

Mechanistic insights into the interaction between gut bacteria polysaccharides and lectins of the host innate immune system

Dimitra Lamprinaki

A thesis submitted for the degree of Doctor of Philosophy (PhD)

To the University of East Anglia

Quadram Institute Bioscience

Gut Microbes and Health

Norwich Research Park

Norwich

NR4 7UQ

January 2021



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

Abstract

Colorectal cancer (CRC) is one of the most common cancer types worldwide, with cases showing the dominance of *Bacteroides fragilis* and *Fusobacterium nucleatum* species. While *B. fragilis* is involved in tumorigenesis through the toxin that it secretes, *F. nucleatum* is implicated in CRC progression through the recruitment of tumour infiltrating immune cells. However, the mechanisms underpinning the interaction between *F. nucleatum* ssp. and immune cells remain to be determined. Mammalian lectins expressed by immune or epithelium cells are glycan-binding proteins playing an important role in host cell physiology. The Siglec superfamily and Galectin-3 have been shown to play a role in tumour progression, and Dectin-2 has been shown to activate immune response. Here, we tested the hypothesis that the interaction of cell surface glycoconjugates present on *B. fragilis* and *F. nucleatum* ssp. with host lectins may mediate immune response and tumour progression.

Our results showed that, among *B. fragilis* and the three *F. nucleatum* subspecies (ATCC 10953, ATCC 25586, ATCC 51191) tested, *F. nucleatum* ssp. bound to Siglecs and specifically to Siglec-7. Further, we showed that the lipopolysaccharide (LPS) derived from *F. nucleatum* ssp. bound to Siglec-7 through novel sugar epitopes contained in the LPS structures and we also uncovered a novel LPS structure for *F. nucleatum* ATCC 51191 strain. Moreover, we purified and characterised the outer membrane vesicles (OMVs) from *F. nucleatum* ssp. and showed that they interact with Siglec-7, as we demonstrated with the whole bacteria or LPS. We then showed that *F. nucleatum* ssp. and their derived LPS and OMVs induced a pro-inflammatory phenotype in dendritic cells and a tumour-promoting phenotype in macrophages. Depletion of Siglec-7 in human myeloid cells led to a change in bacteria internalisation and immune response induced by *F. nucleatum* ssp.

Finally, we showed that *F. nucleatum* ssp. abundance was elevated in the on-tumour as compared to the off-tumour site across CRC clinical samples tested. In addition, our preliminary data reported the presence of anti-*F. nucleatum* ssp. IgG antibodies and high levels of the Siglec-7 protein in the serum of CRC patients, which further support *F. nucleatum*-Siglec-7 interaction as a novel underlying mechanism implicated in CRC.

Access Condition and Agreement

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

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

Acknowledgements

I would like to thank DTP-BBSRC for funding me and for the additional financial support due to the COVID-19 work disruption.

During this adventurous PhD journey, I am grateful to have interacted with such great international people and to have made valuable collaborations.

First and foremost, I was lucky to have had Prof. Nathalie Juge as my primary supervisor, who she supported me throughout my PhD journey, gave me opportunities to attend to trainings and conferences and provided to me invaluable career advices. Furthermore, it was an honour to have had Prof. Paul Crocker in my supervisory team and I would like to thank him and his group in Dundee for their valuable help with Siglecs. I am extremely grateful for all the help that Prof. Kristian Bowles provided to my project, specifically, in coordinating the blood sample collection and the general human tissue work, as well as the clinicians for providing the blood samples in a regular basis. Furthermore, a big thanks to the surgeon Dr James Hernor and the biorepository team, Dr Roxanne Brunton-Sim and Dr Rachael Stanley who provided to me the human resected tissues and Dr Antonietta Hayhoe for her valuable help in ethics applications. Of course, I would not be doing this PhD without Dr Norihito Kawasaki, who was my initial primary supervisor, he introduced me to the glycobiology field, and together with Dr Alexandra Wittmann, former post-doc in Norihito's group trained me to a range of techniques.

I truly appreciate the work of our collaborators, Prof. Cristina De Castro, Prof. Alba Silipo and their group members, especially Ms Pilar Garcia del Vello Moreno, for the LPS structural characterisation, which gave valuable insights into my project, Prof. Matthew Macauley for kindly providing to us the U937 cell lines, and Prof. Kurt Drickamer and Ms Sabine Jegouso for training me on human Dectin-2 expression and purification. Moreover, I thank Dr Kathryn Gotts and Dr Catherine Booth for conducting the TEM analysis, Dr Falk Hildebrand and Ms Angela Del-Castillo-Izquierdo for the 16S sequencing data, Prof. Alastair Watson and Dr Devina Divekar for providing to me the APC mouse, Dr John Walshaw and Dr George Savva for the help with the bioinformatics and statistics, respectively, Dr Nikolaus Wellner for the technical support with microscopy and flow cytometry, and Dr Prabhjoat Chana for the technical support with imaging

flow cytometry.

I would also like to thank Prof. Arjan Narbad and Prof. Joy Burchell for agreeing to be my viva assessors.

Furthermore, a big thanks to all the members of Juge's group for the fun times and especially to the former post-doc Dr Devon Kavanaugh for training me on Galectin-3 expression and purification and Dr Dimitrios Latousakis for being always there to help.

Last but not least, I am deeply grateful to have a supportive family including my mother Mrs Anastasia Karampasi and my father Mr Nikolaos Lamprinakis, my dearest friends Dr Alexandra Collet, Ms Chrysanthi Vazitari and Mrs Mariza Giasafaki who were always remotely beside me, believing in me and encouraging me all of these years.

Table of Contents

Abstract	1
Acknowledgements	2
Abbreviations	7
List of Figures	12
List of Tables	15
CHAPTER 1 - Introduction	16
1.1 Gut anatomy and immunity	17
1.1.1 <i>Mucosa</i>	17
1.1.2 <i>Immune niches along the gut</i>	21
1.1.3 <i>Tumour-associated innate immune cells</i>	24
1.2 Intestinal lectins in immunity and cancer	27
1.2.1 <i>C-type lectins (CTLs)</i>	27
1.2.2 <i>Galectins</i>	29
1.2.3 <i>Sialic acid-binding immunoglobulin-like lectins (Siglecs)</i>	30
1.3 Gut microbiota	35
1.3.1 <i>Structure and function of the human gut microbiota</i>	35
1.3.2 <i>Cell surface glycosylation of bacteria</i>	37
1.4 Colorectal cancer (CRC) and associated microbiota	40
1.4.1 <i>Bacteroides fragilis</i>	42
1.4.2 <i>Fusobacterium nucleatum</i>	43
Hypothesis and Aims	47
CHAPTER 2 - Materials and Methods	48
2.1 Microbiology	49
2.1.1 <i>Strains and Materials</i>	49
2.1.2 <i>Bacterial and Yeast growth and preparation of PFA-fixed bacteria</i>	49
2.1.3 <i>Microbial density and size quantification</i>	50
2.1.4 <i>Outer membrane vesicle (OMV) purification from F. nucleatum ssp.</i>	50
2.1.5 <i>Nanoparticle Tracking Analysis (NTA)</i>	51
2.1.6 <i>Transmission electron microscopy (TEM)</i>	51
2.1.7 <i>LPS extraction from F. nucleatum ssp.</i>	52
2.1.8 <i>Semi-quantitative analysis of LPS in OMVs</i>	52
2.2 Protein expression.....	54
2.2.1 <i>Materials</i>	54
2.2.2 <i>Co-transformation of E. coli BL21DE(3) with Dectin-2-CRD-biotin-tag and BirA plasmids</i>	54
2.2.3 <i>Expression and purification of recombinant Dectin-2-CRD and tetramer formation</i>	55
2.2.4 <i>Expression and purification of recombinant Galectin-3</i>	56
2.2.5 <i>Expression and purification of recombinant Siglec-Fc proteins</i>	57
2.2.6 <i>SDS-PAGE</i>	58
2.2.7 <i>Protein quantification</i>	58

2.3 Binding assays	59
2.3.1 Materials	59
2.3.2 Red Blood Cells binding assays	59
2.3.3 Microbial flow cytometry binding assays	60
2.3.4 ELISA-based binding assays	61
2.3.5 Biolayer Interferometry (BLI) binding assays	61
2.3.6 Binding of <i>F. nucleatum</i> to mammalian cells	61
2.3.7 STD NMR analysis	62
2.4 Immunological assays	63
2.4.1. Materials	63
2.4.2 Generation of moDCs and moM ϕ s from human blood	63
2.4.3 Siglec-7 RNA silencing in immune cells	64
2.4.4 Stimulation of human primary myeloid cells with <i>F. nucleatum</i> ssp	66
2.4.5 Stimulation of U937-PMA cells with <i>F. nucleatum</i> ssp	66
2.4.6 Cell surface marker analysis by flow cytometry	66
2.4.7 Internalisation assays by Imaging Flow Cytometry	67
2.4.8 Microscopy	67
2.4.9 Cytokine analysis	68
2.5 Mammalian tissue processing	69
2.5.1 Materials and Ethics	69
2.5.2 Colonic lamina propria leukocyte isolation	69
2.5.3 Colonic lamina propria leukocyte cell staining by flow cytometry	70
2.5.4 DNA and RNA extraction from human colonic tissues and cDNA synthesis	70
2.5.5 qPCR analysis	71
2.5.6 16S rRNA Sequencing	73
2.5.7 Serological assays	74
2.6 Bioinformatic analysis	75
2.7 Statistical analyses	75
CHAPTER 3 - Screening of CRC-associated bacteria binding to human recombinant lectins	76
3.1 Introduction	77
3.2 Results	78
3.2.1 Interaction of <i>B. fragilis</i> and <i>F. nucleatum</i> ssp. with human Dectin-2, Galectin-3 and Siglec-7	78
3.2.2 <i>F. nucleatum</i> ssp. bound specifically to Siglec-7-Fc	88
3.3 Discussion	98
CHAPTER 4 - Investigating the role of <i>F. nucleatum</i> ssp. cell surface polysaccharides in the interaction with Siglec-7	100
4.1 Introduction	101
4.2 Results	104
4.2.1 <i>F. nucleatum</i> ssp. binds to Siglec-7 V-set in a sialic acid-independent manner	104
4.2.2 <i>F. nucleatum</i> ssp. LPS bound to Siglec-7	107
4.3 Discussion	117

CHAPTER 5 - Deciphering the effects of <i>F. nucleatum</i> ssp. on myeloid cell phenotype.....	120
5.1 Introduction	121
5.2 Results.....	122
5.2.1 <i>Effect of F. nucleatum ssp. on myeloid cells.....</i>	122
5.2.2 <i>Effect of F. nucleatum OMVs or LPS on myeloid cells</i>	142
5.2.3 <i>Siglec-7 expression in immune cells</i>	146
5.3 Discussion	149
CHAPTER 6 - Role of Siglec-7 in <i>F. nucleatum</i> modulation of host immune response.....	151
6.1 Introduction	152
6.2 Results.....	154
6.2.1 <i>Effect of Siglec-7 on F. nucleatum-induced U937 cell response.....</i>	154
6.2.2 <i>Effect of Siglec-7 in F. nucleatum-induced myeloid cell response</i>	159
6.3 Discussion	168
CHAPTER 7 - Siglec-7 expression and <i>Fusobacterium</i> spp. abundance in human CRC tissues	170
7.1 Introduction	171
7.2 Results.....	175
7.2.1 <i>F. nucleatum spp. relative abundance is increased in on-tumour site of CRC tissues</i>	175
7.2.2 <i>Colorectal cancer cases showed high anti-F. nucleatum ssp. IgG levels</i>	178
7.2.3 <i>Expression of Siglec-7 in tumour tissues shows heterogeneity.....</i>	181
7.2.4 <i>Soluble Siglec-7 levels are increased in colorectal cancer cases</i>	186
7.3 Discussion	187
CHAPTER 8 - Key findings and General Discussion	189
8.1 Key findings.....	190
8.2 General Discussion.....	192
References	200

Abbreviations

AMP	Antimicrobial peptide
APC	Antigen-presenting cell
APC	Allophycocyanin
ATCC	American type culture collection
BCCP	Biotin carboxyl carrier protein
BHI	Brain heart infusion
BLI	Biolayer interferometry
BSA	Bovine serum albumin
CA19-9	Carbohydrate antigen 19-9
CCL	Chemokine (C-C motif) ligand
CD	Cluster of differentiation
CEA	Carcinoembryonic antigen
CHO	Chinese hamster ovary
CMAH	CMP- <i>N</i> -acetylneuraminic acid hydroxylase
CPS	Capsular polysaccharide
CRC	Colorectal cancer
CRD	Carbohydrate recognition domain
CTL	C-type lectin
CTLA-4	Cytotoxic T-lymphocyte-associated protein 4 (CTLA-4)
CTP	Cytidine triphosphate
CTV	Cell trace violet
CX3CR	CX3C chemokine receptor
DAPI	4',6-diamidino-2-phenylindole
DC	Dendritic cell
Dectin	Dendritic cell-associated C-type lectin
EBF	Enerotoxigenic <i>Bacteroides fragilis</i>
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(b-aminoethyl ether)- <i>N,N,N',N'</i> -tetraacetic acid
EPS	Extracellular polysaccharides
esgp	Excitation sculpting with gradient pulses
FadA	<i>Fusobacterium</i> adhesin A
Fap2	<i>Fusobacterium</i> autotransporter protein 2

FBS	Fetal bovine serum
FISH	Fluorescence in situ hybridisation
FITC	Fluorescein isothiocyanate
FOXP3	Forkhead box P3
FPLC	Fast protein liquid chromatography
FSC	Forward scatter
GalNAc	N-acetylgalactosamine
GALT	Gut-associated lymphoid tissues
GBS	Group B <i>Streptococcus</i>
GC-MS	Gas chromatography mass spectrometry
GD3	Disialoganglioside with three glycosyl groups
GI	Gastrointestinal
GlcNAc	N-acetylglucosamine
GM-CSF	Granulocyte macrophage colony stimulating factor
GMEM	Glasgow modified essential medium
Gram-	Gram negative
Gram+	Gram positive
GS	Glutamine synthetase
HBSS	Hanks' balanced salt solution
hDC-2 ₍₄₎ - Alexa488	Tetramer human Dectin-2-Alexa488
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HI	Heat inactivated
HIV	Human immunodeficiency virus
HRP	Horseradish peroxidase
HSV	Herpes simplex virus
IBD	Inflammatory bowel disease
IEC	Intestinal epithelial cell
IEL	Intraepithelial lymphocytes
IFC	Imaging flow cytometry
IL	Interleukin
IMAC	Immobilized metal affinity chromatography
ITAM	Immunoreceptor tyrosine-based activation motif
ITIM	Immunoreceptor tyrosine-based inhibitory motif
KDO	3-deoxy-2-keto-D-manno-octulosonic acid

L-6dAltNAc	2-acetamido-2,6-dideoxy-L-altrose
LB	Luria-Bertani
Leg	Legionaminic acid
Lgr5	Leucine-rich repeat-containing G protein-coupled receptor 5
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
M	Microfold
MAMP	Microbial associated molecular pattern
ManNAc	N-acetylmannosamine
M-CSF	Macrophage colony stimulating factor
MDSC	Myeloid-derived suppressor cell
MFI	Mean fluorescence intensity
MHCII	Major histocompatibility complex II
moDC	Monocyte-derived dendritic cell
MOI	Multiplicity of infection
moM ϕ	Monocyte-derived macrophage
MSX	Methionine sulfoximine
MWCO	Molecular weight cut-off
MyD88	Myeloid differentiation primary response 88
NAB	Nonulosonic acid biosynthesis
NCTC	National collection of type cultures
Neu5Ac	Acetylneuraminic acid/ sialic acid
Neu5Gc	N-glycolylneuraminic acid
NeuB	Neu5Ac synthase
NeuC	UDP-GlcNAc 2-epimerase
NeuD	Neu5Ac O-acetyltransferase
NeuO	Polysialic acid O-acetyltransferase
NeuS	Polysialyltransferase
NF- κ B	Nuclear factor- κ B
NK	Natural killer
NLR	Nucleotide-binding oligomerization domain-like receptor
NO	Nitric oxide
NOD	Nucleotide-binding oligomerization domain

Non5Am	5-acetimidoylamino-3,5,9-trideoxygluco-non-2-ulosonic acid
Non5Am7Ac	5,7-diamino-3,5,7,9-tetra-deoxynon-2-ulosonic acid
NTA	Nanoparticle tracking analysis
NuIO	Nonulosonic acid
o/n	Overnight
OD _{600nm}	Optical density in 600 nm
OM	Outer membrane
OMV	Outer membrane vesicle
OPS	O-antigen
OTU	Operational taxonomic unit
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffered saline
PD-1	Programmed cell death 1
PD-L1	Programmed death ligand 1
PE	Phycoerythrin
PEP	Phosphoenolpyruvate
PFA	Paraformaldehyde
PG	Peptidoglycans
PGK1	Phosphoglycerate Kinase 1
PI	Propidium iodide
pIgR	Polymeric immunoglobulin receptor
PMA	Phorbol 12-myristate 13-acetate
PRR	Pattern recognition receptor
PS	Polysaccharide
Pse	Pseudaminic acid
qPCR	Quantitative PCR
R	Arginine
RBC	Red blood cell
R-MFI	Relative MFI
ROS	Reactive oxygen species
RT	Room temperature
SA	Streptavidin
SCFA	Short chain fatty acid

SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SHP	SRC homology domain 2-containing tyrosine phosphatase
sIgA	Secretory immunoglobulin A
Siglec	Sialic acid-binding immunoglobulin-like lectin
SILT	Small isolated lymphoid tissue
siRNA	Small interfering RNA
SLE	Systemic lupus erythematosus
SOC	Super optimal broth with catabolite suppression
SSC	Side scatter
STAT-3	Signal transducer and activator of transcription 3
Syk	Spleen tyrosine kinase
TAM	Tumour-associated macrophage
TCR	T cell receptor
TEM	Transmission electron microscopy
Tfh	T follicular helper
TGF- β	Transforming growth factor beta
Th	T helper
TIGIT	T cell immunoreceptor with Ig and ITIM domains
TLR	Toll-like receptor
TMB	3,3',5,5' Tetramethylbenzidine
TNF α	Tumour necrosis factor α
TNM	Tumour invasion-node-metastasis
Tregs	T regulatory cells
TRIF	TIR domain-containing adaptor-inducing interferon- β
UDP	Uridine diphosphate
VZV	Varicella-zoster virus
WT	Wild type

List of Figures

Figure 1 Anatomy of small (left) and large (right) intestines.....	20
Figure 2 Immune niches along mouse intestine.....	23
Figure 3 Oncogenic evolution and immunity.....	24
Figure 4 C-type lectins in myeloid cells.	28
Figure 5 Human Galectins.	29
Figure 6 Human Siglecs and their intracellular motifs.	31
Figure 7 Siglec-7 structure and its mediated signalling.....	32
Figure 8 Bacteria distribution along and across the lower GI tract.....	35
Figure 9 Glycoconjugates in bacterial cell membranes.....	38
Figure 10 The microbial hypotheses in CRC progression.	41
Figure 11 LPS structure of <i>F. nucleatum</i> strains.....	46
Figure 12 Dectin-2-CRD and streptavidin tetrameric complex.....	78
Figure 13 Analysis of recombinant Dectin-2-CRD-biotin-tag colonies.....	79
Figure 14 SDS-PAGE analysis.	80
Figure 15 Analysis of recombinant Dectin-2-Streptavidin complex.....	81
Figure 16 Purification of hDC-2(4)-Alexa488 tetramer by FPLC.	82
Figure 17 Histogram of hDC-2(4)-Alexa488 binding to <i>S. cerevisiae</i>	83
Figure 18 SDS-PAGE analysis of his-tag Galectin-3 purification.	84
Figure 19 Galectin-3 binding to <i>K. pneumoniae</i> O1.....	85
Figure 20 <i>F. nucleatum</i> ssp. and <i>B. fragilis</i> NCTC 9343 binding to human lectins by flow cytometry.	87
Figure 21 SDS-PAGE analysis of recombinant Siglecs produced in CHO cells.....	89
Figure 22 Histogram of Siglec-Fc binding to RBCs.	90
Figure 23 Human IgG-Fc domain binding to <i>F. nucleatum</i> ssp.	92
Figure 24 Histogram of Siglec-Fc binding to <i>F. nucleatum</i> ssp.	93
Figure 25 Histogram of Siglec-Fc binding to <i>R. gnavus</i>	94
Figure 26 Siglec-7 expression on U937 cells and <i>F. nucleatum</i> ssp. association assays.	96
Figure 27 Siglec-7 expression on CHO cells and <i>F. nucleatum</i> ssp. association assays.	97
Figure 28 Nonulosonic acid family.....	102

Figure 29 De novo sialylation pathway in bacteria.....	102
Figure 30 Histograms of <i>F. nucleatum</i> ssp. and Siglec-7-Fc inhibition assays with GD3.	104
Figure 31 Histograms of SNA binding to <i>F. nucleatum</i> ATCC 51191 or ATCC 10953. ...	105
Figure 32 Table showing the presence of the Neu5Ac biosynthetic cluster enzymes.	106
Figure 33 Histograms of Siglec-Fc binding to sialidase treated <i>F. nucleatum</i> ssp.	107
Figure 34 O-antigen characterisation from <i>F. nucleatum</i> ssp.-derived LPS.	108
Figure 35 ELISA-based binding assays of <i>F. nucleatum</i> -derived LPSs and Siglec-7-Fc.	109
Figure 36 Histograms of <i>F. nucleatum</i> ssp. and Siglec-7-Fc inhibition assays with LPS.	110
Figure 37 Chromatograms of <i>F. nucleatum</i> ssp. and Siglec-7-Fc binding by STD-NMR.	111
Figure 38 <i>F. nucleatum</i> -derived OMV particle analysis.	113
Figure 39 Chromatograms of <i>F. nucleatum</i> -derived OMVs from GC-MS analysis.....	114
Figure 40 <i>F. nucleatum</i> -derived OMVs binding to Siglec-7-Fc.	116
Figure 41 <i>F. nucleatum</i> cell quantification.....	124
Figure 42 <i>F. nucleatum</i> cell quantification by imaging flow cytometry.....	125
Figure 43 <i>F. nucleatum</i> ssp. show differences in their morphology.	127
Figure 44 Monocyte isolation and differentiation.....	129
Figure 45 <i>F. nucleatum</i> ssp. associates with moDCs and moMφs.....	131
Figure 46 <i>F. nucleatum</i> ssp. internalisation into myeloid cells by IFC.	132
Figure 47 Cytokine production of moDCs and moMφs stimulated with <i>F. nucleatum</i> ssp.	135
Figure 48 Cytokine production of U937-PMA stimulated with <i>F. nucleatum</i> ssp.	136
Figure 49 Histograms of cell surface marker expression in moDCs and moMφs.....	137
Figure 50 Flow cytometry analysis of the proportion of dead moDCs or moMφs at unstimulated state.	139
Figure 51 Flow cytometry analysis of the proportion of dead moDCs after stimulation with <i>F. nucleatum</i> ssp.	140
Figure 52 Flow cytometry analysis of the proportion of dead moMφs after stimulation with <i>F. nucleatum</i> ssp.	141
Figure 53 Cytokine production of moDCs and moMφs stimulated with <i>F. nucleatum</i> ssp.- derived OMVs or LPSs.....	143
Figure 54 Histograms of cell surface marker expression in moDCs and moMφs following <i>F. nucleatum</i> stimulation.	145

Figure 55 Histogram of Siglec-7 expression on myeloid cells.....	147
Figure 56 Siglec-E expression in mouse colonic lamina propria leukocytes.	147
Figure 57 Histograms of Siglec-7 expression in moDCs and moMφs following <i>F. nucleatum</i> stimulation.	148
Figure 58 Schematic representation of Siglec-7 structure on the mammalian cell membrane.	153
Figure 59 Histograms of protein expression on U937 cell lines.....	155
Figure 60 Cytokine production in PMA-U937 cells stimulated with <i>F. nucleatum</i> ssp.....	157
Figure 61 Internalisation of <i>F. nucleatum</i> ssp. into U937 cells.....	158
Figure 62 moDCs or moMφs cytokine production in presence of anti-Siglec-7 Ab and stimulation with <i>F. nucleatum</i> ssp.	160
Figure 63 Histograms of Siglec-7 expression and cell death of moDCs and moMφs.....	162
Figure 64 Cytokine production in Siglec-7 silenced moDCs stimulated with <i>F. nucleatum</i> ssp.	164
Figure 65 Histograms of marker expression on Siglec-7 silenced moDCs.....	166
Figure 66 Internalisation of <i>F. nucleatum</i> ssp. into moDCs.....	167
Figure 67 Colorectal tumour invasion stages.....	171
Figure 68 Siglec-7 gene expression in healthy human tissues.....	173
Figure 69 Siglec-7 gene expression levels in colorectal adenocarcinomas.....	173
Figure 70 <i>Fusobacterium</i> spp. relative abundance in the on- and off-tumour sites.....	177
Figure 71 <i>Fusobacterium</i> spp. relative abundance using 16S RNA sequence.	177
Figure 72 anti- <i>F. nucleatum</i> -IgG levels in serum of healthy or colorectal cancer human samples.....	180
Figure 73 Gating strategy of the multicolour stained gut lamina propria leukocytes.	182
Figure 74 RNA electrophoresis and RNA integrity numbers by TapeStation.	184
Figure 75 Relative Siglec-7 mRNA expression in the on- and off-tumour tissue sites of colorectal cancer cases.....	185
Figure 76 Soluble serum Siglec-7 levels in colorectal cancer or healthy samples.	186

List of Tables

Table 1 Galectin-3, dectin-2, and siglec-7: Glycan or pathogen recognition, and cell expression	34
Table 2 Mannose-sepharose column buffer preparation	55
Table 3 IMAC buffer compositions.....	57
Table 4 Pre-designed probes used for Siglec-7 siRNA and silencing conditions	65
Table 5 Primers used in qPCR.....	72
Table 6 qPCR cycling conditions	72
Table 7 List of the colorectal cancer cases and <i>Fusobacterium</i> spp. abundance.....	176
Table 8 List of colorectal cancer cases with the <i>F. nucleatum</i> abundance and Siglec-7 expression in the tested gut lamina propria cell subsets.....	183

CHAPTER 1

-

Introduction

1.1 Gut anatomy and immunity

The human gastrointestinal (GI) tract primary is known for its role in food digestion and nutrient uptake, although it also contains the majority of immune and microbial cells which populate the human body. The gut covers the small and large intestine, two sites with significant anatomical and physiological differences.

The small intestine of approximately 7 m long, consists of the duodenum, jejunum and ileum and its main function is to digest food and absorb dietary nutrients. The large intestine of approximately 1.5 m long reabsorbs the water, and consists of the caecum, the right side-proximal colon (ascending, transverse colon) and the left side-distal colon (descending, sigmoid colon), and rectum. The colon is the most densely populated microbial ecosystem (gut microbiota) in the body. The gut microbiota varies in composition and abundance along the GI tract and in the cross-sectional axes of the gut¹, as discussed in more detail in section 1.3.

The architecture of the intestine consists of a mucosa, submucosa, and a muscle layer (Fig. 1). The mucosa is the most complex compartment and contains the lumen, mucus, an epithelial cell layer, and the lamina propria, as described below (section 1.1.1). The submucosa, underneath the mucosa, contains the gut-associated lymphoid tissues (GALT) as further described in section 1.1.2.

1.1.1 Mucosa

The lumen of the mucosa harbours a large microbial community and contains microbial metabolites as well as host antibodies and antimicrobial peptides (AMPs). The mucosa is covered by a protective mucus layer produced by specialised epithelial cells called goblet cells. In the small intestine, mucus occurs as a loose single-layered, while in the colon mucus is found as a double-layer² composed of an inner mucus layer attached to the epithelium and an outer loose mucus layer facing the lumen^{2,3} (Fig. 1). Both small intestine and large intestine mucus layers are based on the large glycoprotein, MUC2. The loose mucus layer is a habitat for the gut microbiota, while the inner colonic mucus layer is virtually free of microbes under homeostatic conditions⁴.

The epithelium appears as a physical barrier of one cellular layer. The absorptive surface area of the small intestine is significantly increased by numerous finger-like protrusions that point towards the lumen, the so-called villi, and invaginations

extending to the muscularis mucosae known as the crypts of Lieberkühn⁵. The mucosa of the large intestine lacks villi; crypts are bigger compared to the small intestine (Fig. 1). The intestinal epithelium is made of a wide range of cell types, each with a distinct function, as described below.

At the base of the crypt are found the leucine-rich repeat-containing G protein-coupled receptor 5 (Lgr5) stem cells, which are continuous dividing cells and progenitors of the absorptive (enterocytes), secretory (Paneth cells, goblet cells, tuft cells and enteroendocrine cells)⁶, and microfold (M) cells⁷. The differentiated cells migrate towards the top of the crypt where they die after few days of their differentiation (Fig. 1).

Enterocytes are the main epithelial cell type and they function as nutrient absorptive cells. Their apical site forms microvilli in the small intestine which increase the absorption area, but they also express charged glycoproteins which act as a barrier (known as a the glycocalyx)⁸. Enterocytes can also induce inflammatory responses as they express a range of pattern recognition receptors (PRRs) which interact with exogenous antigens⁹.

Paneth cells are the only long-lived epithelial cell types¹⁰, they are mainly found in small intestine and are located at the crypt base and in proximity to stem cells. Their main function is the secretion of AMPs such as α -defensins and lysozyme upon sense of an exogenous stimulus¹¹.

Goblet cells are mainly known as mucin-secreting cells, leading to the production of mucus which serves as a protective layer for the gut epithelium¹². Recently, goblet cells have been involved in host immunomodulation through cytokine secretion and by transferring luminal antigens to the lamina propria antigen-presenting cells (APCs)¹³. Both small and large intestines contain goblet cells, but their abundance in the large intestine is 10% higher compared to small intestine¹⁴.

Enteroendocrine cells are found in approximately 1% of the total epithelial cells and their main function is to secrete gut hormones in response to stimuli¹⁵. The produced hormones are involved in food digestion and absorption, regulation of appetite and insulin secretion¹⁶.

The M cells that cover the lymphoid Peyer's patches (discussed in section 1.1.2) are present in the follicle-associated epithelium which overlies GALT. Their main role is to sample antigens, such as pathogenic or commensal microbes, from the lumen to the lamina propria by phagocytosis or transcytosis^{17,186,17}. Upon antigen

sampling, immune cells located in the follicle below M cells, recognise the foreign antigens and mount an immune response. This function of M cells provides the first step of immune activation and therefore points to the development of M cell-targeted vaccines¹⁹.

Finally, one of the lesser-studied gut epithelial cell types are the Tuft cells, which at a molecular level are very similar to taste cells in the oral cavity⁷. Recent studies showed their involvement in mucosal immunity and initiation of type 2 immunity; characterised by the differentiation of CD4+ T cells into T helper (Th)2 and a distinct cytokine secretion profile²⁰; to helminth infection²¹.

Below the epithelial cell layer is found the lamina propria which hosts a range of innate and adaptive immune cells such as CD103+ dendritic cells (DCs), macrophages, CD8+ T cells and T regulatory cells (see section 1.1.2). The lamina propria also contains structural cells (mesenchymal cells)²², neural networks²³ and rarely epithelium cells.

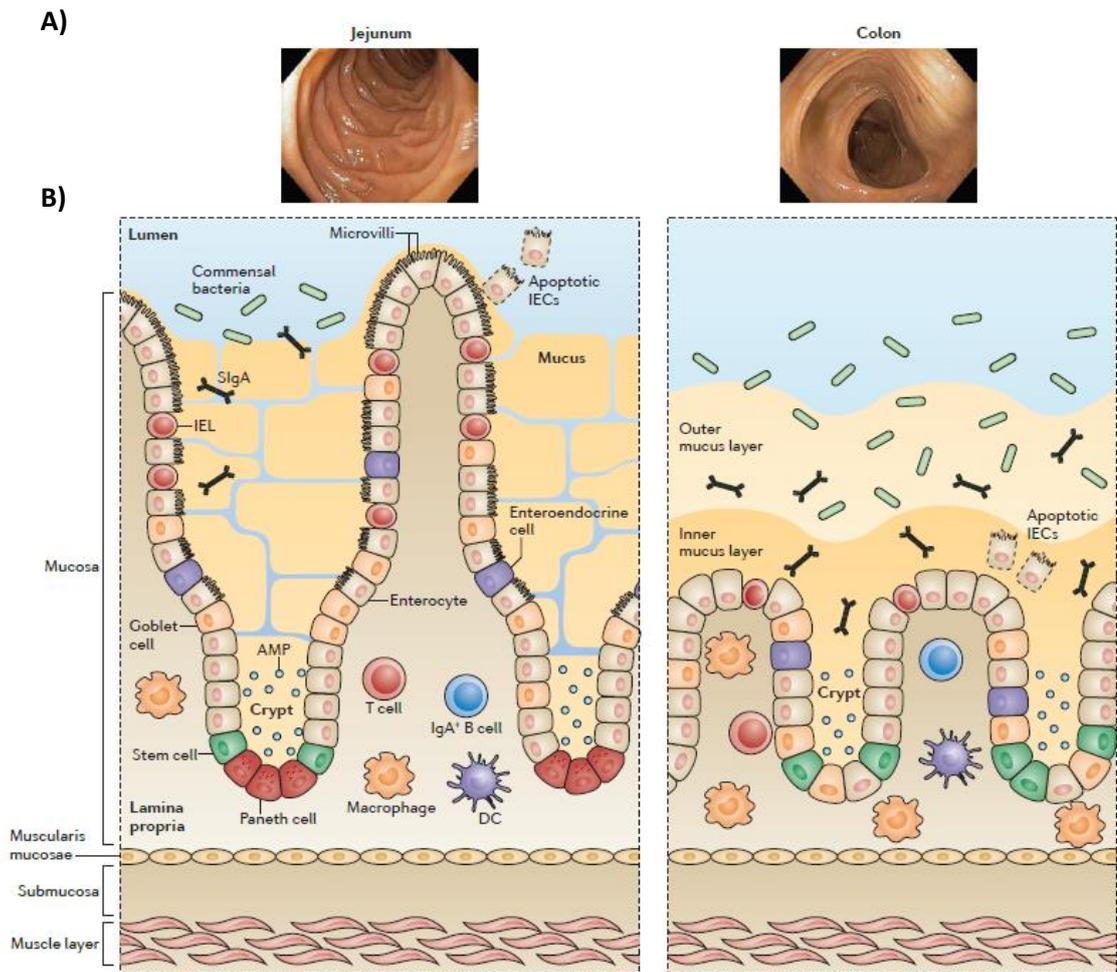


Figure 1 | Anatomy of small (left) and large (right) intestines.

A) Endoscopic view of jejunum (small intestine) and colon. B) Small (left) and large (right) intestine structure showing the mucosa containing the mucus, the epithelial layer, and lamina propria, and underneath the muscularis mucosae, submucosa and muscle layer. IEC, intestinal epithelial cells; IEL, intraepithelial lymphocytes; DC, dendritic cell; SlgA, secretory immunoglobulin A. (Taken from Mowat *et al.* 2014¹⁴).

1.1.2 Immune niches along the gut

In the small and large intestines, the first line of immunity is provided by the epithelium. Further, innate and adaptive immunity is provided by immune cells located in the lamina propria, specifically in the GALT which contains the macroscopic lymphoid follicles, Peyer's patches in the small intestine, caecal and colonic patches, but also smaller follicles of cryptopatches, and isolated lymphoid follicles²⁴ (Fig. 2). The immune cell signature along the GI tract differs largely depending on their location in the gut (Fig. 2) and their originating species (e.g. between human and mouse).

Epithelial cells express PRRs such as toll-like receptors (TLR) or nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs) that recognise microbial components. Upon interaction with microbes, these receptors induce specific intracellular signals through the recruitment of adaptor proteins like myeloid differentiation primary response 88 (MyD88)/ TIR domain-containing adaptor-inducing interferon- β (TRIF) and activation of the transcription nuclear factor- κ B (NF κ B), leading to the production of inflammatory proteins²⁵. Intraepithelial lymphocytes (IELs) can be found interspersed between epithelial cells mainly in the small intestine (Fig. 2) with the majority of them forming cluster of differentiation (CD)8⁺ T cells which are activated after interaction with an APC^{26,27}. IELs are categorised into two groups, type A or B. Type A IELs, T cell receptor (TCR) $\alpha\beta$ +CD8 $\alpha\beta$ + T cells and TCR $\alpha\beta$ +CD4+ T cells, acquire their phenotype after interacting with exogenous stimuli, while, type B IELs, TCR $\gamma\delta$ + T cells, are phenotypically developed independently of the exogenous environment²⁸.

Peyer's patches or colonic patches contain the majority of immune cells. In the lamina propria, the major innate immune cell populations are macrophages and CD103+ dendritic cells (Fig. 2). It is thought that the sampling of antigens can be through the M cell entry, which internalises the antigen and transfer it to the follicle below, in the lamina propria, or by transepithelial dendritic cells which dendrites elongate into the lumen^{29,30}. In mice, the main subsets of dendritic cells are the migratory CD103+CD11b+ and CD103+CD11b- DCs, mainly found in the small intestine, while the equivalent cell subsets in humans are the CD103+Sirpa+ and CD103+Sirpa- DCs^{31,32}, showing high proportion in small intestine and large

intestine, respectively³³. These DCs recognise and internalise antigens and further present them in the major histocompatibility complex II (MHCII).

The macrophage subset found in intestines are the CX3C chemokine receptor (CX3CR)1 macrophages³⁴ which proportion increases from proximal to distal colon¹⁴. These macrophages are highly phagocytic but not migratory, therefore they are not considered as primers of naïve T cells³⁴. However, studies have shown that intestinal macrophages can promote the survival and activity of forkhead box P3 (FOXP3)+ T regulatory cells (Tregs)³⁵. Macrophages and DCs interact with ligands or microbes via the PRRs, such as lectins (as further described in section 1.2) and after MHCII presentation, they migrate to the lymph nodes to activate T cells³⁴.

The intestinal adaptive immune cells in their majority are CD4+ T cells and plasma cells with a small fraction of CD8+ T cells³⁶. Upon interaction of CD4+ T cells with APCs, T cells are differentiated into various subsets, including Tregs, mainly FOXP3 Tregs, essential for immune tolerance³⁷, and T helper cells, Th1, Th2, T follicular helper (Tfh) 2, and Th17. The function of the Th cell subsets is regulated by environmental stimuli³⁸. In mice, Treg cell numbers are decreasing from the colon towards the proximal small intestine whereas Th17 cells are increasing from the colon towards the proximal small intestine (Fig. 2), whereas the limited studies in humans, have shown that both Treg and Th17 cell numbers are higher in the colon as compared to the ileum³⁹. These specific locations of T cells are partially due to the variations in the luminal content along the gut⁴⁰.

In their vast majority, intestinal plasma cells (differentiated B cells) produce IgA with approximately 90% of the total production taking place in the colon. There is an increased B cell accumulation from the cecum to the sigmoid colon which is associated with an increased number of distinct bacterial species⁴¹. Secretion of IgA into the intestinal lumen is triggered by the presence of microbes and mediated by the polymeric immunoglobulin receptor (pIgR), reflecting the regional specificity of plasma cells and microbiota composition in a healthy state⁴¹. pIgR is produced by intestinal epithelial cells and its function is to transfer intraepithelial IgA to the intestinal lumen⁴². In humans, there are two IgA isotypes, IgA1 and IgA2, mainly found in the small intestine and the colon, respectively. The production of the IgA1 subset is driven by protein-antigen interactions while

the IgA2 subset is driven by polysaccharide-antigen interactions⁴³. In addition to the canonical binding between an antibody and its antigen, which includes the Fab variable region, the interaction between IgA and microbes can occur via the noncanonical path, which includes the IgA constant region. This noncanonical binding between IgA and microbes is shown to be mainly mediated by the IgA-glycans⁴⁴.

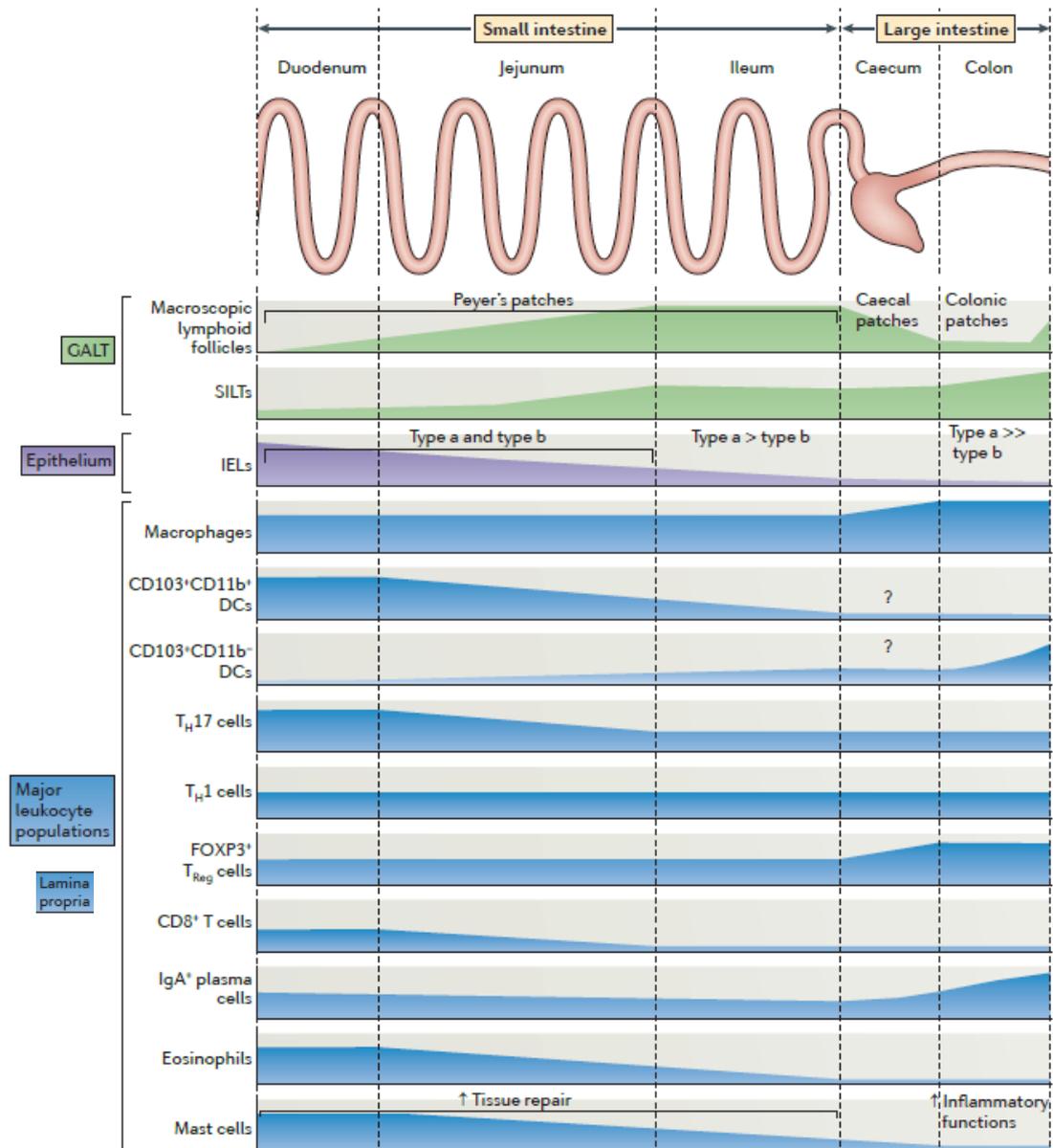


Figure 2 | Immune niches along mouse intestine.

Immune cells are mainly concentrated in the gut-associated lymphoid tissues (GALT) lamina propria, and can also be found in the epithelium as intraepithelial lymphocytes (IELs). The question mark indicates non-characterised regions. SILTs, small isolated lymphoid tissues; DC, dendritic cell; FOXP3, forkhead box P3; IgA, immunoglobulin A. (Taken from Mowat *et al.* 2014¹⁴).

Innate immune cells play a key role in host defence by recognising antigens, a process which then activates adaptive immunity. However, in response to immunoregulatory ligands, innate immune cells can also adopt a tumour promoting phenotype.

1.1.3 Tumour-associated innate immune cells

The immune system can have a dual role in malignant transformation, by eliminating or promoting tumour, which is termed as ‘immunoediting’. During cancer immunoediting the immune system determines tumour’s fate in three steps (Fig. 3): (i) killing of the tumour cells (elimination), (ii) control of tumour expansion and metastasis, a dormant stage where tumour cells are present but often not detectable (equilibrium), and (iii) escape from the malignant microenvironment leading to tumour progression and metastasis (escape)^{45,46}.

Innate immune cells involved in the “escape” tumour mechanisms are called tumour infiltrating immune cells, including tumour-associated macrophages (TAMs), myeloid-derived suppressor cells (MDSCs) and regulatory DCs.

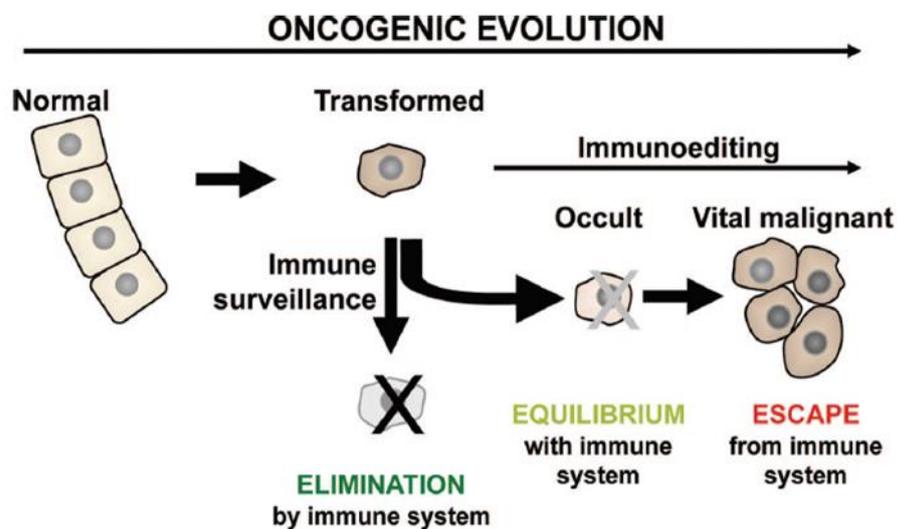


Figure 3 | Oncogenic evolution and immunity.

During cancer immunoediting the immune system is involved in tumour cell elimination (elimination), the control of the malignant cells (equilibrium) and tumour progression (escape). (Taken from Bremnes *et al.* 2011⁴⁵).

Tumour Associated Macrophages (TAMs)

Upon ligand recognition, macrophages can polarise into an M1 or M2 phenotype, with M1 causing an inflammatory response and tumour suppression, while M2 causes an immune-suppression phenotype and a pro-tumour response. As it promotes tumour growth, the M2 phenotype is also called TAM⁴⁷.

When presenting an M1 phenotype, macrophages show induced antigen presentation ability, induction of pro-inflammatory interleukin (IL)-12, IL-23 cytokines and enhanced nitric oxide (NO) and reactive oxygen species intermediates (ROS) production⁴⁸. Furthermore, the M1 phenotype triggers Th1 induction which can suppress tumour progression⁴⁹.

When the M2 phenotype is acquired, macrophages produce elevated levels of IL-10, IL-4, IL-13 and lower levels of IL-12 cytokines and show a decreased production of NO and ROS⁴⁸. Furthermore, TAMs have been shown to induce a Treg phenotype, which is characterised by the suppression of T effector cells (CD8+ T cells)⁴⁷.

Myeloid-Derived Suppressor Cells (MDSCs)

MDSCs are a heterogeneous cell population consisting of myeloid progenitors and immature cells including macrophages, granulocytes, and DCs. In homeostatic conditions, these immature myeloid cells lack suppressive activity and their location is restricted to the bone marrow⁵⁰. Infiltration of MDSCs in certain tissues is associated with a disease state, mostly cancer⁵¹. After MDSC stimulation, the cells adopt a suppressor phenotype with the production of NOS, ROS and expression of programmed death-ligand 1 (PD-L1) leading to further inhibition of T and natural killer (NK) cell action^{52,53}. Additionally, MDSCs are producing transforming growth factor beta (TGF- β) and IL-10 which induce activation of Treg and TAM cells⁵⁴.

Regulatory DCs

DCs, upon exposure to a specific environment, can adapt a suppressive phenotype by inhibition of innate or adaptive immunity. It has been suggested that the level of DC maturation is critical for the adaptation of a suppression phenotype⁵⁵. Regulatory DCs, specifically immature DCs, can infiltrate tumour sites, induce the production of TGF- β and further promote Treg proliferation⁵⁶. Maturation of DCs induces an inflammatory response through the overexpression of MHCII, co-stimulatory molecules and pro-inflammatory cytokines⁵⁷.

1.2 Intestinal lectins in immunity and cancer

Lectins are carbohydrate-recognition proteins and function as PRRs. In the human gut, lectins are mainly expressed in leukocytes but some are also produced by epithelial cells^{58,59}. Mammalian lectins recognise carbohydrate antigens expressed endogenously by the host, or exogenously present in the diet or on microbes⁶⁰. The protein-carbohydrate interaction occurs through the carbohydrate recognition domain (CRD) of lectins. This type of interaction is mostly specific but occurs with low affinity and requires multivalency derived from both lectins and carbohydrate antigens to achieve avidity of binding⁶¹. Host lectins can be secreted or bound to the cell membrane (transmembrane proteins). Secreted lectins can pass through the endothelium into the blood circulation and be involved in cell-cell adhesion via their CRD binding to the carbohydrate antigens expressed on host cells⁶². The transmembrane lectins can be involved in cell adhesion, antigen endocytosis, and intracellular signalling for immunomodulation⁶⁰. Lectins are classified based on their amino acid sequence homology into superfamilies including C-type lectins (CTLs), Galectins, and Siglecs.

1.2.1 *C-type lectins (CTLs)*

CTLs or Ca²⁺-type lectins, are named after their requirement for calcium to stabilise and bind to the carbohydrate hydroxyl group of the antigen^{63,64}. While the majority of CTL-antigen binding is calcium dependent, there are examples within this family of CTLs binding to ligands in a calcium independent manner⁶⁵. CTLs are involved in immune homeostasis and immune response upon pathogen recognition⁶⁶ and bind to a range of carbohydrates via their conserved amino acid recognition motifs, “EPN” (Glutamic acid-Proline-Asparagine) or “WND” (Tryptophan-Asparagine-Aspartic acid). The “EPN” motif binds to glucose, mannose, fucose, and N-acetylglucosamine (GlcNAc) while “WND” binds to galactose and N-acetylgalactosamine (GalNAc). In humans, there are at least 17 types of CTLs expressed by immune or tissue cells such as endothelial and epithelial cells (see Fig. 4 for a list of CTLs expressed by myeloid cells). Myeloid CTLs include a variety of proteins which possess different cell signalling patterns, and result in inflammatory activation or suppression⁶⁷. Cellular inhibition occurs when the immunoreceptor tyrosine-based inhibitory motif (ITIM, canonical

sequence S/I/V/LxYxxI/V/L) is activated, while activation occurs when the immunoreceptor tyrosine-based activation motif (ITAM, Yxx(L/I)₆₋₁₂ Yxx(L/I)) is activated (Fig. 4). Dendritic cell-associated C-type lectin (Dectin)-1 CTL, constitutes an exception as the binding to its ligands is calcium independent^{68,69}.

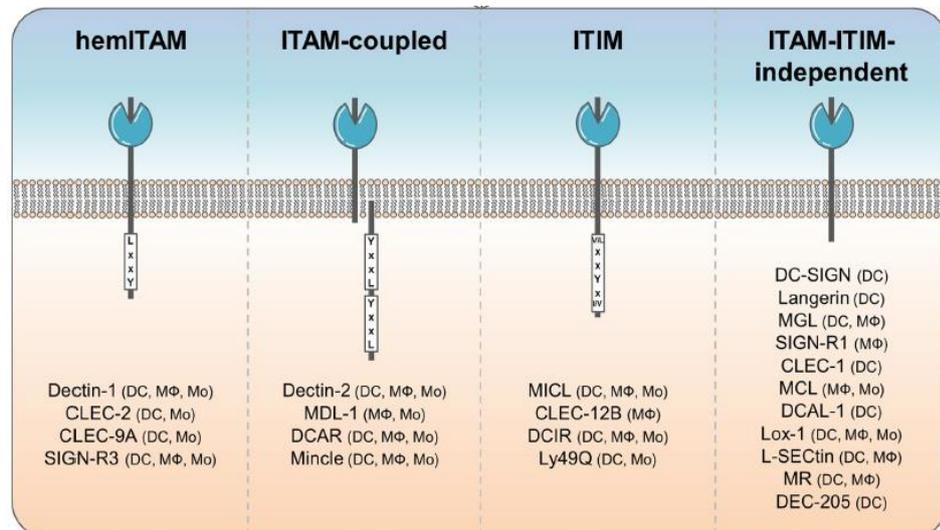


Figure 4 | C-type lectins in myeloid cells.

Myeloid C-type lectins and their intracellular signalling. HemITAM are characterised by a single ITAM motif. ITAM, immunoreceptor tyrosine-based activation motif; ITIM, immunoreceptor tyrosine-based inhibition motif. (Taken from Sabine Mayer *et al.* 2016⁶⁷).

1.2.1.1 Dectin-2

The CTL, Dectin-2 or Clec6A, was first found to be expressed in Langerhans cells, an immature form of DCs residing in epidermis and mucosal tissues⁷⁰ but later, Dectin-2 was shown to be expressed in a range of cell types including macrophages and dendritic cells⁷¹. Dectin-2 interacts with mannose residues that are presented on pathogenic bacteria, such as *Streptococcus pneumoniae*⁷², *Hafnia alvei*⁷³, and *Salmonella enterica*⁷³, fungi such as *Candida albicans*⁷⁴, *Aspergillus fumigatus*⁷⁵ and *Saccharomyces cerevisiae*⁷⁶ (Table 1). Upon binding with its ligand, Dectin-2 couples with the signalling adaptor molecule FcR γ which bears ITAM (Fig. 4). After activation, the spleen tyrosine kinase (Syk) is phosphorylated and initiates a cascade of signalling which results in the activation of transcription factors such as NF κ B⁷⁷. Dectin-2 has also been shown to

contribute to liver metastasis suppression by mediating tumour cell phagocytosis in Kupffer cells⁷⁸.

1.2.2 Galectins

In humans, galectins (former S-type lectins) cover a family of 11 proteins (Fig. 5A) known for their self-recognition ligands for host developmental processes such as cell differentiation and host immune homeostasis⁷⁹. In addition, galectins have been linked to diseases such as fibrosis⁸⁰, cancer (tumour metastasis)⁸¹, and heart disease⁸². Most galectins are non-glycosylated soluble proteins and are found either in the cytosol or extracellularly⁸³. Secretion of galectins occurs in a non-canonical pathway via their incorporation into vesicles such as exosomes⁸⁴. They typically recognise β -galactosides via their CRD, while Galectin-3 has also shown binding to β 1,2-linked oligomannans expressed on fungi⁸⁵. Galectins are classified according to their structure as prototype (single CRD and potential homodimer formation), tandem repeat (two CRD domains linked with a small peptide domain) or chimeric type (single CRD with large amino-terminal domain which can form oligomers; only Galectin-3) (Fig. 5A).

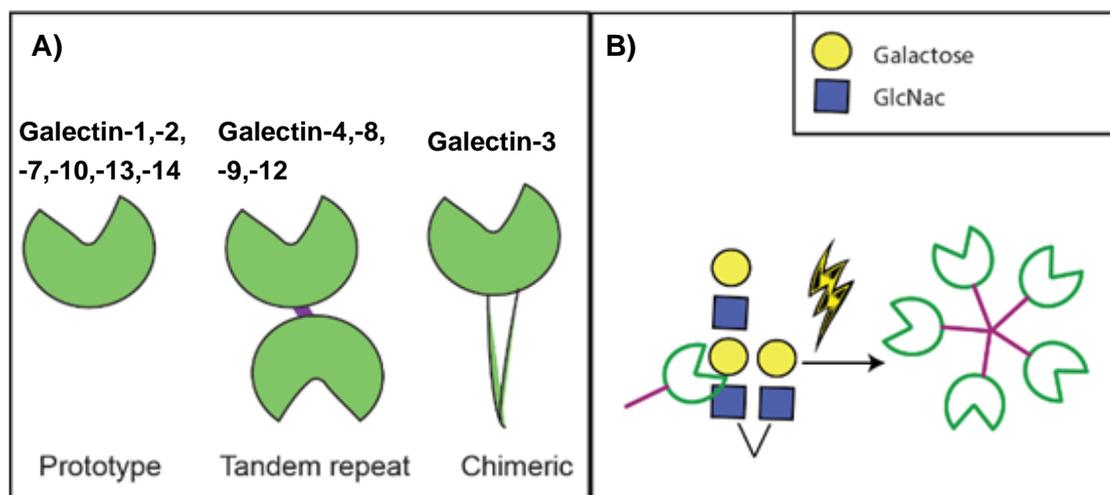


Figure 5 | Human Galectins.

A) The different family members of galectins, prototype, tandem repeat and chimeric.
B) The pentamer formation after binding of monomer Galectin-3 to its ligand.

1.2.2.1 Galectin-3

Galectin-3 is the unique member of the family with a chimera type structure, in which the N-terminus region can polymerise after activation of CRD, to form a pentamer (Fig. 5B). Galectin-3 is expressed by dendritic cells, macrophages, epithelial cells and tumour cells and is mainly localised in tumour cells, epithelial cells, fibroblasts, macrophages and activated T cells. Galectin-3 has pleiotropic biological functions such as influencing cell growth, cell adhesion, cell-cell interaction, and as a pre-mRNA splicing factor⁸⁶. Furthermore, Shan *et al.* 2013 showed that gut homeostasis is enhanced via the complex formation between MUC2, Galectin-3 and the CTL Dectin-1⁸⁷. Extracellular Galectin-3 has been shown to interact with a range of pathogenic bacteria, as listed in Table 1, and further regulate innate immune response. Galectin-3 can also recognise commensal bacteria as shown with *Bifidobacterium longum*⁸⁸. Galectin-3 expressed in macrophages has also been proposed to play a role in distinguish pathogenic from commensal fungi⁸⁹. Interestingly, the study from Kohatsu *et al.* 2006 showed a direct killing of *Candida albicans* after binding to Galectin-3, through a non-typical galectin ligand, β 1,2-linked oligomannans, due to the morphological cell change⁸⁵.

1.2.3 Sialic acid-binding immunoglobulin-like lectins (Siglecs)

In humans, Siglecs encompass a superfamily of 15 transmembrane proteins⁹⁰ (Fig. 6) which are involved in biological processes such as endocytosis, immune modulation and cell-cell interaction⁹¹. In addition, Siglecs have been extensively studied *in vitro* for their involvement in cancer progression⁹² and other conditions, through immune suppression. More recently, researchers are focusing on developing therapeutic approaches by targeting Siglecs⁹³⁻⁹⁵. The extracellular region of these lectins consists of an amino terminal CRD V-set domain that recognises *cis* (ligands expressed by the same cell) or *trans* (ligands expressed by other cells) sialosides⁹¹. The extracellular region of Siglecs also carries a variable number of immunoglobulin domains referred to as the C2-set. In addition, all Siglecs except for Siglec-1 and -4, bear either an activation or an inhibition domain in their intracellular regions which motif varies depending on the Siglec. Siglec-14, -15 and -16 bear an intracellular ITAM while the other members of the Siglec superfamily (except Siglec-1 and -4) contains an ITIM (Fig. 6)⁹⁰. Based on

DNA sequence similarities, Siglecs can be subdivided into two groups: a CD33-related Siglec group including Siglec-3, -5, -6, -7, -8, -9, -10, -11, -12, -13, -14, -16 and a group including Siglec-1, -2, -4 and -15⁹¹.

Some Siglecs have been shown to interact with the cell surface microbe-associated molecular patterns (MAMPs) of pathogenic bacteria, with sialic acid-glycoconjugates⁹⁶ as typical ligands, although, other ligands have been reported. For example, Group B *Streptococcus* (GBS) has been shown to interact with Siglec-5 and Siglec-7 in a sialic acid-independent manner through its β -protein^{97,98}. In addition, Stephenson *et al.* 2014 showed that Siglec-10 bind to a sialic acid-like molecule, pseudaminic acid, present in *Campylobacter jejuni*-flagella⁹⁹.

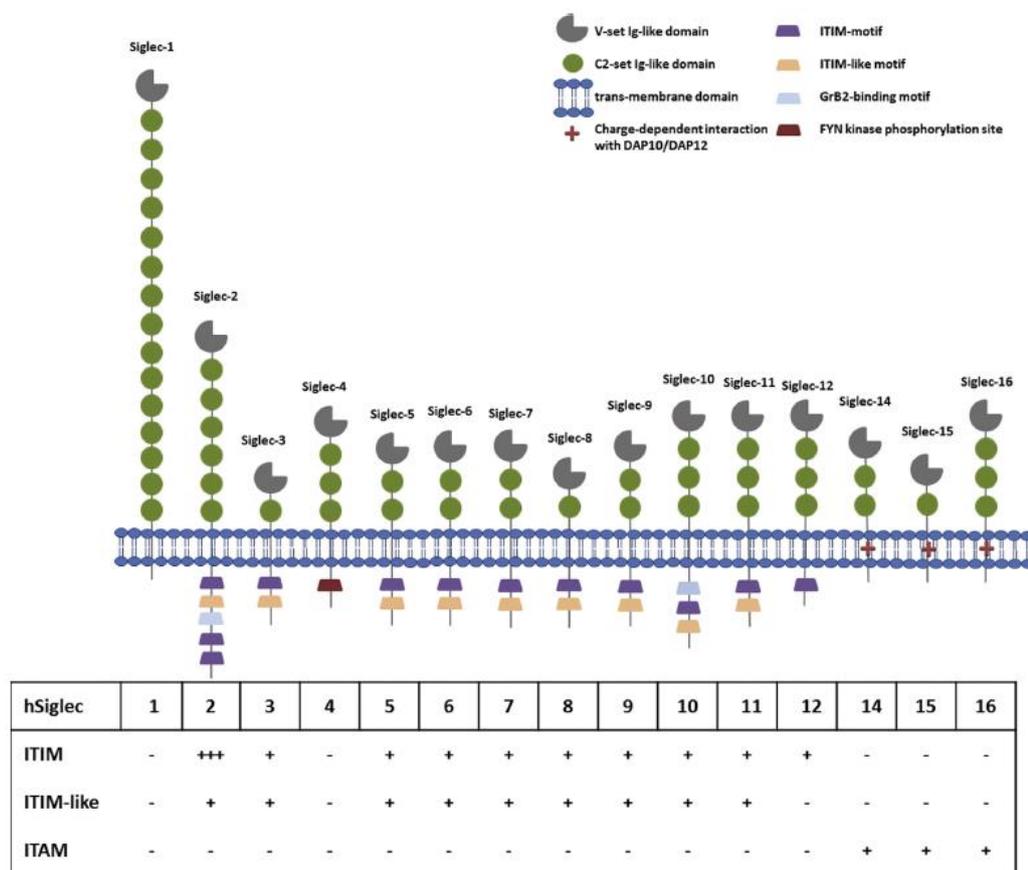


Figure 6 | Human Siglecs and their intracellular motifs.

Structure of the 15 human transmembrane Siglec proteins (taken from Bornhöfft *et al.* 2018⁹⁰).

1.2.3.1 Siglec-7

Siglec-7 (or CD328) is expressed by NK cells, macrophages and DCs of peripheral blood and colonic lamina propria cells. While in peripheral blood Siglec-7 is highly expressed by NK cells, Miyazaki *et al.* 2012 showed that, in the gut lamina propria, Siglec-7 is expressed predominantly in monocytes/macrophages¹⁰⁰, reflecting the tissue specificity of Siglec expression. Siglec-7 consists of an N-terminal CRD V-set Ig domain, two C2-set Ig regions, an intracellular membrane-proximal ITIM and a distal ITIM-like motif¹⁰¹ (Fig. 7). The CRD domain of Siglec-7 preferentially recognises α 2,8-disialylated and branched α 2,6-sialylated glycoconjugates but also α 2,3 linked sialic acids¹⁰². Upon ligand recognition of the extracellular domain, the tyrosine-based ITIM's are phosphorylated by SRC kinases and further recruit SRC homology domain 2-containing tyrosine phosphatase-1 and 2 (SHP1 and SHP2), resulting in the suppression of the inflammatory signalling¹⁰³. The membrane-proximal ITIM is more potent than ITIM-like in SHP1 and SHP2 recruitment and in the inhibitory signalling¹⁰⁴ (Fig. 7).

