1230	33	n	BM	13	NNUH	Probiotic
P61C	M	V	562	24	y	NA	12	NNUH	Probiotic
P61F	M	V	562	24	n	BM	39	NNUH	Probiotic
P61N	M	V	562	24	n	BM	64	NNUH	Probiotic
P62A	F	C	940	30	n	BM	2	NNUH	Probiotic
P62F	F	C	940	30	n	BM	8	NNUH	Probiotic
P62H	F	C	940	30	n	BM_F	13	NNUH	Probiotic
P62L	F	C	940	30	n	F	24	NNUH	Probiotic
P63C	M	V	1374	30	n	BM	6	NNUH	Probiotic
P63I	M	V	1374	30	n	BM	11	NNUH	Probiotic
P63N	M	V	1374	30	n	F	28	NNUH	Probiotic
P64E	M	V	992	30	n	BM_DBM	7	NNUH	Probiotic
P64G	M	V	992	30	n	BM_DBM	11	NNUH	Probiotic
P64H	M	V	992	30	n	F	14	NNUH	Probiotic
P64M	M	V	992	30	n	F	28	NNUH	Probiotic
P64O	M	V	992	30	n	F	28	NNUH	Probiotic
P65C	F	V	1477	30	n	BM_DBM	6	NNUH	Probiotic
P65H	F	V	1477	30	n	F	11	NNUH	Probiotic
P65J	F	V	1477	30	n	F	13	NNUH	Probiotic
P65N	F	V	1477	30	n	F	28	NNUH	Probiotic
P65Q	F	V	670	30	n	F	28	NNUH	Probiotic
P66C	F	C	670	31	y	BM	0	NNUH	Probiotic
P66D	F	C	670	31	y	BM	10	NNUH	Probiotic
P66J	F	C	670	31	n	F	23	NNUH	Probiotic
P66M	F	C	670	31	n	F	23	NNUH	Probiotic
P67F	F	C	800	25	n	BM_F	14	NNUH	Probiotic
P67J	F	C	800	25	y	BM	36	NNUH	Probiotic
P67K	F	C	800	25	n	NA	37	NNUH	Probiotic
P68B	M	C	858	25	y	BM	9	NNUH	Probiotic
P68E	M	C	858	25	n	BM	17	NNUH	Probiotic
P68I.1	M	C	858	25	n	BM	37	NNUH	Probiotic
P68I.2	M	C	858	25	n	BM	37	NNUH	Probiotic
P69A	M	C	1520	31	n	BM	5	NNUH	Probiotic
P69C	M	C	1520	31	n	BM	8	NNUH	Probiotic
P69E	M	C	1520	31	n	BM	11	NNUH	Probiotic
P69I	M	C	1520	31	n	NA	18	NNUH	Probiotic
P70H	M	C	1255	30	n	BM	12	NNUH	Probiotic
P70M	M	C	1255	30	n	F	28	NNUH	Probiotic
P71G	F	C	1065	31	n	BM	8	NNUH	Probiotic
P71N	F	C	1065	31	n	NA	17	NNUH	Probiotic
P71O	F	C	1065	31	n	NA	17	NNUH	Probiotic
P72D	M	C	1454	30	y	BM_F	3	NNUH	Probiotic

P72G	M	C	1454	30	n	BM_F	7	NNUH	Probiotic
P73C	F	C	1504	30	n	BM	4	NNUH	Probiotic
P73D	F	C	1504	30	n	BM	9	NNUH	Probiotic
P74B	M	C	1382	31	y	BM	2	NNUH	Probiotic
P74D	M	C	1382	31	n	BM	7	NNUH	Probiotic
P74F	M	C	1382	31	n	BM	13	NNUH	Probiotic
P74H	M	C	1382	31	n	NA	19	NNUH	Probiotic
P75B	M	V	1262	27	y	BM	5	NNUH	Probiotic
P75D	M	V	1262	27	n	BM	12	NNUH	Probiotic
P75K	M	V	1262	27	n	BM	44	NNUH	Probiotic
P75M	M	V	1262	27	n	BM_DBM	44	NNUH	Probiotic
P76B	M	V	1409	27	y	BM	3	NNUH	Probiotic
P76E	M	V	1409	27	n	BM	5	NNUH	Probiotic
P76H	M	V	1409	27	n	BM	11	NNUH	Probiotic
P76K	M	V	1409	27	n	BM	23	NNUH	Probiotic
P76N	M	V	1409	27	n	BM	42	NNUH	Probiotic
P76P	M	V	1409	27	n	BM	42	NNUH	Probiotic
P77H	F	C	1110	30	n	BM	8	NNUH	Probiotic
P77O	F	C	1110	30	n	NA	17	NNUH	Probiotic
P78C	F	C	1236	32	y	BM_DBM	4	NNUH	Probiotic
P78G	F	C	1236	32	n	F	10	NNUH	Probiotic
P79B	M	C	1544	30	y	BM_DBM	2	NNUH	Probiotic
P79D	M	C	1544	30	n	F	7	NNUH	Probiotic
P79G	M	C	1544	30	n	F	13	NNUH	Probiotic
P79J	M	C	1544	30	n	NA	24	NNUH	Probiotic
P80A	M	V	831	25	y	BM	3	NNUH	Probiotic
P80B	M	V	831	25	n	BM	11	NNUH	Probiotic
P80D	M	V	831	25	n	BM	13	NNUH	Probiotic
P80J	M	V	831	25	n	BM	49	NNUH	Probiotic
P80L	M	V	831	25	n	BM	62	NNUH	Probiotic
P81B	M	C	960	26	n	BM	6	NNUH	Probiotic
P81F	M	C	960	26	n	BM	13	NNUH	Probiotic
P81L	M	C	960	26	n	BM	39	NNUH	Probiotic
P81N	M	C	960	26	n	NA	48	NNUH	Probiotic
P82A	M	C	920	28	y	BM	2	NNUH	Probiotic
P82F	M	C	920	28	n	BM	9	NNUH	Probiotic
P82O	M	C	920	28	n	F	39	NNUH	Probiotic
P82R	M	C	920	28	y	F	39	NNUH	Probiotic
P83B	F	C	1030	32	y	BM	5	NNUH	Probiotic
P83D	F	C	1030	32	n	BM	9	NNUH	Probiotic
P83G	F	C	1030	32	n	BM	12	NNUH	Probiotic
P85B	M	C	1224	30	y	BM	7	NNUH	Probiotic
P85D	M	C	1224	30	n	BM	9	NNUH	Probiotic
P85E	M	C	1224	30	n	BM	11	NNUH	Probiotic
P85H	M	C	1224	30	n	NA	15	NNUH	Probiotic
P86E.1	M	C	1400	31	n	BM	9	NNUH	Probiotic
P86E.2	M	C	1400	31	n	BM	9	NNUH	Probiotic
P86J	M	C	1400	31	n	BM	18	NNUH	Probiotic
P86L	M	C	1400	31	n	NA	18	NNUH	Probiotic
P91E	F	V	700	24	y	BM	14	NNUH	Probiotic

P91K	F	V	700	24	n	BM	40	NNUH	Probiotic
P91N	F	V	700	24	n	BM	58	NNUH	Probiotic
P95C	F	C	1100	30	n	BM	7	NNUH	Probiotic
P95G	F	C	1100	30	n	BM	14	NNUH	Probiotic
P95N	F	C	1100	30	n	BM	27	NNUH	Probiotic
P96D	M	V	694	25	n	BM	12	NNUH	Probiotic
P96G	M	V	694	25	n	BM	20	NNUH	Probiotic
P96M	M	V	694	25	n	NA	27	NNUH	Probiotic
P96N	M	V	694	25	n	NA	27	NNUH	Probiotic
P97O	M	V	530	24	n	BM	30	NNUH	Probiotic
P97T	M	V	530	24	n	BM	49	NNUH	Probiotic
P98C	M	V	685	24	n	BM	14	NNUH	Probiotic
P98J	M	V	685	24	y	BM	32	NNUH	Probiotic
P98N	M	V	685	24	n	BM	50	NNUH	Probiotic
P9G	M	V	964	26	y	BM_DBM	13	NNUH	Probiotic
P9P	M	V	964	26	n	BM_DBM	21	NNUH	Probiotic
Q1.1	M	C	1125	26	n	BM_DBM	NA	Imperial	Control
Q1.2	M	C	1125	26	n	BM_DBM	NA	Imperial	Control
Q1.3	M	C	1125	26	n	BM	NA	Imperial	Control
Q101.1	M	C	1035	28	n	BM	NA	Imperial	Control
Q101.2	M	C	1035	28	n	BM	NA	Imperial	Control
Q101.3	M	C	1035	28	n	BM	NA	Imperial	Control
Q101.4	M	C	1035	28	n	BM	NA	Imperial	Control
Q105.1	F	C	1000	26	y	BM_DBM	NA	Imperial	Control
Q105.2	F	C	1000	26	y	BM_DBM	NA	Imperial	Control
Q105.3	F	C	1000	26	n	BM_DBM	NA	Imperial	Control
Q105.4	F	C	1000	26	n	BM_F	NA	Imperial	Control
Q107.1	M	C	905	26	y	BM_DBM	NA	Imperial	Control
Q107.2	M	C	905	26	n	BM_DBM	NA	Imperial	Control
Q107.3	M	C	905	26	n	BM_DBM	NA	Imperial	Control
Q109.1	F	V	1005	27	y	BM	NA	Imperial	Control
Q109.2	F	V	1005	27	n	BM	NA	Imperial	Control
Q112.1	F	C	1450	31	n	BM_F	NA	Imperial	Control
Q112.2	F	C	1450	31	n	BM	NA	Imperial	Control
Q113.1	F	C	815	26	y	BM_DBM	NA	Imperial	Control
Q113.2	F	C	815	26	y	BM	NA	Imperial	Control
Q113.3	F	C	815	26	n	BM	NA	Imperial	Control
Q113.4	F	C	815	26	n	BM	NA	Imperial	Control
Q116.1	M	V	1200	29	y	BM_DBM	NA	Imperial	Control
Q116.2	M	V	1200	29	n	BM_DBM	NA	Imperial	Control
Q117.1	F	V	625	25	n	BM_DBM	NA	Imperial	Control
Q117.2	F	V	625	25	n	BM_DBM	NA	Imperial	Control
Q117.3	F	V	625	25	n	BM	NA	Imperial	Control
Q119.1	F	V	1540	31	n	BM_DBM	NA	Imperial	Control
Q119.2	F	V	1540	31	n	BM_DBM	NA	Imperial	Control
Q121.1	F	V	560	25	y	BM_DBM	NA	Imperial	Control
Q121.2	F	V	560	25	n	BM	NA	Imperial	Control
Q121.3	F	V	560	25	n	BM	NA	Imperial	Control
Q122.1	F	V	690	25	y	BM_DBM	NA	Imperial	Control
Q122.2	F	V	690	25	y	BM	NA	Imperial	Control

Q122.3	F	V	690	25	n	BM	NA	Imperial	Control
Q128.1	M	C	1215	28	n	BM_DBM	NA	Imperial	Control
Q128.2	M	C	1215	28	n	BM_DBM	NA	Imperial	Control
Q128.3	M	C	1215	28	y	BM	NA	Imperial	Control
Q129.1	M	V	570	23	n	BM_F	NA	Imperial	Control
Q129.2	M	V	570	23	n	BM_F	NA	Imperial	Control
Q13.1	F	C	770	29	n	BM_DBM	NA	Imperial	Control
Q13.2	F	C	770	29	n	BM_DBM	NA	Imperial	Control
Q13.3	F	C	770	29	n	BM	NA	Imperial	Control
Q13.4	F	C	770	29	n	BM	NA	Imperial	Control
Q131	M	V	1565	29	y	BM	NA	Imperial	Control
Q131.1	M	V	1565	29	y	BM_DBM	NA	Imperial	Control
Q133.2	F	V	805	26	n	BM_DBM	NA	Imperial	Control
Q133.3	F	V	805	26	n	BM_DBM	NA	Imperial	Control
Q133.4	F	V	805	26	n	BM	NA	Imperial	Control
Q133.5	F	V	805	26	n	BM	NA	Imperial	Control
Q142.1	M	V	1800	31	n	BM_DBM	NA	Imperial	Control
Q142.2	M	V	1800	31	n	BM	NA	Imperial	Control
Q142.3	M	V	1800	31	n	BM	NA	Imperial	Control
Q144.1	M	C	1270	31	y	BM	NA	Imperial	Control
Q144.2	M	C	1270	31	n	BM	NA	Imperial	Control
Q15.1	F	V	630	25	n	BM	NA	Imperial	Control
Q15.2	F	V	630	25	n	BM	NA	Imperial	Control
Q157.1	F	C	835	26	y	BM	NA	Imperial	Control
Q157.2	F	C	835	26	y	BM	NA	Imperial	Control
Q157.3	F	C	835	26	n	BM	NA	Imperial	Control
Q157.4	F	C	835	26	n	BM	NA	Imperial	Control
Q159.1	M	V	890	26	y	BM_DBM	NA	Imperial	Control
Q159.2	M	V	890	26	n	BM	NA	Imperial	Control
Q159.3	M	V	890	26	n	BM	NA	Imperial	Control
Q159.4	M	V	890	26	n	BM	NA	Imperial	Control
Q161.1	F	V	860	25	y	BM_DBM	NA	Imperial	Control
Q161.2	F	V	860	25	n	BM_DBM	NA	Imperial	Control
Q161.3	F	V	860	25	n	BM_F	NA	Imperial	Control
Q164.1	M	V	640	24	n	BM_DBM	NA	Imperial	Control
Q164.2	M	V	640	24	n	BM	NA	Imperial	Control
Q167.1	F	V	865	26	n	BM_DBM	NA	Imperial	Control
Q167.2	F	V	865	26	n	BM	NA	Imperial	Control
Q167.3	F	V	865	26	n	BM	NA	Imperial	Control
Q167.4	F	V	865	26	n	BM	NA	Imperial	Control
Q168.1	F	V	880	26	y	BM_DBM	NA	Imperial	Control
Q168.2	F	V	880	26	n	BM	NA	Imperial	Control
Q168.3	F	V	880	26	n	BM	NA	Imperial	Control
Q168.4	F	V	880	26	n	BM	NA	Imperial	Control
Q169.1	F	C	1665	31	y	BM_DBM	NA	Imperial	Control
Q169.2	F	C	1665	31	n	BM	NA	Imperial	Control
Q170.1	F	C	1525	31	y	BM_DBM	NA	Imperial	Control
Q170.2	F	C	1525	31	n	BM	NA	Imperial	Control
Q175.1	M	V	755	24	y	BM	NA	Imperial	Control
Q175.2	M	V	755	24	n	BM	NA	Imperial	Control

Q175.3	M	V	755	24	n	BM	NA	Imperial	Control
Q175.4	M	V	755	24	n	BM	NA	Imperial	Control
Q182.1	F	C	1150	27	n	BM	NA	Imperial	Control
Q182.2	F	C	1150	27	n	BM	NA	Imperial	Control
Q183.1	M	V	510	23	n	BM_DBM	NA	Imperial	Control
Q183.2	M	V	510	23	n	BM_DBM	NA	Imperial	Control
Q187.1	F	V	1550	30	n	BM_DBM	NA	Imperial	Control
Q187.2	F	V	1550	30	n	BM	NA	Imperial	Control
Q189.1	M	V	750	25	n	BM_DBM	NA	Imperial	Control
Q189.2	M	V	750	25	n	BM_DBM	NA	Imperial	Control
Q189.3	M	V	750	25	n	BM_DBM	NA	Imperial	Control
Q189.4	M	V	750	25	n	BM	NA	Imperial	Control
Q200.1	M	V	730	25	y	BM_DBM	NA	Imperial	Control
Q200.2	M	V	730	25	n	BM_DBM	NA	Imperial	Control
Q200.3	M	V	730	25	n	BM	NA	Imperial	Control
Q216.1	M	C	1150	28	y	BM_DBM	NA	Imperial	Control
Q219.1	M	C	810	26	y	BM_DBM	NA	Imperial	Control
Q219.2	M	C	810	26	n	BM	NA	Imperial	Control
Q219.3	M	C	810	26	n	BM	NA	Imperial	Control
Q22.1	F	C	1060	31	n	BM_DBM	NA	Imperial	Control
Q22.2	F	C	1060	31	n	BM_DBM	NA	Imperial	Control
Q226.1	M	C	660	24	y	DBM	NA	Imperial	Control
Q226.2	M	C	660	24	y	DBM	NA	Imperial	Control
Q226.3	M	C	660	24	n	DBM	NA	Imperial	Control
Q23.1	F	C	1690	31	n	F	NA	Imperial	Control
Q25.1	M	V	1400	28	n	BM	NA	Imperial	Control
Q25.2	M	V	1400	28	n	BM	NA	Imperial	Control
Q26.1	M	V	840	26	n	BM	NA	Imperial	Control
Q26.2	M	V	840	26	n	BM	NA	Imperial	Control
Q27.1	M	V	600	24	n	BM_DBM	NA	Imperial	Control
Q27.2	M	V	600	24	n	BM_DBM	NA	Imperial	Control
Q28.1	F	C	860	30	n	BM_DBM	NA	Imperial	Control
Q28.2	F	C	860	30	n	BM_DBM	NA	Imperial	Control
Q28.3	F	C	860	30	n	BM	NA	Imperial	Control
Q29.1	F	C	950	30	n	BM_DBM	NA	Imperial	Control
Q29.2	F	C	950	30	n	BM	NA	Imperial	Control
Q3.1	F	V	1885	31	y	BM_DBM	NA	Imperial	Control
Q3.2	F	V	1885	31	y	BM_DBM	NA	Imperial	Control
Q31.1	M	V	960	25	y	BM_DBM	NA	Imperial	Control
Q31.2	M	V	960	25	n	BM	NA	Imperial	Control
Q31.3	M	V	960	25	n	BM	NA	Imperial	Control
Q33.1	F	V	1000	26	y	BM_DBM	NA	Imperial	Control
Q33.2	F	V	1000	26	y	BM_DBM	NA	Imperial	Control
Q33.3	F	V	1000	26	n	BM	NA	Imperial	Control
Q33.4	F	V	1000	26	n	BM	NA	Imperial	Control
Q48.1	F	C	1200	28	n	BM_DBM	NA	Imperial	Control
Q48.2	F	C	1200	28	n	BM	NA	Imperial	Control
Q48.3	F	C	1200	28	n	BM	NA	Imperial	Control
Q49.1	M	V	770	23	y	BM_DBM	NA	Imperial	Control
Q49.2	M	V	770	23	y	BM_DBM	NA	Imperial	Control

Q49.3	M	V	770	23	n	BM	NA	Imperial	Control
Q49.4	M	V	770	23	n	BM	NA	Imperial	Control
Q5.1	F	C	780	31	y	BM_DBM	NA	Imperial	Control
Q52.1	F	V	990	26	n	BM_DBM	NA	Imperial	Control
Q52.4	F	V	990	26	n	BM_DBM	NA	Imperial	Control
Q55.2	M	V	765	25	y	BM_DBM	NA	Imperial	Control
Q55.3	M	V	765	25	n	BM_DBM	NA	Imperial	Control
Q58.1	M	C	920	26	n	BM_DBM	NA	Imperial	Control
Q58.2	M	C	920	26	n	BM_DBM	NA	Imperial	Control
Q58.3	M	C	920	26	n	BM_F	NA	Imperial	Control
Q60.1	M	V	650	24	y	BM_DBM	NA	Imperial	Control
Q60.2	M	V	650	24	y	BM_DBM	NA	Imperial	Control
Q60.3	M	V	650	24	n	BM	NA	Imperial	Control
Q60.4	M	V	650	24	n	BM	NA	Imperial	Control
Q65.1	M	V	915	26	n	BM_DBM	NA	Imperial	Control
Q65.2	M	V	915	26	n	BM	NA	Imperial	Control
Q70.1	F	V	535	26	n	BM_DBM	NA	Imperial	Control
Q70.2	F	V	535	26	n	BM_DBM	NA	Imperial	Control
Q70.3	F	V	535	26	n	BM	NA	Imperial	Control
Q70.4	F	V	535	26	n	BM	NA	Imperial	Control
Q83.1	F	V	850	25	y	BM_DBM	NA	Imperial	Control
Q83.2	F	V	850	25	n	BM_DBM	NA	Imperial	Control
Q83.3	F	V	850	25	n	BM_DBM	NA	Imperial	Control
Q83.4	F	V	850	25	n	BM	NA	Imperial	Control
Q87.1	F	C	1025	27	n	BM_DBM	NA	Imperial	Control
Q87.2	F	C	1025	27	n	BM_DBM	NA	Imperial	Control
Q87.3	F	C	1025	27	n	BM	NA	Imperial	Control
Q89.1	F	C	1246	31	n	BM_F	NA	Imperial	Control
Q89.2	F	C	1246	31	n	BM_F	NA	Imperial	Control
Q95.3	M	V	1695	31	n	BM_F	NA	Imperial	Control
Q95.4	M	V	1695	31	n	BM_F	NA	Imperial	Control

