THE ROLE OF GUT FLORA IN EPITHELIAL BARRIER FUNCTION AND IMMUNITY

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology Medicine and Health

2016

School of Biological Sciences

MARIA GLYMENAKI

Contents

Chapter 6: Supplementary materials and methods

Word count: 52819

List of figures

Chapter 1: Introduction

Chapter 2: Compositional changes in the gut mucus microbiota precede the onset of colitic inflammation

Chapter 3: Stability in metabolite and metagenome profiles before the onset of colitis-type inflammation

Chapter 4: Characterization of antimicrobial protein expression prior to colitis development in mdr1a knockout mice

Chapter 5: Summary discussion

Chapter 6: Supplementary materials and methods

Appendix

List of tables

Chapter 2: Compositional changes in the gut mucus microbiota precede the onset of colitic inflammation

Chapter 3: Stability in metabolite and metagenome profiles before the onset of colitis-type inflammation

Chapter 4: Characterization of antimicrobial protein expression prior to colitis development in mdr1a knockout mice

Chapter 6: Supplementary materials and methods

Abstract

Inflammatory bowel disease (IBD) is associated with an inappropriate immune response to the gut microbiota and disruption of intestinal homeostasis. IBD patients and experimental animal models have consistently shown alterations in the gut microbiota composition. However, these studies have mainly focused on faecal microbiota samples taken after the onset of inflammation and IBD establishment. The colonic microbiota inhabits both the gut lumen and the mucus layer covering the intestinal epithelium. Thus, information about mucus-resident microbiota is not necessarily conveyed in the routine microbiota analyses of faecal samples. To address potential changes in microbial composition and function before the onset of IBD, we compared both mucus and faecal microbiota in the *mdr1a*$^{-/-}$ spontaneous model of colitis over times that we histologically defined as before onset of colitis, during and after colitis onset.

We showed that alterations in microbiota composition preceded the onset of intestinal inflammation and that these changes were evident in the mucus, but not in faeces. This altered microbiota composition was coupled with a reduced inner mucus layer, indicating a compromised mucus barrier prior to colitis development. Upon emergence of inflammation, compositional differences were found in both mucus and faecal microbial communities. Spatial segregation of microbiota with intestinal mucosa was also disrupted on disease onset which we hypothesise contributes to a more severe intestinal pathology. Therefore, our data indicate that microbial changes start locally in the mucus and then proceed to the faecal matter concomitantly with colitis development.

Next, we examined whether microbial gene functional potential and endogenous metabolite profiles followed alterations in gut microbiota taxonomic composition. Our findings showed that the microbial gene content was similar between *mdr1a*$^{-/-}$ mice and wild-type littermate controls, demonstrating stability of the gut microbiome at the face of ensuing gut inflammation. In further support of these findings, urinary metabolite analysis revealed that metabolite profiles were unaffected by intestinal inflammation. Metabolites previously reported to change in IBD were similar between *mdr1a*$^{-/-}$ and wild-type mice at stages preceding and during inflammation. We also found that changes in metabolite profiles did not correlate with colitis scores. However, metabolite changes could discriminate *mdr1a*$^{-/-}$ mice from wild-type controls, suggesting they could have value in predicting risk of IBD with a potential clinical use in at least a subset of individuals with MDR1A polymorphisms.

To assess whether changes in antimicrobial proteins (AMPs) accounted for observed differences in mucus microbiota composition, we also investigated the expression of regenerating islet-derived protein 3 γ (Reg3γ), angiogenin 4 (Ang4), β-defensin 1 and resistin-like molecule beta (Relm-β) in the colon. We found similar levels of these AMPs as well as IgA-producing plasma cells between *mdr1a*$^{-/-}$ and wild-type mice, suggesting that other factors contribute to alterations in microbiota composition.

Overall, our data indicate that the *mdr1a*$^{-/-}$ is a good model of colitis, as it enables us to look at pre-clinical changes in the gut microbiota. This work suggests the importance of mucus sampling for sensitive detection of microbiota changes. Furthermore, metabolite profiling may be a helpful way to discriminate genetic susceptibility to disease.

Declaration

I declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/display.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library's regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses.

Acknowledgements

First of all, I would like to thank my supervisors Sheena Cruickshank, Kathryn Else, Andrew McBain and Geoff Warhurst for their help, guidance and support. Especially I would like to thank Sheena and Kathryn for supporting and encouraging me throughout my PhD. I would also like to thank all the members of Cruickshank and Else groups, past and present, for their help and advice: Larisa, Maff, Becky, Emma, Munirah, Rinal, Michael, Szu-Wei, Tarfa, Becky, Helen, Rachel, Ruth and Gurdeep. Szu-Wei, I will never forget the cakes you brought me when I was working long hours in the laboratory during the weekend. Gurdeep, I really enjoyed our talks about data analysis and thanks for helping me with computer programming. Also, thanks to my advisor Mark Travis for his advice during my PhD.

I would like to thank BBSRC for funding my project and the University of Manchester for offering me a scholarship to cover my living expenses.

I want to acknowledge the help and advice I received from Core Facilities from the University of Manchester. Thanks to Andy, Stacey and Claire from Core Genomics Facility for helping me with my sequencing experiments and qPCR assays. A special thank you to Roger and Steve from the Bioimaging Facility for their tremendous help in microscopy and image processing. I would also want to thank Peter from Histology Facility for all his advice and help on tissue processing. I want to thank all those people from 4th floor of AV Hill, who assisted me during my PhD.

A massive thanks to Maria Constantinou for her help in programming and for her emotional support during my PhD. I would like to thank my friend Jing for her positive thinking. Our yearly expeditions were really something to look for. I would also want to thank my cousin, Nick Petroulakis, for his help in graphic designing.

Finally, I am eternally grateful to my parents for always encouraging me to achieve my dreams and fulfill my goals in life. Mum and dad, thanks for always being there for me and constantly supporting me these four years I have been far away from home.

Abbreviations

ABC: avidin-biotin complex

AD: atopic dermatitis

ADAM17: a disintegrin and metalloproteinase

AID: activation-induced cytidine deaminase

AIEC: adherent invasive *Escherichia coli*

AMPs: antimicrobial proteins

ANG4: angiogenin 4

APRIL: a proliferation-inducing ligand

ATG16L1: autophagy related 16-like 1

ATP: adenosine triphosphate

BAFF: B-cell activating factor

BSA: bovine serum albumin

^{13}C: carbon 13

CCR6: CC-chemokine receptor 6

CD: Crohn's disease

CE: capillary electrophoresis

CECs: colonic epithelial cells

CFUs: colony forming units

CRS1: cryptidin-related sequence 1

CX3CR1: CX3C chemokine receptor 1

DAB: 3, 3'-diaminobenzidine

DAPI: 4',6-diamidino-2-phenylindole

DCs: dendritic cells

DGGE: denaturing gradient gel electrophoresis

DSS: dextran sulfate sodium

EPEC: enteropathogenic *Escherichia coli*

EGFR: epidermal growth factor receptor

ER: endoplasmic reticulum

FDR: false discovery rate

FISH: fluorescence in situ hybridization

Foxp3: forkhead box P3

FT-IR: fourier transform infrared spectroscopy

FUT2: fucosyltransferase 2

FXR: farnesoid X receptor

GALT: gut-associated lymphoid tissue

GC: gas chromatography

GI: gastrointestinal

GIT: gastrointestinal tract

GPBAR1: G-protein-coupled bile acid receptor 1

GPR: G-protein coupled receptor

GWAS: genome wide association studies

^{1}H: hydrogen 1

H&E: haematoxylin and eosin

HBD1: human β-defensin 1

HF/HS: high fat/ high sugar

HIP/PAP: hepatointestinal pancreatic/pancreatitis-associated protein

HRP: horseradish peroxidase

H/UPLC: high/ultra performance liquid chromatography

IBD: inflammatory bowel disease

IECs: intestinal epithelial cells

IFN-γ: interferon-gamma

IgA: immunoglobulin A

IL: interleukin

ILCs: innate lymphoid cells

IT-TOF: time-of-flight mass spectrometer

KEGG: Kyoto Encyclopaedia of Genes and Genomes

KO: knockout

LC: liquid chromatography

LDA: linear discriminant analysis

LRR: leucine-rich repeat

LPS: lipopolysaccharide

M cells: microfold cells

MaAsLin: multivariate association with linear models

MALDI: matrix assisted laser desorption ionization

MDP: muramyl dipeptide

Mdr1a: multidrug resistance gene 1 a

MS: mass spectrometry

MUC2: mucin 2

MyD88: myeloid differentiation primary response gene 88

m/z ratio: mass/charge ratio

NF-κB: nuclear factor kB

NLR: Nod-like receptor

NLRP6: Nod-like receptor pyrin domain-containing protein 6

NMDS: non-metric multidimensional scaling

NMR: nuclear magnetic resonance spectroscopy

NOD2: nucleotide oligomerization domain 2

NR1H4: nuclear receptor subfamily 1 group H member 4

NSTI: nearest sequenced taxon index

OCT: optimal cutting temperature compound

OTU: operational taxonomic unit

PAP: pancreatitis associated protein

PAS: periodic acid- Schiff's

PBS: phosphate-buffered saline

PCA: principal component analysis

PCoA: principal coordinate analysis

Pgp: P-glycoprotein

PICRUSt: phylogenetic investigation of communities by reconstruction of unobserved states

pIgR: polymeric immunoglobulin receptor

PLS: Partial least squares

PRRs: pattern recognition receptors

PSA: polysaccharide A

QC: quality control

QIIME: Quantitative Insights Into Microbial Ecology

qPCR: quantitative PCR

Rag2$^{-/-}$: recombination-activating gene 2 deficient

RF: random forests

REG: regenerating islet-derived protein

Relm-β: resistin-like molecule beta

RELM/FIZZ: resistin-like molecules found in inflammatory zone

RiPPs: ribosomally synthesized post-translationally modified peptides

RNase: ribonuclease

rRNA: ribosomal RNA

RORγt: retinoic-related orphan receptor gamma t

RSD: relative standard deviation

RT: room temperature

RT-qPCR: real time quantitative polymerase chain reaction

S. aureus: Staphylococcus aureus

SCFAs: short chain fatty acids

SFB: segmented filamentous bacteria

Sox9: SRY (Sex Determining Region Y)-Box 9

SPF: specific, pathogen-free

STAMP: STatistical Analysis of Metagenomic Profiles

T1D: type 1 diabetes

TCA: tricarboxylic acid

TCF4: T-cell specific transcription factor 4

TEMED*: N,N,N′,N′*-Tetramethylethylenediamine

TGF-β: transforming growth factor-β

T_H: T helper cell

TLR: Toll-like receptor

TNBS: trinitrobenzene sulfonic acid

TNF-α: tumor necrosis factor-alpha

TOMMs: thiazole/oxazole-modified microcins

Treg: regulatory T cell

TRUC*: T-bet*$^{-/-}$ x *Rag2*$^{-/-}$ ulcerative colitis

TSA: Tyramide Signal Amplification

UC: ulcerative colitis

UPGMA: Unweighted Pair Group Method with Arithmetic Mean

VIP: variable importance in projection

WT: wild-type

ZO: zonula occludens

Chapter 1

Introduction

1.1 Overview

In humans, more than 100 trillion microorganisms, mostly bacteria, inhabit the gastrointestinal (GI) tract with the majority of them colonizing the large intestine (Backhed et al., 2005; Gill et al., 2006; Kamada et al., 2013). Gut microbes form a symbiotic relationship by aiding in food digestion and performing beneficial functions necessary for host physiology. However, the existence of such a high number of bacteria in the gut poses a risk of infection and inflammation, and thus demands a vigilant immune system, whereby bacterial entry into tissues or in the bloodstream should be prohibited or disabled while tolerance to commensal bacteria should be maintained. The intestinal epithelium plays a central role in maintaining tissue homeostasis by mediating the interactions between bacterial cells in the luminal contents and the underlying mucosal immune cells. Disruption of tolerance to gut microbes results in an inappropriate exaggerated immune response with detrimental outcomes like in the case of inflammatory bowel disease (IBD). IBD is a condition of chronic inflammation in the gut, whose aetiology remains unknown despite its increasing prevalence. Interestingly, mice genetically engineered to develop IBD do not exhibit disease symptoms when raised in germ-free conditions and faecal diversion in individuals suffering from IBD results in remission in the excluded gut segments. Additionally, a number of studies in IBD patients and experimental mouse models have reported altered microbial balance during IBD, referred to as dysbiosis. These findings implicate gut microbes as a contributory factor to IBD pathogenesis.