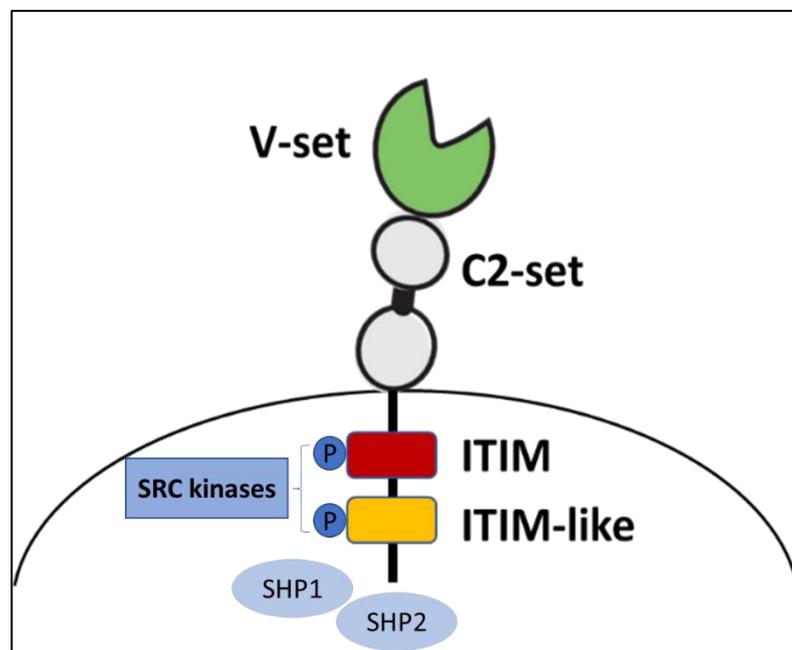


Figure 7 | Siglec-7 structure and its mediated signalling.

The extracellular domain of Siglec-7 consists of a V-set and two C2-set domains. Upon activation, the intracellular ITIM motifs are phosphorylated by SRC kinases and further recruit the SHP1 and 2 phosphatases. ITIM, immunoreceptor tyrosine-based inhibitory motif; SHP, SRC homology domain 2-containing tyrosine phosphatase.

Siglec-7 has been shown to interact with bacteria *C. jejuni*¹⁰⁵, *E. coli* strains¹⁰⁶ and *Pseudomonas aeruginosa*¹⁰⁷ through their surface sialic acids, while as mentioned above, the interaction with GBS is sialic acid-independent through its β -protein⁹⁸. Studies also reported interactions of Siglec-7 with the yeast *C. albicans* and *S. cerevisiae*¹⁰⁶ and more specifically with zymosan yeast particles¹⁰⁶ (Table 1).

Siglec-7 and Siglec-9 have been shown to function as inhibitory receptors similar to the well-known immune checkpoint receptors programmed cell death 1 (PD-1) and cytotoxic T-lymphocyte-associated protein 4 (CTLA-4); inhibitors of T cell activation¹⁰³. This triggered increased interests in investigating Siglecs as target for cancer immunotherapy, as recently reviewed^{94,95}. Recently, a genome-wide CRISPR screen revealed the glycoprotein CD43 expressed on leukemia cells as a highly specific ligand for Siglec-7 and blocking of the interaction between CD43 and Siglec-7 induced the immune killing activity against leukemia cells, which further supports the proposed role of Siglec-7 as an immune checkpoint receptor¹⁰⁸. Alternative strategies to block or prevent Siglec-7-sialoside interactions include the editing of the glycocalyx by sialidases¹⁰⁹ for the development of new cancer immune therapies¹¹⁰. Recently, Yamada *et al.* 2021 showed that Siglec-7 expression in macrophages present in the tumour microenvironment could serve as a biomarker for the efficacy of the immunotherapy against metastatic colorectal cancer¹¹¹.

Table 1 | Galectin-3, Dectin-2, and Siglec-7: Glycan or pathogen recognition, and cell expression

Lectin	Family	Ligand specificity	Cell expression	Bacteria recognition
Galectin-3	Galectins	β -galactosides, β -1,2-oligomannan	Epithelial cells, macrophages, dendritic cells, tumour cells,	<i>Neisseria gonorrhoeae</i> ¹¹² , <i>C. albicans</i> ⁸⁵ , <i>Toxoplasma gondii</i> ¹¹³ , <i>Schistosoma mansoni</i> ¹¹⁴ , <i>Trypanosoma cruzi</i> ¹¹⁵ , <i>Klebsiella pneumoniae</i> ¹¹⁶
Dectin-2	C-type lectins	α -mannans	Dendritic cells, macrophages, monocytes	<i>S. pneumoniae</i> ⁷² , <i>A. fumigatus</i> ⁷⁵ , <i>H. alvei</i> ⁷³ , <i>S. enterica</i> ⁷³ , <i>C. albicans</i> ⁷⁴ , <i>S. cerevisiae</i> ⁷⁶
Siglec-7	Siglecs	α 2,8-disialic acids, branched α 2,8- or α 2,3-sialylates, β -protein	NK cells, Dendritic cells, macrophages, monocytes, T lymphocytes	<i>C. jejuni</i> ¹⁰⁵ , <i>E. Coli</i> ¹⁰⁶ , <i>P. aeruginosa</i> ¹⁰⁷ , GBS ⁹⁸ , yeasts (<i>C. albicans</i> , <i>S. cerevisiae</i>) ¹⁰⁶

1.3 Gut microbiota

1.3.1 Structure and function of the human gut microbiota

The human gut is the most microbe-populated organ in the body containing trillion of microbes, known as the gut microbiota, covering bacteria, fungi, virus and archaea. In a healthy state, the gut microbiota is unique for every individual and is highly diverse¹¹⁷. It consists of a wide range of bacterial genera mainly *Bacteroides* and anaerobic cocci with lower abundance of species from the *Clostridium*, *Bifidobacterium*, *Eubacterium*, *Lactobacillus*, and *Streptococcus* genera^{14,118}. The diversity and abundance of bacteria vary along and across the GI tract and increase from the small intestine to the colon (Fig. 8)^{1,14}. The diversity of microbes is dependent on several factors such as the mode of delivery¹¹⁹, geographic location¹²⁰, diet¹²¹, age¹¹⁹, antibiotic usage¹²² and host genetics¹²³. The commensal bacteria have established a symbiotic relationship with the human host, providing beneficial metabolites, protection from pathogenic microorganisms and maturation of host's immunity, while, in return, benefiting from the nutrient-rich host environment¹²⁴.

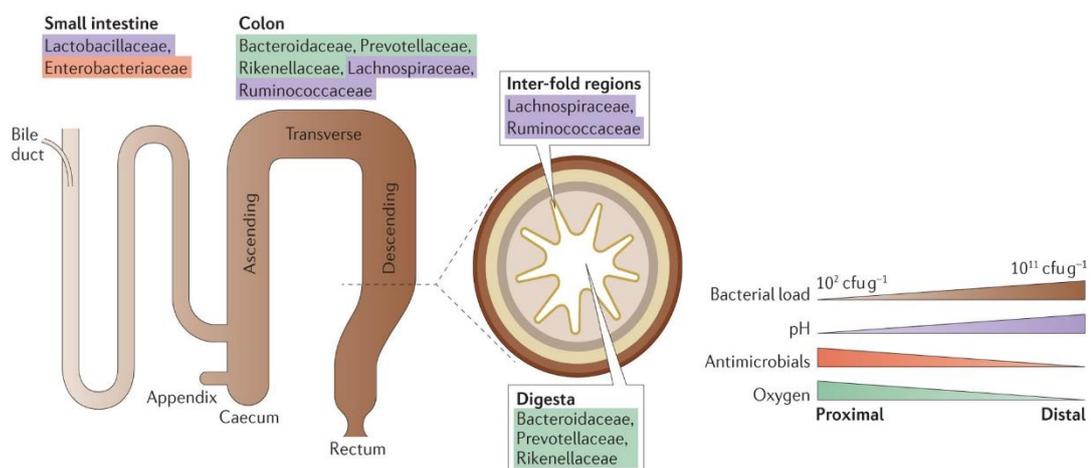


Figure 8 | Bacteria distribution along and across the lower GI tract.

Changes in physicochemical conditions along the gut (pH, antimicrobial peptides and oxygen) lead to a regional specificity in bacterial family dominance, with colon carrying the higher bacterial load when compared to the small intestine. The cross-section shows the bacterial family dominance in the digesta and the inter-fold regions of the lumen. CfU, colony-forming units. (Taken from Donaldson *et al.* 2015¹).

Humans metabolise simple sugars such as sucrose but are lacking the enzyme required for the digestion of complex polysaccharides, such as cellulose or resistant starch. Bacteria produce carbohydrate-active enzymes (CAzymes, www.cazy.org) that digest these sugars and produce short chain fatty acids (SCFAs), including acetate, butyrate, and propionate, which are beneficial for the host¹²⁵. Other compounds such as dietary polyphenols, reach the colon where bacteria can metabolise them into compounds that can pass through the intestinal barrier¹²⁶. Some members of the gut microbiota synthesise vitamins, such as vitamin K, for their survival, which are also beneficial to the host¹²⁷.

Commensal bacteria protect the host from infection through the production of antimicrobial compounds such as bacteriocins¹²⁸, secondary bile acids¹²⁹ and lactic-acid¹³⁰ which can act directly by targeting the physiology or integrity of the pathogens or indirectly by competing with the pathogens for binding sites or nutrients¹²⁹.

Host immune cells mature through their constant interactions with the gut microbiota¹³¹. This mechanism is mediated by the interaction between PRRs present on the surface of immune cells (see section 1.2) and the microbial associated molecular patterns (MAMPs) on the gut microbial cell surface, such as lipopolysaccharides (LPSs) or capsular polysaccharides (CPSs)^{132,133} (see section 1.3.2). The resulting interaction between host cells and the gut microbes can modulate cytokine and chemokine production as well as the production of secretory immunoglobulin A (sIgA), which plays an important role in gut homeostasis⁴² (see section 1.1.2).

Alterations in the gut microbiota composition or niche can disrupt the symbiotic relationship with the host and result in adverse health outcomes, a state called dysbiosis. Dysbiosis can lead to infectious diseases or chronic conditions such inflammatory bowel disease (IBD)¹³⁴, obesity¹³⁵, diabetes¹³⁶, cardiovascular disease¹³⁷, or cancer development¹³⁸. In particular, Gram negative (Gram-) bacteria are a main public health target as they can cause disease and show multi-resistance to antibiotics^{139,140}.

1.3.2 Cell surface glycosylation of bacteria

Glycosylation of bacteria can affect host-bacteria interactions and innate immune responses^{141–147}. Bacteria can synthesise a variety of glycoconjugates such as peptidoglycan (PG) in the periplasm, extracellular polysaccharide (EPS), CPS, glycoproteins and LPS¹⁴⁸ through complex biosynthetic pathways.

Gram positive (Gram+) bacteria contain a thick peptidoglycan layer (~90% of the wall) decorated by glycosylated teichoic acids¹⁴⁹ (Fig. 9A) while Gram- contain a thinner layer (~10% of the wall)¹⁵⁰. Gram- bacteria contain an additional outer membrane (OM) made up of LPS and lipooligosaccharide (LOS), absent in Gram+ (Fig. 9A), covering ~80% of the bacterial surface and providing an extra barrier to the cell¹⁵¹.

Both Gram- LPS and LOS function as a defence system against environmental stress such as toxic molecules and antibiotics. LPS contributes to cell permeability and has been shown to affect negatively hydrophobic antibiotic efficacy¹⁵². LPS is also called endotoxin and has been extensively studied for its immunomodulatory function¹⁵³.

Structurally, LPS and LOS share the lipid A and the core oligosaccharide regions, with LPS having an additional region, the O-antigen polysaccharide (Fig. 9B). The lipid A domain from pathogen-derived LPS binds to TLR-4 receptors, resulting in a pro-inflammatory host response¹⁵⁴. In addition, TLR-4 receptors can act synergistically with lectins for immune response enhancement^{73,91}. The O-antigen is highly diverse among species and together with the CPS, determines the serotype of the strain.

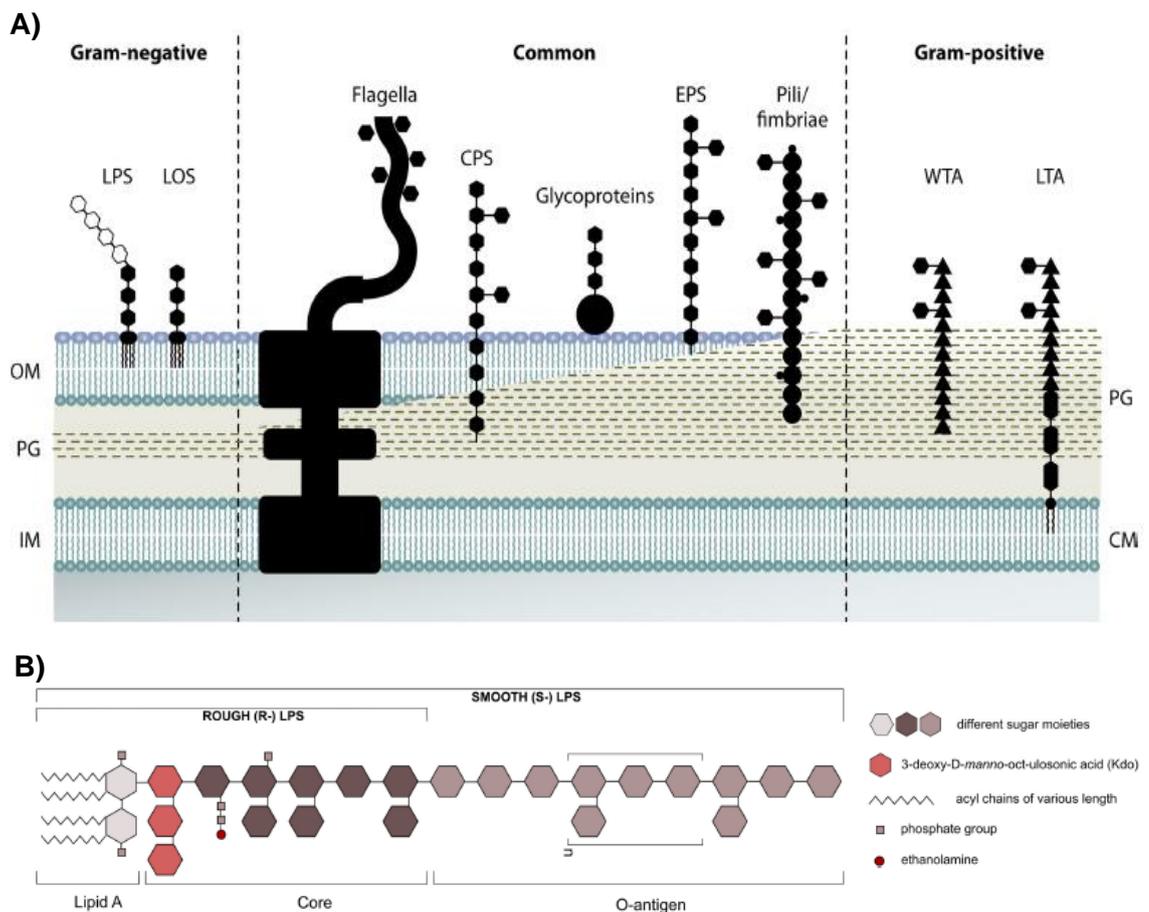


Figure 9 | Glycoconjugates in bacterial cell membranes.

A) Gram+ bacteria constituted of a cytoplasmic membrane, covered by a thick layer of peptidoglycans. Gram- bacteria are constituted of an inner membrane, periplasm, and outer membrane displaying glycoconjugates such as LPS (lipopolysaccharides). B) LPS structure showing the distinct regions corresponding to lipid A, core oligosaccharide and O-antigen. LOS (lipooligosaccharides); CPS (capsular polysaccharides); EPS (extracellular polysaccharides); WTA (wall teichoic acids); LTA (lipoteichoic acids); OM (outer membrane); PG (peptidoglycans); IM (inner membrane); CM (cytoplasmic membrane). (Taken from Tytgat, H. L. P. & Lebeer, S.A. 2014¹⁴⁸).

Gram- and Gram+ bacteria release nano-sized membrane vesicles into the extracellular environment either in a constitutive or in a regulated manner¹⁵⁵. Recent progress in this field supports the pathophysiological functions of these vesicles in bacteria-host interactions¹⁵⁶. Gram- bacteria produce outer membrane vesicles (OMVs) formed in the OM, as buds, with sizes ~20-250 nm, which are released into the luminal environment¹⁵⁷. Their surface shares the molecular characteristics as the OM (containing LPS and proteins presented on the OM), and internally contain several types of molecules such as hydrolases, or

misfolded proteins that can otherwise be toxic for the bacteria, or bacterial DNA¹⁵⁸. In addition, OMVs can carry antibiotic-degrading enzymes against β -lactam antibiotics¹⁵⁹ which makes them a potential target for antimicrobial strategies. OMVs also interact with the host, in the oral cavity for plaque formation, and in the gut with epithelial and immune cells¹⁵⁶. In recent years, the immunomodulatory properties of OMVs are being exploited for the development of bioengineered OMVs applicable to the biomedical field, especially in vaccine development, as OMVs have high adjuvanticity^{157,160,161}.

1.4 Colorectal cancer (CRC) and associated microbiota

CRC is the 3rd most common type of cancer worldwide, it commonly appears in adults over 50 years old although recent data showed an increase in the prevalence of CRC in younger adults^{162–165}.

The identified causes of CRC include the interactions between genetic factors (such as chromosomal instability pathways) and physical (such as radiation), chemical (such as food contaminants) or biological (such as certain microbes) carcinogen exposures. The majority of colon cancer cases are sporadic, with no family history or hereditary genetic alterations¹⁶³ which highlights the importance of lifestyle factors that can be modulated to reduce colon cancer incidence. In addition, it is now well-established that the gut microbiota plays a critical role in tumour development and therapeutic efficacy¹⁶⁶. The CRC tissues show a reduced microbial diversity compared to healthy tissues and harbour a distinct microbiota profile characterised by a high abundance of bacterial species such as *Fusobacterium nucleatum*, *Bacteroides fragilis*, *E. coli*^{167,168} and colibactin-producing bacteria¹⁶⁹. The CRC microbiota profile evolves during cancer development, with significant differences in bacterial species dominance between adenoma and cancer stage¹⁷⁰.

In addition, metagenomic studies reported that tumour tissues are enriched in LPS biosynthesis genes^{171,172}, as well as high circulating LPS levels¹⁷³, suggesting that LPS may be a contributor in CRC progression.

Several pathways¹⁷⁴ have been proposed for the involvement of the gut microbiota to CRC including the apha-bug, driver-passenger, biofilm, microbiota adaptation and bystander effect theories, as described below (Fig. 10).

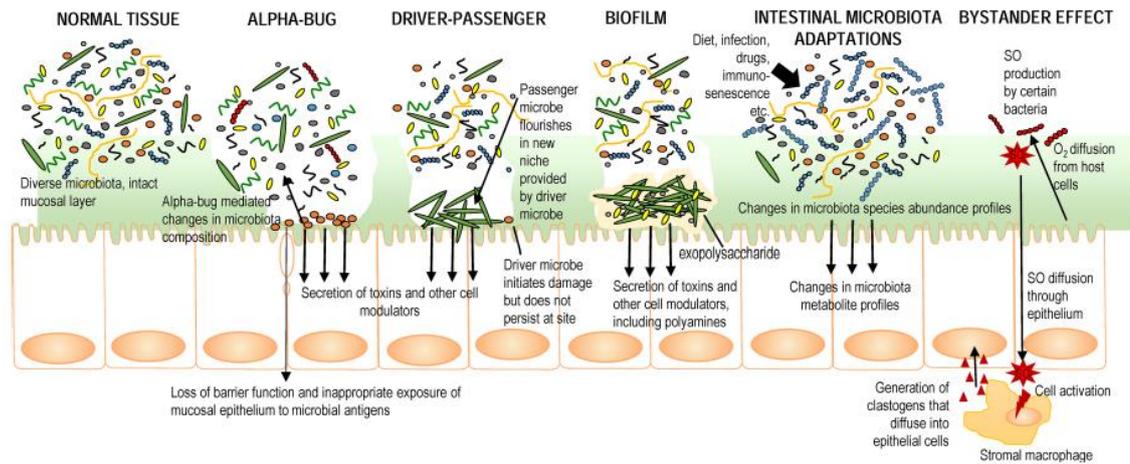


Figure 10 | The microbial hypotheses in CRC progression.

The proposed pathways (alpha-bug, driver-passenger, biofilm, microbiota adaptation and bystander effect) of microbiota involvement in CRC progression. (Taken from Van Raay *et al.* 2017¹⁷⁴).

The alpha-bug hypothesis suggests that certain bacterial species are able to invade the host gut barrier by altering microbiota composition and epithelial barrier integrity, through the secretion of molecules, causing a pro-tumour host response¹⁷⁵.

In the driver-passenger hypothesis, initially an alpha-bug microbe disrupts the microenvironment, leading to altered conditions that then create a metabolic environment favoring a passenger bacterial species to overgrow, compete with the alpha-bug microbe, and drive a pro-tumour host response¹⁷⁶. Recently, computational analyses showed that the metabolic signature in a tumour microenvironment benefits CRC-associated bacteria such as *F. nucleatum*¹⁷⁷.

In the biofilm hypothesis, bacteria such as *E. coli*, *B. fragilis* or *F. nucleatum* form biofilms^{178,179} (polymicrobial communities coated by a polymer structure) adhering to the gut epithelium and altering the host's metabolism¹⁸⁰.

In the microbiota adaptation hypothesis, environmental components (e.g. diet or drugs) cause alteration in the microbiota composition with a distinct metabolite profile¹⁸¹. The newly established microbial niche and their products can either have a pro- or anti-tumour effect. Mehta *et al.* 2017 also showed that prudent diet

(high in dietary fibers and whole grains) has been associated with lower *F. nucleatum*-associated CRC risk¹⁸².

In the bystander effect, bacteria produce certain metabolites which interact with host cells and further promote tumourigenesis, an example of this pathway is *Enterococcus faecalis* which produces superoxide as a metabolite¹⁸³.

In addition to the above pathways, certain bacterial species from the Enterobacteriaceae family or *Sphingomonas* genus are suggested to contribute to CRC through chronic inflammation of the gut epithelium, known as colitis-associated cancer¹⁸⁴.

1.4.1 *Bacteroides fragilis*

B. fragilis is a Gram- bacterial species from the Bacteroidaceae family, Bacteroidales order and Bacteroidetes phylum. As described above, enterotoxigenic *B. fragilis* (or EBF) strains have been shown to act as “alpha-bugs” by producing a toxin that can cause tumour progression in CRC¹⁸⁵. EBF has also been implicated together with *F. nucleatum* in biofilm formation, particularly in the human proximal colon¹⁸⁶. *B. fragilis* expresses a metalloprotease toxin (referred to as BFT) that induces a pro-inflammatory response and cleaves E-cadherin (a transmembrane protein that maintain epithelial cells together)¹⁸⁷. This, in turn, results in a loosening of the epithelium and an increased exposure to pro-oncogenic luminal antigens. E-cadherin binds to β -catenin, and the cleavage of the former by BFT causes increase levels of intracellular β -catenin leading to an enhance transcription of pro-oncogenic genes¹⁸⁸.

Enterotoxigenic strains of *B. fragilis* have also been shown to induce the signal transducer and activator of transcription 3 (STAT-3), a transcription factor which induces Th17 immune response and contributes to tumour progression¹⁸⁹. At the structural level, *B. fragilis* is decorated by a CPS containing the polysaccharide (PS) A and B¹⁹⁰. To date, PS has been shown to have a beneficial effect in mouse immunity development¹⁹¹, and colitis prevention¹⁹².

1.4.2 *Fusobacterium nucleatum*

F. nucleatum is a Gram- bacterial species from the Fusobacteriaceae family, Fusobacteriales order and Fusobacteria phylum. It is an obligate anaerobic, non-spore and non-motile bacterium, which colonises the oral cavity of humans¹⁹³. *Fusobacterium nucleatum* species have a rod-shaped morphology and are divided into 5 subspecies, *F. nucleatum* ssp. *nucleatum*, *F. nucleatum* ssp. *polymorphum*, *F. nucleatum* ssp. *animalis*, *F. nucleatum* ssp. *vincentii* and *F. nucleatum* ssp. *fusiforme*, (refers as *F. nucleatum* ssp. in the rest of the thesis). *F. nucleatum* ssp. have been involved in many pathological conditions such as halitosis, pregnancy complications and colon cancer¹⁹⁴. They are found to be abundant in faeces and biopsies derived from patients with CRC^{168,184,195} with *F. nucleatum* subspecies *animalis* being most prevalent¹⁹⁶. The association of this species with CRC is consistent with the “driver-passenger” hypothesis (section 1.4), as *F. nucleatum* ssp. are normally poor colonisers of the gut. The presence of *F. nucleatum* correlates with CRC development as it is mainly found in sporadic tumour cases¹⁸⁴. In CRC tissue specimens, *F. nucleatum* ssp. are mainly present in the tumour site as compared to adjacent normal site and they are over-abundant at the adenocarcinoma stage¹⁹⁷. *F. nucleatum* ssp. are mainly found in the right-side of the colon as single planktonic species^{198,199} or in multi-microbial biofilms¹⁸⁰, reflecting regional microbial diversity along the GI tract. *F. nucleatum* is also associated with sessile serrated adenomas, a type of adenoma with higher risk to progress to tumour^{198,200} and poor CRC prognosis²⁰¹. Bullman *et al.* 2017 showed an association of *F. nucleatum* with metastases, suggesting that these bacteria can travel from the primary tumour site together with the cancer cells to a distal site²⁰², and recently Chen *et al.* 2020 showed that metastasis could be induced after autophagy activation promoted by *F. nucleatum* species²⁰³. *In vitro* experiments using human tumour spheroids showed that live *F. nucleatum* bacteria formed aggregates in the tumour microenvironment, while heat-killed bacteria were internalised²⁰⁴. Together these studies provide compelling evidence for the importance of *F. nucleatum* ssp. in tumorigenesis.

Investigations of the underpinning mechanisms point towards a role in innate immune cell modulation as described in experimental studies below. Transcriptomic analyses of CRC specimens by RNA-seq showed that *F. nucleatum* ssp. are associated with an expansion of myeloid-derived immune

cells¹⁹⁵. More specifically, deep transcriptomic analysis of patient samples harboring non IBD-associated tumours showed that *Fusobacterium*-CRC tumours are associated with an increased expression of genes associated with MDSCs (CD33, IL6), TAMs (CD209, CD206, IL-6, IL-8 and CXCL-10) and DCs (CD209, TNF α , CD80)¹⁹⁵. The same tumour-associated immune cell phenotype was also observed in *F. nucleatum*-fed Apc^{Min/+} mouse model¹⁹⁵. The expression of these genes are implicated in a pro-inflammatory response induced by the NF- κ B transcription factor²⁰⁵. It is also reported that high levels of NF- κ B expression was associated with high abundance of *Fusobacterium* in CRC specimens¹⁹⁵. In contrast, there was no significant differences in CD8+ and CD4+ T cells between *F. nucleatum*-fed and control Apc^{Min/+} mice¹⁹⁵. Furthermore, PD-L1 which is involved in tumour progression²⁰⁶, was shown to be increased in *Fusobacterium*-enriched CRC tissues¹⁹⁵. Ye *et al.* 2017 showed that *F. nucleatum*-CRC tumour sites have a significant increase in the chemokine (C-C motif) ligand (CCL)20, IL-17A and tumour necrosis factor α (TNF α) cytokine levels as compared to adjacent tissues¹⁹⁶. This study also showed *in vitro*, that co-culture of *F. nucleatum* ATCC 51191 with monocytes (THP-1) induced CCL20 expression, as well as monocyte activation and migration. Chen *et al.* 2018 showed that co-culture of *F. nucleatum* strains with macrophages (RAW 264.7), induced a M2 macrophage phenotype also called TAM²⁰⁷ (see section 1.1.3).

Several bacterial receptors have been identified that mediate *F. nucleatum ssp.* ability to reduce the capacity of the host's immunity to eradicate the abnormal tumour cells. To date, three *F. nucleatum* cell surface molecules, Fap2, FadA and LPS have been implicated in colorectal tumour development. *Fusobacterium* autotransporter protein 2 (Fap2) has been shown to bind to Gal-GalNAc sugars which are highly expressed by colorectal cancer cells²⁰⁸. This protein has also been found to suppress NK and T cells by interacting with the cell surface T cell immunoreceptor with Ig and ITIM domains (TIGIT) protein²⁰⁹. *Fusobacterium* adhesin A (FadA) binds to the transmembrane protein E-cadherin, and further upregulate Annenix A1 which modulates Wnt/ β -catenin signaling in colonic epithelial cells²¹⁰. *F. nucleatum*-derived LPS has been shown to activate macrophage-like cells²¹¹, and particularly, the lipid A domain of *F. nucleatum*-derived LPS was shown to be specifically recognised by soluble CD14²¹². In addition, the interaction of *F. nucleatum* with TLR-4 caused an increased

proliferation of CRC cells²¹³ and induced autophagy²¹⁴ through the induction of miRNA expression. Recent studies also characterised the protein signature of *F. nucleatum*-derived OMVs, which were shown to contain the same virulence factors as the OM of bacteria²¹⁵. In addition, *F. nucleatum*-derived OMVs could modulate gut epithelial cell immunity by interacting with the TLR-2 receptor²¹⁶. However, the role of *F. nucleatum*-derived OMVs on the immune response of human immune cells has not been investigated.

At the structural level, *F. nucleatum* ssp. display strain-specific LPS on the cell surface. Vinogradov *et al.* conducted NMR and MS analysis of the LPS O-antigen of four *F. nucleatum* strains, ATCC 23726 (ssp. *nucleatum*)²¹⁷, MJR7757B31²¹⁸, ATCC 10953 (ssp. *polymorphum*)²¹⁹ and ATCC 25586 (ssp. *nucleatum*)²²⁰. The purified O-antigens displayed a trisaccharide repeat unit which composition varies across strains (Fig. 11A). *F. nucleatum* ATCC 23726 trisaccharide contains a novel nonulosonic sugar, 5,7-diamino-3,5,7,9-tetra-deoxynon-2-ulosonic acid (Non5Am7Ac), *F. nucleatum* ATCC 10953 trisaccharide contains sialic acid/ acetylneuraminic acid (Neu5Ac), and *F. nucleatum* ATCC 25586 trisaccharide contains two novel sugars, 2-acetamido-2,6-dideoxy-L-altrose (L-6dAltNAc) and the nonulosonic acid 5-acetimidoylamino-3,5,9-trideoxygluco-non-2-ulosonic acid (Non5Am) later named fusaminic acid (Fig. 11B).

A)

F. nucleatum ATCC 23726 : -4-β-Non5Am7Ac-4-β-D-GlcNAcyl3NFoAN-3-β-d-FucNAc4N-

F. nucleatum MJR7757B : -3-β-D-ManNAc4Lac-4-β-D-Glc6OAc-3-β-D-FucNAc4N-

F. nucleatum ATCC 10953 : -α-Neup5Ac-(2->4)-β-D-Galp-(1->3)-α-D-FucpNAc4NAc-

F. nucleatum ATCC 25586 : -4-β-Nonp5Am-4-α-L-6dAltpNAc3PCho-3-β-D-QuipNAc-

B)

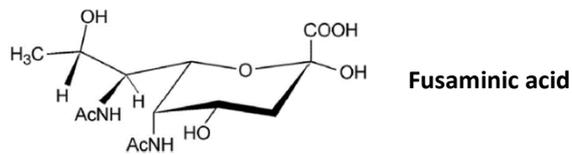


Figure 11 | LPS structure of *F. nucleatum* strains.

A) O-antigen structures of *F. nucleatum* ATCC 23726, 10953, 25586 and MJR7757B.

B) Fusaminic acid structure. (Taken from Vinogradov *et al.* 2016²²⁰).

Hypothesis and Aims

This project aims to test the hypothesis that in CRC, immunoregulation is mediated by the interactions of colorectal cancer-associated bacterial cell surface glycoconjugates with lectins expressed on innate immune cells. This interaction may promote the adaptation of a tumour-associated innate immune cell phenotype leading to tumour progression.

The specific objectives addressing this hypothesis are described below:

1. Analysis of CRC-associated bacteria binding to recombinant human lectins *in vitro*.

We analysed the binding of *B. fragilis* NCTC 9343, *F. nucleatum* ATCC 25586, ATCC 51191, ATCC 10953 strains, as well as *F. nucleatum* ssp.-derivatives (OMV and LPS) to human recombinant Dectin-2, Galectin-3 and to a panel of recombinant human Siglec proteins using flow cytometry, ELISA-based binding assay and biolayer interferometry.

2. Determine the effect of bacteria binding to lectins on myeloid cells

We used human monocyte derived dendritic cells (moDCs) and macrophages (moMφs) or cell lines to determine (1) *F. nucleatum* ssp.-immune cell association and internalisation and (2) the secreted cytokine profile as well as the expression of cell activation and tumour promoting cell surface markers. Further we silenced Siglec-7 in moDCs and we investigated the effect on *F. nucleatum* ssp. in the immune response.

3. Investigation of *Fusobacterium* spp. abundance and Siglec-7 expression in CRC

We sampled on-tumour and off-tumour specimens from resected tissues after colorectal cancer surgery and collected matched blood sample to test (1) *Fusobacterium* spp. abundance and Siglec-7 expression by qPCR, and (2) anti-*F. nucleatum* IgG and soluble Siglec-7 levels in serum, respectively.

CHAPTER 2

-

Materials and Methods

2.1 Microbiology

2.1.1 Strains and Materials

Fusobacterium nucleatum ssp. *polymorphum* ATCC 10953, *F. nucleatum* ssp. *animalis* ATCC 51191, *F. nucleatum* ssp. *nucleatum* ATCC 25586 and *Ruminococcus gnavus* ATCC 29149 and *Klebsiella pneumoniae* O1 (ATCC 43816) type strains from clinical isolates were obtained from American type culture collection (ATCC) in partnership with LGC standards Ltd.. *R. gnavus* E1 isolated from healthy individual's faeces²²¹. *Bacteroides fragilis* National collection of type cultures (NCTC) 9343 was kindly provided by Dr Regis Stentz (Quadram Institute Bioscience, Norwich). *Saccharomyces cerevisiae* isolated from human feces (#2966) was obtained from the National Collection of Yeast Culture (Quadram Institute Bioscience, Norwich, UK). All reagents were purchased from Sigma unless otherwise stated.

2.1.2 Bacterial and Yeast growth and preparation of PFA-fixed bacteria

F. nucleatum and *R. gnavus* strains were cultured in an anaerobic cabinet (Don Whitley) containing 85% N₂, 10% H₂ and 5% CO₂ gas mix at 37°C. *F. nucleatum* strains were cultured in tryptic soy broth medium (Becton Dickinson) supplemented with 5 µg/ml hemin and 1 µg/ml menadione. *R. gnavus* E1 and ATCC 29149 strains were cultured in brain heart infusion (BHI) medium, as previously described²²¹. *S. cerevisiae* was cultured in yeast medium as reported previously²²². *K. pneumoniae* O1 was cultured in Luria-Bertani (LB) medium. For subsequent *in vitro* assays, bacteria were centrifuged at 15,000 x g for 5 min, and the cells were fixed with 4% paraformaldehyde (PFA) (Electron Microscopy Sciences/CN Technical Services Ltd) for 45 min at room temperature (RT), in the dark, followed by two washes in phosphate-buffered Saline (PBS) (Lonza).

2.1.3 Microbial density and size quantification

Yeast density was quantified using an hemacytometer at an Optical Density at 600 nm (OD_{600nm}) of 0.7, monitored by spectrophotometry. For *F. nucleatum* ssp. and *R. gnavus* strains the bacteria density was quantified based on OD_{600nm} of 1 corresponding to 10⁹ cells/ml and for *K. pneumoniae* O1 based on OD_{600nm} of 1 corresponding to 5 x 10⁸ cells/ml. Alternative methods for cell quantification included the use of the Bacteria Counting Kit (Thermo Fisher Scientific) following the manufacturer's instructions, the use of flow cytometry (BD LSRFortessa), or imaging flow cytometry (Amnis ImageStream^x Mk II). For bacteria density and cell size quantification by ImageStream^x Mk II, PFA-fixed bacteria (section 2.1.2) at 10,000 events were collected and processed by IDEAS software. Bacteria density were found by selecting "objects/ ml" in the bright field channel (M04) of the Aspect Ratio_M04 versus Area_M04 dot plot. Bacteria cell size was quantified by first selecting the low intensity for side scatter laser (Channel 6) population (exclusion of control beads), and then applying the "length" feature in bright field.

Another method used was based on DNA quantification. Here, the genomic DNA of the PFA-fixed *F. nucleatum* strains was extracted using the GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific) following the manufacturer's instructions and quantified by Nanodrop (Thermo Fisher Scientific).

2.1.4 Outer membrane vesicle (OMV) purification from *F. nucleatum* ssp.

F. nucleatum derived OMVs were collected from cell culture supernatant as described previously by Liu *et al.* 2019²¹⁵ with some modifications. Briefly, *F. nucleatum* ssp. cells were cultured until reaching an OD_{600nm} of approximately 1.2 (for *F. nucleatum* ATCC 25586 and ATCC 10953) or 0.7 (for *F. nucleatum* ATCC 51191). Cells were centrifuged at 8,500 x g for 15 min at 4°C. Supernatants were collected and vacuum filtered through a 0.22 µm membrane. The filtered supernatant was concentrated by spin-filtration at 5,500 x g and 4°C using a 100K molecular weight cut-off filter unit (MWCO) (Sartorius). OMVs were recovered from the filter using sterile phosphate buffered saline (PBS) and further purified by density gradient ultra-centrifugation. For the gradient, Optiprep medium (60% w/v) was diluted in 0.85% w/v NaCl and 10 mM Tricine-NaOH pH 7.4 solution to

make up 35%, 30%, 25% and 20% densities. In addition, OMVs in PBS were used to make a 40% Optiprep solution. At the bottom of a 13.2 ml Ultra-clear tube (Beckman Coulter) were inserted the 2 ml fraction of 40% density followed by 2 ml of the fractions of decreasing density, forming 5 layers in total. The tube was ultra-centrifuged at 135,000 x g for 16 h at 4°C with minimum acceleration and deceleration using the SW41 Ti rotor (Beckman Coulter). From the top to the bottom fractions were collected and analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (see section 2.2.6). The OMV-containing fractions (1 ml) were diluted 10 times (vol/vol) with sterile PBS and ultra-centrifuged at 200,500 x g for 2 h at 4°C using Type 45 Ti rotor (Beckman Coulter). OMVs were resuspended in sterile PBS and filtered through a 0.22 µm membrane. Concentrated in PBS OMVs were stored at -80°C.

2.1.5 Nanoparticle Tracking Analysis (NTA)

Purified OMVs were quantified and measured in terms of particle size using the NanoSight LM12 (Malvern Panalytical). OMVs were diluted 100 times in PBS and the diluted samples were loaded into a syringe. The syringe was adjusted to the instrument's chamber and the sample was slowly released. The particle size of each OMV sample corresponded to the mean of triplicates. Instrument settings used were camera shutter 1035, camera gain 680, capture duration 60 sec.

2.1.6 Transmission electron microscopy (TEM)

TEM analysis of *F. nucleatum*-derived OMVs was performed by Dr Kathryn Gotts as follows. A 10 µl drop of OMV suspension was placed onto a formvar/carbon coated copper TEM grid, left for 1 min, and then the excess liquid was wiped off. Immediately after, a 10 µl drop of 1.5% uranyl acetate solution was placed onto the grid and left for 1 min. The excess stain was removed using filter papers and the grids were left to dry completely before imaging. The grids were analysed in a FEI Talos F200C transmission electron microscope at 200kV.

2.1.7 LPS extraction from *F. nucleatum* ssp.

The extraction of *F. nucleatum*-derived lipopolysaccharide (LPS) or its derivatives (Lipid A, O-antigen) was performed by Profs Cristina De Castro's and Alba Silipo's groups (University of Naples) as follows.

Bacteria were harvested by centrifugation, lyophilised, and extracted by hot phenol/water method²²³. Phases were dialysed against distilled water; phenol removal; lyophilised, and analysed by SDS-PAGE following silver nitrate staining²²⁴. LPS purified from *F. nucleatum* ATCC 10953 or 51191 strains was detected in the water phase, and the LPS derived from *F. nucleatum* ATCC 25586 was detected in the phenol phase. LPS was further purified by enzymatic digestion (DNAse, RNAse and proteinase K)²²³, the suspension centrifuged at 6,000 rpm for 30 min at 4°C followed by ultracentrifugation at 30,000 rpm for 4 h at 4°C. For O-antigen (OPS) and lipid A domain separation, LPS was subjected to a mild acid hydrolysis using 1% acetic acid (100°C, 2-3 h). The OPS domain was further partially depolymerised and purified by gel filtration chromatography.

2.1.8 Semi-quantitative analysis of LPS in OMVs

The semi-quantitative LPS analysis in *F. nucleatum*-derived OMVs was performed by Prof. Cristina De Castro's group (University of Naples) as follows.

LPS in *F. nucleatum*-derived OMVs was evaluated by gas chromatography-mass spectrometry (GC-MS) analysis of the fatty acids content. C14:0 (myristic acid) was considered as the reporter group for LPS²²⁵, while C16:0 (palmitic acid) and C18:0 (stearic acid) were the reporters for phospholipids. *F. nucleatum*-derived OMVs at 1 mg were treated with HCl/MeOH (1 ml, 1.25 M, 80°C, 16 h) and lipids, were extracted with hexane²²³. The amount of each fatty acid (C14:0, C16:0 and C18:0) was determined by correcting the areas of the corresponding peaks by using standard solutions and C16:0 as internal standard. After hexane extraction, the methanol was analysed for the presence of 3-deoxy-2-keto-D-manno-octulosonic acid (KDO) and L-glycero-D-manno-heptose sugars, both markers of LPS. Fatty acids or the monosaccharide constituents were identified by comparing the retention time and the fragmentation pattern of each peak to a standard. The analysis was carried out using a GC-MS Agilent Technologies 7820A (Santa Clara, CA, USA) equipped with a mass selective detector 5977B

and a HP-5ms capillary column Agilent, Italy (30 m x 0.25 mm i.d., 0.25 μm as film thickness, flow rate 1 ml/min, and He as carrier gas). Electron impact mass spectra were recorded with ionization energy of 70 eV and an ionising current of 0.2 mA. The temperature program used was: 150 °C for 5 min, 150-300°C at 10°C/min, 300°C for 12 min.

2.2 Protein expression

2.2.1 Materials

The plasmid encoding the carbohydrate recognition domain (CRD) of human Dectin-2 with a C-terminus biotinylation sequence (pT5T-Dectin-2 CRD-biotin-tag), and the plasmid expressing the biotin ligase BirA were obtained from Prof. Kurt Drickamer (Imperial College London). The transfection host strain used was *E. coli* BL21DE(3). *E. coli* BL21DE(3) harbouring the human Galectin-3-his-tag plasmid was obtained in house as previously reported²²⁶. CHO-expressing Siglec-Fc cell lines were a kind gift from Prof. Paul Crocker (University of Dundee). All other reagents were purchased from Sigma unless otherwise stated.

2.2.2 Co-transformation of *E. coli* BL21DE(3) with Dectin-2-CRD-biotin-tag and BirA plasmids

Competent *E. coli* BL21DE(3) cells were co-transformed with the Dectin-2-CRD-biotin-tag and BirA plasmids carrying genes for ampicillin and chloramphenicol antibiotic resistance, respectively, following the heat shock protocol²²⁷. Briefly, competent cells were thawed on ice and the DNA plasmids (1 μ l each) was added. The mixture (27 μ l) was incubated on ice for 30 min followed by a heat shock at 42°C for 30 sec. Super optimal broth with catabolite suppression (SOC) (NEB) (500 μ l) was added to the cells and incubated at 37°C for 1 h with vigorously shaking at 250 rpm. The suspension was spread on LB plates containing 50 μ g/ml carbenicillin and 20 μ g/ml chloramphenicol for colony formation. Colonies were selected and tested for the presence of the Dectin-2-CRD and BirA plasmids by PCR using specific primers (5'-GAC AGC AAA TGG GT-3' and 5'-CTT ACT AGT CGG TA-3') and (5'- TTA AGC TGA CGG CAT -3' and 5'- ACT GGG TGA GCA CGA -3'), respectively, using TaqDNA polymerase (NEB) for the DNA amplification. The PCR products were analysed by DNA agarose gel (1%). Glycerol stocks for the positive colonies were stored in -80°C.

2.2.3 Expression and purification of recombinant Dectin-2-CRD and tetramer formation

For Dectin-2-CRD expression, *E. coli* BL21DE(3) cells containing BirA and Dectin-2-CRD-biotin-tag plasmids were cultured in 6 l of LB containing 50 µg/ml carbenicillin and 20 µg/ml chloramphenicol to maintain selection and 12.5 µg/ml biotin to achieve biotinylation. Cells were harvested by centrifugation at 4,000 x g, at 4°C for 15 min and washed with 10 mM Tris-HCl pH 7.8 buffer. Pellets were suspended in 200 ml of cold 10 mM Tris-HCl pH 7.8 buffer and sonicated in 4 x 1 min bursts. After sonication, the solution was centrifuged at 10,000 x g at 4°C for 15 min and the pellets were dissolved in 6 M Guanidine-HCl, 100 mM Tris-HCl pH 7.0 with 14.3 M β-mercaptoethanol, followed by incubation at 4°C for 30 min. The solution was then centrifuged at 28,000 x g at 4°C for 30 min and 1% v/v of Triton 100-X was added. The protein sample was then dialysed against a 0.5 M NaCl solution in a dialysis membrane (Visking, MWCO 12-14 KDa). The buffer was changed after 3 h and the dialysis continued overnight (o/n). The next day, the buffer was changed again and after 3 h, the dialysed solution was centrifuged at 15,000 x g at 4°C for 30 min. The supernatant was passed through a 0.2 µm membrane.

The recombinant Dectin-2-CRD protein was purified by affinity column chromatography using a D-mannose-sepharose column (column volume = 10 ml). The column was first equilibrated with 20 ml of loading buffer. The solution containing Dectin-2-CRD-biotin-tag was loaded onto the column, and the column washed with low NaCl loading buffer (Table 2). Bound Dectin-2 protein was then eluted with the elution buffer (Table 2) and fractions (1 ml) were collected and analysed by SDS-PAGE. Fractions containing Dectin-2-CRD-biotin-tag were collected and dialysed against H₂O.

Table 2 | Mannose-sepharose column buffer preparation

SOLUTION	Tris-HCl	NaCl	CaCl ₂	EDTA
Loading Buffer	25 mM	0.5	25 mM	-
Low NaCl loading buffer		150 mM		
Eluting buffer			-	2.5 mM

For the tetramer complex preparation, purified Dectin-2-CRD was incubated o/n at 4°C with streptavidin (SA)-Alexa488 (ThermoFischer) in a 2- to 2.5-fold excess of Dectin-2-CRD over SA-Alexa488 as determined by weight, and 50 mM of ethylenediaminetetraacetic acid (EDTA) was added to the suspension. The tetramer complex was purified by Fast Protein Liquid Chromatography (FPLC) gel filtration using a HiLoad 16/600 Superdex 200 column (GE Healthcare) at a flow rate of 1 ml/min, with 100 mM NaCl, 10 mM Tris-HCl (pH 7.8) and 2.5 mM EDTA buffer and 2 ml volume fractions were collected.

Biotinylation of the recombinant Dectin-2-CRD was determined using a gel shift assay. Briefly, 5 µg of recombinant Dectin-2-CRD was incubated with SA (4.41 µg) for 1 h at 4°C (molar ratio of 4:1) and the protein sample subjected to SDS-PAGE analysis in non-reducing conditions (see 3.2.6).

2.2.4 Expression and purification of recombinant Galectin-3

E. coli BL21DE(3) cells expressing Galectin-3-his-tag were cultured in 1 l of LB containing 50 µg/ml carbenicillin. Cells were harvested by centrifugation at 6,000 × *g* at 4°C for 15 min and following a wash with PBS the pellets kept at -80°C until use. The pellets were suspended in cold PBS (35 ml) and sonicated 10 times for 15 sec each with 30 sec incubation on ice in between each sonication step. After sonication, the solution was centrifuged at 8,000 × *g* at 4°C for 15 min, and the process was repeated 3 times. After each centrifugation, the supernatant was collected and passed through a 0.2 µm filter. Galectin-3 was first purified by immobilized metal affinity chromatography (IMAC) using a Nickel-immobilized column using the buffers listed in Table 3. The eluate was then subjected to a second affinity chromatography using a lactose-agarose column using PBS and lactose buffer (150 mM lactose in PBS) as a washing buffer and elution buffer, respectively, and fractions (1 ml) were collected.

IMAC SOLUTION	Tris-HCl	NaCl	Imidazole
Binding Buffer	20mM	500mM	5mM
Wash Buffer 1			20mM
Wash Buffer 2			40mM
Wash buffer 3			60mM
Elution Buffer			1M

2.2.5 Expression and purification of recombinant Siglec-Fc proteins

CHO cells transfected with Glutamine Synthetase (GS) transcription cassette and cDNA encoding Siglec-Fc recombinant protein vector were cultured in Glasgow Modified Essential Medium (GMEM) without L-glutamine media supplemented with 10% Fetal Bovine Serum (FBS) (Thermo Scientific Gibco), 100 U/ml penicillin and 100 µg/ml streptomycin (Lonza) and 50X GS supplements until 80-90% confluence was achieved. Adherent cells were washed twice with PBS and protein expression induced by culturing the cells with GMEM without L-glutamine media, supplemented with 200X FetalClone II (Thermo Fisher Scientific), 100 U/ml penicillin and 100 µg/ml streptomycin, 50X GS supplements and 100 µg/ml, and L-Methionine sulfoximine (MSX). After 4 days, the supernatant was harvested and centrifuged at 300 x g for 5 min at RT.

The Siglec-Fc purification was carried out using a gravity-flow column (BIO-RAD) prepared by addition of protein A-Sepharose (300 µl) and washed with PBS (5 ml). Then the harvested CHO supernatant was loaded onto the column and the column washed with PBS (10 ml). To elute Siglec-Fc, a solution of 0.1 M glycine pH 3 was added to the column and fractions (500 µl) were collected in 1.5 ml Eppendorf tubes containing 1 M Tris pH 8 (for neutralisation).

2.2.6 SDS-PAGE

Protein samples (5 μ l) were mixed with 4X Sample Buffer (NuPAGE LDS or Laemmli) and 10X Reducing Agent (NuPAGE or β -mercaptoethanol). For non-reduced conditions, the reducing agent was replaced by water. The mixture (reduced or non-reduced) was heated for 10 min at 70°C and loaded onto a 4-12% polyacrylamide protein gels (NuPAGE or Mini-PROTEAN-TGX) and run for 30 min at 200 V constant, following the manufacturer's instructions. The SDS protein gel was stained with Coomassie-based staining solution InstantBlue™ (Expedeon). The reagents were from Thermo Fisher Scientific or BIO-RAD unless otherwise stated.

2.2.7 Protein quantification

Fractions from each protein purification were concentrated using a Vivaspin 2 or Millipore concentrator of 10 KDa MWCO and the concentrated proteins were quantified by Nanodrop (Thermo Fisher Scientific).

2.3 Binding assays

2.3.1 Materials

U937 (WT, and Siglec-7^{+/+} or Siglec-7^{-/-} clones after CRISPR-Cas9 genome editing) and Chinese hamster ovary (CHO) (WT and Siglec-7) cell lines were a kind gift from Prof. Matthew Macauley (University of Alberta) and Prof. Paul Crocker (University of Dundee), respectively. *E. coli* O111:B4 LPS used as a control. Red blood cells (RBCs) was obtained from the whole blood of haemochromatosis patients undergoing a therapeutic venesection at the Norfolk and Norwich University Hospital (Norwich, UK). The blood collection was approved by the Faculty of Medicine and Health Sciences Research Ethics Committee REC reference number 2013/2014 -14HT (University of East Anglia). All other reagents were purchased from Sigma unless otherwise stated.

2.3.2 Red Blood Cells binding assays

The functionality of Siglec-Fcs was confirmed using RBCs isolated from whole blood as described in section 2.4.2. RBCs (2×10^7 cells) were incubated with pre-complexed Siglec-Fc (except Siglec-9-Fc) and α -Fc(-PE (phycoerythrin) or -FITC) for 30 min at 4°C or with Siglec-9-Fc immobilised on beads and analysed by flow cytometry. For the pre-complex preparation, 4 μ g/ml of Siglec-Fc protein were incubated with 4 μ g/ml α -Fc(-PE or -FITC) for 30 min at 4°C. For the Siglec-9-Fc-bead preparation, 45 μ l of μ MACS Protein A MicroBeads (MACS) were blocked with 1% BSA in PBS for 1 h at 4°C followed by incubation with Siglec-9-Fc at 2 μ g/ml for 1 h at 4°C. The unbound Siglec-9-Fc were washed off by μ MACS separator (MACS) using 2% BSA in PBS and recovered using the same buffer. For RBC de-sialylation, 10^7 cells were incubated with *Vibrio cholerae* neuraminidase (1:50) for 1 h at 37°C in the neuraminidase buffer. Sialidase-treated RBCs were washed twice with 2% BSA in PBS.

2.3.3 Microbial flow cytometry binding assays

For the binding assays, bacteria or yeast (10^7 cells) were incubated for 1 h at 37°C with the recombinant lectins: 1) tetramer Dectin-2-CRD-Alexa488 (2 µg/ml) in FACS buffer (Hanks' Balanced Salt Solution (HBSS) (Lonza) containing 25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) and 0.1% BSA in the presence of 2 mM CaCl₂ or 10 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), 2) Galectin-3 (10 µg/ml) in PBS or 3) Siglec-Fcs or human-Fc (control) (4 µg/ml) (mouse Siglec-1, -2 and Siglec-E and the human Siglec-3, Siglec-5, Siglec-7, Siglec-9 and -10) in PBS, or 4) *Sambucus nigra* I lectin (SNA)-FITC at 4 µg/ml. Following centrifugation at 16,000 x *g* at 4°C for 5 min, the cells were harvested and washed with their respective incubation buffers. For Siglec-Fcs or control, cells were incubated with the pre-complexed Siglec-Fc and α-Fc-PE (0.2 µg/ml) for 1 h at 4°C. For Galectin-3 binding, cells were incubated with mouse anti-His-tag antibody (1 µg/ml) (Millipore) for 1 h at 4°C, centrifuged at 16,000 x *g* at 4°C for 5 min, then washed and finally incubated with the secondary antibody goat anti-mouse-IgG- APC (allophycocyanin) antibody (0.3 µg/ml) (BioLegend) for 1 h at 4°C. For the Siglec-7-Fc inhibition assays, Siglec-7-Fc and α-Fc-PE Ab pre-complex was first incubated with GD3 (disialoganglioside with three glycosyl groups) at 50 µg/ml for 30 min at 4°C. For de-sialylation, bacteria (10^7 cells) were treated with 20 U of sialidase α2-3,6,8,9 Neuraminidase A in its 1X GlycoBuffer I (NEB) or control treated in 1X GlycoBuffer I alone, o/n at 37°C.

All binding assays were analysed by Fortessa (BD Biosciences) flow cytometer. The setting parameters were collected on a log scale and the threshold of FSC parameter was set to 1000. The collected data were processed by FlowJo. The mean fluorescence intensity (MFI) is referred to the geometric mean fluorescence intensity value. The relative MFI (R-MFI) is referred to the value obtained after subtraction of the control MFI.

2.3.4 ELISA-based binding assays

For binding assays, bacteria (10^7 cells) or bacteria-derived molecules (LPS or OMVs) in either PBS or 100 mM NaHCO₃ and 33 mM Na₂CO₃ in H₂O (pH 9) (coating buffer) solution were coated on a 96-well plate, o/n at 4°C. Following 3 times washing with 0.05% tween in PBS (washing buffer), the plate was incubated with 1% bovine serum albumin (BSA) for 1 h at RT. Following 3 times washing with washing buffer, the plate was incubated with pre-complexed Siglec-Fc (see section 2.3.3) and α -Fc-HRP (horseradish peroxidase) for 2 h at RT. For the pre-complex preparation, Siglec-Fc protein (4 μ g/ml) was incubated with 50,000X α -Fc-HRP (Abcam) for 1 h at RT. Following 3 times washing with washing buffer, the plate was incubated with 3,3',5,5'-Tetramethylbenzidine (TMB) (BioLegend) until colour change. Colour development was stopped by the addition of 2 N H₂SO₄ and the absorbance was measured at 450 nm with reference at 570 nm using a FLUOStar (BMG Labtech).

2.3.5 Biolayer Interferometry (BLI) binding assays

For binding assays, Siglec-Fc proteins (5 μ g/ml in PBS) were loaded onto anti-human Fc biosensors (AHC) (Sartorius) for 600 sec. Soluble bacteria-derived OMVs at 10^{10} particles/ml or GD3 (control) at 10 μ g/ml in PBS were incubated with the immobilised Siglec-Fc for 1000 sec (association step) followed by incubation with PBS for 1000 sec (dissociation step). Assays were performed on the Octet Red96 (ForteBio) instrument at 30°C with shaking at 1000 RPM. Data were processed by Data Analysis 7 software (ForteBio).

2.3.6 Binding of *F. nucleatum* to mammalian cells

U937 monocytic cell lines (WT or Siglec-7^{-/-} or Siglec-7^{+/+}) or CHO (WT and Siglec-7) were cultured in RPMI 1640 medium (Lonza) supplemented with 25 mM HEPES, 10% Heat inactivated (HI) FBS (Thermo Scientific Gibco), 55 μ M of 2-mercaptoethanol, 100 U/ml penicillin and 100 μ g/ml streptomycin (Lonza), 2 mM of L-glutamine (Lonza) and 1 mM Sodium Pyruvate (Lonza) or Hams F-12 medium (ThermoFisher) supplemented with 100 U/ml penicillin and 100 μ g/ml streptomycin and 10% FBS, respectively. Mammalian cells (10^7 cells) were stained with 100,000X of Cell Trace Violet (CTV) (ThermoFisher) for 15 min at RT in the dark. *F. nucleatum* (10^8 cells) were fluorescently-labelled with 10 μ g/ml

of fluorescein isothiocyanate (FITC) for 30 min at RT in the dark. Both mammalian cells and bacteria were washed and resuspended in the mammalian cell medium described above. U937 or CHO-CTV (5×10^5 cells/well in 100 μ l) were then incubated with *F. nucleatum*-FITC (5×10^6 cells/well in 200 μ l) in a U-shape 96-well for 1 h at 4°C. Following centrifugation at 500 x g for 3 min, the pelleted cells were resuspended in FACS buffer and analysed by Fortessa flow cytometry and data were processed by FlowJo. The mean fluorescence intensity (MFI) is referred to the geometric mean fluorescence intensity value.

2.3.7 STD NMR analysis

The STD NMR analysis of Siglec-7-Fc and *F. nucleatum*-derived LPS interactions was performed by Prof. Alba Silipo's group (University of Naples) as follows.

The partial depolymerised OPS derived from *F. nucleatum* ssp. (see section 2.1.7) was prepared in deuterated PBS buffer (20 mM PBS, NaCl 150mM, pH 7.4), using Siglec-7-Fc-ligand ratios varying from 1: 20 to 1: 80 with 15 μ M of Siglec-7-Fc protein and analysed by Bruker 600 MHz AVANCE NEO equipped with a cryo probe and the data processed using the TOPSPIN 4.1.0 software. The Siglec-7-Fc resonances were saturated by applying 40 Gauss pulses with a length of 50 ms, setting the on-resonance pulse at aromatic region (7.5/6-5 ppm) and the off-resonance pulse frequency at 100 ppm. Very low residual signals were observed in some spectra for the ligands in the free state which were considered during data processing. To suppress the water signal, an excitation sculpting with gradient pulses (esgp) was applied. For reducing the NMR signals of Siglec-7-Fc, a spin-lock filter (20 ms) was used.

2.4 Immunological assays

2.4.1. *Materials*

Human peripheral blood was obtained from haemochromatosis patients undergoing a therapeutic venesection at the Norfolk and Norwich University Hospital (Norwich, UK). Blood collection in this study was approved by the Faculty of Medicine and Health Sciences Research Ethics Committee REC reference number 2013/2014 -14HT (University of East Anglia). U937 (WT, Siglec-7^{+/+}, and Siglec-7^{-/-}) cell lines were a kind gift from Prof. Matthew Macauley (University of Alberta).

2.4.2 *Generation of moDCs and moMφs from human blood*

Human blood monocytes were isolated following centrifugation of whole blood sample using a Ficoll-Paque gradient (GE Healthcare) for 40 min, at 400 x g and RT with minimum acceleration and deceleration speed. The peripheral blood mononuclear cells (PBMCs), located at the interface of the serum in the top fraction and blood cells and granulocytes at the bottom fraction, were collected. PBMCs were diluted 3-fold with HBSS supplemented with 3% FBS and 10 mM EDTA and the solution centrifuged at 270 x g, RT for 7 min. Monocytes (CD14⁺ cells) were isolated from PBMCs by magnetic bead separation, using CD14 positive selection microbeads (StemCell technologies or ThermoFisher) according to the manufacturer's instructions.

Monocyte-derived dendritic cells (moDCs) or macrophages (moMφs) were generated by culturing 10⁶ cells/ml of fresh isolated monocytes (CD14⁺ cells) in Mercedes medium (RPMI 1640 medium (Lonza) supplemented with 25 mM HEPES, 10% HI FBS (Thermo Scientific Gibco), 55 μM 2-mercaptoethanol, 100 U/ml penicillin and 100 μg/ml streptomycin (Lonza), 2 mM glutamine (Lonza), 1 mM non-essential amino acids (Lonza) and 1 mM Sodium Pyruvate (Lonza)) supplemented with cytokines, granulocyte-macrophage colony-stimulating factor (GM-CSF) and IL-4 (PeproTech) (25 ng/ml) or macrophage colony-stimulating factor (M-CSF) (PeproTech) (25 ng/ml) for differentiation of monocytes to moDCs or moMφs, respectively. The cells were incubated for 7 days at 37°C in a 5% CO₂ incubator (Sanyo), with change of medium and cytokine supplementation on day 3.

2.4.3 Siglec-7 RNA silencing in immune cells

MoDCs or moMφs were transfected with one, two or three pre-designed small interfering RNA (siRNA) *Silencer Select* SIGLEC7 probes or with the scramble siRNA (Invitrogen) with reverse transfection, as previously described²²⁸. Briefly, 3×10^5 moDCs or moMφs in a 24-well plate were incubated with a complex of two or three probes (Table 4) to a final 200 nM concentration or with the negative control (scramble) and 1% HiPerFect transfectant (Qiagen) in warm RPMI 1640 (non-supplemented) medium for two, three or four days (Table 4).

To determine Siglec-7 expression on moDCs and moMφs, 10^5 cells were incubated with human Fc-block (BioLegend) for 15 min at RT followed by an incubation with mouse anti-human Siglec-7 APC conjugated at 1:50 dilution, for 1 h at 4°C. To detect cell death, the cells were incubated with propidium iodide (PI) at 1 µg/ml. The cells were then analysed by Fortessa flow cytometry.

Table 4 | Pre-designed probes used for Siglec-7 siRNA and silencing conditions

Probe name	siRNA ID #	Nucleotide Sequence (5'->3')	2-Day silencing	3-Day silencing	3-Day silencing	4-Day silencing
Probe-1	s25729	GGAGGAGUCUGACCCUGUAtt	+	+	+	+
Probe-2	s25730	GACAGAAGAGUAACCGGAAtt	+	+	+	+
Probe-3	s25731	GGAAUGAUUAAGCUGGAAtt			+	
Scramble	Negative control 1 / non-target	Information not available	+	+	+	+

*2.4.4 Stimulation of human primary myeloid cells with *F. nucleatum* ssp.*

For moDCs and moMφs, 10^6 cells/ml were cultured in a 96-well plate in the Mercedes medium as described above and stimulated with PFA-fixed *F. nucleatum* ssp. at MOI of 50 or 5 at 37°C for 2 h, 6 h or 18 h, or with *F. nucleatum* ssp.-derived LPS or control *E. coli* O111:B4 at 10 or 1 μg/ml, or *F. nucleatum* ssp.-derived OMVs at 5×10^7 particles/ml for 18 h. On the next day, cells were centrifuged at 500 x g for 3 min and the supernatant collected and stored at -80°C for later use (cytokine analysis) while the cell pellet was used immediately for cell surface marker expression analysis.