Appendix 3

Metadata all infants samples

Study No.	Hospital	Gestation (weeks)	Weight (gr)	Ab. BenPen + Gent	# Days of Ab trea.	Mode of Delivery	Prob (y/n)	1st week days from birth	2nd week days from birth	3rd week days from birth	1st week BM or DBM? (y/n)	2nd week BM or DBM? (y/n)	3rd week BM or DBM? (y/n)
M15	St Mary's	30	1370	y	3	C section	n	4	14	21	y	y	Y
M16	St Mary's	29	1340	y	3	C section	n	7	14	21	y	y	Y
M17	St Mary's	29	1230	y	3	C section	n	7	13	21	y	y	Y
M26	St Mary's	28	1322	y	6	C section	n	5	14	18	y	y	Y
M29	St Mary's	28	934	y	3	C section	n	9	14	22	y	y	Y
M36	St Mary's	31	426	y	6	C section	n	8	14	20	y	y	Y
M37	St Mary's	31	697	y	2	C section	n	7	14	21	y	y	Y
M38	St Mary's	29	500	y	3	C section	n	5	12	20	y	y	Y
M39	St Mary's	28	870	y	5	NVD	n	8	14	22	y	y	Y
M7	St Mary's	31	627	n	0	NVD	n	5	11	21	y	y	Y
M40	St Mary's	30	1320	n	0	C section	n	10	13	21	y	y	Y
M54	St Mary's	28	980	n	0	C section	n	5	14	16	y	y	Y
M55	St Mary's	28	1330	n	0	C section	n	5	13	16	y	y	Y
M56	St Mary's	28	1200	n	0	C section	n	5	14	16	y	y	Y
M59	St Mary's	28	980	n	0	C section	n	4	14	22	y	y	Y
M60	St Mary's	28	1100	n	0	C section	n	5	14	22	y	y	Y
M103	St Mary's	28	1380	n	0	C section	n	8	14	21	y	y	N
Q48	Queen Charlotte's	28	770	n	0	C section	n	6	15	18	y	y	Y
Q29	Queen Charlotte's	30	800	n	0	C section	n	5	14	21	y	y	Y
P80	NNUH	25	831	y	3	NVD	y	3	13	20	y	y	Y
P63	NNUH	26	1374	y	2	NVD	y	7	15	19	y	y	N
P65	NNUH	27	1477	y	2	NVD	y	6	13	20	y	y	N
P42	NNUH	29	1384	y	2	C section	y	7	15	22	y	y	Y
P69	NNUH	30	1420	y	3	C section	y	6	12	22	y	y	Y
P74	NNUH	31	1382	y	3	C section	y	5	15	23	y	y	Y
P79	NNUH	30	1444	y	3	C section	y	7	14	20	y	y	Y
P46	NNUH	30	1400	y	3	NVD	y	8	14	23	y	y	Y
P75	NNUH	31	1262	y	4	NVD	y	5	12	24	y	y	Y
P76	NNUH	31	1409	y	4	NVD	y	5	14	17	y	y	Y
P70	NNUH	30	1255	y	5	C section	y	9	13	20	y	y	Y
P48	NNUH	30	1425	y	8	NVD	y	6	11	20	y	y	N
P60	NNUH	33	1230	n	0	C section	y	5	15	17	y	y	Y
P62	NNUH	30	940	n	0	C section	y	6	15	25	y	y	Y
P77	NNUH	30	1110	n	0	C section	y	5	14	20	y	y	Y

Appendix 4

AMR genes detected using ´walk-out´ analysis

Premature infant P205

ReadId	HostHit	CARDhit	PercentId
a6230d05-2515-4dba-b7a8-ca2c1c8641f2	Enterobacter cloacae	ARO:3001830|ACT-7	92.49
21ccabe7-78be-4380-980c-1711baa4262a	Enterobacter cloacae	ARO:3001847|ACT-27	84.07
06711712-3445-4f0e-aa08-d5871266f0f2	Enterobacter cloacae	ARO:3000263|marA	80.55
5d6c65f0-4052-4757-b78c-d03a289bed10	Enterobacter cloacae	ARO:3000518|CRP	80.22
e9737391-a8da-4065-a05b-9b4d68ae6fae	Enterobacter cloacae	ARO:3001847|ACT-27	87.8
97c715ce-77be-42ec-8abd-818d0639f5ba	Enterobacter cloacae	ARO:3001847|ACT-27	92.42
f5749dfb-899e-485c-ac14-36d0b1604f78	Enterobacter cloacae	ARO:3000216|acrB	81.27
7ac77a00-ef2f-4ad1-b8ad-21e3a560f46b	Enterobacter cloacae	ARO:3000518|CRP	84.28
2d78108b-a656-4131-9971-f8c9bdbe8fcb	Enterobacter cloacae	ARO:3001847|ACT-27	83.73
e7f46d19-4b9b-4d78-8ca1-ffffc95393d5	Enterobacter cloacae	ARO:3000216|acrB	80.23
5e153150-2529-4306-8158-ba7d5d0364ba	Enterobacter cloacae	ARO:3000216|acrB	85.65
72877490-fab6-4758-b03c-f776ef0b8bf2	Enterobacter cloacae	ARO:3000207|acrA	82.16
422f7d5f-bb7b-4974-a87f-1cae04d268eb	Enterobacter cloacae	ARO:3001852|ACT-32	85.86
36f70b32-15fd-4e0b-9c94-c72571398d84	Enterobacter cloacae	ARO:3000825|robA	87.63
fd8df6a0-46fd-46ac-999c-fa4360c28137	Enterobacter cloacae	ARO:3000207|acrA	81.88
596959bb-1080-4e0e-8e85-4a4680f494f6	Enterobacter cloacae	ARO:3000216|acrB	82.89
6c2e0f17-ace8-4b8a-ac1a-6e8641d31b20	Staphylococcus aureus	ARO:3000178|tetK	87.33
62390dfc-a944-4800-8d46-0cba10eb984b	Enterobacter cloacae	ARO:3000216|acrB	81.74
b05b78e9-11d5-4c14-8f19-1b454be2cee1	Enterobacter cloacae	ARO:3000207|acrA	80.79
6eb576bc-f12b-4ad8-8254-1bc558df12f3	Enterobacter cloacae	ARO:3000518|CRP	86.62
2e9be510-cee3-4d37-a9f9-9b6b0ac33ab1	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.75
c2c6c3fe-bc78-4f68-ab48-33c374dab7b6	Enterobacter cloacae	ARO:3003923|oqxB	80
7a7673a6-76bf-4fb1-bee0-a32fe7984e2b	Enterobacter cloacae	ARO:3000825|robA	87.98
9e58203e-11f7-4af4-ba02-a2ae863e2dae	Enterobacter cloacae	ARO:3000825|robA	90.8
2430d2e3-5d3b-48fb-ac0e-93ba1c3aef61	Enterobacter cloacae	ARO:3000216|acrB	86.97
d0f48c2e-1fce-4e96-91d8-242cb5d5dd99	Enterobacter cloacae	ARO:3003830|aminocoumarin	81.35
dfd129e3-110b-4182-8256-371a78664592	Staphylococcus epidermidis	ARO:3000391|norA	91.76
0d203ae2-92e8-4857-a62e-b0b2261ca109	Enterobacter cloacae	ARO:3000830|cpxA	82.08
4959c7c0-b8a8-4d3a-9349-366a80503caa	Enterobacter cloacae	ARO:3000825|robA	86.56
0a2a5933-d089-4141-a799-62b1793485da	Enterobacter cloacae	ARO:3000207|acrA	81.55
0a2a5933-d089-4141-a799-62b1793485da	Enterobacter cloacae	ARO:3000216|acrB	86.05
80c018ef-42bd-4065-b9e8-2d72be33fe22	Enterobacter cloacae	ARO:3001847|ACT-27	87.9
53f66feb-0c21-4350-91d5-7336b0a3e1a6	Enterobacter cloacae	ARO:3000516|emrR	80.23

c7ae6143-4c0a-4461-8144-3ba574a28392	Enterobacter cloacae	ARO:3000263|marA	81.34
88cd3c99-d367-4081-b159-84d0fc4723ba	Enterobacter cloacae	ARO:3000216|acrB	84.58
59a120c7-7304-41f6-a29a-35b850dcb777	Enterobacter cloacae	ARO:3000516|emrR	80.86
2444d78b-0d6a-42f2-9890-6ee409f14642	Enterobacter cloacae	ARO:3000216|acrB	82.73
74838f01-7aa8-41bb-a0cc-615fdce65038	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	83.52
d976c554-a142-4e6f-a1fa-b1572c550824	Enterobacter cloacae	ARO:3000216|acrB	81.86
61006e14-9728-41bb-9dc7-bee206cbb447	Enterobacter cloacae	ARO:3001830|ACT-7	87.91
da80e898-1db6-486e-89c5-8c2f4492c07c	Enterobacter cloacae	ARO:3000216|acrB	85.98
441bd110-51c2-43cd-a7c6-2f020a67c665	Enterobacter cloacae	ARO:3001838|ACT-17	90.18
5344e7b2-4f4b-40b7-a8d9-246df4ff16f8	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.07
e85a7508-990e-4b04-9d3d-0b96f9997fdc	Enterobacter cloacae	ARO:3000825|robA	92.09
9acedb62-5294-484d-9607-85ef39e904c9	Enterobacter cloacae	ARO:3001830|ACT-7	87.03
9a625860-7ef5-463a-ac2b-db3be186d263	Enterobacter cloacae	ARO:3000518|CRP	80.27
3990f624-29a8-43ea-aa60-5bc8a055cb31	Enterobacter cloacae	ARO:3000825|robA	87.84
3984c77b-b22e-425e-99b4-450ac4d99f9d	Staphylococcus epidermidis	ARO:3000178|tetK	81.15
891b762f-f208-4ae6-a277-4ae7e5c1a218	Enterobacter cloacae	ARO:3003923|oqxB	81.27
d3435b65-399a-4fcb-96cc-bc48622d7d33	Enterobacter cloacae	ARO:3000024|patA	82.03
573348b9-cffb-4031-9a2a-de6f0d031565	Enterobacter cloacae	ARO:3000831|cpxR	82.18
23d8a370-8c05-49d6-b76f-98e5236e7f5c	Enterobacter cloacae	ARO:3003923|oqxB	85.25
1f4ac2fe-2bda-4199-84ba-a3cd11ce9905	Enterobacter cloacae	ARO:3000216|acrB	86.6
5980533b-417e-4ed2-89b5-ca5915ae781b	Enterobacter cloacae	ARO:3000518|CRP	81.05
8b2d4d7b-033e-4d8e-8dc0-ae5fc2f162c0	Enterobacter cloacae	ARO:3000216|acrB	82.09
a9b8a9d2-4169-439f-b81f-86f7aeb26104	Enterobacter cloacae	ARO:3000516|emrR	80.81
ea17b316-c02a-4d3c-8cfe-6400d4637d49	Enterobacter cloacae	ARO:3003830|aminocoumarin	81.86
ff1a6349-9052-428e-99a9-ee815bec7fa8	Enterobacter cloacae	ARO:3003923|oqxB	82.45
a931a60e-9b0c-4060-b63a-cbdd5b18aafd	Enterobacter cloacae	ARO:3000825|robA	89.91
c5a78e7c-cfb8-4a9d-a908-ef0661a9f1bf	Enterobacter cloacae	ARO:3000216|acrB	81.25
eedee14d-0334-4b3b-bb7a-548af8313df8	Enterobacter cloacae	ARO:3000216|acrB	87.16
6c7790e5-113b-4776-8547-e26c385f0612	Enterobacter cloacae	ARO:3003923|oqxB	85.98
6a2d759b-46e1-4bef-90a2-e6403c80c827	Enterobacter cloacae	ARO:3000825|robA	88.62
c6c66581-6bc0-4963-9359-2f2637ef15e5	Enterobacter cloacae	ARO:3000207|acrA	82.96
e182c38e-78ae-4ce0-9b09-02ae9354c9af	Enterobacter cloacae	ARO:3003923|oqxB	80.56
552ed409-7858-440a-a17e-d1faeab3291b	Enterobacter cloacae	ARO:3000378|mexB	80.52
3beef42b-228a-4e1f-944c-2825884f2317	Enterobacter cloacae	ARO:3000825|robA	88.62
45066790-3a2a-46be-b8d3-ee3ccad20174	Enterobacter cloacae	ARO:3000518|CRP	80.07
6c7924ba-266c-467c-9538-a8e629318353	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	83.02
84de44f2-b58e-452e-b33b-620895bdcdd7	Enterobacter cloacae	ARO:3001847|ACT-27	88.09
bd244e08-c8f8-4d77-ada3-7c0a3571b679	Enterobacter cloacae	ARO:3000825|robA	82.52

a8809e59-3177-493f-b265-2b7d339917c1	Enterobacter cloacae	ARO:3000825|robA	83.1
a17e8faf-3be1-440c-ab75-bf1f1cbae7ce	Enterobacter cloacae	ARO:3000676|H-NS	83.33
72cf1ebe-0519-402c-bf31-cbbfc327d1f9	Enterobacter cloacae	ARO:3000825|robA	86.28
b6d3d7d8-e4b7-47d7-811a-63d1a944b5fc	Staphylococcus agnetis	ARO:3000621|PC1	86.74
0ce42433-1841-4a60-85c7-cd3114db09ba	Enterobacter cloacae	ARO:3000216|acrB	89.41
0ce42433-1841-4a60-85c7-cd3114db09ba	Enterobacter cloacae	ARO:3000207|acrA	85.41
e1a9361f-d0bd-436d-9d31-68182782a221	Enterobacter cloacae	ARO:3001847|ACT-27	89.56
698d6aaf-2b12-4f07-98f9-fac76da7c22b	Enterobacter cloacae	ARO:3000216|acrB	83.76
88b26a17-c557-4b40-8bf2-1df1979591ce	Enterobacter cloacae	ARO:3000825|robA	86.08
9f0bf21d-3357-4b65-bd50-55315e5af285	Enterococcus faecalis	ARO:3003551|emeA	86.67
f9ed9409-c4ac-41d4-909a-4c9b05b7965e	Enterobacter cloacae	ARO:3000825|robA	90.83
93817556-55d5-4cf0-9708-8a5b4a2e97ca	Enterobacter cloacae	ARO:3000825|robA	86.05
2a12d450-7bb0-43ff-94b9-6b33fc351a8d	Enterobacter cloacae	ARO:3000825|robA	87.28
e7a5ee68-9551-44b7-b576-040c50773ee4	Enterobacter cloacae	ARO:3000074|emrB	80.12
813b78aa-9691-4572-83e2-f04b7a6feb0c	Enterobacter cloacae	ARO:3000216|acrB	83.3
abc27985-629b-4861-9cb3-00c465b66422	Enterobacter cloacae	ARO:3003923|oqxB	83.04
4f2147e4-54ac-4794-88a1-e0062ab48502	Enterobacter cloacae	ARO:3000207|acrA	80.02
fad1cbe8-6a16-473a-bd5e-03bf823f9128	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.56
3cf4b3aa-3ec5-4c85-998d-a1adba583564	Enterobacter cloacae	ARO:3001830|ACT-7	89.61
3cf4b3aa-3ec5-4c85-998d-a1adba583564	Enterobacter cloacae	ARO:3001835|ACT-13	80.46
51eb719d-a46f-4a39-8296-47440423b9cc	Enterobacter cloacae	ARO:3000216|acrB	80.13
d2672dc7-9169-48f3-bf3c-f7fad231db4b	Enterobacter cloacae	ARO:3000516|emrR	80.36
96ed2db9-787f-439c-86bd-5f212a780596	Enterobacter cloacae	ARO:3000676|H-NS	81.44
d87de989-8812-4a9e-b7fb-b35ff1721361	Enterobacter cloacae	ARO:3000216|acrB	85.32
3b188190-c84d-4998-9e12-47c4f0daab23	Enterobacter cloacae	ARO:3003922|oqxA	82.35
02b911f0-96f3-43e8-bf8f-be1aae5620ab	Enterobacter cloacae	ARO:3000518|CRP	85.54
8ff0da3b-babc-423c-9d88-0d4d8f1e96e1	Enterobacter cloacae	ARO:3001847|ACT-27	86.31
8dcb6bea-f718-4473-ab70-511ef5ed9e50	Enterobacter cloacae	ARO:3001847|ACT-27	86.08
f0dbc62f-f25a-457c-88b0-c01b8c50c264	Enterobacter cloacae	ARO:3000825|robA	90.17
82975174-7bde-43ca-ae07-a366167ceb5e	Enterobacter cloacae	ARO:3003922|oqxA	80.39
0c24ee70-b1ab-4f09-93ac-ed4517ef95c6	Enterobacter cloacae	ARO:3001830|ACT-7	87.18
cbd94876-73a6-400c-8f3c-e5c0d795579c	Enterobacter cloacae	ARO:3000676|H-NS	81.75
f9f76a9d-3bf6-4fd1-a939-d52d8c70e381	Enterobacter cloacae	ARO:3000207|acrA	83.7
11e16b29-62b8-4fb1-9da0-5d689f43e7bf	Enterobacter cloacae	ARO:3000074|emrB	80.25
7b7873b0-6280-4194-9166-9f1505bd8a6a	Enterobacter cloacae	ARO:3000518|CRP	85
bb402264-f0dd-450f-a713-def0d2210a07	Enterobacter cloacae	ARO:3000825|robA	91.09
42d724bd-72a4-4679-b71f-66654fb9df3b	Enterobacter cloacae	ARO:3000216|acrB	83.94
7eb7c099-3079-41e3-8b9a-ed2b40083767	Enterobacter cloacae	ARO:3000074|emrB	80.2
40274626-3298-42e5-b1f0-fe296e7dde61	Enterobacter cloacae	ARO:3001847|ACT-27	89.01

6c498d0d-19a9-449d-933a-acfbf4696ffd	Enterobacter cloacae	ARO:3000825|robA	92.28
04c9220b-d30e-4c69-b912-31c322043a02	Enterobacter cloacae	ARO:3003830|aminocoumarin	81.3
7e3ca015-5728-4fcf-b856-7c10861f8bd4	Enterobacter cloacae	ARO:3000518|CRP	80.18
f610f670-9ea4-48ea-8b4f-4ded4c579c56	Enterobacter cloacae	ARO:3000825|robA	94.12
999a2f97-a9e1-4d55-9fb4-8bc9991c8b06	Enterobacter cloacae	ARO:3001847|ACT-27	82.06
52780a08-4745-4498-ae5b-8bd292ef56c7	Enterobacter cloacae	ARO:3003923|oqxB	84.25
52780a08-4745-4498-ae5b-8bd292ef56c7	Enterobacter cloacae	ARO:3003922|oqxA	82.74
055e51e4-603c-4639-980c-82489d65189d	Enterobacter cloacae	ARO:3000825|robA	93.38
a511e3c9-8d55-48f8-bb62-1795a94f91c8	Enterobacter cloacae	ARO:3000825|robA	87.73
1b5209ae-f1af-4009-9e2a-7c7c3222fbbd	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.63
916b2c80-ca3a-436d-a9e2-246e5c0374dc	Enterobacter cloacae	ARO:3000207|acrA	82.09
2a4c6971-558c-4021-967c-fef1e19746a0	Enterobacter cloacae	ARO:3001847|ACT-27	84.45
3847bb05-4cd5-451e-89f6-c4474b75f0a4	Staphylococcus aureus	ARO:3000319|mphC	87.69
f451ef19-b6d3-4ce6-9e19-0112ad00db52	Enterobacter cloacae	ARO:3001841|ACT-20	89.43
f451ef19-b6d3-4ce6-9e19-0112ad00db52	Enterobacter cloacae	ARO:3001844|ACT-24	87.94
0e9e7263-d1a2-4024-b9b3-f856c59e8f0d	Enterobacter cloacae	ARO:3001330|mdtD	80.91
32eccf9d-f2cd-40b4-a508-5743a6009e80	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.75
f6b46792-98fc-4c27-96d1-cb400672018f	Enterobacter cloacae	ARO:3003922|oqxA	80.52
ae965e8e-7975-493d-bed7-d9eb13ae852e	Staphylococcus sp.	ARO:3002865|dfrC	94.64
7f94bcbb-b7d1-4f84-95bd-7f295a19d1ec	Enterobacter cloacae	ARO:3000825|robA	84.75
aaab3d6d-81c4-46c8-bd1b-87fe239140a7	Enterobacter cloacae	ARO:3000825|robA	87.56
fb8543ec-677f-4b4a-9376-07197f73ac79	Enterobacter cloacae	ARO:3000825|robA	89.16
667800f2-afa0-4220-9571-2f7346f31557	Enterobacter cloacae	ARO:3001838|ACT-17	87.6
21565350-9b24-4ff1-9aab-2cc8b0fb1da7	Enterobacter cloacae	ARO:3003922|oqxA	80
28e7293a-6a72-48e6-a53f-7002e7de7e41	Enterobacter cloacae	ARO:3001847|ACT-27	90.17
ea7670e5-33f0-440f-a055-cbd162dab44e	Enterobacter cloacae	ARO:3000024|patA	86.14
f5ca607d-ae8a-48a1-ba56-4b650f847f8a	Enterobacter cloacae	ARO:3000516|emrR	80.92
9d7d8110-2d57-4ddd-852f-4adda279f7b2	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.59
9ef4cc10-940a-4479-9bc9-6c4f1b6df541	Enterobacter cloacae	ARO:3000207|acrA	82.53
9ef4cc10-940a-4479-9bc9-6c4f1b6df541	Enterobacter cloacae	ARO:3000216|acrB	85.03
80d9ae32-56ac-4c89-a9b9-20bf8d36a7c2	Staphylococcus epidermidis	ARO:3000391|norA	84.42
b5d68c2a-d46f-4e96-a91f-23dd4b5b3254	Enterobacter cloacae	ARO:3000216|acrB	82.46
d332f063-53f6-4841-bdad-58119c9cd320	Enterobacter cloacae	ARO:3000518|CRP	88.28
bbe4fa20-1e81-4dba-aedd-6c7d76309d27	Enterobacter cloacae	ARO:3000825|robA	89.79
6073cb30-9d34-4283-8c1f-58d5164dbb64	Staphylococcus aureus	ARO:3000489|sav1866	89.47
2570266c-977e-4e30-8550-219d6054d660	Enterobacter cloacae	ARO:3000825|robA	89.7
eea3c8b5-e265-4e89-97c1-b156f9dfa19e	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.63
96abd111-ae02-4dfb-9b77-5c15f420632b	Enterobacter cloacae	ARO:3003923|oqxB	82.95