Studies on microbial composition in IBD have focused on microbial changes in faeces during onset of clinical symptoms before treatment or during disease establishment and management. However, the gut microbiota is diverse along the length of the GI tract and there are specific niches across the gut. In fact very little is known about the bacteria inhabiting niches other than the gut lumen such as the mucus, as most analysis focuses on faeces. This is surprising considering that the mucus layer is most relevant when addressing gut microbe-epithelial crosstalk. Key unanswered questions are therefore whether a shift in microbial composition is present preceding IBD onset and whether alterations in microbial communities are present in both mucus and faeces or in either one of them. How altered microbial composition affects subsequent disease progression, at what extent it is associated with bacterial gut localisation and how intestinal epithelium impacts on shaping the gut ecosystem are further points that need to be addressed. Dysbiosis also affects microbial

community function. Indeed, microbial function and metabolism are altered in IBD. However, due to the fact that current studies on microbial function and metabolism have focused on changes after the onset of IBD, it is unclear whether such changes precede IBD or are the result of it.

The aim of this thesis is to use a mouse model to perform a systematic analysis of the microbial census found in faeces and mucus at stages preceding the onset of any detectable inflammation and at disease manifestation. Moreover, this thesis aims to identify potential metabolic changes of host or microbial origin that could strongly influence the host-microbiota crosstalk during the course of colitis progression. Finally, the role of the intestinal epithelial barrier in moulding microbial composition is considered. To address these aims, I will analyse composition, gene encoding potential and gut location of faecal and mucus microbial communities at pre- and post- inflammation time points using a mouse model of spontaneous colitis- the multidrug resistance gene 1 a (*mdr1a*$^{-/-}$) model. I will also determine metabolic profiles in urinary samples and expression patterns of markers associated with the gut barrier such as antimicrobial proteins.

1.2 The intestinal epithelial barrier function

The GI tract consists of a monolayer of polarized epithelial cells that prevents inadvertent entry of bacterial antigens in underlying sterile parts of the body. Epithelial cells adopt a polarized structure distinguishing apical and basal surfaces. The apical surface of the epithelium is exposed to the external environment, while the basolateral surface faces internal structures (Abreu, 2010; Wilson et al., 2011). The basal surface covers a layer of connective tissue, called the basal lamina, which consists of a network of fibrous substances such as collagen, polysaccharides and matrix components including fibronectin and laminin (Abreu, 2010; Veldhoen and Brucklacher-Waldert, 2012).

The intestine is the largest mucosal surface area in the body where bacterial communities and host cells are juxtaposed (Abreu, 2010; Cerf-Bensussan and Gaboriau-Routhiau, 2010; Peterson and Artis, 2014). In addition, the GI tract is a major target for pathogens via ingested substances. Thus, the mucosal epithelium is constantly challenged with potential bacteria colonization and invasion from both commensal and pathogenic organisms. This has

led to the development of a formidable array of mechanisms including physical and biochemical barriers between the external environment encompassing dietary and bacterial antigens in the gut lumen, and the intestinal mucosa.

Intestinal epithelial cells (IECs) establish and maintain a continuous barrier to avert potential microbial incursion. Epithelial cells that line the gut lumen are tightly packed with tight junctions and adherens junctions, forming the apical junctional complex, as well as desmosomes (Henderson et al., 2011; Peterson and Artis, 2014; Turner, 2009). Tight junctions are found nearest to the intestinal lumen and neighbouring adherens junctions are beneath them. A network of actin and myosin supports both of these types of junctions, which are composed of transmembrane adhesion proteins (such as cadherin in the case of adherens junctions, and claudin and occludin in the case of tight junctions), scaffolding proteins (like zonula occludens 1 (ZO1) and ZO2 for tight junctions) and junctional molecules including catenin and different kinases (Henderson et al., 2011; Turner, 2009).

Tight junctions seal the gaps between adjacent epithelial cells controlling permeability of luminal components. However, the barrier is dynamic responding to physiologic and immunologic stimuli such as pro-inflammatory cytokines and as a consequence its permeability can be adapted (Henderson et al., 2011; Turner, 2009). Defective tight junction function leads to increased influx of microbial products and potential food allergens across the barrier and this enhanced permeability characterizes the 'leaky' gut phenotype (Turner, 2009). In Crohn's disease (CD), a form of IBD, patients in remission that had increased intestinal permeability relapsed in just one year (Wyatt et al., 1993). This evidence alongside a series of observations in experimental mouse models (Madsen et al., 1999; Su et al., 2009) implicates gut permeability as a contributor to IBD progression. Importantly, isolated barrier dysfunction *per se* is insufficient to cause IBD, as immunoregulatory mechanisms are activated upon barrier dysfunction to suppress inflammation (Boirivant et al., 2008; Iliev et al., 2009). Therefore, IECs are of paramount importance in maintaining intestinal homeostasis by regulating the precarious balance between inflammatory and anti-inflammatory responses.

The diversity of functions carried out by IECs is integral to the maintenance of homeostasis, as they physically separate commensal bacteria, sense and respond to the external environment and regulate the outcome of immune response either to inflammation or

immunoregulation (Henderson et al., 2011; Peterson and Artis, 2014; Rescigno, 2011). The epithelial monolayer is composed of four main cell types: absorbent enterocytes, which make up most of the IECs, mucus-producing goblet cells, hormone-producing enteroendocrine cells and antimicrobial proteins (AMPs) secretory Paneth cells (Abreu, 2010; Cerf-Bensussan and Gaboriau-Routhiau, 2010; Veldhoen and Brucklacher-Waldert, 2012) (Figure 1). The luminal secretion of mucin glycoproteins, which are the major components of mucus, and AMPs promotes microbial exclusion from the intestinal epithelial surface (Abraham and Medzhitov, 2011; Peterson and Artis, 2014; Saleh and Elson, 2011). Immunoglobulin A (IgA) produced by plasma B cells is secreted in the mucus layer further contributing to the antibacterial defence (Sartor, 2006).

Specialized IEC lineages are derived from a common stem cell progenitor located in the base of intestinal crypts (Crosnier et al., 2006). Other specialized IECs, microfold cells (M cells), sample luminal antigens and whole bacteria in order to present them to closely associated mononuclear phagocytes, such as dendritic cells (DCs) and macrophages (Mabbott et al., 2013). M cells are concentrated in follicle-associated epithelium overlying distinct lymphoid structures including Peyer's patches in the small intestine and isolated lymphoid follicles in the colon (Mabbott et al., 2013; Peterson and Artis, 2014). It was also shown that goblet cells themselves deliver soluble antigens directly to lamina propria tolerogenic $CD103^{+}$ DCs in the small intestine (McDole et al., 2012). In addition to luminal antigen sampling by specialized IECs, CX3C chemokine receptor 1 (CX3CR1) positive cells of the phagocytic lineage, which are interspersed in the epithelial monolayer, uptake bacterial antigens through extension of transepithelial processes (Niess et al., 2005; Rescigno et al., 2001; Rossini et al., 2014). Despite the establishment of an intestinal epithelial barrier, it is specifically adapted to allow limited and controlled bacteria-immune system interactions directing appropriate tolerogenic or anti-pathogen immune responses. Therefore, the intestinal epithelium acts as the gatekeeper of homeostatic functions in the gut, being selectively permeable only to allow tolerance to commensal bacteria colonisation and to mount an effective inflammatory response against potential pathogens.