*2.4.5 Stimulation of U937-PMA cells with *F. nucleatum* ssp.*

For U937 (WT or Siglec-7^{-/-} or Siglec-7^{+/+}) monocytic cell lines, 5×10^5 cells/ml were differentiated using phorbol 12-myristate 13-acetate (PMA) at 100 ng/ml for 36 h in the medium described in section 2.3.6. U937-PMA (10^5 cells) were then stimulated with PFA-fixed *F. nucleatum* ssp. (5×10^5 cells) at 37°C for 18 h. On the next day, cells were centrifuged at 500 x g for 3 min and the supernatant collected and stored at -80°C for later use (cytokine analysis).

2.4.6 Cell surface marker analysis by flow cytometry

MoDC and moMφ cell pellets were first incubated with human Fc block for 15 min at RT and then incubated with anti-human Abs for 30 min at 4°C. Siglec-7-APC at 1:100 Programmed Death-Ligand 1 (PD-L1)-PE at 1:50 dilution, CD80-PE at 1:100 dilution, CD86-Alexa488 at 1:200 dilution, CD206-APC/Cy7 at 1:200 dilution, CD163-PE at 1:100 dilution, isotype controls mouse IgG1-PE, κ at 1:100 or IgG2b-PE, κ at 1:25 dilution and PI or 4',6-diamidino-2-phenylindole (DAPI) at 1 μg/ml for dead cell staining. The cells were then analysed by flow cytometry. All antibodies were purchased from BioLegend.

2.4.7 Internalisation assays by Imaging Flow Cytometry

Mammalian cells (5×10^5), primary or cell lines, were co-cultured with (5×10^6) FITC-labelled *F. nucleatum* ssp. cells in a U-shape 96-well for 4 h at 37°C. Cells were then washed with FACS buffer and the plate centrifuged at 500 x g for 3 min. The cell pellets were then resuspended in FACS buffer and analysed by imaging flow cytometry ImageStream^x Mk II (Amnis). A total number of 5,000 positive to FITC cells were collecting using INSPIRE (Amnis) software. For the data processing, the percentage of internalised bacteria were identified using the “internalization” wizard with erode mask function at 7 number of pixels, using IDEAS 6.2 software.

2.4.8 Microscopy

MoDCs and moMφs were differentiated onto coverslips (VWR) in 24-well plates from human purified CD14+ cells in Mercedes medium supplemented with cytokines, as described above (section 2.4.2). After 7 days of culture, moDCs or moMφs were washed with PBS and incubated with 2.5×10^7 PFA-fixed *F. nucleatum* ssp. for 5 h. MoDCs or moMφs were washed and fixed with 4% PFA followed by a 45 min incubation at RT in the dark. The cells were then washed and incubated in fluorescence *in situ* hybridisation (FISH) buffer (20 mM Tris-HCl, 0.9 M NaCl, 0.1% w/v SDS) at 50°C, o/n, in the dark with 10 ng/μl of the FISH-probe (5' → 3' GCTGCCTCCCGTAGGAGT) fluorescently-labelled at the 5' end with Alexa555 (Eurofins). The next day, cells were incubated for 20 min with the FISH washing buffer (20 mM Tris-HCl, 0.9 M NaCl) at 50°C and then washed 3 times with PBS. The cells were then incubated with DAPI (300 nM) for 5 min in the dark, washed 3 times with PBS, mounted in fluoromount (eBioscience) and imaged using Zeiss (Carl Zeiss Ltd.) fluorescence microscope followed by image processing using Fiji processing package.

2.4.9 Cytokine analysis

Cell culture supernatant was monitored for IL-1 β , MMP3, MMP9, IP10, IL-6, IL-10, IL-13, IFN γ , GM-CSF, TNF α , M-CSF, VEGF-A, and IL-8 production by either ELISA (BioLegend) or ProcartaPlex-12 plex (ThermoFisher) analysis according to the manufacturers' instructions. Multi-plex readings were performed on Luminex® 100/200™.

2.5 Mammalian tissue processing

2.5.1 Materials and Ethics

Tissue collection in this study was approved by the Faculty of Medicine and Health Sciences Research Ethics Committee REC reference number 2012/2013-78 (University of East Anglia). Colorectal cancer (CRC) cases with prior antibiotic use or with inflammatory bowel disease (IBD) were excluded. Human specimens (on-tumour or off-tumour site) and whole blood from CRC patients were obtained from Norfolk and Norwich University Hospital (NNUH). Colons from C57BL/6 WT or Cre-lox tamoxifen inducible APC/KRAS mice (without being tamoxifen induced) were obtained from Prof. Alastair Watson at the disease model unit (University of East Anglia, Norwich, UK).

2.5.2 Colonic lamina propria leukocyte isolation

Mammalian colonic tissues were transferred into a tube containing PBS supplemented with 2% FBS, and vortexed to ensure removal of faeces and mucus. For epithelial cell removal, the colon was first cut into pieces and transferred into a tube containing pre-warmed PBS supplemented with 5 mM EDTA and 10 mM HEPES (Lonza) and incubated in a shaker at 37°C at 240 rpm for 15 min, the medium was replaced with fresh pre-warmed supplemented PBS and the incubation repeated once more. For lamina propria cell release, the tissue first was transferred into cold PBS supplemented with 2% FBS and vortexed (for EDTA removal), then cut into smaller pieces and transferred into a tube containing RPMI 1640 supplemented with 2 mM CaCl₂, 0.25 mg/ml DNase I and Collagenase. The suspension was then incubated in a shaker at 240 rpm, at 37°C for 20 min. The remaining tissue was disaggregated with a needle and centrifuged at 787 x g at 4°C for 10 min. The pellet was suspended with 10 ml of RPMI 1640 and layered onto 30% percoll (Thermo Fisher Scientific) (which underlays 100% percoll). The gradient was centrifuged at 670 x g at RT for 30 min with break set to "0" and acceleration to "3". The immune cells were isolated from the interlayer between 100% and 30% percoll, washed with cold RPMI 1640 supplemented with 25 mM HEPES, 2% FBS, 100 U/ml penicillin and 100 µg/ml streptomycin, 2 mM glutamine, 1 mM non-essential amino acids and centrifuged

at 787 x g at 4°C for 10 min, resuspended in FACS buffer and used immediately for subsequent analyses.

2.5.3 Colonic lamina propria leukocyte cell staining by flow cytometry

To monitor Siglec-7 expression in fresh isolated colonic lamina propria leukocytes (from on-tumour or off-tumour), cells were first incubated with human Fc block for 15 min at RT and then subjected to multi-colour staining for 30 min at 4°C. The antibodies used were Siglec-7-APC, CD45-BUV395, CD11b-BV711, CD103-FITC at 1:100 dilution.

To identify Siglec-E expression in mice, colonic lamina propria immune cells were first incubated with mouse Fc block for 15 min at RT and then incubated with anti-mouse Abs (multicolour staining) for 30 min at 4°C. Siglec-E-PE at 1:50 dilution or isotype control rat IgG2a, κ -PE, κ at 1:50, IA^b-APC at 1:100 dilution, CD11b-PeCy7 at 1:500 dilution, F4/80-Alexa700 at 1:200 dilution, CD103-FITC at 1:200 dilution.

All antibodies were purchased from BioLegend. Cells were analysed by Fortessa flow cytometer and data processing was performed using the FlowJo software.

2.5.4 DNA and RNA extraction from human colonic tissues and cDNA synthesis

DNA or RNA was extracted from human colonic specimens (on-tumour or off-tumour site) (20-50 mg). For DNA extraction, QIAamp DNA Mini (Qiagen) or Monarch Genomic DNA Purification (NEB) kits were used following the manufacturers' instructions. For RNA extraction, the RNA RNeasy Lipid Mini (Qiagen) kit was used following the manufacturer's instructions, and RNA quality was tested by TapeStation (Agilent). For the cDNA production, the extracted RNA (800 ng) was reverse transcribed using the QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's instructions.

2.5.5 qPCR analysis

For real-time quantitative PCR (qPCR), 20 ng of DNA or cDNA were used in a 10 μ l QuantiFast SYBR Green PCR kit (Qiagen) reaction using primers listed in Table 5, and analysed using the ABI7500 TaqMan (ThermoFisher) with the thermal cycler profile described in Table 6. Relative *Fusobacterium* spp. abundance and Siglec-7 expression was calculated by ΔC_T (C_T *Fusobacterium* spp. 16S - C_T Eubacteria 16S) and (C_T Siglec-7 - C_T Phosphoglycerate Kinase 1 (PGK1)) method, respectively.

Table 5 | Primers used in qPCR

Target	Forward Primer (5'→3')	Reverse Primer (5'→3')
PGK1 ²²⁹	GTGGAATGGCTTTTACCTTCC	CTTGGCTCCCTCTTCATCAA
SIGLEC7 ²³⁰	GGCCTGTATCAGGAGTGTTGCT	GCAGCCAGGCCATGGTG
<i>Fusobacterium</i> spp. ¹⁹⁵	GGATTTATTGGGCGTAAAGC	GGCATTCTACAAATATCTACGAA
Eubacteria 16S ¹⁹⁵	GGTGAATACGTTCCCGG	TACGGCTACCTTGTTACGACTT

Table 6 | qPCR cycling conditions

	Step	Time (min)	Temperature (°C)
	PCR initial heat activation	5	95
2-step Cycle (40 cycles)	Denaturation	00:10	95
	Combined Annealing/extension	00:35	60

2.5.6 16S rRNA Sequencing

The 16S rRNA sequencing was performed by Dr Falk Hildebrand's group (Quadram Institute Bioscience) as follows.

The DNA extracted from CRC tissues was normalised to 5 ng/μl with EB (10 mM Tris-HCl). A PCR master mix was made according to the instructions of the Kap2G Robust PCR kit Sigma using the V4 hypervariable region of the 16S rRNA primers²³¹ (515F: GTGYCAGCMGCCGCGGTAA, 806R: GGACTACNVGGGTWTCTAAT), and 19 μl of the mix added to each well of a 96-well plate. The PCR conditions were 95°C for 5 min, 30 cycles of 95°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec followed by a final 72°C for 5 min. The PCR product was treated with a 0.7X SPRI using KAPA Pure Beads (Roche) and eluted in 20 μl of EB (10 mM Tris-HCl). A second PCR master mix was made according to the instructions of the Kap2G Robust PCR kit, and 11 μl of the mix added to each well of a 96-well plate. 2 μl of each P7 and P5 of Nextera XT Index Kit v2 primers (Illumina) were added to each well. Finally, the 5 μl of the clean specific PCR mix was added and mixed. The PCR was run at 95°C for 5 min, followed by 10 cycles of 95°C for 30 sec, 55°C for 30 sec and 72°C for 30 sec and by a final run at 72°C for 5 min. The libraries were quantified using the high sensitivity Quant-iT dsDNA Assay Kit (ThermoFisher) by measuring the fluorescence (excitation/emission maxima ~510/527 nm) using FLUOStar (BMG Labtech). Libraries were pooled following quantification in equal quantities. The final pool was cleaned using 0.7X SPRI KAPA Pure Beads (Roche) according to the manufacturer's instructions. The final pool was quantified on a Qubit 3.0 instrument and on a High Sensitivity D1000 ScreenTape (Agilent) using the Agilent TapeStation 4200 to calculate the final library pool concentration.

The pool was run at a final concentration of 8 pM on an Illumina MiSeq instrument using MiSeq Reagent Kit v3 (600 cycle) (Illumina) following the Illumina denaturation and loading instructions, which included a 20% PhiX spike in (PhiX Control v3 Illumina).

The V4 16S rRNA region amplicon sequences were processed using a LotuS pipeline (v 2.0)²³². The steps included demultiplexing, quality-filtering, clustering of the reads with uparse²³³ at 97% identity and taxonomic classification. Host contamination was removed by aligning the reads to the human reference genome GRCh38.p13 using minimap²³⁴. Chimeric operational taxonomic units

(OTUs) were filtered out with uchime2²³⁵. Reads were aligned to their respective databases and the LotuS LCA algorithm using lambda aligner²³⁶ against reference database SILVA (v 123)²³⁷.

Sample analyses were performed using R (v 3.6.1). First, less than 1,000 reads of 16S rRNA were removed from the analysis and the sample depth was rarefied 10 times to 4,281 reads per sample with R-package rtk (v 0.2.6.1)²³⁸ using the smallest number of sequences per sample observed. Additionally, 16S taxa counts were normalised by the number of 16S copies for each sample.

2.5.7 Serological assays

For serum generation, plasma from CRC patients were treated with 20% w/v CaCl₂ (1:50 CaCl₂: Plasma) at 4°C for 2 days followed by centrifugation at 8,000 rpm for 5 min. Serum specific anti-Fn-IgG level was determined in the supernatant by an ELISA-based method.

For the detection of anti-*F. nucleatum* ssp. IgG in the serum, *F. nucleatum* ssp. or *R. gnavus* E1/ ATCC 29149 (controls) bacteria (10⁷ cells/well), *F. nucleatum* ssp.-derived LPS or *E. coli* O111:B4 LPS (control) in PBS (100 µl) were coated in an enhanced binding 96-well plate (ThermoFisher) o/n at 4°C. Wells were washed three times with washing buffer (as described in section 2.3.4). Wells were then blocked with 200 µl of 1% BSA in PBS at RT for 2 h. Wells were washed three times with washing buffer and then incubated with 100 µl of serum diluted at 1:800 in the above blocking buffer at 37 °C for 1 h. The plate was washed three times, and 100 µl of α-Fc-HRP diluted at 1:20,000 were added to the wells. Following an incubation at 37 °C for 30 min, the plate was washed three times, and TMB solution (BioLegend) was added to the wells. Colour development was stopped by the addition of 2N H₂SO₄ and the absorbance was measured at 450 nm with reference at 570 nm using FLUOStar (BMG Labtech).

For the detection of soluble Siglec-7 in serum, the Siglec-7/CD328 ELISA kit (R&D Systems) was used with serum diluted at 1:2 according to the manufacturer's instructions.

2.6 Bioinformatic analysis

Identification of genes encoding putative enzymes involved in the sialic acid biosynthetic pathway of *F. nucleatum* ssp. was conducted using the tblastn (ver. 2.7.1+) program. The genomes of *F. nucleatum* ATCC 25586 (accession # AE00995), *F. nucleatum* ATCC 10953 (accession # NZ_CM000440), and *F. nucleatum* ATCC 51191 (accession # GL985141) were accessible from National Center for Biotechnology Information (NCBI). The amino acid sequences encoding NeuA (N-acylneuraminate cytidyltransferase, NCBI accession # WP_005897387.1), NeuB (N-acetylneuraminate synthase, NCBI accession # WP_005897390.1). The genes, NeuC (UDP-N-acetylglucosamine 2-epimerase, tr|A0A2C6C850) from *F. nucleatum* ATCC 10953, as well as NeuS (Poly-alpha-2,8 sialosyl sialyltransferase, tr|Q1R749) and NeuD (acyltransferase, tr|Q1R744) from *Escherichia coli* strain UTI89 and NeuO (Polysialic acid O-acetyltransferase, sp|A1ADJ6) from *E. coli* O1:K1/ APEC were accessible from UniProt and were used as query. Hits with an E-value less or equal to zero were considered significant.

2.7 Statistical analyses

One-way ANOVA followed by Tukey's test or t test were used for multiple or two-group comparisons, respectively, on Prism software (GraphPad).

For the statistical analyses of 16S sequencing data, the rarefactions were used to calculate the mean Shannon's diversity index of each sample from 16S OTU frequencies. Richness was calculated as the total number of observed taxa in the samples. In a paired test design, the Wilcoxon signed-rank test was used to compare Shannon index, OTU richness and different taxa between on-tumour and off-tumour sites. $P < 0.05$ was considered as statistically significant. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; n.s., not statistically difference.

CHAPTER 3

-

Screening of CRC-associated bacteria binding to human recombinant lectins

3.1 Introduction

Commensal or pathogenic bacteria can interact with PRRs expressed on the surface of mammalian immune cells, regulating immunity. Among PRRs, lectins specifically recognise carbohydrate structures (glycoconjugates) on host or microbial cells through their carbohydrate recognition domain (CRD). The interaction of host lectins with glycoconjugates displayed by commensal or pathogenic microbes mediates homeostasis or inflammation, respectively, although lectins can also be involved in host's cell-cell recognition. Lectins, such as Siglecs¹¹⁰ or Galectin-3²³⁹ have been involved in tumour progression through their interaction with host glycans overexpressed on cancer cell surface, while Dectin-2 has been implicated in suppression of liver metastasis by inducing phagocytosis⁷⁸.

To date the role of lectins in tumour progression has been demonstrated to be mediated by their interaction with the host glycoconjugates displayed by malignant cells but their role in sensing tumour-associated bacteria has not been investigated.

Here, we assessed the binding of the dominant bacterial species implicated in colorectal tumour progression, *Fusobacterium nucleatum* ssp. and *Bacteroides fragilis* to human recombinant lectins, Dectin-2, Galectin-3 and Siglecs by flow cytometry.

The binding of *F. nucleatum* strains (ATCC 25586, 10953, 51191) or *B. fragilis* NCTC 9343 was first investigated against recombinant human Dectin-2 (CRD) and human full-length Galectin-3 proteins expressed in *E. coli*, or commercial Siglec-7-Fc. Dectin-2 was further purified as a fluorescence tetramer. Next, *F. nucleatum* strains (ATCC 25586, 10953, 51191) were tested for binding to a range of Siglecs expressed in CHO-cells.

3.2 Results

3.2.1 Interaction of *B. fragilis* and *F. nucleatum* ssp. with human Dectin-2, Galectin-3 and Siglec-7

3.2.1.1 Human recombinant Dectin-2 production

Human Dectin-2 is a transmembrane protein with a short cytoplasmic tail, a transmembrane segment and an extracellular domain which consists of a CRD and a stalk region. The CRD (aa 64-209) has a molecular weight of 17 kDa and contains an EPN (Glu-Pro-Asn) motif involved in the binding of Dectin-2 to glycans in a calcium-dependent manner. Here, we attempted to construct a Dectin-2-CRD oligomer to increase the Dectin-2-CRD and ligand binding avidity. The strategy employed was to co-express in *E. coli* the plasmids encoding human Dectin-2-CRD-biotin-tag and biotin ligase (BirA); the enzyme which catalyses the biotinylation of biotin carboxyl carrier protein (BCCP); in order to enzymatically biotinylate Dectin-2. The recombinant Dectin-2-CRD protein with a C-terminal biotinylated region will allow the formation of a tetramer complex after binding to streptavidin-tetramers (4 moles of biotinylated Dectin-2 protein bind to 1 mole of streptavidin) (Fig. 12).

Briefly, after transformation of *E. coli* BL21DE(3) with both the pT5T Dectin-2-CRD-biotin-tag plasmid (harbouring the carbenicillin resistance gene) and the BirA plasmid harbouring the chloramphenicol resistance gene), the colonies were selected from LB agar plates supplemented with carbenicillin and chloramphenicol and the presence of both plasmids confirmed by colony PCR with specific primers (Materials and Methods, section 2.2.2). The amplified DNA fragments showed sizes of ~555 bp and ~1 kbp, corresponding to those amplified from the control pT5T Dectin-2-CRD-biotin-tag plasmid (Fig. 13A) and BirA (Fig. 13B) plasmid, respectively, confirming successful transformation.

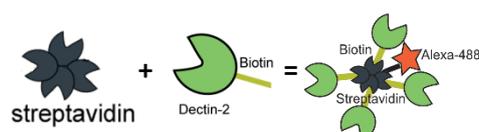


Figure 12 | Dectin-2-CRD and streptavidin tetrameric complex.

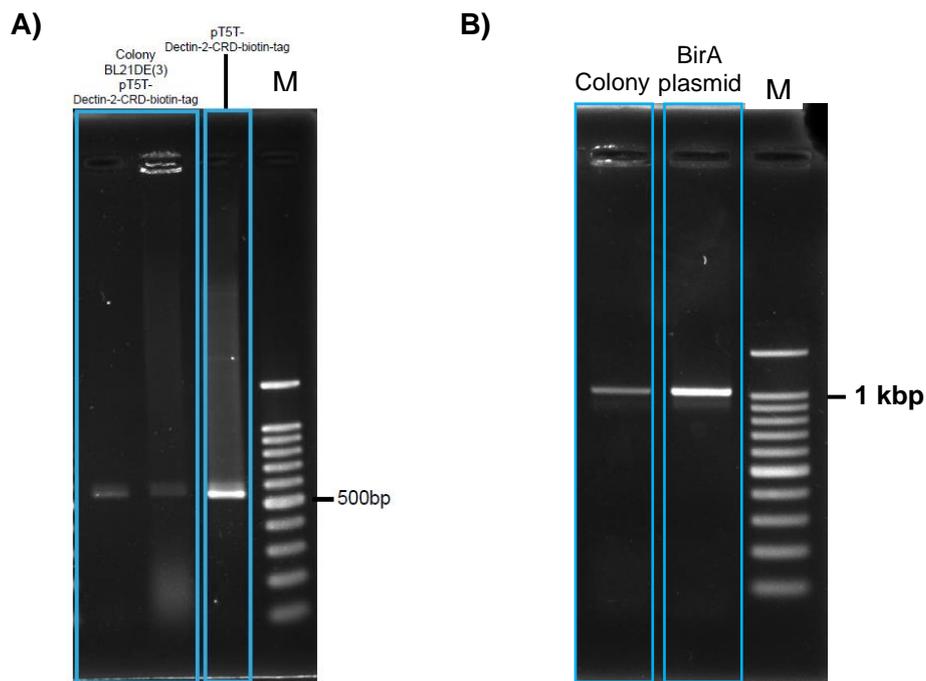


Figure 13 | Analysis of recombinant Dectin-2-CRD-biotin-tag colonies.

The positive colonies were subjected to PCR and the amplicon was analysed on a 1% agarose gel. A) Dectin-2-CRD-biotin-tag DNA is shown at 555 bp. The pT5T-Dectin-2-CRD-biotin-tag plasmid was used as a control. B) BirA insert is shown at 1 kbp. The plasmid containing the BirA gene was used as a control. As marker (M) was used a 100 bp ladder.

The recombinant protein was purified from *E. coli* BL21DE(3) by affinity chromatography using a mannose-sepharose column as described in Materials and Methods (section 2.2.3). Fractions were analysed by SDS-PAGE and those corresponding to the expected size for biotinylated human Dectin-2 CRD (hDC-2-CRD-bio) (approximately 17 kDa) (Fig. 14) were pooled (fractions 8 and 9), dialysed against PBS and concentrated. The total purification yield of hDC-2-CRD-bio was approximately 0.33 mg/l of bacterial culture. The protein eluting before the 8th fraction correspond to a proportion of Dectin-2 not re-folded successfully, resulting in low affinity for the mannose-sepharose column and early elution.

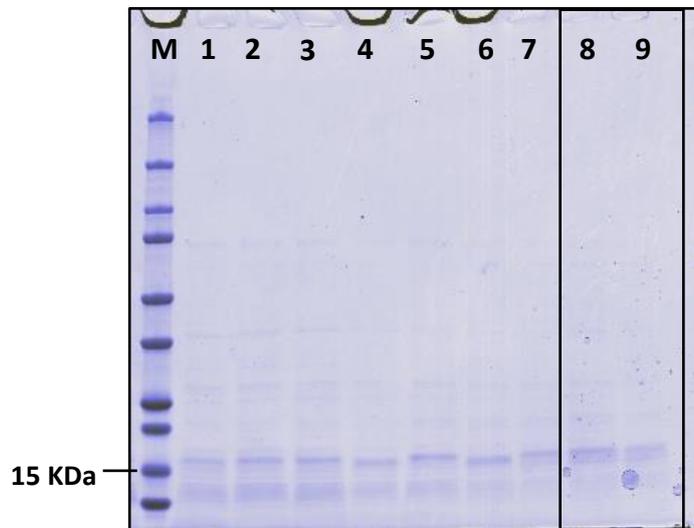
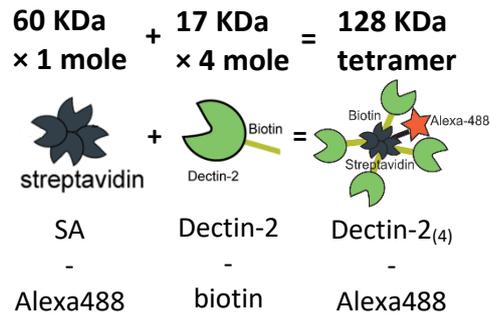


Figure 14| SDS-PAGE analysis.

hDC-2-CRD-bio was purified by affinity chromatography. Fractions eluted from a mannose-sepharose column were analysed on a 17.5% SDS-PAGE gel. Fractions 8 and 9 were pooled, dialysed and concentrated. M, marker (precision Plus Protein Dual Colour Standards, BIO-RAD).

In order to test biotinylation of the purified protein, Dectin-2-CRD (~17 kDa) protein was incubated with streptavidin (~60 kDa) and resolved in a non-reducing SDS-PAGE. An additional band of approximately 50 kDa was observed after the reaction with streptavidin, suggesting the formation of a complex between streptavidin and biotinylated Dectin-2-CRD (Fig. 15A). However, the apparent size of the complex was smaller than the expected size (128 kDa) (Fig. 15A) which corresponds to the formation of a full tetrameric complex between streptavidin and Dectin-2-CRD (Fig. 15B). This may be due to the non-reducing SDS-PAGE conditions affecting the migration of the streptavidin-Dectin-2 complex in the gel. In addition, a substantial amount of Dectin-2-CRD and streptavidin remained unbound, suggesting that a proportion of Dectin-2-CRD protein was not biotinylated (Fig. 15B).

A)



B)

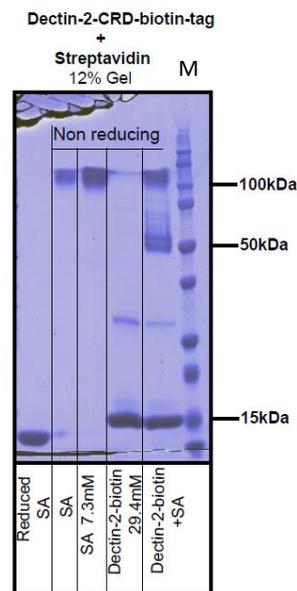


Figure 15 | Analysis of recombinant Dectin-2-Streptavidin complex.

A) The calculated size of Dectin-2-CRD and streptavidin complex is 128 kDa.

B) Non-reduced SDS-PAGE. SA, Streptavidin; M, marker (precision Plus Protein Dual Colour Standards, BIO-RAD).

Next, the recombinant protein was incubated with fluorescently labelled streptavidin (SA-Alexa488) and the complex purified by FPLC using a Superdex 200 column. The FPLC chromatogram showed two protein peaks, suggesting the formation of complexes with different degrees of oligomerisation (Fig. 16A). Peak 1 corresponds to the theoretical size of the tetramer human Dectin-2-Alexa488 (hDC-2₍₄₎-Alexa488), as confirmed by native-PAGE analysis (Fig. 16B). The size of the complex in peak 2 could not be determined due to low protein concentration.

The tetrameric hDC-2₍₄₎-Alexa488 complex showed in peak 1 was used in the rest of the study.

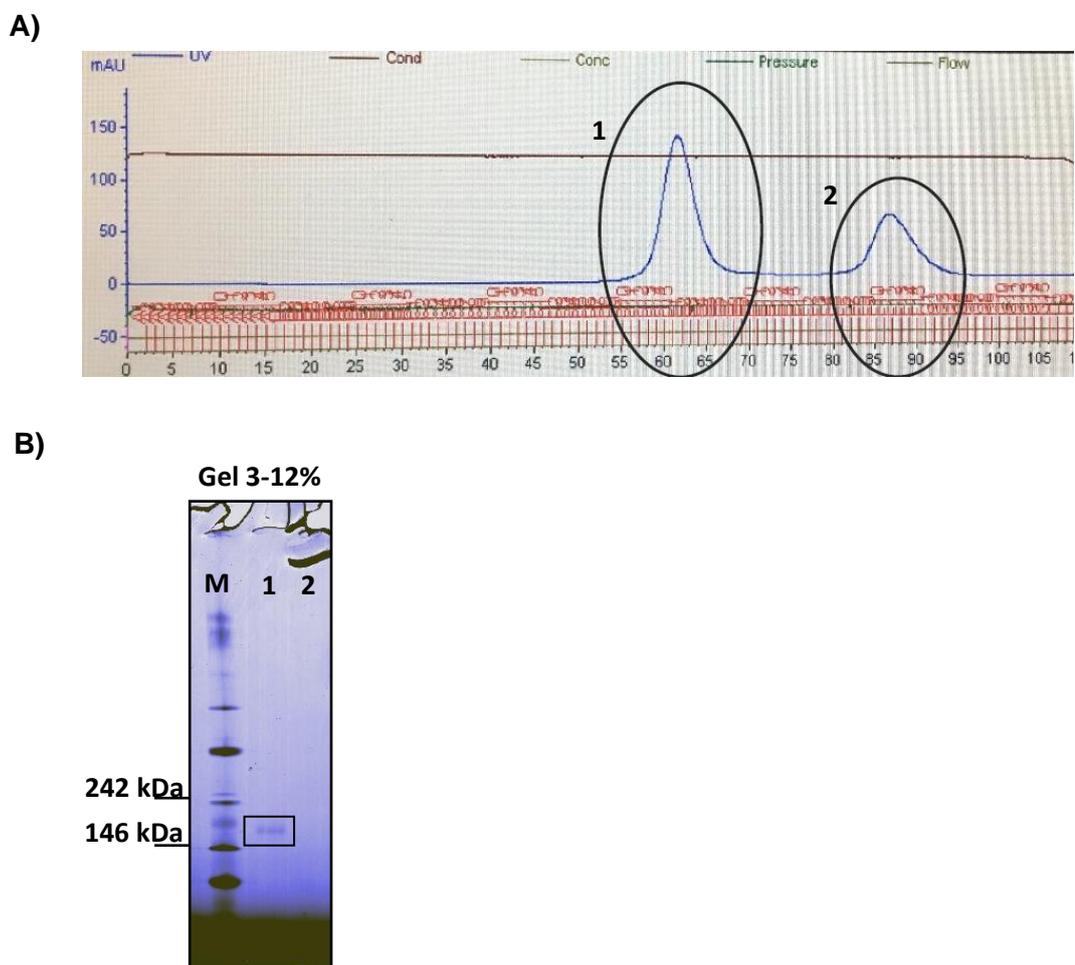


Figure 16 | Purification of hDC-2₍₄₎-Alexa488 tetramer by FPLC.

A) FPLC chromatogram showing two peaks corresponding to complexes of hDC-2₍₄₎-Alexa488 oligomers with different degrees of polymerisation, B) Native-PAGE of the fractions corresponding to the two peaks. The calculated size of the tetramer is 128 kDa. M, Marker NativeMark, ThermoFisher).

Saccharomyces cerevisiae is known for binding to Dectin-2-CRD in a calcium dependent manner²²². Here, *S. cerevisiae* was used as a control to assess the functionality of the hDC-2-Alexa488 tetramer. Binding of *S. cerevisiae* was tested to the purified hDC-2₍₄₎-Alexa488 in the presence of calcium or EGTA (a calcium chelator) using flow cytometry. The results showed a clear shift of the yeast population to positive values in the presence of calcium, while binding was abolished in the presence of EGTA (Fig. 17). This indicates binding of hDC-2₍₄₎-Alexa488 to *S. cerevisiae* in a calcium dependent manner via its CRD (Fig. 17), confirming the functionality of the recombinant tetrameric protein.

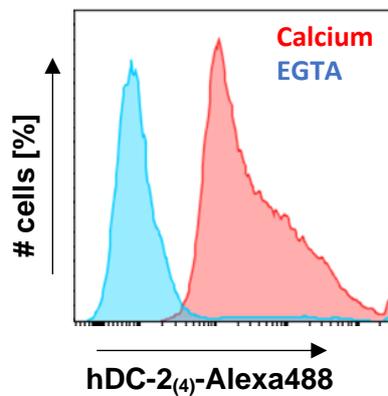


Figure 17 | Histogram of hDC-2₍₄₎-Alexa488 binding to *S. cerevisiae*.

S. cerevisiae binding to the tetramer human Dectin-2 was analysed by flow cytometry in the presence of calcium (in red) and the binding is abolished in the presence of EGTA (in blue).

3.2.1.2 Human recombinant Galectin-3 production

Recombinant full-length Galectin-3 proteins were expressed in *E. coli* and purified by affinity chromatography using a two-step purification process as described in Materials and Methods (section 2.2.4) (Fig. 18).

In the first step, Galectin-3 was eluted from the IMAC column based on metal-histidine affinity, and fractions 3 to 7 were collected (Fig. 18A) and buffer exchanged in PBS. In the second step, the sample was subjected to affinity chromatography using a lactose column, and fractions 6 to 12 were collected (Fig. 18B) and buffer exchanged in PBS. The expected molecular weight of the purified Galectin-3-full length-his-tag recombinant protein was approximately 28 kDa as shown by SDS-PAGE (Fig. 18). The purification yield of Galectin-3-full length-his-tag protein was approximately 1.28 mg/l of bacterial culture.

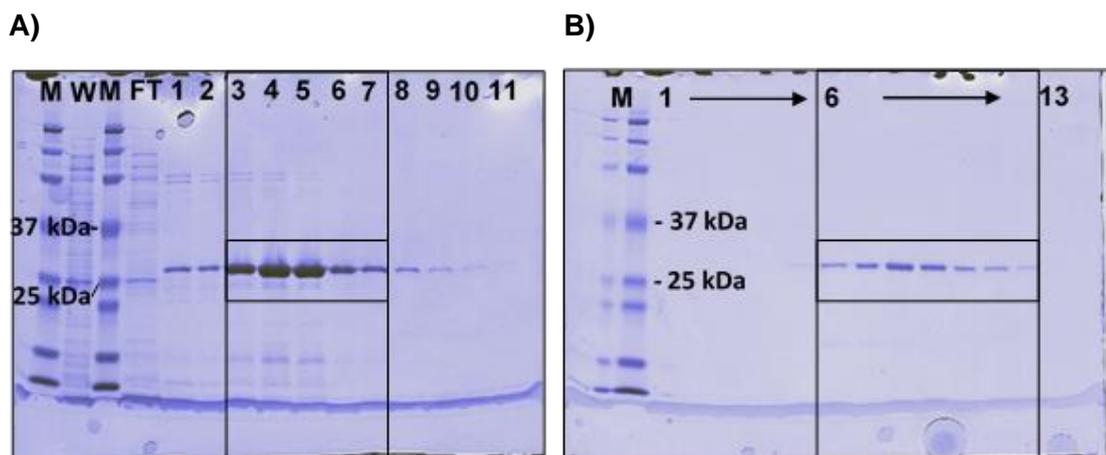


Figure 18 | SDS-PAGE analysis of his-tag Galectin-3 purification.

Fractions from the A) IMAC and B) lactose affinity chromatography were analysed by electrophoresis using a 12% SDS-PAGE gel. W, wash; FT, flow through; M, marker (precision Plus Protein Dual Colour Standards, BIO-RAD).

Here, *Klebsiella pneumoniae* O1, a known Galectin-3 binding bacterium¹¹⁶, was used as a control to assess the functionality of the human Galectin-3 by flow cytometry. Bacteria were incubated with the recombinant histidine tagged Galectin-3, and the binding detected using a mouse anti-histidine antibody and a fluorescently labelled anti-mouse-APC antibody (Fig. 19A). The results showed a shift of the bacterial population to positive intensity values when incubated with Galectin-3 (Fig. 19B), confirming the functionality of the recombinant human Galectin-3.

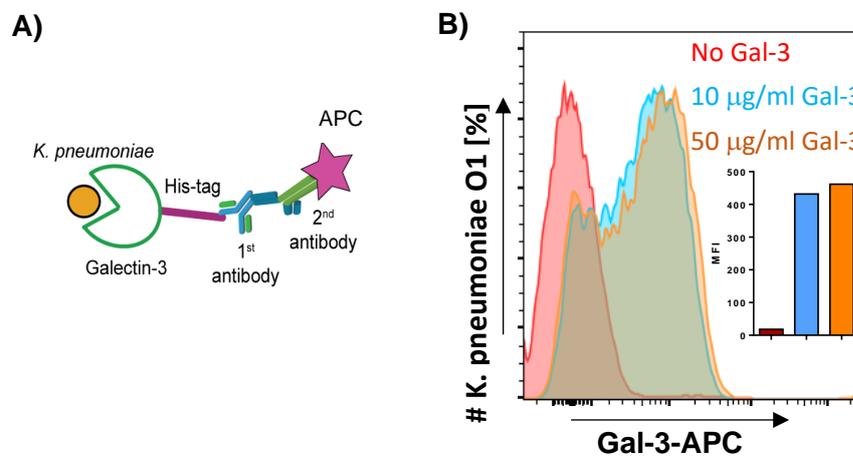


Figure 19 | Galectin-3 binding to *K. pneumoniae* O1.

A) Galectin-3 labelling approach for binding detection by flow cytometry. 1st antibody was mouse anti-his-tag, 2nd antibody was the anti-mouse IgG-APC. B) Histogram and the mean fluorescence value (MFI) of Galectin-3 at 10 µg/ml (in orange) or 50 µg/ml (in blue) binding to bacteria. His-tag, histidine-tagged; Gal-3, Galectin-3. As a control was used bacteria with no Galectin-3 (in red).

3.2.1.3 Binding assays of *B. fragilis* and *F. nucleatum* ssp. to Dectin-2, Galectin-3 and Siglec-7

F. nucleatum strains (ATCC 25586, 10953, 51191) and *B. fragilis* NCTC 9343 were screened for their binding to recombinant hDC-2(4)-Alexa488 and Galectin-3 produced above, as well as to commercial Siglec-7-Fc using flow cytometry (Fig. 20). The strains were grown at Optical Density in 600 nm (OD_{600nm}) of 0.7-1.3 and normalised to 10^7 cells (see Materials and Methods, section 2.1.3). Flow cytometry was used to monitor binding as shown above for hDC-2(4)-Alexa488 and Galectin-3. Siglec-7-Fc was first pre-complexed with α -Fc-PE Ab, to achieve multivalency, followed by incubation with the bacteria. The data showed strong binding of all *F. nucleatum* strains tested to Siglec-7-Fc as indicated by the shift of the bacterial population from zero (control) to positive intensity values (Fig. 20A). In contrast, only a proportion of the *F. nucleatum* ssp. population bound to Galectin-3: 20% for *F. nucleatum* ATCC 51191 and ATCC 10953, and 40% for *F. nucleatum* ATCC 25586, respectively (Fig. 20A). *F. nucleatum* ssp. showed partial binding to hDC-2(4)-Alexa488 in the presence of calcium and the binding was reduced by approximately 10% following addition of EGTA (Fig. 20A).

B. fragilis NCTC 9343 showed binding to hDC-2(4)-Alexa488, but the addition of EGTA did not reduce the binding (Fig. 20B). Also, *B. fragilis* showed only approximately 10% binding to Galectin-3 and no binding to Siglec-7-Fc (Fig. 20B).

Together these results show that among the bacterial species, *B. fragilis* and *F. nucleatum* and the recombinant lectins, Siglec-7-Fc, Galectin-3, and hDC-2(4)-Alexa488 tested, strong binding was observed between *F. nucleatum* ssp. and Siglec-7-Fc.

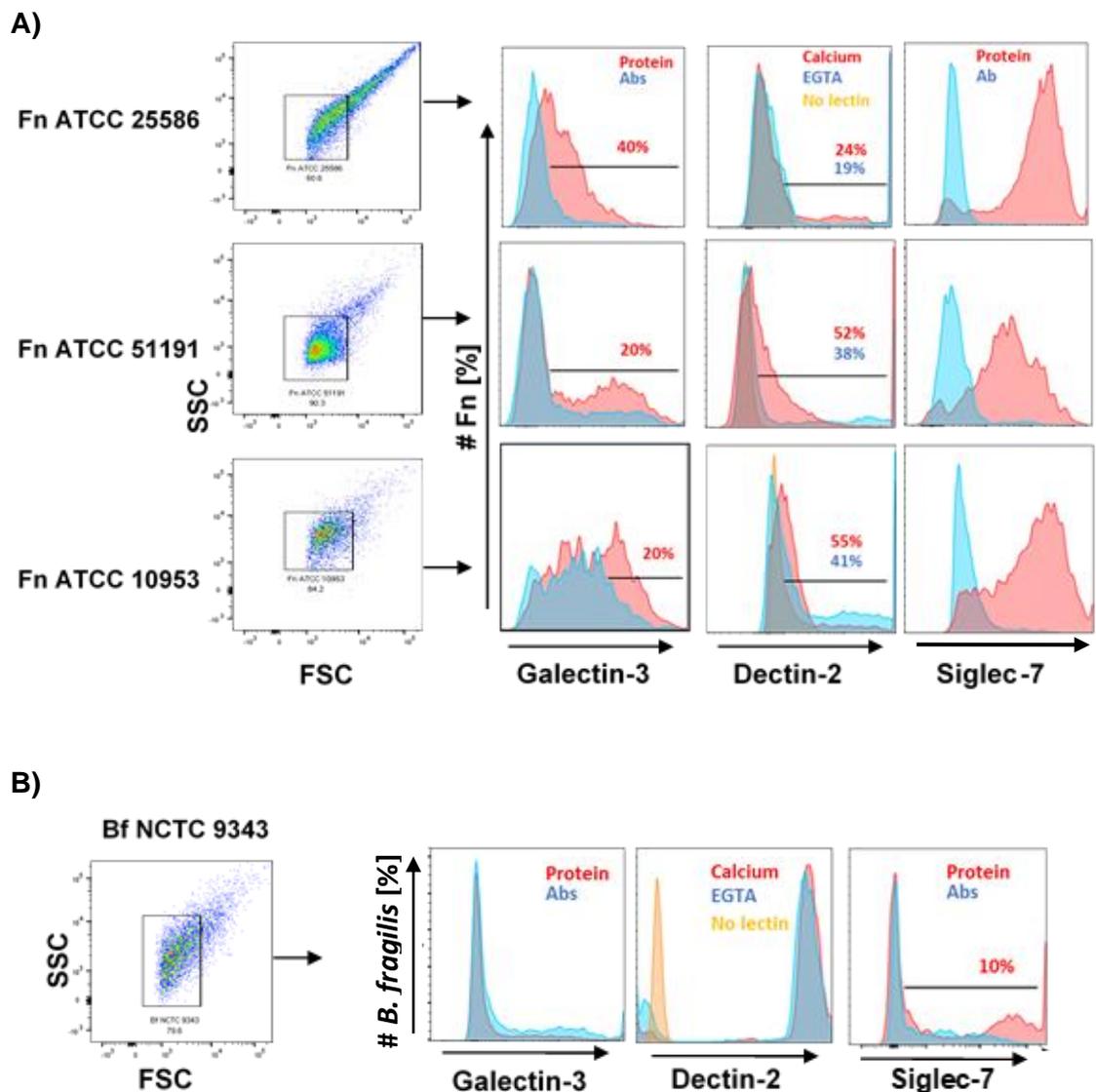


Figure 20 | *F. nucleatum* ssp. and *B. fragilis* NCTC 9343 binding to human lectins by flow cytometry.

A) *F. nucleatum* binding analysis to Galectin-3 (in red), hDC-2₍₄₎-Alexa488 in presence of calcium (in red) or EGTA (in blue), and Siglec-7 (in red). B) *B. fragilis* NCTC 9343 binding analysis to Galectin-3, hDC-2₍₄₎-Alexa488 and Siglec-7-Fc. Fn, *F. nucleatum*; Bf, *B. fragilis*, FSC, forward scatter; SSC, side scatter. As controls were used: for hDC-2₍₄₎-Alexa488 binding assays, the bacteria alone with no lectin (in orange), and for Galectin-3 or Siglec-7-Fc binding assays, the bacteria incubated with the antibodies and no lectin (in blue). Data shown are from one representative experiment reproduced in three independent experiments.

3.2.2 *F. nucleatum* ssp. bound specifically to Siglec-7-Fc

Having shown that all three *F. nucleatum* strains (ATCC 25586, 10953, 51191) bound to Siglec-7-Fc, we investigated the binding of these strains towards a panel of recombinant Siglec-Fcs including human Siglec-(2, 3, 5, 7, 9, 10, and E)-Fcs and mSiglec-1-Fc expressed in CHO cells.

3.2.2.1 Recombinant Siglec-Fc production

The recombinant Siglec-7-Fc and mSiglec-1-Fc proteins were produced in CHO cell lines established by Prof. Paul Crocker, using the Glutamine Synthetase (GS) expression-based selection system, as described in Materials and Methods (section 2.2.5). In this system, to achieve high yield expression of recombinant Siglecs, CHO cells were transfected with a vector carrying a cDNA encoding Siglec-Fc and GS (L-glutamate) genes. GS is catalysing the glutamate and ammonia reaction for glutamine production which is an important metabolic process for cell survival. When methionine sulphoximine MSX (a glutamate analogue) is added to the media, it causes inhibition of the endogenous glutamine synthesis in CHO cells. Furthermore, it induces transcription of the transfected GS cassette-vector²⁴⁰ which results in high yield production of Siglec-Fc proteins in the supernatant.

Following expression in CHO cells, Siglec-Fc proteins were then purified by affinity chromatography using a protein A-sepharose column. Fractions were analysed by SDS-PAGE (Fig. 21), and those containing Siglec-Fc, fractions 1 to 6 for Siglec-7-Fc (Fig. 21A), and fractions 2 to 4 for mSiglec-1-Fc (Fig. 21B), were pooled and buffer-exchanged in PBS. The purification yield of Siglec-7-Fc, and mSiglec-1-Fc was approximately 1.5 mg/l, and 7.6 mg/l of CHO culture, respectively.

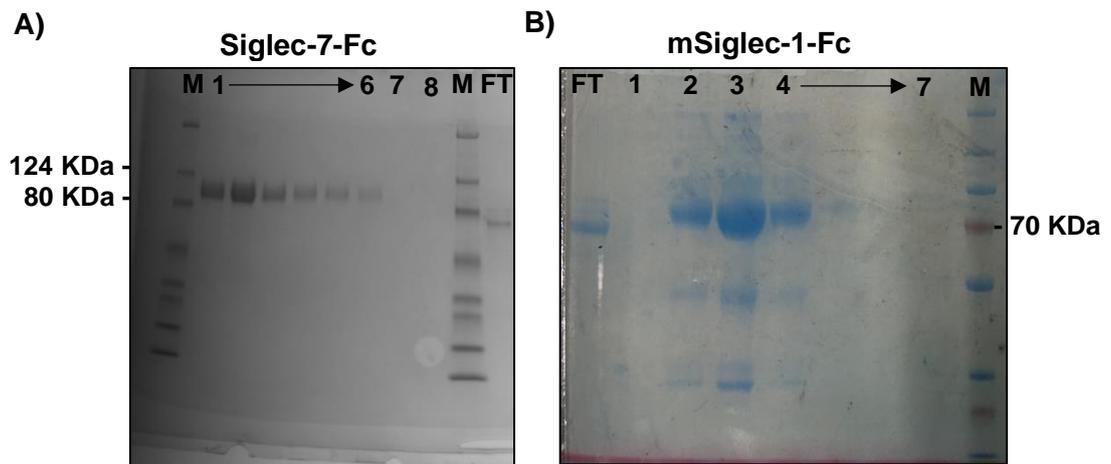
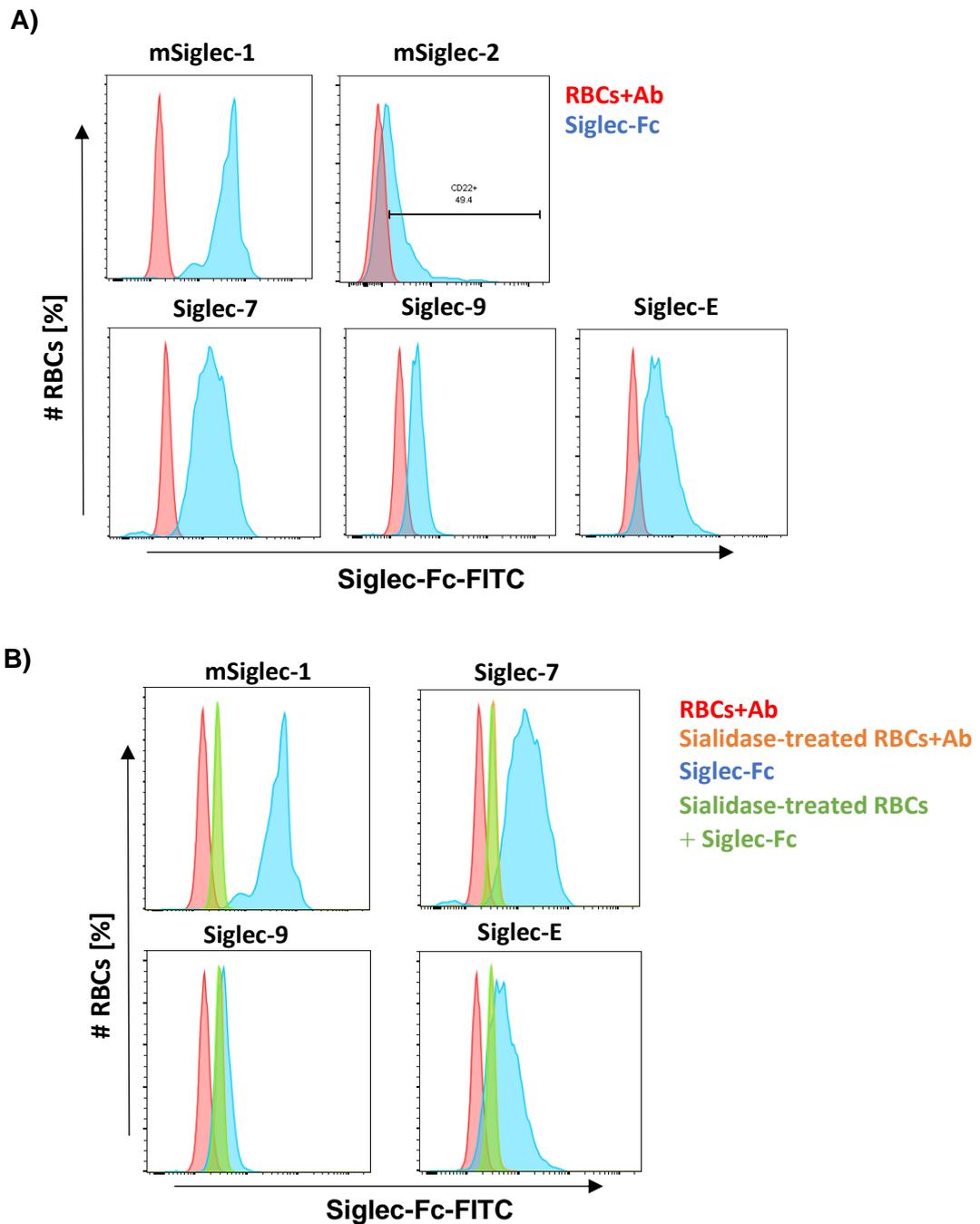


Figure 21 | SDS-PAGE analysis of recombinant Siglecs produced in CHO cells. Fractions from protein A column of A) Siglec-7-Fc and B) mSiglec-1-Fc, were analysed by electrophoresis in a 4-12% SDS-PAGE gel. M, marker (pre-stained SDS-PAGE broad range Standards, BIO-RAD) and right the PageRuler Prestained, ThermoFisher; FT, flow through.

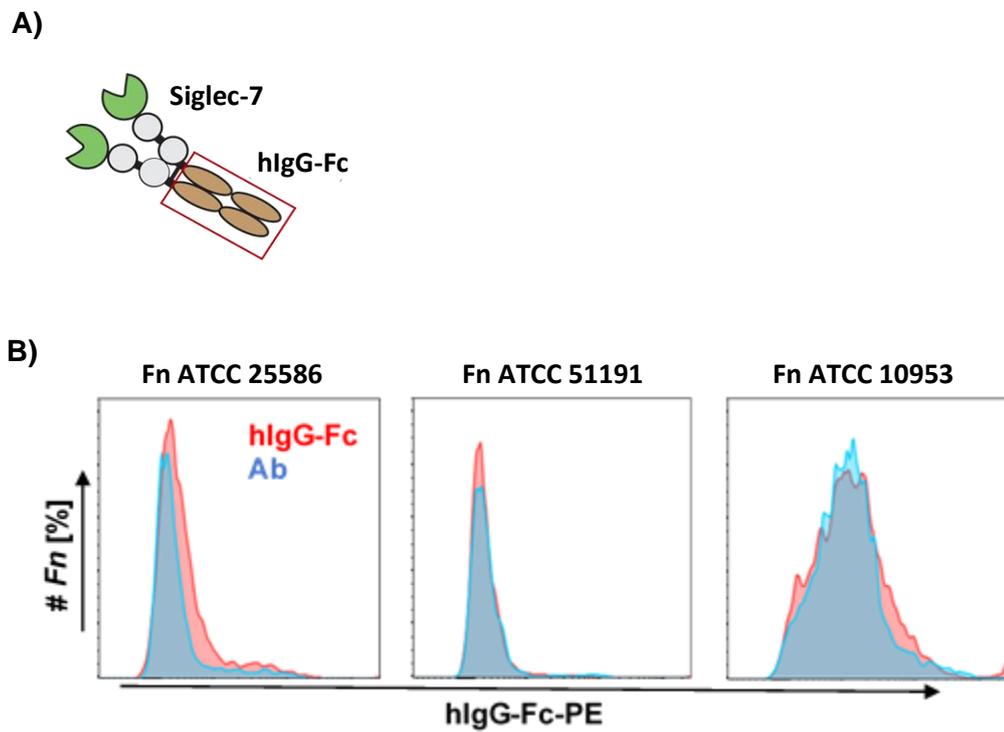
Next, we tested the functionality of the recombinant Siglec-Fc proteins produced in CHO cells by conducting binding assays with human red blood cells (RBCs). RBCs express glycoproteins which carry sialic acids²⁴¹, such as the known ABO blood group glycoproteins²⁴², which are ligands of Siglec-1, Siglec-2, Siglec-5, Siglec-7, Siglec-9 and Siglec-E. Due to the low affinity binding of Siglecs to sialic acids²⁴³, binding of RBCs was tested against recombinant Siglec-Fcs pre-complexed with α -Fc-FITC Ab to increase multivalency (as described in Materials and Methods section 2.3.2) apart for mSiglec-2 which was used in a non pre-complexed form, and Siglec-9 which was immobilised onto protein A beads. The results of the flow cytometry binding assays showed a clear shift of the RBC population towards higher intensity values after binding with mSiglec-1, or hSiglec-7, -9 and -E (Fig. 22A), confirming the functionality of the recombinant Siglec-Fc proteins. mSiglec-2 showed binding to around 50% of the RBC population, which may be due to the lack of multivalence for this Siglec type. Further, we showed that treating RBCs with sialidase abolished the binding of mSiglec-1, -E or human Siglec-7 and -9 to the cells, confirming the specificity of the binding (Fig. 22B).



3.2.2.2 Binding of *F. nucleatum* strains to recombinant Siglecs

Here, we tested the binding of *F. nucleatum* ATCC 51191, ATCC 10953 and ATCC 25586 strains to the panel of recombinant Siglecs purified from CHO cultures by flow cytometry. *Ruminococcus gnavus* E1 was chosen as a control which does not express sialic acids on its surface (unpublished data). Recombinant Siglec-Fcs consist of two domains, a Siglec homodimer and a human IgG-Fc conjugate (Fig. 23A). When bacteria were incubated with the pre-complexed human IgG-Fc and α -Fc-PE Ab, no binding was observed to the human IgG-Fc domain (Fig. 23B), confirming the specificity of the interaction between *F. nucleatum* and the Siglec region of the recombinant Fc conjugated proteins.

The incubation of bacteria with the pre-complexed Siglec-Fcs and α -Fc-PE Ab, resulted in a strong binding of all *F. nucleatum* strains tested to Siglec-7-Fc, as shown by the shift of the bacterial population to higher intensity values, across all strains (Fig. 24A). Interestingly, only *F. nucleatum* ATCC 25586 showed binding, of approximately 76% of the population to Siglec-9-Fc (Fig. 24A), a Siglec with close homology and binding specificity to Siglec-7. In addition, a proportion of *F. nucleatum* ATCC 25586 population of approximately 79% showed binding to Siglec-5-Fc, and approximately 60% of *F. nucleatum* ATCC 51191 population bound to Siglec-5-Fc (Fig. 24A). *F. nucleatum* strains did not show binding to mSiglec-E-Fc, the mouse ortholog of Siglec-9 and -7 (Fig. 24B).



A) Schematic representation of a recombinant Siglec-Fc (Siglec-7-Fc), showing the conjugation with the hlgG-Fc domain. B) Histograms of *F. nucleatum* ssp. binding to hlgG-Fc protein (in red) only. hlgG-Fc, human Immunoglobulin G-Fragment crystallizable; Fn, *F. nucleatum*. As a control was used the bacteria incubated with antibody only (in blue).

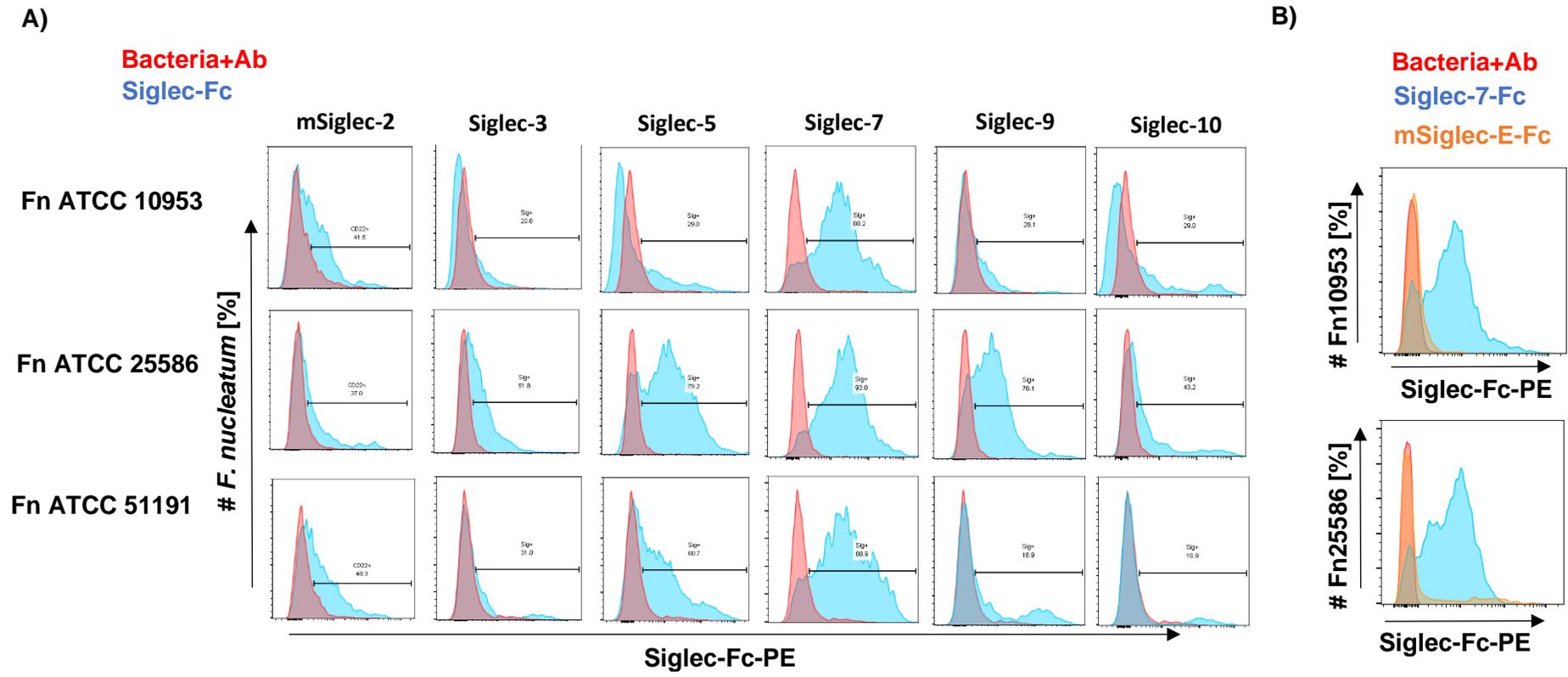


Figure 24 | Histogram of Siglec-Fc binding to *F. nucleatum* ssp.
 A) Binding assays of *F. nucleatum* ATCC 10953, ATCC 25586, ATCC 51191 to human Siglec-3, -5, -7, -9 or -10 or mouse Siglec -2 (in blue), and B) binding assays of *F. nucleatum* ATCC 10953 or ATCC 25586 to Siglec-7 (in blue) or mSiglec-E (in orange) using flow cytometry. Fn, *F. nucleatum*. As a control was used the bacteria incubated with antibody only (in red). Data shown are from one representative experiment reproduced in three independent experiments.

R. gnavus E1 did not bind to any Siglec-Fc lectin (Fig. 25), supporting the specificity of *F. nucleatum* ssp. binding to Siglec-Fc proteins.

Bacteria+Ab
Siglec-Fc

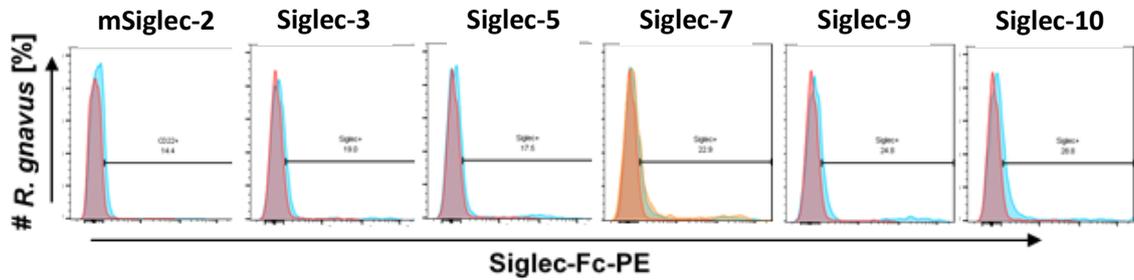


Figure 25 | Histogram of Siglec-Fc binding to *R. gnavus*.

R. gnavus binding to Siglec-Fc lectins (in blue) was analysed by flow cytometry. As a control was used the bacteria incubated with antibody only (in red).

3.2.2.3 Siglec-7 is involved in *F. nucleatum* ssp. association with mammalian cells

In order to investigate whether *F. nucleatum* binds to Siglec-7 displayed on mammalian cells, flow cytometry binding assays were carried out between *F. nucleatum* ATCC 10953, ATCC 25586, or ATCC 51191 strains and the human U937 monocytic cell lines expressing Siglec-7 (WT) or Siglec-7 deficient (Siglec-7^{-/-}) (Fig. 26A).

The results showed that approximately 23-40% of *F. nucleatum* population was associated with the WT U937 cells (Fig. 26B), and that association of *F. nucleatum* ATCC 25586, ATCC 10953, and ATCC 51191 to Siglec-7 deficient U937 cells was reduced by approximately by 14%, 6% and 4.5%, respectively, when compared to the WT U937 cells (Fig. 26B).

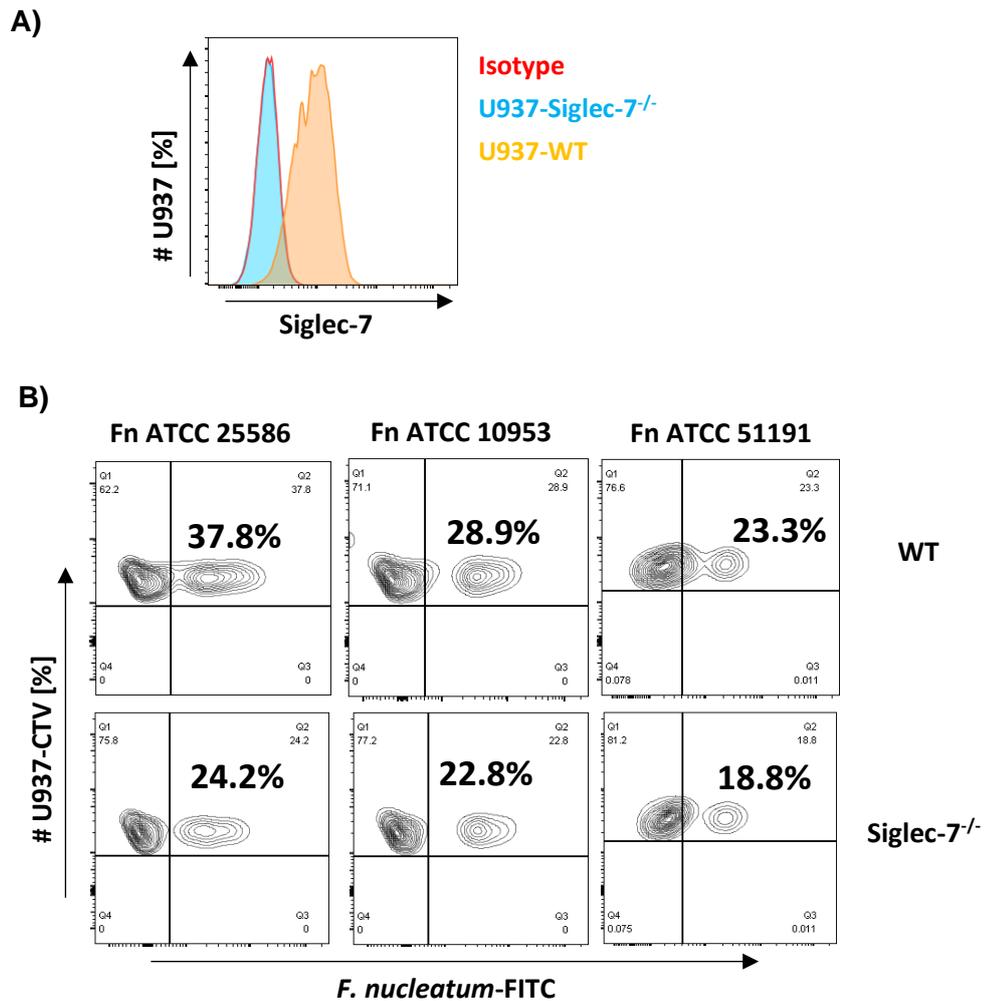


Figure 26 | Siglec-7 expression on U937 cells and *F. nucleatum* ssp. association assays.