72bfe536-8263-4dac-a12a-d57f179d91e2	Enterobacter cloacae	ARO:3001847|ACT-27	88.32
f3e2a04b-bde0-4bc2-915e-d52d16ba6829	Enterobacter cloacae	ARO:3000825|robA	82.25
bae0dbaf-4a1c-4eed-b72f-e66963ceff20	Enterobacter cloacae	ARO:3001830|ACT-7	85.95
8199e22a-86a3-4be6-9b20-d77bca024021	Enterobacter cloacae	ARO:3000518|CRP	81.07
24b127e6-2e4f-48e0-9b4f-f2d641262248	Enterobacter cloacae	ARO:3000216|acrB	80.42
bc5d6bc1-d7bc-49a8-a761-a330fe41fa6f	Staphylococcus epidermidis	ARO:3000391|norA	91.07
2d741a20-0401-4f99-80a9-44f2bfd1f75f	Enterobacter cloacae	ARO:3000676|H-NS	80.39
b43fa09d-204c-44b1-b942-787b5e7f608a	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	81.65
39440cad-8dc2-4270-985e-22195485f52d	Enterobacter cloacae	ARO:3001852|ACT-32	81.47
db38add8-c5ab-4d41-9ea6-7b8d41d9386f	Enterobacter cloacae	ARO:3000216|acrB	85.21
4a1e86cb-99c4-4701-8bd2-f5e87fffe5f0	Enterobacter cloacae	ARO:3000825|robA	89.85
a30d4421-adac-458e-b4b9-3b1774dc8b79	Enterobacter cloacae	ARO:3000825|robA	87.5
f2bb02f0-36a9-4190-b2d1-a1ab69a12c34	Enterobacter cloacae	ARO:3000518|CRP	80.91
25ef2467-a8d4-4273-b93b-5218a4ee7bf6	Enterobacter cloacae	ARO:3000516|emrR	80.27
663e81ab-fa1b-4a8e-9d1e-8487430932b5	Enterobacter cloacae	ARO:3000825|robA	92.94
22d84bd0-0701-4e17-878d-079c7bb3f732	Enterobacter cloacae	ARO:3001837|ACT-15	87.33
e2db8a3e-a993-412d-aa46-08a2c24850a0	Enterobacter cloacae	ARO:3001847|ACT-27	85.31
4663a1e4-cc8d-488d-926d-4d4840b22917	Enterobacter cloacae	ARO:3001830|ACT-7	88.47
350aae2c-9ee6-4378-96ea-7753fa4085ca	Enterobacter cloacae	ARO:3003052|smeB	82.4
69dc61e4-e5d6-4fda-80f8-5e91f81f3e9f	Enterobacter cloacae	ARO:3000825|robA	89.94
e2ed961e-ccb8-457e-b509-ac5663aa5869	Enterobacter cloacae	ARO:3000825|robA	86.55
fddab6a1-bc90-4149-ae60-1a0e4f5bd9da	Staphylococcus aureus	ARO:3003552|fusB	92.91
94ab3fd2-1393-4742-abc1-a1eca3c8c506	Enterobacter cloacae	ARO:3003923|oqxB	81.44
8a57793b-71be-41ca-ab08-e57601d1c93b	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.75
d037d0c0-261a-4a3a-baa4-687799f8674a	Enterobacter cloacae	ARO:3000516|emrR	82.25
bf062aa4-6aaa-4527-b85b-11a87084f149	Enterobacter cloacae	ARO:3000825|robA	89.06
5b169490-bffe-4247-83c1-b5575cbb9182	Enterobacter cloacae	ARO:3000518|CRP	81.02
5344a661-caa2-4cb6-9277-33a1e6345b66	Enterobacter cloacae	ARO:3000825|robA	82.75
b89ea282-c02d-4335-8c1a-5a6b654835d4	Enterobacter cloacae	ARO:3003830|aminocoumarin	82.78
348b56b4-09fe-4eae-9479-7746f85efd33	Enterobacter cloacae	ARO:3001847|ACT-27	88.97
6ad64be2-e0c7-442f-b018-bc068f8f1c56	Enterobacter cloacae	ARO:3000518|CRP	84.82
f80f4175-6173-4bb4-bcef-7ed04494653f	Enterobacter cloacae	ARO:3003830|aminocoumarin	85.56
8c0fee91-03d9-441a-a7d1-3c4529680bf6	Enterobacter cloacae	ARO:3001838|ACT-17	88.33
e205f7ce-93e0-4d85-8c84-c0ffe6a059e9	Enterobacter cloacae	ARO:3000825|robA	87.12
8255e212-2060-469d-89bd-af605d518772	Enterobacter cloacae	ARO:3000676|H-NS	84.51
bf433851-a61c-4c49-ab8f-dcfd5619f7ad	Enterobacter cloacae	ARO:3000216|acrB	84.56
c66ba70b-9d35-49aa-8e3b-37a67c4a2a05	Enterobacter cloacae	ARO:3000216|acrB	81.12
23769754-d4b8-4f9b-8dc5-7c8d4dbe4a1f	Enterobacter cloacae	ARO:3000825|robA	94.58
c24c3b2d-f48d-4d50-ae13-b664b9d9e184	Bifidobacterium animalis	ARO:3003730|Bifidobacterium ileS	80.37

4dc956d4-f9d7-450b-a15c-e143908b28f8	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.44
ba5e8406-a038-42ff-8a2b-09b3aec83d0a	Enterobacter cloacae	ARO:3000825|robA	87.21
bda016a7-8e2d-4e3e-aee0-22846b0f1022	Enterobacter cloacae	ARO:3000216|acrB	87.78
b6620a1d-fa42-48c6-80f9-57d7c7163ec9	Enterobacter cloacae	ARO:3000825|robA	88.85
5655384a-97b1-4bca-92ba-d76d470c2287	Enterobacter cloacae	ARO:3000024|patA	83.77
bac878dc-aa43-4d93-8a50-104b52e2537d	Enterobacter cloacae	ARO:3000825|robA	85.29
3eb2bf64-0680-4e96-8159-0cdf8b4a287e	Enterobacter cloacae	ARO:3001847|ACT-27	86.64
f44572cc-46a0-43e1-b8fa-f1fa12094dae	Enterobacter cloacae	ARO:3003922|oqxA	81.38
c4836dc2-48c1-4847-bc8e-4861d13044ff	Enterobacter cloacae	ARO:3001837|ACT-15	81.79
4c0b7765-1ebd-4d9f-a0a6-cceb205e74e9	Enterobacter cloacae	ARO:3001841|ACT-20	89.74
1272efbe-6724-4f1d-8087-675badae6e3c	Enterobacter cloacae	ARO:3000518|CRP	80.98
89c31618-d87e-44de-9e93-8743ce2dc961	Staphylococcus aureus	ARO:3000565|tet38	92.77
1665e7ce-f95c-4310-a952-23b5b7c5537e	Enterobacter cloacae	ARO:3002986|bacA	80
9ba6d334-7fea-4d5c-ad8c-c9b15b3385b4	Enterobacter cloacae	ARO:3000825|robA	91.47
dc9cf254-2fce-467f-9147-74f71595fcfc	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	85.24
80c110d8-8bdd-4bbb-bd03-12900deb0757	Enterobacter cloacae	ARO:3000825|robA	90.44
de03e5d0-a059-4d98-aba8-016e9428028e	Enterobacter cloacae	ARO:3000216|acrB	83.33
c7e936d7-e9b7-4d23-b6be-16bfcb6448ad	Enterobacter cloacae	ARO:3000216|acrB	85.03
a44dd817-ecba-47b9-a22f-1165f37986c6	Enterobacter cloacae	ARO:3000825|robA	89.93
c7867f2a-1239-45ee-959b-3b3c37dd1e22	Enterobacter cloacae	ARO:3000516|emrR	80
42443bf3-32d2-4409-a5a8-ddd8e1363561	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	83.2
c03abea3-3ba6-435f-81e6-db802d5db2cf	Enterobacter cloacae	ARO:3000216|acrB	86.79
c03abea3-3ba6-435f-81e6-db802d5db2cf	Enterobacter cloacae	ARO:3000207|acrA	80.39
49e126a3-c81a-43a9-9f1e-fb87b01bbfae	Enterobacter cloacae	ARO:3000263|marA	80.69
260f2442-25a9-49b8-ae71-a4d3bd5a90da	Enterobacter cloacae	ARO:3000024|patA	82.11
f9cd6825-da96-41f6-8f24-1038042ea252	Enterobacter cloacae	ARO:3000216|acrB	84.97
032f84d1-d86e-4023-9a4c-3e542f676097	Enterobacter cloacae	ARO:3001830|ACT-7	89.78
1e3f4985-b652-4e4d-8985-9cabaeedbfd5	Enterobacter cloacae	ARO:3001847|ACT-27	90.34
2c1c704a-c4a8-4076-bcd3-cbf5ce788f13	Staphylococcus aureus	ARO:3000746|mepR	90.75
9651c04a-198a-4f2a-a834-e33f201d35e1	Staphylococcus epidermidis	ARO:3000815|mgrA	81.51
4109a90c-dbdc-4054-94dc-5dd360da0bea	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	87.78
46932b00-da18-4937-b2ce-a3a9a07c1cee	Enterobacter cloacae	ARO:3000216|acrB	84.3
b1751ba0-2e2f-4d7c-82d8-97ca49dcc3dd	Enterobacter cloacae	ARO:3000216|acrB	85.04
9d19f1a9-b303-4063-ae32-01fda880743f	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	80.63
42975b41-891d-41f4-b299-080e31550e5d	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	80.71
e39c2c7f-060f-4424-9c1b-b069503b5ca6	Enterobacter cloacae	ARO:3003830|aminocoumarin	82.39
03f91857-86da-4acd-b36a-95b39e4d0eca	Enterobacter cloacae	ARO:3000216|acrB	84.55
d3a11e7f-30fe-4d10-a998-5cd9882dbe96	Enterobacter cloacae	ARO:3003830|aminocoumarin	82.33

b6b3bb67-bc02-49f4-9aa5-fa7fb3335e5e	Enterobacter cloacae	ARO:3000207|acrA	80.74
07f4777d-4693-439c-a26f-22de2f5eae01	Enterobacter cloacae	ARO:3001847|ACT-27	89.91
9acc8906-dfdc-4e80-bbf6-b3f115fcd24a	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	83.91
e69f3a9c-dad7-4b15-a734-3ce351cb4041	Enterobacter cloacae	ARO:3000024|patA	84.48
76344f68-6cb5-40bd-88bb-5d5778326e33	Enterobacter cloacae	ARO:3000074|emrB	81.48
d0265621-ea04-4c18-a0f7-2a097d6b3e70	Enterobacter cloacae	ARO:3000825|robA	91.45
dbe51358-3c9f-4f29-a0e7-5c195b3f4d7f	Enterobacter cloacae	ARO:3003923|oqxB	80.83
ab83a28c-d4d7-46ee-89b3-db32bb6ae83f	Enterobacter cloacae	ARO:3001847|ACT-27	81.37
d01f0f55-340b-4545-b89d-f8d80c654f69	Enterobacter cloacae	ARO:3003830|aminocoumarin	82.86
8862d213-54d0-4363-9c31-84f7498bd1d5	Staphylococcus epidermidis	ARO:3000391|norA	85.11
935800ce-aed8-42f2-8c19-a5fe69f5ed69	Enterobacter cloacae	ARO:3000216|acrB	86.86
30c1fd4b-1101-432a-8fbc-c706eadb14b7	Enterobacter cloacae	ARO:3000830|cpxA	84.08
5bab8604-4d1c-420a-bfff-93d87c814b4a	Enterobacter cloacae	ARO:3001847|ACT-27	85.97
180b8ec6-9b19-4990-b23f-f1bf20173fa3	Enterobacter cloacae	ARO:3000825|robA	90.08
8acb01c0-2599-4c62-a958-896ccd28395f	Staphylococcus epidermidis	ARO:3000391|norA	91.31
a01c2d35-c91c-4e21-aa1d-f7c6dcb1de38	Enterobacter cloacae	ARO:3001847|ACT-27	91.24
fb91fac1-49fc-46fc-9ccc-bf86ddbec371	Enterobacter cloacae	ARO:3000216|acrB	87.08
ac97b774-3adb-4e48-bb3d-cee5a4465fd3	Enterobacter cloacae	ARO:3000216|acrB	84.24
ee6a2d79-036d-4177-83c3-7816b8c57128	Enterobacter cloacae	ARO:3000825|robA	86.64
7b8963fc-94c2-47e2-94f1-32a6e1821c78	Enterobacter cloacae	ARO:3000518|CRP	80.46
ff6f2b7b-aa79-4d5b-9547-c1a0c3b231e1	Enterobacter cloacae	ARO:3001827|ACT-16	92.42
af55976d-db21-48ed-905a-db10c84ee82c	Enterobacter cloacae	ARO:3000825|robA	92.48
76966873-8e08-4f38-9992-41069b0503b9	Enterobacter cloacae	ARO:3000825|robA	92.08
baf97fe0-89da-4b16-b49e-e9c07019828a	Enterobacter cloacae	ARO:3000516|emrR	81.89
d372165f-1305-44b6-8d0b-286ed95d032b	Enterobacter cloacae	ARO:3000216|acrB	82.51
9d990e0b-32cb-4c5b-8b16-691186de4218	Enterobacter cloacae	ARO:3000216|acrB	86.82
7c856173-37f4-48b9-b4c1-d9b3b9a5ec55	Enterobacter cloacae	ARO:3000825|robA	87.39
86ce881e-c448-47f1-b038-b32af1b2043d	Enterobacter cloacae	ARO:3000216|acrB	84.02
b11eada9-5296-4d7c-b1b8-baf321341dfd	Enterobacter cloacae	ARO:3003830|aminocoumarin	86.55
acd97355-ff91-4e95-aad1-05828db01de1	Enterobacter cloacae	ARO:3000825|robA	87.08
2a7a928c-1fd7-4f20-9f7c-dfa23313f87e	Enterobacter cloacae	ARO:3000825|robA	83.28
49714ad4-15c2-4351-9634-2c006d74365e	Enterobacter cloacae	ARO:3003052|smeB	80.17
a03a7ac4-cf42-48ab-9a09-fed171602b86	Enterobacter cloacae	ARO:3003922|oqxA	82.15
88f7a13d-8793-48de-a437-dc31e9837f56	Enterobacter cloacae	ARO:3000825|robA	91.7
ff2f8d31-d44c-4227-956e-52bbbebe8c6e	Enterobacter cloacae	ARO:3001847|ACT-27	90.13
b599750a-1355-4d21-b728-a79104395d6c	Enterobacter cloacae	ARO:3001830|ACT-7	86.99
bdfbf434-7986-4863-b446-b60391be475e	Enterobacter cloacae	ARO:3000825|robA	86.9
7a750b24-299e-479e-b775-d0778f969d16	Enterobacter cloacae	ARO:3001830|ACT-7	81.64

87b8c612-a900-452e-8546-feb1e6adcb34	Enterobacter cloacae	ARO:3000825|robA	85.08
08606748-9de9-4d6c-ae26-7178df5ba919	Staphylococcus epidermidis	ARO:3000815|mgrA	80.13
7f1d8e27-c945-4a0f-90fe-0722c29cf966	Enterobacter cloacae	ARO:3003830|aminocoumarin	82.25
3c1fa6d5-ea7d-42ea-9270-64d14bcd8340	Enterobacter cloacae	ARO:3001847|ACT-27	85.35
745225f0-3508-47fe-ad8c-ee032f8ef516	Enterobacter cloacae	ARO:3003922|oqxA	81.36
3ad29102-931b-4e4a-9566-c35bfa3c4676	Enterobacter cloacae	ARO:3003923|oqxB	81.29
70605f8e-989e-4cd0-bb46-3f08841d7f38	Enterobacter cloacae	ARO:3003830|aminocoumarin	83.1
2ac12f94-4e03-41b4-9b95-359ea5d50b25	Enterobacter cloacae	ARO:3001847|ACT-27	84.27
4261f469-47d5-447e-b361-76c4f96f730e	Enterobacter cloacae	ARO:3000518|CRP	80.78
60acb22c-06a3-4f6f-9a31-665f617e738b	Enterobacter cloacae	ARO:3000825|robA	86.04
f69e6290-a2d9-435c-a97e-468eca40df09	Enterobacter cloacae	ARO:3000216|acrB	83.9
3d233b38-67e1-4b3e-b387-9754620b1895	Enterobacter cloacae	ARO:3001830|ACT-7	86.59
45ebefff-41a3-468a-a556-7e8acbc2d09c	Enterobacter cloacae	ARO:3000216|acrB	83.06
50695ab2-0773-47b5-93de-690f31615b5b	Enterobacter cloacae	ARO:3001830|ACT-7	84.22
bdf8fd43-64c1-4a98-83fd-31f970aae167	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.14
2955d096-c9c6-4016-8f4e-523fe99def25	Enterobacter cloacae	ARO:3000825|robA	88.57
3870b25d-5ca8-495c-893b-1179784cb2ea	Enterobacter cloacae	ARO:3000825|robA	87.68
cd6b68bb-24b1-4686-8bc8-805302c964c2	Enterobacter cloacae	ARO:3000216|acrB	82.35
cd6b68bb-24b1-4686-8bc8-805302c964c2	Enterobacter cloacae	ARO:3000207|acrA	80.12
0ed7ee3a-2798-4990-834b-e381b57cfd41	Enterobacter cloacae	ARO:3003923|oqxB	80.76
102a872f-06ce-44d3-aadc-7bd1d23bf4c7	Enterobacter cloacae	ARO:3000207|acrA	82.51
1de11fb5-bd40-4bd5-9e06-7e4c8974cd6b	Enterobacter cloacae	ARO:3000216|acrB	82.2
55b45c2d-ae52-4806-bf34-deeac9db9eb9	Enterobacter cloacae	ARO:3001830|ACT-7	84.04
7102b733-1118-40e6-8674-a2a747a7215b	Enterobacter cloacae	ARO:3000216|acrB	83.38
e809e1dd-be57-47cd-9720-11b48700c186	Enterobacter cloacae	ARO:3000676|H-NS	80.62
2ed6045d-cb9c-4c66-85c5-a38f84c27a9a	Staphylococcus epidermidis	ARO:3000391|norA	91.6
10c0c348-7d1d-481a-b473-70903097cd54	Enterobacter cloacae	ARO:3000024|patA	82.25
dee76079-bec5-42f2-b7fd-0dbd3af894e2	Enterobacter cloacae	ARO:3000825|robA	88.63
123fdf7f-0e40-4614-8d8a-ce446a9a18b9	Enterobacter cloacae	ARO:3001847|ACT-27	85.81
f1738fbb-082c-455f-a37f-c53d4a02e635	Enterobacter cloacae	ARO:3000825|robA	82.1
47add2d8-3a27-493b-88a2-aab83d6c8e7a	Enterobacter cloacae	ARO:3000516|emrR	85.02
3cd2cf89-4dd3-4b06-9503-bd0b82012a06	Staphylococcus sp.	ARO:3002865|dfrC	93.53
e2bc3aaa-d70b-4515-b93d-e0d9fb686c59	Enterobacter cloacae	ARO:3001855|ACT-35	90.64
df45a006-9a7a-4903-a3c4-fbfa9a8bffaa	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.87
3b5a5719-3959-467b-b449-d61b5670be0d	Enterobacter cloacae	ARO:3001830|ACT-7	86.41
13b27338-2936-4d77-9de1-92fc43ce9647	Enterobacter cloacae	ARO:3000216|acrB	88.04
13b27338-2936-4d77-9de1-92fc43ce9647	Enterobacter cloacae	ARO:3000207|acrA	80.98
92097703-993f-4510-84f8-d1d77adc92fb	Enterobacter cloacae	ARO:3000825|robA	88.26
a3236a89-6486-47ba-9d9a-08bda0948368	Enterobacter cloacae	ARO:3001847|ACT-27	84.71