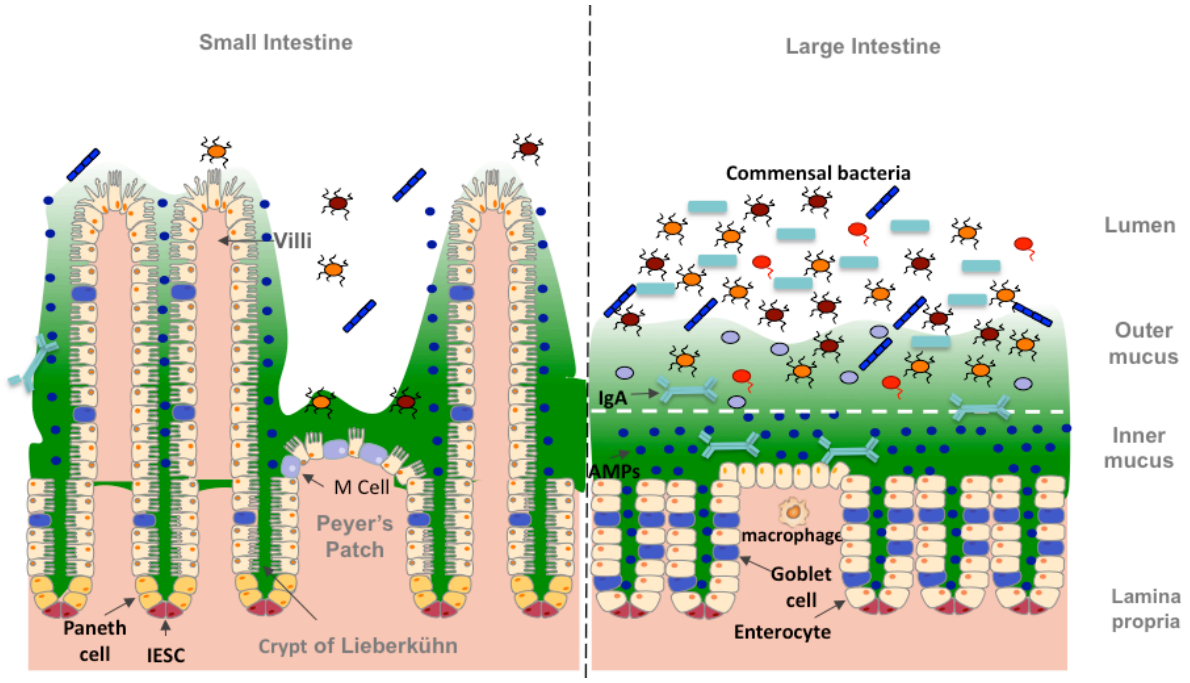


Figure 1. Landscape of the intestinal mucosa. A single layer of epithelial cells sequesters commensal microbiota from the underlying lamina propria, thereby functioning as a physical barrier. The intestinal epithelium comprises enterocytes, mucus-producing goblet cells, enteroendocrine cells and antimicrobial proteins (AMPs) secretory Paneth cells. Paneth cells and intestinal stem cells are located at the base of small intestinal crypts, called crypts of Lieberkühn. Enterocytes and goblet cells can also secrete AMPs. The mucus layer covering the intestinal epithelium limits bacterial access to the epithelial surface, providing another line of defence. AMPs and secretory IgA are concentrated in the mucus layer and contribute to the mucus barrier function. In the large intestine, the mucus consists of an inner layer (dotted line) that is devoid of bacteria and a thicker outer layer that commensal bacteria reside. In contrast, a single mucus layer is present in the small intestine. Lamina propria, a connective tissue layer, is beneath the epithelial barrier and encompasses immune cells. Intestinal immune cells in the lamina propria may be organized in distinct lymphoid structures such as Peyer's patches. Microfold cells (M cells) reside above Peyer's patches and sample the instestinal lumen. Commensal bacteria inhabit the outer mucus layer and the lumen. The density and diversity of commensal bacteria increase from the small intestine to the large intestine.

1.2.1 Mucus layer

Mucus is a viscous highly hydrated layer that covers the GI tract and provides the first line of defence against potential microbial invasion in the host epithelium (Johansson et al., 2011a; Pelaseyed et al., 2014). Mucins and especially mucin 2 (MUC2) are the main structural components of the mucus barrier (Gallo and Hooper, 2012; McGuckin et al., 2011). Mucins are heavily O-glycosylated glycoproteins that are tightly packed in goblet cells and expand upon secretion to form enormous net-like structures that characterize the mucus layer (Birchenough et al., 2015; Pelaseyed et al., 2014). Apart from the gel-forming mucins, there are also membrane bound mucins produced by enterocytes. These transmembrane mucins form the glycocalyx extending about a micrometer on the apical surface of enterocytes and major components comprise MUC3, MUC12 and MUC17 (Pelaseyed et al., 2014). The trasmembrane mucins also have a role in protecting the host from microbial pathogenic attacks (McAuley et al., 2007; McGuckin et al., 2011).

Mucus organization differs along the GI tract between the small intestine and the large intestine (Figure 1). The mucus in the colon consists of an inner stratified dense layer that is firmly attached to the epithelium and is impermeable to bacteria, and a thicker outer loose layer that provides a habitat for commensal bacteria (Johansson et al., 2008). Therefore, the inner mucus layer keeps bacteria at a distance from the epithelial surface in order to maintain the delicate balance between host and microbiota (Gallo and Hooper, 2012; McGuckin et al., 2011). The inner mucus layer, which is about 50μm thick in the mouse colon and 200-300μm in humans, is proteolytically cleaved and converted to the outer mucus layer that has an expanded volume reaching about 100μm in mice and 700μm in humans (Atuma et al., 2001; Johansson et al., 2014; Johansson et al., 2011b; Johansson et al., 2008). Bacteria residing in the expanded outer mucus layer gain access to the glycan-rich mucins, which are an important energy source (Johansson et al., 2011b; Sonnenburg et al., 2005). There is a great diversity of glycan patterns including fucosylated, sulphated and sialylated glycans (Holmen Larsson et al., 2013) that can be utilized by bacteria possessing appropriate enzymes for their degradation (Pelaseyed et al., 2014). Furthermore, the MUC2 O-glycans not only serve as nutrients, but also provide attachment sites for commensal and pathogenic bacteria (Juge, 2012). Thus, the host mucin glycan content can affect the composition of the gut microbiota with implications on host physiology and susceptibility to inflammation (Sommer et al.,

2014).

The small intestine lacks the two-layered mucus system present in the colon and instead it has a single mucus layer (Pelaseyed et al., 2014). The mucus layer of the small intestine is not attached to the epithelium and is permeable to bacteria (Ermund et al., 2013). Since an inner mucus layer is absent, Paneth cell secretions and enterocyte produced antimicrobial factors are involved in restricting epithelial contact with bacteria in the small intestine (Johansson and Hansson, 2011; Vaishnava et al., 2011). Additionally, this non-attached mucus gel is continuously renewed by secretion from goblet cells and moves with peristaltic propulsion in the distal direction, so that bacteria that have successfully colonized mucus are continuously washed out (Johansson and Hansson, 2011; Pelaseyed et al., 2014). The turnover rate of the inner colonic mucus layer is also high enough to keep it free from bacteria (Johansson et al., 2008).

The intestinal mucus is a dynamic system and its properties can be influenced by microbial signals and the immune system (Birchenough et al., 2015; Jakobsson et al., 2014b; Johansson et al., 2015; McGuckin et al., 2011). Germ-free mice exhibited mucus attached to the epithelium in the small intestine and penetrable inner mucus layer and lower Muc2 amounts in the colon compared with conventionally raised animals (Johansson et al., 2015; Rodriguez-Pineiro et al., 2013). Secretion of Muc2 was enhanced in infected mice probably as a mechanism to flush enteric pathogens from the mucosal surface (Bergstrom et al., 2010). Moreover, bacteria have been shown to penetrate the inner mucus layer in mice that develop spontaneous colitis and in patients with active ulcerative colitis (UC), a form of IBD (Johansson et al., 2014; Johansson et al., 2010). Some UC patients in remission also have penetrable mucus (Johansson et al., 2014). Furthermore, mouse models in which critical components of the mucus have been depleted are predisposed to the development of colitis (Fu et al., 2011; Johansson et al., 2008; Van der Sluis et al., 2006). In *Muc2*$^{-/-}$ mice, which lack the protective mucus layer, bacteria were found in direct contact with the epithelium and deep down in the crypts, while in some cases bacteria were also found inside epithelial cells (Johansson et al., 2008). Therefore, once the inner mucus layer becomes penetrable to bacteria or is depleted, bacteria will reach the epithelial surface in large numbers, triggering the immune response and driving inflammation.

1.2.2 Antimicrobial proteins (AMPs)

AMPs provide an innate immune mechanism to augment host defence against microbial encroachment. A diverse array of AMPs is produced reflecting the complexity of gut microbes and the continuous challenge faced by the intestinal surface. AMPs comprise representatives of several major protein families including defensins, cathelicidins, C-type lectins (such as regenerating islet-derived protein (REG) family) and ribonucleases (RNases such as angiogenin 4 (ANG4)) (Bevins and Salzman, 2011; Gallo and Hooper, 2012). The types of AMPs differ along the GI tract, highlighting heterogeneous distribution of sources of origin as well as differential microbial composition (Bowcutt et al., 2014).

Paneth cells are a specialized IEC lineage found at the base of intestinal crypts, called crypts of Lieberkühn, in the small intestine, and constitute an important source of AMPs (Abreu, 2010; Bevins and Salzman, 2011). Paneth cell granules rich in AMPs including α-defensins (referred to as cryptdins in mice) (Ganz, 2003), C-type lectins (such as REG3γ in mice and its orthologue hepatointestinal pancreatic/pancreatitis-associated protein (HIP/PAP) in humans) (Cash et al., 2006; Christa et al., 1996), lysozyme C, secretory phospholipase A2 (Bevins and Salzman, 2011; Gallo and Hooper, 2012) and RNases such as ANG4 (Hooper et al., 2003) are emptied in the crypt lumen, creating a high local concentration of antimicrobials and thus protecting from microbial invasion. AMPs diffuse to the mucus layer, where they remain concentrated close to the epithelial surface providing an antibacterial barrier (Meyer-Hoffert et al., 2008). Some of these AMPs are constitutively expressed, while others are transcriptionally induced indicating that the pattern of these effector molecules can change in response to relevant stimuli (Cash et al., 2006; O'Neil et al., 1999; Wehkamp et al., 2003).

1.2.2.1 Defensins

Defensins are a large family of AMPs, typically 30-40 amino acids in length and contain many subtypes, which differ in their amino acid composition and structure (Ganz, 2003). They are classified as α or β depending on the patterns of the three disulphide bonds formed between their six cysteine residues (Bevins and Salzman, 2011; Ganz, 2003). The α-defensins are the most abundant AMPs in the human intestine and are highly expressed in Paneth cells, although their expression is also observed in neutrophils and macrophages (Bevins and

Salzman, 2011; Gallo and Hooper, 2012). Notably, α-defensins have not been detected in neutrophils in mice (Ganz, 2003). The α-defensins are largely constitutively expressed, as their T-cell specific transcription factor 4 (TCF4) is part of the Paneth cell maturation pathway (van Es et al., 2005). In contrast, β-defensins are induced in enterocytes in the large intestine with the exception of human β-defensin 1 (HBD1) that presents constitutive expression (O'Neil et al., 1999). Both α- and β-defensins have a broad range of antimicrobial effects targeting Gram-positive and Gram-negative bacteria, parasites and fungi; however the activity spectrum depends on the type of defensin (Bevins and Salzman, 2011; Gallo and Hooper, 2012). Apart from their antimicrobial activity, β-defensins can act as mediators of inflammation by attracting CC-chemokine receptor 6 (CCR6) expressing DCs and T cells (Yang et al., 1999).

Defensins perform their bactericidal role by disrupting the cell membrane on targeted bacterial cells. Their cationic charged region promotes attraction to the negatively charged bacterial surface, resulting in the formation of a channel into the lipid-rich bacterial membrane and ultimately bacterial lysis (Wilson et al., 2011). The basic mechanism of their activity is the depolarization of bacterial membranes and subsequently the loss of proton potential, which is vital for bacterial cell survival (Kagan et al., 1990). In a similar mode of action, positively charged cathelicidins, which are encountered in the large intestine, kill their targeted microorganisms (Bals and Wilson, 2003). In addition to membrane permeabilization, defensins can interfere with bacterial cell wall synthesis by binding to building block lipid II such as HBD3 (Sass et al., 2010) or self-assemble creating a network of fibrils that entraps potent bacterial intruders (Chu et al., 2012).