A) Expression of Siglec-7 on WT (in orange) or Siglec-7^{-/-} (in blue) cells using anti-Siglec-7 antibody, with the isotype (in red) as a control, and B) association assays of *F. nucleatum* ATCC 25586, ATCC 10953, ATCC 51191 to U937 (WT or Siglec-7^{-/-}) cells, using flow cytometry. Fn, *F. nucleatum*. Data shown are from one representative experiment reproduced in three independent experiments.

These results are also consistent with binding assays conducted using WT-CHO cells (which do not express Siglec-7) or Siglec-7-expressing engineered CHO cells (Fig. 27A), showing that *F. nucleatum* ATCC 10953 or ATCC 25586 association to CHO cells was increased in Siglec-7-expressing CHO cells by 38% and 35%, respectively, when compared to WT CHO cells (Fig. 27B).

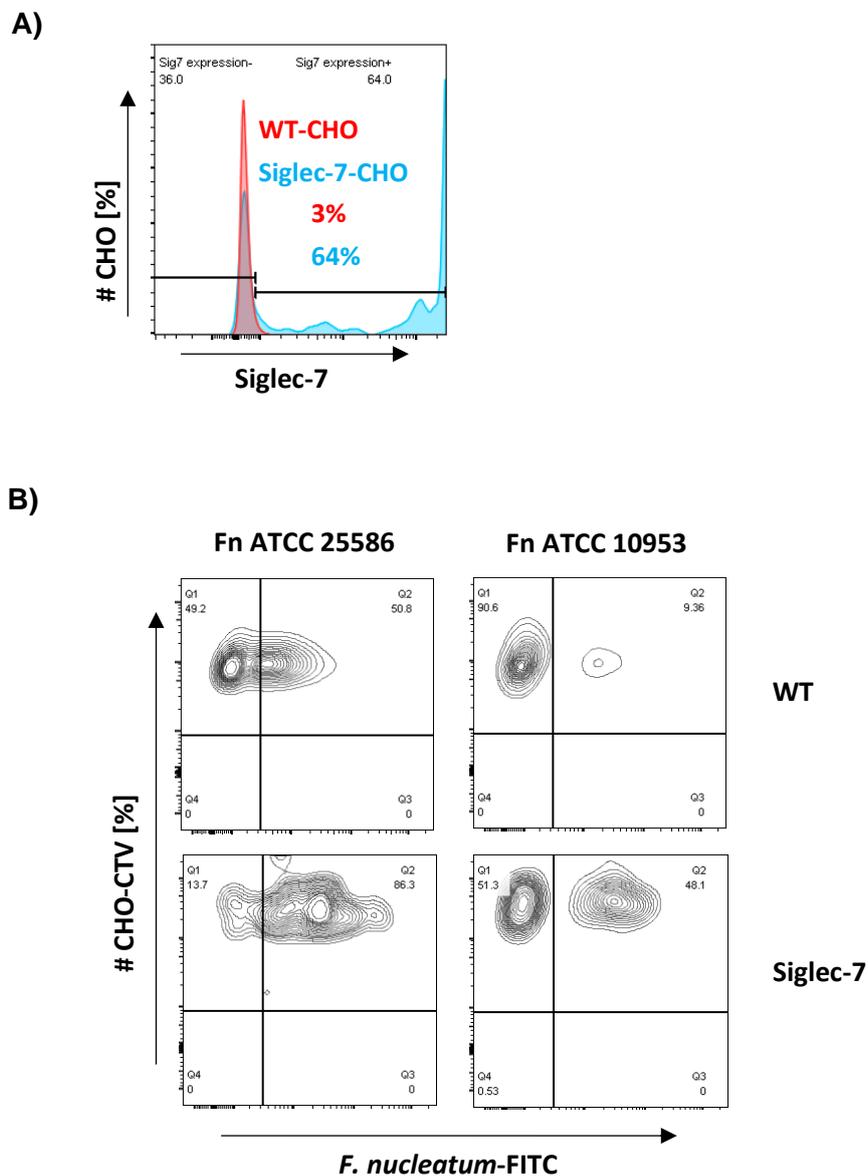


Figure 27 | Siglec-7 expression on CHO cells and *F. nucleatum* ssp. association assays.

A) WT-CHO cells (in red) did not express Siglec-7, while approximately the 64% of Siglec-7 CHO cell population (in blue) express Siglec-7. B) Association assays of *F. nucleatum* ATCC 25586, ATCC 10953 to CHO (WT or Siglec-7-expressed) using flow cytometry. Fn, *F. nucleatum*. Data shown are produced from one experiment.

3.3 Discussion

Here, colorectal cancer associated bacteria, *F. nucleatum* ssp. ATCC 10953, ATCC 25586, ATCC 51191, and *B. fragilis* NCTC 9343, were screened for their binding specificity to a panel of host lectins. Our results showed that among the different bacterial strains and lectins tested, the main interaction was observed between *F. nucleatum* ssp. and recombinant Siglec-7-Fc, and this interaction was also confirmed on Siglec-7 expressing cell lines.

B. fragilis NCTC 9343 showed only partial binding to Siglec-7-Fc, but strong binding to the recombinant Dectin-2, although this interaction was not through the calcium-dependent Dectin-2-CRD domain. The binding could be the result of protein-protein interaction or unspecific charged interactions with a non-canonical site of the Dectin-2-CRD protein. *B. fragilis* has a capsular polysaccharide which is exposed to the extracellular environment containing the polysaccharide (PS) A and B. The chemical characterisation of PS A and B sugar repeating units showed the presence of charged groups both positively and negatively¹⁹⁰, including β -galactosides, although the presence of sialic acids or sialic-acid molecules has not been reported. Our results that *B. fragilis* NCTC 9343 did not bind to Galectin-3 in the conditions tested, are surprising as Galectin-3 is a β -galactoside lectin. This may be due to a reduced accessibility of Galectin-3 to the *B. fragilis* β -galactoside structures, which occupy internal positions within PS.

Siglecs have been reported for binding to sialic acids on host cells¹⁰⁴ but also on pathogens⁹⁶. Our results are the first to demonstrate that Siglec-7 binds to colorectal cancer-associated bacteria, *F. nucleatum* ssp. Human Siglec proteins including Siglec-1, -5, -7, -9, or -14 have been shown to interact with *Campylobacter jejuni*, *Neisseria meningitidis*, Group B *Streptococcus* (GBS), *E. coli*, and modulate host immunity^{96,106}. To date, interactions of Siglecs with pathogens have mainly been reported to involve sialic acids or sialic acid-like molecules presented on the pathogen cell surface but other physiological ligands remain unknown. Siglec-7 has been shown to interact with *E. coli* strains (sialylated or not)¹⁰⁶, and with *C. jejuni*¹⁰⁵ in a sialic acid-dependent manner, while binding to GBS was through the bacterial β -protein in a sialic acid-independent manner⁹⁸. Siglec-1 and Siglec-7 have also been shown to interact with human

immunodeficiency virus (HIV)^{244,245}, and Siglec-4 with Varicella-zoster virus (VZV), or herpes simplex virus (HSV)²⁴⁶ for infection enhancement. Binding of Siglec-7 to virus was in a sialic acid-dependent manner^{245,246}.

F. nucleatum ssp. outer membrane contains lipopolysaccharides (LPS) showing variable chemical composition of the O-antigen repeating unit across bacterial strains. In the O-antigen domain, *F. nucleatum* ATCC 10953 presents internal sialic acids²¹⁹ and *F. nucleatum* ATCC 25586 contains a sialic acid-like molecule called fusaminic acid²²⁰, while together with our collaborator we recently characterised and published the O-antigen sugar structure of *F. nucleatum* ATCC 51191, which contains uronic acids²⁴⁷, an active negatively charged sugar group, but no sialic acid or sialic-acid like sugar. Our results that all *F. nucleatum* strains tested bind to Siglec-7 may suggest that Siglec-7 recognise non-sialic acid ligands on the cell surface of *F. nucleatum* strains, a hypothesis that will be explored in the next chapter. It should be noted that *F. nucleatum* strains did not show binding to Siglec-E-Fc, the mouse ortholog of Siglec-9 and -7. Despite the sequence similarities between Siglec-E, Siglec-9 and -7, these Siglecs appear to have different ligand specificities¹⁰⁴. Recently, Wisnovsky *et al.* 2021 showed that Siglec-7 but not Siglec-9 recognises sialoglycans presented on CD43-expressing leukemia cells¹⁰⁸. Furthermore, none of the *F. nucleatum*-derived LPS from the strains tested here contain mannose, the ligand of Dectin-2, and only the *F. nucleatum* ATCC 10953 strain presented a β -galactoside, in line with our results showing no significant binding of the strains to Dectin-2 and Galectin-3.

Our results that Siglec-7 binds to *F. nucleatum* ssp., may contribute to the immunomodulation properties of these strains and their role in promoting tumour development. In line with this, we showed that *F. nucleatum* ssp. could bind to Siglec-7 displayed by mammalian cells and in particular the human U937 monocytic cell line. The next chapters investigate in more detail the molecular receptors of the interaction between Siglec-7 and *F. nucleatum* and the impact of the interaction on the host immune response.

CHAPTER 4

-

Investigating the role of *F. nucleatum* ssp. cell surface polysaccharides in the interaction with Siglec-7

4.1 Introduction

The sialic acid immunoglobulin-like lectin 7 (Siglec-7) is expressed by leukocytes and induces immune suppression upon sialic acid recognition¹⁰⁴. Siglec-7 has mainly been studied for its interaction with host sialosides, showing a binding specificity towards α 2,8-disialylated structures, such as the ganglioside GD3, and branched α 2,6-sialyl residues¹⁰². The crystal structure of Siglec-7 has been determined, showing a V-set domain which includes an Arginine residue (R) at position 124, and a C-C' loop region, both involved in the binding of sialylated residues²⁴⁸. While Siglec-7 has 80% sequence similarities with Siglec-9, the latter does not bind α 2,8-disialyl residues, which may be due to differences in the C-C' loop²⁴⁸. Recently, Yamakawa *et al.* 2020 identified a novel binding site of Siglec-7 close to the canonical binding site and these two sites bind di- and tri-sialic acids in a similar way²⁴⁹.

Sialic acids are part of the large family of nonulosonic acids (NuLOs), sugars with a 9-carbon backbone α -keto acid structure, which also includes legionaminic and pseudaminic acids structurally similar to sialic acids (Fig. 28). Legionaminic and pseudaminic acids are mainly expressed by bacteria and found in bacterial capsular polysaccharides²⁵⁰ and/or in the O-antigen domain of lipopolysaccharide (LPS)²⁵¹. In mammals, N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc) are the most common sialic acids, while in humans Neu5Gc is absent due to the loss of function of the sialic acid-modifying enzyme CMP-N-acetylneuraminic acid hydroxylase (CMAH)^{252,253}.

Although Neu5Ac has mainly been studied and characterised in mammals, it is also expressed by pathogenic bacteria²⁵⁴, presumably through an alteration of the legionaminic acid biosynthesis pathway, as suggested by phylogenetic analyses²⁵⁵. In bacteria, the pathways for the *de novo* biosynthesis and modifications of sialic acid (Neu5Ac) have been well described in the model organism *E. coli* K1²⁵⁶ (Fig. 29). The genes involved in Neu5Ac biosynthesis and modification are part of a cluster comprising *neuA*, *neuB*, *neuC*, *neuD*, *neuS*, and *neuO* (Fig. 29). The biosynthesis begins when uridine diphosphate N-acetylglucosamine (UDP-GlcNAc) 2-epimerase (NeuC) converts UDP-GlcNAc to N-acetylmannosamine (ManNAc) from which sialic acid is formed by the action of Neu5Ac synthase (NeuB). Next, the sialic acid can be further modified in different ways by the action of Neu5Ac O-acetyltransferase (NeuD), or CMP-sialic

acid synthetase (NeuA) which produces an intermediate that could be used from polysialyltransferase (NeuS) for polysialic acid production which this can further be modified by polysialic acid O-acetyltransferase (NeuO) (Fig. 29).

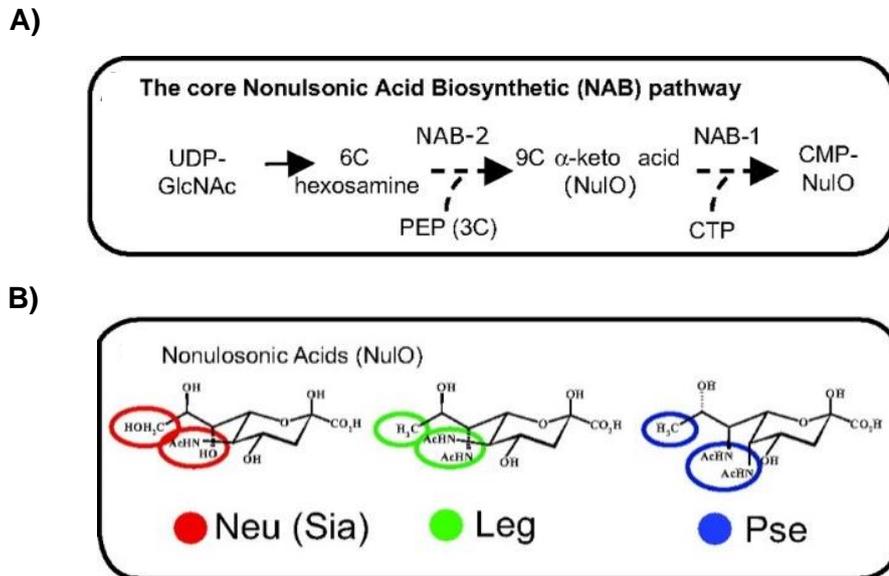


Figure 28 | Nonulosonic acid family.

A) The core biosynthetic pathway of nonulosonic acids. NAB-1 and NAB-2 are the general enzymes involved in the nonulosonic acid biosynthesis. Nonulosonic acid biosynthesis (NAB), uridine diphosphate *N*-acetylglucosamine, (UDP-GlcNAc), phosphoenolpyruvate (PEP), cytidine triphosphate (CTP), cytidine monophosphate nonulosonic acid (CTP-NuIO). B) Three members of the nonulosonic family, sialic acid (Neu), legionaminic acid (Leg) and pseudaminic acid (Pse). Ovals indicate their structural differences. (Taken from Lewis *et al.* 2009²⁵⁵).

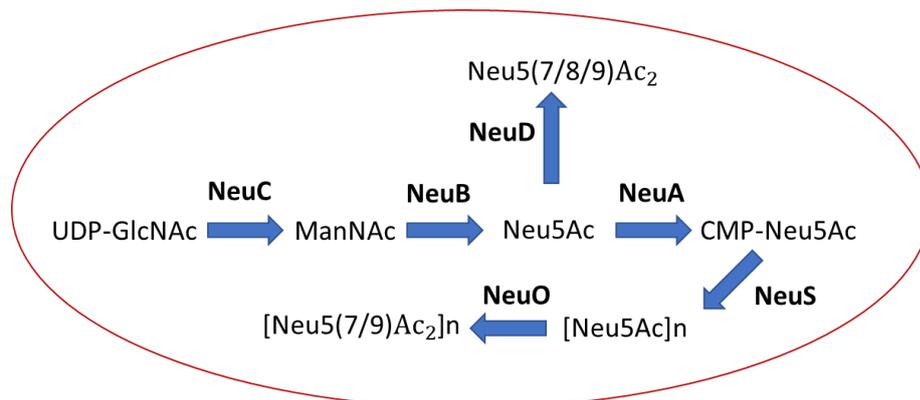


Figure 29 | De novo sialylation pathway in bacteria.

The sialic acid synthesis and modifications in *E. coli* K1 bacterial model. UDP-GlcNAc, uridine diphosphate-*N*-acetylglucosamine; NeuC, hydrolyzing UDP-GlcNAc epimerase; ManNAc, *N*-acetylmannosamine; NeuB, Neu5Ac synthase; Neu5Ac, *N*-acetylneuraminic acid or sialic acid; NeuA, CMP-Neu5Ac synthetase CMP-Neu5Ac, cytidine monophospho-*N*-acetylneuraminic acid; NeuD, Neu5Ac O-acetyltransferase; NeuS, polysialyltransferase; NeuO, polysialic acid O-acetyltransferase.

Siglec-7 has been shown to interact with *Campylobacter jejuni* in a sialic acid-dependent manner through the rough type of LPS which lacks the O-antigen, called lipooligosaccharides (LOS)¹⁰⁵, and with the β -protein of GBS in a non-sialic acid manner⁹⁸. However, most studies have focused on unravelling the molecular interactions of Siglec-7 with host ligands, while the molecular interactions with bacterial ligands and their effect on the host response remain largely unknown.

Having shown that the colorectal cancer-associated bacteria, *F. nucleatum* ssp. bound to Siglecs and specifically to Siglec-7-Fc, this chapter investigated the molecular mediators of this interaction focusing on *F. nucleatum* ssp. LPS and outer membrane vesicles (OMVs).

4.2 Results

4.2.1 *F. nucleatum* ssp. binds to Siglec-7 V-set in a sialic acid-independent manner

Here we used flow cytometry to investigate the type of interaction between recombinant Siglec-7-Fc and *F. nucleatum* ATCC 25586, ATCC 10953, and ATCC 51191 strains, as described in Materials and Methods (section 2.3.3). First, we conducted inhibition assays using the ganglioside GD3 (Neu5Ac2,8Neu5Ac2,3Gal β 1,4Glc), a known ligand of Siglec-7 which binds to the V-set domain¹⁰², with *F. nucleatum* ATCC 25586 and ATCC 51191 strains. The results showed that GD3 led to a significant reduction of *F. nucleatum* ATCC 51191 and ATCC 25586 binding to Siglec-7-Fc by approximately 92% and 87%, respectively, as shown by the relative mean fluorescence intensity (R-MFI) values (Fig. 30). These results suggest that the V-set domain of Siglec-7 is involved in the interaction with *F. nucleatum* ssp.

Bacteria+Ab
Siglec-7-Fc
Siglec-7-Fc + GD3

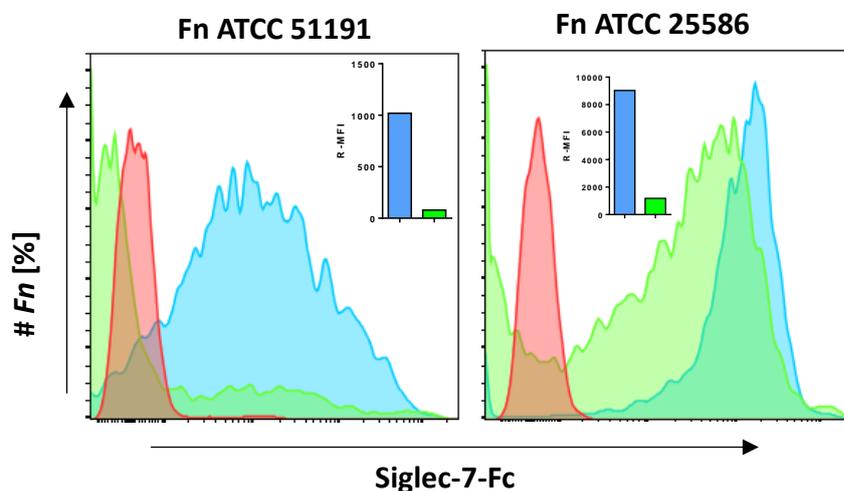


Figure 30 | Histograms of *F. nucleatum* ssp. and Siglec-7-Fc inhibition assays with GD3.

Binding assays of *F. nucleatum* ATCC 25586 or ATCC 51191 to Siglec-7-Fc in the presence of GD3 (in green) or not (in blue) using flow cytometry. Fn, *F. nucleatum*; R-MFI, relative mean fluorescence intensity. The bacteria incubated with secondary antibody only was used as the control (in red).

We next investigated the role of sialic acid as a potential bacterial ligand of Siglec-7 interaction. We first used the plant lectin *Sambucus nigra* I lectin (SNA), a sialic acid-recognition lectin which shows specificity towards α 2,6-sialylated structures²⁵⁷, to detect the presence of sialic acids on *F. nucleatum* ATCC 51191 or ATCC 10953 cell surface. Our results showed that approximately 57% and 31% of the *F. nucleatum* ATCC 51191 and ATCC 10953 population, respectively, bound to SNA lectin (Fig. 31), suggesting the potential presence of α 2-6-sialylated glycoconjugates on the cell surface.

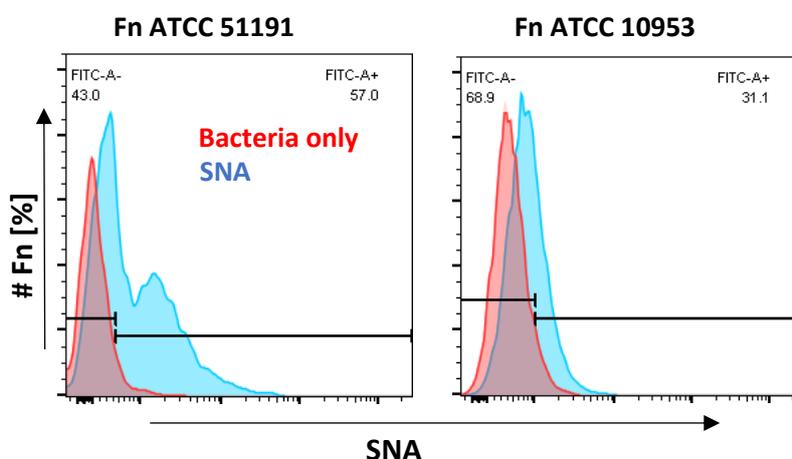


Figure 31 | Histograms of SNA binding to *F. nucleatum* ATCC 51191 or ATCC 10953.

Binding assays of *F. nucleatum* ATCC 10953 or ATCC 51191 to SNA lectin (in blue) using flow cytometry. Fn, *F. nucleatum*. As a control was used the bacteria in the absence of lectin (in red).

Our bioinformatic analysis of the Neu5Ac biosynthetic pathway in *F. nucleatum* ssp. genomes (Materials and Methods, section 2.6), showed the presence of all Neu5Ac biosynthetic enzymes (*neuC*, *neuB*) and the *neuA*, *neuD* modification enzymes for *F. nucleatum* ATCC 10953, while only *neuB* was found in *F. nucleatum* ATCC 25586 genome (Fig. 32). However, no genes involved in Neu5Ac biosynthetic could be identified in *F. nucleatum* ATCC 51191 partial genome (Fig. 32). These results are consistent with the presence of internal sialic acid in *F. nucleatum* ATCC 10953 LPS²⁵⁸, whereas, *F. nucleatum* ATCC 25586 LPS contains a sialic acid-like molecule called fusaminic acid²²⁰ and *F. nucleatum* ATCC 51191 contains uronic acids²⁴⁷, but no sialic acid or sialic-acid like sugar.

	NeuO	NeuS	NeuC	NeuA	NeuB	neuD	
Fn 10953	Red	Red	Green	Green	Green	Green	Not identified
Fn 25586	Red	Red	Red	Red	Green	Red	YES - identified
Fn 51191	Red	Red	Red	Red	Red	Red	

Figure 32 | Table showing the presence of the Neu5Ac biosynthetic cluster enzymes.

Bioinformatic analysis of *F. nucleatum* ATCC 10953, ATCC 25586, and ATCC 51191 genomes was conducted by BLAST to identify the presence or absence of Neu genes (colour-coded). Fn, *F. nucleatum*.

To further investigate whether the binding of *F. nucleatum* ssp. to Siglec-7 was due to the presence of sialic acid on the bacterial cell surface, the *F. nucleatum* ATCC 25586, ATCC 51191, and ATCC 10953 cells were treated with Neuraminidase A, a sialidase with broad specificity to (α 2-3,6,8,9) sialyl linkages, cleaving linear and branched non-reducing terminal sialic acid residues from glycoconjugates²⁵⁹. The sialidase treatment only led to a small reduction in the binding of *F. nucleatum* ssp. to Siglec-7 as indicated by the mean fluorescence intensity (MFI) values (Fig. 33). The enzymatic treatment together with our bioinformatics results may suggest that, in these strains, sialic acid may not be accessible to the sialidase, or that Siglec-7 can recognise other ligands on the bacterial cell surface such as sialic acid-like molecules or other sugars, which would be unaffected by the sialidase treatment.

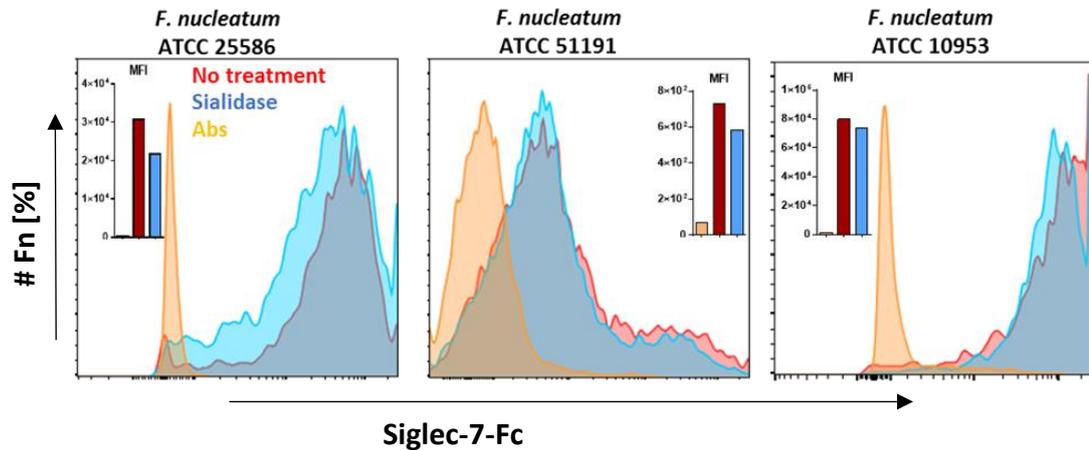


Figure 33 | Histograms of Siglec-7-Fc binding to sialidase treated *F. nucleatum* ssp. Binding assays of *F. nucleatum* ATCC 25586, ATCC 10953, or ATCC 51191 to Siglec-7-Fc following sialidase treatment (in blue) or untreated (in red) using flow cytometry. Fn, *F. nucleatum*; MFI, mean fluorescence intensity. As a control was used the bacteria incubated with antibody only (in orange).

4.2.2 *F. nucleatum* ssp. LPS bound to Siglec-7

To further identify the Siglec-7 ligand(s) present on *F. nucleatum* cell surface, LPS was purified from *F. nucleatum* strains and tested for binding to recombinant Siglec-7-Fc by ELISA-based binding assay, as described in Materials and Methods (section 2.3.4).

In collaboration with Profs Alba Silipo and Cristina De Castro (University of Napoli), *F. nucleatum* ATCC 51191, ATCC 25586 and ATCC 10953-derived LPS were extracted and purified by hot phenol/water and enzymatic digestion. The LPS structure of *F. nucleatum* ATCC 10953²¹⁹ and *F. nucleatum* ATCC 25586²²⁰ but not *F. nucleatum* ATCC 51191 is known. Therefore the structural characterisation of *F. nucleatum* ATCC 51191-derived LPS²⁴⁷ was carried out using a combination of GC-MS, MALDI and NMR analyses. We found that *F. nucleatum* ATCC 51191 O-antigen repeating unit contains glucose, fucose, and uronic acids, and shows 60% acetylation (Fig. 34A). However, no sialic acid residues were identified. This in contrast to *F. nucleatum* ATCC 10953²¹⁹ and *F. nucleatum* ATCC 25586²²⁰, which LPS structures contain internal sialic acids or sialic acid-like sugars, respectively. In addition, LPS analysis by silver stain SDS-PAGE of LPS extracted from the 3 *F. nucleatum* strains showed that the apparent size of *F. nucleatum*-derived O-antigen from all 3 *F. nucleatum* strains was

smaller compared to *E. coli*-derived O-antigen, with *F. nucleatum* ATCC 51191-derived O-antigen showing the smallest size and least heterogeneity, as compared to the *F. nucleatum* ATCC 25586 and ATCC 10953 O-antigens (Fig. 34B).

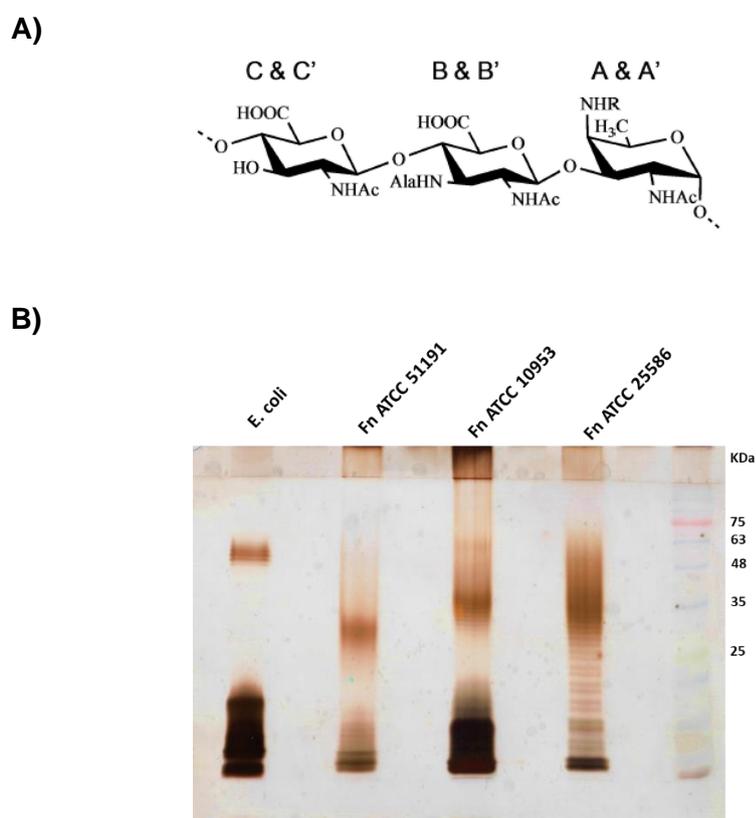


Figure 34 | O-antigen characterisation from *F. nucleatum* ssp.-derived LPS. A) O-antigen monosaccharide characterisation of *F. nucleatum* ATCC 51191-derived LPS, 4)- β -d-GlcpNAcA-(1-4)- β -d-GlcpNAc3NAlaA-(1-3)- α -d-FucpNAc4NR-(1, structure, A, B and C is assigned to R=acetyl residues, and A', B' and C' is assigned to R=H residues (taken from Garcia-Vello *et al.* 2020). B) LPS were analysed on a 12% SDS-PAGE and silver-stained. The gel is indicating the apparent sizes of the O-antigen extracted from *E. coli* O127:B8 (Lane 1), *F. nucleatum* ATCC 51191 (Lane 2), ATCC 10953 (Lane 3), or ATCC 25586 (Lane 4) LPS. As a marker was used the BLUeye Prestained Protein Ladder (GeneDireX).

Further, the LPS purified from *F. nucleatum* ATCC 51191 and ATCC 10953 was tested for its binding to recombinant Siglec-7-Fc by ELISA-based assay. The results showed that both *F. nucleatum* ATCC 51191 and ATCC 10953-derived LPS bound to Siglec-7-Fc but not to the Siglec-9-Fc used as a negative control (Fig. 35).

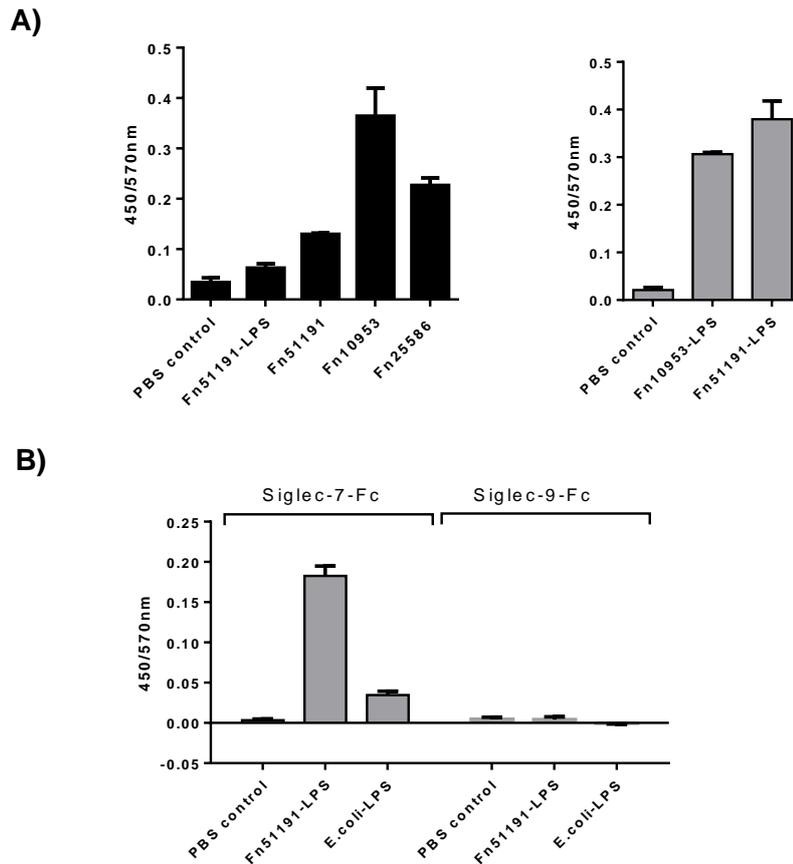


Figure 35 | ELISA-based binding assays of *F. nucleatum*-derived LPSs and Siglec-7-Fc.

F. nucleatum ATCC 51191-derived LPS, whole *F. nucleatum* ATCC 51191, ATCC 10953, ATCC 25586, or *F. nucleatum* ATCC 10953-derived LPS were tested against A) Siglec-7-Fc and B) Siglec-7-Fc or Siglec-9-Fc. Fn, *F. nucleatum*; PBS only or *E. coli*-derived LPS were used as negative controls.

To further confirm the contribution of LPS in the interaction between *F. nucleatum* ssp. and Siglec-7, we investigated the inhibition ability of *F. nucleatum* ATCC 51191-derived LPS in the binding of Siglec-7-Fc to the whole *F. nucleatum* ATCC 25586 or 10953 strains. While pre-incubation of Siglec-7-Fc with *F. nucleatum* ATCC 51191-derived LPS, led to a reduction of *F. nucleatum* ATCC 10953 binding to Siglec-7 by only 20% (Fig. 36), the binding of Siglec-7 to *F. nucleatum* ATCC 25586 showed a reduction by approximately 60%, based on the R-MFI values (Fig. 36), suggesting a role of LPS as a ligand of Siglec-7.

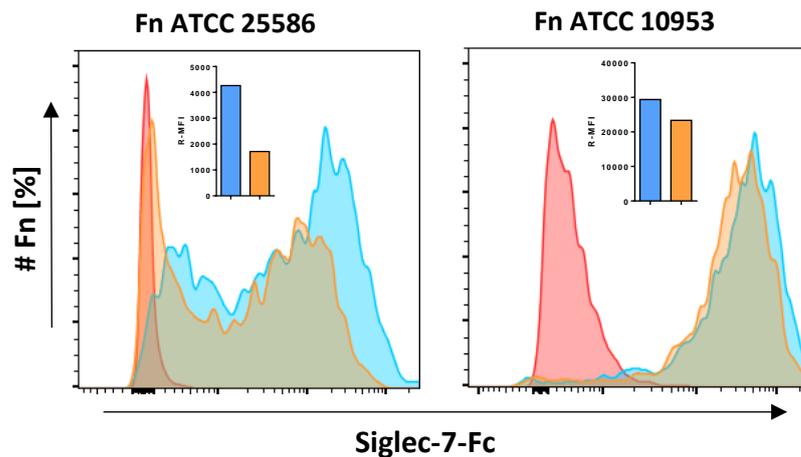


Figure 36 | Histograms of *F. nucleatum* ssp. and Siglec-7-Fc inhibition assays with LPS.

Binding assays of *F. nucleatum* ATCC 10953 or ATCC 25586 to Siglec-7-Fc in the absence (in blue) or presence (in orange) of *F. nucleatum* ATCC 51191-derived LPS using flow cytometry. Fn, *F. nucleatum*; R-MFI, relative mean fluorescence intensity. As a control was used the bacteria incubated with antibody only (in red).

To identify the *F. nucleatum* LPS epitopes involved in the recognition of Siglec-7-Fc, qualitative STD-NMR assays were performed by our collaborators at the University of Napoli (Fig. 37). The results showed that Siglec-7-Fc bound to the partially depolymerised O-antigen (OPS) domain of *F. nucleatum*-derived LPS, while the nature of the OPS-derived epitopes varied depending on the *F. nucleatum* strain. For *F. nucleatum* ATCC 10953-derived OPS, Siglec-7-Fc showed binding to the sialic acid (N) and its neighbour fucosamine (B) residue through the *N*-acetyl group and the methyl protons at position 6 (Fig. 37A). Interestingly, all the sugar residues quinovosamine (A), altrose (B) and fusaminic acid (X) contained in *F. nucleatum* ATCC 25586-derived OPS, showed binding to Siglec-7-Fc (Fig. 37B). For *F. nucleatum* ATCC 51191-derived OPS, Siglec-7-Fc showed binding mainly to fucosamine (A) and the alanine-containing sugar (B) (Fig. 37C).

Overall, these results suggest that the O-antigen domain of *F. nucleatum* ATCC 10953, ATCC 25586, and ATCC 51191-derived LPS is a ligand for Siglec-7, and that the binding epitopes are strain-specific involving sialic acid, sialic-acid-like, and non-sialic acid residues.

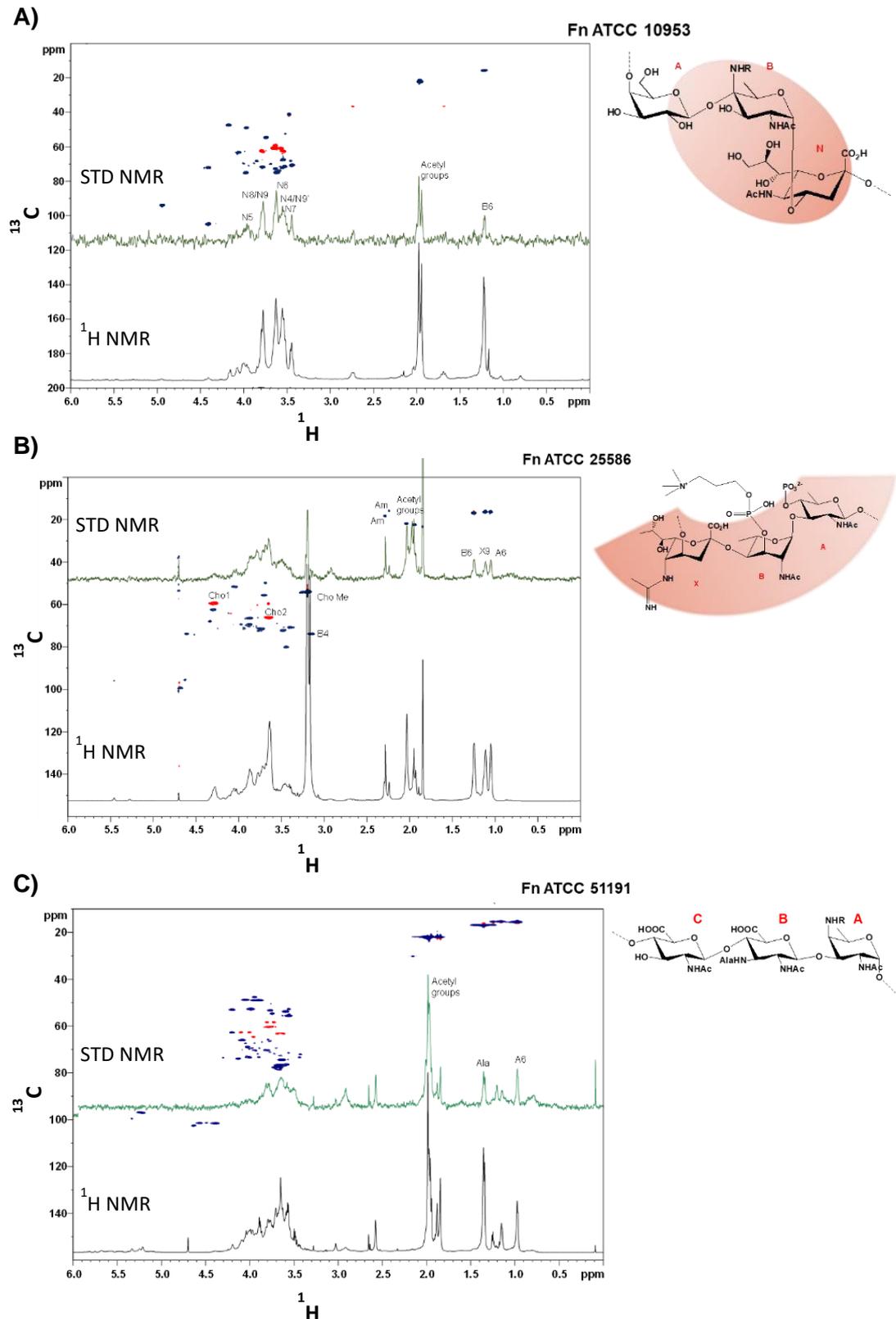


Figure 37 | Chromatograms of *F. nucleatum* ssp. and Siglec-7-Fc binding by STD-NMR.

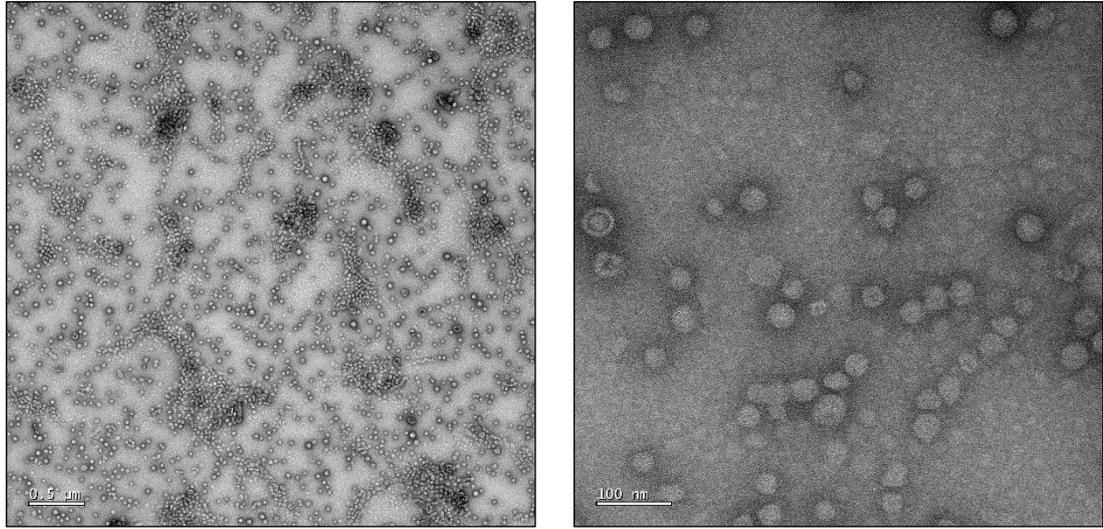
Superimposition of the STD 1D NMR Siglec-7 spectrum (green) in the presence of partially depolymerised OPS and the ^1H NMR spectrum (black) as reference, of *F. nucleatum* A) ATCC 10953, B) ATCC 25586 and C) ATCC 51191. The ^1H - ^{13}C HSQC spectrum of -CH and -CH₃ (in blue) and -CH₂ (in red) groups is also shown in each chromatogram.

4.2.2.1 *F. nucleatum* ssp. *OMVs bound to Siglec-7*

Here, we characterised the *F. nucleatum*-derived OMVs and investigated their binding to recombinant Siglec-7-Fc. *F. nucleatum* ATCC 10953, ATCC 51191 and ATCC 25586-derived OMVs were purified from bacteria grown at OD_{600nm} 0.6-1.2 by density gradient ultracentrifugation and characterised by transmission electron microscopy (TEM) and by nanoparticle tracking analysis using Nanosight. The microscopy images showed that *F. nucleatum* OMVs appeared spherical, as shown for *F. nucleatum* ATCC 25586-derived OMVs in Fig. 38A. Furthermore, the nanoparticle tracking analysis showed that the diameter of all *F. nucleatum*-derived OMVs ranged from 30 to 250 nm (Fig. 38B). The majority of *F. nucleatum* ATCC 10953 and ATCC 51191-derived OMVs showed a diameter of approximately 130-140 nm, while the majority of *F. nucleatum* ATCC 25586-derived OMVs showed a bigger size of approximately 170 nm with some particles reaching up to 390 nm, which may be the result of OMV clustering.

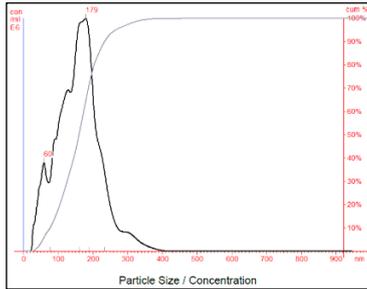
A)

F. nucleatum ATCC 25586

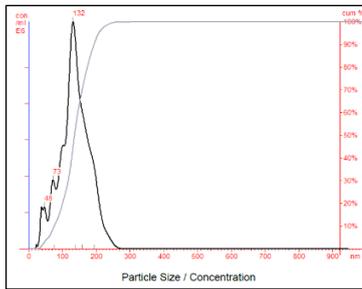


B)

F. nucleatum ATCC 25586



F. nucleatum ATCC 51191



F. nucleatum ATCC 10953

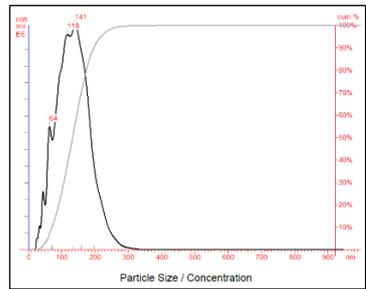


Figure 38 | *F. nucleatum*-derived OMV particle analysis.

A) *F. nucleatum* ATCC 25586-derived OMVs were visualised by TEM and B) *F. nucleatum* ATCC 25586, ATCC 10953 and ATCC 51191-derived OMVs particle size was analysed by NanoSight. The main peak corresponds to the mode of the population.

Since OMVs are derived from the outer membrane of Gram negative bacteria, the presence of LPS in *F. nucleatum* ATCC 10953, ATCC 51191 and ATCC 25586-derived OMVs was determined by our collaborators using GC-MS for acetylated O-methyl glycoside and total fatty acid composition analysis. The monosaccharide compositional analysis showed that all *F. nucleatum*-derived OMVs carried the main LPS-containing glycan, 3-deoxy-d-manno-oct-2-ulosonic acid (KDO) (Fig. 39A). For the lipid compositional analysis, the myristic acid, C14:0, contained in Lipid A domain of LPS was semi-quantified in relation to phospholipids, palmitic acid, C16:0, and stearic acid, C18:0 (Fig. 39B), by calculating the ratio of C14:0 (LPS) moles to C16:0 and C18:0 (phospholipids) moles. Together, these results showed that OMVs consist of approximately 60-70% LPS in relation to phospholipids.

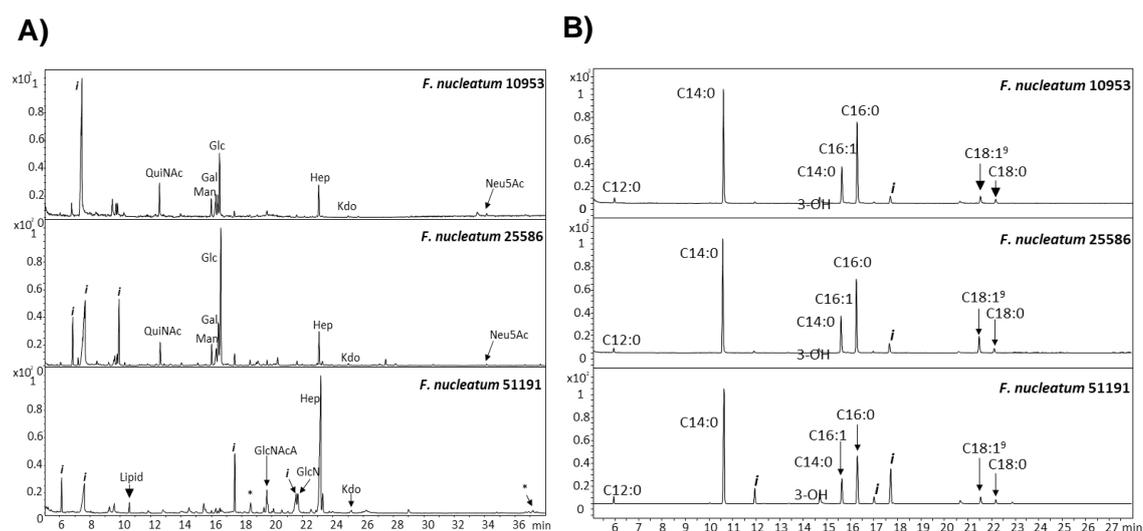
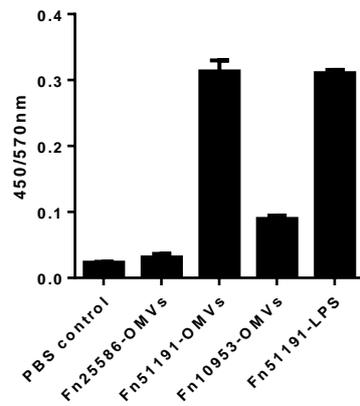


Figure 39 | Chromatograms of *F. nucleatum*-derived OMVs from GC-MS analysis. *F. nucleatum* ATCC 25586, ATCC 10953 and ATCC 51191-derived OMVs A) Acetylated O-methyl glycoside analysis and B) total fatty acid composition analysis. Simple and saturated fatty acids are named as C_x:0, x being the total number of carbons of the fatty acid; C-3 hydroxylated fatty acids as C_x:0 (3-OH). * unknown monosaccharide, *i*: impurity.

Next, the interaction between Siglec-7 and purified OMVs from *F. nucleatum* strains were analysed by ELISA-based binding assays or biolayer interferometry, as described in Materials and Methods in section 2.3.4 or 2.3.5, respectively. The ELISA-based binding assays showed that *F. nucleatum*-derived ATCC 25586, ATCC 51191 and ATCC 10953-derived OMVs bound to recombinant Siglec-7-Fc whereas no significant binding was observed to the negative controls (Fig. 40A). It is of note that *F. nucleatum* ATCC 51191-derived LPS was found to bind to levels comparable to that obtained with the whole *F. nucleatum* ATCC 51191-derived OMVs included in the assay (Fig. 40A).

Binding of *F. nucleatum* ATCC 25586-derived OMVs to Siglec-7-Fc was also investigated by biolayer interferometry (BLI) (Fig. 40B). Binding between soluble OMVs and Siglec-7-Fc immobilised on biosensors resulted in a change of biosensor's thickness and the generation of a wavelength shift (nm) of the reflected white light. A shift in wavelength was also observed for GD3, a known ligand for Siglec-7, indicating binding to Siglec-7-Fc (Fig. 40B). In contrast, a very low wavelength shift was observed using the negative controls, PBS (instead of OMVs) with Siglec-7-Fc immobilised sensors, or the biosensors without immobilised Siglec-7-Fc with *F. nucleatum* ATCC 25586-derived OMVs, supporting the specificity of the interaction between soluble OMVs and immobilised Siglec-7-Fc (Fig. 40B).

A)



B)

GD3 (+ Control)
Fn25586-OMVs
PBS only
(- Control)
Fn25586-OMVs
(- Control)

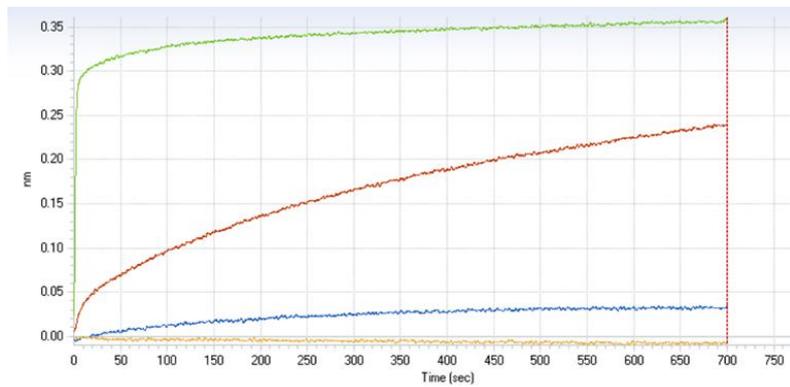


Figure 40 | *F. nucleatum*-derived OMVs binding to Siglec-7-Fc.

A) Binding of *F. nucleatum* (ATCC 10953, ATCC 25586, ATCC 51191)-derived OMVs, and ATCC 51191-derived LPS to soluble Siglec-7-Fc by ELISA. PBS (instead of LPS or OMVs) was used a negative control. B) Binding of *F. nucleatum* ATCC 25586-derived OMVs to immobilised Siglec-7-Fc (in red) or to the control a-hFc biosensors with no immobilised Siglec-7-Fc (in blue) by bi-layer interferometry. PBS (instead of OMVs) (in orange) or GD3 (in green) were used as negative and positive controls, respectively. Fn, *F. nucleatum*.

4.3 Discussion

The interaction of Siglec-7 with host sialylated glycoconjugates has been extensively studied, whereas very little is known of its interaction with bacteria. In this chapter, we investigated the molecular receptors mediating the interaction between Siglec-7-Fc and *F. nucleatum* ssp., including the *F. nucleatum* ssp. derived LPS and OMVs.

Using competitive inhibition assays, we showed that the interaction between *F. nucleatum* ATCC 25586, ATCC 10953 and ATCC 51191 and Siglec-7 occurs through the V-set domain of Siglec-7, the sialic acid-recognition domain for sialylated structures such as GD3²⁴⁹. We next used a combination of enzymatic treatment and structural analysis assays to gain insights into the nature of the bacterial ligands recognised by Siglec-7.

We showed that treating the bacteria with a sialidase had little effect on the binding of the 3 *F. nucleatum* strains to Siglec-7-Fc. This could be due to the nature of *F. nucleatum* LPS since internal sialic acid and fusaminic acid are present in LPS of strains ATCC 10953²¹⁹ and ATCC 25586²²⁰, respectively, therefore these sugars may not be recognised or being accessible to the sialidase. To date the structure of *F. nucleatum* ATCC 51191 LPS had not been reported. Here, *F. nucleatum* ATCC 51191 LPS was purified and the O-antigen characterised using chemical and NMR/MS approaches. We showed that the *F. nucleatum* ATCC 51191 LPS-derived O-antigen did not contain any sialic acid or sialic acid-like sugar but contained the active sugar group, uronic acid, consistent with the results of the binding assays following sialidase treatment. *F. nucleatum* ATCC 51191 O-antigen, [-4)- β -D-GlcpNAcA-(1-4)- β -D-GlcpNAc3NAIaA-(1-3)- α -D-FucpNAc4NR-(1-)]²⁴⁷, represents a new structure for bacterial LPS which could contribute to a distinct host immune response. In tumour tissues, *F. nucleatum* ATCC 51191 is present in higher levels when compared to other strains¹⁹⁶, therefore, further investigation is needed to identify the pro-tumour potentials of *F. nucleatum* ATCC 51191 O-antigen.

Using ELISA-based and BLI binding assays, we confirmed that *F. nucleatum* ATCC 51191 or ATCC 10953-derived LPS could bind to Siglec-7-Fc. The STD-NMR analysis provided further insights into the bacterial O-antigen epitopes involved in the recognition of *F. nucleatum* ssp. by Siglec-7, showing that the

bacterial sugar epitopes were strain-dependent. For *F. nucleatum* ATCC 10953 and ATCC 25586, sialic acid and fusaminic acid were among the sugars found to interact with Siglec-7 while for *F. nucleatum* ATCC 51191, the interaction appeared to be mediated by alanine-containing uronic acid. In addition, fucosamine, common to both *F. nucleatum* ATCC 10953 and ATCC 51191 LPS was also found to interact with Siglec-7. Together these data confirmed that the O-antigen domain of *F. nucleatum* ssp. LPS interacts with Siglec-7 and identified novel bacterial ligands for Siglec-7, including the fusaminic acid, fucosamine, quinovosamine, and altrose, in addition to the known sialic acid ligand. Moreover, the O-antigen characterisation of *F. nucleatum* extracted LPS, showed variations in the apparent sizes and patterns across all three *F. nucleatum* strains as compared to *E. coli* O-antigen. These structural differences may contribute to differences in the Siglec-7-mediated immunomodulatory properties of the strains. Environmental changes in bacterial growth can cause alteration of lipid A and O-antigen length of LPS and a concomitant acceleration of OMV production²⁶⁰.

Since LPS is a known component of OMVs in Gram negative (Gram-) bacteria, we purified OMVs from *F. nucleatum* ATCC 25586, ATCC 10953 and ATCC 51191 strains, and confirmed that LPS was present in *F. nucleatum* OMVs. To our knowledge LPS abundance in OMVs from Gram- bacteria is unknown. Here we showed, using semiquantitative MS analysis, that LPS constitutes around 60-70% of *F. nucleatum* OMVs. The diameter-range of OMVs was similar across *F. nucleatum* ATCC 51191, ATCC 10953 and ATCC 25586 strains (30-250 nm), in agreement with the expected typical OMV size range for Gram negative bacteria¹⁵⁸. To date, only OMVs from *F. nucleatum* ssp. *animalis* strain²¹⁵ have been characterised showing a smaller diameter-range (40-100 nm), as compared to the *F. nucleatum* strains studied here. The same study conducted proteomic analysis of OMVs revealing the presence of multiple virulence factors which could be involved in *F. nucleatum*-derived pathogenicity²¹⁵. In addition, a recent study showed that OMVs derived from *F. nucleatum* ssp. *nucleatum* DSM 15643 - ATCC 23726 strain interact with epithelial cells in a TLR-2 depending manner, resulting in NF- κ B transcription factor activation triggering the immune response²¹⁶. Here, we showed for the first time that OMVs extracted from *F. nucleatum* ssp. could interact with host lectins. *F. nucleatum* ssp. OMVs bound

to recombinant Siglec-7-Fc, supporting our data showing the binding of Siglec-7 to the whole bacteria and isolated LPS.

Together the flow cytometry, ELISA and STD-NMR binding assays confirmed the interaction of *F. nucleatum* ssp. to Siglec-7 and identified LPS epitopes as novel ligands for Siglec-7. In addition, we characterised for the first time the OMVs extracted from *F. nucleatum* ATCC 25586, ATCC 10953 and ATCC 51191 and showed interaction with Siglec-7, recapitulating the LPS-mediated binding observed with the whole bacteria.

In the next chapter, we will investigate the immunomodulatory properties of *F. nucleatum* ssp. and their derived LPS and OMVs on human immune cells.

CHAPTER 5

-

**Deciphering the effects of *F. nucleatum*
ssp. on myeloid cell phenotype**

5.1 Introduction

F. nucleatum ssp. have been implicated in tumour progression through their ability to recruit tumour infiltrating immune cells¹⁹⁵. To date, research on the interplay between *F. nucleatum* and the host in CRC has mainly focused on the interaction with host's epithelial cells, including colonic cell lines^{196,213,261} or spheroids²⁰⁴. These studies investigated *F. nucleatum* ssp. attachment and invasion, induction of colonic cell proliferation, and the bacteria effect on pro-inflammatory cytokine secretion which can be involved in metastasis.

Studies on *F. nucleatum* ssp. and immune cell interactions are limited, and the results are dependent on the myeloid cell type and bacterial strain tested. In human monocytic cell lines, *F. nucleatum* ssp. *animalis* (ATCC 51191), the most prevalent strain in CRC, has been shown to induce CCL20 expression and monocyte migration¹⁹⁶, and in human neutrophil-like cell lines, the same subspecies induce a pro-inflammatory cytokine profile²⁶². Xue *et al.* 2018 showed that human THP-1-derived macrophages acquire a M1-macrophage phenotype following *F. nucleatum* ATCC 25586 stimulation²⁶³, while Chen *et al.* 2018 showed that stimulation of the mouse macrophage cell line RAW 264.7 with *F. nucleatum* ATCC 10953 led to an induction of M2-macrophage polarisation²⁰⁷. *F. nucleatum* ATCC 25586-derived LPS has been shown to induce IL-8 and β -defensin expression in oral epithelial cells²⁶⁴, and a pro-inflammatory phenotype in macrophage-like cells²¹¹. To date, only a few studies have investigated the effect of *F. nucleatum* ssp. on human immune primary cells, showing that *F. nucleatum* ssp. *nucleatum* promotes peripheral blood mononuclear cell (PBMC) death^{265,266} by causing PBMC aggregation²⁶⁶.

In this chapter, we investigate the effect of *F. nucleatum* ATCC 10953, ATCC 51191 and ATCC 25586 strains and their derived outer membrane vesicles (OMVs) and LPS, on the cytokine profile and cell surface marker expression of human primary myeloid cells or the U937 macrophage cell line.

5.2 Results

5.2.1 Effect of *F. nucleatum* ssp. on myeloid cells

In order to investigate the immunomodulatory properties of *F. nucleatum* ATCC 25586, ATCC 10953 and ATCC 51191 strains, we first sought to identify the most accurate way to quantify *F. nucleatum* cells. We investigated three different methods for counting bacteria cells besides the common spectrophotometry approach (based on OD_{600nm} reading) where an OD_{600nm} at 1 corresponds to 10⁹ bacteria/ml²⁶⁷. Additionally, since *F. nucleatum* strain variations have been reported to impact host response²⁶², we characterised the morphology of *F. nucleatum* ATCC 25586, ATCC 10953 and ATCC 51191 strains. We then determined the internalisation of *F. nucleatum* ssp. in immune cells and the effect of the strains on cell cytokine production and cell surface marker expression in host immune cells.