9eb5afec-7fa1-4c29-a2bd-a0979bddef80	Enterobacter cloacae	ARO:3001830|ACT-7	88.34
d0fc27e0-59fc-4436-864c-0633fb0bc6a7	Enterobacter cloacae	ARO:3000216|acrB	83.83
d9d649ec-f17f-4da4-8ba0-5e1571cade00	Enterobacter cloacae	ARO:3000825|robA	86.22
7378552d-7df5-43ed-af73-aba395002533	Enterobacter cloacae	ARO:3000825|robA	86.6
745ec8b0-e3fc-4804-94cf-73dfbe723565	Enterobacter cloacae	ARO:3000825|robA	89.06
20e0b6df-7ce5-439e-a78e-5e69a4739aca	Enterobacter cloacae	ARO:3003923|oqxB	81.45
cfdbab9f-e466-4fa2-8b52-7095aa47b13e	Enterobacter cloacae	ARO:3001847|ACT-27	90.88
e41d5968-31ba-42c5-9aae-63cf883f38ea	Enterobacter cloacae	ARO:3003923|oqxB	80.28
4cf972fd-5714-457b-8cc8-fb13d69fae73	Enterobacter cloacae	ARO:3000216|acrB	86.86
0f62822f-595f-4ea7-af20-996d2800bdca	Enterobacter cloacae	ARO:3000024|patA	82.04
58a261eb-af50-4f8e-8274-a1d0957f793d	Enterobacter cloacae	ARO:3000216|acrB	85.62
19c6b05c-44f2-4ede-a48a-3b0b1f6f606b	Staphylococcus aureus	ARO:3000178|tetK	93.07
61db4f9d-f2a2-4e2e-8e2b-49d274a47ae8	Enterobacter cloacae	ARO:3001847|ACT-27	85.67
baab5c8f-4c7d-4025-b64b-de6ea8d3b772	Enterobacter cloacae	ARO:3000825|robA	83.05
aee7d4e5-6522-46d2-9166-0c63ac6cd533	Enterobacter cloacae	ARO:3001838|ACT-17	83.43
69818f09-17fe-4ca4-8c14-52fa18f175cd	Staphylococcus epidermidis	ARO:3002865|dfrC	86.8
4d9873dd-95a9-4e5a-90cf-b5dd5f5be000	Enterobacter cloacae	ARO:3000216|acrB	84.77
4183590e-b46d-4e48-a2b2-393f1f8b341a	Enterobacter cloacae	ARO:3000518|CRP	81.59
be59d6ba-efcd-4dac-a125-95b58781e076	Staphylococcus aureus	ARO:3000565|tet38	90.03
6d048e84-e508-49e5-928d-f2ff3be8c7ae	Enterobacter cloacae	ARO:3003052|smeB	80.27
65c6d41f-dc8e-40da-b31d-e39f57c9ee2d	Enterobacter cloacae	ARO:3000024|patA	82.68
4863462e-b815-46ff-ac0b-182fed55a7fa	Enterobacter cloacae	ARO:3001830|ACT-7	89.29
242197b8-c77d-44c8-a5ef-141d1c9a9f85	Enterobacter cloacae	ARO:3003830|aminocoumarin	84.95
b112784e-9b86-48f4-a0e4-7335c0b16089	Staphylococcus aureus	ARO:3000319|mphC	82.05
710d6982-c115-4090-abc1-db1a27ed4da0	Enterobacter cloacae	ARO:3003923|oqxB	80.22
b5392b9d-9fcf-4a6a-bdf6-cab8186af6bc	Enterobacter cloacae	ARO:3003830|aminocoumarin	81.15
4a3e07a7-d0c9-4ce0-82ce-e07d2ca9e58e	Enterobacter cloacae	ARO:3000825|robA	81.56
67709be3-d47f-4fe4-afa2-ca7725a031ec	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.58
0670f02b-a4a3-49d3-91fd-48bd315759f3	Enterobacter cloacae	ARO:3000516|emrR	81.36
45354294-6290-4840-8a6b-7c03c5ae7fff	Enterobacter cloacae	ARO:3000676|H-NS	82.18
e003b557-268d-4bc3-b17f-65c6108bf929	Enterobacter cloacae	ARO:3000216|acrB	85.87
e003b557-268d-4bc3-b17f-65c6108bf929	Enterobacter cloacae	ARO:3000207|acrA	88.4
9e65bad7-a126-4ac7-80f1-7267c38e3ae4	Enterobacter cloacae	ARO:3000825|robA	88.27
53ac5c0c-c464-499f-92da-0a7e147adf9c	Enterobacter cloacae	ARO:3003923|oqxB	83.76
a5681e7a-0fa7-43cb-8821-c2fb5f4d5374	Staphylococcus aureus	ARO:3000839|arlS	85.11
4614b18f-91d8-4560-983c-0444d3922207	Enterobacter cloacae	ARO:3000825|robA	90.77
49c03996-e325-49df-b620-22b1f9e19dec	Enterobacter cloacae	ARO:3000825|robA	87.74
a8db2f99-abab-44f4-bf41-017b7993094b	Enterobacter cloacae	ARO:3001827|ACT-16	86.76
980bf7ff-d76f-4cbb-b9fc-717ae22bb8aa	Enterobacter cloacae	ARO:3000024|patA	82.87

62cb336b-92c5-46a9-a052-e4e549016d85	Enterobacter cloacae	ARO:3000825|robA	87.11
15bc9687-72f4-49ea-8bc6-4db2c5651d5a	Enterobacter cloacae	ARO:3001830|ACT-7	90
d4f01cfe-4f9a-417f-9868-35ef35294de7	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	80.02
81b1f5ca-68d5-4ae7-99a8-f40dc421f1eb	Staphylococcus aureus	ARO:3000565|tet38	86.88
3d3693d5-d69e-424e-bdd2-ecd6ef8e422d	Enterobacter cloacae	ARO:3000518|CRP	80.27
d937ef46-bacc-4b72-8a99-255341bb587a	Enterobacter cloacae	ARO:3000825|robA	89.48
39f7ceb5-2708-409b-89d3-ffdb67c06ee0	Enterobacter cloacae	ARO:3000024|patA	82.38
c60da8ee-50ef-4811-862c-a9fdec598c84	Enterobacter cloacae	ARO:3000825|robA	86.75
6a569535-ad42-4068-a36e-bb44865ac3ea	Staphylococcus pasteuri	ARO:3003552|fusB	90.04
4245efaa-15f2-4462-8d52-4a9612939542	Enterobacter cloacae	ARO:3000825|robA	80.92
a5cc147f-d5cd-4d05-b5ec-2f6da0868ef4	Enterobacter cloacae	ARO:3000830|cpxA	80.53
a7f31cf0-5682-441b-a0c5-8791a124de30	Enterobacter cloacae	ARO:3000825|robA	85.77
4d8e3717-0326-46d2-bb92-ddab30826b5e	Enterobacter cloacae	ARO:3000825|robA	89.42
4cf7b469-1788-4431-9361-086cfda3c0b5	Enterobacter cloacae	ARO:3000825|robA	89.84
c657130d-1345-4dd4-b47b-1d3e393959ac	Bifidobacterium animalis	ARO:3003730|Bifidobacterium ileS	80.6
caa98cc3-4d83-445a-9b21-87f7cbbd807f	Enterobacter cloacae	ARO:3001852|ACT-32	86.27
1b4fbcdf-9518-4e6f-9ebb-93d9bdcd1343	Enterobacter cloacae	ARO:3003923|oqxB	80.71
7eb1cbc3-6d15-4f29-a7e5-b07096f9d7f2	Staphylococcus epidermidis	ARO:3000391|norA	85.27
0c6d0673-37d2-497b-9fbc-657f4a6b8258	Enterobacter cloacae	ARO:3000825|robA	90.58
421853b7-bded-4a09-ae38-b04eed8ad581	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	80.44
b3bc9126-59f2-4803-92df-06f30c688632	Enterobacter cloacae	ARO:3000216|acrB	83.09
26480c28-d561-4343-b7d8-9550b44ff7dc	Enterobacter cloacae	ARO:3001852|ACT-32	85.96
1265726e-fa61-4a8c-880c-94d2d3af12a0	Enterobacter cloacae	ARO:3000825|robA	90.18
f1f41081-bfec-43ad-af7b-0311f02d1839	Enterobacter cloacae	ARO:3000207|acrA	82.35
8c4c03b3-f421-4621-aa84-4e6bed8d1b48	Enterobacter cloacae	ARO:3001830|ACT-7	87.89
c86abdc8-e6ed-42da-aa05-fa26178ec3c0	Enterobacter cloacae	ARO:3000831|cpxR	80.18
d511931e-6f13-46b6-a26c-bb79acd406e3	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.69
31762543-e0f0-4da7-b9bc-edb98dc03462	Enterobacter cloacae	ARO:3000216|acrB	84.11
b081d4f8-6ef2-46bd-b200-5213b5ca5c57	Enterobacter cloacae	ARO:3000518|CRP	85.71
7479f87b-31f5-476c-8b43-be7123e3689a	Enterobacter cloacae	ARO:3003171|ACT-36	80
2a70f844-6a7e-4556-9490-fdc2c062b8ca	Enterobacter cloacae	ARO:3000518|CRP	82.61
d2930048-e3d1-4baa-91a1-d3272a55e912	Enterobacter cloacae	ARO:3003830|aminocoumarin	82.05
6619eb6a-42a0-43d8-bad7-9f9f0da357c0	Enterobacter cloacae	ARO:3001847|ACT-27	92.11
e9af0f71-1965-4689-87b7-f59d348114bc	Enterobacter cloacae	ARO:3003830|aminocoumarin	81.07
4f41b849-d87b-4237-ac8e-b52c4e806175	Enterobacter cloacae	ARO:3003923|oqxB	81.7
b4c1a038-b2ce-40cb-bacd-c2bb605e28ed	Enterobacter cloacae	ARO:3000216|acrB	84.36
7eec9347-b96e-4720-b737-11cf0114322f	Enterobacter cloacae	ARO:3000216|acrB	84.19
6529bb19-76fa-4a74-9d4b-794a50008839	Staphylococcus aureus	ARO:3000026|mepA	87.09

b32e718f-8dc4-43a4-a62b-9d56c89ebb81	Enterobacter cloacae	ARO:3000216|acrB	86.72
d7968db9-937a-49e7-9e37-e244b4a57767	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.42
7a52b746-26ba-4dde-95c9-e72bfff5c093	Enterobacter cloacae	ARO:3000216|acrB	81.65
0db6077b-5ddb-4df8-b779-7bca33e7e18e	Enterobacter cloacae	ARO:3003923|oqxB	82.44
27109828-c59f-4e9e-a200-f211c404c85d	Enterobacter cloacae	ARO:3000216|acrB	85.61
f83b40e5-62b4-4c1e-947c-c0c768fdf53e	Enterobacter cloacae	ARO:3000216|acrB	85.37
125484c0-cca3-46c3-a9bf-f49a5c2bd294	Enterobacter cloacae	ARO:3003923|oqxB	82.61
1bfb46ef-14dd-4478-b9fd-943356018e6b	Enterobacter cloacae	ARO:3000216|acrB	86.06
84de5e6c-1fdc-4043-956d-79bf2be5d8b9	Enterobacter cloacae	ARO:3000825|robA	88.87
edbbf562-2ba5-44aa-b760-d2e8bd7872b0	Enterobacter cloacae	ARO:3003923|oqxB	80.48
144dba86-2cec-4573-a6f6-8b926dc202b0	Enterobacter cloacae	ARO:3000216|acrB	88.81
a5bceeda-4362-4b17-b712-e42fc3a37d4f	Enterobacter cloacae	ARO:3000216|acrB	83.56
20d2ac74-c0b4-4c1b-b216-d754f2ceac7f	Enterobacter cloacae	ARO:3000216|acrB	82.95
dfaa8bb2-2b00-4fdf-a9b1-1255575056d5	Enterobacter cloacae	ARO:3000207|acrA	80.15
6fa36aaf-fc72-4d38-a579-4d310483c327	Enterobacter cloacae	ARO:3001847|ACT-27	81.4
c11e80a0-44d8-4555-a40d-10d47faa8095	Enterobacter cloacae	ARO:3003923|oqxB	83.75
bf16f014-d3e7-4950-8bbc-f3e9ca5b83a4	Enterobacter cloacae	ARO:3000024|patA	80.36
2723a479-637b-41e3-83a2-0e615ae6d38a	Enterobacter cloacae	ARO:3003923|oqxB	81.35
47605658-1db2-461a-bf67-7befaf114f97	Enterobacter cloacae	ARO:3000216|acrB	83.21
0610ae51-57a5-4a17-9a99-626932d55e12	Enterobacter cloacae	ARO:3003923|oqxB	82.23
1e725a77-f61e-4ec0-bc85-109fb5229773	Enterobacter cloacae	ARO:3000216|acrB	84.97
f513ca5a-4342-4342-9857-e8dcf6d52d1a	Enterobacter cloacae	ARO:3000216|acrB	87.15
8b19fec1-eab8-4d7b-b733-8506cca134f8	Enterobacter cloacae	ARO:3000216|acrB	81.16
e00c03e8-07e5-45bb-afa3-cc1ebac7ba94	Enterobacter cloacae	ARO:3000216|acrB	80.5
a43cd2f2-ea3b-43c2-9384-c40d72f4f9f7	Enterobacter cloacae	ARO:3003922|oqxA	82.48
d5a3c3c1-e95d-41f7-80ca-116133be327b	Enterobacter cloacae	ARO:3000216|acrB	82.77
e64515af-074f-4b21-99a2-918fff1f0532	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.37
cafd7216-9deb-41e8-a7cf-c3a700833c58	Enterobacter cloacae	ARO:3000825|robA	83.79
3da9f4c9-3835-41f0-94ec-7a296ac16436	Enterobacter cloacae	ARO:3000216|acrB	81.11
653a0230-6546-4b55-914e-6af0fa10d75f	Enterobacter cloacae	ARO:3000216|acrB	85.6
800c9247-0dc5-498a-ae8a-731b75f06a03	Enterobacter cloacae	ARO:3001830|ACT-7	85.43
f65b4bff-98a0-4dcb-8704-970c4c851b08	Enterobacter cloacae	ARO:3003923|oqxB	81.75
b31e93a0-2641-4721-ac5b-18658a63a855	Enterobacter cloacae	ARO:3003923|oqxB	80.06
12cbe5a0-88a7-4cbf-b3b9-e170324add88	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	81.05
b24288ba-c34f-4931-917a-b31d9ab501f4	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.75
b2427d87-99fa-4b44-b21a-42f4aaede0b8	Enterobacter cloacae	ARO:3000825|robA	83.66
f79b4a65-46c7-415a-a304-99c758e5c6c8	Enterobacter cloacae	ARO:3000216|acrB	83.96
badb38e1-69bc-4807-b5d2-2b1333e61b66	Enterobacter cloacae	ARO:3000216|acrB	87.76

ae9d6662-eb70-4b1a-abef-e982010716bb	Enterobacter cloacae	ARO:3000216|acrB	81.49
adda52a3-a077-4191-a284-8e62bb1395e7	Enterobacter cloacae	ARO:3000216|acrB	83.37
dd98abf7-d944-47f2-9839-2331ca2f4204	Enterobacter cloacae	ARO:3001847|ACT-27	91.7
7e545135-07d6-4d9c-b8ca-9de190ed67da	Enterobacter cloacae	ARO:3001830|ACT-7	84.53
58a4171a-0b66-4c21-8bc1-b58fe461fe26	Enterobacter cloacae	ARO:3000216|acrB	82.15
88ba283a-d701-4c61-ae8d-784ce73da985	Enterobacter cloacae	ARO:3003923|oqxB	80.87
8bffc7b7-a5cf-479c-ba2e-501a4222ea24	Enterobacter cloacae	ARO:3001852|ACT-32	88.77
27f74a85-bd12-4890-acf6-8ae5183080a2	Enterobacter cloacae	ARO:3000207|acrA	82.02
b5e0e619-054a-4919-803f-0981f0d95604	Enterobacter cloacae	ARO:3000216|acrB	82.25
04375fc5-2671-436a-97e7-529b75895cd1	Enterobacter cloacae	ARO:3000216|acrB	85.23
3755ce18-9476-40d5-b285-8e74f74844eb	Enterobacter cloacae	ARO:3003923|oqxB	82.34
fc459e79-f868-4325-9740-06af9b3e3302	Enterobacter cloacae	ARO:3000216|acrB	81.86
d49c37cd-9ce3-49a1-8ad0-2b7ecfc03d82	Enterobacter cloacae	ARO:3000024|patA	82.47
2a167c6c-91b2-4ad4-8994-6b33d55b35b7	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	85.97
1e797756-36ec-4c9c-844f-91c5782ccc3a	Enterobacter cloacae	ARO:3000216|acrB	85.1
a01f724f-d205-4dab-b5a4-06b5f18a29d9	Enterobacter cloacae	ARO:3001830|ACT-7	83.84
06c775d3-f1f8-45f8-92aa-a6ab47293992	Enterobacter cloacae	ARO:3000216|acrB	81.47
955e1891-4872-4bc4-bcca-4bbf2e92d19b	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	84.55
5d1d775d-656a-425b-a5e5-f6b290918f14	Enterobacter cloacae	ARO:3000825|robA	89.18
689348de-dc9d-454c-bd1c-36d3d09949e9	Enterobacter cloacae	ARO:3000216|acrB	83.65
f7854cff-4f7f-40ac-825b-2ecc9d2db956	Enterobacter cloacae	ARO:3000216|acrB	80.5
16d08a0e-32a3-4835-bb10-3d4ca03452d5	Enterobacter cloacae	ARO:3000216|acrB	84.76
58b4fdee-617b-4dc0-bf9f-1a7fd62d3255	Enterobacter cloacae	ARO:3000216|acrB	81.49
3f7e3627-c812-4162-aae1-6baf223e9372	Enterobacter cloacae	ARO:3000216|acrB	86.6
d94758df-80a0-4c4d-8c27-43c88beaeeda	Enterobacter cloacae	ARO:3000024|patA	80.98
58a6d5f6-9170-4435-b273-2ea02c0580e3	Staphylococcus epidermidis	ARO:3000391|norA	86.04
3b940043-2f74-48cb-9412-38c9b3a46327	Enterobacter cloacae	ARO:3000216|acrB	83.67
029eba72-9c14-44ad-80ec-42682ea29933	Enterobacter cloacae	ARO:3000216|acrB	83.06
ebf228e4-4e18-42c1-8af9-248780e86ca1	Enterobacter cloacae	ARO:3000216|acrB	85.57
41c5ea5f-278b-4122-8b40-39345fde3b0b	Enterobacter cloacae	ARO:3000216|acrB	82.47
60f5eae3-22ed-4b25-9ba9-c15bf81333e5	Enterobacter cloacae	ARO:3000216|acrB	81.94
88dcbacf-d778-4bec-ac39-448b38eecfe3	Enterobacter cloacae	ARO:3000216|acrB	86.24
63af0068-4de6-4ae5-89a2-d769f419b561	Enterobacter cloacae	ARO:3000216|acrB	80.97
34f8eec5-5ef8-4190-ab0c-e4991294b680	Enterobacter cloacae	ARO:3000216|acrB	81.24
94f4e554-7141-4801-a447-27af41d757cf	Enterobacter cloacae	ARO:3000216|acrB	83.1
ca192605-d280-43a2-9fad-c8ad24ae021f	Enterobacter cloacae	ARO:3000216|acrB	81.09
909675c8-d491-44c3-b2bc-15b997fc9111	Enterobacter cloacae	ARO:3003923|oqxB	80.85
6a6ed231-5efd-499d-ac09-03d55d2233b2	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	83.4