Expression of some members of the defensins AMP family requires microbial signalling, whereas others, mostly α-defensins are constitutively expressed (O'Neil et al., 1999; Putsep et al., 2000). HBD2 is a main representative of β-defensin family and its expression in IECs is regulated by inflammatory and bacterial stimuli (O'Neil et al., 1999; Vora et al., 2004). Interestingly, HBD2 expression is upregulated in inflamed colonic epithelial tissue obtained from patients suffering from UC (O'Neil et al., 1999). IECs can be induced to secrete β-defensin 2 in response to Toll-like receptor (TLR) -4 and -2 activation (Vora et al., 2004). Of note, TLRs belong to a family of pattern recognition receptors (PRRs) that recognizes conserved microbial motifs and have a critical role in the initiation of innate immune defence mechanisms (Abreu, 2010).

Emerging evidence shows that β-defensin expression is controlled by mucins in addition to the TLR signalling pathway. Colonic MUC2 mucus stimulates the expression of β-defensin 2 in combination with the pro-inflammatory interleukin (IL)-1β, whereas it abrogates its antimicrobial activity (Cobo et al., 2015). MUC2 protects enteric bacteria against the antimicrobial effects of β-defensin 2, favouring their growth. *Ex vivo* culture of enteropathogenic *Escherichia coli* (EPEC) showed increased numbers in the presence of colonic contents from *Muc2*$^{+/-}$ and *Muc2*$^{+/+}$ mice, but EPEC levels remained stable and even decreased when co-cultured with colonic contents from *Muc2*$^{-/-}$ mice (Cobo et al., 2015). Furthermore, the levels of β-defensin expression were reduced in *Muc2*$^{-/-}$ mice.

Other host PRRs that regulate the expression of bacterially induced defensins include nucleotide oligomerization domain 2 (NOD2). NOD2 is found intracellularly in Paneth cells, DCs and macrophages in the small intestine, where it recognizes muramyl dipeptide (MDP), a peptidoglycan-derived fragment common in Gram-positive and Gram-negative bacterial cell wall (Abreu et al., 2005; Biswas et al., 2012; Kobayashi et al., 2005). NOD2 activation by MDP binding induces antimicrobial activity in the intestinal crypts (Petnicki-Ocwieja et al., 2009). NOD2 is required for the expression of α-defensins, as NOD2 deficient mice displayed approximately a 100-fold reduction of AMPs α-defensin 4 and α-defensin-related sequence 10 (Kobayashi et al., 2005). *Nod2*$^{-/-}$ mice also have altered microbiota composition and are more susceptible to GI pathogen colonization (Kobayashi et al., 2005; Petnicki-Ocwieja et al., 2009). However, a recent study using *Nod2*$^{-/-}$ mice on pure genetic background (C57BL/6) suggests that AMP expression and microbiota composition is NOD2 status independent (Shanahan et al., 2014). The original study of *Nod2*$^{-/-}$ mice used embryonic stem cells of a different mouse strain injected into C57BL/6 blastocysts to create these mice, leading to a mixed genetic background (Kobayashi et al., 2005). Expression profile of AMPs including α-defensins was similar between *Nod2*$^{-/-}$ mice and co-housed littermate controls; the only antimicrobial found to be different was cryptidin-related sequence 1 (CRS1) (Shanahan et al., 2014). Interestingly, no major differences in faecal microbial communities were found between *Nod2*$^{-/-}$ and wild-type co-housed animals, suggesting that environmental impact is greater than host genetics on regulation of microbial communities (Shanahan et al., 2014).

To validate the role of NOD2 in shaping microbiota composition and driving AMPs expression, a subsequent study was performed using *Nod2*$^{-/-}$ mice on a C57BL/6 background

but separately housed from wild-type controls (Ramanan et al., 2014). Levels of antimicrobials such as Reg3β, resistin-like molecule beta (Relm-β), β-defensin 1 and phospholipase A2 were increased in small intestinal crypts in response to expansion of the commensal *Bacteroides vulgatus* in the lumen of *Nod2*$^{-/-}$ mice. Apart from heightened host defence expression, the intestinal epithelium carried further abnormalities in these mice including high inflammatory gene expression (Ramanan et al., 2014). No littermates were used in this study, so it does not rule out any potential confounding environmental effects. Collectively, these results indicate that NOD2-induced AMPs can modulate microbial composition restricting the pathogenic expansion of a member of the commensal microbiota to maintain homeostasis.

1.2.2.2 Reg3γ

Reg3γ is a C-type lectin that belongs to the REG family and is predominantly expressed in the small intestine (Cash et al., 2006), although it is also produced in the large intestine albeit at low levels (Dieckgraefe et al., 2002). Expression levels of Reg3γ in the distal colon are almost undetectable in contrast to the levels in the proximal colon (Burger-van Paassen et al., 2012). Reg3γ recognizes carbohydrate moieties of peptidoglycan on the surface of Gram-positive bacteria and acts by incurring damage to their bacterial cell wall (Cash et al., 2006; Lehotzky et al., 2010). Reg3γ functions by disrupting bacterial membrane through the formation of hexameric oligomeric pores (Mukherjee et al., 2014). Reg3β, another REG family member, is usually co-expressed with Reg3γ, but it targets Gram-negative along with Gram-positive bacteria (Miki et al., 2012).

The presence of intestinal bacteria is necessary for triggering the expression of Reg3γ in Paneth cells and enterocytes (Cash et al., 2006; Ogawa et al., 2003; Vaishnava et al., 2008). Goblet cells also express Reg3γ and Reg3β as was shown in *Muc2*$^{-/-}$ mice (Burger-van Paassen et al., 2012). Studies in germ-free mice showed reduced Reg3γ expression before bacterial reconstitution from conventionally raised mice (Cash et al., 2006). Deficiency of the myeloid differentiation primary response gene 88 (MyD88), an adaptor protein in the TLR signalling pathway, resulted in diminished expression of Reg3γ and increased susceptibility to *Listeria monocytogenes* infection in the gut (Brandl et al., 2007). Therefore, epithelial cell intrinsic MyD88-mediated signalling induces Reg3γ expression, which in turn prevents the

dissemination of pathogens and commensal bacteria across the mucosal barrier into the underlying intestinal tissue (Vaishnava et al., 2008; Vaishnava et al., 2011).

In addition to microbial colonization, epithelial cell expression of Reg3γ is dependent on IL-22 derived from group 3 innate lymphoid cells (ILCs) residing in the lamina propria (Sanos et al., 2009; Zheng et al., 2008). Group 3 ILCs are innate immune cells that express the transcription factor retinoic-related orphan receptor gamma t (RORγt) and produce the cytokines IL-17A, IL-17F and IL-22 (Tait Wojno and Artis, 2012). Neutralization of IL-22 and depletion of the ILC population in recombination-activating gene 2 deficient (*Rag2*$^{-/-}$) mice, which lack mature lymphocytes, resulted in low levels of Reg3γ and Reg3β expression (Sanos et al., 2009). In the colon, the epithelial leucine-rich repeat (LRR) C19 receptor also mediates the expression of REG3 proteins through the nuclear factor kB (NF-κB) signalling pathway (Cao et al., 2016).

Reg3γ is a key AMP that functions to circumvent a potential intestinal barrier breach and preserves barrier integrity by controlling the numbers of bacteria in contact with the intestinal epithelium (Cash et al., 2006; Loonen et al., 2013; Vaishnava et al., 2011). However, the activity of Reg3γ is limited to mucosa-associated bacteria and it is not thought to affect luminal bacterial composition and quantity (Vaishnava et al., 2011). Furthermore, Reg3γ targets specifically Gram-positive mucosa-associated bacteria. Thus, Reg3γ limits contact of bacteria with the intestinal epithelial surface establishing a zone of approximately 50μm (Vaishnava et al., 2011). Recent evidence shows that REG3 family members also reduce the translocation of commensal bacteria to extra-intestinal tissues such as mesenteric lymph nodes and liver following ethanol administration (Wang et al., 2016). Moreover, Reg3γ regulates mucus properties, as *Reg3γ*$^{-/-}$ mice have altered mucus distribution (Loonen et al., 2013). Increased expression of Reg3γ during inflammation may function as a mechanism to counteract microbial colonization and subsequent intestinal epithelium penetration (Brandl et al., 2007; Burger-van Paassen et al., 2012; Dessein et al., 2009; Loonen et al., 2013; Ramanan et al., 2014; van Ampting et al., 2009). Reg3γ and its human homolog HIP/PAP expression are greatly enhanced in dextran sulfate sodium (DSS)- induced colitis and in inflamed colon of IBD patients respectively (Dieckgraefe et al., 2002; Mizoguchi et al., 2003; Ogawa et al., 2003).

1.2.2.3 Ang4

Ang4, a member of the angiogenins family, is an antimicrobial RNase, which is expressed in Paneth cells with a role in host defence (Crabtree et al., 2007; Hooper et al., 2003). In addition to restricting bacterial access to the host epithelium, it is involved in vascular network formation in the intestine (Crabtree et al., 2007; Stappenbeck et al., 2002). Ang4 has selective bactericidal activity against both Gram-positive and Gram-negative bacteria (Hooper et al., 2003), forming pore-like structures on the bacterial cell membrane (Walker et al., 2013). Ang4 expression is microbially regulated as shown by experiments using mono-colonised germ free mice (Hooper et al., 2003) and conventionally raised mice during weaning (Burger-van Paassen et al., 2012; Hooper et al., 2003). Microbial challenge can also elicit Ang4 production in an indirect way. Upon exposure to *Salmonella enterica* serovar Typhimurium or commensal bacteria, IECs secrete IL-23, which promotes IL-22 release from γδ intraepithelial lymphocytes triggering Ang4 expression in Paneth cells (Walker et al., 2013). However, the extent of IL-22 contribution in Ang4 production is yet to be defined.

The cellular source of Ang4 in the large intestine is goblet cells as shown in a study of parasitic *Trichuris muris* infection (Forman et al., 2012). Its expression varies along the length of the large intestine with higher levels in the proximal colon (Burger-van Paassen et al., 2012). Ang4 expression was increased in response to *Trichuris muris* in resistant mice and was dependent on IL-13, a T helper cell (T_H) 2 cytokine (Forman et al., 2012). The role of IL-13 in driving Ang4 expression is further supported by data in IL-9 transgenic mice, in which IL-9 drives upregulation of Ang4 in Paneth cells through IL-13 (Steenwinckel et al., 2009). Moreover, IL-25, a T_H2 cytokine secreted by IECs following microbial stimulation, mediates induction of Ang4 via the downstream activity of IL-13 (Noor et al., 2016). Taken together, these data suggest that Ang4 expression is under microbial control and can be regulated by cytokine pathways.

1.2.2.4 Relm-β

Relm-β is a goblet-cell specific protein that belongs to the small cysteine-rich resistin-like molecules found in inflammatory zone (RELM/FIZZ) family (He et al., 2003; Steppan et al., 2001). It is expressed throughout the large intestine with highest levels of expression in the

distal colon and caecum and lower expression in the proximal colon (He et al., 2003). It is also minimally expressed in the ileum, but it is absent from the rest of the GI tract (McVay et al., 2006; Steppan et al., 2001). Gut commensal microbes stimulate Relm-β expression and secretion in the intestinal lumen, as germ-free mice exhibit essentially undetectable levels of this protein (He et al., 2003). Its induction relies on MyD88 activation in response to TLR recognition of bacterial antigens (Vaishnava et al., 2008). Expression of Relm-β is further dependent on the T_H2 cytokine IL-13, as its levels were robustly upregulated after GI nematode infection in resistant mice (Artis et al., 2004). In support to the role of IL-13 in triggering the expression of this effector molecule, a study in DSS-treated mice showed that increased levels of Relm-β were downstream of IL-13 signaling (Hogan et al., 2006).