5.2.1.1 Optimisation of *F. nucleatum* strain quantification

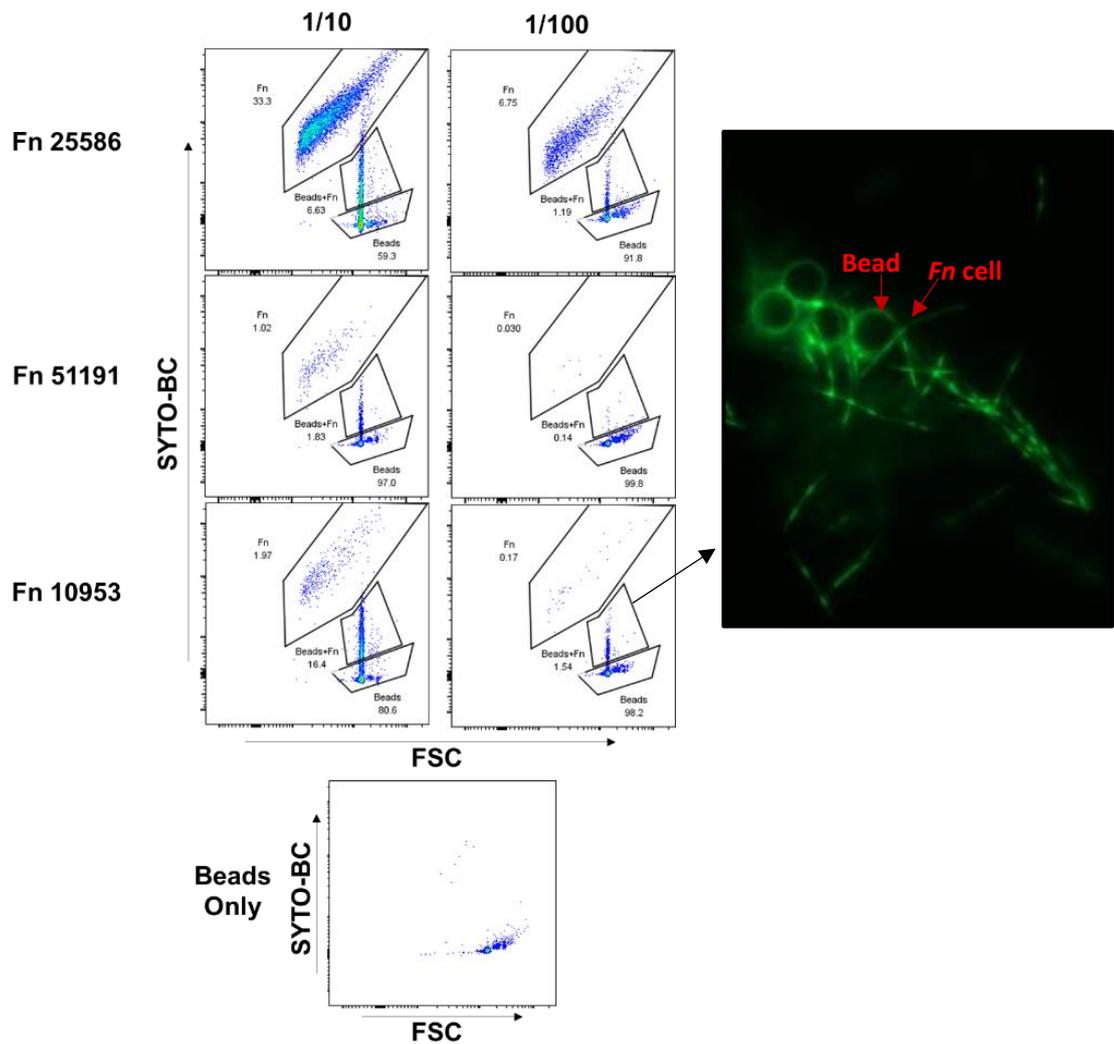
Three approaches were tested to select a reliable method for the quantification of *F. nucleatum* cells grown *in vitro*. We first used flow cytometry, in which SYTO-BC fluorescently-labelled bacteria and non-stained beads (of known density) were mixed and analysed by flow cytometry. Here, bacteria (SYTO-BC+) and beads (SYTO-BC-) should appear as two distinct populations in the SYTO-BC versus forward side scatter (FSC) dot plot. However, our results showed an intermediate population when using different bacterial dilutions, suggesting binding of bacteria to the beads (Fig. 41A). This observation was confirmed by fluorescence microscopy (Fig. 41A), preventing the use of this method for quantification of *F. nucleatum* cells.

Next, we attempted to quantify *F. nucleatum* cell numbers by DNA quantification. Here, bacterial genomic DNA was extracted from *F. nucleatum* ssp. PFA-fixed, and the extracted DNA was quantified as described in Materials and Methods (section 2.1.3). The *F. nucleatum* cell numbers were determined by applying the equation shown in Fig. 41B showing similar values to that obtained based on the OD_{600nm} measurements (Fig. 41B).

Finally, imaging flow cytometry (IFC) was used to identify non-stained bacteria densities in bright field (Fig. 42A). This technique quantifies the bacterial cells passing in a fluid stream and identifies the bacterial density, based on the specific

volume and controls beads the system withdraws. The results showed *F. nucleatum* cell numbers similar to that obtained using the OD_{600nm} (Fig. 42B). Based on these results, in the rest of the study, quantification of *F. nucleatum* cells was primarily based on the OD_{600nm} value but in some cases where several bacterial washes were needed, the IFC values were used for confirmation.

A)



B)

$$\text{Bacterial cells: } \frac{6.022 \times 10^{23} \frac{\text{molecules}}{\text{mol}}}{\text{Chromosome length (bp)} \times \text{MW of bp} \left(\frac{\text{bp}}{\text{mol}} \right)} \times \text{extracted DNA (g)}$$

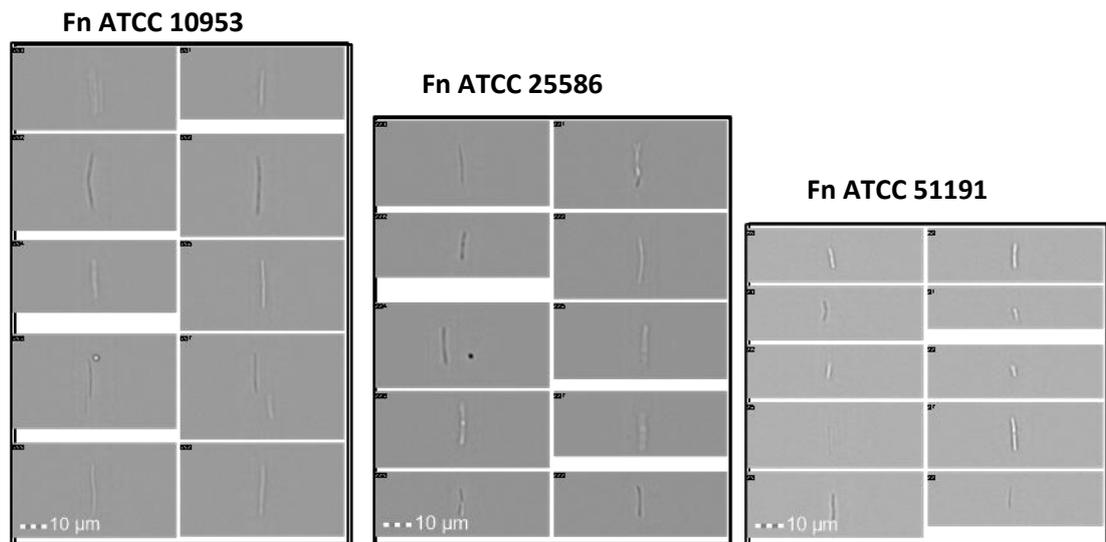
MW of bp: 650 g/mol (mean value)
 Avogadro: 6.022×10^{23} molecules/mol

Bacterial strain	#Cells after DNA extraction ($\times 10^8$)	#Cells based on OD _{600nm} ($\times 10^8$)
<i>Fn</i> ATCC 25586	5.45	2.70
<i>Fn</i> ATCC 10953	3.94	2.50
<i>Fn</i> ATCC 51191	4.10	3.00

Figure 41 | *F. nucleatum* cell quantification.

A) Flow cytometry analysis of SYTO-BC-*F. nucleatum* strains or the beads (dot plot of FSC versus SYTO-BC), and fluorescence microscopy image showing the binding of the strains to the beads. B) Comparison of *F. nucleatum* cell quantification values based on the extracted genomic DNA and OD_{600nm}. *Fn*, *F. nucleatum*.

A)



B)

Bacterial strain	#Cells based on IFC ($\times 10^8$)	#Cells based on OD _{600nm} ($\times 10^8$)
<i>Fn</i> ATCC 25586	3.00	5.00
<i>Fn</i> ATCC 10953	1.00	4.20
<i>Fn</i> ATCC 51191	0.60	2.00

Figure 42 | *F. nucleatum* cell quantification by imaging flow cytometry.

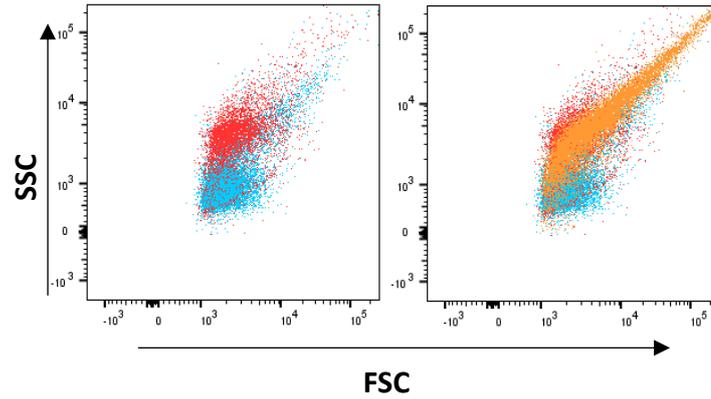
A) Bright field images of *F. nucleatum* strains. B) Comparison of *F. nucleatum* cell quantification between IFC and the OD_{600nm} values. *Fn*, *F. nucleatum*.

5.2.1.2 *F. nucleatum* strains show morphological differences

Flow cytometry can be applied to identify differences in the morphology of cells by analysing the plot of forward scatter (FSC) versus the side scatter (SSC)^{268,269}. In mammalian cells, the intensity of the FSC is an indicator of cell size, while the intensity of the SSC informs on cell granularity. Here, we observed morphological differences between *F. nucleatum* ATCC 10953, ATCC 25586, and ATCC 51191 strains, as indicated by the dot plot of FSC versus SSC (Fig. 43A). These differences suggest that *F. nucleatum* strains differ in size, while the observed population tail of strain *F. nucleatum* ATCC 25586 suggests auto-aggregation (Fig. 43A), as reported previously for this strain using spectrophotometry measurements at different time points²⁷⁰. These results were also confirmed by IFC and light microscopy (Fig. 43B). Further, using IFC we showed that *F. nucleatum* ATCC 10953 cell size ranged from 11 to 119 μm with a median size at 23 μm , while *F. nucleatum* ATCC 25586 cell size ranged from 4 to 51 μm with a median size at 15.5 μm , and *F. nucleatum* ATCC 51191 cell size ranged from 1 to 86.5 μm with a median size at 19.5 μm (Fig. 43B). These morphological differences between strains could be of importance for their immunogenic function.

A)

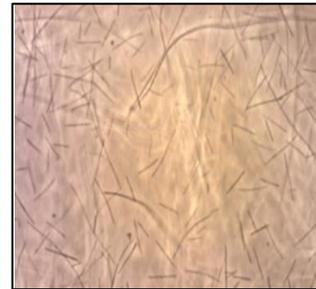
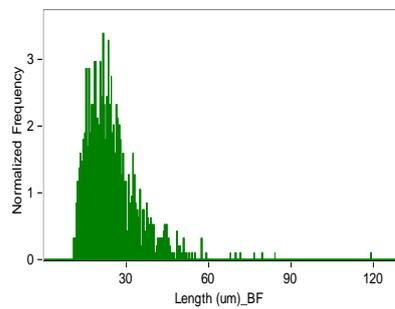
Fn ATCC 10953
Fn ATCC 51191



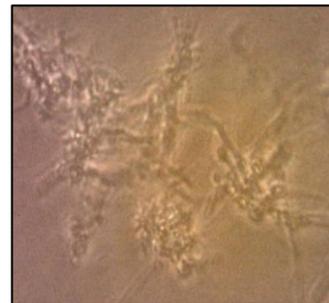
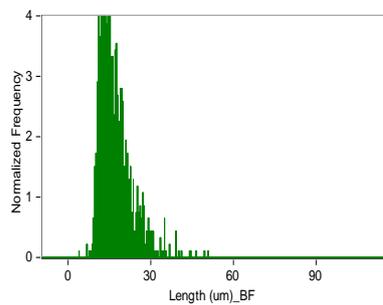
Fn ATCC 10953
Fn ATCC 51191
Fn ATCC 25586

B)

Fn ATCC 10953



Fn ATCC 25586



Fn ATCC 51191

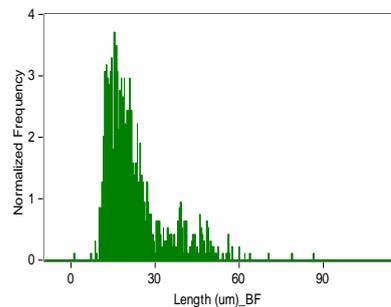


Figure 43 | *F. nucleatum* ssp. show differences in their morphology.

A) Superimposed FSC vs SSC dot plots of *F. nucleatum* ATCC 10953 (in red), ATCC 51191 (in blue), and ATCC 25586 (in orange) after flow cytometry analysis. B) Histograms of cell length measurements using IFC together with light microscopy images (40X objective) of the corresponding *F. nucleatum* strains. Fn, *F. nucleatum*.

5.2.1.3 *F. nucleatum* strains are internalised into myeloid cells

Here, we first investigated the capacity of *F. nucleatum* ATCC 10953, ATCC 25586 and ATCC 51191 to be internalised by primary myeloid strains. Human monocytes (CD14) were isolated from PBMCs obtained from human blood as described in Materials and Methods (section 2.4.2). Monitoring CD14 expression in isolated CD14+ and washed cells (CD14-) fractions by flow cytometry (Fig. 44A) showed high monocyte isolation efficacy as no CD14+ cells were detected in the fraction of washed cells (Fig. 44A). We also stained the crude PBMCs fraction before CD14+ isolation and our analyses showed that monocytes accounted for approximately 7% of the total PBMCs, close to the value of 10-20% reported in healthy humans²⁷¹.

The isolated monocytes were then differentiated into dendritic cells (moDCs) or macrophages (moMφs), as described in Materials and Methods (section 2.4.2). Differentiated cells have a distinct CD14 and DC-SIGN marker expression²⁷². Herein, we confirmed cell differentiation by flow cytometry, showing that the collected adherent moMφs expressed both CD14 and DC-SIGN, while both floating and adherent moDCs showed high DC-SIGN expression and downregulation of CD14 expression (Fig. 44B).

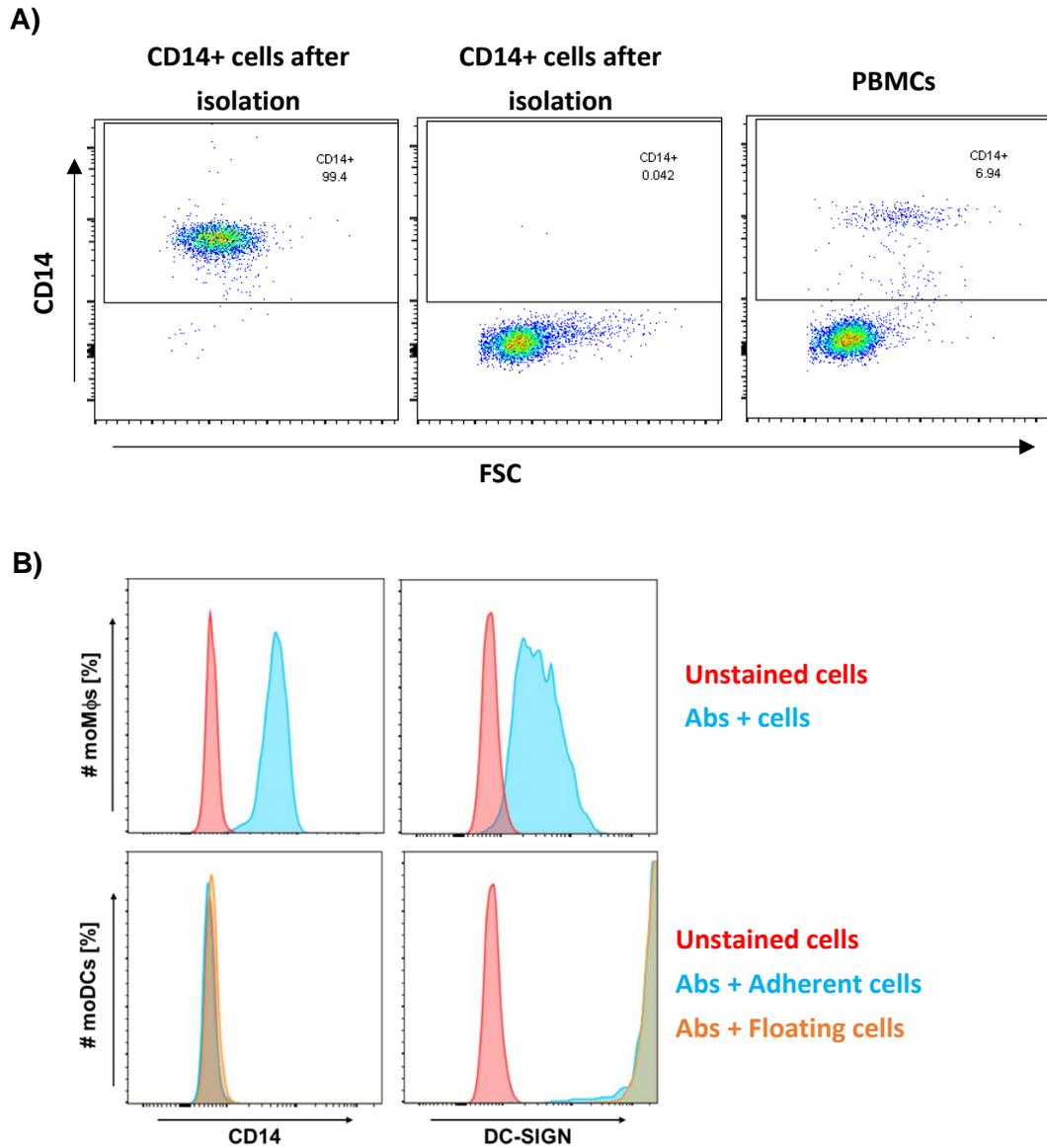
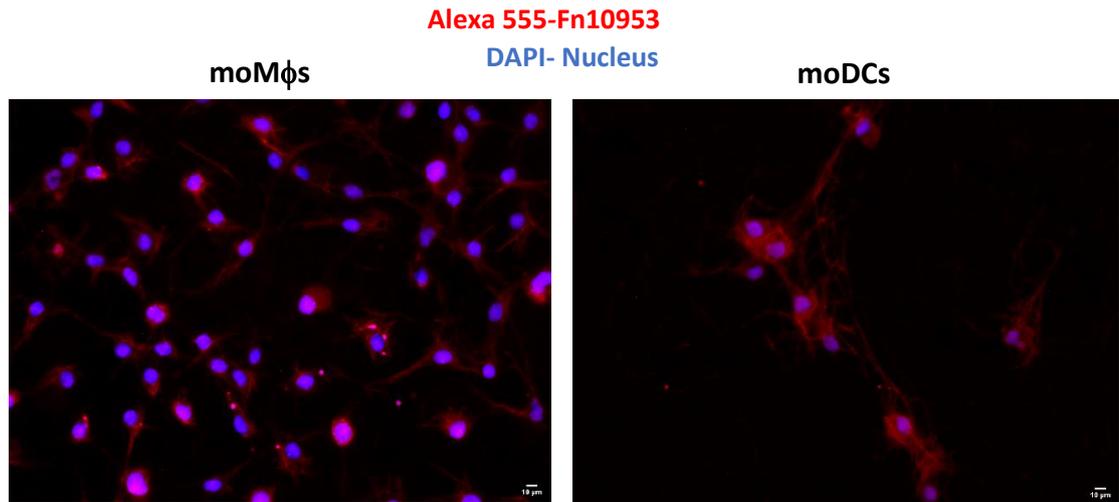


Figure 44 | Monocyte isolation and differentiation.

Expression of the CD14, monocyte marker, in PBMCs, and cell subsets (CD14+ or CD14-) after CD14+ kit isolation. B) Expression of CD14 and DC-SIGN on macrophages (moMφs) (in blue), and monocyte-derived dendritic cells (moDCs) adherent (in blue) or floating (in orange). Abs, antibodies. Unstained cells in red.

We then investigated the association and internalisation of *F. nucleatum* strains into moDCs and moMφs by fluorescence microscopy or IFC. The gut symbiont *R. gnavus* ATCC 29149, considered as a non-invasive strain, was used as a control. Bacterial cells stained with FISH for fluorescence microscopy, or labelled with FITC for IFC analysis, showed association with both moDCs or moMφs (Fig. 45). In addition, IFC analysis showed *F. nucleatum* internalisation (60-90%) in both primary immune cells (Fig. 46B). Internalisation of *F. nucleatum* was approximately 10-15% higher in moMφs (85-90%) as compared to moDCs (~75%), while *R. gnavus* ATCC 29149 showed low levels of internalisation (~15-20%) in both myeloid cells (Fig. 46A). Interestingly, the morphology of the internalised *F. nucleatum* strains varied across strains with *F. nucleatum* ATCC 10953 adopting a circular shape (Fig. 46B) which could be due to the length of this particular strain, as we previously reported (see section 5.2.1.2).

A)



B)

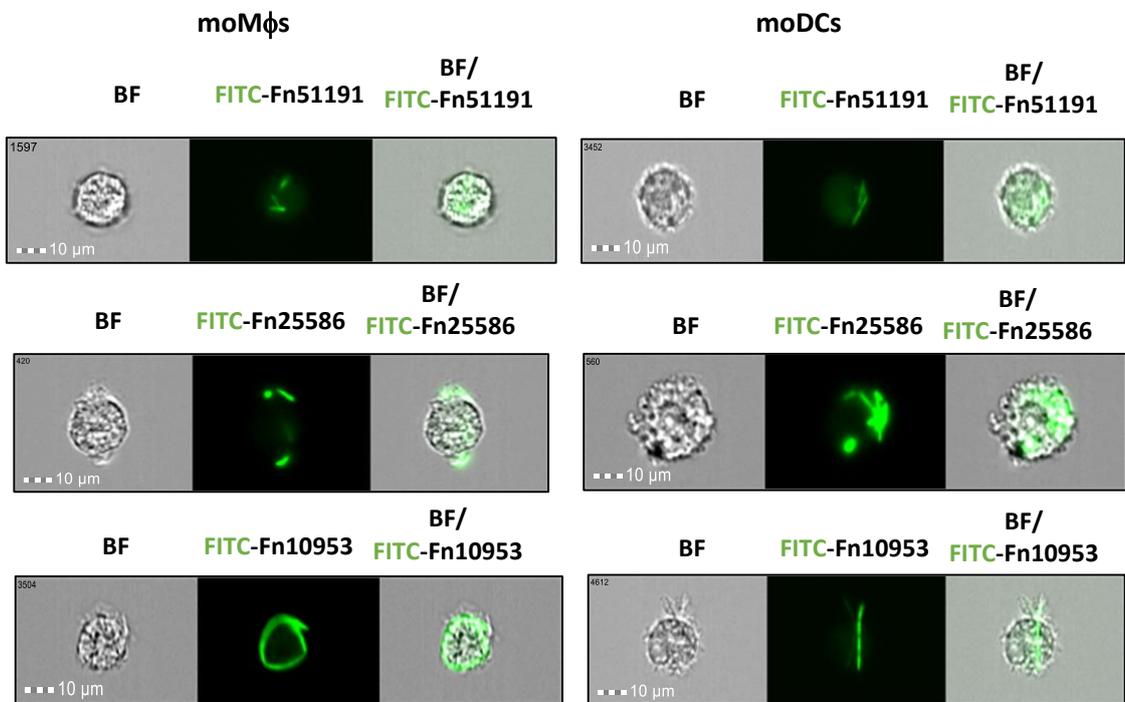


Figure 45 | *F. nucleatum* ssp. associates with moDCs and moMφs.

A) Microscopy images of *F. nucleatum* ATCC 10953 association with moDCs and moMφs. *F. nucleatum* ATCC 10953 is shown in red and moDCs or moMφs nucleus in blue. DAPI was used for nucleus staining and FISH-Alexa555 for bacterial staining. B) IFC images of *F. nucleatum* ssp. association with moMφs or moDCs. Microscopy and IFC images were taken with 40X objective. Fn, *F. nucleatum*.

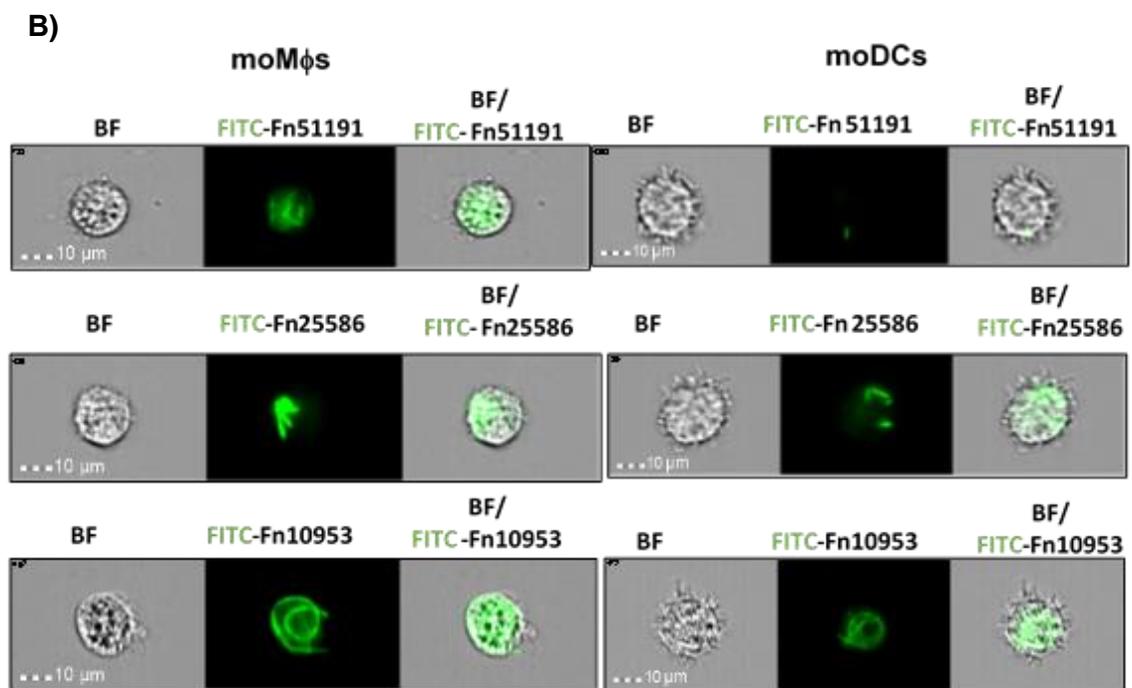
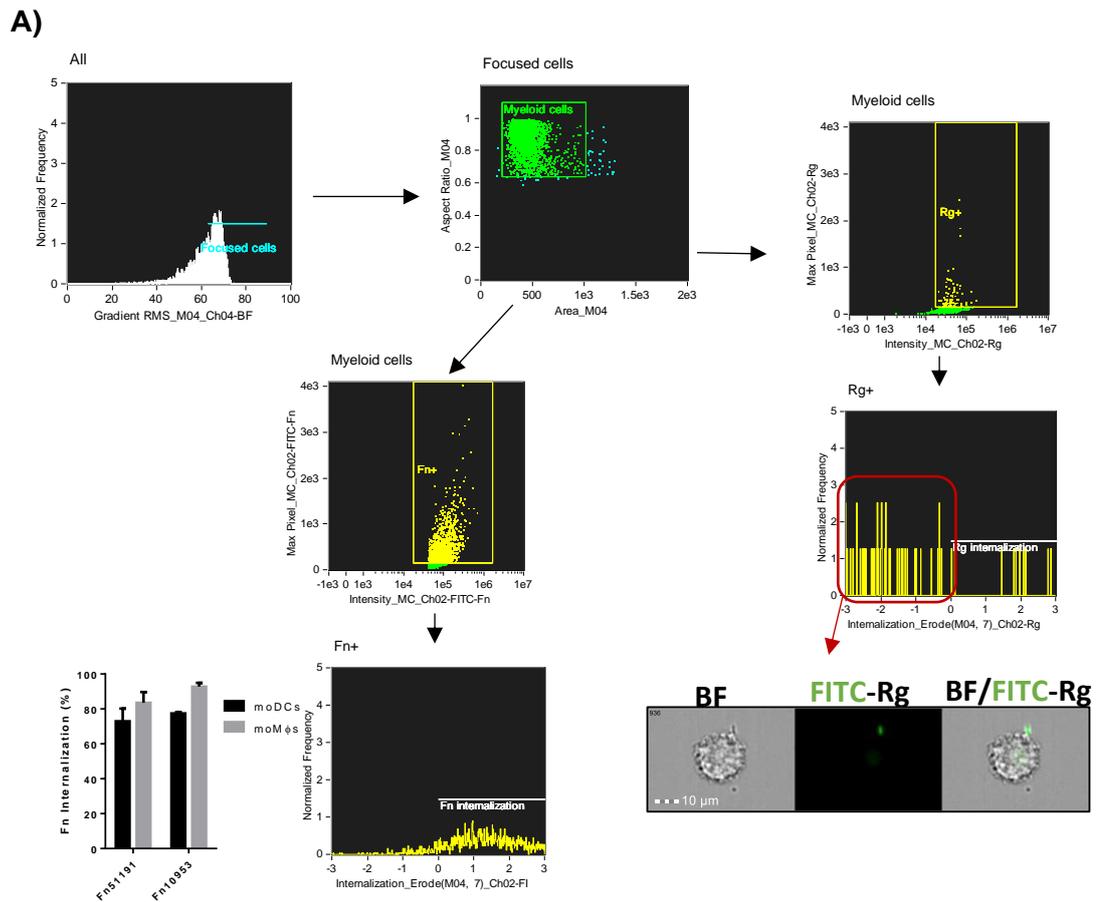


Figure 46 | *F. nucleatum* ssp. internalisation into myeloid cells by IFC.

A) The focused single myeloid cell population was gated and the internalised bacteria were identified in the high intensity of "maximum FITC pixel" population using IDEAS software. *R. gnavus* ATCC 29149 was used as a control. B) IFC images of the internalised *F. nucleatum* ssp. All images were taken with 40X objective. Fn, *F. nucleatum*; Rg, *R. gnavus*.

5.2.1.4 *F. nucleatum* strains modulate host cell phenotype in a cell subset manner

Next, we investigated the effects of *F. nucleatum* ATCC 25586, ATCC 10953 and ATCC 51191 strains on cytokine production and cell surface marker expression in moDCs and moMφs. The human-derived primary myeloid cells were stimulated with *F. nucleatum* bacterial cells at different time points (2 h, 6 h or 18 h) with a multiplicity of infection (MOI) of 5 or 50.

Stimulation of moDCs with *F. nucleatum* ssp. for 2 h did not significantly induce TNF α , IL-10 or IL-6 production (Fig. 47A). However, after a 6 h stimulation of moDCs with *F. nucleatum* ATCC 10953 or ATCC 25586 at both MOI 50 and 5, and with *F. nucleatum* ATCC 51191 at MOI 50, IL-6 production was significantly induced ($p < 0.0001$) as compared to the unstimulated control (Fig. 47B). Additionally, moDCs stimulation with *F. nucleatum* ATCC 10953 or ATCC 25586 at MOI 50 significantly induced TNF α production ($p < 0.0001$) as compared to the unstimulated medium control (Fig. 47B). Moreover, following a 18 h stimulation of moDCs at MOI 50 and 5 with *F. nucleatum* ATCC 25586, ATCC 10953 or ATCC 51191 strains led to a significant increase in the production of TNF α ($p < 0.0001$), and IL-6 ($p < 0.0001$) in a dose-dependent manner while at MOI 50 the 3 strains significantly induced IL-10 as compared to the unstimulated control (Fig. 47B).

Stimulation of moMφs with *F. nucleatum* ssp. at MOI 50 and 5 led to a significant induction of IL-10 production ($p < 0.0001$) after 6 h and 18 h stimulation (Fig. 47B) across all 3 strains, and for *F. nucleatum* ATCC 10953 and ATCC 25586, IL-10 was produced in a dose-dependent manner after 6 h, while only *F. nucleatum* ATCC 25586 showed dose-dependency after 18 h (Fig. 47B). Production of IL-6 was significantly induced ($p < 0.0001$) following stimulation at both MOI 50 and 5 with *F. nucleatum* ATCC 10953 for 6 h and 18 h (Fig. 47B), and with *F. nucleatum* ATCC 51191 after 18 h (Fig. 47B), while *F. nucleatum* ATCC 51191 showed a significant induction of IL-6 production ($p < 0.001$) at MOI 5 but not 50 after 6 h stimulation as compared to the unstimulated control (Fig. 47B). Interestingly, stimulation with *F. nucleatum* ATCC 10953 for 6 h at MOI 5 showed a significant increase in IL-6 levels ($p < 0.0001$) when compared to the stimulation at MOI 50 (Fig. 47B). No statistically significant induction of TNF α was observed following

stimulation of moM ϕ s with either of the three *F. nucleatum* strains as compared to the unstimulated control (Fig. 47B).

Additionally, stimulation of moDCs or moM ϕ s with *F. nucleatum* strains for 18 h led to a significant induction of IL-8 production as compared to the unstimulated control (data not shown).

In addition to human-derived moDCs and moM ϕ s, a macrophage-like monocytic cell line (U937) was used after differentiation with phorbol 12-myristate 13-acetate (PMA). Stimulation of PMA-differentiated U937 with *F. nucleatum* ATCC 10953, ATCC 51191, or ATCC 25586 at MOI 5 for 18 h showed a significant induction of IL-10 production while TNF α was produced at low levels (Fig. 48).

Overall, our results showed that all *F. nucleatum* strains at both MOI 50 or 5 induced IL-6 after 18 h stimulation in both moDCs or moM ϕ s, while TNF α production was only significantly induced in moDCs. IL-10 was significantly induced in moM ϕ s and only at MOI 50 in moDCs. A cytokine profile similar to that obtained in moM ϕ s was also observed in the *F. nucleatum*-stimulated macrophage-like monocytic cell lines (U937).

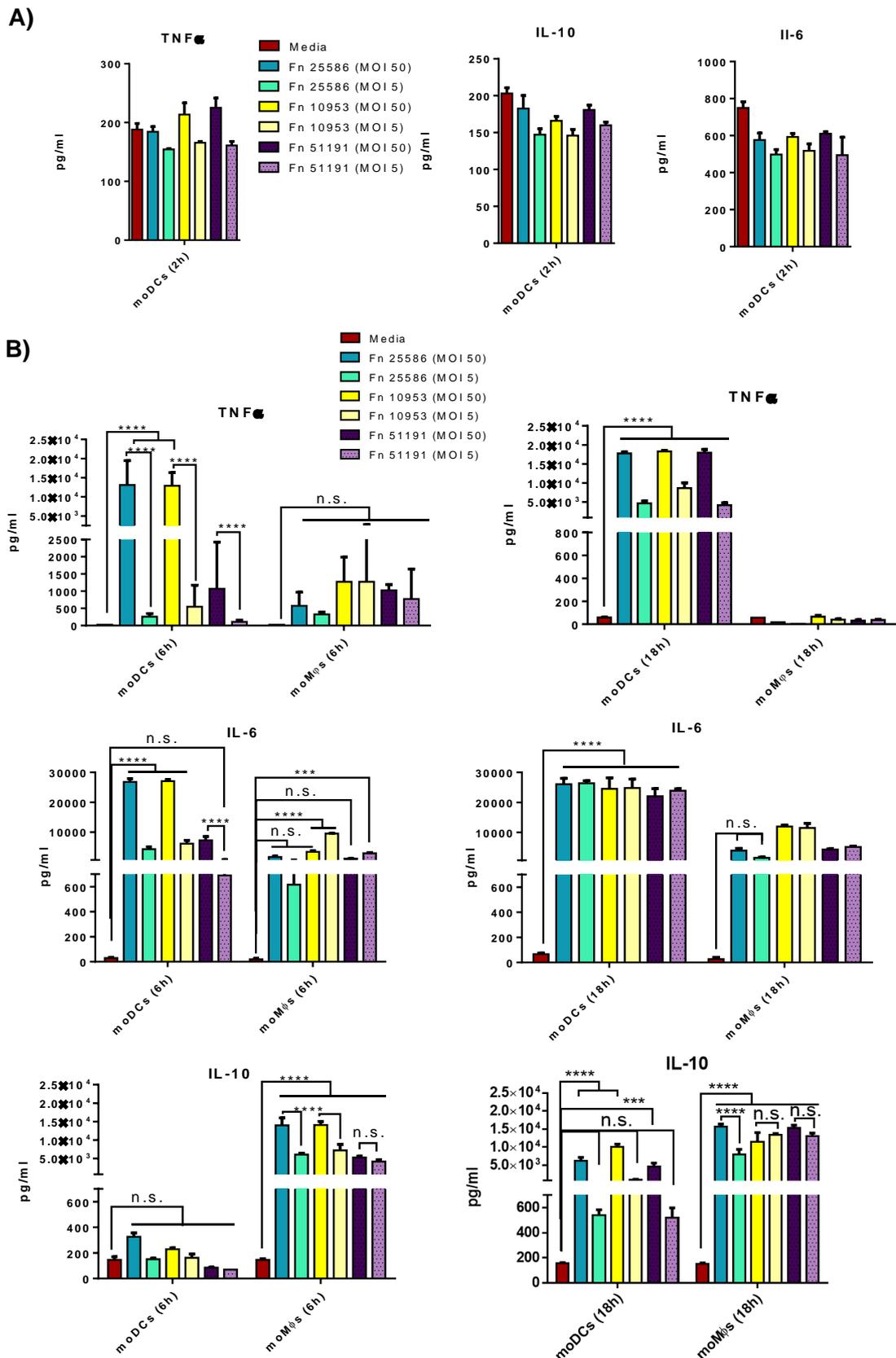


Figure 47 | Cytokine production of moDCs and moM ϕ s stimulated with *F. nucleatum* ssp.

A) Stimulation of moDCs for 2 h. B) Stimulation of moDCs or moM ϕ s for 6 h and 18 h. Fn, *F. nucleatum* strains were used at MOI of 5 or 50 as colour-coded. Unstimulated cells (Media in red) was used as a control. Data shown are the mean of triplicates from one representative experiment reproduced in two independent experiments. Statistics: *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001; n.s., not statistically difference.

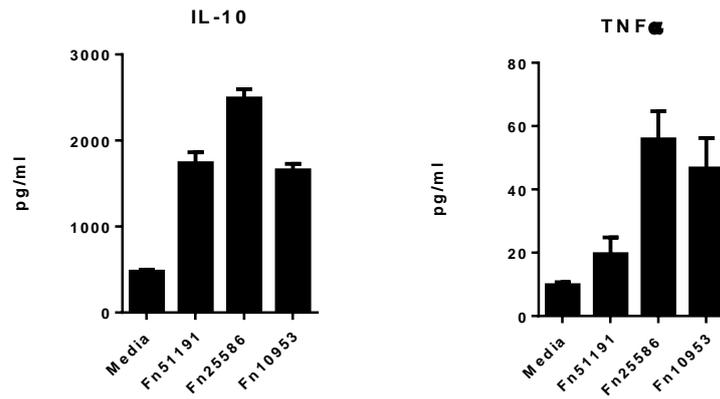


Figure 48 | Cytokine production of U937-PMA stimulated with *F. nucleatum* ssp. IL-10 and TNF α cytokine analysis of the macrophage-differentiated monocytic cell line U937 stimulated with *F. nucleatum*. Unstimulated cells (media) was used as a control. Fn, *F. nucleatum*. Data shown are the mean of triplicates from one representative experiment reproduced in three independent experiments.

We next investigated the expression of cell surface markers (CD80, CD86, PD-L1, CD163 and CD206) after stimulation of moDCs or moM ϕ s with *F. nucleatum* ATCC 25586, ATCC 10953 or ATCC 51191 strains at MOI 50 and 5 (Fig. 49). Stimulation of moDCs with *F. nucleatum* ssp. showed an induction of CD80, CD86 and PD-L1 marker expression for all 3 strains tested after 6 h and 18 h stimulation, with MOI 50 showing higher levels of marker expression when compared to MOI 5 (Fig. 49A).

Stimulation of moM ϕ s with *F. nucleatum* ssp. induced PD-L1 expression for all 3 strains tested after 6 h and 18 h stimulation, with MOI 50 showing increased PD-L1 levels as compared to MOI 5 (Fig. 49B). A significant reduction of CD86 expression was observed after 18 h for all three *F. nucleatum* strains at both MOI 50 and 5, as compared to the unstimulated control (Fig. 49B). The unstimulated moM ϕ s showed a reduction of CD163 and CD206 expression after 18 h, while stimulation with *F. nucleatum* ATCC 51191 at MOI 5 for 6 h showed the highest induction of CD163 and CD206 when compared to MOI 50, the other two strains, or the unstimulated control (Fig. 49B). Furthermore, CD206 expression was decreased following stimulation for 18 h with *F. nucleatum* ATCC 25586 or ATCC 10953 at MOI 5 and 50, or at MOI 50, respectively (Fig. 49B). Additionally, *F. nucleatum* ATCC 10953 induced the expression of CD206 only at MOI 5 when moM ϕ s were stimulated for 6 h (Fig. 49B).

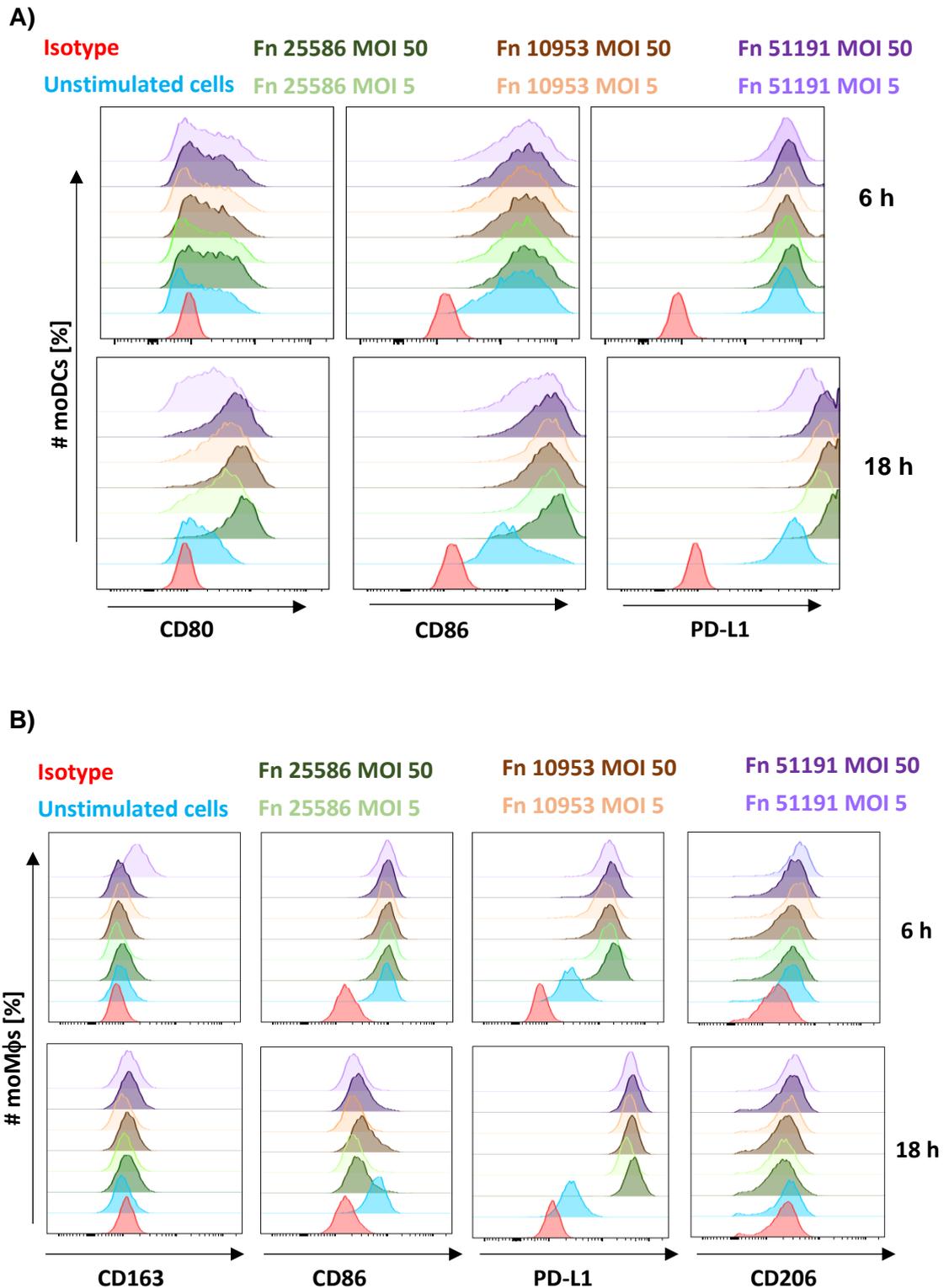


Figure 49 | Histograms of cell surface marker expression in moDCs and moMφs.
 A) moDCs and B) moMφs stimulated with *F. nucleatum* ATCC 10953, ATCC 25586 and ATCC 51191 strains at MOI 50 or 5 for 6 h or 18 h as colour-coded. Isotype control was used in the unstimulated cells (in red). Fn, *F. nucleatum*. Data shown are from one representative experiment reproduced in two independent experiments.

Given the observed differences in immune response depending on MOI, we next investigated the effect of *F. nucleatum* ssp. at MOI 50 or 5 on myeloid cell viability using propidium iodide (PI), an indicator of cell death, by flow cytometry. Under unstimulated state, moDCs showed an approximately 2-fold higher level of cell death as compared to moMφs (Fig. 50). Following *F. nucleatum* ssp. stimulation at MOI 50 or 5, the proportion of PI positive cells increased in moDCs when compared to the unstimulated control, while cell death remained unaffected following stimulation of moMφs. Stimulation with *F. nucleatum* ATCC 10953 showed the highest percentage of moDC death (Fig. 51) with a 14% increase between MOI 5 and MOI 50 (Fig. 51) while *F. nucleatum* ATCC 51191 or ATCC 25586 showed an approximately 5% increase in moDCs death at MOI 50 compared to MOI 5 (Fig. 51). These results are consistent with previous reports showing strain-dependent variations in cell death and function of neutrophils when treated with *F. nucleatum* ATCC 25586, ATCC 10953, or ATCC 49256 strains²⁶². In contrast to moDCs, stimulation of moMφs with *F. nucleatum* ssp. had little or no effect on cell viability (Fig. 52), as reported previously for *F. nucleatum* ATCC 25586 strain using the THP-1-derived macrophage cell line²⁶³. Based on this analysis, *F. nucleatum* strains were used at MOI 5 for 18 h in the rest of the study on human immune cells.

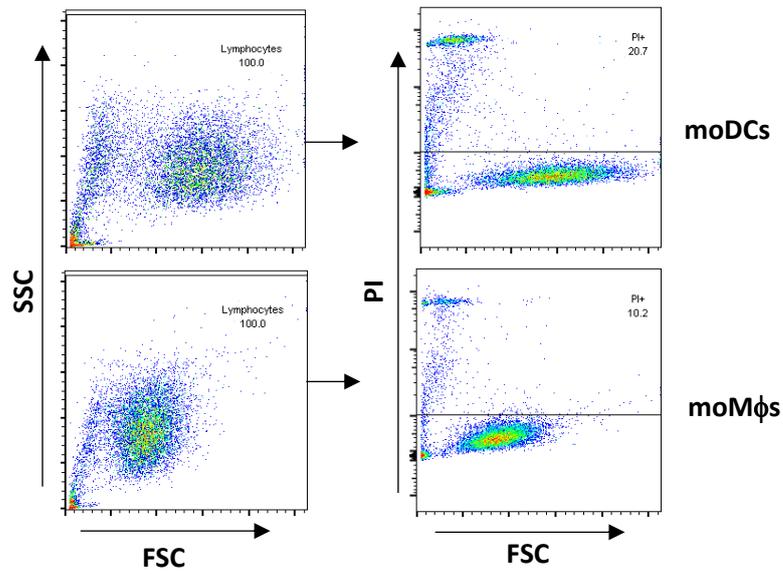


Figure 50 | Flow cytometry analysis of the proportion of dead moDCs or moMφs at unstimulated state.

On the left, the SSC versus FSC dot plots discriminate populations based on their granularity and size. On the right, the PI versus FSC dot plot demonstrates the proportions of dead and alive cells following staining with propidium iodide (PI). SSC, side scatter; FSC, forward scatter. Data shown are from one representative experiment reproduced in three independent experiments.

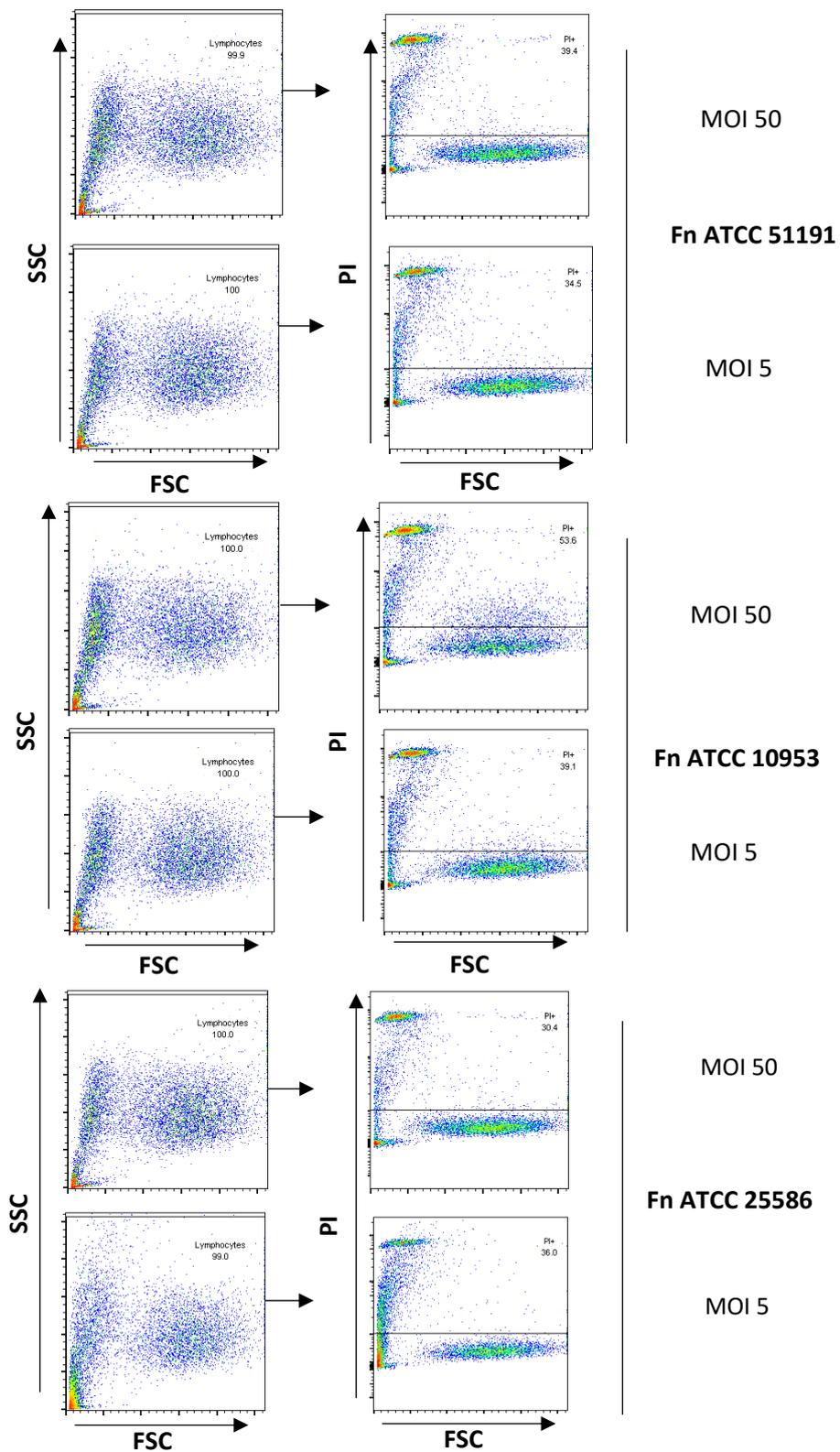


Figure 51 | Flow cytometry analysis of the proportion of dead moDCs after stimulation with *F. nucleatum* ssp.

MoDCs were stimulated with *F. nucleatum* ATCC 51191 or ATCC 10953 at MOI 50 or 5 and dead moDCs were analysed by flow cytometry following PI staining. Fn, *F. nucleatum*. Data shown are from one representative experiment reproduced in three independent experiments.

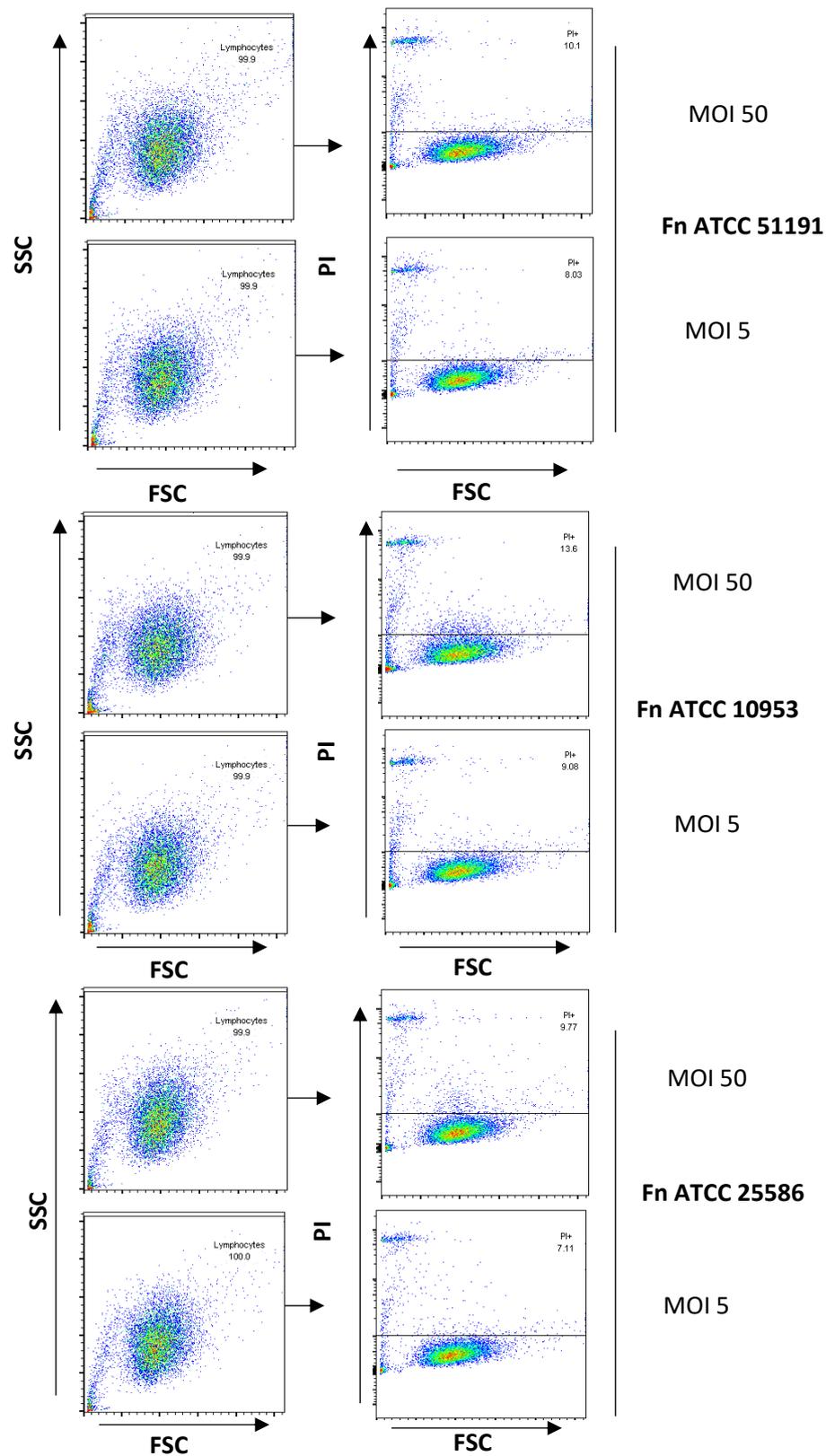


Figure 52 | Flow cytometry analysis of the proportion of dead moMφs after stimulation with *F. nucleatum* ssp.

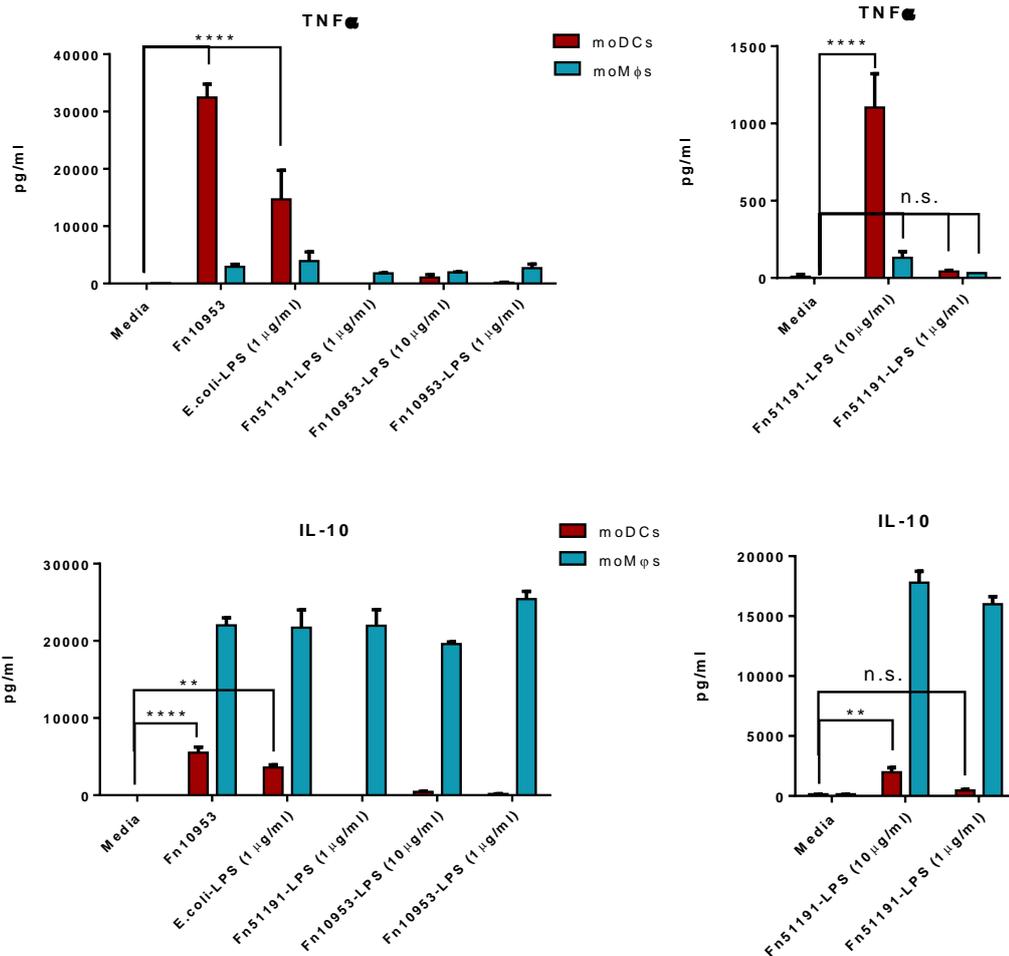
MoMφs were stimulated with *F. nucleatum* ATCC 51191 or ATCC 10953 at MOI 50 or 5 and dead moDCs were analysed by flow cytometry following PI staining. Fn, *F. nucleatum*. Data shown are from one representative experiment reproduced in three independent experiments.

5.2.2 Effect of *F. nucleatum* OMVs or LPS on myeloid cells

In order to gain insights into the molecular mediators of *F. nucleatum* interaction with immune cells, we tested the capacity of *F. nucleatum* ATCC 51191 or ATCC 10953-derived LPS or OMVs derived from *F. nucleatum* ATCC 51191, ATCC 10953, or ATCC 25586 strains to modulate human primary immune cell response.

We showed that a 18 h stimulation of moDCs with *F. nucleatum* ATCC 51191-derived LPS led to a significant induction of TNF α ($p < 0.0001$) and IL-10 ($p < 0.01$) at 10 $\mu\text{g/ml}$ but not 1 $\mu\text{g/ml}$ (Fig. 53A), while *F. nucleatum* ATCC 10953-derived LPS did not significantly induced TNF α and IL-10 cytokine production using the same conditions when compared to the unstimulated control (Fig. 53A). In contrast, stimulation of moM ϕ s with *F. nucleatum* ATCC 25586, ATCC 10953 or ATCC 51191-derived OMVs or LPS at 10 $\mu\text{g/ml}$ or 1 $\mu\text{g/ml}$, showed a significant induction of IL-10 only (Fig. 53). The observed cytokine profile was similar to that previously obtained with the whole bacteria at MOI 5 (Fig. 47, section 5.2.1.4). Additionally, stimulation of moM ϕ s with the control *E. coli*-derived LPS at 1 $\mu\text{g/ml}$ showed a cytokine profile similar to that obtained with *F. nucleatum*-derived LPS. In contrast to the *F. nucleatum*-derived LPS, stimulation of moDCs with *E. coli*-derived LPS induced the expression of TNF α ($p < 0.001$) and IL-10 ($p < 0.01$) (Fig. 53A).

A)



B)

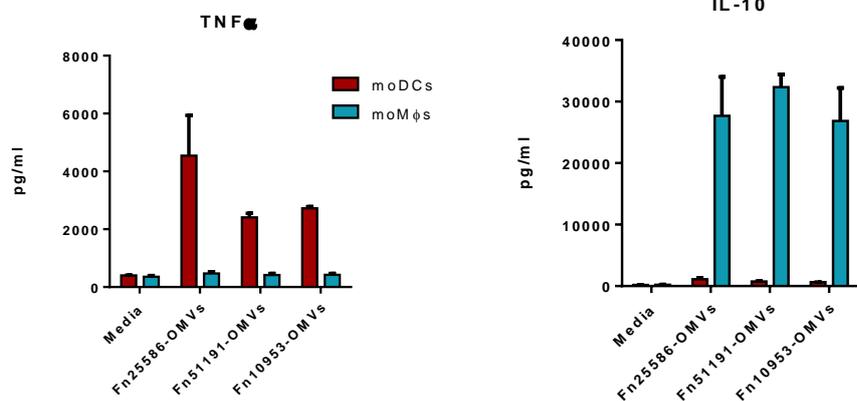


Figure 53 | Cytokine production of moDCs and moM ϕ s stimulated with *F. nucleatum* ssp.-derived OMVs or LPSs.

A) Myeloid cells, moDCs (in red), moM ϕ s (in blue) were stimulated with *F. nucleatum* ATCC 51191 or 10953-derived LPS at 10 or 1 μ g/ml or with B) *F. nucleatum*-ssp.-derived OMVs. Fn, *F. nucleatum*; Unstimulated cells (Media) was used as negative control. Data shown are the mean of triplicates from one representative experiment reproduced in three independent experiments. Statistics: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; n.s., not statistically difference.

We then investigated the effect of *F. nucleatum* ATCC 10953, ATCC 25586, or ATCC 51191-derived OMVs or LPS derived from *F. nucleatum* ATCC 10953, ATCC 51191 at 1 or 10 $\mu\text{g/ml}$ on cell surface marker expression of myeloid cells. The CD80 cell surface marker expression was induced when moM ϕ s or moDCs were treated with *F. nucleatum* ATCC 10953 or ATCC 51191-derived LPS or OMVs from all three strains, as compared to the unstimulated control, recapitulating the effect observed using the whole bacteria (Fig. 54).

In moDCs, stimulation with *F. nucleatum*-derived LPS showed an induction of CD86 expression, as also observed with the whole bacteria (Fig. 54A), whereas stimulation with *F. nucleatum*-derived OMVs showed a reduction of CD86 expression compared to the unstimulated control (Fig. 54A).

Stimulation of moM ϕ s with *F. nucleatum* ATCC 51191-derived LPS led to a reduction in CD86 expression, as also observed with the whole bacteria (Fig. 54B). In contrast, no consistent results could be obtained across biological replicates following stimulation of moM ϕ s with OMVs or *F. nucleatum* ATCC 10953-derived LPS with regards to CD86 expression.