0290b8da-bb8e-4d6c-b14b-c78970871e2c	Enterobacter cloacae	ARO:3000516|emrR	80.21
7e4c40e2-7b03-4c07-b351-3fede5e6ddbe	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	80.56
60e4a1ea-40a9-4d63-9b83-3e835d5fdf84	Enterobacter cloacae	ARO:3000216|acrB	84.91
c5c1bbed-496c-4291-85e5-62d35ec11beb	Enterobacter cloacae	ARO:3000216|acrB	83.31
5f30e900-427f-4e12-ad5c-af6c0809578c	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	81.33
b1b4254b-e8b9-41d4-9045-fe5fe24e00dd	Enterobacter cloacae	ARO:3003923|oqxB	83.77
0cede423-40c1-471f-a362-8aedf69b76fd	Enterobacter cloacae	ARO:3000216|acrB	87.21
5e9f6456-c5c5-4e71-acb6-ab7ad52b82c3	Enterobacter cloacae	ARO:3001852|ACT-32	88.25
d6cb969f-9f0e-42fc-8890-c4a55129f9d4	Enterobacter cloacae	ARO:3000825|robA	90.09
96b0d7b3-a9c8-4866-baed-4ced7e97d4d1	Enterobacter cloacae	ARO:3000216|acrB	81.97
e3d2db0d-d00f-4ed2-96e3-4ec1f9f864aa	Enterobacter cloacae	ARO:3000216|acrB	84.05
64adbbf8-5968-460b-99d1-368385f10f18	Enterobacter cloacae	ARO:3000216|acrB	82.47
db76f47e-8c85-474f-bbca-81c5af8bb9d7	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.31
197551e6-aeb4-42e3-adc0-3925e8562aef	Enterobacter cloacae	ARO:3000216|acrB	83.48
a44ab211-0537-48b2-91a4-17f4f69ba469	Enterobacter cloacae	ARO:3000216|acrB	83.66
5b88fcbb-9026-484f-80d3-7e9173702499	Enterobacter cloacae	ARO:3000207|acrA	80.93
08f39e9a-82de-43ef-b3e7-4b6bc747d16a	Enterobacter cloacae	ARO:3003922|oqxA	81.42
bb2b2f5a-3986-476f-a643-1c717fa3499e	Enterobacter cloacae	ARO:3001830|ACT-7	87.05
568b229a-f9e2-4fe7-91ba-9ee419bc9507	Enterobacter cloacae	ARO:3001852|ACT-32	84.03
fecfff03-0a45-46db-8c6e-a38f759d59dc	Enterobacter cloacae	ARO:3000825|robA	91.16
426bdc5c-496e-4597-85e3-8aebba6375d5	Enterobacter cloacae	ARO:3001830|ACT-7	87.58
acdc7a1d-41bd-47d9-a383-9176c8140332	Enterobacter cloacae	ARO:3000216|acrB	84.51
184026b4-96b2-445c-b0cd-c31f97b17a8d	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	80.86
e890a10e-17b0-40ef-b20a-2fe6994e58e0	Enterobacter cloacae	ARO:3001830|ACT-7	88.23
05e64a0e-8f0a-480d-95a4-093627237006	Enterobacter cloacae	ARO:3000216|acrB	86.55
aed638fc-bbb2-43c7-8e6d-862eb3e14f62	Enterobacter cloacae	ARO:3001847|ACT-27	84.93
af429be7-0130-47ba-8fba-d4f1504c0044	Enterobacter cloacae	ARO:3000216|acrB	86.52
f2e3c00c-6a48-4b8b-808b-8f4bf4f178c6	Enterobacter cloacae	ARO:3000825|robA	89.84
5867ad01-119b-4235-ae53-14eb085ca7e9	Enterobacter cloacae	ARO:3000216|acrB	84.54
d9068b6b-bbff-410f-96ee-b01e2df6d6f7	Enterobacter cloacae	ARO:3000216|acrB	83.29
0b7c84ce-b53a-4208-ba4c-6288986f8cf7	Enterobacter cloacae	ARO:3000216|acrB	83.66
a399b430-3463-48e7-b357-9e9c5312cfdc	Enterobacter cloacae	ARO:3000216|acrB	80.01
c3d55761-568b-46ba-9288-250d2467fe5b	Enterobacter cloacae	ARO:3000216|acrB	84.14
a709fa42-0809-4bd9-ae9a-6a67b5e29431	Enterobacter cloacae	ARO:3000216|acrB	86.82
35121759-754d-4173-8fe7-e571d2d1eec8	Enterobacter cloacae	ARO:3003922|oqxA	80.34
27ef056f-15cb-4810-a7d6-1aa0ca6cc9e6	Enterobacter cloacae	ARO:3000216|acrB	85.32
1874c0be-96f8-47dd-9dfb-3853c4b8d89a	Enterobacter cloacae	ARO:3000216|acrB	87.57
bf2eb1e2-9eae-4d73-bf08-33458cdabe1b	Enterobacter cloacae	ARO:3000216|acrB	83.11

e468c043-fb0e-4698-a135-d18dbd7734e5	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.68
907c2a82-8183-4ae4-be49-35dbe0288623	Enterobacter cloacae	ARO:3001855|ACT-35	80.73
25b050ff-a629-46d8-a5e2-1ee0f8514cdf	Enterobacter cloacae	ARO:3000216|acrB	84.2
4b2a1a60-84c3-46e0-8cfb-d63f67dd9d4d	Enterobacter cloacae	ARO:3000216|acrB	84.14
8aec7a27-3a35-4f46-b22c-ad6574d56f49	Staphylococcus epidermidis	ARO:3000178|tetK	84.34
ae2af016-f0a4-4a7e-b2b9-13ac65cc5bee	Enterobacter cloacae	ARO:3000216|acrB	86.46
be9a070d-569e-4e2b-ba5e-ad86adabbcbb	Enterobacter cloacae	ARO:3000216|acrB	83
ba87282c-1121-4f7d-896d-efdc5755720b	Enterobacter cloacae	ARO:3000216|acrB	82.43
a548cef3-4979-4af1-85df-2c25b4909e7c	Enterobacter cloacae	ARO:3000825|robA	85.42
fbcbb9d1-0261-465e-9fe6-564c967ff51f	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	85.69
d7395d9b-ff74-47ed-9a94-8ba7a2ec0830	Enterobacter cloacae	ARO:3003923|oqxB	80.56
c7d0534e-7fa6-4ab6-b8bf-f5fb6b22e669	Enterobacter cloacae	ARO:3000216|acrB	85.71
2c09ea17-fcea-40b5-a776-fd0ada899f05	Enterobacter cloacae	ARO:3000216|acrB	81.7
2cd4b54e-cf09-487d-b1b5-81d4610e5809	Enterobacter cloacae	ARO:3000216|acrB	87.98
4dd7d0e8-5634-4cf1-98cf-1b0db485731b	Enterobacter cloacae	ARO:3001845|ACT-25	81.32
594a9df9-fc45-4eb5-a624-144aa175532c	Enterobacter cloacae	ARO:3000074|emrB	81.4
f0bf0a87-d629-4c3a-891a-9f97569d9d45	Enterobacter cloacae	ARO:3000216|acrB	84.52
90567112-b392-4905-98a9-74edc3933d5b	Enterobacter cloacae	ARO:3000825|robA	91.29
4f7baa88-b0b7-4151-b752-ccb473b6026e	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.49
5ce6987c-fdda-419b-9d41-1d312c3b2918	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	84.03
5aed02bf-ef31-4a6c-9bdb-599530d1fe4b	Enterobacter cloacae	ARO:3000216|acrB	84.85
b9c05677-b9c0-48df-95b3-70565447185e	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.29
076cb5e0-ff78-42ad-a198-f9735133bd16	Staphylococcus epidermidis	ARO:3000178|tetK	90.42
8842d1b9-81de-4a4c-aa91-5200ce8f7160	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	81.76
15bfa662-21ba-47ae-8edf-adbef9d53cfc	Enterobacter cloacae	ARO:3000216|acrB	83.1
6e5a871c-021c-4513-8f47-87cf32c3994e	Enterobacter cloacae	ARO:3000216|acrB	80.52
e5d34cf2-1baa-4418-a19e-9fbb09d069ba	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	81.2
dbe07992-4f20-44a4-a799-1aefbc3c839e	Enterobacter cloacae	ARO:3003922|oqxA	81.49
b06dd381-1fc1-4efc-8658-ae775041b46d	Enterobacter cloacae	ARO:3000825|robA	91.07
32a9e278-67e3-4227-8637-b089a89f48f0	Enterobacter cloacae	ARO:3000216|acrB	82.28
e34b20ec-5fd4-4d5c-8c18-9aff9e401e63	Enterobacter cloacae	ARO:3000825|robA	83.56
b268b1dd-625f-4aaa-b931-0a3d1f903bbc	Enterobacter cloacae	ARO:3001852|ACT-32	87.67
7ace70d6-76ab-40df-9927-e204cef94f24	Enterobacter cloacae	ARO:3000216|acrB	82.84
0a3f4c73-9541-4996-8a8f-c170c3a9ace8	Enterobacter cloacae	ARO:3000216|acrB	85.75
2f6bf428-e746-4b64-b943-e344c8623313	Enterobacter cloacae	ARO:3000216|acrB	85.3
0566c1ff-5793-4d1a-b2ca-d76ffab2205f	Enterobacter cloacae	ARO:3003923|oqxB	82.79
ea06077c-a7e5-4f01-8ecc-2a0e6486b761	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	82.4
c216f707-71df-427c-9d45-5838ad4e7927	Enterobacter cloacae	ARO:3001830|ACT-7	84.12

84408555-b1d4-4e3e-982c-2a8900005b09	Enterobacter cloacae	ARO:3000216|acrB	82.41
53a58acc-1918-4e28-bace-bea515b5f114	Enterobacter cloacae	ARO:3000216|acrB	83.33
74b0c43f-7b62-4a31-81ad-4ffc77166a60	Enterobacter cloacae	ARO:3001847|ACT-27	84.68
ec5a5ffa-c7e1-4dea-b558-81e91294cb18	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	80.44
b60e0c65-e3c6-41d4-8893-a358e0bb0c13	Enterobacter cloacae	ARO:3000216|acrB	81.25
22cf8878-9c88-42bf-9660-376a4b9e3f5f	Enterobacter cloacae	ARO:3000216|acrB	81.67
90293e18-41a8-4517-8101-12970195fc48	Enterobacter cloacae	ARO:3000216|acrB	84.69
28c6e144-4a0d-4a1e-847e-257b7538b0a9	Enterobacter cloacae	ARO:3000216|acrB	81.82
eadbe995-831f-4b80-91ae-92c66d1768ac	Enterobacter cloacae	ARO:3000825|robA	90.29
ac207589-ed6b-4e26-9520-3727099b22d2	Enterobacter cloacae	ARO:3003923|oqxB	81.29
405f0ec9-ca8d-4894-a4ca-219b82733461	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	80.51
d3d36f7a-a92f-448c-a22f-9915dbc2208c	Enterobacter cloacae	ARO:3001852|ACT-32	86.89
6777697d-b1fc-4d9c-84f6-db09a7afc16a	Enterobacter cloacae	ARO:3000825|robA	85.13
b6a35cce-f46b-4182-8546-b3126b375bda	Enterobacter cloacae	ARO:3003923|oqxB	81.72
ca8c9d57-ae2a-4cd8-9d64-97a97c7a6e35	Enterobacter cloacae	ARO:3000216|acrB	80.74
de2b9b47-61a6-4c3b-a6eb-ef1161c2a917	Enterobacter cloacae	ARO:3001852|ACT-32	83.78
23eb4ae5-8bce-4dee-a436-83d7ea392278	Enterobacter cloacae	ARO:3000216|acrB	83.36
d105fc30-fdb8-4fa2-b870-5499b47e74f0	Enterobacter cloacae	ARO:3001838|ACT-17	86.33
7c465b29-8292-4a97-b466-776f9f5ace35	Enterobacter cloacae	ARO:3000216|acrB	81.98
d6188072-1394-4fb0-a53e-e7174de29c86	Enterobacter cloacae	ARO:3001827|ACT-16	84.59
60a30b56-86fc-4f41-b7ff-08ee18a68d3b	Enterobacter cloacae	ARO:3001852|ACT-32	88.15
fd057855-7daa-4d52-b867-078648529096	Bifidobacterium longum	ARO:3003730|Bifidobacterium ileS	80.12
582cb57c-a2ed-4817-bdde-eeedd3cf5a2a	Enterobacter cloacae	ARO:3003922|oqxA	82.78
71e35480-d15b-45c0-9f10-5b4972837000	Enterobacter cloacae	ARO:3001845|ACT-25	82.89
0704e2c8-2c06-4ddb-bb3c-dfe9d00866d8	Enterobacter cloacae	ARO:3001852|ACT-32	83.74
fdc55421-36eb-4851-99d6-1bf5e5d04aea	Enterobacter cloacae	ARO:3000825|robA	84.45
e1fdbeeb-8985-46b5-bc5e-feda07db0662	Enterobacter cloacae	ARO:3000216|acrB	85.48
dd87e8f6-562c-4e37-b090-61dbe1802a9f	Enterobacter cloacae	ARO:3000216|acrB	80.9
a01f8751-9111-43cf-b4a4-a4489492f13f	Enterobacter cloacae	ARO:3001852|ACT-32	91.07
a7a45e38-8776-491c-8bd8-4bb40da5e369	Enterobacter cloacae	ARO:3001847|ACT-27	87.81
2f4fbe0a-ced8-42f6-a5c7-8ff2f8c1233c	Enterobacter cloacae	ARO:3003830|aminocoumarin	80.15
c3332952-d03a-457f-93f7-d877e4927941	Enterobacter cloacae	ARO:3000216|acrB	83.16
9169da3f-118c-4075-902f-2bb0c192ecca	Enterobacter cloacae	ARO:3000216|acrB	83.84
b3859b91-e83d-448e-aed8-0372e75059d5	Enterobacter cloacae	ARO:3000216|acrB	83.04
949e2260-6be9-4566-9a45-a1c05e940c23	Enterobacter cloacae	ARO:3000825|robA	86.09
f84def54-c041-4b40-8e7c-130bcd0467e5	Enterobacter cloacae	ARO:3000216|acrB	85.5
8c92e865-e976-44fb-8ace-242086bb26fc	Enterobacter cloacae	ARO:3001830|ACT-7	84.56
8c92e865-e976-44fb-8ace-242086bb26fc	Enterobacter cloacae	ARO:3001839|ACT-18	83.49
a5af37d7-00a4-4ffd-9831-da6f2d8417f4	Enterobacter cloacae	ARO:3000216|acrB	82.52

bd94a22d-d554-4322-a3a9-244cd4a2e435	Enterobacter cloacae	ARO:3000825|robA	87.03
6e772918-2d01-4cb9-9cdb-27c468d77b3d	Enterobacter cloacae	ARO:3000216|acrB	80.09
335dabc5-a1de-4dda-b3cc-20ee1d971c2a	Enterobacter cloacae	ARO:3000825|robA	90.36
fa605de9-c429-49c3-9ca8-3c39702c7516	Enterobacter cloacae	ARO:3001830|ACT-7	89.31
8822c019-a7e2-43a3-9850-8069fe36b569	Enterobacter cloacae	ARO:3001847|ACT-27	88.75
91df06cb-0c7a-4c68-9998-eea922f182ee	Enterobacter cloacae	ARO:3003923|oqxB	81.17
5588dae4-9d57-428c-a953-6cdfae47ca07	Enterobacter cloacae	ARO:3003923|oqxB	82.75
ebe65798-8ba2-413d-bf8f-b0124494c0ea	Enterobacter cloacae	ARO:3000216|acrB	81.93

Premature infant P49

ReadId	HostHit	CARDhit	PercentId
9f0f9a0b-6fe1-4a75-938f-dfc23d4db4d2	Escherichia coli	ARO:3001397|OXA-2	80.46
43962585-c65f-4aea-8206-b51480c82537	Enterobacter cloacae	ARO:3000825|robA	81.34
9f9cc129-8097-4d1f-97c8-aa0ccbbe756c	Enterobacter cloacae	ARO:3000825|robA	90.61
92eaf722-df44-4365-a8db-07d259356b37	Klebsiella pneumoniae	ARO:3001397|OXA-2	88.02
1e8a4a88-f0b8-423f-aba1-de54058ae3ca	Enterobacter cloacae	ARO:3000216|acrB	90.54
1f44caa1-0583-439a-b131-d9702fd434cc	Klebsiella pneumoniae	ARO:3001397|OXA-2	90.07
1f44caa1-0583-439a-b131-d9702fd434cc	Klebsiella pneumoniae	ARO:3001397|OXA-2	86.15
ba11cd47-e6ca-40d9-ba63-4ff556a8aa26	Enterobacter cloacae	ARO:3000825|robA	89.84
8e2be451-e215-42b9-8c61-688cd4e1bf9f	Klebsiella michiganensis	ARO:3000024|patA	80.45
2800067b-11af-4835-9d8a-accb104478d7	Enterobacter cloacae	ARO:3001839|ACT-18	90.83
f75b2cb7-dab3-4976-a535-30323fc6ddb5	Enterobacter cloacae	ARO:3001839|ACT-18	84.65
8f706682-d0dc-4ecd-9c6a-5c83b424fbd2	Enterobacter cloacae	ARO:3000074|emrB	81.17
f93c2543-62e6-4571-a016-6159520b5bbd	Klebsiella pneumoniae	ARO:3001397|OXA-2	81.7
e43018b6-dcda-4ad4-aa71-62142beb26c1	Enterobacter cloacae	ARO:3003923|oqxB	82.51

Premature infant P8

ReadId	HostHit	CARDhit	PercentId
91c109a9-4197-4db2-b4de-1899ed8405b1	Klebsiella pneumoniae	ARO:3003923|oqxB	83.18
91c109a9-4197-4db2-b4de-1899ed8405b1	Klebsiella pneumoniae	ARO:3003922|oqxA	84.34
3dbb4270-a451-4088-8514-3113ad8b11fd	Klebsiella pneumoniae	ARO:3002831|vgaC	87.07
e491ce14-95f8-40a3-979f-cfbe9fc3bd2c	Klebsiella pneumoniae	ARO:3003154|SHV-187	85.02
9b9162e2-097a-42b4-9eb5-d24ab62aca19	Escherichia coli	ARO:3002831|vgaC	85.34
ea828b01-5736-4bd9-b047-ef49d4a66c1f	Klebsiella pneumoniae	ARO:3002831|vgaC	86.53
4b76be2a-5061-44e3-a74c-9a3ed625231b	Klebsiella pneumoniae	ARO:3001181|SHV-137	83.95
a4d9e191-0bb6-4d4e-9537-bb2de59452a5	Klebsiella pneumoniae	ARO:3003923|oqxB	82.64
ef9078d6-de7f-44a8-b76b-3c3f63987f33	Klebsiella pneumoniae	ARO:3003923|oqxB	85.54
e4274fd5-7b82-4b49-8803-432a1109055e	Klebsiella pneumoniae	ARO:3003923|oqxB	88.68
deefa3bf-74ae-40ab-9f63-d257e4954b09	Klebsiella pneumoniae	ARO:3000676|H-NS	80.1
291250ff-f781-4fc5-9393-6fbf4248f720	Escherichia coli	ARO:3003922|oqxA	86.17

9eeabd98-9050-436f-b757-7ff5445b4c2f	Proteus mirabilis	ARO:3002670|cat	89.12
fabbb72e-615b-487a-83b4-786a6e57bb84	Klebsiella pneumoniae	ARO:3003923|oqxB	86.84
5708db1c-d312-4c9e-b710-b83e16df0824	Klebsiella pneumoniae	ARO:3001191|SHV-152	83.18
06a44d06-9a6e-4d99-bec5-014ea00c4e1a	Escherichia coli	ARO:3002831|vgaC	85.71
36efa0b2-29f5-49b9-81d5-bd63e854d92b	Shigella flexneri	ARO:3002601|aadA	86.25
36efa0b2-29f5-49b9-81d5-bd63e854d92b	Shigella flexneri	ARO:3002854|dfrA1	89.63
36efa0b2-29f5-49b9-81d5-bd63e854d92b	Shigella flexneri	ARO:3002895|sat-1	87.11
8fa05c8e-7b7b-47e8-8c94-e04760daffcc	Klebsiella pneumoniae	ARO:3003209|FosA5	84.04
978d00a7-e56f-4734-85c7-8f6468eb31fc	Klebsiella pneumoniae	ARO:3001068|SHV-9	84.33
94107794-e610-483e-a9b0-262b3a3b5ce1	Klebsiella pneumoniae	ARO:3003923|oqxB	83.43
11fe1704-d265-497c-b7d1-82b3177655b0	Shigella flexneri	ARO:3002601|aadA	92.4
11fe1704-d265-497c-b7d1-82b3177655b0	Shigella flexneri	ARO:3002895|sat-1	93.93
11fe1704-d265-497c-b7d1-82b3177655b0	Shigella flexneri	ARO:3002854|dfrA1	91.25
11dd0010-f065-4ef1-b1ee-b4a87a1e0234	Klebsiella pneumoniae	ARO:3003923|oqxB	87.22
11dd0010-f065-4ef1-b1ee-b4a87a1e0234	Klebsiella pneumoniae	ARO:3003922|oqxA	87.64
f420290b-60c0-44ba-b403-570d157ad714	Klebsiella pneumoniae	ARO:3000024|patA	82.44
8eda44d2-eb8f-4ca5-92fa-199d2b2bb191	Klebsiella pneumoniae	ARO:3003830|aminocoumarin	81.89
c418092f-f317-4930-b1eb-9333a3b50f8d	Klebsiella pneumoniae	ARO:3003923|oqxB	84.94
70249fee-7fda-405f-a88a-492e9daa3825	Klebsiella pneumoniae	ARO:3002831|vgaC	80.87
58e8f118-0279-4884-b815-4b1dad5cdabd	Klebsiella pneumoniae	ARO:3003922|oqxA	83.54
803acbc6-b004-42cb-8810-ff01ecfb6608	Proteus mirabilis	ARO:3002670|cat	84.75
7ae51fb0-8476-4f73-9648-4d5199e8f3f6	Klebsiella pneumoniae	ARO:3003922|oqxA	87.29
deb6ae0c-247e-48a4-ac54-61d9ffaa32d7	Klebsiella pneumoniae	ARO:3002831|vgaC	87.92
416fdf9a-524c-4c18-b0cd-1d6726577c2e	Klebsiella pneumoniae	ARO:3003922|oqxA	86.97
416fdf9a-524c-4c18-b0cd-1d6726577c2e	Klebsiella pneumoniae	ARO:3003923|oqxB	87.2
30a17c45-3589-4892-a21e-72656458044f	Escherichia fergusonii	ARO:3002831|vgaC	87.07
ae9eec6d-b8b7-488e-9c73-9a6792094653	Klebsiella pneumoniae	ARO:3001191|SHV-152	83.92
26f62185-131f-4784-8f0e-4e0fbd57d5bb	Escherichia coli	ARO:3002831|vgaC	88.28
2ef1633d-f513-4cd3-b8c0-e9b3ef7367a7	Klebsiella pneumoniae	ARO:3003922|oqxA	88.95
4ca686b0-7954-44d2-9bff-b810c912b6ed	Klebsiella pneumoniae	ARO:3000518|CRP	81.1
68aee99e-e2ca-43ae-98e4-afae4c8b5964	Klebsiella pneumoniae	ARO:3003923|oqxB	84.76
68aee99e-e2ca-43ae-98e4-afae4c8b5964	Klebsiella pneumoniae	ARO:3003922|oqxA	88.11
bd102361-6156-42ac-a750-353389c29ef5	Escherichia coli	ARO:3002831|vgaC	83.9
0274146f-d932-4ab9-8056-fd88a879dc56	Proteus mirabilis	ARO:3000177|tetJ	89.11
0213c7e0-4fc4-4c35-9f29-b1bc5d57bf76	Proteus mirabilis	ARO:3000177|tetJ	90.5
0e58708e-36be-4a0e-a6c6-c1b7420640cb	Escherichia coli	ARO:3002831|vgaC	82.82
d3e58f03-ef82-437d-84ea-d9ec9be82baa	Klebsiella pneumoniae	ARO:3003922|oqxA	84.68
2ecd147d-e3e2-49af-8e08-1c74ea752712	Escherichia coli	ARO:3002831|vgaC	84.48