Relm-β has a crucial role in maintaining epithelial barrier integrity and regulating colonic inflammation. Relm-β deficiency was associated with increased intestinal permeability concomitant with an altered gene transcriptional profile including reduced Reg3γ and Reg3β levels (Hogan et al., 2006; Morampudi et al., 2016). Additionally, Relm-β promotes mucin expression and secretion, thus strengthening the epithelial defence barrier (Krimi et al., 2008). As Relm-β has a role in intestinal barrier function, it may be assumed that its depletion increases susceptibility to inflammation; however, the absence of Relm-β was protective for DSS-treatment or nematode infection-induced inflammation (Hogan et al., 2006; McVay et al., 2006; Nair et al., 2008). Nevertheless, data are conflicting as Relm-β deficiency in trinitrobenzene sulfonic acid (TNBS) and infectious colitis was associated with increased disease severity (Baumler et al., 2015; Hogan et al., 2006; Krimi et al., 2008). Taken together, the disparity in these models of inflammation regarding the role of Relm-β suggests that its biological function may be context dependent.

Relm-β affects susceptibility to inflammation by regulating innate and adaptive immune responses (Baumler et al., 2015; McVay et al., 2006; Nair et al., 2008). It has been reported that it activates macrophages to produce the pro-inflammatory cytokine tumor necrosis factor-alpha (TNF-α) in DSS-induced colitis, aggravating mucosal injury (McVay et al., 2006). The following studies have demonstrated that Relm-β also augments interferon-gamma (IFN-γ) production by $CD4^+$ T cells during chronic nematode infection, contributing to the persistence and severity of infection (Nair et al., 2008). Importantly, Relm-β drives the recruitment of $CD4^+$ T cells during pathogen infection to increase the proliferation of IECs, thereby limiting tissue damage suffered by infection (Baumler et al., 2015). Collectively,

these data indicate that Relm-β functions as a mediator of the inflammatory response with an impact on the host immune system. Given the induction of Relm-β expression in response to inflammatory stimuli, it remains to be determined whether it can recruit $CD4^+$ T cells in different colitis settings.

Recent evidence shows that Relm-β promotes colitis in *Muc2*$^{-/-}$ mice by altering commensal microbial communities (Morampudi et al., 2016). Relm-β is strongly upregulated in *Muc2*$^{-/-}$ mice in concert with colitis development. As *Muc2*$^{-/-}$ mice have a dysbiotic microbiota, Relm-β functions by exaggerating dysbiosis and worsening the clinical course of disease. Relm-β dramatically induces the expression and secretion of antimicrobial lectins including Reg3β, and Reg3β in turn exerts its bactericidal effect mostly on Gram-positive *Lactobacillus* species, resulting in their selective loss (Morampudi et al., 2016). Therefore, Relm-β modifies gut microbial ecology indirectly by regulating the expression of other AMPs and in particular Reg3β.

1.2.3 IgA

IgA, the most abundant antibody isotype in mucosal secretions, is produced by plasma cells in the lamina propria and secreted in the intestinal lumen, where it contributes to innate defence mechanisms by restricting bacterial access to the intestinal epithelium (Peterson and Artis, 2014). DCs conditioned by IEC-derived signals promote the maturation of naïve B cells into IgA-secreting plasma cells and their gut homing through nitric oxide, IL-10, retinoic acid and transforming growth factor-β (TGF-β) (Macpherson and Uhr, 2004; Mora et al., 2006; Mora and von Andrian, 2008). B cell maturation entails heavy chain class-switch recombination and is dependent on cognate $CD4^+$ T cell stimulation. In the absence of T-cell help, IECs influence local IgA production by releasing factors such as B-cell activating factor (BAFF) and a proliferation-inducing ligand (APRIL) or induce mucosal DCs to secrete these factors (Gutzeit et al., 2014; Mora and von Andrian, 2008; Peterson and Artis, 2014). Intestinal bacteria have an integral role in eliciting the differentiation of IgA producing plasma cells, thus increasing the secretion of IgA in the lumen (Fritz et al., 2012; Peterson et al., 2007; Suzuki et al., 2010). For example, bacteria-derived flagellin upregulates retinoic acid production in DCs driving the generation of IgA positive B cells (Peterson et al., 2007).

IECs mediate the transport of secretory IgA in the mucus layer. Dimeric IgA complexes initially bind to the polymeric immunoglobulin receptor (pIgR) on the basolateral membrane of IECs and then are actively transcytosed to the lumen (Henderson et al., 2011; Peterson and Artis, 2014). IgA opsonises bacterial components and soluble antigens, impeding a possible interaction with intestinal epithelial cells and immune cells (Brown et al., 2013). Additional actions of IgA in protecting the host include alterations of IgA-bound bacteria such as reduction of bacterial epitope expression, and bacterial fitness by altering gene expression implicated in oxidative stress response (Peterson et al., 2007). IgA can also impact directly on the host dampening pro-inflammatory gene expression (Peterson et al., 2007). Therefore, IgA production protects from a damaging potent immune response to commensal microbiota. Studies in activation-induced cytidine deaminase (AID) deficient mice, which lack class-switch recombination and somatic hypermutation of the Ig genes and as a result are IgA-deficient, have shown a dramatic expansion of the Firmicutes member, epithelial-associated segmented filamentous bacteria (SFB), in the small intestine and over-activation of the mucosal immune system (Suzuki et al., 2004). Furthermore, mice selectively deficient in somatic hypermutation present an expansion in anaerobic bacteria in the small intestine, immune hyperactivation and compromised defence, indicating the importance of a diverse IgA repertoire to control the complex gut microbiota (Wei et al., 2011). Collectively, secretory IgA provides an adaptive immune component to the epithelial barrier regulating gut microbial communities to preserve immune homeostasis.

1.3 Gut microbiota

The human gut is home to an outstanding amount of microbes, which surpass in number human cells by an order of magnitude (Gill et al., 2006). Gut microbiota is composed mainly of bacteria, but it also includes archaea, viruses, fungi and parasites (Backhed et al., 2005; Gill et al., 2006; Qin et al., 2010). The gut microbiota is dominated by anaerobic bacteria and consists of approximately 500-1000 bacterial species, which belong to a small fraction of the known bacterial phyla (2012; Peterson et al., 2008; Qin et al., 2010; Sommer and Backhed, 2013). The most abundant phyla in the gut are Bacteroidetes and Firmicutes (Eckburg et al., 2005; Gill et al., 2006; Nielsen et al., 2003; Suau et al., 1999; Wang et al., 2003). The principal classes in Firmicutes phylum are Clostridia and Bacilli, whereas the predominant

class in Bacteroidetes is Bacteroides. Other bacterial phyla such as Fusobacteria, Actinobacteria, Proteobacteria and Verrucomicrobia are less frequently encountered (De Filippo et al., 2010; Eckburg et al., 2005; Nielsen et al., 2003; Qin et al., 2010; Wang et al., 2003). The density and diversity of microbial communities increase along the GI tract with the highest concentration found in the large intestine (Ahmed et al., 2007; Peterson et al., 2008) (Figure 1). The distribution of the gut microbiota also differs across the gut. Complex and diverse microbial populations reside in the gut lumen and in the outer mucus layer of the intestinal mucosa (Johansson et al., 2008). Despite the fact that bacteria are in close proximity with the intestinal immune system, they have managed to accomplish a mutualistic relationship with the host.

A long co-evolutionary history between the host and gut microbiota has led to the development of a symbiotic relationship in which microbiota perform many beneficial functions and in turn, the host provides nutrient resources and niches for their survival (Kamada and Nunez, 2013; Kamada et al., 2013). Commensal bacteria contribute to the fermentation of complex carbohydrates, synthesis of vitamins, development of gut-associated lymphoid tissue (GALT), inhibition of intestinal colonization by pathogens and modulation of immune responses (Diehl et al., 2013; Hooper, 2004; Kamada and Nunez, 2013; Kamada et al., 2013; Lathrop et al., 2011). The diversity of functions performed by the gut microbiota is illustrated in the number of genes it encodes. The collective number of genes encoded by microbes' genome (the microbiome) outnumbers at least 100 times those of the host (2012; Backhed et al., 2005; Gill et al., 2006). A metagenomic analysis of the gut microbiome revealed that the majority of genes encoded by bacterial species are implicated in the metabolism of amino acids, polysaccharides and xenobiotics, as well as in the methanogenesis process and vitamin synthesis including the 2-methyl-D-erythritol-4-phosphate pathway (Gill et al., 2006). Thus, the microbiome provides biochemical and metabolic functions complementing host's physiology.

1.3.1 Gut microbiota and host immune system

Gut microbiota contributes to the development and education of the immune system and in turn immune responses regulate the microbial composition in the intestine (Hooper et al., 2012; Kamada and Nunez, 2013). Therefore, a bi-directional relationship exists between the

gut microbiota and the host with a crucial role in intestinal homeostasis. Studies in germ-free mice have shown that the gut microbiota is required for GALT development (Bouskra et al., 2008; Kamada et al., 2013; Pabst et al., 2006; Sommer and Backhed, 2013). GALT comprises lymphoid structures such as Peyer's patches, crypt patches and isolated lymphoid follicles that have a critical role in eliciting an inflammatory or a tolerogenic immune response (Kamada and Nunez, 2013; Sommer and Backhed, 2013). Accumulating evidence indicates that the gut microbiota regulates the generation of specialized lymphocyte subsets of the CD4 lineage (Fritz et al., 2012; Geuking et al., 2011; Ivanov et al., 2009). Under steady-state conditions, Th17 cells are found in low numbers in the intestine in mice and humans. A specific type of clostridia-related bacteria, SFB, induces Th17 cell generation (Ivanov et al., 2009). Moreover, luminal adenosine triphosphate (ATP) produced only by commensal bacteria can drive Th17 cell development (Atarashi et al., 2008). Regulatory T cells (Tregs) also accumulate in the intestine for the maintenance of tolerance (Kamada and Nunez, 2013; Sommer and Backhed, 2013). The Gram-negative *Bacteroides fragilis* induces the differentiation of forkhead box P3 ($Foxp3^+$) Tregs and this is mediated by its outer membrane polysaccharide A (PSA) (Round et al., 2011; Round and Mazmanian, 2010), which acts by signaling through TLR2 on $CD4^+$ T cells (Round et al., 2011). PSA protected mice from experimental colitis induced by *Helicobacter hepaticus* by driving the production of IL-10 producing Tregs, which inhibited inflammatory cytokine expression (Mazmanian et al., 2008). Additionally, it has been shown that a mixture of 46 *Clostridium* spp. cluster IV and XIVa strains and altered Schaedler flora, which consists of eight defined commensals, induce the generation of $Foxp3^+$ Tregs in the lamina propria (Atarashi et al., 2013; Atarashi et al., 2011; Geuking et al., 2011). Oral inoculation with *Clostridium* species during early life resulted in attenuation of experimental colitis and allergy in mice (Atarashi et al., 2011). Clostridia induce differentiation and expansion of Tregs through the secretion of short chain fatty acids (SCFAs) and in particular butyrate (Furusawa et al., 2013; Smith et al., 2013). SCFAs, which are by-products of bacterial fermentation of dietary fibres, perform their role by enhancing histone H3 acetylation of the *Foxp3* locus (Arpaia et al., 2013; Furusawa et al., 2013; Smith et al., 2013). Collectively, these data suggest that the regulation of immune cells by commensal bacteria is necessary for immune homeostasis.