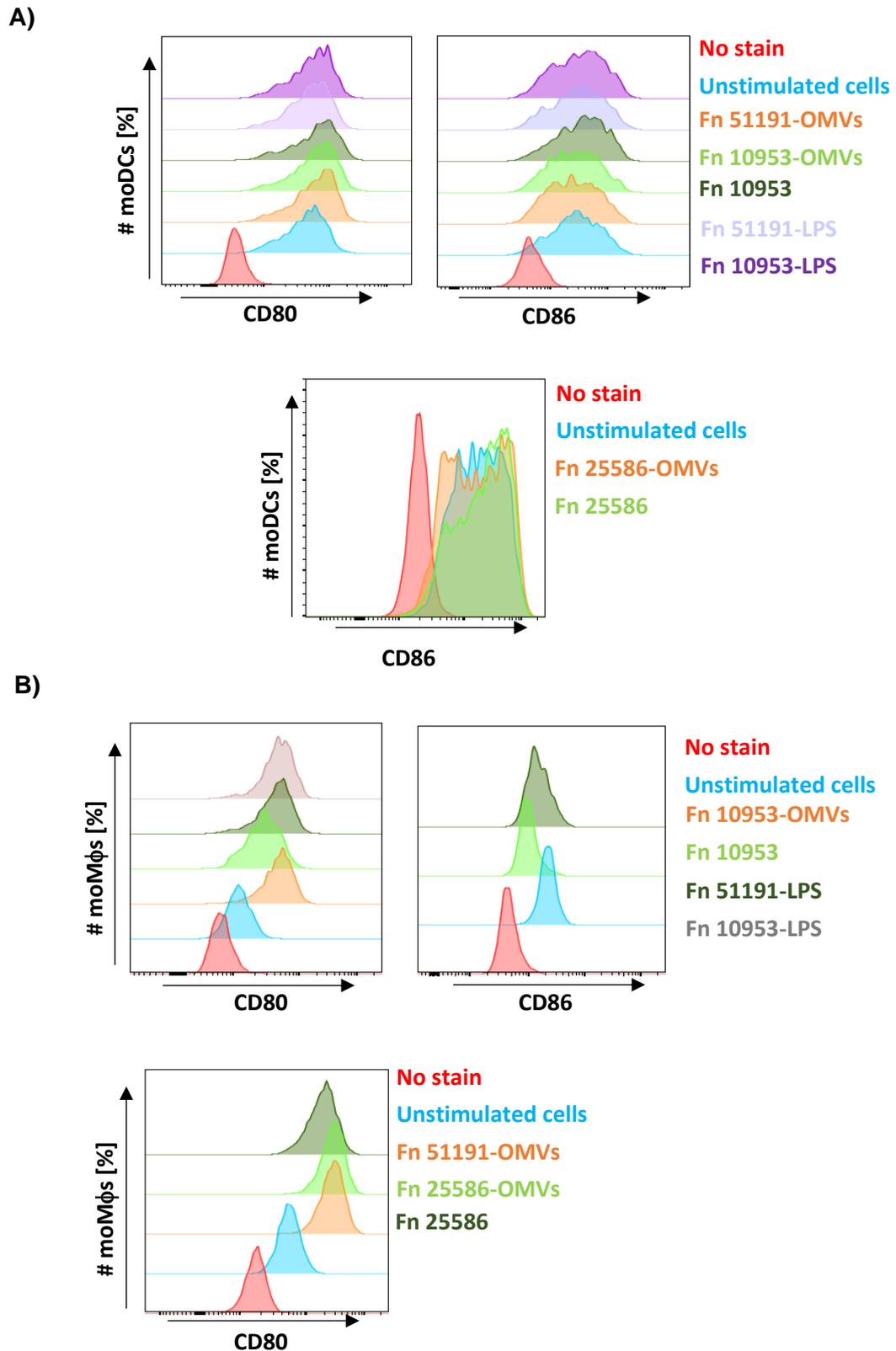


Figure 54 | Histograms of cell surface marker expression in moDCs and moMφs following *F. nucleatum* stimulation.

A) MoDCs and B) moMφs stimulated with *F. nucleatum* strains or *F. nucleatum*-derived LPSs or OMVs as colour-coded. Unstimulated cells (in blue) and isotype (in red) were used as controls. Fn, *F. nucleatum*. Data shown are one representative experiment reproduced in three independent experiments.

5.2.3 Siglec-7 expression in immune cells

Having shown that *F. nucleatum* ssp. and their derivatives modulate host immune response and previously that *F. nucleatum* ssp. bind to Siglec-7 (Chapter 3), we investigated the expression of Siglec-7 and its mouse ortholog Siglec-E in human primary cells and lamina propria leukocytes, respectively, and the effect of *F. nucleatum* ssp. on Siglec-7 expression in human primary cells.

5.2.3.1 Investigation of Siglec-7 and -E expression in immune cells

We first investigated Siglec-7 expression in human monocytes, moDCs or moMφs by flow cytometry, and we showed a distinct population shift in the presence of anti-Siglec-7 antibody, indicating the presence of Siglec-7 on both moDC and moMφ cell surface (Fig. 55).

To date, studies have reported Siglec-E expression in mouse splenic dendritic cells and neutrophils of peripheral blood and their precursors in bone marrow²⁷³, however, Siglec-E expression in the mouse gut is not known. To this aim, leukocytes isolated from mouse colonic lamina propria were analysed by flow cytometry as described in Materials and Methods (2.5.3). To identify the Siglec-E positive cells, the gating strategy was to first select live single cells positive to MHC-II. Further, the CD103 positive cells and the subpopulation of CD11b (positive and negative) cells were tested for Siglec-E expression (Fig. 56). Our results showed that Siglec-E is expressed in colonic CD11b+ cells and the expression is higher in F4/80+ (macrophages) cells (Fig. 56).

However, since our flow cytometry binding assays did not show binding between Siglec-E and *F. nucleatum* (Chapter 3, section 3.2.2.2), the experiments performed in this work were conducted only using human-derived primary cells or cell lines (Chapters 5,6) or human resected tissues (Chapter 7).

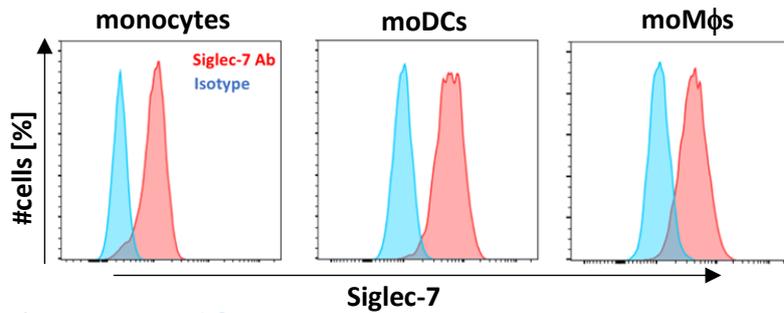


Figure 55 | Histogram of Siglec-7 expression on myeloid cells.

Expression of Siglec-7 (in red) on human primary monocytes, monocyte-derived dendritic cells (moDCs), and monocyte-derived macrophages (moMφs). Isotype control in blue.

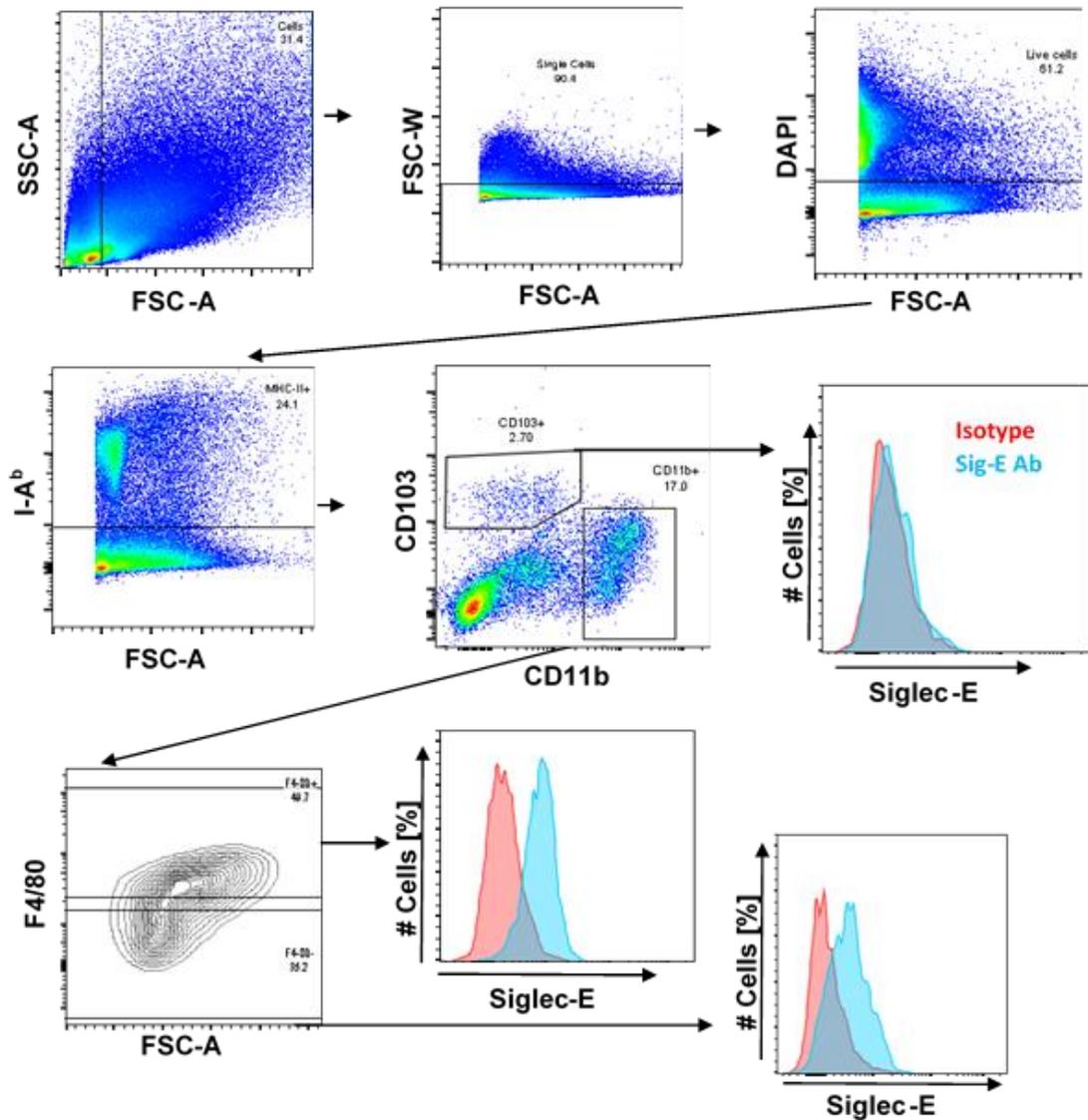


Figure 56 | Siglec-E expression in mouse colonic lamina propria leukocytes.

The leukocyte population was selected in the side scatter (SSC) versus forward scatter (FSC) dot plot, following single cell selection on FSC-W versus FSC-A dot plot the negative in DAPI population (living cells) and positive to I-A^b (MHC-II) population was selected. Siglec-E expression (in blue) was identified in the CD103 positive population and the subpopulations (F4/80 positive and negative) of CD11b. Isotype control in red.

5.2.3.2 Effect of *F. nucleatum* on Siglec-7 expression in moDCs and moMφs

To identify the effect of *F. nucleatum* ssp. on Siglec-7 expression, moDCs or moMφs were stimulated with *F. nucleatum* ATCC 10953, *F. nucleatum* ATCC 10953 or ATCC 51191-derived OMVs, and *F. nucleatum* ATCC 10953 or ATCC 51191-derived LPS at 10 μg/ml and Siglec-7 expression was analysed by flow cytometry.

In moDCs, stimulation with *F. nucleatum* ATCC 10953, LPS or OMVs did not show differences in Siglec-7 cell surface expression as compared to the unstimulated control (Fig. 57A). However, stimulation of moMφs with *F. nucleatum*-derived LPS or OMVs or whole *F. nucleatum* ATCC 10953 resulted in the reduction of Siglec-7 cell surface expression compared to the unstimulated media control (Fig. 57B).

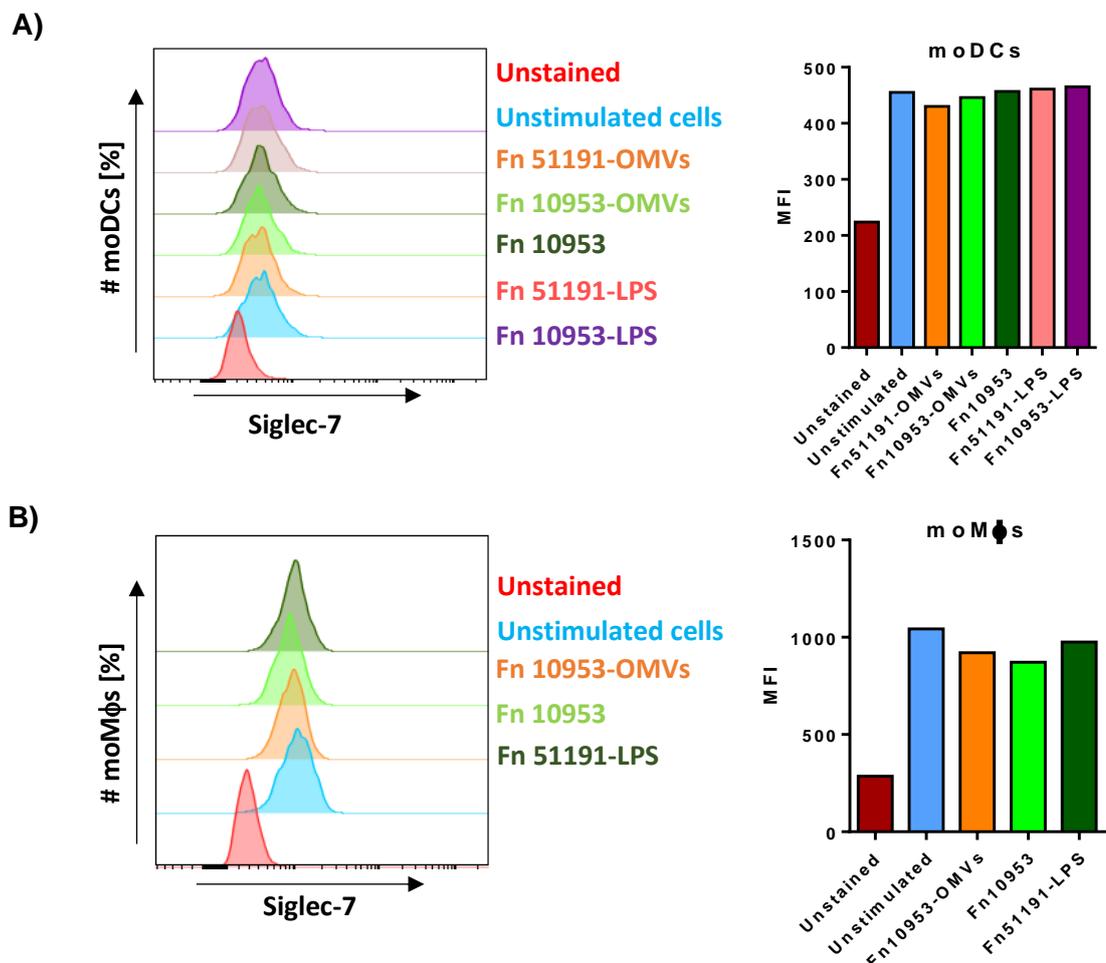


Figure 57 | Histograms of Siglec-7 expression in moDCs and moMφs following *F. nucleatum* stimulation.

A) MoDCs and B) moMφs stimulated with *F. nucleatum* strains or *F. nucleatum*-derived LPS or OMVs as colour-coded. Unstimulated cells (in blue) and unstained (in red) were used as controls. Fn, *F. nucleatum*. Data shown are one representative experiment reproduced in three independent experiments.

5.3 Discussion

F. nucleatum's involvement in colorectal tumour progression has mainly been associated with its ability to modulate host immunity but the mediators of this interaction remain largely unknown. Here, we report the effects of *F. nucleatum* ATCC 10953, ATCC 25586 and ATCC 51191 strains and of their derived OMVs or LPS on human innate immune response.

Our results showed that in moDCs, all three *F. nucleatum* strains, their derived OMVs, and the *F. nucleatum* ATCC 51191-derived LPS at 10 µg/ml induced a pro-inflammatory cytokine profile characterised by an increase in TNF α and decrease in IL-10 production. In addition, all three *F. nucleatum* strains induced the expression of CD80 and CD86 cell surface markers. These results are consistent with studies showing an increase in pro-inflammatory cytokines such as TNF α , IL-24 and IL-6 in CRC mouse models gavaged with *F. nucleatum* as compared to the non-treated mice group and with transcriptomic data of CRC patients with high load of *Fusobacterium* as compared to patients with low load of *Fusobacterium*, respectively¹⁹⁵. It is of note that stimulation of moDCs with *E. coli*-derived LPS induced both TNF α and IL-10 production at concentration 10-fold lower than the immunogenic concentration of *F. nucleatum* ATCC 51191-derived LPS used, indicating a species-specific LPS effect on immune response. Interestingly, in contrast to *F. nucleatum* ATCC 51191-derived LPS, no effect on TNF α or IL-10 cytokine production was observed when moDCs were stimulated with *F. nucleatum* ATCC 10953-derived LPS, while CD80 and CD86 cell surface marker expression was induced, indicating a strain-specific effect of *F. nucleatum* LPS on the immune phenotype of these cells. Variation in the expression of the cell surface marker CD86 of moDCs was also observed between *F. nucleatum* and their derivatives, OMVs or LPS, which may be due to the presentation, concentration, size or nature of the immunogenic epitopes. Only recently *F. nucleatum* OMVs have been purified from *F. nucleatum* ssp. *animalis* and characterised for their proteomic profile²¹⁵ and interaction with epithelial cells²¹⁶ but their biological roles with respect to their interactions with immune cells have not been investigated. To our knowledge there are no reports comparing the effects between whole bacteria and their derived OMVs on the host immune response.

Stimulation of moMφs with *F. nucleatum* strains showed a M2-macrophage polarisation characterised by the induction of IL-10, IL-6, and the cell surface marker CD163⁴⁷. In addition, *F. nucleatum* strains induced the expression of the cell surface marker PD-L1, known as an immune checkpoint protein and the reduction of the cell surface marker CD86, associated with T cell activation. Induction of IL-10 but not TNF α was also observed when the macrophage-like cell line U937 was stimulated with the *F. nucleatum* strains. Similarly, moMφs stimulated with *F. nucleatum* derived LPS or OMVs showed induction of IL-10 and CD80 but the CD86 expression was variable. These results are in agreement with previous studies showing an infiltration of M2-macrophages in *F. nucleatum* ssp. positive clinical CRC specimens²⁷⁴ or the development of a M2 acquired phenotype after stimulation of macrophage-like cell lines with *F. nucleatum* ATCC 10953 and isolates²⁰⁷.

In conclusion, our collective data of moDCs and moMφs responses following stimulation with *F. nucleatum* ssp. are in agreement with reports correlating *Fusobacterium* spp. positive CRC cases to elevated local IL-10 and TNF α levels^{195,275}. We also showed that moMφs had higher degree of *F. nucleatum* ssp. internalisation and cell viability when compared to moDCs and, interestingly, *F. nucleatum* ATCC 10953 strain, the longest in size across strains, showed the highest increase of moDCs death compared to the rest strains, suggesting a potential unique evasion strategy for this strain. The same strain has been previously shown to significantly induce cell apoptosis and necrosis in neutrophils²⁶². Our findings of the morphological and myeloid cell-specific immune modulation properties of *F. nucleatum* ssp. pave the way for further research on the interactions between *F. nucleatum* and immune cells.

Interestingly, Siglec-7 expression was reduced in moMφs after stimulation with *F. nucleatum* ssp. or the derived LPS or OMVs, raising questions about the contribution of Siglec-7 in the immune modulation induced by *F. nucleatum* ssp., which will be addressed in the next chapter.

CHAPTER 6

-

Role of Siglec-7 in *F. nucleatum* modulation of host immune response

6.1 Introduction

Siglec-7 is expressed on the surface of leukocytes, mainly in peripheral blood NK cells, but its expression has also been reported in human colonic lamina propria leukocytes mainly in monocytes/macrophages¹⁰⁰. Intracellularly, Siglec-7 bears an immunoreceptor tyrosine-based inhibition motif (ITIM) proximal and an ITIM-like motif distal to the cell membrane (Fig. 58) which upon ligand binding becomes tyrosine phosphorylated, leading to the recruitment of the phosphatases SH2-domain-containing protein tyrosine phosphatase 1 (SHP-1) and 2 (SHP-2). ITIM-containing motifs have the ability to suppress immunoreceptor tyrosine-based activation motif (ITAM)-induced immune responses⁹⁰.

Siglec-7 is mainly known for its interaction with host's malignant cells from numerous tissue origins, resulting in immune suppression⁹¹. Human cell lines which show a strong association with Siglec-7 are derived from melanoma, hematological malignancies, cervical cancer or colon adenocarcinoma^{276,277}. For example, Jandus *et al.* 2014, showed that the interaction of HeLa (cervical cancer) or cell lines derived from hematological malignancies with Siglec-7 expressed on human primary NK cells inhibits anti-tumour responses²⁷⁷. Recently, Wisnovsky *et al.* 2021 showed that Siglec-7 binds to a glycopeptide presented on CD43 which expressed on leukemia cells, and blocking of the interaction between Siglec-7 and CD43 induced immune cell activity¹⁰⁸.

Furthermore, microorganisms such as Group B *Streptococcus* (GBS)⁹⁸, HIV²⁴⁵, or zymosan yeast particles¹⁰⁶ have been shown to modulate the host immune response in a Siglec-7-dependent manner. In NK cells, the interaction of GBS with Siglec-7 results in immune evasion by inhibiting the pyroptotic NK phenotype⁹⁸. Interestingly, the interaction of zymosan yeast particles and human primary monocytes showed a Siglec-7-dependent pro-inflammatory immune response¹⁰⁶, in contrast to the classical immunosuppressive response of Siglec-7. Additionally, it has been reported that in human monocyte-derived macrophages, Siglec-7 contributes to HIV internalisation²⁴⁵. *C. jejuni* has also been shown to interact with Siglec-7 through its cell surface lipooligosaccharides (LOS), although the associated effect on the host immune response has not been studied¹⁰⁵.

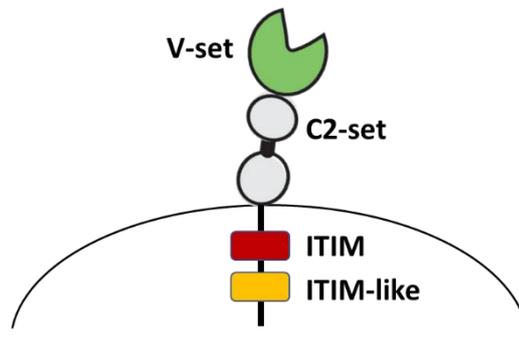


Figure 58 | Schematic representation of Siglec-7 structure on the mammalian cell membrane. The transmembrane Siglec-7 lectin consists of a V-set and 2 C2-set extracellular domains, and intracellularly consist of an ITIM and an ITIM-like motif.

F. nucleatum ssp. have been shown to interact with innate immune cells through the host receptors, TIGIT²⁰⁹, TLR-4, CD14²¹² but there is no report of interaction between *F. nucleatum* ssp. and lectins expressed on human immune cells.

In this chapter, following our findings that *F. nucleatum* ssp. bind to Siglec-7 and that *F. nucleatum* ssp. modulate the primary cell immune response, we investigated the contribution of Siglec-7 on the host immune response induced by *F. nucleatum* ssp. and their derivatives: *F. nucleatum* ATCC 51191 or ATCC 10953-derived lipopolysaccharide (LPS), and *F. nucleatum* ATCC 51191-derived outer membrane vesicles (OMVs).

6.2 Results

6.2.1 Effect of Siglec-7 on *F. nucleatum*-induced U937 cell response

We first investigated the contribution of Siglec-7 on *F. nucleatum*-induced U937 response by comparing the effect of *F. nucleatum* strains on cytokine production and internalisation between WT or Siglec-7-positive cells sorted after CRISPR/Cas9-mediated Siglec-7 deletion (Siglec-7^{+/+} cells), and Siglec-7 depleted (Siglec-7^{-/-}) cells. We confirmed that Siglec-7 is expressed in both WT and Siglec-7^{+/+} but not in Siglec-7^{-/-} cells (Fig. 59A). The U937 cell lines were then differentiated into a macrophage-like phenotype using phorbol 12-myristate 13-acetate (PMA), as shown by the induction of CD11b and CD11c cell surface markers (Fig. 59B). PMA-U937 differentiated cell lines (WT or Siglec-7^{+/+}, and Siglec-7^{-/-}) were then used in the rest of this work.

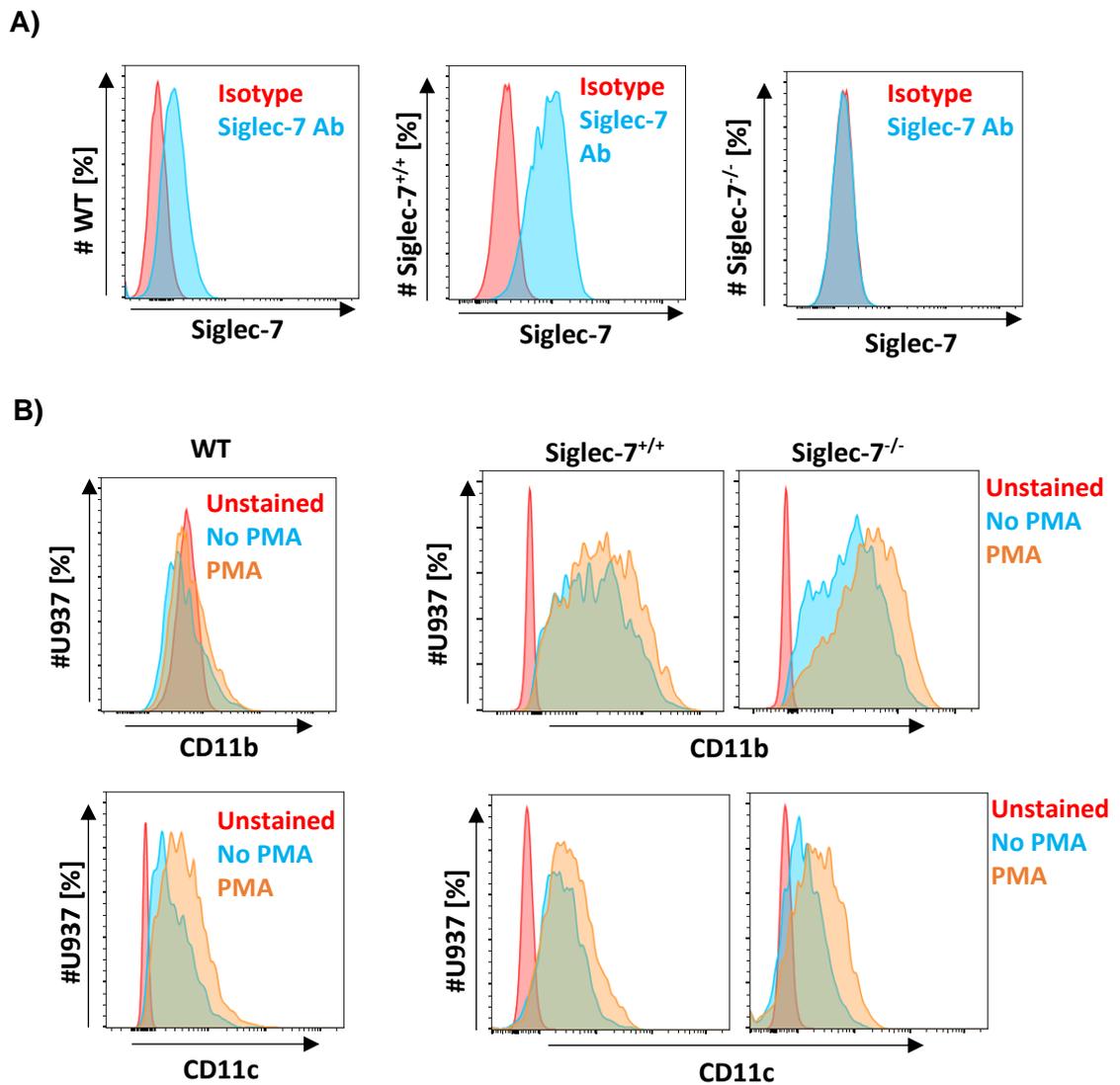


Figure 59 | Histograms of protein expression on U937 cell lines.

Expression of A) Siglec-7 (in blue) or isotype (in red) on U937 (WT, Siglec-7^{+/+}, Siglec-7^{-/-}) cells, and B) CD11b and CD11c on U937 (in blue) or PMA differentiated U937 (in orange) (WT, Siglec-7^{+/+}, Siglec-7^{-/-}), and unstained cells (in red).

Next, we treated the PMA-U937 cells with *F. nucleatum* ssp. or their derived LPS and OMVs and cytokine production was quantified. When Siglec-7^{-/-} cells were stimulated with *F. nucleatum* ATCC 10953, ATCC 25586 or ATCC 51191, a statistically significant induction of TNF α and IL-10 cytokine production was observed as compared to the WT (Fig. 60A,B) or Siglec-7^{+/+} cells (Fig. 60B). In addition, IL-10 production was significantly induced in Siglec-7^{-/-} cells following stimulation with *F. nucleatum* (ATCC 10953 or ATCC 51191)-derived LPS at 10 or 1 μ g/ml or *F. nucleatum* ATCC 51191-derived OMVs ($p < 0.0001$) as compared to the stimulated-Siglec-7^{+/+} cells (Fig. 60C). Interestingly, when Siglec-7^{+/+} cells were stimulated with *F. nucleatum* ATCC 51191-derived LPS at 10 or 1 μ g/ml, no significant induction of IL-10 was detected when compared to the unstimulated control (Fig. 60C).

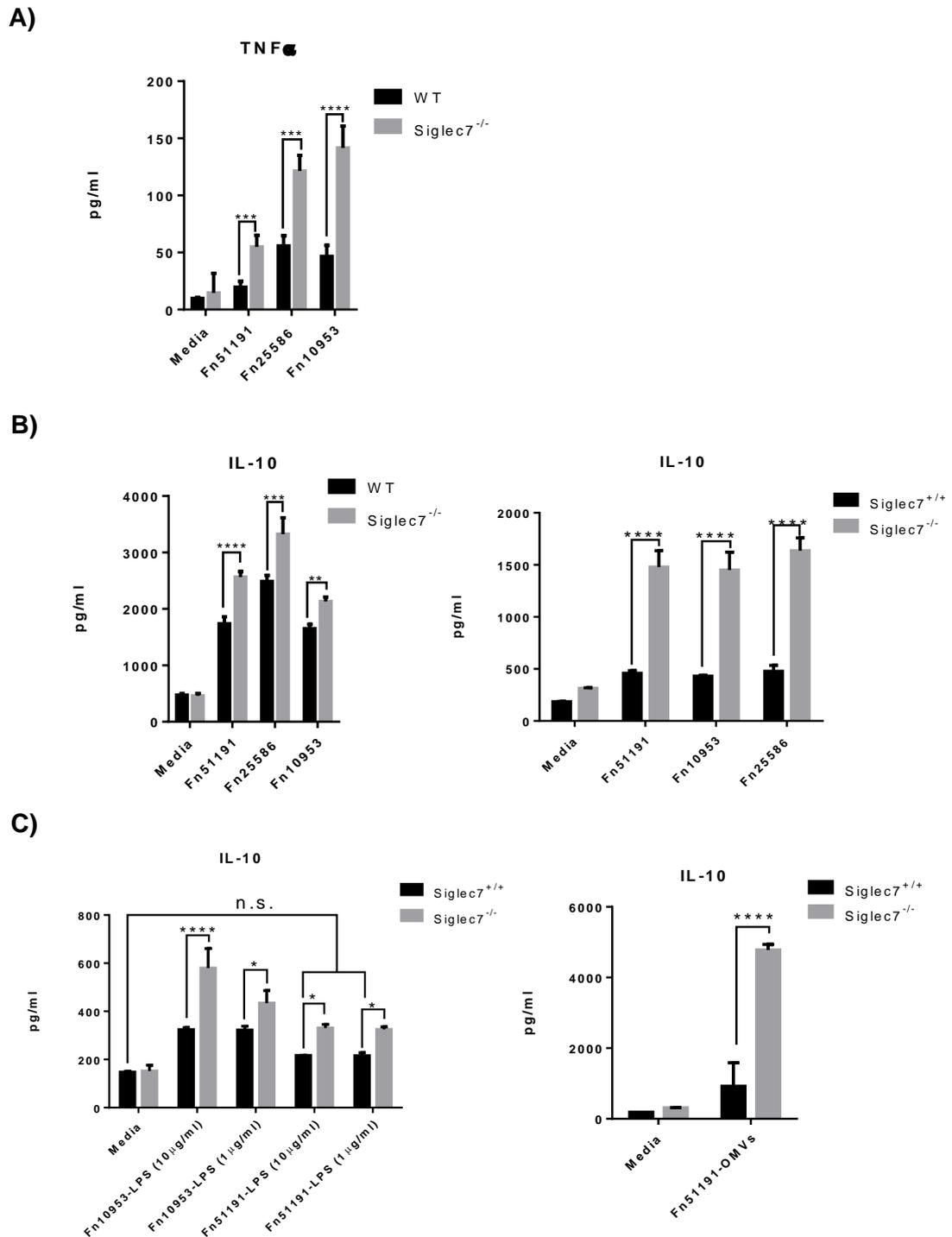


Figure 60 | Cytokine production in PMA-U937 cells stimulated with *F. nucleatum* ssp. U937 (WT or Siglec-7^{+/+}, Siglec-7^{-/-})-derived A) TNF α and B) IL-10 production after stimulation with *F. nucleatum* ATCC 51191, ATCC 10953, ATCC 25586 or C) IL-10 production after stimulation with *F. nucleatum* (ATCC 51191 or ATCC 10953)-derived LPS at 10 or 1 μ g/ml, and *F. nucleatum* ATCC 51191-derived OMVs. Fn, *F. nucleatum*. Unstimulated cells (media) was used as a negative control. Data shown are the mean of triplicates from one representative experiment reproduced in three independent experiments. Statistics: *p<0.05; **p<0.01; *p<0.001; ****p<0.0001; n.s., not statistically difference.**

We then investigated the internalisation of *F. nucleatum* ATCC 10953, ATCC 25586 or ATCC 51191 into PMA-U937-Siglec-7^{+/+} or Siglec-7^{-/-} cells by imaging flow cytometry (IFC) using FITC-labelled bacteria. The proportion of internalised bacteria were identified as described in Materials and Methods (section 2.4.7). Siglec-7^{-/-} cells showed a significant increase in internalisation of *F. nucleatum* ATCC 10953 (p<0.05) and ATCC 25586 (p<0.01) (Fig. 61) as compared to Siglec-7^{+/+} cells. In contrast, no statistically significant difference was observed for *F. nucleatum* ATCC 51191 internalisation in Siglec-7^{-/-} cells when compared to Siglec-7^{+/+} cells (Fig. 61).

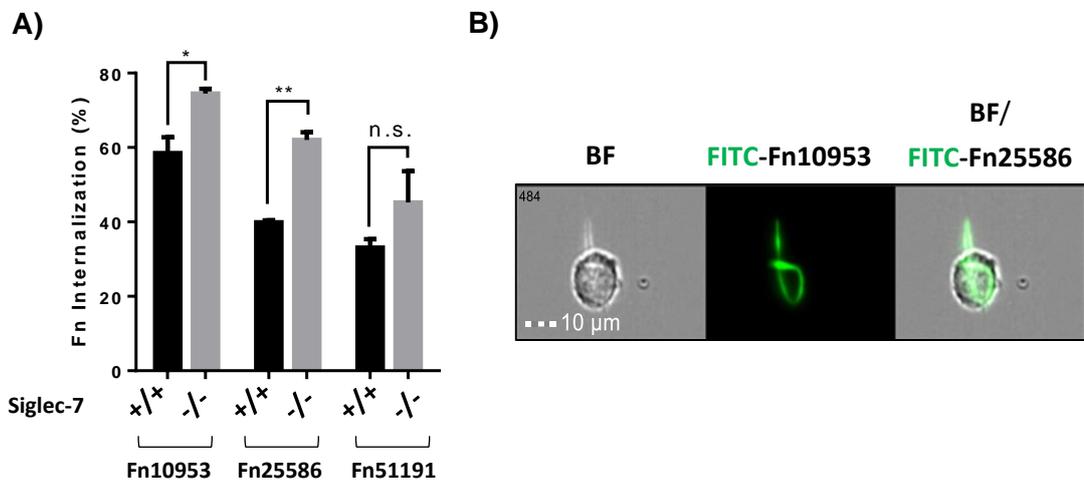


Figure 61 | Internalisation of *F. nucleatum* ssp. into U937 cells.

A) U937-Siglec-7^{+/+} or Siglec-7^{-/-} were treated with FITC-labelled *F. nucleatum* ATCC 10953, ATCC 25586, or ATCC 51191. B) Image of internalised *F. nucleatum* ATCC 10953 and ATCC 25586, into U937-Siglec-7^{-/-} cell. BF, bright field; Fn, *F. nucleatum*. Data shown are the mean of duplicates from one representative experiment reproduced in three independent experiments. Statistics: *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001; n.s., not statistically difference.

6.2.2 Effect of Siglec-7 in *F. nucleatum*-induced myeloid cell response

Here we investigated the contribution of Siglec-7 on *F. nucleatum* ssp.-induced moDCs or moMφs immune responses using two approaches, i) by blocking Siglec-7 interaction with *F. nucleatum* ssp. and ii) by silencing Siglec-7 expression in moDCs or moMφs. Next, we determined the effect of *F. nucleatum* ssp. on cytokine production, cell surface marker expression and internalisation in human primary cells.

6.2.2.1 anti-Siglec-7 antibody modulates myeloid cell response

Here we used anti-Siglec-7 antibody (Sig7Ab), as an attempt to block the interaction between Siglec-7 expressed by moDCs or moMφs and *F. nucleatum* ATCC 51191 or ATCC 10953. However, the presence of the antibody *per se* led to a significant induction of IL-10 ($p < 0.05$) and IL-8 ($p < 0.0001$) production in moMφs (Fig. 62). Also, we showed that pre-treatment of moMφs with Sig7Ab followed by *F. nucleatum* ssp. stimulation significantly induced the production of TNF α and IL-10 but had no effect on the production of IL-8 when compared to the stimulated untreated or treated with the antibody isotype (IsoAb) cells (Fig. 62). Pre-treatment of moDCs with Sig7Ab followed by *F. nucleatum* ssp. stimulation did not significantly affect TNF α production as compared to the stimulated untreated or IsoAb-treated cells (Fig. 62), whereas the production of IL-10 was significantly reduced when compared to the stimulated untreated cells but not to IsoAb-treated cells (Fig. 62). Production of IL-8 showed a significant reduction when Sig7Ab pre-treated moDCs were stimulated with *F. nucleatum* ATCC 51191 ($p < 0.001$), but no significant difference was observed when cells were stimulated with *F. nucleatum* ATCC 10953, as compared to the stimulated untreated cells (Fig. 62). The production of IL-8 in isoAb-pre-treated moDCs was significantly reduced when stimulated with *F. nucleatum* ATCC 10953 ($p < 0.05$) but no significant difference was observed with *F. nucleatum* ATCC 51191 when compared to stimulated Sig7Ab-treated cells (Fig. 62).

Together these results showed that the effects of Sig7Ab on myeloid cell response hampered the use of the Siglec-7 antibody strategy to explore the involvement of Siglec-7 in *F. nucleatum* ssp.-induced immune response.

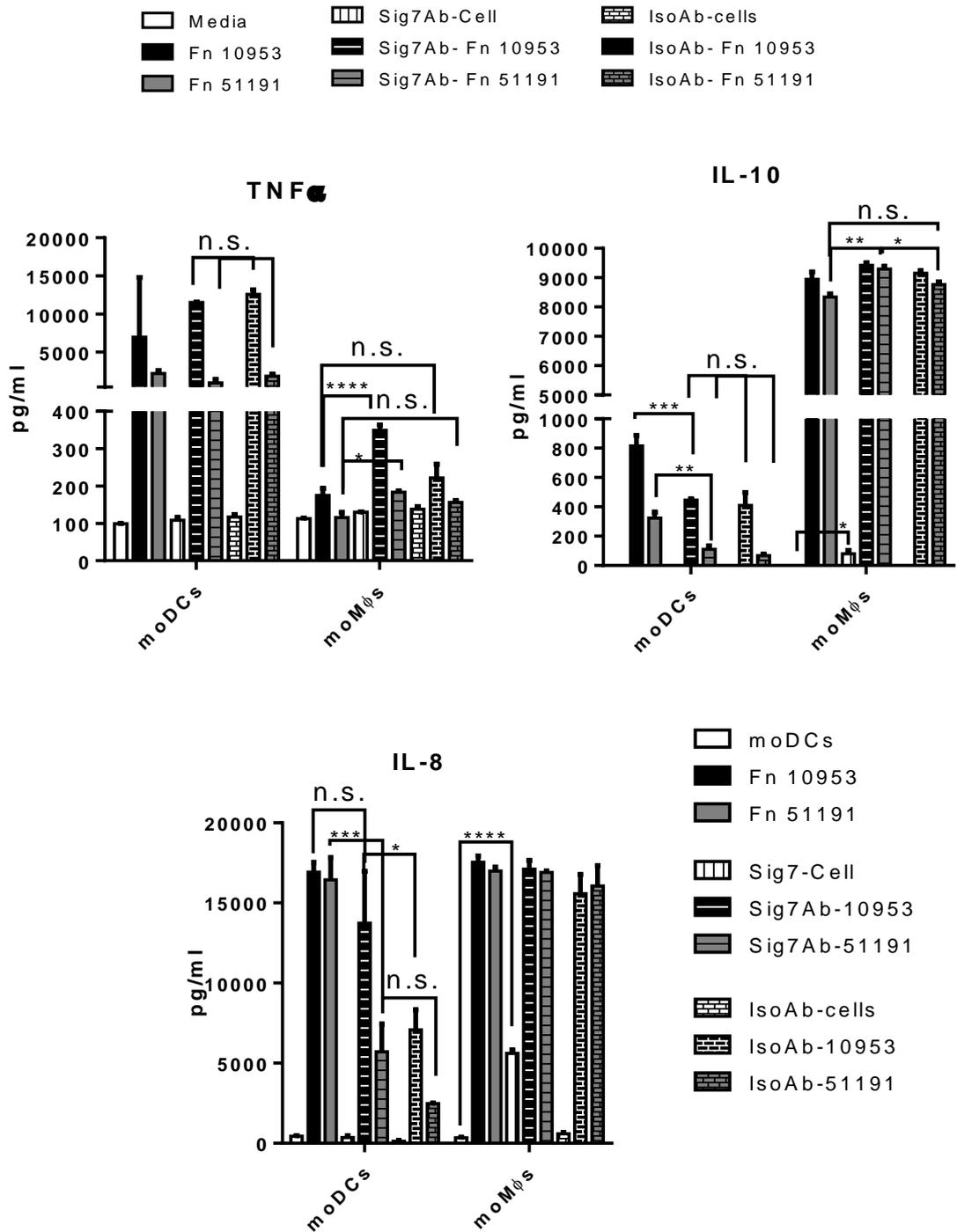


Figure 62 | moDCs or moMφs cytokine production in presence of anti-Siglec-7 Ab and stimulation with *F. nucleatum* ssp.

MoDCs or moMφs were pre-treated with anti-Siglec-7 Ab (horizontal line pattern) or the isotype antibody (cubic pattern) and stimulated with *F. nucleatum* ATCC 10953 (in black) or ATCC 51191 (in grey). Unstimulated cells (in white) was used as a negative control. Fn, *F. nucleatum*; anti-Siglec-7 Ab (Sig7Ab); isotype antibody (IsoAb). Data shown are the mean of triplicates from one representative experiment reproduced in two independent experiments. Statistics: *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001; n.s., not statistically difference.

6.2.2.2 Optimisation of Siglec-7 silencing in moDCs and moMφs

Here, we aimed to reduce Siglec-7 expression in primary myeloid cells by mRNA silencing. MoDCs or moMφs were incubated with the pre-designed Siglec-7 mRNA silencing probes (Sig7 probes) or with scramble probe which were used as a negative control. The scramble probe contains a nucleotide sequence which does not target any gene present in the recipient cells. The protocol for optimised siRNA in human primary cells from Troegeler *et al.* 2014²²⁸ was adapted to Siglec-7, as described in Materials and Methods (section 2.4.3).

For moDCs, optimal silencing was achieved following incubation with a combination of two Sig7 probes, resulting in 68% reduction of the relative mean fluorescence intensity (R-MFI) value of Siglec-7, whereas the use of 3 Sig7 probes led to a 41% reduction in the Siglec-7 R-MFI value (Fig. 63A). When moDCs were treated with the scramble control or Sig7 probes, we observed an approximately 15% increase in cell death as compared to the non-treated cells (Fig. 63B). While the incubation of moDCs with 2 Sig7 probes for 4 days resulted in levels of Siglec-7 silencing similar to that obtained after 3 days incubation, we observed an approximately 37% increase of dead cells in both Sig7 or scramble pre-treatment cells when compared to the non-treated cells (Fig. 63C). In moMφs, incubation with 3 Sig7 probes for 3 days showed around 35% reduction of the Siglec-7 R-MFI value (Fig. 63A) and an approximately 2-3% increase in cell death as compared to the non-treated cells (Fig. 63B). The requirement for prolonged incubation period could be due to the long Siglec-7 half-life in moDCs and moMφs.

Based on the above results, the best conditions for Siglec-7 mRNA silencing were found to require the incubation of moDCs with two probes for 3 days. These conditions were used in the rest of the study.

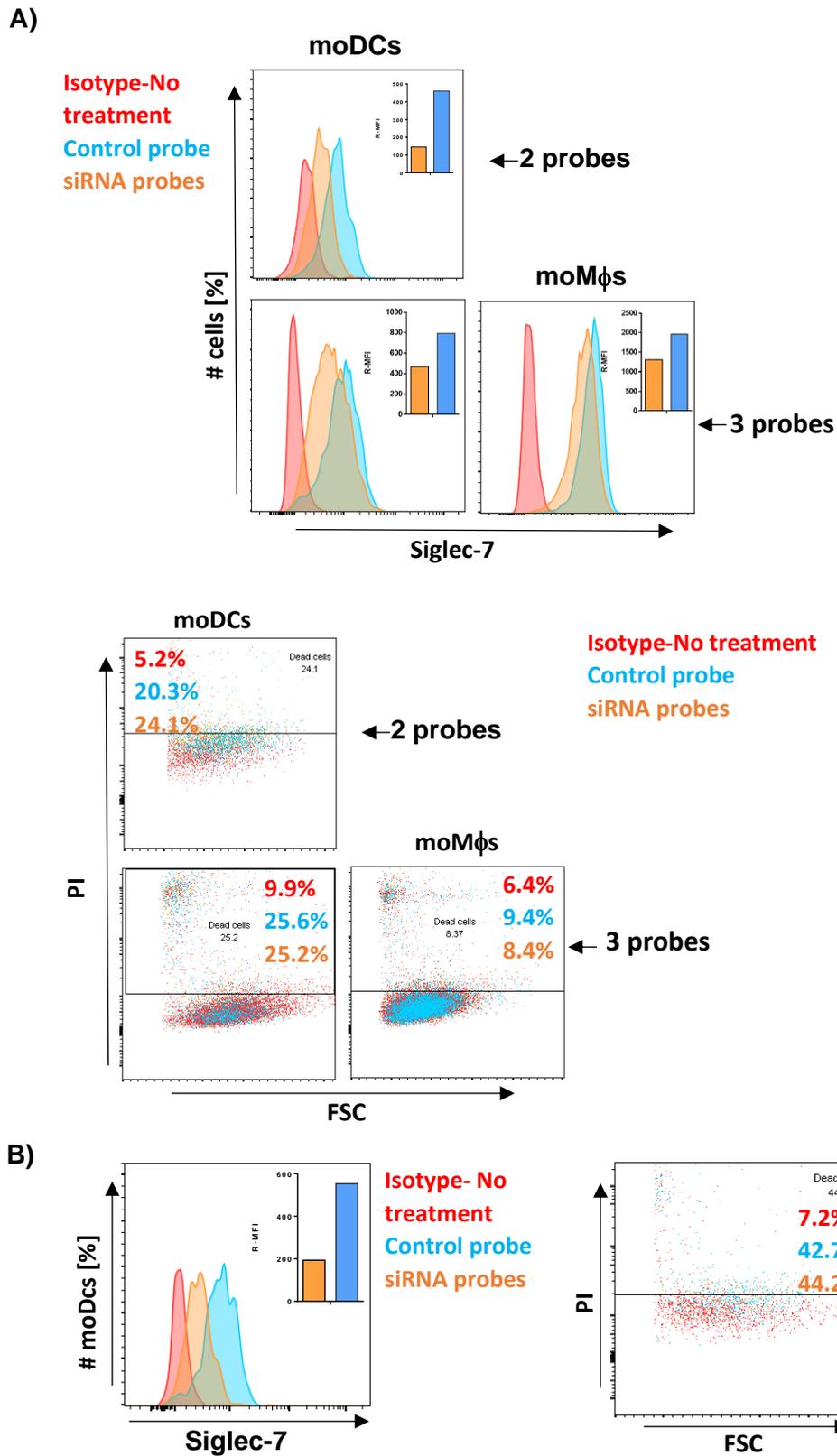


Figure 63 | Histograms of Siglec-7 expression and cell death of moDCs and moMφs. Expression of Siglec-7 on cells and the proportion of dead cells treated with the scramble control probe (in blue) or Siglec-7 mRNA silenced probes (in orange) for A) 3 days in moDCs with 2 or 3 probes or moMφs with 3 probes or for B) 4 days in moDCs with 2 probes. Isotype in red. R-MFI, relative mean fluorescence intensity.

6.2.2.3 *F. nucleatum* modulates moDC response in a Siglec-7-dependent manner

To investigate the involvement of Siglec-7 in the immune response induced by *F. nucleatum* ssp., moDCs treated with Siglec-7 mRNA silencing (Sig7 probes) or scramble probes, as described above, were stimulated with *F. nucleatum* ATCC 10953, ATCC 25586 or ATCC 51191 strains or with *F. nucleatum* ATCC 51191-derived LPS at 10 or 1 μ g/ml and the cell response determined with respect to cytokine production and cell surface marker expression. The results showed a significant reduction in TNF α production when Sig7-silenced moDCs were stimulated with *F. nucleatum* ATCC 10953 ($p < 0.05$) or ATCC 51191 ($p < 0.0001$) compared to the stimulated scramble control moDCs (Fig. 64A). Similarly, IL-10 production was also significantly reduced when the silenced moDCs were stimulated with *F. nucleatum* ATCC 10953 ($p < 0.01$) compared to the stimulated scramble control moDCs (Fig. 64A). In contrast, when moDCs were stimulated with *F. nucleatum* ATCC 51191-derived LPS, no statistically significant difference in TNF α production was observed between the silenced and scramble control moDCs (Fig. 64B). While IL-10 production was induced in Siglec-7 silenced moDCs following *F. nucleatum* ATCC 51191-derived LPS stimulation at 10 μ g/ml, there was no statistically significant difference at 1 μ g/ml, when compared to the stimulated scramble control moDCs (Fig. 64B).

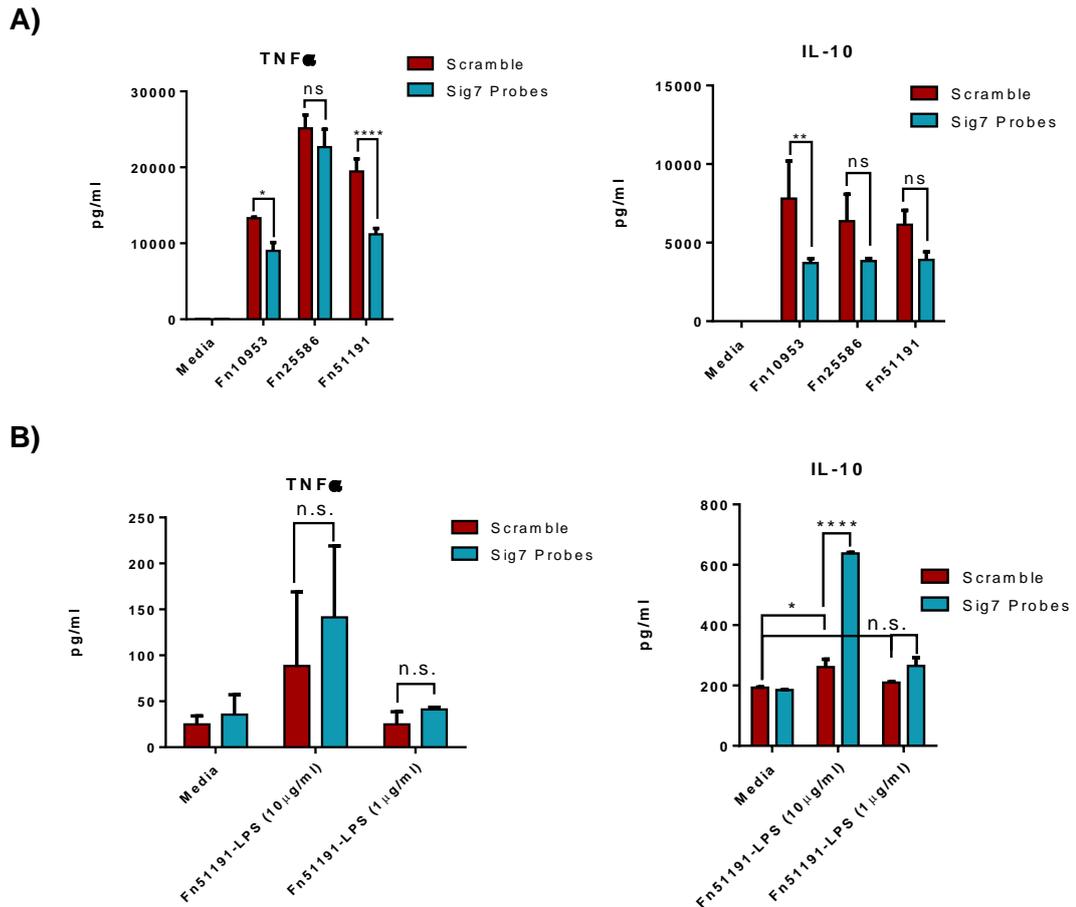


Figure 64 | Cytokine production in Siglec-7 silenced moDCs stimulated with *F. nucleatum* ssp.

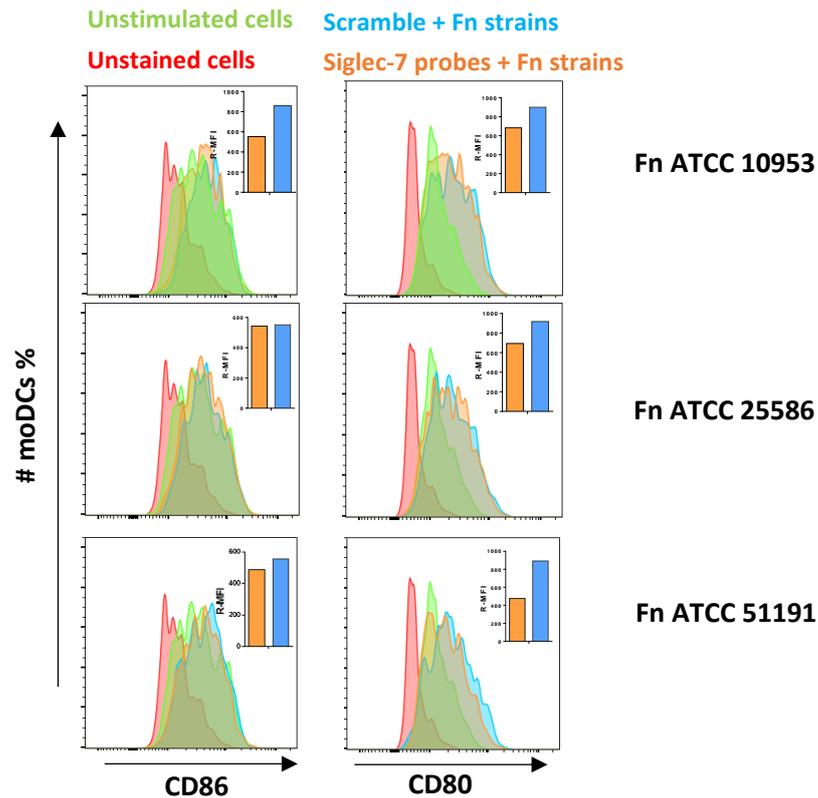
IL-10 and TNF α production by moDCs pre-treated with Sig7 (in blue) or scramble (in red) probes and stimulated with A) *F. nucleatum* ATCC 51191, ATCC 10953, ATCC 25586 or B) *F. nucleatum* ATCC 51191-derived LPS at 10 or 1 μ g/ml. Unstimulated cells were used as a control (Media). Fn, *F. nucleatum*. Data shown are the mean of triplicates from one representative experiment reproduced in three independent experiments. Statistics: *p<0.05; **p<0.01; ***p<0.001; ****p<0.0001; n.s., not statistically difference.

Next, we investigated the effect of *F. nucleatum* ssp. on cell surface marker expression in Siglec-7 silenced or scramble control moDCs by flow cytometry. Here, the R-MFI value is referred to the difference between the MFI of the *F. nucleatum*-stimulated Sig-7 silenced or scramble control moDCs and the unstimulated Sig-7 silenced or scramble control moDCs.

The results showed a reduction in CD86 and CD80 R-MFI values, indicating a reduced expression of these markers when Siglec-7 silenced moDCs were stimulated with *F. nucleatum* ATCC 10953, ATCC 25586 or ATCC 51191 strains, as compared to the stimulated scramble moDCs (Fig. 65A). When stimulated with *F. nucleatum* ATCC 51191-derived LPS at 10 µg/ml, CD86 expression was reduced in Siglec-7 silenced moDCs while PD-L1 expression was induced as compared to the LPS-stimulated scramble moDCs (Fig. 65B). Furthermore, our preliminary internalisation assays of *F. nucleatum* ATCC 25586 into moDCs, showed an approximately 10% increase of *F. nucleatum* internalisation into Siglec-7 silenced moDCs as compared to scramble moDCs (Fig. 66).

Overall, we showed that in *F. nucleatum* ssp. stimulated moDCs, Siglec-7 was involved in mainly TNF α but also IL-10 cytokine production, and the co-stimulatory for T cell activation markers, CD80, CD86, expression. In addition, in moDCs stimulated with *F. nucleatum* ATCC 51191-derived LPS at 10 µg/ml, the presence of Siglec-7 showed reduced IL-10 production and PD-L1 expression but induced CD86 expression.

A)



B)

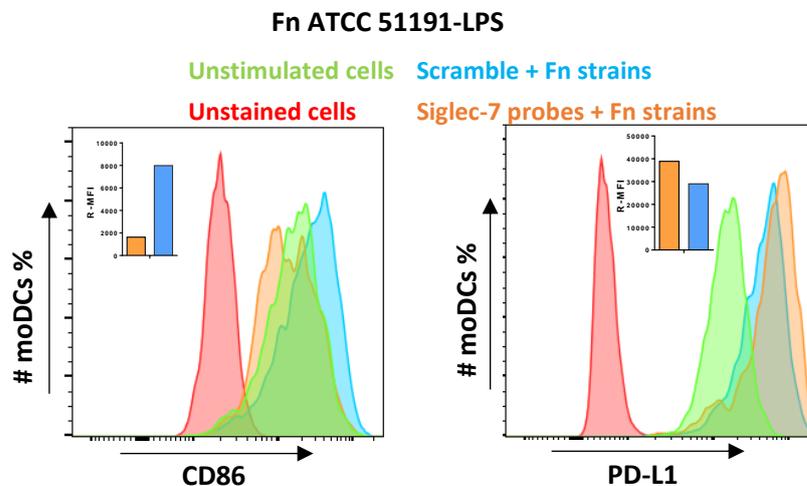
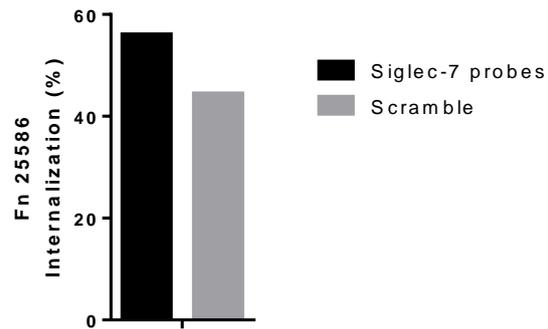


Figure 65 | Histograms of marker expression on Siglec-7 silenced moDCs.

Histograms and the relative to unstimulated control (in green) mean fluorescence intensity (R-MFI) values of A) CD86 and CD80 on Siglec-7 silenced (in orange) or scramble control (in blue) moDCs followed by *F. nucleatum* ATCC 10953, ATCC 25586 or ATCC 51191 stimulation, or B) CD86 and PD-L1 on Siglec-7 silenced (in orange) or scramble control (in blue) moDCs followed by *F. nucleatum* ATCC 51191-derived LPS stimulation at 10 $\mu\text{g/ml}$. Unstained cells (in red) were used as a control. Fn, *F. nucleatum*. Data shown are from one representative experiment reproduced in three independent experiments.

A)



B)

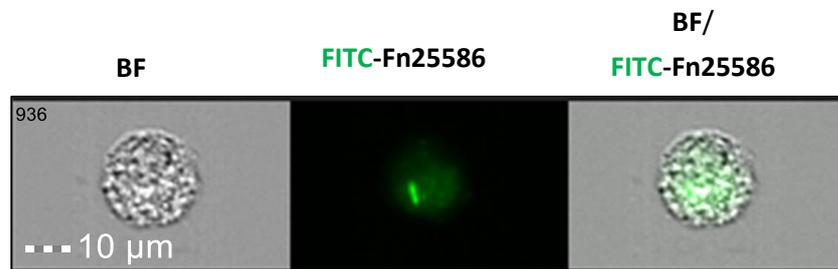


Figure 66 | Internalisation of *F. nucleatum* ssp. into moDCs.

A) Siglec-7 silenced (in black) or scramble (in grey) moDCs were stimulated with *F. nucleatum* ATCC 25586. B) Image of internalised FITC-labelled *F. nucleatum* ATCC 25586 into Siglec-7 silenced moDC. Fn, *F. nucleatum*; BF, bright field. Data shown are produced from one experiment.

6.3 Discussion

In this chapter, we investigated the contribution of Siglec-7 on the innate immune cell response induced by *F. nucleatum* ssp. or their derived LPS. We first showed that anti-Siglec-7 antibody *per se* resulted in moMφs stimulation and production of anti-inflammatory IL-10 and pro-inflammatory IL-8 cytokines. Similarly, Varchetta *et al.* 2012 showed that anti-Siglec-7 antibody treatment of PBMCs could induce a pro-inflammatory cytokine production¹⁰⁶. In addition, we showed that pre-treatment of moDCs or moMφs with the isotype of Siglec-7 antibody could also affect the *F. nucleatum* ssp. induced immune response. Therefore, we concluded that the Siglec-7 antibody blocking approach was not suitable to examine the contribution of Siglec-7 in *F. nucleatum* induced cell response.

Using Siglec-7 mRNA silencing, we showed that Siglec-7 contributed to *F. nucleatum*-induced cell response in terms of cytokine production, cell marker expression and internalisation. The effects differed between freshly isolated primary monocyte-derived dendritic cells (moDCs) and the macrophage-like cell line (U937).

Following binding to their ligands, Siglecs have been shown to modulate internalisation of the receptors. For example, Chang *et al.* 2014 showed that Siglec-E deficiency induced phagocytosis and a pro-inflammatory cytokine profile in macrophages challenged with GBS *in vivo*²⁷⁸ while Siglec-5 was shown to enhance *Neisseria meningitidis* internalisation into CHO cells²⁷⁹. Here, we showed that Siglec-7 deletion promotes internalisation of *F. nucleatum* ssp. (ATCC 10953, ATCC 25586, ATCC 51191) into U937 cells, and the same trend was observed in our preliminary results using Siglec-7 silenced moDCs with *F. nucleatum* ATCC 25586. CD33-related Siglecs have been shown to mediate endocytosis of sialylated glycans and this is connected mainly with the ITIM phosphorylation²⁸⁰. To our knowledge, there are no studies reporting the role of Siglec-7 in internalisation of bacteria. However, Siglec-7 has been implicated in the internalisation of HIV in NK cells²⁸¹ and moMφs cells²⁴⁵, suggesting a role of Siglec-7 in promoting the entry of the virus into immune cells and enhancing the sensitivity of the host to infection by the virus.

Siglec-7 silencing in moDCs showed a reduction in cytokine production, mainly TNF α , following stimulation with *F. nucleatum* (ATCC 10953, ATCC 25586, ATCC 51191), suggesting a contribution of Siglec-7 in the pro-inflammatory

phenotype observed in these primary cells. In contrast to the whole bacteria, Siglec-7 silencing resulted in an induction of IL-10 by moDCs stimulated with *F. nucleatum* ATCC 51191-derived LPS at 10 µg/ml. However, Siglec-7 deletion in U937 cells showed induction of IL-10 production after stimulation with *F. nucleatum* ATCC 51191 or ATCC 10953-derived LPS at both 10 and 1 µg/ml or *F. nucleatum* ATCC 51191-derived OMVs.