9bf41674-83a4-409b-a5be-472b5c2ef163	Klebsiella pneumoniae	ARO:3003152|SHV-185	86.28
db145b6c-51bc-4701-8f89-defb8c008ee4	Klebsiella pneumoniae	ARO:3003923|oqxB	89.52
db145b6c-51bc-4701-8f89-defb8c008ee4	Klebsiella pneumoniae	ARO:3003922|oqxA	87.22
4b8cc3ab-656f-4027-ac04-e01ddf92c0f3	Escherichia coli	ARO:3002831|vgaC	89.22
c96c2a28-9d08-4572-80fc-da193ac5a80e	Klebsiella pneumoniae	ARO:3003923|oqxB	88.56
b519efda-0873-4fc1-91a9-8ad9bbe68f57	Klebsiella pneumoniae	ARO:3003209|FosA5	86.56
ca566715-e763-4e21-8aa9-796d6b3336f0	Klebsiella pneumoniae	ARO:3003923|oqxB	82.37
ca566715-e763-4e21-8aa9-796d6b3336f0	Klebsiella pneumoniae	ARO:3003922|oqxA	81.67
f8456b52-980d-48ac-bc09-b18e04a88a76	Klebsiella pneumoniae	ARO:3003209|FosA5	85
185577b4-5519-4674-a2db-5fcb3835bbb1	Klebsiella pneumoniae	ARO:3001096|SHV-38	86.98
730c231c-cee8-474d-857a-f09d4d085038	Klebsiella pneumoniae	ARO:3003922|oqxA	84.62
936730d0-438a-4a85-9274-75d172ef4c56	Proteus mirabilis	ARO:3002670|cat	87.08
d1a03b1e-69b0-4633-9168-8c11dbd21948	Klebsiella pneumoniae	ARO:3003922|oqxA	81.4
20a7ea01-94de-4074-9d51-db7e46a4ce28	Klebsiella pneumoniae	ARO:3001193|SHV-154	85.45
cb52655d-1a93-4570-ada9-f1c84a81b719	Klebsiella pneumoniae	ARO:3003209|FosA5	84.47
bf34b2b3-efca-455c-8c39-165c930b8e6a	Klebsiella pneumoniae	ARO:3003922|oqxA	87.39
bf34b2b3-efca-455c-8c39-165c930b8e6a	Klebsiella pneumoniae	ARO:3003923|oqxB	89.15
a05233b7-df27-41c3-b1f9-5f2598d9c26d	Klebsiella pneumoniae	ARO:3001137|SHV-83	81.03
ab12799c-a61e-46ba-9047-b0ad5e709718	Escherichia coli	ARO:3002831|vgaC	80.33
6f63c035-b433-4c90-95f6-025bbfdab1ea	Escherichia coli	ARO:3002831|vgaC	91.14
0272afc5-e974-4c2c-af3c-cff5b69919e2	Shigella flexneri	ARO:3002601|aadA	89.26
0272afc5-e974-4c2c-af3c-cff5b69919e2	Shigella flexneri	ARO:3002854|dfrA1	94.14
0272afc5-e974-4c2c-af3c-cff5b69919e2	Shigella flexneri	ARO:3002895|sat-1	86.58
85f7edfd-25d9-423e-ae59-b1d3f56bc263	Klebsiella pneumoniae	ARO:3002831|vgaC	91.6
f8d9bb9d-342c-474d-8be4-475cf72e3142	Klebsiella pneumoniae	ARO:3003155|SHV-188	84.44
be929281-3ee0-4beb-80ce-8b34da29dd8e	Klebsiella pneumoniae	ARO:3001191|SHV-152	85.16
6bcc0b9e-8f4e-4085-9f91-542c2018ee30	Klebsiella pneumoniae	ARO:3003922|oqxA	85.8
20034209-49a0-4c16-91bb-3c537f7d180b	Klebsiella pneumoniae	ARO:3003923|oqxB	80.59
20034209-49a0-4c16-91bb-3c537f7d180b	Klebsiella pneumoniae	ARO:3003922|oqxA	82
db271f06-e250-4e53-991a-7da4664b0bb1	Klebsiella pneumoniae	ARO:3003922|oqxA	81.61
613712fe-44d8-4dd6-935d-26b741a6194f	Escherichia coli	ARO:3002831|vgaC	88.41
f4a007b4-8e36-4a48-a154-8bdeea83f10b	Klebsiella pneumoniae	ARO:3003209|FosA5	84.65
3e415653-15e5-4d4c-86c6-c23ddadd3983	Salmonella sp.	ARO:3002621|aadA24	81.57
ea16dcc5-0a3c-4fbf-a26b-65000580f23a	Escherichia coli	ARO:3003923|oqxB	80.15
c6c91a73-365d-4e57-9822-c34a69794e5c	Klebsiella pneumoniae	ARO:3003209|FosA5	91.9
24e83392-0351-4b93-974d-3e4b479aa90f	Klebsiella pneumoniae	ARO:3000676|H-NS	82.33
28329d1f-39b8-4222-b803-75fb1fcec0cb	Pasteurella multocida	ARO:3002621|aadA24	85.54
ef24885e-3c4e-43c5-ae38-db3817ffa528	Klebsiella pneumoniae	ARO:3002831|vgaC	91.45
78bb2cb5-91c1-46af-8edf-8a3dfecfe0cf	Klebsiella pneumoniae	ARO:3001128|SHV-74	84.81

aaecf687-1cf0-494f-b15f-747a33782d67	Klebsiella pneumoniae	ARO:3002831|vgaC	88.16
48e15219-cc6c-44d4-b464-e3f0caefea66	Klebsiella pneumoniae	ARO:3003923|oqxB	80.45
e7610d2e-0e08-4f3b-acac-b588ad67b167	Klebsiella pneumoniae	ARO:3003923|oqxB	83.67
e7610d2e-0e08-4f3b-acac-b588ad67b167	Klebsiella pneumoniae	ARO:3003922|oqxA	81.17
e44ac170-6230-44a9-bb6a-8e26aa15ec2d	Escherichia coli	ARO:3002831|vgaC	89.29
74cf7c41-f97b-4390-98f4-af250139b1d7	Escherichia coli	ARO:3002831|vgaC	85.96
4181ee03-3184-445c-92ba-d4d32994e116	Escherichia fergusonii	ARO:3002831|vgaC	89.61
f6ecc73a-74d8-46c2-befc-793c923da76e	Klebsiella pneumoniae	ARO:3003923|oqxB	87.67
4afb2689-dfe4-4d52-b897-068d961bd28a	Shigella dysenteriae	ARO:3002854|dfrA1	80
464d8021-1a1b-420c-a538-107c5a10e25e	Klebsiella pneumoniae	ARO:3003209|FosA5	82.4
5b49a546-1311-4f04-98f0-91b55e9e4e08	Klebsiella pneumoniae	ARO:3001070|SHV-11	89.9
9d48e2fc-98e5-48c3-9050-0cf1081d24f3	Shigella flexneri	ARO:3002601|aadA	85.85
9d48e2fc-98e5-48c3-9050-0cf1081d24f3	Shigella flexneri	ARO:3002895|sat-1	82.59
9d48e2fc-98e5-48c3-9050-0cf1081d24f3	Shigella flexneri	ARO:3002854|dfrA1	83.57
062d12bf-c862-4a7c-b0c5-08a2707c51ba	Klebsiella pneumoniae	ARO:3001137|SHV-83	90.38
72d04edc-31ab-4dee-989e-10c138a4955f	Klebsiella pneumoniae	ARO:3003923|oqxB	84.43
5236051a-e03c-4b52-a6b7-d1e493d0ff53	Klebsiella pneumoniae	ARO:3003209|FosA5	85.34
b2868101-9c5e-4871-aea9-bf64436b6a98	Klebsiella pneumoniae	ARO:3002831|vgaC	88.46
a1c2e5ee-7fb7-4d09-ba29-9aa275115479	Shigella flexneri	ARO:3002601|aadA	81
a1c2e5ee-7fb7-4d09-ba29-9aa275115479	Shigella flexneri	ARO:3002854|dfrA1	83.2
a1c2e5ee-7fb7-4d09-ba29-9aa275115479	Shigella flexneri	ARO:3002895|sat-1	81.26
c0c55578-ccdd-4b27-95ba-ebd35df3e7bf	Klebsiella pneumoniae	ARO:3003923|oqxB	84.02
59183d11-cc29-4f16-8d41-5ea53ca3474b	Klebsiella pneumoniae	ARO:3001096|SHV-38	89.29
87adc2b6-fb40-46d2-a2a1-81f87086bd10	Klebsiella pneumoniae	ARO:3001123|SHV-69	80.6
131d1fb8-557c-428a-9f63-8e81849d3398	Klebsiella pneumoniae	ARO:3003209|FosA5	85.29
1bf87f12-2c88-4c19-8d73-0403b1e97f57	Escherichia coli	ARO:3002831|vgaC	90.13
85967a6a-cb34-4adb-9612-1988c01b8251	Klebsiella pneumoniae	ARO:3001123|SHV-69	84.14
e4bef367-82cf-4044-ac93-8fadc6a327a9	Klebsiella pneumoniae	ARO:3003154|SHV-187	88.59
05c1a531-ab36-4d4d-b87d-552cabe85db9	Klebsiella pneumoniae	ARO:3000794|mdtC	81.46
3139b264-7368-48e8-a01a-6466301330d1	Klebsiella pneumoniae	ARO:3001092|SHV-34	90.11
2d2e272a-0162-4328-9f25-d96168dd23d5	Escherichia coli	ARO:3002831|vgaC	91.95
8d336d92-9b3e-46ba-b4a1-ece8f6c4a07d	Proteus mirabilis	ARO:3002670|cat	80.92
d8bcfbf9-e353-4fdc-9f28-9ee62d09aece	Klebsiella pneumoniae	ARO:3001096|SHV-38	92.05
4928531a-fdc9-4fa9-9dac-c90fe4c356f6	Klebsiella pneumoniae	ARO:3003209|FosA5	82.52
91d8975a-10e2-40b6-9fb4-fc3e83047ae4	Shigella flexneri	ARO:3002601|aadA	89.34
906a5f6b-6477-4b4e-ac26-b4e3c7c2f109	Klebsiella pneumoniae	ARO:3002831|vgaC	87.65
a94be529-7648-4ba2-bd17-028d4d71189c	Klebsiella pneumoniae	ARO:3003922|oqxA	80.17
0ffb3dcd-e9e1-4e2f-849a-6b16103f6e44	Proteus mirabilis	ARO:3002670|cat	84.66
5e97ce63-356a-482c-9900-5b2ffed2336b	Klebsiella pneumoniae	ARO:3003154|SHV-187	90.02

1a6a2f2b-9d56-41bc-a02f-ffea55be4c29	Klebsiella pneumoniae	ARO:3003922|oqxA	83.85
5a6d3107-59ff-46f0-937f-4a248b43c5cc	Klebsiella pneumoniae	ARO:3003922|oqxA	85.19
ae154d69-70b6-443c-94b1-381d546ae864	Klebsiella pneumoniae	ARO:3003209|FosA5	86.73
11b36bea-b0da-4532-8b69-bd1359b5cb33	Klebsiella pneumoniae	ARO:3003922|oqxA	83.49
89b10ac5-4fe8-4ae9-b5a6-3b651e361945	Proteus mirabilis	ARO:3000177|tetJ	88.46
d71e1959-8cda-478d-bad2-41e03c59e741	Klebsiella pneumoniae	ARO:3003154|SHV-187	83.35
8e94e7ea-2d84-4c91-a3f6-739773b63a4e	Klebsiella pneumoniae	ARO:3003209|FosA5	83.95
644d0352-ee02-46c1-a934-9d8d28cf7198	Escherichia coli	ARO:3002831|vgaC	82.76
30cd6edf-265d-40a1-90c0-8f8d23a385df	Klebsiella pneumoniae	ARO:3003923|oqxB	88.54
a7b949ea-4723-4aef-9a17-093b094d9091	Klebsiella pneumoniae	ARO:3001161|SHV-112	80.83
78775584-6065-496f-88b9-5a568da14c97	Klebsiella pneumoniae	ARO:3003922|oqxA	82.19
2b4f4c3d-761c-4784-a4df-261abc84524b	Klebsiella pneumoniae	ARO:3003923|oqxB	87.96
2b4f4c3d-761c-4784-a4df-261abc84524b	Klebsiella pneumoniae	ARO:3003922|oqxA	93.67
24116c54-5aee-4baa-b830-c8946ce0fa22	Klebsiella pneumoniae	ARO:3000676|H-NS	80.55
91d184fa-5909-443f-bf87-a8230c2e4701	Klebsiella pneumoniae	ARO:3003209|FosA5	86.16
dcd773ab-ca81-470b-92c5-d2a9d4f80634	Klebsiella pneumoniae	ARO:3003209|FosA5	85.91
5dd293cd-91e7-47c5-a73a-d35b7f861e78	Klebsiella pneumoniae	ARO:3003923|oqxB	84.53
5d1fdce7-c78f-4cea-8cac-0ade6cb16e27	Klebsiella pneumoniae	ARO:3003209|FosA5	87.79
d7b4149d-7d7b-443e-9171-4f7329f8f8f6	Klebsiella pneumoniae	ARO:3002831|vgaC	86.13
7230612b-5daa-4f00-996a-be3604667bbe	Escherichia coli	ARO:3002831|vgaC	88.36
c473d5d3-2810-4b2d-8fa5-91c7b3cab428	Klebsiella pneumoniae	ARO:3003922|oqxA	86.74
d27bb6a8-8cd2-4304-bba5-d4c2e4c10193	Proteus mirabilis	ARO:3002670|cat	85.44
cb2b7ce6-286f-4c80-9536-a62e1cba3788	Klebsiella pneumoniae	ARO:3001137|SHV-83	88.33
26d78d67-dc51-432a-8aae-a5f583a9e438	Klebsiella pneumoniae	ARO:3003209|FosA5	81.75
eac54dfe-c098-4966-8332-bb0f899345a4	Klebsiella pneumoniae	ARO:3002831|vgaC	86.67
4755fb9a-ca22-4e07-bf0b-1c0e1ebfe473	Klebsiella pneumoniae	ARO:3003923|oqxB	84.75
b8d5b736-7871-4ac5-8bdb-faaba65cf6c5	Klebsiella pneumoniae	ARO:3003922|oqxA	82.04
22ae98e3-bd7a-4850-a8b1-1f6c91831f9b	Klebsiella pneumoniae	ARO:3002831|vgaC	87.87
e94c6a7c-df7e-40f9-8083-312b7d4f218a	Escherichia coli	ARO:3002831|vgaC	86.58
5e0a1236-4891-466e-a75e-ea1977c51d0b	Escherichia coli	ARO:3003923|oqxB	90.37
34ec617c-7b1b-46cf-be91-b8e5076475b7	Klebsiella pneumoniae	ARO:3003923|oqxB	80.49
8e9889d2-af51-454b-ab44-70c9c73a7822	Klebsiella pneumoniae	ARO:3003923|oqxB	82.42
8e9889d2-af51-454b-ab44-70c9c73a7822	Klebsiella pneumoniae	ARO:3003922|oqxA	80.31
d5374a8c-99e2-4689-8ac0-4d47d4378e21	Klebsiella pneumoniae	ARO:3003923|oqxB	82.93
4ba8db7b-fda5-41de-8e7c-d539e9845f5c	Klebsiella pneumoniae	ARO:3003923|oqxB	85.62
0366488f-579f-41e2-b29d-301350bc5f25	Klebsiella pneumoniae	ARO:3000024|patA	82.38
f57c3e0a-0117-4b7f-9853-8944dd18dc80	Escherichia coli	ARO:3002831|vgaC	86.64
97652b2b-ea18-456c-bba8-ccc27b43cf91	Klebsiella pneumoniae	ARO:3003923|oqxB	88.45

ec8556bb-3e71-4c47-8d88-30b84554d51e	Klebsiella pneumoniae	ARO:3003154|SHV-187	82.11
445b6f7d-94cc-4ec7-bc6b-5782daf10c7b	Klebsiella pneumoniae	ARO:3002831|vgaC	81.4
d04a054e-4de4-4e3e-94c8-fa0d3bae0904	Klebsiella pneumoniae	ARO:3000216|acrB	80.16
1abf83b8-cca3-48a7-99a3-a0d60e81bb8e	Escherichia coli	ARO:3003923|oqxB	88.32
dd526b48-a811-486a-91ea-0c2e8fa76b47	Klebsiella pneumoniae	ARO:3003923|oqxB	88.2
54a89240-f72b-4dd1-adea-5f6a1219bc0a	Klebsiella pneumoniae	ARO:3003923|oqxB	87.56
c6ab4c1e-a65f-46c6-9f5c-56606f406288	Klebsiella pneumoniae	ARO:3000793|mdtB	81.42
a54dff89-39cb-4a5a-8ed7-2f7322894cdf	Klebsiella pneumoniae	ARO:3000216|acrB	80.86
263b363a-581d-4929-8172-bcbce75abc1e	Klebsiella pneumoniae	ARO:3003923|oqxB	88.47
06588c97-8d0c-448b-a662-7ed7468df794	Klebsiella pneumoniae	ARO:3003922|oqxA	89.99
9860e432-7ac5-48e7-8391-97ea6181dd55	Klebsiella pneumoniae	ARO:3003923|oqxB	85.01
d084769d-7727-43a1-84cc-eadd5f778972	Klebsiella pneumoniae	ARO:3003923|oqxB	86.66
a28b8f7d-7a13-4b12-b89a-8a975d0bd8e8	Escherichia coli	ARO:3003923|oqxB	90.83
225d2d92-d924-42d4-86d6-56586dce52fb	Klebsiella pneumoniae	ARO:3003923|oqxB	88.49
6d12b038-0afd-4290-a35c-8da3dcf4dfde	Klebsiella pneumoniae	ARO:3003923|oqxB	85.5
591b3413-be96-42ee-9a63-71b1c42abfd9	Klebsiella pneumoniae	ARO:3000216|acrB	80.66
e58361b7-e4e4-4af5-8983-a2174d588b4a	Klebsiella pneumoniae	ARO:3003923|oqxB	89.96
407531b5-5749-4efd-b40a-4952a8ea3b74	Klebsiella pneumoniae	ARO:3003923|oqxB	83.76
ee541880-521a-4289-b2fc-b07712bbe554	Klebsiella pneumoniae	ARO:3003923|oqxB	84.77
0248ad22-c428-462c-9c6a-61fff7172b06	Klebsiella pneumoniae	ARO:3003923|oqxB	80.05
05bdc6c2-8333-458f-b05b-ebf8145be288	Klebsiella pneumoniae	ARO:3003923|oqxB	82.31
7380dc40-6398-4665-822a-497a2f60c482	Klebsiella pneumoniae	ARO:3003923|oqxB	89.89

References

1. health, G.m.f. (2019).
2. Taroncher-Oldenburg, G. (2018).
3. Evans, J.M., Morris, L.S. & Marchesi, J.R. The gut microbiome: the role of a virtual organ in the endocrinology of the host. *The Journal of endocrinology* **218**, R37-47 (2013).
4. Marchesi, J.R. et al. The gut microbiota and host health: a new clinical frontier. *Gut* **65**, 330-339 (2016).
5. Riaz Rajoka, M.S. et al. Interaction between diet composition and gut microbiota and its impact on gastrointestinal tract health. *Food Science and Human Wellness* **6**, 121-130 (2017).
6. Blander, J.M., Longman, R.S., Iliev, I.D., Sonnenberg, G.F. & Artis, D. Regulation of inflammation by microbiota interactions with the host. *Nature immunology* **18**, 851-860 (2017).
7. Lewis, B.B. & Pamer, E.G. Microbiota-Based Therapies for Clostridium difficile and Antibiotic-Resistant Enteric Infections. *Annual review of microbiology* **71**, 157-178 (2017).
8. Lane, E.R., Zisman, T.L. & Suskind, D.L. The microbiota in inflammatory bowel disease: current and therapeutic insights. *Journal of Inflammation Research* **10**, 63-73 (2017).
9. Turta, O. & Rautava, S. Antibiotics, obesity and the link to microbes - what are we doing to our children? *BMC Medicine* **14** (2016).
10. Candon, S. et al. Antibiotics in Early Life Alter the Gut Microbiome and Increase Disease Incidence in a Spontaneous Mouse Model of Autoimmune Insulin-Dependent Diabetes. *PLoS ONE* **10** (2015).
11. Ayukekbong, J.A., Ntemgwa, M. & Atabe, A.N. The threat of antimicrobial resistance in developing countries: causes and control strategies. *Antimicrobial Resistance & Infection Control* **6**, 47 (2017).
12. Lemon, K.P., Armitage, G.C., Relman, D.A. & Fischbach, M.A. Microbiota-targeted therapies: an ecological perspective. *Sci Transl Med* **4**, 137rv135-137rv135 (2012).
13. Sanders, M.E. Probiotics: Definition, Sources, Selection, and Uses. *Clinical Infectious Diseases* **46**, S58-S61 (2008).
14. Wilson, B.C., Vatanen, T., Cutfield, W.S. & O'Sullivan, J.M. The Super-Donor Phenomenon in Fecal Microbiota Transplantation. *Front Cell Infect Microbiol* **9**, 2 (2019).
15. Day, R.L., Harper, A.J., Woods, R.M., Davies, O.G. & Heaney, L.M. Probiotics: current landscape and future horizons. *Future science OA* **5**, Fso391 (2019).
16. Tamburini, S., Shen, N., Wu, H.C. & Clemente, J.C. The microbiome in early life: implications for health outcomes. *Nat Med* **22**, 713-722 (2016).
17. Jimenez, E. et al. Is meconium from healthy newborns actually sterile? *Research in microbiology* **159**, 187-193 (2008).
18. Arboleya, S. et al. Establishment and development of intestinal microbiota in preterm neonates. *FEMS Microbiol Ecol* **79**, 763-772 (2012).
19. Penders, J. et al. Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118**, 511-521 (2006).
20. Walker, W.A. The importance of appropriate initial bacterial colonization of the intestine in newborn, child, and adult health. *Pediatr Res* **82**, 387-395 (2017).
21. Mulder, I.E. et al. Restricting microbial exposure in early life negates the immune benefits associated with gut colonization in environments of high microbial diversity. *PLoS One* **6**, e28279 (2011).
22. Nagpal, R. et al. Evolution of gut Bifidobacterium population in healthy Japanese infants over the first three years of life: a quantitative assessment. *Scientific reports* **7**, 10097-10097 (2017).