1.3.2 Bacterial metabolites

Gut microbiota process otherwise indigestible compounds (for example xylan, pectin and arabinose rich foods such as plant polysaccharides), host-derived glycans such as mucins and xenobiotics giving rise to nutrients and metabolites necessary for host metabolism (Brestoff and Artis, 2013; Nicholson et al., 2012). Metabolites include amino acids, bile acids, lipids, vitamins and SCFAs and are considered to be diet-dependent, as they are directly linked to diet and digestion. Other microbial products comprise lipopolysaccharide (LPS) and peptidoglycan that are diet-independent (Brestoff and Artis, 2013). Bacterial metabolites are either excreted in faeces or absorbed though the mucosa; therein they enter systemic circulation and can be further modified by host metabolic pathways (De Preter and Verbeke, 2013). In addition, host metabolites can be directed to the gut though the biliary pathway and be metabolized by resident bacteria. A well-known example of host metabolites modified further by microbiota is bile acids.

1.3.2.1 Bile acids

Bile acids are a family of cholesterol-derived amphipathic molecules that are synthesized in the liver and function to solubilise fat for the digestion and absorption of dietary lipids and lipid-soluble vitamins in the small intestine (Brestoff and Artis, 2013). Bile acids are subsequently secreted in the gallbladder, where they remain concentrated, and in response to diet-dependent signals they are emptied in the small intestine lumen to perform their detergent-like role (Brestoff and Artis, 2013; Ridlon et al., 2006). They are conjugated to glycine or taurine in the liver forming primary bile acids, and this process increases the efficiency of reabsorption in following cycles of digestion. Bile acids are reabsorbed via active transport in the terminal ileum and are directed back in the liver, where they are reconjugated with taurine or glycine and secreted in the bile, thus entering the enterohepatic circulation (Martinez-Augustin and Sanchez de Medina, 2008).

During enterohepatic circulation, the gut microbiota participates in bile acid metabolism by deconjugation of taurine or glycine residues and hydroxy group oxidation (Brestoff and Artis, 2013; Ridlon et al., 2006). Despite the fact that the transport of bile acids is highly efficient (~95%), a small amount of bile acids (400-800mg) is not reabsorbed at this stage and

becomes substrate for microbial metabolic reactions in the large intestine (Ridlon et al., 2006). Primary bile acids (for example cholic and chenodeoxycholic acids) derived from the liver are converted into secondary bile acids (for example deoxycholic, lithocholic and muricholic acids) (Brestoff and Artis, 2013). Secondary bile acids produced by dexydroxylation reactions are passively reabsorbed via portal circulation and returned to the liver (Ridlon et al., 2006). Secondary bile acids undergo further modifications such as sulfation and consequently are less effectively reabsorbed and ultimately excreted in the faeces (Brestoff and Artis, 2013; Ridlon et al., 2006). Highlighting the role of microbiota in bile acid synthesis, studies in germ-free mice and in mice treated with broad-spectrum antibiotics have demonstrated lower levels of secondary bile acids and higher amounts of conjugated bile acids compared with conventionally raised animals (Claus et al., 2011; Duboc et al., 2013; Swann et al., 2011). Therefore, luminal bile acid pool mostly comprises secondary bile acids and low levels of primary bile acids under physiological conditions.

In addition to their roles in digestion, bile acids can act as signalling molecules regulating metabolism and the immune system (Brestoff and Artis, 2013). Some bile acids can bind to their respective receptors, namely G-protein-coupled bile acid receptor 1 (GPBAR1) and nuclear receptor subfamily 1 group H member 4 (NR1H4; also known as farnesoid X receptor, FXR), which are highly expressed in monocytes and macrophages as well as some other immune cells, and exert anti-inflammatory actions involving the inhibition of NF-kB (Fiorucci et al., 2009; Vavassori et al., 2009; Wang et al., 2011). Interestingly, the activity of NR1H4 was decreased in the ileum of CD patients (Nijmeijer et al., 2011). However, the bile acid taurocholic acid promoted the growth of pathogenic bacteria in IL-10$^{-/-}$ mice (Devkota et al., 2012). Secondary bile acids may contribute to the establishment of certain bacterial niches in the gut, as they can be toxic to bacteria sensitive to these molecules (Ridlon et al., 2006).

1.3.2.2 SCFAs

Acetate, butyrate and propionate are the main SCFAs produced by the gut microbiota during fermentation of dietary fibres (Hamer et al., 2008). SCFAs are absorbed in the colonic mucosa and less than 5% remains in the lumen (Topping and Clifton, 2001). They are 1-6 carbons in length and are implicated in the regulation of host metabolism and immune system

homeostasis (Brestoff and Artis, 2013). Acetate enters the peripheral circulation and contributes to cholesterol synthesis, whereas propionate is utilized for glucose metabolism in the liver (Koropatkin et al., 2012). Acetate can also be used as a substrate for butyrate formation by some bacterial species (Duncan et al., 2004). Butyrate is the main energy source for colonocytes and also exerts an anti-inflammatory role by inhibiting NF-kB activation (Hamer et al., 2008). Furthermore, butyrate is implicated in the physiological function of the intestinal mucosa by stimulating mucus secretion and regulating vascular flow, motility and sodium and water absorption (Mortensen and Clausen, 1996). Therefore, depletion of butyrate or inhibition of its oxidation leads to loss of epithelial barrier integrity (Ahmad et al., 2000; De Preter et al., 2012). Impaired butyrate uptake and utilization are characteristics of UC (Le Gall et al., 2011).

SCFAs regulate the immune system in many ways including binding to G-protein coupled receptor (GPR) 43 (Maslowski et al., 2009; Smith et al., 2013), inhibition of histone deacetylases or regulation of autophagy (Arpaia et al., 2013; Donohoe et al., 2011; Furusawa et al., 2013). SCFAs promote neutrophil chemotaxis and effector functions including reactive oxygen species and phagocytic activity by activating the GPR43 or GPR41 signalling pathway (Sina et al., 2009; Trompette et al., 2014). Moreover, SCFAs induce Treg cell differentiation to maintain tolerance in the gut (Arpaia et al., 2013; Furusawa et al., 2013; Smith et al., 2013). Thus, depletion of GPR43 or addition of exogenous SCFAs ameliorated colonic inflammation in DSS-induced colitis (Brestoff and Artis, 2013; Maslowski et al., 2009).

SCFAs exert anti-inflammatory roles in additional immune cells such as macrophages, DCs, T cells and even IECs (Brestoff and Artis, 2013). SCFAs inhibited the expression of monocyte chemotactic protein-1 and IL-10 in LPS stimulated monocytes (Cox, 2009). In DCs, treatment with butyrate was associated with reduced expression of IL-12, IFN-γ and increased expression of IL-10 and IL-23 (Berndt et al., 2012; Liu et al., 2012). SCFAs may also directly regulate $CD4^+$ T cells by inhibiting their proliferation while promoting their apoptosis (Brestoff and Artis, 2013). Moreover, SCFAs can regulate the production of pro-inflammatory cytokines in IECs, as Caco-2 cells treated with butyrate and IL-1β produced less IL-8 than Caco-2 cells treated with IL-1β alone (Huang et al., 1997).

SCFAs and especially butyrate regulate the development and function of Tregs, contributing to the suppression of inflammation (Furusawa et al., 2013; Smith et al., 2013). The action of SCFAs is not limited to the maintenance of normal colonic function, but also they exert a systemic effect to peripheral tissues. Mice fed a high-fiber diet had increased levels of SCFAs and were protected from allergic airway inflammation, whereas a low-fiber diet decreased SCFAs and worsened inflammation in the lung (Trompette et al., 2014).

1.3.2.3 Amino acids, lipids and vitamins

Gut microbiota are important for the extraction, synthesis and absorption of amino acids. Amino acids serve as building blocks of proteins, substrates in metabolic pathways (such as glutamine in citric acid cycle), signalling molecules (such as glycine) and contribute to the function of the immune system (Brestoff and Artis, 2013). Fermentation of proteins produces various metabolites including phenolic compounds, branched chain fatty acids, S-containing compounds, amines and ammonia (De Preter et al., 2015). For instance, the protein fermentation compounds p-cresol and 3-methyl-1H-indole are products of microbial metabolism of the amino acids tyrosine and tryptophan respectively (De Preter et al., 2015). Mice fed an unrestricted tryptophan diet demonstrated increased indole-3-aldehyde levels, which upregulated IL-22 expression and triggered an immune response providing colonization resistance to the fungal pathogen *Candida albicans* and protection from DSS-induced colitis (Zelante et al., 2013).

Findings in germ-free mice demonstrate altered lipid metabolism and low fat mass, indicating the microbial effect on fatty acids (Claus et al., 2011). Colonization with human faeces partially rescued lipid abnormalities in gem-free mice (Martin et al., 2007). Furthermore, colonization of germ-free mice with *Bacteroides thetaiotaomicron* and/or *Bifidobacterium longum* increased arachidonic acid derived lipids and downstream metabolites such as prostaglandins (Dorrestein et al., 2014). Importantly, fatty acids can influence immune cell function (Brestoff and Artis, 2013). T_H1, T_H2 and T_H17 cells rely on glucose metabolism, whereas fatty acid oxidation promotes Tregs (Michalek et al., 2011; Pearce et al., 2009; Shi et al., 2011). Gut microbiota are also a source of essential vitamins especially of the B and K groups. There are 13 essential vitamins in humans: the lipid-soluble vitamins A, D, E, and K,

and the water-soluble vitamins B_1, B_2, B_3, B_5, B_6, B_7, B_9, B_{12} and C (Brestoff and Artis, 2013).

1.3.2.4 Antibacterial natural products

Gut microbiota due to its high amount compete for resources such as nutrient availability and niche colonization. Among the bacterial natural products secreted to mediate these competitive interactions are ribosomally synthesized post-translationally modified peptides (RiPPs), which as their name implies bear extensive post-translational modifications (Donia and Fischbach, 2015). The production of RiPPs is widespread in bacterial species and a great diversity is encountered including lantibiotics, bacteriocins, microcins, thiazole/oxazole-modified microcins (TOMMs) and thiopeptides (Arnison et al., 2013; Donia et al., 2014; Nunes-Alves, 2014). These classes of RiPPs exert antibacterial activity against a limited number of bacterial species closely related to the bacterial producer.