To date most studies have shown that Siglec-7, following ligand recognition induces a suppressor phenotype in immune cells and mainly NK cells^{98,104}. Our findings that Siglec-7 is involved in a pro-inflammatory phenotype in moDCs upon *F. nucleatum* ssp. interaction is in line with the reported effect of *F. nucleatum* on tumour development. To our knowledge, only one study by Varchetta *et al.* 2012, using Siglec-7 silencing in monocytes, also showed association of Siglec-7 with the pro-inflammatory cytokine production of TNF α and IL-1 α upon interaction with yeast particles in a sialic acid-independent manner¹⁰⁶. These differences in cell immune response could be attributed to the mode of recognition, the nature of the interactions (cell-cell or cell-microbe), or the heterogeneity in PRR expression in the different cell subsets used in the *in vitro* studies and between primary and immortalised cells. Indeed, other PRRs may act synergistically with Siglec-7 to contribute to a distinct immune response. For example, TLR-4, a receptor with an intracellular activation motif^{282,283}, has been shown to establish a direct interaction with Siglecs, including Siglec-7²⁸⁴, and particularly, with Siglec-9 and Siglec-E which have also been shown to negatively regulate TLR-4 mediated pathways^{230,284}. Therefore, considering that TLR-4 is expressed in both U937²⁸⁵ and moDCs²⁸⁶, and that *F. nucleatum* ssp. have shown to interact with TLR-4²⁸⁷, our findings could result from a synergetic effect between Siglec-7 and TLR-4 in these cells. More work is needed to determine the signalling pathways affected by the interaction between Siglec-7 and *F. nucleatum* in innate immune cells.

Overall, our findings showed a novel biological function of Siglec-7 upon interaction with colorectal cancer associated-bacterial species, *F. nucleatum* ssp., which opens the way for further investigation on the role of Siglec-7 in modulating immune response and cancer development.

CHAPTER 7

-

Siglec-7 expression and *Fusobacterium* spp. abundance in human CRC tissues

7.1 Introduction

CRC tissues are characterised by an increased abundance of *F. nucleatum* species with a high prevalence in adenocarcinoma cases as compared to the adenoma stage²⁷⁵. Moreover, studies investigating the correlation of *F. nucleatum* and tumour invasion-node-metastasis (TNM) evolution, showed an association of *F. nucleatum* abundance in the advanced overall TNM stage^{199,288}. Tumour invasion stages (T) are classified in 4 levels (Fig. 67). T1 describes the tumour invasion in the submucosa, T2 corresponds to the stage where the tumour has invaded the muscularis propria, T3 is when the tumour has developed across the muscularis propria to reach the serosa, and T4 corresponds to the stage where the tumour has spread to either the peritoneum (T4a), or to other organs of the body (T4b). *F. nucleatum* DNA levels were found to be significantly higher in tumour invasion stages T3-T4 compared to T1-T2 stages^{288,289}.

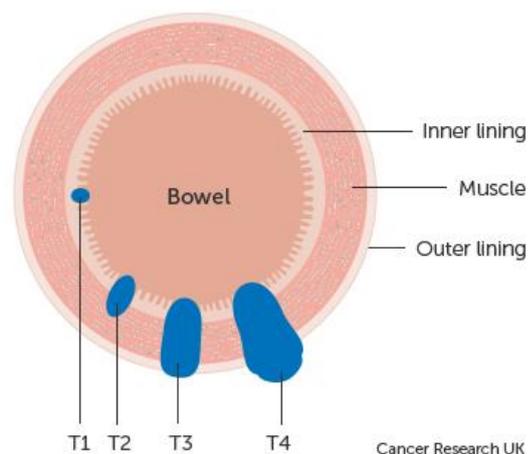


Figure 67 | Colorectal tumour invasion stages.
The T1-4 tumour invasion stages. (Taken from CancerResearchUK.org).

Specific microbial ligands can induce humoral immunity including the production of microbe-specific antibodies detectable in the blood. Furthermore, studies have investigated the presence of anti-*F. nucleatum* antibodies in plasma/serum of CRC cases. A study from Wang *et al.* 2016 including 258 CRC cases, 150 benign colon disease and 200 healthy individuals, showed high levels of anti-*F. nucleatum* ATCC 25586 serum antibodies (IgA and IgG) in CRC cases compared to the other groups²⁹⁰. The same study also showed that CRC detection could be

improved in terms of sensitivity when using anti-*F. nucleatum* antibodies in combination with the common CRC biomarkers, carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA19-9). Recently, Alkharaan *et al.* 2020 showed that anti-*F. nucleatum* antibodies (against the whole isolate strain or the *F. nucleatum*-derived Fap2 protein) found in plasma and saliva could also serve as a biomarker for pancreatic cancer, which remains asymptomatic until late stages of cancer²⁹¹.

Siglecs have also been proposed to serve as potential disease biomarkers. Siglecs are mainly known as transmembrane proteins, though, a few Siglecs can be found in soluble form in human serum and this has been linked to certain human diseases^{245,292–294}. For example, rheumatoid arthritis patients showed elevated levels of soluble Siglec-9 in serum and synovial fluid compared to healthy donors²⁹². Also, high soluble Siglec-8 levels in serum have been shown to be associated with hypereosinophilic diseases²⁹⁴. In addition, elevated soluble Siglec-1 levels in plasma/serum have been associated with systemic lupus erythematosus (SLE), with European SLE patients showing higher soluble Siglec-1 levels compared to non-Europeans, suggesting an ethnic background link²⁹³. Higher levels of soluble Siglec-7 were found in serum of AIDS patients compared to healthy donors²⁴⁵. In humans, Siglec-7 is mainly known to be expressed in NK peripheral blood cells²⁹⁵, although its gene expression has also been reported in different tissues, including the colon, according to the Human protein atlas project (Fig. 68)²⁹⁶. With regards to tumour tissues, there is a variation in Siglec-7 gene expression in the on-tumour compared to the off-tumour tissues depending on the intestinal region, with colon and rectum showing lower levels of Siglec-7 gene expression in the on-tumour compared to the off-tumour tissue samples (Fig. 69). Recently, Siglec-7 expression in macrophages in tumour tissue has been proposed as predictive biomarker for the efficacy of cancer vaccination against metastatic CRC¹¹¹.

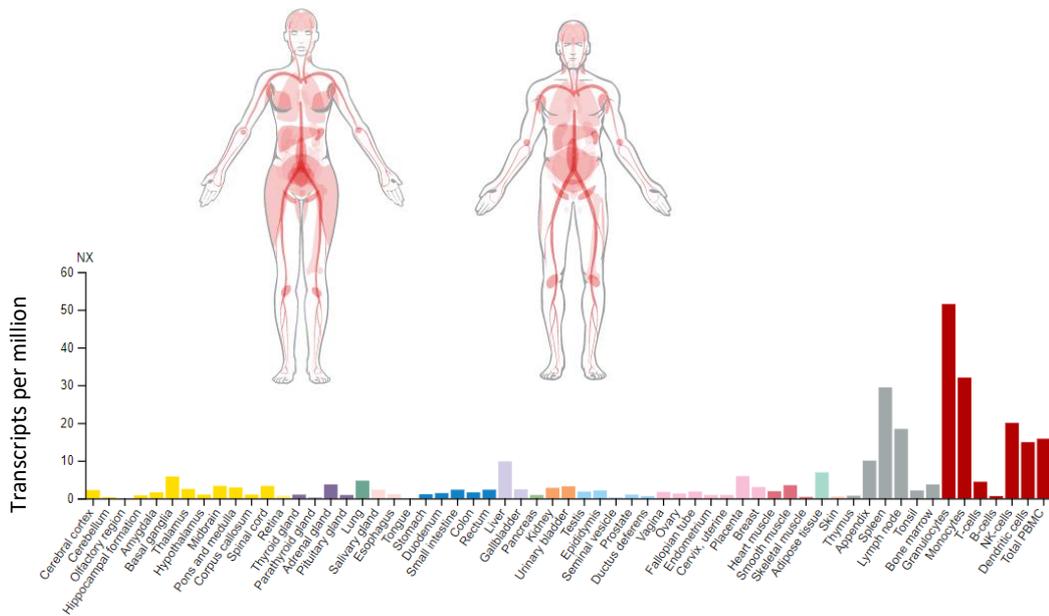


Figure 68 | Siglec-7 gene expression in healthy human tissues. Normalised expression (NX) levels of 55 different tissues and 6 blood cell subsets. (Extracted from the Human Protein Atlas project²⁹⁶).

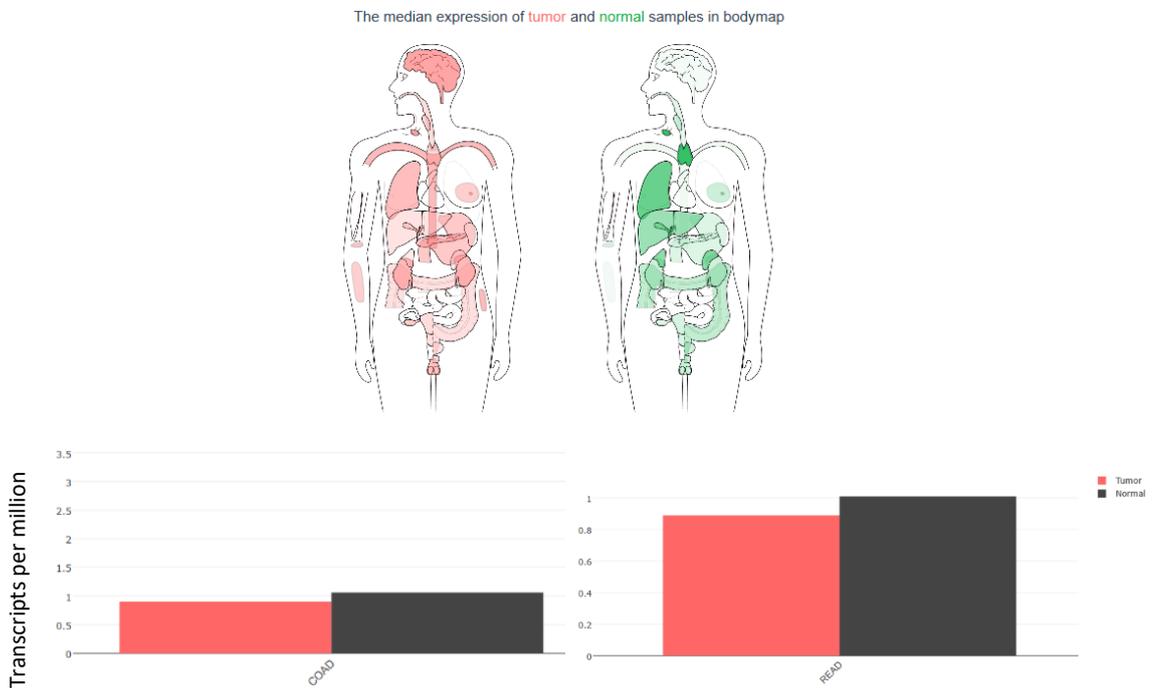


Figure 69 | Siglec-7 gene expression levels in colorectal adenocarcinomas. Siglec-7 transcriptomic profile in the on-tumour (referred as tumor in red) and off-tumour (referred as normal in black) tissues. For the colon adenocarcinomas (COAD), the sample size were 275 on-tumour tissues and 349 off-tumour tissues. For the rectum adenocarcinomas (READ), the sample size were 92 on-tumour tissues and 318 off-tumour tissues. (Extracted from Tang *et al.* 2017, GEPIA web server).

Following our findings that Siglec-7 is involved in the immune response induced by *F. nucleatum* ssp. (Chapter 6), we first investigated *Fusobacterium* spp. abundance and Siglec-7 expression in colorectal cancer specimens derived from patients at the Norfolk and Norwich University Hospital (NNUH). Next, we sought to identify soluble Siglec-7 levels and anti-*F. nucleatum* ssp. IgG antibodies in the matched serum samples from the CRC cases.

7.2 Results

7.2.1 *F. nucleatum* spp. relative abundance is increased in on-tumour site of CRC tissues

A total of 13 CRC cases were examined for the presence of *Fusobacterium* spp. in the on- and off-tumour sites of tissues from patients with CRC and their relative abundance (compared to all bacteria-Eubacteria) determined by qPCR. The examined samples were considered positive for *Fusobacterium* spp. when the Ct value was lower than the non-template control value of 35. The age range of the cases was 69-83 years old, and the tumour location and stage/grade of tumour determined by the NNUH pathology unit (Table 7). Our results showed that all CRC cases were *Fusobacterium* spp. positive at the on-tumour site, and three cases (#368, #011, #250) were *Fusobacterium* spp.-negative at the off-tumour site (Table 7). Interestingly, the three off-tumour *Fusobacterium* spp.-negative cases were left-sided colon cancer types whereas the off-tumour *Fusobacterium* spp.-positive cases varied in terms of tumour location (Table 7). *Fusobacterium* spp. relative abundance was significantly higher at the on-tumour compared to off-tumour site ($p < 0.0001$) of the cases as calculated with pooled samples using 2-way ANOVA (Fig. 70A). Multiple comparison analyses showed significant higher relative abundance of *Fusobacterium* spp. in the on-tumour compared to the off-tumour site only in the cases referred previously as off-tumour-*Fusobacterium* spp. negative (#368, #011, #250) ($p < 0.0001$) and in the cases with advanced, T4a (#337) or T4b (#349) invasion stage ($p < 0.05$) (Fig. 70B).

The bacterial DNA extracted from the CRC tissue for the qPCR analysis was also used for the 16S ribosomal RNA (or 16S rRNA) gene sequencing analysis conducted in-house by our collaborator at QIB using Illumina sequencing. At the genus level, the 16S analysis showed that *Fusobacterium* spp. was the main genus showing statistically significant higher abundance in the on-tumour site compared to the off-tumour site (referred as normal), in agreement with the qPCR results (Fig. 71A). Operational taxonomic unit (OTU) clustering of *Fusobacterium* spp. showed that OTU3, which is assigned to *F. nucleatum*, was significantly increased in the on-tumour compared to the off-tumour site (referred as normal) (Fig. 71B).

Table 7 | List of the colorectal cancer cases and *Fusobacterium* spp. abundance

Case #	Age	Gender	Tumour location	Invasion Stage	<i>Fusobacterium</i> spp. negative at off-tumour site
19BR 337	82	Male	Middle Rectal	T4a	
19BR 338	83	Male	Caecum	T3	
19BR 349	71	Male	Caecum	T4b	
19BR 367	76	Female	Upper Rectal	T1	
19BR 368	76	Female	Middle transverse colon/splenic flexure	T3	✓
19BR 380	70	Female	Upper caecum / ascending colon	T1	
19BR 394	71	Male	Upper Rectal	T3	
20BR 011	81	Male	Rectum	T1	✓
20BR 049	72	Male	Ascending colon	T3	
20BR 050	74	Female	Caecum	T4a	
20BR 112	81	Female	Caecum	T4a	
20BR 191	63	Male	Ascending colon	T3	
20BR 250	67	Female	Sigmoid colon	T2	✓

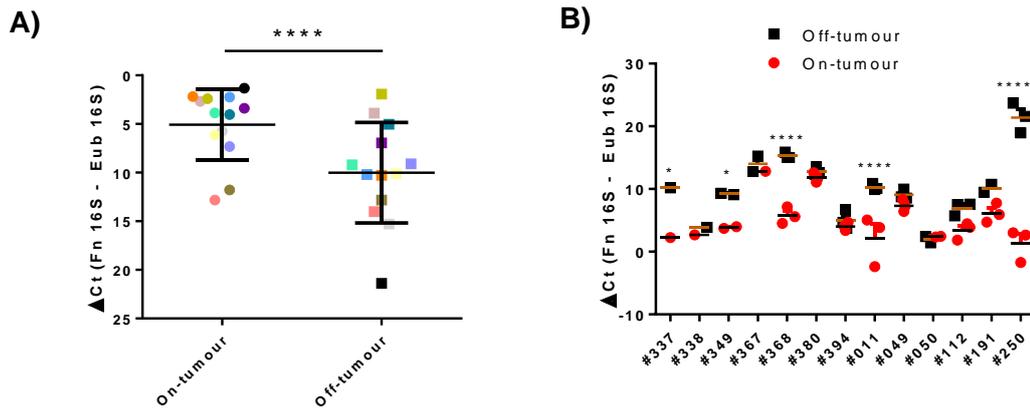


Figure 70 | *Fusobacterium* spp. relative abundance in the on- and off-tumour sites. A) All 13 colorectal cancer cases. Each colour represents the mean ΔCt of technical replicates of each case. B) ΔCt values of on-tumour (in red) and off-tumour (in black) sites of each case in replicates. Fn, *F. nucleatum*. Statistics: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; n.s., not statistically difference.

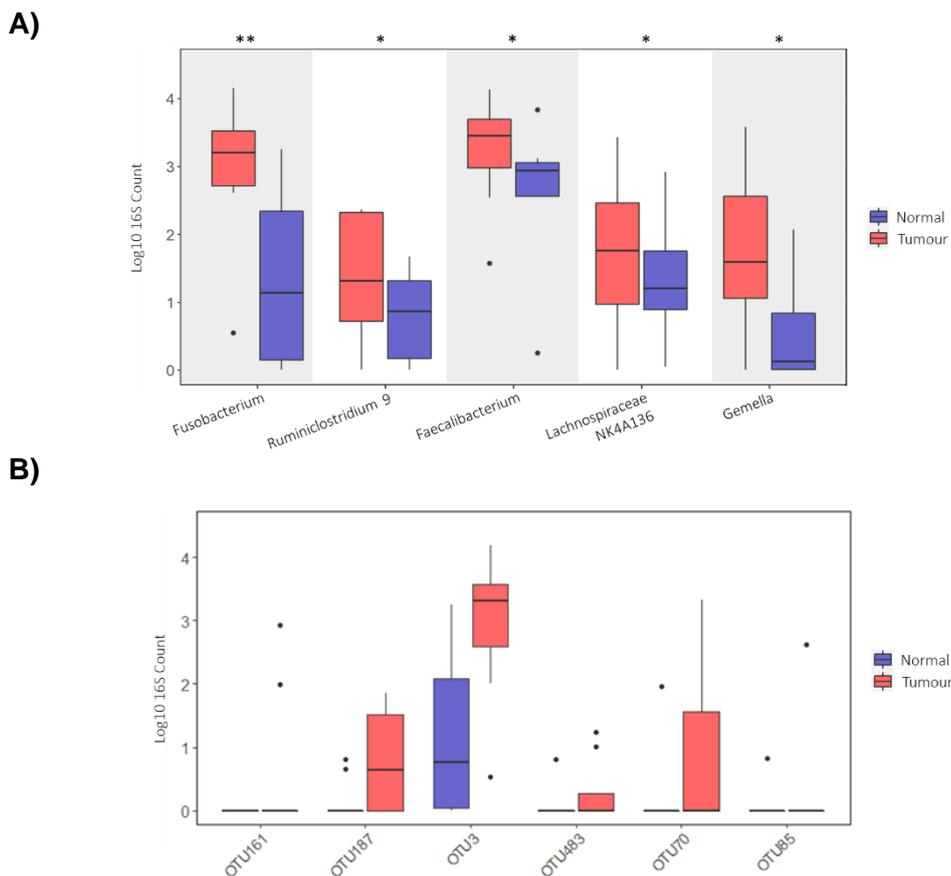


Figure 71 | *Fusobacterium* spp. relative abundance using 16S RNA sequence. A) The 5 genera found to be statistically significant different when compared the on-tumour (referred as tumour, in red) to the off-tumour (referred as normal, in purple) site. B) *Fusobacterium* OTUs showing the main differences in on-tumour compared to the off-tumour sites.

7.2.2 Colorectal cancer cases showed high anti-*F. nucleatum* ssp. IgG levels

Having shown the high abundance of *F. nucleatum* in CRC tissues, we next investigated whether we could detect variations in the level of anti-*F. nucleatum* IgG in human serum of CRC cases, as compared to healthy donors, using an ELISA-based assay. Given the limited accessibility to human samples (due to COVID-19 outbreak), the analysis was performed with 3 of the CRC samples. The ELISA plate was coated with either the *F. nucleatum* strains or *F. nucleatum*-derived LPS, considered as antigens for the interaction with the potential anti-*F. nucleatum* IgG antibodies in the human serum. The controls included PBS, *R. gnavus* ATCC 29149, or *E. coli*-derived LPS. Anti-*F. nucleatum* IgG antibody levels were detected spectrophotometrically as described in Materials and Methods (section 2.5.7).

Among the three CRC cases (#112, #191, #250) tested, case #112 (T4a caecum cancer) and case #250 (T2 sigmoid colon cancer) showed a significant increase in anti-*F. nucleatum* IgG levels, particularly when *F. nucleatum*-ATCC 25586 was used as antigen ($p < 0.0001$) as compared to the PBS control or healthy donors (Fig. 72A,B). When *F. nucleatum*-ATCC 10953 was used, IgG levels showed a significant increase ($p < 0.0001$) for case #112 when compared to the PBS and healthy #175 controls (Fig. 72A), and for case #250 ($p < 0.05$) when compared to the PBS and the healthy #180 or #014 controls (Fig. 72C). In the presence of whole *F. nucleatum*-ATCC 51191, IgG levels showed a significant increase for both #112 and #250 cases ($p < 0.05$) when compared to the PBS control, as well as for case #112 ($p < 0.005$) and case #250 ($p < 0.01$) when compared to the healthy #175 and #179 controls, respectively (Fig. 72A,B).

When *F. nucleatum* ssp.-derived LPS was used, only *F. nucleatum* ATCC 10953-derived LPS showed a significant increase in the IgG levels for both of #112 ($p < 0.01$) and #250 ($p < 0.0001$) CRC cases, when compared to the PBS control (Fig. 72C). However, only the #250 CRC case showed significant higher IgG levels ($p < 0.0001$) when compared to the healthy #180 or #013 or #014 controls (Fig. 72C). Interestingly, the #250 CRC case also showed a significant increase in IgG levels against *E. coli*-LPS, used as a control, when compared to the PBS ($p < 0.005$) or healthy ($p < 0.0001$) controls (Fig. 72C). These results suggest that anti-*F. nucleatum* ssp. IgG antibody recognition is strain/molecule-specific.

While all samples tested, CRC or healthy, showed a statistically significant increase in IgG levels against the control species, *R. gnavus* ATCC 29149, when compared to the PBS control, there was no statistically significant difference in the anti-*R. gnavus* IgG levels between healthy and CRC cases (Fig. 72C).

Taken together, these data indicate that two (#112, #250) out of the three CRC cases tested showed increased anti-*F. nucleatum* IgG levels as compared to PBS control and only the #250 case showed statistically significant high levels of anti-*F. nucleatum* ATCC 10953-LPS IgG when compared to the healthy controls. Interestingly, the #250 CRC case was the only one among the three CRC cases tested that had high *Fusobacterium* spp. levels in the on-tumour tissue and was off-tumour *Fusobacterium* spp.-negative (see section 7.2.1.). These results suggest that the production of anti-*F. nucleatum* IgGs may be associated with an on-tumour localised *Fusobacterium* spp. enrichment, although the analysis of a bigger sample size is required to support this finding along with the assessment of different tumour locations and invasion stages which could be factors involved in this association.

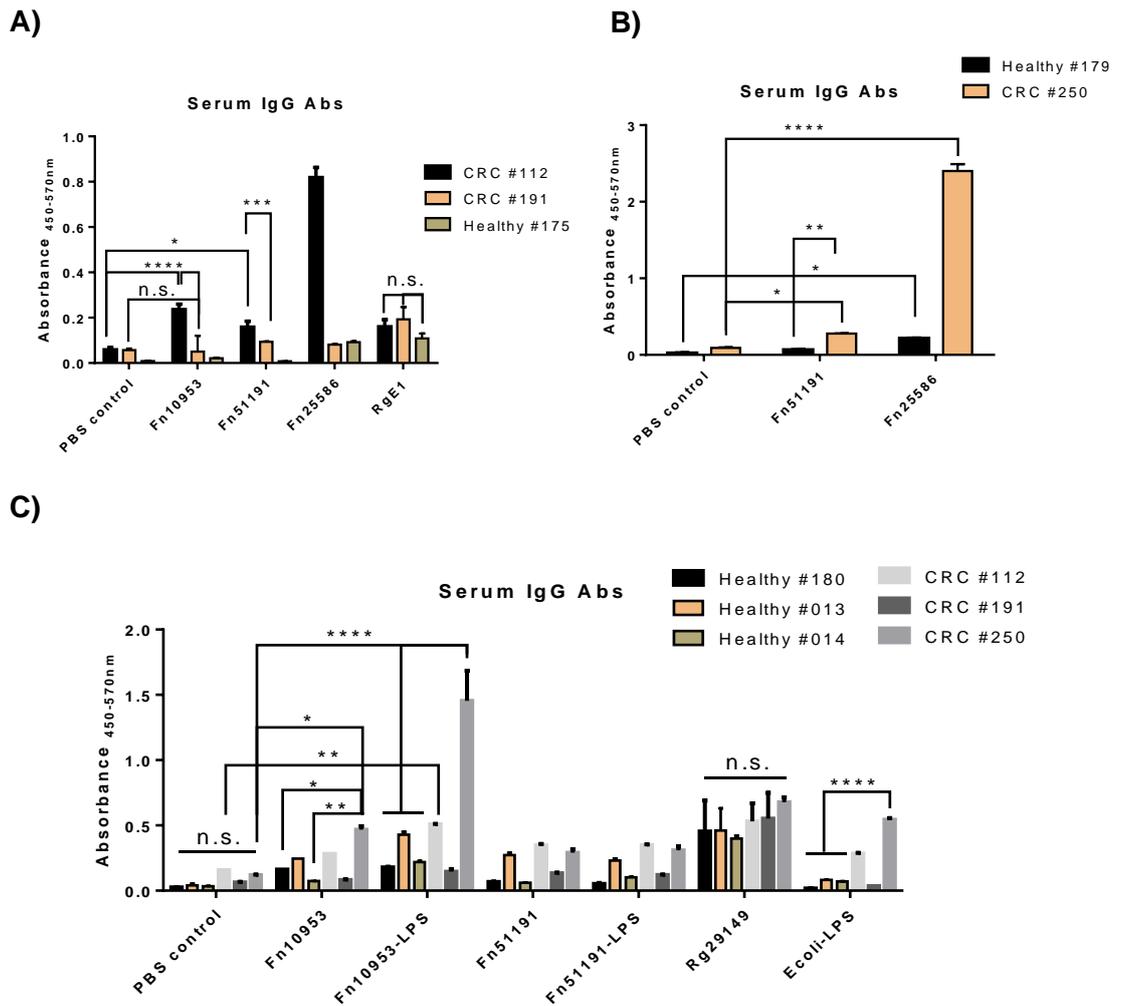


Figure 72 | anti-*F. nucleatum*-IgG levels in serum of healthy or colorectal cancer human samples.

Whole *F. nucleatum* ATCC 10953, ATCC 51191, ATCC 25586 or *R. gnavus* E1, ATCC 29149, *F. nucleatum* ATCC 10953 or ATCC 25586-derived LPS were used as antigens for anti-*F. nucleatum*-IgG identification in A) the CRC cases #112 and #191 including as a control healthy donor #175 sample and B) the CRC case #250 including as a control healthy donor #179 sample. C) All three CRC cases (#112, #191, #250) and healthy controls (#180, #013, #014). Fn, *F. nucleatum*; Rg, *R. gnavus*. Statistics: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; n.s., not statistically difference.

7.2.3 Expression of Siglec-7 in tumour tissues shows heterogeneity

To determine Siglec-7 expression in the on-tumour and off-tumour sites, two approaches were followed: i) cell surface Siglec-7 expression in lamina propria leukocytes was determined by flow cytometry, and ii) Siglec-7 mRNA relative abundance (ΔCt) (compared to the housekeeping gene PGK1) was determined by qPCR.

The isolated lamina propria leukocytes from the on-tumour and off-tumour sites were stained and gated as shown in Fig. 73 and the cell surface Siglec-7 expression was analysed in the CD45+CD103(+/-)CD11b(+/-) cell populations by flow cytometry. The results showed that in 7 out of the 9 cases tested, Siglec-7 expression was higher in the off-tumour sites compared to the on-tumour site, with CD45+CD103-CD11b+ being the main cell population showing Siglec-7 expression. Cases #367 and #368 showed a cell subset specificity in Siglec-7 expression between the on-tumour and off-tumour sites (Table 8).

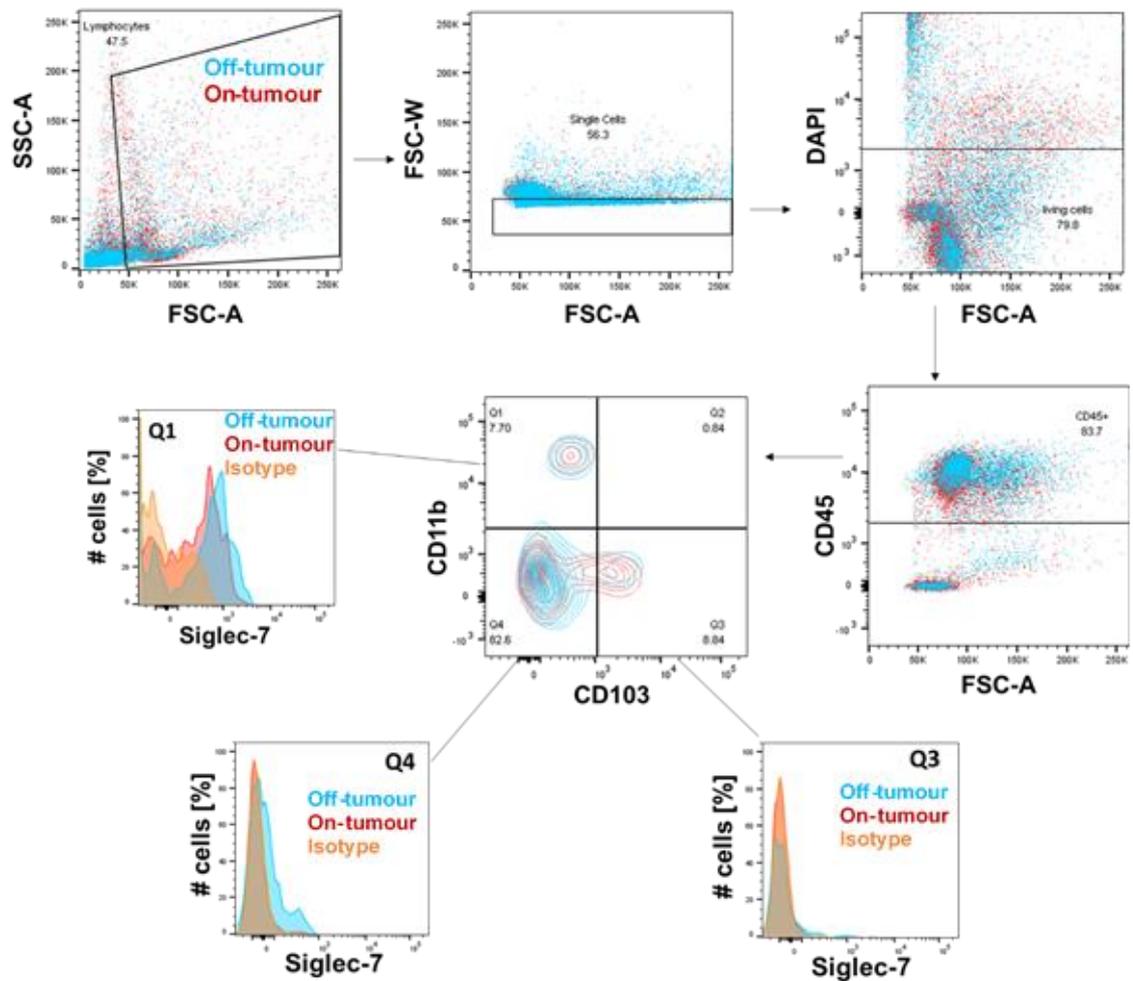


Figure 73 | Gating strategy of the multicouleur stained gut lamina propria leukocytes.

The leukocyte population was selected in the side scatter (SSC) versus forward scatter (FSC) dot plot, following single cell selection on FSC-W versus FSC-A dot plot the negative in DAPI population (living cells) and positive to CD45 population was selected. Siglec-7 expression on on-tumour (in red), off-tumour (in blue) was identified in the CD45 subpopulations (CD103+/- and CD11b+/-). Isotype control in orange.

Table 8 | List of colorectal cancer cases with the *F. nucleatum* abundance and Siglec-7 expression in the tested gut lamina propria cell subsets

Case #	Tumour location	Invasive Stage	<i>F. nucleatum</i> abundance	Siglec-7 expression			
				CD103-CD11b +	CD103-CD11b -	CD103 + CD11b +	CD103 + CD11b -
19BR 349	Caecum	T4b	On- > off-tumour (p<0.05)	Off- > on-tumour	No expression	No population	No expression
19BR 367	Upper Rectal	T1	n.s.	On- > off-tumour	Off- > on-tumour	On- > off-tumour	Off- > on-tumour
19BR 368	Middle transverse colon/splenic flexure	T3	On- > off-tumour (p<0.0001)	Off- > on-tumour	No expression	On- > off-tumour	Off- > on-tumour
19BR 380	Upper caecum / ascending colon	T1	n.s.	No difference	Off- > on-tumour	No difference	No difference
19BR 394	Upper Rectal	T3	n.s.	Off- > on-tumour	No difference	No difference	No difference
20BR 049	Ascending colon	T3	n.s.	Off- > on-tumour	Off- > on-tumour	No population	No expression
20BR 050	Caecum	T4a	n.s.	Off- > on-tumour	Off- > on-tumour	No population	No expression
20BR 112	Caecum	T4a	n.s.	Off- > on-tumour	Off- > on-tumour	No population	No population
20BR 191	Ascending colon	T3	n.s.	Off- > on-tumour	No difference	No population	No expression

Next, we investigated the mRNA Siglec-7 expression levels in the resected tissues (on-tumour versus off-tumour). We first attempted to extract RNA from the lamina propria leukocytes isolated from the CRC tissues but, despite several attempts, low yields and mRNA degradation (Fig. 74A) prevented us to continue with this approach.

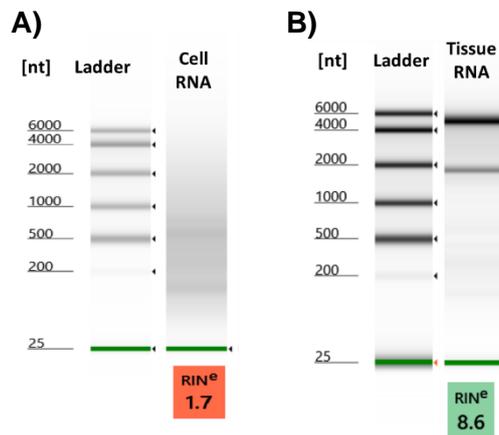
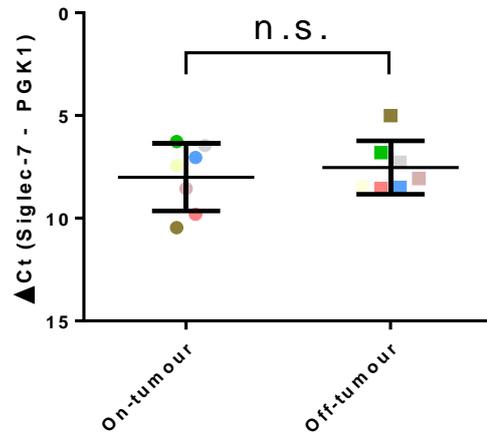


Figure 74 | RNA electrophoresis and RNA integrity numbers by TapeStation. RNA extracted from A) gut lamina propria leukocytes and B) gut tissue.

Instead, the qPCR analysis was carried out using RNA extracted from the whole tissues (using on- or off-tumours) of 7 cases, as approximately 30 μ g material could be obtained with RNA integrity number (RIN) values ranging from 6.5 to 8.6 (Fig. 74B). The pooled data of Δ Ct values did not show any statistically significant differences between the on-tumour and off-tumour sites (Fig. 75A). When each case was compared individually for Siglec-7 mRNA expression between the on-tumour and off-tumour sites, significant differences in Siglec-7 expression were observed in 2 out of the 7 cases, with case #011 showing higher Siglec-7 expression in the on-tumour site compared to off-tumour ($p < 0.05$) and case #250 showing higher Siglec-7 expression in the off-tumour site compared to on-tumour ($p < 0.05$) (Fig. 75B).

A)



B)

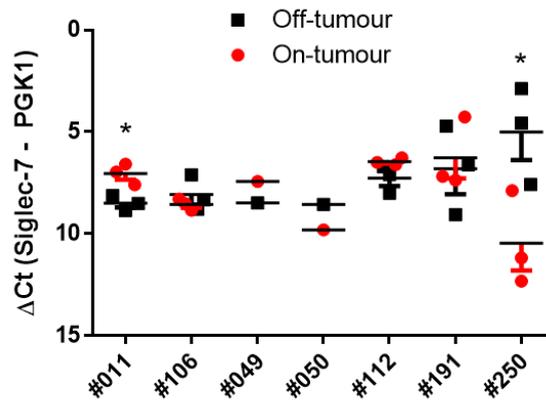


Figure 75 | Relative Siglec-7 mRNA expression in the on- and off-tumour tissue sites of colorectal cancer cases.

A) Each colour represents the mean ΔCt of technical replicates of each case. B) ΔCt values of on-tumour (in red) and off-tumour (in black) sites of each case in replicates. Statistics: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; **** $p < 0.0001$; n.s., not statistically difference.

7.2.4 Soluble Siglec-7 levels are increased in colorectal cancer cases

Here we sought to investigate the levels of secreted serum Siglec-7 in serum of the 3 CRC cases mentioned above as compared to healthy donor serum samples, using a commercially available kit which is based on the sandwich ELISA system. The results showed statistically significant higher Siglec-7 serum levels ($p < 0.05$) in CRC cases as compared to the healthy controls (Fig. 76). The three CRC samples corresponded to distinct cancer type and invasion stages (T4a caecum, T3 ascending colon or T2 sigmoid colon cancers).

While our results clearly showed elevated soluble Siglec-7 serum levels in CRC cases when compared to healthy donors, a bigger sample size is required to support the data and reduce the effect of other factors, such as age, that could affect the levels of protein expression.

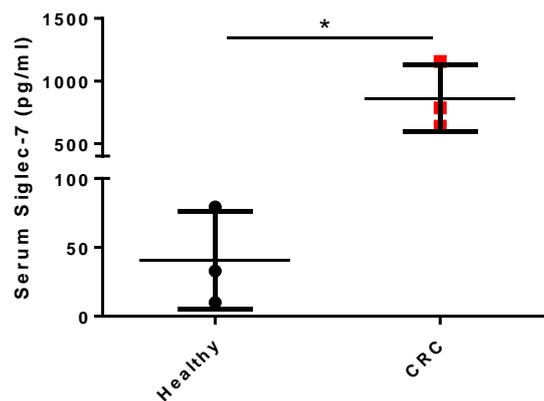


Figure 76 | Soluble serum Siglec-7 levels in colorectal cancer or healthy samples.

Soluble Siglec-7 were quantified in three colorectal cancer cases (in red) or three healthy donors (in black). Each point represents the mean of duplicates of each case. Statistics: * $p < 0.05$.

7.3 Discussion

The gut microbiota has been associated with colorectal cancer development as well as other types of cancers, and appears to play a role in the efficacy of immunotherapy¹⁶⁶. Anticancer therapies aiming at reinstating immunosurveillance have revolutionised the treatment of melanoma but are profoundly influenced in their efficacy by the gut microbiota²⁹⁷. Recently, Baruch *et al.* 2020 showed that in melanoma patients, receiving a faecal microbiota transplantation (FMT), leading to a beneficial alteration of the host immune cell profile in both gut lamina propria and tumour site²⁹⁸. In addition, *F. nucleatum* appears to contribute to the chemoresistance of CRC^{214,299,300}. It is therefore important to develop specific and non-invasive tools which can detect *F. nucleatum*.

Our sequencing analyses on DNA extracted from fresh 13 CRC tissues revealed that all cases were *Fusobacterium* spp. positive with the dominance of the OTU3 cluster, which includes the *F. nucleatum* spp.³⁰¹. In addition, we showed a statistically significant increase of *Fusobacterium* spp. in the on-tumour compared to the off-tumour tissue, as it has been previously reported³⁰². The cases showing high *Fusobacterium* spp. abundance in the on-tumour site were characterised by an advanced invasion stage T4, as previously reported^{288,289}. Here the cases showing off-tumour-negative to *Fusobacterium* spp. were all left-sided colorectal cancer types. This difference in location will be worth exploring further (with a larger sample set) as a study by Mima *et al.* 2016 showed an increase of *Fusobacterium* spp. abundance from rectum to caecum¹⁹⁹ while a study from Gao *et al.* 2017 did not show significant differences in *Fusobacterium* spp. distribution along colon¹⁷².

In addition, we showed increased anti-*F. nucleatum* IgG and anti-*F. nucleatum* ATCC 10953-LPS IgG levels against the #250 case when compared to the healthy controls. The #250 case, sigmoid colon cancer with T2 invasive stage, was characterised by high *Fusobacterium* spp. levels in the on-tumour tissue while off-tumour was *Fusobacterium* spp.-negative. Therefore, *F. nucleatum* ATCC 10953-LPS could be used as a specific diagnostic marker for the detection of IgG levels in *Fusobacterium* spp. positive CRC cases.

Recently, Siglec-7 was proposed as a predictive biomarker for immunotherapy efficacy in colorectal cancer, where survival of metastatic CRC patients undergoing immunotherapy was shown to be shorter in the patients with high Siglec-7 expression in the tumour localised macrophages as compared to the low Siglec-7 expression patients¹¹¹. Here, within the limitations of our study (in terms of numbers), no significant difference in Siglec-7 RNA expression was observed between on- and off-tumour tissues, while the cell surface Siglec-7 expression on colonic lamina propria cells showed heterogeneity across CRC cases. However, soluble Siglec-7 levels were significantly higher in the three CRC serum tested compared to the healthy serum samples. These CRC cases were all positive to *Fusobacterium* spp. The presence of Siglec-7 in serum was previously reported in patients infected with HIV showing a positive association between increased Siglec-7 levels and HIV infection²⁴⁵. This is the first report showing elevated lectin levels in CRC, and further research is warranted, including a bigger sample size to explore whether soluble Siglec-7 could be used as a potential diagnostic marker for CRC.

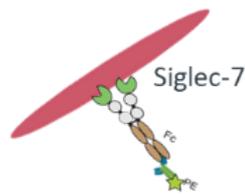
Overall, our preliminary results using clinical human samples from patients with CRC support the need for a closer investigation of *Fusobacterium* spp.-positive cases in the CRC development, and point towards Siglec-7 as a novel therapeutics/diagnostics target.

CHAPTER 8

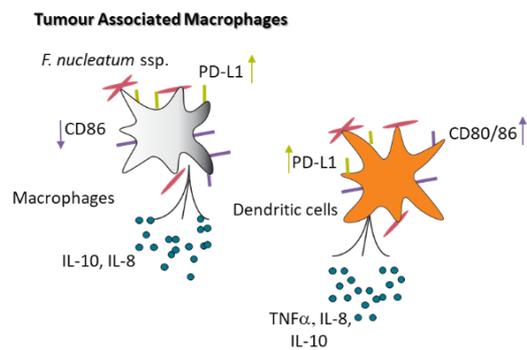
-

Key findings and General Discussion

8.1 Key findings

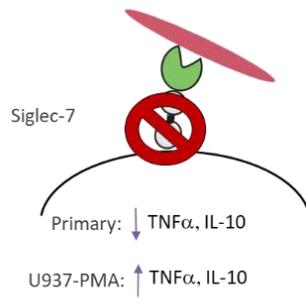


- *F. nucleatum* ATCC 10953, ATCC 25586 and ATCC 51191 and their derived OMVs bound specifically to Siglec-7.
- *F. nucleatum* ATCC 10953 and ATCC 51191-derived LPS bound to Siglec-7.
- *F. nucleatum* ssp. bind to Siglec-7 V-set in a sialic acid independent manner.

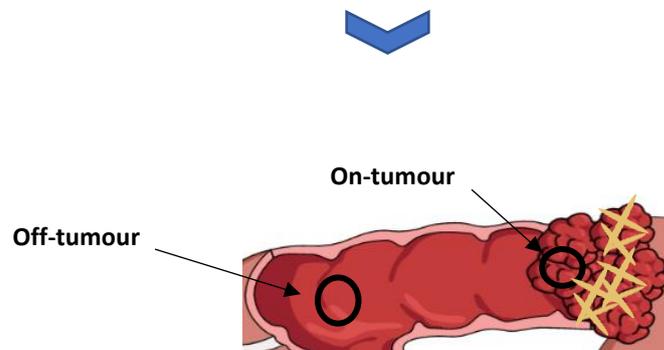


- Siglec-7 is involved in *F. nucleatum* ssp. association with mammalian cells.
- *F. nucleatum* ssp. and their derived LPS and OMVs modulate host cell phenotype in a cell subset manner.





- Anti-Siglec-7 antibody modulates myeloid cell response.
- *F. nucleatum* ssp. modulate moDCs response in a Siglec-7-dependent manner.



- *F. nucleatum* spp. relative abundance is increased in on-tumour site of CRC tissues
- Expression of Siglec-7 in tumour tissues shows heterogeneity
- Soluble Siglec-7 levels are increased in colorectal cancer cases
- Colon cancer cases showed high anti-*F. nucleatum* ssp. IgG levels

8.2 General Discussion

In 1994, the international agency for research on cancer announced the first carcinogenic bacterium, *Helicobacter pylori*³⁰³. This bacterium which colonises half of the world's population can also cause infection and ultimately gastric cancer³⁰⁴. Since then, the field of microbiota and cancer has grown extensively with metagenomic studies revealing an association between distinct bacteria populations and the development of a range of cancer types³⁰⁵. Further, the gut microbiota has been shown to have a direct effect on GI cancers but also an indirect effect on liver cancer³⁰⁶ and the treatment of melanoma¹⁶⁶.

In the past years, colorectal cancer (CRC)'s incidence has increased in young adults¹⁶⁴, and since the highest proportion of cases are sporadic with no hereditary background, these studies highlight the importance of preventable environmental factors in the development of CRC¹⁶³. It is now clear that CRC is associated with a distinct bacterial signature consisting mainly of *B. fragilis* and *F. nucleatum* bacterial species^{171,307}. While *B. fragilis* has mainly been associated with CRC through its toxin production¹⁸⁵, *F. nucleatum* has been involved in CRC progression mainly through its ability to recruit tumour infiltrating immune cells¹⁹⁵. Efforts to identify the proteins involved in *F. nucleatum* associated tumour progression, have focused on the interaction between *F. nucleatum* ssp. and host epithelial cells, revealing the role of the bacterial proteins FadA and Fap2 in these interactions³⁰⁸. However, the mediators of the interaction between *F. nucleatum* and human immune cells remain largely unknown.

Mammalian lectins fulfil several physiological functions, such as adhesion to other cells, endocytosis, and when expressed on immune cells can also mediate immune modulation through their intracellular signaling³⁰⁹. Among the diversity of human lectins, several types have been implicated in cancer progression and were selected as targets of this study. One such type is the Siglec family which has been associated with tumour progression through interactions with glycans expressed on host malignant cells, resulting in pro-tumour immunity⁹¹. Galectin-3, expressed extra- or intra-cellularly, has also been shown to be increased in CRC³¹⁰, and involved in cancer progression and mainly metastasis through its ability to enhance the cell adhesion³¹¹. Dectin-2, a transmembrane protein

bearing an intracellular signaling which activates immunity, is mainly known for its induction of a pro-inflammatory phenotype upon interactions with endotoxin or yeast as we have previously shown^{73,222}, and has also been implicated in an anti-tumour response by internalising cancer cells⁷⁸.

To date, interactions between CRC-associated bacteria and host lectins remain undiscovered. Here, we showed that among *B. fragilis* NCTC 9343 and the three subspecies of *F. nucleatum* ATCC 10953, ATCC 25586 and ATCC 51191 tested, only *F. nucleatum* strains showed binding to Siglecs with specificity to Siglec-7, when compared to the Galectin-3 and Dectin-2 lectins. We showed that the interaction occurred via the V-set domain of Siglec-7 and involves *F. nucleatum* LPS as bacterial ligand. The V-set domain of Siglec-7 is known to bind to mammalian sialic acids with high affinity to α 2,8-disialyl and branched α 2,6-sialyl residues expressed on host cells¹⁰², through the Arg-124 residue located on the F β strand and its C-C' loop structure²⁴⁸.

Sialic acids found in terminal position of glycoconjugate structures are mainly known for their biological roles in mammalian cells, but bacteria also evolved to express sialic acids either in terminal or internal positions in LPS or capsular polysaccharides molecules³¹², possibly for mammalian cell mimicry and ultimately evasion from the host immune response. In line with this, binding to Siglec-7 has been reported for *Campylobacter jejuni* via its LPS in a sialic acid-dependent manner¹⁰⁵, but Siglec-7 has also been reported binding to GBS-derived β -protein⁹⁸, yeast particles and *E. coli* in a non-sialic acid manner¹⁰⁶. We therefore hypothesised that the interaction between Siglec-7 and *F. nucleatum* strains may be mediated by sialic structures present in *F. nucleatum* LPS.

Sialic acid expression in *F. nucleatum* ssp. is strain-dependent with *F. nucleatum* ATCC 10953 LPS carrying internal sialic acids²¹⁹, whereas ATCC 25586 LPS displays a sialic acid-like molecule named fusaminic acid²²⁰. Here, we characterised the structure of *F. nucleatum* ATCC 51191 LPS which revealed a novel repeat of the O-antigen structure of the LPS, [\rightarrow 4)- β -d-GlcpNAcA-(1 \rightarrow 4)- β -d-GlcpNAc3NAAlaA-(1 \rightarrow 3)- α -d-FucpNAc4NR-(1 \rightarrow], (R=acetylated 60 %), and a bis-phosphorylated hexa-acylated lipid A moiety²⁴⁷. The O-antigen of *F. nucleatum* ATCC 51191 LPS showed the presence of the active sugar group

uronic acid but no sialic acid or sialic acid derivatives. According to the Bacterial Carbohydrate Structure Database³¹³, this trisaccharide represents a novel O-antigen structure for bacterial LPS. These findings were also consistent with our bioinformatic analyses showing that only the *F. nucleatum* ATCC 10953 strain carries the genes encoding the sialic acid biosynthetic and modification enzymes (NeuA, NeuB, NeuC, NeuD). STD-NMR analyses further revealed that Siglec-7-Fc bound to the O-antigen sugars in *F. nucleatum* ssp.-derived LPS including sialic acid for *F. nucleatum* ATCC 10953 LPS, but also novel ligands such as fusaminic acid for *F. nucleatum* ATCC 25586 LPS, the alanine-containing uronic acid for *F. nucleatum* ATCC 51191 LPS, and fucosamine, a sugar common to LPS from both *F. nucleatum* ATCC 10953 and ATCC 51191 strains. The identification of these sugar epitopes as ligands for Siglec-7 has not been reported before, and more work is needed to determine their mode and site of interaction. The long and extended C-C' loop present in the V-set domain of Siglec-7²⁴⁸, may allow the broad ligand specificity of this lectin and its interactions with internal glycans. Recently, Yamakawa *et al.* 2020 showed a novel Siglec-7 region involved in the recognition of sialic acids²⁴⁹. It will therefore be interesting to determine the relative contribution of these binding sites in mediating these novel non-sialic acid sugar interactions with Siglec-7.

LPS is also found on the outer membrane vesicles (OMVs) produced by Gram-negative bacteria, such as *F. nucleatum* ssp. These vesicles have nanoscale sizes and can contribute to bacterial pathogenicity¹⁵⁸. In addition, *F. nucleatum* OMVs have recently been reported to induce host immune response in epithelial cells²¹⁶. Having demonstrated that *F. nucleatum* ssp. bound to Siglec-7 via LPS, we investigated to which extent this interaction could be recapitulated using *F. nucleatum* OMVs. The OMVs we purified from *F. nucleatum* ssp. displayed 60-70% of LPS and a diameter ranging from 30 to 250 nm. Importantly, we showed that *F. nucleatum* OMVs bound to Siglec-7 as demonstrated with LPS or the whole bacteria. We further demonstrated that Siglec-7 contributes to the host immune response induced by *F. nucleatum* ssp., *F. nucleatum* ssp.-derived LPS or OMVs, suggesting that OMVs could act as mediators of *F. nucleatum*-induced immune response. The use of advanced models that mimic the human physiology such as the gut-on-chip system³¹⁴ will help investigate the mechanisms by which

F. nucleatum OMVs reach the gut lamina propria for immune cell interaction in CRC or healthy state and the role they play in tumour progression or mediation.

Following binding to their ligands, members of the Siglec family have been shown to modulate internalisation of the receptors upon ligand binding. For example, in the case of CD33-related Siglecs expressed on immune cells, endocytosis of the protein together with its sialylated antigen is mainly mediated by the intracellular domain involving the ITIM phosphorylation²⁸⁰. Siglec-7 expressed on moMφs and NK cells has been shown to induce HIV internalisation resulting in an enhancement of immune activation^{245,281} but to date there are no data on the role of Siglec-7 in internalisation of bacteria. Here, using microscopy and imaging flow cytometry, we showed that *F. nucleatum* could associate with monocyte-derived dendritic cells (moDCs) and human macrophage-like U937 cells expressing Siglec-7 and that internalisation was affected in the Siglec-7-silenced and Siglec-7 depleted cells, respectively. These data therefore suggest a role of Siglec-7 in internalisation of *F. nucleatum* in both DCs and macrophages. U937 cells engulfment of bacteria is expected as macrophages are specialised in phagocytosis for ultimate elimination of the phagocytosed foreign particles in their high lysosomal proteolytic activity environment³¹⁵. DCs' prevalent function is to present antigens and activate adaptive immunity, and it has been suggested that phagocytosis of foreign particles, such as bacteria, occur to the degree needed for immune response initiation³¹². Further research is needed to understand the underlying mechanisms involved in Siglec-7-mediated *F. nucleatum* internalisation by macrophages and dendritic cells and the potential role of the Siglec-7 ITIM domain in mediating endocytosis.

Fusobacterium spp. associated with CRC modulates host immune response by recruiting tumour infiltrating immune cells and inducing local IL-10 and TNF α production^{195,275}. In this work, we showed that, in moDCs, *F. nucleatum* induced a pro-inflammatory phenotype characterised by the induction of TNF α , IL-8, CD86 and CD80. We also observed that *F. nucleatum* ATCC 10953, the longest across the strains, caused the highest level of moDC death compared to the ATCC 25586 and ATCC 51191 strains. These results are in line with a previous study showing that *F. nucleatum* ssp. affect cell viability by inducing neutrophil apoptosis and necrosis in a strain-specific manner²⁶², suggesting a

potential strain-specificity in the immune activation and evasion strategy induced by *F. nucleatum*. In contrast to moDCs, *F. nucleatum* induced monocyte-derived macrophages (moMφs) acquired an M2 phenotype characterised by the induction of IL-10, IL-8 and PD-L1 expression and reduction of CD86 expression, in agreement with previous studies using macrophage cell lines²⁶³. In addition, moMφs showed a higher proportion of internalised *F. nucleatum* ssp. when compared to moDCs, which could be due to the acquired M2 phenotype, which is characterised by increased endocytosis levels when compared to M1 macrophages³¹⁶.

Next, we showed that Siglec-7 modulated the *F. nucleatum* ssp.-induced host immune response in a cell subset-specific manner. Siglec-7 depleted U937 cells showed induction of cytokine (TNF α and IL-10) production, while Siglec-7 silencing in moDCs showed decreased production of mainly TNF α but also IL-10. This difference may also be due to the use of primary versus cancer cell lines. To date, only one study showed a role of Siglec-7 in the induction of TNF α by human primary monocytes upon interaction with zymosan yeast particles, in a non-sialic acid manner¹⁰⁶. The contribution of Siglec-7 in promoting a pro-inflammatory response in human monocyte-derived cells upon *F. nucleatum* stimulation is in contrast to the known Siglec-7 immune suppressive function⁹¹. Siglecs have shown to act collaboratively with other pattern recognition receptors (PRRs), such as TLR-4, and to inhibit the activation of the immunostimulatory motif of TLR-4 or other PRRs^{282,283}. In addition, a recent study showed that Siglec-9 agonists could prevent hyperinflammation caused by PRRs upon COVID-19 infection³¹⁷, supporting the evidence of a synergetic effect between Siglecs and other PRRs. It will therefore be important in the future to investigate the molecular pathways induced by the interaction of *F. nucleatum* with Siglec-7 and whether this is a result of a synergetic effect of Siglec-7 with other PRRs and/or if it is dependent on the type of ligand and/or binding mode to the canonical or a non-canonical Siglec-7 region.

LPS is a well-established endotoxin which activates host immunity in a species- and strain-specific manner³¹⁸. In *in vitro* studies, LPS has been shown to influence host immune response in concentrations ranging from 1 to 100 ng/ml. Here, *F. nucleatum* ssp.-derived OMVs as well as *F. nucleatum* ATCC 51191-

derived LPS induced an immune response in moDCs at 10 µg/ml to the same level as that obtained with the whole bacteria, while *F. nucleatum* ATCC 10953-derived LPS showed no stimulation of moDCs. It should be noted that the *E. coli* LPS used as a control showed a significant induction of moDCs at 10-fold lower concentration. Therefore, the effect of high concentrations of *F. nucleatum* ATCC 51191 LPS on the response of moDCs should be taken with caution. In contrast to the profile obtained with moDCs, *F. nucleatum* ATCC 10953 or ATCC 51191-derived OMVs and LPSs (at 10 and 1 µg/ml) led to a M2 macrophage phenotype in moMφs by inducing IL-10, as observed with the whole bacteria. In the future, it will be of interest to determine the capacity of DCs and macrophages to internalise *F. nucleatum*-derived OMVs using imaging flow cytometry, as we have done with the whole bacteria.

Structurally, the O-antigen of LPS extracted from *F. nucleatum* ATCC 10953, ATCC 25586 or ATCC 51191 strains showed high heterogeneity and smaller size as compared to *E. coli*, with *F. nucleatum* ATCC 51191-derived LPS showing the least heterogeneity and smaller size. These differences in LPS O-antigen sugar composition, heterogeneity and size could contribute to the differences in the observed immune responses³¹⁹. CRC development has been also induced by *F. nucleatum* ssp.-associated biofilm formation¹⁸⁶. Moreover, *F. nucleatum* LPS could be involved in CRC development through the contribution of the O-antigen domain to biofilm formation³²⁰. Exploring the strain-specific diversity of *F. nucleatum* O-antigen structures and sizes may provide novel insights into the ability of specific strains of *F. nucleatum* to form biofilms in CRC while providing molecular leads for disrupting or blocking Siglec-7 receptors using specific glycans.

In the last part of the PhD project, we investigated the relevance of our findings in the context of CRC using samples obtained from patients with CRC. *F. nucleatum* ssp. abundance has mainly been reported in CRC cases of advanced cancer stage^{199,288} and in the proximal colon¹⁹⁹. Our qPCR analyses of 13 CRC tissue samples showed that all the on-tumour samples from these cases were *Fusobacterium* spp. positive. CRC clinical studies across the world showed variation in the positive to *Fusobacterium* spp. CRC cases ranging from 13-85%, which may in part be due to geographical location^{307,321}. We showed that *F.*

nucleatum abundance was increased in the on-tumour compared to the off-tumour tissues, in line with metagenomic studies comparing on- versus off-tumour³²² or healthy versus CRC tissues¹⁷⁰. Moreover, our sequencing data showed the dominance of the OTU3 cluster, which includes the *F. nucleatum* spp.³⁰¹. Interestingly, cases positive to both on- and off-tumour *Fusobacterium* spp. were characterised by T4 advanced invasion stage, similar to the observation of studies using qPCR analyses, showing an association of *F. nucleatum* spp. abundance and an advanced TNM stage in CRC^{288,289}. The presence of *Fusobacterium* spp. at the off-tumour site could be indicative of an advanced CRC tumour stage. However, in CRC tissues, and given the small sample size, it was not possible to correlate *F. nucleatum* abundance with Siglec-7 expression.

Microbial ligands can induce humoral immunity and the production of microbe-specific antibodies can be detected in the blood. In agreement with previous studies reporting high anti-*F. nucleatum* IgG and IgA levels in CRC serum samples³²³, two of the three *F. nucleatum* positive CRC cancer cases tested here showed high anti-*F. nucleatum* ssp. IgG levels. While in the published study, the ELISA-based assays were based on the use of *F. nucleatum* bacteria, we used the extracted LPS from *F. nucleatum* as an antigen for IgG detection. Interestingly, high IgG levels were detected in the CRC cases with high on-tumour *F. nucleatum* abundance when using LPS from *F. nucleatum* ATCC 10953. Although preliminary and limited by the small numbers of samples tested, our results may point towards the use of specific *F. nucleatum* LPS as a marker for the detection of IgG levels in support of the development of diagnostic tools. In addition, our preliminary data showed that the levels of circulating Siglec-7 were increased in the serum of CRC patients as compared to healthy donors. Increased Siglec-7 levels have been reported in serum of patients infected with HIV as compared to healthy individuals' serum²⁴⁵ but this is the first time that elevated circulating Siglec levels have been reported in CRC. However large size cohorts will be required to correlate these data with Siglec-7 expression in CRC tissues. This is an area requiring further investigation as despite the role of Siglec-sialic acid axis in cancer supported by the hypersialylation of cancer cells³²⁴, only limited data are available on Siglec expression in CRC. Recently, Yamada *et al.* 2021 showed that, after immunotherapy, high Siglec-7 expression in CRC tissue

macrophages was associated with shorter survival, suggesting that Siglec-7 could be a predictive marker for CRC immunotherapy efficacy in metastatic CRC cases¹¹¹. Our data showing the interaction of Siglec-7 with *F. nucleatum* and that Siglec-7 and *F. nucleatum* are increased in the serum of CRC patients may open the way to the development of non-invasive diagnostic tools with improved specificity e.g. through the combination of biomarkers incorporating Siglec-7 and anti-*F. nucleatum* antibody levels in serum samples.

Immunotherapy has been recognised as a new effective cancer treatment and FDA has approved the use of PD-1 and CTLA-4 immune checkpoint inhibitors for CRC treatment. However, this treatment seems to be mainly effective in a small proportion of CRC cases showing defective DNA mismatch repair pathway³²⁵. Due to the association of Siglecs to pathological conditions such as cancer, infectious diseases and neurological disorders, this lectin superfamily has been proposed as a novel target for therapeutic development^{94,95}. Our results that OMVs derived from *F. nucleatum* ssp. result in activation of the host immune response through the interaction of LPS with Siglec-7, support further research in the development of *F. nucleatum* OMVs as vaccination agents in the treatment of CRC. Indeed, due to their immunogenic properties, there is a lot of interest in developing bacterial OMVs for application as vaccine agents³²⁶ and an OMV-based vaccine has been FDA-approved for the prevention of *Neisseria meningitidis* serogroup B infection³²⁷. Our results suggesting a role of Siglec-7-*F. nucleatum* interaction in immune response provide new molecular leads for the development of clinical tools in CRC for prevention, e.g. by developing specific Siglec-7-targeted *F. nucleatum*-OMV based vaccines but also for treatment, e.g. by targeting Siglec-7 protein with non-immunogenic antagonists such as *F. nucleatum*-derived LPS glycans or OMVs.

In conclusion, our study showed for the first time the role of Siglec-7 in the interaction of colorectal cancer-associated bacteria, *F. nucleatum* ssp., with human immune cells, opening new avenues of research to unravel the *F. nucleatum* tumour promoting mechanisms. The Siglec-7-*F. nucleatum* axis represents a novel target to be exploit for developing new CRC therapeutic treatments.