23. Lee, J.-H. & O'Sullivan, D.J. Genomic insights into bifidobacteria. *Microbiol Mol Biol Rev* **74**, 378-416 (2010).
24. Bashiardes, S., Thaiss, Christoph A. & Elinav, E. It's in the Milk: Feeding the Microbiome to Promote Infant Growth. *Cell Metabolism* **23**, 393-394 (2016).
25. Lamberti, L.M., Fischer Walker, C.L., Noiman, A., Victora, C. & Black, R.E. Breastfeeding and the risk for diarrhea morbidity and mortality. *BMC public health* **11 Suppl 3**, S15 (2011).
26. Sorbara, M.T. & Pamer, E.G. Interbacterial mechanisms of colonization resistance and the strategies pathogens use to overcome them. *Mucosal Immunology* **12**, 1-9 (2019).
27. Bull, M.J. & Plummer, N.T. Part 1: The Human Gut Microbiome in Health and Disease. *Integrative medicine (Encinitas, Calif.)* **13**, 17-22 (2014).
28. Rowland, I. et al. Gut microbiota functions: metabolism of nutrients and other food components. *Eur J Nutr* **57**, 1-24 (2018).
29. den Besten, G. et al. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* **54**, 2325-2340 (2013).
30. Roy, C.C., Kien, C.L., Bouthillier, L. & Levy, E. Short-chain fatty acids: ready for prime time? *Nutrition in clinical practice : official publication of the American Society for Parenteral and Enteral Nutrition* **21**, 351-366 (2006).
31. Pickard, J.M., Zeng, M.Y., Caruso, R. & Núñez, G. Gut microbiota: Role in pathogen colonization, immune responses, and inflammatory disease. *Immunol Rev* **279**, 70-89 (2017).
32. Dzidic, M., Boix-Amorós, A., Selma-Royo, M., Mira, A. & Collado, M.C. Gut Microbiota and Mucosal Immunity in the Neonate. *Med Sci (Basel)* **6**, 56 (2018).
33. Fujimura, K.E. et al. Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med* **22**, 1187-1191 (2016).
34. Medina, M., Izquierdo, E., Ennahar, S. & Sanz, Y. Differential immunomodulatory properties of Bifidobacterium logum strains: relevance to probiotic selection and clinical applications. *Clinical and experimental immunology* **150**, 531-538 (2007).
35. Nicola, S. et al. Searching for the Perfect Homeostasis: Five Strains of Bifidobacterium longum From Centenarians Have a Similar Behavior in the Production of Cytokines. *J Clin Gastroenterol* **50 Suppl 2, Proceedings from the 8th Probiotics, Prebiotics & New Foods for Microbiota and Human Health meeting held in Rome, Italy on September 13-15, 2015**, S126-s130 (2016).
36. Park, J. et al. Short-chain fatty acids induce both effector and regulatory T cells by suppression of histone deacetylases and regulation of the mTOR-S6K pathway. *Mucosal Immunol* **8**, 80-93 (2015).
37. Dzidic, M., Boix-Amoros, A., Selma-Royo, M., Mira, A. & Collado, M.C. Gut Microbiota and Mucosal Immunity in the Neonate. *Med Sci (Basel)* **6** (2018).
38. WHO Preterm birth. (2018).
39. Lee, J.H. Catheter-related bloodstream infections in neonatal intensive care units. *Korean J Pediatr* **54**, 363-367 (2011).
40. Been, J.V. et al. Preterm birth and childhood wheezing disorders: a systematic review and meta-analysis. *PLoS Med* **11**, e1001596 (2014).
41. Haataja, P. et al. Asthma and atopic dermatitis in children born moderately and late preterm. *European journal of pediatrics* **175**, 799-808 (2016).
42. Shulhan, J., Dicken, B., Hartling, L. & Larsen, B.M. Current Knowledge of Necrotizing Enterocolitis in Preterm Infants and the Impact of Different Types of Enteral Nutrition Products. *Adv Nutr* **8**, 80-91 (2017).

43. Pammi, M. & Weisman, L.E. Late-onset sepsis in preterm infants: update on strategies for therapy and prevention. *Expert review of anti-infective therapy* **13**, 487-504 (2015).
44. Alfirevic, Z., Milan, S.J. & Livio, S. Caesarean section versus vaginal delivery for preterm birth in singletons. *Cochrane Database of Systematic Reviews* (2013).
45. Wampach, L. et al. Birth mode is associated with earliest strain-conferred gut microbiome functions and immunostimulatory potential. *Nat Commun* **9**, 5091-5091 (2018).
46. D'Agata, A.L. et al. Effects of early life NICU stress on the developing gut microbiome. *Developmental psychobiology* **61**, 650-660 (2019).
47. Staude, B. et al. The Microbiome and Preterm Birth: A Change in Paradigm with Profound Implications for Pathophysiologic Concepts and Novel Therapeutic Strategies. *Biomed Res Int* **2018**, 7218187 (2018).
48. Dahl, C. et al. Preterm infants have distinct microbiomes not explained by mode of delivery, breastfeeding duration or antibiotic exposure. *International journal of epidemiology* **47**, 1658-1669 (2018).
49. Korpela, K. et al. Intestinal microbiota development and gestational age in preterm neonates. *Scientific Reports* **8**, 2453 (2018).
50. Warrier, I., Du, W., Natarajan, G., Salari, V. & Aranda, J. Patterns of drug utilization in a neonatal intensive care unit. *Journal of clinical pharmacology* **46**, 449-455 (2006).
51. Tripathi, N., Cotten, C.M. & Smith, P.B. Antibiotic use and misuse in the neonatal intensive care unit. *Clin Perinatol* **39**, 61-68 (2012).
52. van den Anker, J. & Allegaert, K. Rational Use of Antibiotics in Neonates: Still in Search of Tailored Tools. *Healthcare (Basel)* **7** (2019).
53. Wang, M. et al. Reduced diversity in the early fecal microbiota of infants with atopic eczema. *J Allergy Clin Immunol* **121**, 129-134 (2008).
54. Bisgaard, H. et al. Childhood asthma after bacterial colonization of the airway in neonates. *New England Journal of Medicine* **357**, 1487-1495 (2007).
55. Gibson, M.K. et al. Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nature Microbiology* **1**, 16024 (2016).
56. Willing, B.P., Russell, S.L. & Finlay, B.B. Shifting the balance: antibiotic effects on host–microbiota mutualism. *Nat Rev Micro* **9**, 233-243 (2011).
57. Hu, Y., Yang, X., Lu, N. & Zhu, B. The abundance of antibiotic resistance genes in human guts has correlation to the consumption of antibiotics in animal. *Gut microbes* **5**, 245-249 (2014).
58. Forslund, K., Sunagawa, S., Coelho, L.P. & Bork, P. Metagenomic insights into the human gut resistome and the forces that shape it. *BioEssays : news and reviews in molecular, cellular and developmental biology* **36**, 316-329 (2014).
59. van Schaik, W. The human gut resistome. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* **370**, 20140087 (2015).
60. Lu, N. et al. DNA microarray analysis reveals that antibiotic resistance-gene diversity in human gut microbiota is age related. *Sci Rep* **4**, 4302 (2014).
61. Gregory, K.E. & Walker, W.A. Immunologic Factors in Human Milk and Disease Prevention in the Preterm Infant. *Curr Pediatr Rep* **1**, 10.1007/s40124-40013-40028-40122 (2013).
62. Xu, W. et al. Systematic Review of the Effect of Enteral Feeding on Gut Microbiota in Preterm Infants. *J Obstet Gynecol Neonatal Nurs* **47**, 451-463 (2018).
63. Underwood, M.A. Human milk for the premature infant. *Pediatr Clin North Am* **60**, 189-207 (2013).

64. Quigley, M. & McGuire, W. Formula versus donor breast milk for feeding preterm or low birth weight infants. *The Cochrane database of systematic reviews*, Cd002971 (2014).
65. Xu, W. et al. Systematic Review of the Effect of Enteral Feeding on Gut Microbiota in Preterm Infants. *J Obstet Gynecol Neonatal Nurs* **47**, 451-463 (2018).
66. Silverman, M.A., Konnikova, L. & Gerber, J.S. Impact of Antibiotics on Necrotizing Enterocolitis and Antibiotic-Associated Diarrhea. *Gastroenterology clinics of North America* **46**, 61-76 (2017).
67. Hosny, M., Cassir, N. & La Scola, B. Updating on gut microbiota and its relationship with the occurrence of necrotizing enterocolitis. *Human Microbiome Journal* **4**, 14-19 (2017).
68. Claud, E.C. & Walker, W.A. Hypothesis: inappropriate colonization of the premature intestine can cause neonatal necrotizing enterocolitis. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **15**, 1398-1403 (2001).
69. Dobbler, P.T. et al. Low Microbial Diversity and Abnormal Microbial Succession Is Associated with Necrotizing Enterocolitis in Preterm Infants. *Frontiers in Microbiology* **8** (2017).
70. Zhang, C. et al. Paneth cell ablation in the presence of Klebsiella pneumoniae induces necrotizing enterocolitis (NEC)-like injury in the small intestine of immature mice. *Disease models & mechanisms* **5**, 522-532 (2012).
71. Schonherr-Hellec, S. et al. Clostridial Strain-Specific Characteristics Associated with Necrotizing Enterocolitis. *Appl Environ Microbiol* **84** (2018).
72. Ward, D.V. et al. Metagenomic Sequencing with Strain-Level Resolution Implicates Uropathogenic E. coli in Necrotizing Enterocolitis and Mortality in Preterm Infants. *Cell reports* **14**, 2912-2924 (2016).
73. Cortese, F. et al. Early and Late Infections in Newborns: Where Do We Stand? A Review. *Pediatrics & Neonatology* **57**, 265-273 (2016).
74. Cortese, F. et al. Early and Late Infections in Newborns: Where Do We Stand? A Review. *Pediatrics and neonatology* (2015).
75. Chmielarczyk, A. et al. Bloodstream Infections due to Enterobacteriaceae Among Neonates in Poland--Molecular Analysis of the Isolates. *Polish journal of microbiology / Polskie Towarzystwo Mikrobiologow = The Polish Society of Microbiologists* **64**, 217-225 (2015).
76. Ali, G.Y., Algohary, E.H., Rashed, K.A., Almoghanum, M. & Khalifa, A.A. Prevalence of Candida colonization in preterm newborns and VLBW in neonatal intensive care unit: role of maternal colonization as a risk factor in transmission of disease. *The journal of maternal-fetal & neonatal medicine : the official journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstet* **25**, 789-795 (2012).
77. Aceti, A. et al. Probiotics for prevention of necrotizing enterocolitis in preterm infants: systematic review and meta-analysis. *Italian journal of pediatrics* **41**, 89 (2015).
78. Aceti, A. et al. Probiotics Prevent Late-Onset Sepsis in Human Milk-Fed, Very Low Birth Weight Preterm Infants: Systematic Review and Meta-Analysis. *Nutrients* **9** (2017).
79. Milani, C. et al. Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. *PLoS One* **8**, e68739 (2013).
80. Yarza, P. et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Micro* **12**, 635-645 (2014).

81. Sharpton, T.J. An introduction to the analysis of shotgun metagenomic data. (2014).
82. Sharpton, T.J. An introduction to the analysis of shotgun metagenomic data. *Front Plant Sci* **5**, 209 (2014).
83. Clemente, J.C. et al. The microbiome of uncontacted Amerindians. *Science advances* **1** (2015).
84. Clemente, J.C. et al. The microbiome of uncontacted Amerindians. *Science advances* **1**, e1500183 (2015).
85. Didelot, X., Walker, A.S., Peto, T.E., Crook, D.W. & Wilson, D.J. Within-host evolution of bacterial pathogens. *Nature Reviews Microbiology* (2016).
86. Mwangi, M.M. et al. Tracking the in vivo evolution of multidrug resistance in Staphylococcus aureus by whole-genome sequencing. *Proceedings of the National Academy of Sciences* **104**, 9451-9456 (2007).
87. Takii, T. et al. Whole-genome sequencing-based epidemiological analysis of anti-tuberculosis drug resistance genes in Japan in 2007: Application of the Genome Research for Asian Tuberculosis (GReAT) database. *Scientific Reports* **9**, 12823 (2019).
88. Marcobal, A. et al. A metabolomic view of how the human gut microbiota impacts the host metabolome using humanized and gnotobiotic mice. *Isme j* **7**, 1933-1943 (2013).
89. Turi, K.N., Romick-Rosendale, L., Ryckman, K.K. & Hartert, T.V. A review of metabolomics approaches and their application in identifying causal pathways of childhood asthma. *J Allergy Clin Immunol* **141**, 1191-1201 (2018).
90. Zierer, J. et al. The fecal metabolome as a functional readout of the gut microbiome. *Nature Genetics* **50**, 790-795 (2018).
91. Bridgman, S.L. et al. Fecal Short-Chain Fatty Acid Variations by Breastfeeding Status in Infants at 4 Months: Differences in Relative versus Absolute Concentrations. *Frontiers in nutrition* **4**, 11 (2017).
92. Tan, J. et al. in Advances in Immunology, Vol. 121. (ed. F.W. Alt) 91-119 (Academic Press, 2014).
93. Alcon-Giner, C. et al. Optimisation of 16S rRNA gut microbiota profiling of extremely low birth weight infants. *BMC genomics* **18**, 841-841 (2017).
94. WHO (2015).
95. Gritz, E.C. & Bhandari, V. The human neonatal gut microbiome: a brief review. *Frontiers in pediatrics* **3**, 17-17 (2015).
96. Rooks, M.G. & Garrett, W.S. Gut microbiota, metabolites and host immunity. *Nat Rev Immunol* **16**, 341-352 (2016).
97. Rodríguez, J.M. et al. The composition of the gut microbiota throughout life, with an emphasis on early life. *Microbial Ecology in Health and Disease* **26**, 10.3402/mehd.v3426.26050 (2015).
98. Sherman, M.P., Zaghouani, H. & Niklas, V. Gut microbiota, the immune system, and diet influence the neonatal gut-brain axis. *Pediatric research* **77**, 127-135 (2015).
99. Groer, M.W. et al. Development of the preterm infant gut microbiome: a research priority. *Microbiome* **2**, 38 (2014).
100. Samuels, N., van de Graaf, R.A., de Jonge, R.C.J., Reiss, I.K.M. & Vermeulen, M.J. Risk factors for necrotizing enterocolitis in neonates: a systematic review of prognostic studies. *BMC Pediatrics* **17**, 105 (2017).
101. Wandro, S. et al. The Microbiome and Metabolome of Preterm Infant Stool Are Personalized and Not Driven by Health Outcomes, Including Necrotizing Enterocolitis and Late-Onset Sepsis. *mSphere* **3** (2018).
102. Kwak, M.J. et al. Evolutionary architecture of the infant-adapted group of Bifidobacterium species associated with the probiotic function. *Systematic and applied microbiology* **39**, 429-439 (2016).

103. Mikami, K., Kimura, M. & Takahashi, H. Influence of Maternal Bifidobacteria on the Development of Gut Bifidobacteria in Infants. *Pharmaceuticals* **5**, 629-642 (2012).
104. Yarza, P. et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* **12**, 635-645 (2014).
105. Caporaso, J.G. et al. QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335-336 (2010).
106. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614-620 (2013).
107. Caporaso, J.G. et al. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* **26**, 266-267 (2009).
108. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460-2461 (2010).
109. Gordon, A. & Hannon, G. Fastx-toolkit. *FASTQ/A short-reads preprocessing tools (unpublished) http://hannonlab. cshl. edu/fastx_toolkit* **5** (2010).
110. Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research* **41**, D590-D596 (2012).
111. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *Journal of molecular biology* **215**, 403-410 (1990).
112. Huson, D. et al. (preparation, 2015).
113. Gordon, A. & Hannon, G. Fastx-toolkit. FASTQ/A short-reads pre-processing tools. *Unpublished Available online at: http://hannonlab. cshl. edu/fastx_toolkit* (2010).
114. Pruitt, K.D., Tatusova, T. & Maglott, D.R. NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic acids research* **33**, D501-D504 (2005).
115. Buchfink, B., Xie, C. & Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nature methods* **12**, 59-60 (2015).
116. Ye, J., McGinnis, S. & Madden, T.L. BLAST: improvements for better sequence analysis. *Nucleic acids research* **34**, W6-W9 (2006).
117. Genedoc, S. (2016).
118. Gerasimidis, K. et al. The effect of DNA extraction methodology on gut microbiota research applications. *BMC Research Notes* **9**, 365 (2016).
119. Rintala, A. et al. Gut Microbiota Analysis Results Are Highly Dependent on the 16S rRNA Gene Target Region, Whereas the Impact of DNA Extraction Is Minor. *Journal of biomolecular techniques : JBT* **28**, 19-30 (2017).
120. Shaw, A.G. et al. Late-Onset Bloodstream Infection and Perturbed Maturation of the Gastrointestinal Microbiota in Premature Infants. *PLoS One* **10**, e0132923 (2015).
121. Sela, D.A. Bifidobacterial utilization of human milk oligosaccharides. *International journal of food microbiology* **149**, 58-64 (2011).
122. Fanning, S. et al. Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 2108-2113 (2012).
123. Milani, C. et al. Evaluation of bifidobacterial community composition in the human gut by means of a targeted amplicon sequencing (ITS) protocol. *FEMS microbiology ecology* **90**, 493-503 (2014).
124. Sim, K. et al. Improved detection of bifidobacteria with optimised 16S rRNA-gene based pyrosequencing. *PloS one* **7**, e32543 (2012).
125. Mammedov, T.G. et al. A Fundamental Study of the PCR Amplification of GC-Rich DNA Templates. *Computational biology and chemistry* **32**, 452-457 (2008).

126. Frey, U.H., Bachmann, H.S., Peters, J. & Siffert, W. PCR-amplification of GC-rich regions:'slowdown PCR'. *Nature protocols* **3**, 1312 (2008).
127. Service, C.R.a.D.I. (2015).
128. Pollock, J., Glendinning, L., Wisedchanwet, T. & Watson, M. The Madness of Microbiome: Attempting To Find Consensus "Best Practice" for 16S Microbiome Studies. *Appl Environ Microbiol* **84** (2018).
129. Angly, F.E. et al. CopyRighter: a rapid tool for improving the accuracy of microbial community profiles through lineage-specific gene copy number correction. *Microbiome* **2**, 11 (2014).
130. Stoddard, S.F., Smith, B.J., Hein, R., Roller, B.R. & Schmidt, T.M. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. *Nucleic Acids Res* **43**, D593-598 (2015).
131. Deshpande, G.C., Rao, S.C., Keil, A.D. & Patole, S.K. Evidence-based guidelines for use of probiotics in preterm neonates. *BMC Medicine* **9**, 92-92 (2011).
132. Tarnow-Mordi, W. & Soll, R.F. Probiotic supplementation in preterm infants: it is time to change practice. *J Pediatr* **164**, 959-960 (2014).
133. Turroni, F. et al. Diversity of bifidobacteria within the infant gut microbiota. *PLoS One* **7**, e36957 (2012).
134. Turroni, F. et al. The infant gut microbiome as a microbial organ influencing host well-being. *Italian journal of pediatrics* **46**, 16 (2020).
135. Ruiz, L., Delgado, S., Ruas-Madiedo, P., Sanchez, B. & Margolles, A. Bifidobacteria and Their Molecular Communication with the Immune System. *Front Microbiol* **8**, 2345 (2017).
136. Hartling, L., Shulhan, J., Larsen, B.M. & Dicken, B. Current Knowledge of Necrotizing Enterocolitis in Preterm Infants and the Impact of Different Types of Enteral Nutrition Products. *Advances in Nutrition* **8**, 80-91 (2017).
137. Jacobs, S.E. et al. Probiotic effects on late-onset sepsis in very preterm infants: a randomized controlled trial. *Pediatrics* **132**, 1055-1062 (2013).
138. Olsen, R., Greisen, G., Schroder, M. & Brok, J. Prophylactic Probiotics for Preterm Infants: A Systematic Review and Meta-Analysis of Observational Studies. *Neonatology* **109**, 105-112 (2016).
139. Dermyshi, E. et al. The "Golden Age" of Probiotics: A Systematic Review and Meta-Analysis of Randomized and Observational Studies in Preterm Infants. *Neonatology* **112**, 9-23 (2017).
140. Costeloe, K., Hardy, P., Juszczak, E., Wilks, M. & Millar, M.R. Bifidobacterium breve BBG-001 in very preterm infants: a randomised controlled phase 3 trial. *Lancet (London, England)* **387**, 649-660 (2016).
141. Deshpande, G., Rao, S., Athalye-Jape, G., Conway, P. & Patole, S. Probiotics in very preterm infants: the PiPS trial. *The Lancet* **388**, 655 (2016).
142. Robertson, C. et al. Incidence of necrotising enterocolitis before and after introducing routine prophylactic Lactobacillus and Bifidobacterium probiotics. *Archives of disease in childhood. Fetal and neonatal edition* (2019).
143. Page, A.J. et al. Robust high-throughput prokaryote de novo assembly and improvement pipeline for Illumina data. *Microbial genomics* **2**, e000083 (2016).
144. Brown-Elliott, B.A., Killingley, J., Vasireddy, S., Bridge, L. & Wallace, R.J., Jr. In Vitro Comparison of Ertapenem, Meropenem, and Imipenem against Isolates of Rapidly Growing Mycobacteria and Nocardia by Use of Broth Microdilution and Etest. *J Clin Microbiol* **54**, 1586-1592 (2016).
145. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* **7**, 562 (2012).