Lantibiotics and bacteriocins are the most common RiPPs found in human microbiota. Lantibiotics are short peptides (<40 amino acids) carrying chemical modifications, whereas bacteriocins are longer in length and are usually unmodified (Donia and Fischbach, 2015). Biosynthesis of lantibiotics is widely distributed in the phylum Firmicutes, displaying potent antibacterial activity against a narrow spectrum of other Gram-positive bacteria. For example, ruminococcin A is a lantibiotic produced by *Ruminococcus granus* and *Clostridium nexile* (Dabard et al., 2001). Commensal bacteria are not the sole producers of lantibiotics, as pathogens are able to produce these compounds to compete for resources and establish their colonization (Coburn and Gilmore, 2003).

Microcins carry a broad range of unusual modification such as conversion of cysteine and serine residues to thiazoles and oxazoles (microcin B17) (Arnison et al., 2013; Donia and Fischbach, 2015) and the majority of them is produced by Enterobacteriaceae and specifically *Escherichia coli* strains, however, they are also encountered in other commensal and pathogenic bacterial species (Duquesne et al., 2007; Micenkova et al., 2014). The mode of function of microcins is analogous to lantibiotics in Gram-positive bacteria with the difference that microcins target Gram-negative bacteria (Arnison et al., 2013). TOMMs resemble microcin B17, but they comprise natural products synthesized by both Gram-

positive and Gram-negative bacteria (Donia and Fischbach, 2015). Another group of RiPPs, thiopeptides, are characterized by thiazole rings and rings rich in nitrogen, and inhibit the growth selectively of closely related Gram-positive bacteria (Arnison et al., 2013).

1.3.2.5 Analysis of metabolites

Metabolites are in a dynamic equilibrium with host physiological processes and a disruption is translated in altered metabolic composition in bodily fluids. The serum, urine, faeces and tissue can provide adequate matrices for metabolite analysis. Faecal water extracts provide a more representative picture of bacterial metabolites produced by the gut microbiota (De Preter and Verbeke, 2013). Many microbial metabolites are absorbed and excreted in the urine directly or after modification by host metabolism, the so-called mammalian microbial co-metabolites, thereby urine samples reflect both host and bacterial metabolites. However, analysis of metabolites in serum and tissue samples entails mainly host metabolites confouding the influence of gut microbiota composition and its products.

Many analytical platforms are available for the identification and quantification of metabolites, which are low molecular weight molecules (Yoshida et al., 2012). Gas chromatography (GC), liquid chromatography (LC) and high/ultra performance liquid chromatography (H/UPLC) coupled to mass spectrometry (MS), and nuclear magnetic resonance spectroscopy (NMR) can be used for metabolite analysis (De Preter and Verbeke, 2013). Other analytical technologies include Fourier transform infrared spectroscopy (FT-IR) or capillary electrophoresis (CE) coupled to MS.

MS detects spatially or temporally separated ions in a mass analyzer (Ellis et al., 2007). Ions are initially formed in an ion source, accelerated by an electric field and then are separated according to mass/charge (m/z) ratio (Ellis et al., 2007; Harris, 2010). The identity of each metabolite is determined by its m/z value and the retention time of its extracted ion peak. The amount of a particular metabolite can be compared between samples using the ion peak height or area (Yoshida et al., 2012). GC-MS provides very high resolution in metabolite analysis, enabling the separation of structurally similar compounds (De Preter and Verbeke, 2013; Gromski et al., 2015). However, GC-MS requires chemical compounds to be volatile and thermally stable. So, samples may need to undergo preparation before the

chromatographic separation of their compounds. For example, chemical derivatization renders polar compounds more volatile (De Preter and Verbeke, 2013; Harris, 2010). LC-MS is used alternatively in case compounds are not volatile and cannot be derivatized (De Preter and Verbeke, 2013; Gromski et al., 2015). LC-MS can detect a broad range of polar and non-polar metabolites with specificity and sensitivity (De Preter and Verbeke, 2013; Gromski et al., 2015). CE-MS is a more sensitive technique detecting a broader range of polar metabolites than LC-MS (De Preter and Verbeke, 2013). In fact, CE-MS detects different groups of molecules such as charged, neutral, polar and hydrophobic (Yoshida et al., 2012). Matrix assisted laser desorption ionization (MALDI)-MS has been introduced for imaging MS in order to visualize the spatial distribution of metabolites in tissue samples, providing a powerful tool for examination of cellular processes (Yoshida et al., 2012).

Unlike the aforementioned methods, NMR does not require prior separation of compounds and instead minimal sample preparation is implemented (De Preter and Verbeke, 2013; Gromski et al., 2015; Lin et al., 2011; Nicholson et al., 1999). NMR is non-destructive to the samples, which can be reused for other analysis. It provides information about the molecular structure of the metabolites with high reproducibility, but it has lower sensitivity compared to MS based methods. NMR spectra are unique and well-resolved allowing discrimination of a wide range of metabolites (Ellis et al., 2007; Gromski et al., 2015). However, not all metabolites can be detected if their spectra overlap or if their concentration is low, as their signals may be overshadowed by high abundance metabolites (Lin et al., 2011). NMR detects nuclear transitions when samples are placed in a magnetic field and receive electromagnetic radiation usually at radiofrequency spectrum (Ellis et al., 2007). It basically exploits the magnetic properties of specific atomic nuclei and is based on the principle that every sample contains isotopes whose atomic nuclei possess magnetic spin greater than zero (for example nuclei with odd number of protons or neutrons such as hydrogen 1 (^{1}H) and carbon 13 (^{13}C)) in order the NMR effect to occur (Ellis et al., 2007; Harris, 2010).

FT-IR, a method of vibrational spectroscopy, measures the vibration of compounds that have been exposed to infrared beam (Ellis et al., 2007). Functional groups such as carbonyl groups absorb the infrared radiation and start vibrating in specific ways that correlate with specific chemical compounds. FT-IR offers rapid, high-throughput analysis of metabolites; however it is less specific and sensitive than MS methods (Ellis et al., 2007; Gromski et al., 2015). It is also non-destructive relative to the sample and requires a small amount of biofluid to generate

spectra, making it ideal in cases of scarce samples (Ellis et al., 2007). Some molecular vibrations may not be distinct with FT-IR spectroscopy, thus Raman spectroscopy is used instead (Harris, 2010). Whereas FT-IR measures absorption of energy, Raman spectroscopy measures energy exchange with electromagnetic radiation of a specific wavelength (Ellis et al., 2007). Raman spectroscopy enables direct measurement of metabolites from biological fluids as well as *in vivo*. Some of the drawbacks of this technique are its weak signal making it time-consuming and the background noise; however new methods have been devised to enhance signal intensity and specificity (Ellis et al., 2007). The applicability of these analytical technologies also varies depending on the concentration values of specific metabolites. NMR and FT-IR (mM sensitivity) have higher detection limits than GC-MS (< mM sensitivity) and LC-MS (nM sensitivity) (De Preter and Verbeke, 2013).

1.3.3 Factors affecting the gut microbiota

Genetic background, age, diet as well as lifestyle impact on gut microbiota composition. There is an essential degree of interpersonal variation in the makeup of gut microbial species and strains (2012; Palmer et al., 2007). Fluctuations in the abundance of microbial species are also observed in each individual over time, but within-subject variation is consistently lower than between-subject variation (2012). In fact, a follow-up study of microbial composition in healthy adults revealed long-term stability of the gut microbiota in a per individual basis with 60% of bacterial strains remaining stable over a course of five years (Faith et al., 2013). However, the stability of the gut microbiota was reduced when individuals were subjected to a calorie-restricted diet for some period of time, indicating that the gut microbiota responds to physiologic change (Faith et al., 2013). A shift in the representation of bacterial taxa may influence metabolic pathways or even the robustness of the system, resulting in elevated immune activation and differential susceptibility to infectious pathogens.

1.3.3.1 Genetic factors

The host genetic background and in particular genes determining the immunophenotype may be associated with microbial variability in the intestine (Abraham and Medzhitov, 2011; Spor et al., 2011). Genetic studies based on 16S ribosomal RNA (rRNA) analysis have shown that

there is a higher degree of similarity in the gut microbiota among identical twins compared to fraternal twins and unrelated paired controls (Stewart et al., 2005). In agreement with that, Dicksved et al have found that bacterial profiles of healthy twins resemble each other much more than those of twin pairs discordant for CD (Dicksved et al., 2008). A hallmark study based on intestinal bacteria transplantation between germ-free mice and zebrafish, underlined the influence of host factors on shaping gut microbiota (Rawls et al., 2006). More specifically, the final microbial profile resembled the original source in respect to bacterial lineages. However, the relative representation of bacterial lineage had changed to accommodate to the new host gut habitat (Rawls et al., 2006). Therefore, intraspecies microbial profiles are much more alike than interspecies profiles (Ley et al., 2008). Despite the significant findings of such studies, any extrapolation of conclusions should be cautious, as the confounding effect of environmental impact was not eliminated.

Friswell et al conducted a study, in which the influence of genetic and environmental factors was divided (Friswell et al., 2010). Relocation of young highly isogenic mice as well as uterine implantation of genetically distinct mouse embryos resulted in a dramatic change in intestinal microbiota composition. In fact, the bacterial shift reflected the different environment, implying little interference of the genetic factor on the formation of gut bacteria structure (Friswell et al., 2010). In the same study, comparison of the microbial profiles of isogenic mice reared in different research establishments revealed changes in the microbial census, which responded mainly to environmental influence (Friswell et al., 2010). Therefore, the genetic and environmental arm must both be considered regarding their role in the development, maintenance and stability of gut microbiota.

In line with the genetic implication on the gut microbial infrastructure, TLRs and Nod-like receptors (NLRs) may play a rather significant role (Abreu, 2010; Abreu et al., 2005). Polymorphisms in NOD2 have been associated with an altered microbiota (Biswas et al., 2012; Hansen et al., 2010). *Nod2*$^{-/-}$ mice were more susceptible to oral infection with *Listeria monocytogenes* and had increased bacterial load in the crypts at the terminal ileum (Petnicki-Ocwieja et al., 2009). The amount of *Bacteroides,* Firmicutes and *Bacillus* at the terminal ileum was increased in *Nod2*$^{-/-}$ mice (Petnicki-Ocwieja et al., 2009).

1.3.3.2 Diet

Diet has long been regarded as a major factor impacting on the microbiota composition as well as non-infectious intestinal disorders. From the 1960s- when Burkitt first recognized the protective effect of a low-fat, high fibre diet- until now, important insights have been gained on the role of nutrition on gut microbiota composition (Burkitt, 1973). Murine and human studies have demonstrated that a shift in the ratio of Firmicutes to Bacteroidetes is associated with obesity (Ley et al., 2006; Turnbaugh et al., 2009a; Turnbaugh et al., 2009b). Firmicutes are abundant in obese individuals, whereas Bacteroidetes predominate in healthy individuals. This misbalance is normalized upon introduction of a low-calorie diet (Ley et al., 2006). It would be intriguing to propose that an alteration of this ratio predisposes to obesity and thus it could be used as a biomarker, although it may merely reflect that the diet has shaped the microbiota composition favouring growth of particular bacterial groups.