References

1. Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nat. Rev. Microbiol.* **14**, 20–32 (2015).
2. Li, H. *et al.* The outer mucus layer hosts a distinct intestinal microbial niche. *Nat. Commun.* **6**, 8292 (2015).
3. Johansson, M. E. V, Larsson, J. M. H. & Hansson, G. C. The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl**, 4659–65 (2011).
4. Johansson, M. E. V *et al.* The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 15064–9 (2008).
5. Clevers, H. The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274 (2013).
6. Baulies, A., Angelis, N. & Li, V. S. W. Hallmarks of intestinal stem cells. *Dev.* **147**, 1–12 (2020).
7. Gerbe, F., Legraverend, C. & Jay, P. The intestinal epithelium tuft cells: Specification and function. *Cell. Mol. Life Sci.* **69**, 2907–2917 (2012).
8. Snoeck, V., Goddeeris, B. & Cox, E. The role of enterocytes in the intestinal barrier function and antigen uptake. *Microbes Infect.* **7**, 997–1004 (2005).
9. Miron, N. & Cristea, V. Enterocytes: Active cells in tolerance to food and microbial antigens in the gut. *Clin. Exp. Immunol.* **167**, 405–412 (2012).
10. Ireland, H., Houghton, C., Howard, L. & Winton, D. J. Cellular inheritance of a Cre-activated reporter gene to determine paneth cell longevity in the murine small intestine. *Dev. Dyn.* **233**, 1332–1336 (2005).
11. Bevins, C. L. & Salzman, N. H. Paneth cells, antimicrobial peptides and maintenance of intestinal homeostasis. *Nat. Rev. Microbiol.* **9**, 356–368 (2011).
12. Birchenough, G. M. H., Johansson, M. E. V., Gustafsson, J. K., Bergström, J. H. & Hansson, G. C. New developments in goblet cell mucus secretion and function. *Mucosal Immunol.* **8**, 712–719 (2015).
13. Knoop, K. A. & Newberry, R. D. Goblet cells: multifaceted players in immunity at mucosal surfaces. *Mucosal Immunol.* **11**, 1551–1557 (2018).
14. Mowat, A. M. & Agace, W. W. Regional specialization within the intestinal immune system. *Nat Rev Immunol* **14**, 667–685 (2014).
15. Worthington, J. J., Reimann, F. & Gribble, F. M. Enteroendocrine cells-sensory sentinels of the intestinal environment and orchestrators of mucosal immunity. *Mucosal Immunol.* **11**, 3–20 (2018).
16. Gribble, F. M. & Reimann, F. Function and mechanisms of enteroendocrine cells and gut hormones in metabolism. *Nat. Rev. Endocrinol.* **15**, 226–237 (2019).
17. Corr, S. C., Gahan, C. C. G. M. & Hill, C. M-cells: Origin, morphology and role in mucosal immunity and microbial pathogenesis. *FEMS Immunol. Med. Microbiol.* **52**, 2–12 (2008).
18. Rios, D. *et al.* Antigen sampling by intestinal M cells is the principal pathway initiating mucosal IgA production to commensal enteric bacteria. *Mucosal Immunol.* **9**, 907–916 (2016).
19. Wang, M., Gao, Z., Zhang, Z., Pan, L. & Zhang, Y. Roles of M cells in infection and mucosal vaccines. *Hum. Vaccines Immunother.* **10**, 3544–3551 (2014).
20. Allen, J. E. & Maizels, R. M. Diversity and dialogue in immunity to helminths. *Nat. Rev. Immunol.* **11**, 375–388 (2011).
21. Gerbe, F. *et al.* Intestinal epithelial tuft cells initiate type 2 mucosal immunity to helminth

- parasites. *Nature* **529**, 226–230 (2016).
22. Powell, D. W., Pinchuk, I. V., Saada, J. I., Chen, X. & Mifflin, R. C. Mesenchymal cells of the intestinal lamina propria. *Annu. Rev. Physiol.* **73**, 213–237 (2011).
 23. Yoo, B. B. & Mazmanian, S. K. The Enteric Network: Interactions between the Immune and Nervous Systems of the Gut. *Immunity* **46**, 910–926 (2017).
 24. Agace, W. W. & McCoy, K. D. Regionalized Development and Maintenance of the Intestinal Adaptive Immune Landscape. *Immunity* **46**, 532–548 (2017).
 25. Maldonado-Contreras, A. L. & McCormick, B. A. Intestinal epithelial cells and their role in innate mucosal immunity. *Cell Tissue Res.* **343**, 5–12 (2011).
 26. Sheridan, B. S. & Lefrançois, L. Intraepithelial lymphocytes: To serve and protect. *Curr. Gastroenterol. Rep.* **12**, 513–521 (2010).
 27. Faria, A. M. C., Reis, B. S. & Mucida, D. Tissue adaptation: implications for gut immunity and tolerance. *bioRxiv* (2017).
 28. Mayassi, T. & Jabri, B. Human intraepithelial lymphocytes. *Mucosal Immunol.* **11**, 1281–1289 (2018).
 29. Rescigno, M., Rotta, G., Valzasina, B. & Ricciardi-Castagnoli, P. Dendritic cells shuttle microbes across gut epithelial monolayers. *Immunobiology* **204**, 572–81 (2001).
 30. Schulz, O. & Pabst, O. Antigen sampling in the small intestine. *Trends Immunol.* **34**, 155–161 (2013).
 31. Stagg, A. J. Intestinal Dendritic Cells in Health and Gut Inflammation. *Front. Immunol.* **9**, 2883 (2018).
 32. Watchmaker, P. B. *et al.* Comparative transcriptional and functional profiling defines conserved programs of intestinal DC differentiation in humans and mice. *Nat. Immunol.* **15**, 98–108 (2014).
 33. Mann, E. R. *et al.* Compartment-specific immunity in the human gut: properties and functions of dendritic cells in the colon versus the ileum. *Gut* **65**, 256–270 (2016).
 34. Cerovic, V., Bain, C. C., Mowat, A. M. & Milling, S. W. F. Intestinal macrophages and dendritic cells: What's the difference? *Trends Immunol.* **35**, 270–277 (2014).
 35. Hadis, U. *et al.* Intestinal Tolerance Requires Gut Homing and Expansion of FoxP3+ Regulatory T Cells in the Lamina Propria. *Immunity* **34**, 237–246 (2011).
 36. Sathaliyawala, T. *et al.* Distribution and Compartmentalization of Human Circulating and Tissue-Resident Memory T Cell Subsets. *Immunity* **38**, 187–197 (2013).
 37. Harrison, O. J. & Powrie, F. M. Regulatory T Cells and Immune Tolerance in the Intestine. *Cold Spring Harb. Perspect. Biol.* **5**, a018341–a018341 (2013).
 38. van Wijk, F. & Cheroutre, H. Mucosal T cells in gut homeostasis and inflammation. *Expert Rev. Clin. Immunol.* **6**, 559–566 (2010).
 39. Carrasco, A. *et al.* Regional Specialisation of T Cell Subsets and Apoptosis in the Human Gut Mucosa: Differences Between Ileum and Colon in Healthy Intestine and Inflammatory Bowel Diseases. *J. Crohn's Colitis* **10**, 1042–1054 (2016).
 40. Wang, L., Zhu, L. & Qin, S. Gut Microbiota Modulation on Intestinal Mucosal Adaptive Immunity. *J. Immunol. Res.* **2019**, (2019).
 41. James, K. R. *et al.* Distinct microbial and immune niches of the human colon. *Nat. Immunol.* **21**, 343–353 (2020).
 42. Pabst, O., Cerovic, V. & Hornef, M. Secretory IgA in the Coordination of Establishment and Maintenance of the Microbiota. *Trends Immunol.* **37**, 287–296 (2016).
 43. Lin, M., Du, L., Brandtzaeg, P. & Pan-Hammarström, Q. IgA subclass switch recombination in

- human mucosal and systemic immune compartments. *Mucosal Immunol.* **7**, 511–520 (2014).
44. Pabst, O. & Slack, E. IgA and the intestinal microbiota: the importance of being specific. *Mucosal Immunol.* **13**, 12–21 (2020).
 45. Bremnes, R. M. *et al.* The role of tumor-infiltrating immune cells and chronic inflammation at the tumor site on cancer development, progression, and prognosis: Emphasis on non-small cell lung cancer. *J. Thorac. Oncol.* **6**, 824–833 (2011).
 46. Prendergast, G. C. Immune escape as a fundamental trait of cancer: Focus on IDO. *Oncogene* **27**, 3889–3900 (2008).
 47. Sica, A., Schioppa, T., Mantovani, A. & Allavena, P. Tumour-associated macrophages are a distinct M2 polarised population promoting tumour progression: Potential targets of anti-cancer therapy. *Eur. J. Cancer* **42**, 717–727 (2006).
 48. Mantovani, A. *et al.* The chemokine system in diverse forms of macrophage activation and polarization. *Trends Immunol.* **25**, 677–686 (2004).
 49. Chen, J. J. W. *et al.* Up-regulation of tumor interleukin-8 expression by infiltrating macrophages: Its correlation with tumor angiogenesis and patient survival in non-small cell lung cancer. *Clin. Cancer Res.* **9**, 729–737 (2003).
 50. Kumar, V., Patel, S., Tcyganov, E. & Gabrilovich, D. I. The Nature of Myeloid-Derived Suppressor Cells in the Tumor Microenvironment. *Trends Immunol.* **37**, 208–220 (2016).
 51. Gabrilovich, D. I. & Nagaraj, S. Myeloid-derived suppressor cells as regulators of the immune system. *Nat. Rev. Immunol.* **9**, 162–74 (2009).
 52. Hoehst, B. *et al.* Myeloid derived suppressor cells inhibit natural killer cells in patients with hepatocellular carcinoma via the NKp30 receptor. *Hepatology* **50**, 799–807 (2009).
 53. Srivastava, M. K., Sinha, P., Clements, V. K., Rodriguez, P. & Ostrand-Rosenberg, S. Myeloid-Derived Suppressor Cells Inhibit T-Cell Activation by Depleting Cystine and Cysteine. *Cancer Res.* **70**, 68–77 (2010).
 54. Umansky, V., Blattner, C., Gebhardt, C. & Utikal, J. The Role of Myeloid-Derived Suppressor Cells (MDSC) in Cancer Progression. *Vaccines* **4**, 36 (2016).
 55. Tran Janco, J. M., Lamichhane, P., Karyampudi, L. & Knutson, K. L. Tumor-Infiltrating Dendritic Cells in Cancer Pathogenesis. *J. Immunol.* **194**, 2985–2991 (2015).
 56. Chaudhary, B. & Elkord, E. Regulatory T Cells in the Tumor Microenvironment and Cancer Progression: Role and Therapeutic Targeting. *Vaccines* **4**, 28 (2016).
 57. Dalod, M., Chelbi, R., Malissen, B. & Lawrence, T. Dendritic cell maturation: Functional specialization through signaling specificity and transcriptional programming. *EMBO J.* **33**, 1104–1116 (2014).
 58. Pardo-Camacho, C., González-Castro, A. M., Rodiño-Janeiro, B. K., Pigrau, M. & Vicario, M. Epithelial immunity: Priming defensive responses in the intestinal mucosa. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **314**, G247–G255 (2018).
 59. Gicheva, N., Macauley, M. S., Arlian, B. M., Paulson, J. C. & Kawasaki, N. Siglec-F is a novel intestinal M cell marker. *Biochem. Biophys. Res. Commun.* **479**, 1–4 (2016).
 60. Brown, G. D. & Crocker, P. R. Lectin Receptors Expressed on Myeloid Cells. *Microbiol. Spectr.* **4**, 1–26 (2016).
 61. Roy, R., Murphy, P. & Gabius, H.-J. Multivalent Carbohydrate-Lectin Interactions: How Synthetic Chemistry Enables Insights into Nanometric Recognition. *Molecules* **21**, 629 (2016).
 62. Beulaja Manikandan, S., Manikandan, R., Arumugam, M. & Mullainadhan, P. An overview on human serum lectins. *Heliyon* **6**, e04623 (2020).
 63. Zelensky, A. N. & Gready, J. E. The C-type lectin-like domain superfamily. *FEBS J.* **272**, 6179–6217

(2005).

64. Kurt Drickamer. Database of human proteins containing CTLDs. Available at: <http://www.imperial.ac.uk/research/animalllectins/ctld/mammals/humandata updated.html>.
65. McGreal, E. P., Miller, J. L. & Gordon, S. Ligand recognition by antigen-presenting cell C-type lectin receptors. *Curr. Opin. Immunol.* **17**, 18–24 (2005).
66. Dambuzza, I. M. & Brown, G. D. C-type lectins in immunity: Recent developments. *Curr. Opin. Immunol.* **32**, 21–27 (2015).
67. Mayer, S., Raulf, M.-K. & Lepenies, B. C-type lectins: their network and roles in immunity/pathogen recognition. *Histochem Cell Biol* **in press**, 223–237 (2017).
68. Herre, J., Gordon, S. & Brown, G. D. Dectin-1 and its role in the recognition of β -glucans by macrophages. *Mol. Immunol.* **40**, 869–876 (2004).
69. Ma, J. & Underhill, D. M. β -Glucan Signaling Connects Phagocytosis To Autophagy. *Glycobiology* **23**, 1047–1051 (2013).
70. Ariizumi, K. *et al.* Cloning of a Second Dendritic Cell-associated C-type Lectin (Dectin-2) and Its Alternatively Spliced Isoforms. *J. Biol. Chem.* **275**, 11957–11963 (2000).
71. Taylor, P. R. *et al.* Dectin-2 is predominantly myeloid restricted and exhibits unique activation-dependent expression on maturing inflammatory monocytes elicited in vivo. *Eur. J. Immunol.* **35**, 2163–2174 (2005).
72. Akahori, Y. *et al.* Dectin-2-dependent host defense in mice infected with serotype 3 *Streptococcus pneumoniae*. *BMC Immunol.* **17**, 1 (2016).
73. Wittmann, A. *et al.* Dectin-2 recognizes mannosylated O-antigens of human opportunistic pathogens and augments lipopolysaccharide activation of myeloid cells. *J. Biol. Chem.* **291**, 17629–17638 (2016).
74. Ifrim, D. C. *et al.* The Role of Dectin-2 for Host Defense Against Disseminated Candidiasis. *J. Interferon Cytokine Res.* **36**, 267–76 (2016).
75. Sun, H. *et al.* Activation of NF- κ B and respiratory burst following *Aspergillus fumigatus* stimulation of macrophages. *Immunobiology* **219**, 25–36 (2014).
76. Sato, K. *et al.* Dectin-2 is a pattern recognition receptor for fungi that couples with the Fc receptor γ chain to induce innate immune responses. *J. Biol. Chem.* **281**, 38854–38866 (2006).
77. Kanazawa, N. Dendritic cell immunoreceptors: C-type lectin receptors for pattern-recognition and signaling on antigen-presenting cells. *J. Dermatol. Sci.* **45**, 77–86 (2007).
78. Kimura, Y. *et al.* The innate immune receptor Dectin-2 mediates the phagocytosis of cancer cells by Kupffer cells for the suppression of liver metastasis. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 14097–14102 (2016).
79. Stowell, S. R., Cummings, R. D. & Walker, J. M. *Galectins Series Editor.*
80. Henderson, N. C. *et al.* Galectin-3 regulates myofibroblast activation and hepatic fibrosis. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 5060–5065 (2006).
81. Chou, F. C., Chen, H. Y., Kuo, C. C. & Sytwu, H. K. Role of galectins in tumors and in clinical immunotherapy. *Int. J. Mol. Sci.* **19**, (2018).
82. Dong, R. *et al.* Galectin-3 as a novel biomarker for disease diagnosis and a target for therapy (Review). *Int. J. Mol. Med.* **41**, 599–614 (2018).
83. Hughes, R. C. Secretion of the galectin family of mammalian carbohydrate-binding proteins. *Biochim. Biophys. Acta - Gen. Subj.* **1473**, 172–185 (1999).
84. Bänfer, S. *et al.* Molecular mechanism to recruit galectin-3 into multivesicular bodies for polarized exosomal secretion. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E4396–E4405 (2018).

85. Kohatsu, L., Hsu, D. K., Jegalian, A. G., Liu, F.-T. & Baum, L. G. Galectin-3 Induces Death of *Candida* Species Expressing Specific β -1,2-Linked Mannans. *J. Immunol.* **177**, 4718 LP – 4726 (2006).
86. Dumic, J., Dabelic, S. & Flögel, M. Galectin-3: An open-ended story. *Biochim. Biophys. Acta - Gen. Subj.* **1760**, 616–635 (2006).
87. Shan, M. *et al.* Mucus Enhances Gut Homeostasis and Oral Tolerance by Delivering Immunoregulatory Signals. *Science (80-.)*. **447**, (2013).
88. Kavanaugh, D., Kane, M., Joshi, L. & Hickey, R. M. Detection of galectin-3 interaction with commensal bacteria. *Appl. Environ. Microbiol.* **79**, 3507–3510 (2013).
89. Esteban, A. *et al.* Fungal recognition is mediated by the association of dectin-1 and galectin-3 in macrophages. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 14270–14275 (2011).
90. Bornhöfft, K. F., Goldammer, T., Rebl, A. & Galuska, S. P. Siglecs: A journey through the evolution of sialic acid-binding immunoglobulin-type lectins. *Dev. Comp. Immunol.* **86**, 219–231 (2018).
91. Macauley, M. S., Crocker, P. R. & Paulson, J. C. Siglec-mediated regulation of immune cell function in disease. *Nat. Rev. Immunol.* **14**, 653–66 (2014).
92. Läubli, H. & Varki, A. Sialic acid-binding immunoglobulin-like lectins (Siglecs) detect self-associated molecular patterns to regulate immune responses. *Cell. Mol. Life Sci.* **77**, 593–605 (2020).
93. Duan, S. & Paulson, J. C. Siglecs as Immune Cell Checkpoints in Disease. *Annu. Rev. Immunol.* **38**, 365–395 (2020).
94. Smith, B. A. H. & Bertozzi, C. R. The clinical impact of glycobiology: targeting selectins, Siglecs and mammalian glycans. *Nat. Rev. Drug Discov.* (2021). doi:10.1038/s41573-020-00093-1
95. Manni, M. & Läubli, H. Targeting glyco-immune checkpoints for cancer therapy. *Expert Opin. Biol. Ther.* 14712598.2021.1882989 (2021). doi:10.1080/14712598.2021.1882989
96. Varki, T. A. and A. Siglec Interactions with Pathogens. *Glycosci. Biol. Med.* (2015). doi:10.1007/978-4-431-54841-6_211
97. Carlin, A. F. *et al.* Group B Streptococcus suppression of phagocyte functions by protein-mediated engagement of human Siglec-5. *J. Exp. Med.* **206**, 1691–1699 (2009).
98. Fong, J. J. *et al.* Siglec-7 engagement by GBS β -protein suppresses pyroptotic cell death of natural killer cells. *Proc. Natl. Acad. Sci.* **115**, 10410–10415 (2018).
99. Stephenson, H. N. *et al.* Pseudaminic acid on *Campylobacter jejuni* flagella modulates dendritic cell IL-10 expression via Siglec-10 receptor: A novel flagellin-host interaction. *J. Infect. Dis.* **210**, 1487–1498 (2014).
100. Miyazaki, K. *et al.* Colonic Epithelial Cells Express Specific Ligands for Mucosal Macrophage Immunosuppressive Receptors Siglec-7 and -9. *J. Immunol.* **188**, 4690–4700 (2012).
101. Avril, T., Floyd, H., Lopez, F., Vivier, E. & Crocker, P. R. The membrane-proximal immunoreceptor tyrosine-based inhibitory motif is critical for the inhibitory signaling mediated by Siglecs-7 and -9, CD33-related Siglecs expressed on human monocytes and NK cells. *J. Immunol.* **173**, 6841–9 (2004).
102. Yamaji, T., Teranishi, T., Alphey, M. S., Crocker, P. R. & Hashimoto, Y. A small region of the natural killer cell receptor, Siglec-7, is responsible for its preferred binding to α 2,8-disialyl and branched α 2,6-sialyl residues. A comparison with Siglec-9. *J. Biol. Chem.* **277**, 6324–6332 (2002).
103. Ravetch, J. V. Immune Inhibitory Receptors. *Science (80-.)*. **290**, 84–89 (2002).
104. Crocker, P. R., Paulson, J. C. & Varki, A. Siglecs and their roles in the immune system. *Nat. Rev. Immunol.* **7**, 255–266 (2007).
105. Avril, T., Wagner, E. R., Willison, H. J. & Crocker, P. R. Sialic acid-binding immunoglobulin-like

- lectin 7 mediates selective recognition of sialylated glycans expressed on *Campylobacter jejuni* lipooligosaccharides. *Infect. Immun.* **74**, 4133–4141 (2006).
106. Varchetta, S. *et al.* Engagement of Siglec-7 Receptor Induces a Pro-Inflammatory Response Selectively in Monocytes. *PLoS One* **7**, (2012).
 107. Khatua, B. *et al.* Sialic acids acquired by *Pseudomonas aeruginosa* are involved in reduced complement deposition and siglec mediated host-cell recognition. *FEBS Lett.* **584**, 555–561 (2010).
 108. Wisnovsky, S., Möckl, L., Malaker, S. A., Pedram, K. & Hess, G. T. Genome-wide CRISPR screens reveal a specific ligand for the glycan-binding immune checkpoint receptor Siglec-7. (2021). doi:10.1073/pnas.2015024118/-/DCSupplemental.Published
 109. Xiao, H., Woods, E. C., Vukojicic, P. & Bertozzi, C. R. Precision glycoalyx editing as a strategy for cancer immunotherapy. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 10304–10309 (2016).
 110. Fraschilla, I. & Pillai, S. Viewing Siglecs through the lens of tumor immunology. *Immunol. Rev.* **276**, 178–191 (2017).
 111. Yamada, K. *et al.* Siglec-7 is a predictive biomarker for the efficacy of cancer vaccination against metastatic colorectal cancer. *Oncol. Lett.* **21**, 1–11 (2021).
 112. John, C. M. *et al.* Galectin-3 binds lactosaminylated lipooligosaccharides from *Neisseria gonorrhoeae* and is selectively expressed by mucosal epithelial cells that are infected. *Cell. Microbiol.* **4**, 649–661 (2002).
 113. Debierre-Grockiego, F. *et al.* Binding of *Toxoplasma gondii* glycosylphosphatidylinositols to galectin-3 is required for their recognition by macrophages. *J. Biol. Chem.* **285**, 32744–32750 (2010).
 114. Oliveira, F. L. *et al.* Lack of galectin-3 disturbs mesenteric lymph node homeostasis and B cell niches in the course of *schistosoma mansoni* infection. *PLoS One* **6**, 1–11 (2011).
 115. Kleshchenko, Y. Y. *et al.* Human Galectin-3 Promotes *Trypanosoma cruzi* Adhesion to Human Coronary Artery Smooth Muscle Cells Human Galectin-3 Promotes *Trypanosoma cruzi* Adhesion to Human Coronary Artery Smooth Muscle Cells. *Infect. Immun.* **72**, 6717–6721 (2004).
 116. Mey, A., Leffler, H., Hmama, Z., Normier, G. & Revillard, J. P. The animal lectin galectin-3 interacts with bacterial lipopolysaccharides via two independent sites. *J. Immunol.* **156**, 1572–7 (1996).
 117. Franzosa, E. A. *et al.* Identifying personal microbiomes using metagenomic codes. *Proc. Natl. Acad. Sci.* **112**, E2930–E2938 (2015).
 118. Jandhyala, S. M. *et al.* Role of the normal gut microbiota. *World J. Gastroenterol.* **21**, 8836–8847 (2015).
 119. Rutayisire, E. *et al.* The mode of delivery affects the diversity and colonization pattern of the gut microbiota during the first year of infants' life: a systematic review. *BMC Gastroenterol.* **16**, 86 (2016).
 120. Mancabelli, L. *et al.* Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations. *Environ. Microbiol.* **19**, 1379–1390 (2017).
 121. Zmora, N., Suez, J. & Elinav, E. You are what you eat: diet, health and the gut microbiota. *Nat. Rev. Gastroenterol. Hepatol.* (2018). doi:10.1038/s41575-018-0061-2
 122. Zimmermann, P. & Curtis, N. The effect of antibiotics on the composition of the intestinal microbiota - a systematic review. *J. Infect.* **79**, 471–489 (2019).
 123. Lim, M. Y. *et al.* The effect of heritability and host genetics on the gut microbiota and metabolic syndrome. *Gut* **66**, 1031–1038 (2017).
 124. Sommer, F. & Bäckhed, F. The gut microbiota--masters of host development and physiology. *Nat. Rev. Microbiol.* **11**, 227–38 (2013).

125. Parada Venegas, D. *et al.* Short Chain Fatty Acids (SCFAs)-Mediated Gut Epithelial and Immune Regulation and Its Relevance for Inflammatory Bowel Diseases. *Front. Immunol.* **10**, (2019).
126. Ozdal, T. *et al.* The reciprocal interactions between polyphenols and gut microbiota and effects on bioaccessibility. *Nutrients* **8**, 1–36 (2016).
127. Frick, P. G., Riedler, G. & Brögli, H. Dose response and minimal daily requirement for vitamin K in man. *J. Appl. Physiol.* **23**, 387–389 (1967).
128. Kommineni, S. *et al.* Bacteriocin production augments niche competition by enterococci in the mammalian gastrointestinal tract. *Nature* **526**, 719–722 (2015).
129. Rolhion, N. & Chassaing, B. When pathogenic bacteria meet the intestinal microbiota. *Philos. Trans. R. Soc. B Biol. Sci.* **371**, 20150504 (2016).
130. Alakomi, H., Skyttä, E., Saarela, M. & Helander, I. M. Lactic Acid Permeabilizes Gram-Negative Bacteria by Disrupting the Outer Membrane Lactic Acid Permeabilizes Gram-Negative Bacteria by Disrupting the Outer Membrane. *Appl. Environ. Microbiol.* **66**, 2000–2005 (2005).
131. Kayama, H., Okumura, R. & Takeda, K. Interaction Between the Microbiota , Epithelia , and Immune Cells in the Intestine. (2020).
132. Eberl, G. & Boneca, I. G. Bacteria and MAMP-induced morphogenesis of the immune system. *Curr. Opin. Immunol.* **22**, 448–454 (2010).
133. Zheng, D., Liwinski, T. & Elinav, E. Interaction between microbiota and immunity in health and disease. *Cell Res.* **30**, 492–506 (2020).
134. Caruso, R., Lo, B. C. & Núñez, G. Host–microbiota interactions in inflammatory bowel disease. *Nat. Rev. Immunol.* (2020). doi:10.1038/s41577-019-0268-7
135. Muscogiuri, G. *et al.* Gut microbiota: a new path to treat obesity. *Int. J. Obes. Suppl.* **9**, 10–19 (2019).
136. Patterson, E. *et al.* Gut microbiota, obesity and diabetes. *Postgrad. Med. J.* **92**, 286–300 (2016).
137. Trøseid, M., Andersen, G. Ø., Broch, K. & Hov, J. R. The gut microbiome in coronary artery disease and heart failure: Current knowledge and future directions. *EBioMedicine* **52**, 102649 (2020).
138. Fan, X., Jin, Y., Chen, G., Ma, X. & Zhang, L. Gut microbiota dysbiosis drives the development of colorectal cancer. *Digestion* **318000**, 1–8 (2020).
139. Munson, D. A. & Evans, J. R. Health Care–Acquired Infections in the Nursery. in *Avery’s Diseases of the Newborn* 551–564 (Elsevier, 2012). doi:10.1016/B978-1-4377-0134-0.10040-X
140. Exner, M. *et al.* Antibiotic resistance : What is so special about multidrug-resistant Gram-negative bacteria ? Antibiotikaresistenz : Was ist so besonders an den Gram-negativen. *GMS Hyg. Infect. Control* **12**, 1–24 (2017).
141. Wheeler, R., Chevalier, G., Eberl, G. & Gomperts Boneca, I. The biology of bacterial peptidoglycans and their impact on host immunity and physiology. *Cell. Microbiol.* **16**, 1014–1023 (2014).
142. Foligne, B. *et al.* A key role of dendritic cells in probiotic functionality. *PLoS One* **2**, (2007).
143. Pluddemann, A., Mukhopadhyay, S. & Gordon, S. The interaction of macrophage receptors with bacterial ligands. *Expert Rev Mol Med* **8**, 1–25 (2006).
144. Sukhithasri, V., Nisha, N., Biswas, L., Anil Kumar, V. & Biswas, R. Innate immune recognition of microbial cell wall components and microbial strategies to evade such recognitions. *Microbiol. Res.* **168**, 396–406 (2013).
145. Lebeer, S., Vanderleyden, J. & De Keersmaecker, S. C. J. Host interactions of probiotic bacterial surface molecules: comparison with commensals and pathogens. *Nat. Rev. Microbiol.* **8**, 171–184 (2010).

146. Giorgetti, G. *et al.* Interactions between Innate Immunity, Microbiota, and Probiotics. *J. Immunol. Res.* **2015**, 1–7 (2015).
147. Iborra, S. & Sancho, D. Signalling versatility following self and non-self sensing by myeloid C-type lectin receptors. *Immunobiology* **220**, 175–184 (2015).
148. Tytgat, H. L. P. & Lebeer, S. The sweet tooth of bacteria: common themes in bacterial glycoconjugates. *Microbiol. Mol. Biol. Rev.* **78**, 372–417 (2014).
149. Rini, J., Esko, J. & Varki, A. Glycosyltransferases and Glycan-processing Enzymes. in *Essentials of Glycobiology. 2nd edition.* (2009). doi:NBK1921 [bookaccession]
150. Smith, A. C. & Hussey, M. A. Gram stain protocols. *Am. Soc. Microbiol.* **1**, 14 (2005).
151. Raetz, C. R. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71**, 635–700 (2002).
152. Delcour, A. H. Outer membrane permeability and antibiotic resistance. *Biochim. Biophys. Acta - Proteins Proteomics* **1794**, 808–816 (2009).
153. Rathinam, V. A. K., Zhao, Y. & Shao, F. Innate immunity to intracellular LPS. *Nat. Immunol.* **20**, 527–533 (2019).
154. Park, B. S. & Lee, J.-O. Recognition of lipopolysaccharide pattern by TLR4 complexes. *Exp. Mol. Med.* **45**, e66–e66 (2013).
155. Gill, S., Catchpole, R. & Forterre, P. Extracellular membrane vesicles in the three domains of life and beyond. *FEMS Microbiol. Rev.* **43**, 273–303 (2019).
156. Cecil, J. D., Sirisaengtaksin, N., O'BRIEN-SIMPSON, N. M. & Krachler, A. M. Outer Membrane Vesicle-Host Cell Interactions. *Protein Secret. Bact.* **7**, 201–214 (2019).
157. Qing, G. *et al.* Natural and engineered bacterial outer membrane vesicles. *Biophys. Reports* **5**, 184–198 (2019).
158. Schwechheimer, C. & Kuehn, M. J. Outer-membrane vesicles from Gram-negative bacteria: Biogenesis and functions. *Nat. Rev. Microbiol.* **13**, 605–619 (2015).
159. Kim, S. W. *et al.* Outer membrane vesicles from β -lactam-resistant *Escherichia coli* enable the survival of β -lactam-susceptible *E. coli* in the presence of β -lactam antibiotics. *Sci. Rep.* **8**, 1–13 (2018).
160. Irene, C. *et al.* Bacterial outer membrane vesicles engineered with lipidated antigens as a platform for *Staphylococcus aureus* vaccine. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 21780–21788 (2019).
161. Gerritzen, M. J. H., Martens, D. E., Wijffels, R. H., van der Pol, L. & Stork, M. Bioengineering bacterial outer membrane vesicles as vaccine platform. *Biotechnol. Adv.* **35**, 565–574 (2017).
162. Loomans-Kropp, H. A. & Umar, A. Increasing Incidence of Colorectal Cancer in Young Adults. *J. Cancer Epidemiol.* **2019**, 1–9 (2019).
163. Keum, N. N. & Giovannucci, E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 713–732 (2019).
164. Weinberg, B. A. & Marshall, J. L. Colon Cancer in Young Adults: Trends and Their Implications. *Curr. Oncol. Rep.* **21**, 1–7 (2019).
165. Exarchakou, A., Donaldson, L. J., Girardi, F. & Coleman, M. P. Colorectal cancer incidence among young adults in England: Trends by anatomical subsite and deprivation. *PLoS One* **14**, 1–13 (2019).
166. Gopalakrishnan, V. *et al.* Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science (80-.)*. **359**, 97–103 (2018).
167. Mo, Z. *et al.* Meta-analysis of 16S rRNA Microbial Data Identified Distinctive and Predictive

- Microbiota Dysbiosis in Colorectal Carcinoma Adjacent Tissue. *mSystems* **5**, 1–16 (2020).
168. Wirbel, J. *et al.* Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat. Med.* **25**, (2019).
 169. Arthur, J. C. Microbiota and colorectal cancer: colibactin makes its mark. *Nat. Rev. Gastroenterol. Hepatol.* 4–5 (2020). doi:10.1038/s41575-020-0303-y
 170. Saito, K. *et al.* Metagenomic analyses of the gut microbiota associated with colorectal adenoma. *PLoS One* **14**, 1–18 (2019).
 171. Dai, Z. *et al.* Multi-cohort analysis of colorectal cancer metagenome identified altered bacteria across populations and universal bacterial markers. *Microbiome* **6**, 70 (2018).
 172. Gao, R. *et al.* Mucosa-associated microbiota signature in colorectal cancer. *Eur. J. Clin. Microbiol. Infect. Dis.* **36**, 2073–2083 (2017).
 173. de Waal, G. M., de Villiers, W. J. S., Forgan, T., Roberts, T. & Pretorius, E. Colorectal cancer is associated with increased circulating lipopolysaccharide, inflammation and hypercoagulability. *Sci. Rep.* **10**, 1–20 (2020).
 174. Van Raay, T. & Allen-Vercoe, E. Microbial Interactions and Interventions in Colorectal Cancer. *Microbiol. Spectr.* **5**, (2017).
 175. Sears, C. L. & Pardoll, D. M. Perspective: Alpha-Bugs, Their Microbial Partners, and the Link to Colon Cancer. *J. Infect. Dis.* **203**, 306–311 (2011).
 176. Tjalsma, H., Boleij, A., Marchesi, J. R. & Dutilh, B. E. A bacterial driver–passenger model for colorectal cancer: beyond the usual suspects. *Nat. Rev. Microbiol.* **10**, 575–582 (2012).
 177. Garza, D. R. *et al.* Metabolic models predict bacterial passengers in colorectal cancer. *Cancer Metab.* **8**, 1–13 (2020).
 178. Tomkovich, S. *et al.* Human colon mucosal biofilms from healthy or colon cancer hosts are carcinogenic. *J. Clin. Invest.* **130**, 1699–1712 (2019).
 179. Dejea, C. M. *et al.* Patients with familial adenomatous polyposis harbor colonic biofilms containing tumorigenic bacteria. *Science (80-.)*. **359**, 592–597 (2018).
 180. Li, S., Konstantinov, S. R., Smits, R. & Peppelenbosch, M. P. Bacterial Biofilms in Colorectal Cancer Initiation and Progression. *Trends Mol. Med.* **23**, 18–30 (2017).
 181. Yang, J. & Yu, J. The association of diet, gut microbiota and colorectal cancer: what we eat may imply what we get. *Protein Cell* **9**, 474–487 (2018).
 182. Mehta, R. S. *et al.* Association of Dietary Patterns With Risk of Colorectal Cancer Subtypes Classified by *Fusobacterium nucleatum* in Tumor Tissue. *JAMA Oncol.* **3**, 921 (2017).
 183. Carolina Vieira de Almeida, A. T. and A. A. The controversial role of *Enterococcus faecalis* in colorectal cancer. *Therap. Adv. Gastroenterol.* **11**, 1–11 (2018).
 184. Richard, M. L. *et al.* Mucosa-associated microbiota dysbiosis in colitis associated cancer. *Gut Microbes* **0976**, 1–12 (2017).
 185. Boleij, A. *et al.* The bacteroides fragilis toxin gene is prevalent in the colon mucosa of colorectal cancer patients. *Clin. Infect. Dis.* **60**, 208–215 (2015).
 186. Drewes, J. L. *et al.* High-resolution bacterial 16S rRNA gene profile meta-analysis and biofilm status reveal common colorectal cancer consortia. *npj Biofilms Microbiomes* **3**, (2017).
 187. Wu, S., Rhee, K.-J., Zhang, M., Franco, A. & Sears, C. L. Bacteroides fragilis toxin stimulates intestinal epithelial cell shedding and -secretase-dependent E-cadherin cleavage. *J. Cell Sci.* **120**, 3713–3713 (2007).
 188. Wu, S., Morin, P. J., Maouyo, D. & Sears, C. L. Bacteroides fragilis enterotoxin induces c-Myc expression and cellular proliferation. *Gastroenterology* **124**, 392–400 (2003).

189. Thiele Orberg, E. *et al.* The myeloid immune signature of enterotoxigenic *Bacteroides fragilis*-induced murine colon tumorigenesis. *Mucosal Immunol.* **10**, 421–433 (2017).
190. Tzianabos, A. O. *et al.* The capsular polysaccharide of *Bacteroides fragilis* comprises two ionically linked polysaccharides. *J. Biol. Chem.* **267**, 18230–18235 (1992).
191. Mazmanian, S. K., Cui, H. L., Tzianabos, A. O. & Kasper, D. L. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122**, 107–118 (2005).
192. Chang, Y.-C. *et al.* TLR2 and interleukin-10 are involved in *Bacteroides fragilis*-mediated prevention of DSS-induced colitis in gnotobiotic mice. *PLoS One* **12**, e0180025 (2017).
193. Caselli, E. *et al.* Defining the oral microbiome by whole-genome sequencing and resistome analysis: the complexity of the healthy picture. *BMC Microbiol.* **20**, 120 (2020).
194. Poore, G. D. *et al.* Microbiome analyses of blood and tissues suggest cancer diagnostic approach. *Nature* **579**, 567–574 (2020).
195. Kostic, A. D. *et al.* *Fusobacterium nucleatum* Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment. *Cell Host Microbe* **14**, 207–215 (2013).
196. Ye, X. *et al.* *Fusobacterium nucleatum* subspecies *animalis* influences pro-inflammatory cytokine expression and monocyte activation in human colorectal tumors. *Cancer Prev. Res.* **10**, canprevres.0178.2016 (2017).
197. Bundgaard-Nielsen, C., Baandrup, U. T., Nielsen, L. P. & Sørensen, S. The presence of bacteria varies between colorectal adenocarcinomas, precursor lesions and non-malignant tissue. *BMC Cancer* **19**, 399 (2019).
198. Yu, J. *et al.* Invasive *Fusobacterium nucleatum* may play a role in the carcinogenesis of proximal colon cancer through the serrated neoplasia pathway. *Int. J. Cancer* **139**, 1318–1326 (2016).
199. Mima, K. *et al.* *Fusobacterium nucleatum* in Colorectal Carcinoma Tissue According to Tumor Location. *Clin. Transl. Gastroenterol.* **7**, (2016).
200. Torlakovic, E., Skovlund, E., Snover, D. C., Torlakovic, G. & Nesland, J. M. Morphologic reappraisal of serrated colorectal polyps. *Am. J. Surg. Pathol.* **27**, 65–81 (2003).
201. Kunzmann, A. T. *et al.* *Fusobacterium nucleatum* tumor DNA levels are associated with survival in colorectal cancer patients. *Eur. J. Clin. Microbiol. Infect. Dis.* **38**, 1891–1899 (2019).
202. Bullman, S. *et al.* Analysis of *Fusobacterium* persistence and antibiotic response in colorectal cancer. *Science (80-)*. **358**, 1443–1448 (2017).
203. Chen, Y. *et al.* *Fusobacterium nucleatum* promotes metastasis in colorectal cancer by activating autophagy signaling via the upregulation of CARD3 expression. *Theranostics* **10**, 323–339 (2020).
204. Kasper, S. H. *et al.* Colorectal cancer-associated anaerobic bacteria proliferate in tumor spheroids and alter the microenvironment. *Sci. Rep.* **10**, 1–13 (2020).
205. Hayden, M. S., Ghosh, S., Hayden, M. S. & Ghosh, S. outstanding questions NF- κ B, the first quarter-century : remarkable progress and outstanding questions. 203–234 (2016). doi:10.1101/gad.183434.111
206. Li, Y. *et al.* The Prognostic and Clinicopathological Roles of PD-L1 Expression in Colorectal Cancer: A Systematic Review and Meta-Analysis. *Front. Pharmacol.* **10**, 1–10 (2019).
207. Chen, T. *et al.* *Fusobacterium nucleatum* promotes M2 polarization of macrophages in the microenvironment of colorectal tumours via a TLR4-dependent mechanism. *Cancer Immunol. Immunother.* **67**, 1635–1646 (2018).
208. Abed, J. *et al.* Fap2 Mediates *Fusobacterium nucleatum* Colorectal Adenocarcinoma Enrichment by Binding to Tumor-Expressed Gal-GalNAc. *Cell Host Microbe* **20**, 215–225 (2016).
209. Yu, X. *et al.* The surface protein TIGIT suppresses T cell activation by promoting the generation of mature immunoregulatory dendritic cells. *Nat. Immunol.* **10**, 48–57 (2009).

210. Rubinstein, M. R. *et al.* Fusobacterium nucleatum promotes colorectal cancer by inducing Wnt/ β -catenin modulator Annexin A1 . *EMBO Rep.* e47638 (2019). doi:10.15252/embr.201847638
211. Grenier, D. & Grignon, L. Response of human macrophage-like cells to stimulation by Fusobacterium nucleatum ssp. nucleatum lipopolysaccharide. *Oral Microbiol. Immunol.* **21**, 190–196 (2006).
212. Asai, Y., Makimura, Y., Kawabata, A. & Ogawa, T. Soluble CD14 Discriminates Slight Structural Differences between Lipid As That Lead to Distinct Host Cell Activation. *J. Immunol.* **179**, 7674–7683 (2014).
213. Yang, Y. *et al.* Fusobacterium nucleatum Increases Proliferation of Colorectal Cancer Cells and Tumor Development in Mice by Activating Toll-Like Receptor 4 Signaling to Nuclear Factor- κ B, and Up-regulating Expression of MicroRNA-21. *Gastroenterology* **152**, 851-866.e24 (2017).
214. Yu, T. C. *et al.* Fusobacterium nucleatum Promotes Chemoresistance to Colorectal Cancer by Modulating Autophagy. *Cell* **170**, 548-563.e16 (2017).
215. Liu, J. *et al.* Proteomic characterization of outer membrane vesicles from gut mucosa-derived fusobacterium nucleatum. *J. Proteomics* **195**, 125–137 (2019).
216. Martin-Gallausiaux, C., Malabirade, A., Habier, J. & Wilmes, P. Fusobacterium nucleatum Extracellular Vesicles Modulate Gut Epithelial Cell Innate Immunity via FomA and TLR2. *Front. Immunol.* **11**, (2020).
217. Vinogradov, E., St Michael, F. & Cox, A. D. Structure of the LPS O-chain from Fusobacterium nucleatum strain ATCC 23726 containing a novel 5,7-diamino-3,5,7,9-tetradeoxy-L-gluco-non-2-ulosonic acid presumably having the D-glycero-L-gluco configuration. *Carbohydr. Res.* **468**, 69–72 (2018).
218. Vinogradov, E., St Michael, F. & Cox, A. D. Structure of the LPS O-chain from Fusobacterium nucleatum strain MJR 7757 B. *Carbohydr. Res.* **463**, 37–39 (2018).
219. Vinogradov, E., St. Michael, F., Homma, K., Sharma, A. & Cox, A. D. Structure of the LPS O-chain from Fusobacterium nucleatum strain 10953, containing sialic acid. *Carbohydr. Res.* **440–441**, 38–42 (2017).
220. Vinogradov, E., Michael, F. S. & Cox, A. D. The structure of the LPS O-chain of Fusobacterium nucleatum strain 25586 containing two novel monosaccharides , 2-acetamido-2 , 6- non-2-ulosonic acid. *Carbohydr. Res.* **440–441**, 10–15 (2017).
221. Crost, E. H. *et al.* Utilisation of Mucin Glycans by the Human Gut Symbiont Ruminococcus gnavus Is Strain-Dependent. *PLoS One* **8**, (2013).
222. Lamprinaki, D. *et al.* LC3-associated phagocytosis is required for dendritic cell inflammatory cytokine response to gut commensal yeast Saccharomyces cerevisiae. *Front. Immunol.* **8**, (2017).
223. De Castro, C., Parrilli, M., Holst, O. & Molinaro, A. Microbe-Associated Molecular Patterns in Innate Immunity. in *Methods in Enzymology* 89–115 (2010). doi:10.1016/S0076-6879(10)80005-9
224. Tsai, C.-M. & Frasch, C. E. A sensitive silver stain for detecting lipopolysaccharides in polyacrylamide gels. *Anal. Biochem.* **119**, 115–119 (1982).
225. Hase, S., Hofstad, T. & Rietschel, E. T. Chemical structure of the lipid A component of lipopolysaccharides from Fusobacterium nucleatum. *J. Bacteriol.* **129**, 9–14 (1977).
226. Leclaire, C. *et al.* Molecular basis for intestinal mucin recognition by galectin-3 and C-type lectins. *FASEB J.* **0**, fj.201700619R (2018).
227. Froger, A. & Hall, J. E. Transformation of Plasmid DNA into E. coli Using the Heat Shock Method. *J. Vis. Exp.* 253 (2007). doi:10.3791/253
228. Troegeler, A. *et al.* An efficient siRNA-mediated gene silencing in primary human monocytes,

- dendritic cells and macrophages. *Immunol. Cell Biol.* **92**, 699–708 (2014).
229. Vega-Sanchez, R., Arenas-Hernandez, M., Vazquez-Perez, J. A., Moreno-Valencia, Y. & Gomez-Lopez, N. Evaluation of reference genes for expression studies in leukocytes from term human pregnancy. *Placenta* **36**, 240–245 (2015).
230. Higuchi, H., Shoji, T., Iijima, S. & Nishijima, K. I. Constitutively expressed Siglec-9 inhibits LPS-induced CCR7, but enhances IL-4-induced CD200R expression in human macrophages. *Biosci. Biotechnol. Biochem.* **80**, 1141–1148 (2016).
231. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci.* **108**, 4516–4522 (2011).
232. Hildebrand, F., Tadeo, R., Voigt, A., Bork, P. & Raes, J. LotuS: an efficient and user-friendly OTU processing pipeline. *Microbiome* **2**, 30 (2014).
233. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* **10**, 996–998 (2013).
234. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
235. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* **27**, 2194–2200 (2011).
236. Hauswedell, H., Singer, J. & Reinert, K. Lambda: the local aligner for massive biological data. *Bioinformatics* **30**, i349–i355 (2014).
237. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).
238. Saary, P., Forslund, K., Bork, P. & Hildebrand, F. RTK: efficient rarefaction analysis of large datasets. *Bioinformatics* **33**, 2594–2595 (2017).
239. Farhad, M., Rolig, A. S. & Redmond, W. L. The role of Galectin-3 in modulating tumor growth and immunosuppression within the tumor microenvironment. *Oncoimmunology* **7**, 1–8 (2018).
240. Fan, L., Frye, C. C. & Racher, A. J. The use of glutamine synthetase as a selection marker: recent advances in Chinese hamster ovary cell line generation processes. *Pharm. Bioprocess.* **1**, 487–502 (2013).
241. Aoki, T. A Comprehensive Review of Our Current Understanding of Red Blood Cell (RBC) Glycoproteins. *Membranes (Basel)*. **7**, 56 (2017).
242. Cohen, M., Hurtado-Ziola, N. & Varki, A. ABO blood group glycans modulate sialic acid recognition on erythrocytes. *Blood* **114**, 3668–3676 (2009).
243. O'Reilly, M. K. & Paulson, J. C. Multivalent ligands for siglecs. *Methods Enzymol.* **478**, 343–363 (2010).
244. Zou, Z. *et al.* Siglecs Facilitate HIV-1 Infection of Macrophages through Adhesion with Viral Sialic Acids. *PLoS One* **6**, e24559 (2011).
245. Varchetta, S. *et al.* Sialic acid-binding Ig-like lectin-7 interacts with HIV-1 gp120 and facilitates infection of CD4pos T cells and macrophages. *Retrovirology* **10**, 1–13 (2013).
246. Suenaga, T. *et al.* Myelin-associated glycoprotein mediates membrane fusion and entry of neurotropic herpesviruses. *Proc. Natl. Acad. Sci.* **107**, 866–871 (2010).
247. Garcia-Vello, P. *et al.* Structure of the O-antigen and the lipid A from the lipopolysaccharide of *Fusobacterium nucleatum* ATCC 51191. *ChemBioChem* **cbic.202000751** (2020). doi:10.1002/cbic.202000751
248. Alphey, M. S., Attrill, H., Crocker, P. R. & van Aalten, D. M. F. High Resolution Crystal Structures of Siglec-7. *J. Biol. Chem.* **278**, 3372–3377 (2003).

249. Yamakawa, N. *et al.* Discovery of a new sialic acid binding region that regulates Siglec-7. *Sci. Rep.* **10**, 1–14 (2020).
250. Kenyon, J. J., Marzaioli, A. M., De Castro, C. & Hall, R. M. 5,7-di-N-acetyl-acinetaminic acid: A novel non-2-ulosonic acid found in the capsule of an *Acinetobacter baumannii* isolate. *Glycobiology* **25**, 644–654 (2015).
251. Glaze, P. A., Watson, D. C., Young, N. M. & Tanner, M. E. Biosynthesis of CMP- N , N '-Diacetyllegionaminic Acid from UDP- N , N '-Diacetylbacillosamine in *Legionella pneumophila* †. *Biochemistry* **47**, 3272–3282 (2008).
252. Varki, A. & Schauer, R. Sialic Acids. in (eds. Varki, A. et al.) (2009).
253. Alisson-Silva, F. *et al.* Human evolutionary loss of epithelial Neu5Gc expression and species-specific susceptibility to cholera. *PLOS Pathog.* **14**, e1007133 (2018).
254. Weiman, S. *et al.* Genetic and biochemical modulation of sialic acid O-acetylation on group B *Streptococcus*: Phenotypic and functional impact. *Glycobiology* **19**, 1204–1213 (2009).
255. Lewis, A. L. *et al.* Innovations in host and microbial sialic acid biosynthesis revealed by phylogenomic prediction of nonulosonic acid structure. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13552–13557 (2009).
256. Li, Y. & Chen, X. Sialic acid metabolism and sialyltransferases: Natural functions and applications. *Appl. Microbiol. Biotechnol.* **94**, 887–905 (2012).
257. Song, X. *et al.* A Sialylated Glycan Microarray Reveals Novel Interactions of Modified Sialic Acids with Proteins and Viruses. *J. Biol. Chem.* **286**, 31610–31622 (2011).
258. Lewis, A. L., Robinson, L. S., Agarwal, K. & Lewis, W. G. Discovery and characterization of de novo sialic acid biosynthesis in the phylum *Fusobacterium*. *Glycobiology* **26**, 1107–1119 (2016).
259. Iwamori, M., Ohta, Y., Uchida, Y. & Tsukada, Y. *Arthrobacter ureafaciens* sialidase isoenzymes, L, m1 and m2, cleave fucosyl GM1. *Glycoconj. J.* **14**, 67–73 (1997).
260. Bonnington, K. E. & Kuehn, M. J. Outer membrane vesicle production facilitates LPS remodeling and outer membrane maintenance in *Salmonella* during environmental transitions. *MBio* **7**, 1–15 (2016).
261. Rubinstein, M. R. *et al.* *Fusobacterium nucleatum* Promotes Colorectal Carcinogenesis by Modulating E-Cadherin/ β -Catenin Signaling via its FadA Adhesin. *Cell Host Microbe* **14**, 195–206 (2013).
262. Kurgan, Ş. *et al.* Strain-Specific Impact of *Fusobacterium nucleatum* on Neutrophil Function. *J. Periodontol.* **88**, 380–389 (2017).
263. Xue, Y. *et al.* Indoleamine 2,3-dioxygenase expression regulates the survival and proliferation of *Fusobacterium nucleatum* in THP-1-derived macrophages. *Cell Death Dis.* **9**, (2018).
264. Krisanaprakornkit, S. *et al.* Inducible expression of human β -defensin 2 by *Fusobacterium nucleatum* in oral epithelial cells: Multiple signaling pathways and role of commensal bacteria in innate immunity and the epithelial barrier. *Infect. Immun.* **68**, 2907–2915 (2000).
265. Jewett, a *et al.* Induction of apoptotic cell death in peripheral blood mononuclear and polymorphonuclear cells by an oral bacterium, *Fusobacterium nucleatum*. *Infect. Immun.* **68**, 1893–8 (2000).
266. Huynh, T. *et al.* The role of aggregation in *fusobacterium nucleatum*-induced immune cell death. *J. Endod.* **37**, 1531–1535 (2011).
267. Bachrach, G., Ianculovici, C., Naor, R. & Weiss, E. I. Fluorescence based measurements of *Fusobacterium nucleatum* coaggregation and of fusobacterial attachment to mammalian cells. *FEMS Microbiol. Lett.* **248**, 235–240 (2005).
268. Gant, V. A., Warnes, G., Phillips, I. & Savidge, G. F. The application of flow cytometry to the study of bacterial responses to antibiotics. *J. Med. Microbiol.* **39**, 147–154 (1993).

269. Wickens, H. J., Pinney, R. J., Mason, D. J. & Gant, V. A. Flow Cytometric Investigation of Filamentation, Membrane Patency, and Membrane Potential in *Escherichia coli* following Ciprofloxacin Exposure. *Antimicrob. Agents Chemother.* **44**, 682–687 (2000).
270. Merritt, J., Niu, G., Okinaga, T. & Qi, F. Autoaggregation response of *Fusobacterium nucleatum*. *Appl. Environ. Microbiol.* **75**, 7725–7733 (2009).
271. Kleiveland, C. R. Peripheral Blood Mononuclear Cells. in *The Impact of Food Bioactives on Health: in vitro and ex vivo models* (eds. Verhoeckx, K. et al.) 161–167 (Springer International Publishing, 2015). doi:10.1007/978-3-319-16104-4_15
272. Ohradanova-Repic, A., Machacek, C., Fischer, M. B. & Stockinger, H. Differentiation of human monocytes and derived subsets of macrophages and dendritic cells by the HLDA10 monoclonal antibody panel. *Clin. Transl. Immunol.* **5**, e55-9 (2016).
273. Zhang, J. Q., Biedermann, B., Nitschke, L. & Crocker, P. R. The murine inhibitory receptor mSiglec-E is expressed broadly on cells of the innate immune system whereas mSiglec-F is restricted to eosinophils. *Eur. J. Immunol.* **34**, 1175–1184 (2004).
274. Park, H. E., Kim, J. H., Cho, N. Y., Lee, H. S. & Kang, G. H. Intratumoral *Fusobacterium nucleatum* abundance correlates with macrophage infiltration and CDKN2A methylation in microsatellite-unstable colorectal carcinoma. *Virchows Arch.* **471**, 329–336 (2017).
275. McCoy, A. N. *et al.* *Fusobacterium* Is Associated with Colorectal Adenomas. *PLoS One* **8**, (2013).
276. Hudak, J. E., Canham, S. M. & Bertozzi, C. R. Glycocalyx engineering reveals a Siglec-based mechanism for NK cell immunoevasion. *Nat. Chem. Biol.* **10**, 69–75 (2014).
277. Jandus, C. *et al.* Interactions between Siglec-7/9 receptors and ligands influence NK cell-dependent tumor immunosurveillance. *J. Clin. Invest.* **124**, 1810–1820 (2014).
278. Chang, Y.-C. *et al.* Group B Streptococcus Engages an Inhibitory Siglec through Sialic Acid Mimicry to Blunt Innate Immune and Inflammatory Responses In Vivo. *PLoS Pathog.* **10**, e1003846 (2014).
279. Jones, C., Virji, M. & Crocker, P. R. Recognition of sialylated meningococcal lipopolysaccharide by siglecs expressed on myeloid cells leads to enhanced bacterial uptake. *Mol. Microbiol.* **49**, 1213–1225 (2003).
280. Walter, R. B. *et al.* ITIM-dependent endocytosis of CD33-related Siglecs: role of intracellular domain, tyrosine phosphorylation, and the tyrosine phosphatases, Shp1 and Shp2. *J. Leukoc. Biol.* **83**, 200–211 (2008).
281. Zheng, Y. *et al.* The Roles of Siglec7 and Siglec9 on Natural Killer Cells in Virus Infection and Tumour Progression. *J. Immunol. Res.* **2020**, 1–9 (2020).
282. Kawasaki, T. & Kawai, T. Toll-Like Receptor Signaling Pathways. *Front. Immunol.* **5**, (2014).
283. Sanjuan, M. A. *et al.* Toll-like receptor signalling in macrophages links the autophagy pathway to phagocytosis. *Nature* **450**, 1253–7 (2007).
284. Chen, G.-Y. *et al.* Broad and direct interaction between TLR and Siglec families of pattern recognition receptors and its regulation by Neu1. *Elife* **3**, 1–18 (2014).
285. Greene, C. M., McElvaney, N. G., O’Neill, S. J. & Taggart, C. C. Secretory leucoprotease inhibitor impairs toll-like receptor 2- and 4-mediated responses in monocytic cells. *Infect. Immun.* **72**, 3684–3687 (2004).
286. Novak, N., Koch, S., Allam, J.-P. & Bieber, T. Dendritic cells: Bridging innate and adaptive immunity in atopic dermatitis. *J. Allergy Clin. Immunol.* **125**, 50–59 (2010).
287. Wu, Y. *et al.* *Fusobacterium nucleatum* Potentiates Intestinal Tumorigenesis in Mice via a Toll-Like Receptor 4/p21-Activated Kinase 1 Cascade. *Dig. Dis. Sci.* **63**, 1210–1218 (2018).
288. Sun, Y. *et al.* *Fusobacterium nucleatum* infection is correlated with tumor metastasis and postoperative survival of colorectal cancer patients in China. *Transl. Cancer Res.* **5**, 579–588

- (2016).
289. de Carvalho, A. C. *et al.* Microbiota profile and impact of fusobacterium nucleatum in colorectal cancer patients of barretos cancer hospital. *Front. Oncol.* **9**, 1–11 (2019).
 290. Wang, H. F. *et al.* Evaluation of antibody level against Fusobacterium nucleatum in the serological diagnosis of colorectal cancer. *Sci. Rep.* **6**, 1–10 (2016).
 291. Alkharaan, H. *et al.* Circulating and Salivary Antibodies to Fusobacterium nucleatum Are Associated With Cystic Pancreatic Neoplasm Malignancy. *Front. Immunol.* **11**, 1–10 (2020).
 292. Wang, X. *et al.* Siglec-9 is upregulated in rheumatoid arthritis and suppresses collagen-induced arthritis through reciprocal regulation of Th17-/Treg-cell differentiation. *Scand. J. Immunol.* **85**, 433–440 (2017).
 293. Oliveira, J. J. *et al.* The plasma biomarker soluble SIGLEC-1 is associated with the type I interferon transcriptional signature, ethnic background and renal disease in systemic lupus erythematosus. *Arthritis Res. Ther.* **20**, 1–13 (2018).
 294. Na, H. J., Hamilton, R. G., Klion, A. D. & Bochner, B. S. Biomarkers of eosinophil involvement in allergic and eosinophilic diseases: Review of phenotypic and serum markers including a novel assay to quantify levels of soluble Siglec-8. *J. Immunol. Methods* **383**, 39–46 (2012).
 295. Angata, T. & Varki, A. Siglec-7: a sialic acid-binding lectin of the immunoglobulin superfamily. *Glycobiology* **10**, 431–438 (2000).
 296. The Human Protein Atlas. Available at: <https://www.proteinatlas.org/ENSG00000168995-SIGLEC7/tissue>.
 297. Zitvogel, L., Ayyoub, M., Routy, B. & Kroemer, G. Microbiome and Anticancer Immunosurveillance. *Cell* **165**, 276–287 (2016).
 298. Baruch, E. N. *et al.* Fecal microbiota transplant promotes response in immunotherapy-refractory melanoma patients. *Science (80-.)*. **5920**, eabb5920 (2020).
 299. Ramos, A. & Hemann, M. T. Drugs, Bugs, and Cancer: Fusobacterium nucleatum Promotes Chemoresistance in Colorectal Cancer. *Cell* **170**, 411–413 (2017).
 300. Geller, L. T. *et al.* Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. *Science (80-.)*. **1160**, 1156–1160 (2017).
 301. Kostic, A. D. *et al.* Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Res.* **22**, 292–298 (2012).
 302. Repass, J. *et al.* Replication Study: Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Elife* **7**, (2018).
 303. Schistosomes, liver flukes and Helicobacter pylori. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, 7-14 June 1994. *IARC monographs on the evaluation of carcinogenic risks to humans* **61**, 1–241 (1994).
 304. Wroblewski, L. E., Peek, R. M. & Wilson, K. T. Helicobacter pylori and Gastric Cancer: Factors That Modulate Disease Risk. *Clin. Microbiol. Rev.* **23**, 713–739 (2010).
 305. Xavier, J. B. *et al.* The Cancer Microbiome: Distinguishing Direct and Indirect Effects Requires a Systemic View. *Trends in Cancer* **6**, 192–204 (2020).
 306. Jia, B. & Jeon, C. O. Promotion and induction of liver cancer by gut microbiome-mediated modulation of bile acids. *PLOS Pathog.* **15**, e1007954 (2019).
 307. Lee, S. A., Liu, F., Riordan, S. M., Lee, C. S. & Zhang, L. Global investigations of fusobacterium nucleatum in human colorectal cancer. *Front. Oncol.* **9**, 1–11 (2019).
 308. Brennan, C. A. & Garrett, W. S. Fusobacterium nucleatum — symbiont, opportunist and oncobacterium. *Nat. Rev. Microbiol.* **17**, 156–166 (2019).

309. Taylor ME, Drickamer K, Schnaar RL, et al. Discovery and Classification of Glycan-Binding Proteins. in *Essentials of Glycobiology [Internet]* (ed. Varki A, Cummings RD, Esko JD, et al.) doi:10.1101/glycobiology.3e.028
310. Nakamura, M. *et al.* Involvement of galectin-3 expression in colorectal cancer progression and metastasis. *Int. J. Oncol.* (1999). doi:10.3892/ijo.15.1.143
311. Barrow, H., Rhodes, J. M. & Yu, L. G. The role of galectins in colorectal cancer progression. *Int. J. Cancer* **129**, 1–8 (2011).
312. Stanley P, C. R. Structures Common to Different Glycans. in *Essentials of Glycobiology* (ed. Varki A, Cummings RD, Esko JD, et al.) (2009).
313. Toukach, P. V. Bacterial Carbohydrate Structure Database 3: Principles and Realization. *J. Chem. Inf. Model.* **51**, 159–170 (2011).
314. Low, L. A., Mummery, C., Berridge, B. R., Austin, C. P. & Tagle, D. A. Organs-on-chips: into the next decade. *Nat. Rev. Drug Discov.* (2020). doi:10.1038/s41573-020-0079-3
315. Mendoza-Coronel, E. & Castañón-Arreola, M. Comparative evaluation of in vitro human macrophage models for mycobacterial infection study. *Pathog. Dis.* **74**, ftw052 (2016).
316. Tarique, A. A. *et al.* Phenotypic, Functional, and Plasticity Features of Classical and Alternatively Activated Human Macrophages. *Am. J. Respir. Cell Mol. Biol.* **53**, 676–688 (2015).
317. Delaveris, C. S. *et al.* Synthetic Siglec-9 agonists inhibit neutrophil activation associated with COVID-19. **2019**, (2020).
318. Vatanen, T. *et al.* Variation in Microbiome LPS Immunogenicity Contributes to Autoimmunity in Humans. *Cell* **165**, 842–853 (2016).
319. Alexander, C. & Rietschel, E. T. Bacterial lipopolysaccharides and innate immunity. *J. Endotoxin Res.* **7**, 167–202 (2001).
320. Hathroubi, S. *et al.* Surface Polysaccharide Mutants Reveal that Absence of O Antigen Reduces Biofilm Formation of *Actinobacillus pleuropneumoniae*. *Infect. Immun.* **84**, 127–137 (2016).
321. Yamamura, K. *et al.* *Fusobacterium nucleatum* in gastroenterological cancer: Evaluation of measurement methods using quantitative polymerase chain reaction and a literature review. *Oncol. Lett.* (2017). doi:10.3892/ol.2017.7001
322. Castellarin, M. *et al.* *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. 299–306 (2011). doi:10.1101/gr.126516.111
323. Kurt, M. & Yumuk, Z. Diagnostic accuracy of *Fusobacterium nucleatum* IgA and IgG ELISA test in colorectal cancer. *Sci. Rep.* 1–6 (2021). doi:10.1038/s41598-021-81171-1
324. Holm, M. *et al.* N-glycomic profiling of colorectal cancer according to tumor stage and location. *PLoS One* **15**, e0234989 (2020).
325. Breakstone, R. Colon cancer and immunotherapy—can we go beyond microsatellite instability? *Transl. Gastroenterol. Hepatol.* **6**, 12–12 (2021).
326. Holst, J. *et al.* Vaccines against meningococcal serogroup B disease containing outer membrane vesicles (OMV) Lessons from past programs and implications for the future. *Hum. Vaccines Immunother.* **9**, 1241–1253 (2013).
327. Vernikos, G. & Medini, D. Bexsero® chronicle. *Pathog. Glob. Health* **108**, 305–316 (2014).