146. Beckonert, O. et al. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nature Protocols* **2**, 2692 (2007).
147. Alcon-Giner, C. et al. Optimisation of 16S rRNA gut microbiota profiling of extremely low birth weight infants. *BMC genomics* **18**, 841 (2017).
148. Wang, M. et al. Fecal microbiota composition of breast-fed infants is correlated with human milk oligosaccharides consumed. *Journal of pediatric gastroenterology and nutrition* **60**, 825-833 (2015).
149. Staude, B. et al. The Microbiome and Preterm Birth: A Change in Paradigm with Profound Implications for Pathophysiologic Concepts and Novel Therapeutic Strategies. *BioMed research international* **2018**, 7218187-7218187 (2018).
150. Ting, J.Y. et al. Duration of Initial Empirical Antibiotic Therapy and Outcomes in Very Low Birth Weight Infants. *Pediatrics* **143** (2019).
151. Zmora, N. et al. Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features. *Cell* **174**, 1388-1405.e1321 (2018).
152. Wang, C. et al. Intestinal Microbiota Profiles of Healthy Pre-School and School-Age Children and Effects of Probiotic Supplementation. *Annals of Nutrition and Metabolism* **67**, 257-266 (2015).
153. The European Committee on Antimicrobial Susceptibility Testing - EUCAST 2018. (2018).
154. Pokusaeva, K., Fitzgerald, G.F. & van Sinderen, D. Carbohydrate metabolism in Bifidobacteria. *Genes & nutrition* **6**, 285-306 (2011).
155. Fukuda, S., Toh, H., Taylor, T.D., Ohno, H. & Hattori, M. Acetate-producing bifidobacteria protect the host from enteropathogenic infection via carbohydrate transporters. *Gut microbes* **3**, 449-454 (2012).
156. Gotoh, A. et al. Sharing of human milk oligosaccharides degradants within bifidobacterial communities in faecal cultures supplemented with Bifidobacterium bifidum. *Sci Rep* **8**, 13958 (2018).
157. Henrick, B.M. et al. Elevated Fecal pH Indicates a Profound Change in the Breastfed Infant Gut Microbiome Due to Reduction of Bifidobacterium over the Past Century. *mSphere* **3**, e00041-00018 (2018).
158. Lau, A.S. & Liong, M.T. Lactic Acid Bacteria and Bifidobacteria-Inhibited Staphylococcus epidermidis. *Wounds : a compendium of clinical research and practice* **26**, 121-131 (2014).
159. A. Cheikhyoussef , N.P.a.H.Z. (2017).
160. Dermyshi, E. et al. The “Golden Age” of Probiotics: A Systematic Review and Meta-Analysis of Randomized and Observational Studies in Preterm Infants. *Neonatology* **112**, 9-23 (2017).
161. Alfaleh, K., Anabrees, J., Bassler, D. & Al-Kharfi, T. Probiotics for prevention of necrotizing enterocolitis in preterm infants. *The Cochrane database of systematic reviews*, Cd005496 (2011).
162. Fernandez-Carrocera, L.A. et al. Double-blind, randomised clinical assay to evaluate the efficacy of probiotics in preterm newborns weighing less than 1500 g in the prevention of necrotising enterocolitis. *Archives of disease in childhood. Fetal and neonatal edition* **98**, F5-9 (2013).
163. Wang, Q., Dong, J. & Zhu, Y. Probiotic supplement reduces risk of necrotizing enterocolitis and mortality in preterm very low-birth-weight infants: an updated meta-analysis of 20 randomized, controlled trials. *Journal of pediatric surgery* **47**, 241-248 (2012).
164. P, R.C.S.G.C.R.J.J.M.H.B.E.M.A.H.L.C. Incidence of necrotising enterocolitis before and after

introducing routine prophylactic Lactobacillus and Bifidobacterium probiotics. (2019).
165. Baranowski, J.R. & Claud, E.C. Necrotizing Enterocolitis and the Preterm Infant Microbiome. *Advances in experimental medicine and biology* (2019).
166. Guinane, C.M. et al. Genome sequence of Bifidobacterium breve DPC 6330, a strain isolated from the human intestine. *Journal of bacteriology* **193**, 6799-6800 (2011).
167. O'Mahony, C. et al. Commensal-induced regulatory T cells mediate protection against pathogen-stimulated NF-kappaB activation. *PLoS pathogens* **4**, e1000112-e1000112 (2008).
168. Abdulkadir, B. et al. Routine Use of Probiotics in Preterm Infants: Longitudinal Impact on the Microbiome and Metabolome. *Neonatology* **109**, 239-247 (2016).
169. Ruiz, L., Margolles, A. & Sánchez, B. Bile resistance mechanisms in Lactobacillus and Bifidobacterium. *Frontiers in microbiology* **4**, 396-396 (2013).
170. Ritter, P., Kohler, C. & von Ah, U. Evaluation of the passage of Lactobacillus gasseri K7 and bifidobacteria from the stomach to intestines using a single reactor model. *BMC microbiology* **9**, 87-87 (2009).
171. Moubareck, C., Gavini, F., Vaugien, L., Butel, M.J. & Doucet-Populaire, F. Antimicrobial susceptibility of bifidobacteria. *Journal of Antimicrobial Chemotherapy* **55**, 38-44 (2005).
172. Mukhopadhyay, S., Sengupta, S. & Puopolo, K.M. Challenges and opportunities for antibiotic stewardship among preterm infants. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, fetalneonatal-2018-315412 (2018).
173. Gotoh, A. et al. Sharing of human milk oligosaccharides degradants within bifidobacterial communities in faecal cultures supplemented with Bifidobacterium bifidum. *Scientific Reports* **8**, 13958 (2018).
174. Walker, A. Intestinal colonization and programming of the intestinal immune response. *Journal of clinical gastroenterology* **48 Suppl 1**, S8-S11 (2014).
175. Al-Nedawi, K. et al. Gut commensal microvesicles reproduce parent bacterial signals to host immune and enteric nervous systems. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **29**, 684-695 (2015).
176. Fukuda, S. et al. Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature* **469**, 543-547 (2011).
177. Martinez, F.A.C., Balciunas, E.M., Converti, A., Cotter, P.D. & de Souza Oliveira, R.P. Bacteriocin production by Bifidobacterium spp. A review. *Biotechnology Advances* **31**, 482-488 (2013).
178. Hays, S. et al. Probiotics and growth in preterm infants: A randomized controlled trial, PREMAPRO study. *Clinical nutrition (Edinburgh, Scotland)* **35**, 802-811 (2016).
179. Mathias, A., Pais, B., Favre, L., Benyacoub, J. & Corthesy, B. Role of secretory IgA in the mucosal sensing of commensal bacteria. *Gut microbes* **5**, 688-695 (2014).
180. trust, T.U.s. What is sepsis? (2019).
181. Liu, L. et al. Global, regional, and national causes of child mortality: an updated systematic analysis for 2010 with time trends since 2000. *Lancet (London, England)* **379**, 2151-2161 (2012).
182. Baqui, A.H. et al. Rates, timing and causes of neonatal deaths in rural India: implications for neonatal health programmes. *Bull World Health Organ* **84**, 706-713 (2006).

183. Simonsen, K.A., Anderson-Berry, A.L., Delair, S.F. & Davies, H.D. Early-onset neonatal sepsis. *Clinical microbiology reviews* **27**, 21-47 (2014).
184. Hornik, C.P. et al. Early and late onset sepsis in very-low-birth-weight infants from a large group of neonatal intensive care units. *Early human development* **88 Suppl 2**, S69-74 (2012).
185. Chan, G.J., Lee, A.C., Baqui, A.H., Tan, J. & Black, R.E. Risk of early-onset neonatal infection with maternal infection or colonization: a global systematic review and meta-analysis. *PLoS Med* **10**, e1001502 (2013).
186. Shao, Y. et al. Stunted microbiota and opportunistic pathogen colonization in caesarean-section birth. *Nature* (2019).
187. Fuchs, A., Bielicki, J., Mathur, S., Sharland, M. & Van Den Anker, J.N. Reviewing the WHO guidelines for antibiotic use for sepsis in neonates and children. *Paediatr Int Child Health* **38**, S3-S15 (2018).
188. Yoon, M.Y. & Yoon, S.S. Disruption of the Gut Ecosystem by Antibiotics. *Yonsei Med J* **59**, 4-12 (2018).
189. Gibson, M.K., Crofts, T.S. & Dantas, G. Antibiotics and the developing infant gut microbiota and resistome. *Current opinion in microbiology* **27**, 51-56 (2015).
190. AlFaleh, K. & Anabrees, J. Probiotics for prevention of necrotizing enterocolitis in preterm infants. *The Cochrane database of systematic reviews*, Cd005496 (2014).
191. Sawh, S.C., Deshpande, S., Jansen, S., Reynaert, C.J. & Jones, P.M. Prevention of necrotizing enterocolitis with probiotics: a systematic review and meta-analysis. *PeerJ* **4**, e2429 (2016).
192. Rees, C.M., Hall, N.J., Fleming, P. & Eaton, S. Probiotics for the prevention of surgical necrotising enterocolitis: systematic review and meta-analysis. *BMJ paediatrics open* **1**, e000066 (2017).
193. Lau, C.S. & Chamberlain, R.S. Probiotic administration can prevent necrotizing enterocolitis in preterm infants: A meta-analysis. *Journal of pediatric surgery* **50**, 1405-1412 (2015).
194. Esaiassen, E. et al. Effects of Probiotic Supplementation on the Gut Microbiota and Antibiotic Resistome Development in Preterm Infants. *Frontiers in pediatrics* **6**, 347-347 (2018).
195. Villamor-Martínez, E. et al. Probiotic Supplementation in Preterm Infants Does Not Affect the Risk of Bronchopulmonary Dysplasia: A Meta-Analysis of Randomized Controlled Trials. *Nutrients* **9** (2017).
196. 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* **31**, 1674-1676 (2015).
197. Rowe, W.P.M. et al. Streaming histogram sketching for rapid microbiome analytics. *Microbiome* **7**, 40 (2019).
198. Weisburg, W.G., Barns, S.M., Pelletier, D.A. & Lane, D.J. 16S ribosomal DNA amplification for phylogenetic study. *Journal of bacteriology* **173**, 697-703 (1991).
199. Gavan, T.L. & Town, M.A. A Microdilution Method for Antibiotic Susceptibility Testing: An Evaluation. *American Journal of Clinical Pathology* **53**, 880-885 (1970).
200. Rose, G. et al. Antibiotic resistance potential of the healthy preterm infant gut microbiome. *PeerJ* **5**, e2928 (2017).
201. Zou, Z.-H. et al. Prenatal and postnatal antibiotic exposure influences the gut microbiota of preterm infants in neonatal intensive care units. *Annals of Clinical Microbiology and Antimicrobials* **17**, 9 (2018).
202. Turroni, F. et al. Glycan Utilization and Cross-Feeding Activities by Bifidobacteria. *Trends in microbiology* **26**, 339-350 (2018).

203. Huerta-Cepas, J. et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research* **47**, D309-D314 (2018).
204. Gibson, M.K. et al. Developmental dynamics of the preterm infant gut microbiota and antibiotic resistome. *Nat Microbiol* **1**, 16024 (2016).
205. Aeschlimann, J.R. The role of multidrug efflux pumps in the antibiotic resistance of Pseudomonas aeruginosa and other gram-negative bacteria. Insights from the Society of Infectious Diseases Pharmacists. *Pharmacotherapy* **23**, 916-924 (2003).
206. Rowe, W.P.M. & Winn, M.D. Indexed variation graphs for efficient and accurate resistome profiling. *Bioinformatics* **34**, 3601-3608 (2018).
207. Hammond, D.S., Schooneveldt, J.M., Nimmo, G.R., Huygens, F. & Giffard, P.M. bla(SHV) Genes in Klebsiella pneumoniae: Different Allele Distributions Are Associated with Different Promoters within Individual Isolates. *Antimicrobial Agents and Chemotherapy* **49**, 256-263 (2005).
208. Lutsar, I. et al. Meropenem vs standard of care for treatment of late onset sepsis in children of less than 90 days of age: study protocol for a randomised controlled trial. *Trials* **12**, 215-215 (2011).
209. Santajit, S. & Indrawattana, N. Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens. *BioMed Research International* **2016**, 2475067 (2016).
210. Flannery, D.D. et al. Temporal Trends and Center Variation in Early Antibiotic Use Among Premature Infants. *JAMA Network Open* **1**, e180164-e180164 (2018).
211. Rose, G. et al. Antibiotic resistance potential of the healthy preterm infant gut microbiome. *PeerJ* **5**, e2928 (2017).
212. Gasparrini, A.J. et al. Persistent metagenomic signatures of early-life hospitalization and antibiotic treatment in the infant gut microbiota and resistome. *Nature Microbiology* (2019).
213. Fjalstad, J.W., Esaiassen, E., Juvet, L.K., van den Anker, J.N. & Klingenberg, C. Antibiotic therapy in neonates and impact on gut microbiota and antibiotic resistance development: a systematic review. *Journal of Antimicrobial Chemotherapy* **73**, 569-580 (2017).
214. Dong, Y. & Speer, C.P. Late-onset neonatal sepsis: recent developments. *Archives of Disease in Childhood - Fetal and Neonatal Edition* **100**, F257-F263 (2015).
215. Cohen-Wolkowiez, M. et al. Early and late onset sepsis in late preterm infants. *The Pediatric infectious disease journal* **28**, 1052-1056 (2009).
216. Tsai, C.H., Chen, Y.Y., Wang, K.G., Chen, C.Y. & Chen, C.P. Characteristics of early-onset neonatal sepsis caused by Escherichia coli. *Taiwanese journal of obstetrics & gynecology* **51**, 26-30 (2012).
217. Das, S., Jayaratne, R. & Barrett, K.E. The Role of Ion Transporters in the Pathophysiology of Infectious Diarrhea. *Cell Mol Gastroenterol Hepatol* **6**, 33-45 (2018).
218. Berkes, J., Viswanathan, V.K., Savkovic, S.D. & Hecht, G. Intestinal epithelial responses to enteric pathogens: effects on the tight junction barrier, ion transport, and inflammation. *Gut* **52**, 439-451 (2003).
219. Santajit, S. & Indrawattana, N. Mechanisms of Antimicrobial Resistance in ESKAPE Pathogens. *BioMed research international* **2016**, 2475067-2475067 (2016).
220. D'Accolti, M., Soffritti, I., Mazzacane, S. & Caselli, E. Fighting AMR in the Healthcare Environment: Microbiome-Based Sanitation Approaches and Monitoring Tools. *International journal of molecular sciences* **20** (2019).
221. Trejo, F.M., Minnaard, J., Perez, P.F. & De Antoni, G.L. Inhibition of Clostridium difficile growth and adhesion to enterocytes by Bifidobacterium supernatants. *Anaerobe* **12**, 186-193 (2006).

222. Cotten, C.M. et al. Prolonged duration of initial empirical antibiotic treatment is associated with increased rates of necrotizing enterocolitis and death for extremely low birth weight infants. *Pediatrics* **123**, 58-66 (2009).
223. Cho, I. et al. Antibiotics in early life alter the murine colonic microbiome and adiposity. *Nature* **488**, 621-626 (2012).
224. Wei, S., Weiss, Z.R. & Williams, Z. Rapid Multiplex Small DNA Sequencing on the MinION Nanopore Sequencing Platform. *G3: Genes|Genomes|Genetics* **8**, 1649-1657 (2018).
225. Melville, J.M. & Moss, T.J.M. The immune consequences of preterm birth. *Frontiers in Neuroscience* **7** (2013).
226. Stewart, C.J. et al. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. *Microbiome* **5** (2017).
227. Deshpande, G., Jape, G., Rao, S. & Patole, S. Benefits of probiotics in preterm neonates in low-income and medium-income countries: a systematic review of randomised controlled trials. *BMJ open* **7**, e017638 (2017).
228. Magi, A., Semeraro, R., Mingrino, A., Giusti, B. & D'Aurizio, R. Nanopore sequencing data analysis: state of the art, applications and challenges. *Briefings in bioinformatics* (2017).
229. Leggett, R.M., Heavens, D., Caccamo, M., Clark, M.D. & Davey, R.P. NanoOK: multi-reference alignment analysis of nanopore sequencing data, quality and error profiles. *Bioinformatics* **32**, 142-144 (2016).
230. Jia, B. et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res* **45**, D566-d573 (2017).
231. Luber, P., Bartelt, E., Genschow, E., Wagner, J. & Hahn, H. Comparison of Broth Microdilution, E Test, and Agar Dilution Methods for Antibiotic Susceptibility Testing of Campylobacter jejuni and Campylobacter coli. *Journal of Clinical Microbiology* **41**, 1062-1068 (2003).
232. Chen, H.N., Lee, M.L., Yu, W.K., Lin, Y.W. & Tsao, L.Y. Late-onset Enterobacter cloacae sepsis in very-low-birth-weight neonates: experience in a medical center. *Pediatrics and neonatology* **50**, 3-7 (2009).
233. Swick, M.C., Morgan-Linnell, S.K., Carlson, K.M. & Zechiedrich, L. Expression of Multidrug Efflux Pump Genes acrAB-tolC, mdfA, and norE in Escherichia coli Clinical Isolates as a Function of Fluoroquinolone and Multidrug Resistance. *Antimicrobial Agents and Chemotherapy* **55**, 921-924 (2011).
234. Werner, N.L., Hecker, M.T., Sethi, A.K. & Donskey, C.J. Unnecessary use of fluoroquinolone antibiotics in hospitalized patients. *BMC Infectious Diseases* **11**, 187 (2011).
235. Sim, K. et al. Dysbiosis anticipating necrotizing enterocolitis in very premature infants. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* **60**, 389-397 (2015).
236. Liakopoulos, A., Mevius, D. & Ceccarelli, D. A Review of SHV Extended-Spectrum beta-Lactamases: Neglected Yet Ubiquitous. *Front Microbiol* **7**, 1374 (2016).
237. Brown, B.L., Watson, M., Minot, S.S., Rivera, M.C. & Franklin, R.B. MinION™ nanopore sequencing of environmental metagenomes: a synthetic approach. *GigaScience* **6**, 1-10 (2017).
238. Hodzic, Z., Bolock, A.M. & Good, M. The Role of Mucosal Immunity in the Pathogenesis of Necrotizing Enterocolitis. *Frontiers in pediatrics* **5**, 40 (2017).

239. Poirel, L., Jayol, A. & Nordmann, P. Polymyxins: Antibacterial Activity, Susceptibility Testing, and Resistance Mechanisms Encoded by Plasmids or Chromosomes. *Clinical microbiology reviews* **30**, 557-596 (2017).
240. Jayol, A. et al. Resistance to Colistin Associated with a Single Amino Acid Change in Protein PmrB among Klebsiella pneumoniae Isolates of Worldwide Origin. *Antimicrobial Agents and Chemotherapy* **58**, 4762-4766 (2014).
241. Illumina Specification for MiSEq sequencing. *Illumina website* (2019).
242. Biosciences, P. Template preparation and sequencing guide. *Pacific Bioscience* (2018).
243. Enroth, H. et al. Evaluation of QuickFISH and maldi Sepsityper for identification of bacteria in bloodstream infection. *Infectious diseases (London, England)* **51**, 249-258 (2019).
244. Bailey, D. et al. Use of MALDI-TOF for Diagnosis of Microbial Infections. *Clinical Chemistry* **59**, 1435-1441 (2013).
245. Tyler, A.D. et al. Evaluation of Oxford Nanopore's MinION Sequencing Device for Microbial Whole Genome Sequencing Applications. *Scientific Reports* **8**, 10931 (2018).
246. Lemon, J.K., Khil, P.P., Frank, K.M. & Dekker, J.P. Rapid Nanopore Sequencing of Plasmids and Resistance Gene Detection in Clinical Isolates. *Journal of clinical microbiology* **55**, 3530-3543 (2017).
247. Stewart, C.J. et al. Temporal bacterial and metabolic development of the preterm gut reveals specific signatures in health and disease. *Microbiome* **4**, 67 (2016).
248. Stewart, C.J. et al. Longitudinal development of the gut microbiome and metabolome in preterm neonates with late onset sepsis and healthy controls. *Microbiome* **5**, 75 (2017).
249. Kalliomaki, M. et al. Distinct patterns of neonatal gut microflora in infants in whom atopy was and was not developing. *J Allergy Clin Immunol* **107**, 129-134 (2001).
250. Kalliomaki, M., Collado, M.C., Salminen, S. & Isolauri, E. Early differences in fecal microbiota composition in children may predict overweight. *Am J Clin Nutr* **87**, 534-538 (2008).
251. Martín, R. & Langella, P. Emerging Health Concepts in the Probiotics Field: Streamlining the Definitions. *Frontiers in microbiology* **10**, 1047-1047 (2019).
252. Mancini, N. et al. The Era of Molecular and Other Non-Culture-Based Methods in Diagnosis of Sepsis. *Clinical microbiology reviews* **23**, 235-251 (2010).
253. Technologies, N. PromethION. (2019).