A comparative study of the faecal microbiota between children following a western-type diet and children feeding mainly on plant fibres indicated that dietary components could shape gut microbial composition (De Filippo et al., 2010). In this study, Bacteroidetes and Actinobacteria were enriched in the case of a high-fibre diet, whereas Firmicutes and Proteobacteria were more abundant in the western diet. A subsequent study showed that not only long-term dietary patterns but also short-term dietary intake rapidly altered the structure and metabolic activity of the gut microbiota (David et al., 2013). Human subjects given an animal-based diet for a few days had increased representation of *Alistipes*, *Bilophila* and *Bacteroides* and decreased levels of Firmicutes including *Roseburia*, *Eubacterium rectale* and *Ruminococcus bromii* compared to subjects on a plant-based diet (David et al., 2013). Diversity in the diet can impact on the diversity of the microbiota in the gut as shown in a study on elderly individuals living in community or in residential care centres (Claesson et al., 2012).

The diversification of microbial communities under different nutrition regimes could have implications on the susceptibility to enteric infections and IBD. Enterobacteriaceae such as *Shigella* and *Escherichia* are found in low numbers in response to a high-fibre diet, whereas their relative representation is strongly increased in a western-type diet in humans (De Filippo et al., 2010). *Prevotella* is present at higher numbers in the faecal microbiota of individuals on a high-fibre diet than in those on a western diet (De Filippo et al., 2010;

Yatsunenko et al., 2012). Moreover, a high-fibre diet is characterized by the presence of bacterial species such as *Faecalibacterium prausnitzii*, which have potential anti-inflammatory roles (Burkitt, 1973; De Filippo et al., 2010; Sokol et al., 2008). Mice fed on a high fat/ high sugar (HF/HS) diet showed an overgrowth of Enterobacteriaceae and a decrease in Firmicutes and were more susceptible to DSS-induced colitis (Agus et al., 2016). Faecal transplantation from HF/HS treated mice to germ-free mice increased susceptibility to adherent invasive *Escherichia coli* (AIEC) infection (Agus et al., 2016). Furthermore, plant extracts or phytonutrients promoted the expansion of clostridia species and mucus secretion, protecting against enteric bacterial *Citrobacter rodentium* infection (Wlodarska et al., 2015b). Importantly, depletion of dietary fibres in ex-germ-free mice colonized with defined bacterial communities and in humanized mice resulted in an increase of mucus-degrading bacteria and a shift in microbiota localization towards the inner mucus layer, while the expression of inflammatory markers was heightened (Earle et al., 2015; Sonnenburg and Sonnenburg, 2014). All the above indicate a linkage between diet, gut microbiota and intestinal immunity. Recent evidence demonstrates that diet has a dominant effect in shaping gut microbiota composition, as consumption of a HF/ HS diet incurred reproducible microbial alterations in inbred and outbred mouse strains despite differences in host genotype (Carmody et al., 2015).

1.3.3.3 Antibiotics

Besides an altered diet, the modern lifestyle is characterized by widespread antibiotic administration for the clearance of pathogenic agents, while antibiotics are also commonly added to livestock and thus are potentially present in our diet. Antibiotic usage has a major role in the perturbation of mucosal microbial communities (Ubeda and Pamer, 2012). A human study of oral ciprofloxacin administration demonstrated a significant reduction of the diversity and richness of the mucosal microbiota (Dethlefsen et al., 2008). Ciprofloxacin influenced the abundance of approximately one third of bacterial taxa. Following ciprofloxacin treatment, it took four weeks until the microbial ecosystem was partly recovered, indicating that a disturbance in microbiota can persist following antibiotic treatment (Dethlefsen et al., 2008). Another human study using clindamycin showed a decrease in the population of *Bacteroides*, which did not return to its initial pre-treatment levels for up to two years (Jernberg et al., 2007). In a similar trend, drug administration in

mouse models induced long-lasting changes in the composition of the intestinal microbiota (Ubeda and Pamer, 2012). Even minimal antibiotic doses can have a dramatic effect on microbiota composition and its metabolic capabilities (Cho et al., 2012). Subtherapeutic antibiotic treatments caused a major shift in microbial composition resulting in a bloom of the Firmicutes phylum and its family Lachnospiraceae, and a reduction in Bacteroidetes accompanied by increased SCFAs production, indicating alterations in microbial metabolic status as well (Cho et al., 2012).

1.3.3.4 Early microbial exposure and age

Environmental challenge is greater in the newly acquired microbiota, following birth (Palmer et al., 2007). An infant passes from the relatively sterile environment in uterus to a bacteria dominated world. The intestine is a readily colonisable environment, being exposed to microbial sources from the birth canal, faecal material and breast milk (Friswell et al., 2010; Palmer et al., 2007). Aerobic bacteria, colonizing early in the gut, are replaced by anaerobic bacteria during infant gut microbiome development (Dominguez-Bello et al., 2010; Kostic et al., 2015; Palmer et al., 2007). Vaginally delivered infants were enriched in the genera *Bacteroides*, *Bifidobacterium*, *Parabacteroides* and *Escherichia/Shigella* (Backhed et al., 2015). In contrast, the gut microbiota from caesarean section born neonates resembled their mother's microbiome less and were derived mostly from microbial communities from the skin, mouth and external environment (Backhed et al., 2015). Infants born through caesarean section had decreased bacterial diversity and delayed colonisation of *Bacteroides* members, some of which were still missing at 4 and 12 months (Backhed et al., 2015; Jakobsson et al., 2014a). Therefore, delivery mode shapes gut microbiota composition in early life (Backhed et al., 2015; Dominguez-Bello et al., 2010). Breast-feeding has also an important role in shaping and controlling the maturation of gut microbial communities during the first year of life (Backhed et al., 2015; Penders et al., 2005). Breast-fed infants had increased levels of *Lactobacillus* and *Bifidobacterium* compared to bottle-fed babies (Backhed et al., 2015). Cessation of breast-feeding had profound effects on the microbiota composition, which adopted a mature configuration characterized by the dominance of *Bacteroides* and clostridia species including *Roseburia*, *Clostridium* and *Anaerostipes* (Backhed et al., 2015). A high rate of variability is observed in the first 12 months of life in man, after which point the

microbial composition starts to create a mature adult-like profile (Backhed et al., 2015; Mariat et al., 2009; Palmer et al., 2007).

The intestinal microbiota are relatively stable throughout adult life, however accruing evidence indicates that bacterial composition changes from adult life to old age (Mariat et al., 2009). Indeed, the Firmicutes/ Bacteroidetes ratio increases from birth to adulthood and subsides in elderly individuals. Altogether, these data suggest that the establishment of a complex microbial consortium in the gut is a progressive life-long process.

1.3.4 Surveys of the gut microbiota

Technological advances in the field of microbiome in the recent years have made it possible to delve into the structure and function of the intestinal ecosystem (Morgan and Huttenhower, 2014; Peterson et al., 2008). In combination with a decrease in costs of next-generation sequencing techniques, progress in experimental and analytical tools for studying the microbiome have spurred an uprecedented growth in studies reporting and analyzing sequence-based data from microbial communities (Kuczynski et al., 2012; Morgan and Huttenhower, 2014). Importantly, large-scale surveys of human microbial structure, function and diversity were able to be completed as part of the Human Microbiome Project and MetaHit Project (2012; Qin et al., 2010).

Analysis of bacterial populations used to be performed using culture-based techniques. Despite its usefulness in describing a single bacterium, there are many drawbacks in application of such an approach in the complex microbial gut environment. Not only is it time-consuming and expensive, but also it is unreliable, as culture of the most prevalent microorganisms can be inconsistent in different attempts using the exact same conditions (Manichanh et al., 2012). Most importantly, some bacterial species including anaerobic bacteria cannot be cultivated using conventional media (Kuczynski et al., 2012; Morgan and Huttenhower, 2014). Due to inherent problems, a comprehensive microbial analysis was lacking until the advent of culture-independent, nucleic-acid based methods (Eckburg et al., 2005).

Nucleic-acid methods are currently used to characterize complex microbial populations, giving information about the types of microbes (microbial census) and their metabolic

capacity (metagenomic analysis) (Eckburg et al., 2005; Muyzer et al., 1993; Peterson et al., 2008). Analysis of microbial profiles is based on the amplification of 16S rRNA gene followed by sequencing, or shotgun metagenomic sequencing (Kuczynski et al., 2012; Morgan and Huttenhower, 2014). 16S rRNA is an ideal candidate, as it has regions that are highly conserved among prokaryotic species as well as variable regions, which function as a molecular fingerprint of particular bacterial clades (Eckburg et al., 2005; Muyzer et al., 1993; Zoetendal et al., 2006). Thus, 16S sequencing enables taxonomic assignment of the members of a microbial community by matching 16S sequences to known databases (Kuczynski et al., 2012; Morgan and Huttenhower, 2014). Despite its widespread use due to its low cost, there are certain limitations introduced by copy number variation, and amplification and primer selection bias (Morgan and Huttenhower, 2014).

Shotgun metagenomics involves direct sequencing of the entire nucleotide pool without a prior amplification step and provides information not only about the census of bacteria, but also about their metabolic capabilities and functionality (Frank and Pace, 2008; Kuczynski et al., 2012; Morgan and Huttenhower, 2014). In addition, it identifies bacterial, archaeal, fungal and viral sequences allowing cross-kingdom analysis (Morgan and Huttenhower, 2014). Also, it provides high-resolution description about microbial composition reaching at species and strain level of microbial taxa (Morgan and Huttenhower, 2014). The magnitude of acquired data in accordance with decreasing cost will make metagenomic approaches more widely used in future studies.

Despite the fact that sequencing is the ultimate method for microbial characterization, a less in-depth approach such as denaturing gradient gel electrophoresis (DGGE) can also be applied. In DGGE, total genomic DNA is isolated and subjected to PCR amplification of the 16S rRNA gene (Muyzer et al., 1993). The PCR reaction gives products of approximately 200bp that are subsequently separated by electrophoresis on a denaturing gradient polyacrylamide gel. The different banding patterns on the gel are due to different base composition of the variable regions of the 16S rRNA gene (Muyzer et al., 1993). A comparison of the banding patterns of different samples provides an assessment on the similarity of different microbial communities. DGGE provides qualitative data in relation to microbial census; however, no quantitative data regarding relative abundance is provided.

The advent of massively parallel sequencing such as 454 pyrosequencing and Illumina sequencing has revolutionized the field of gut microbiota research, spawning the era of metagenomics (Morgan and Huttenhower, 2014; Peterson et al., 2008). However, high-throughput sequencing methods are limited by the fact that they cannot distinguish between active and inactive bacterial members or between resident and transient members. Another limitation is that they do not provide information about the individual contribution of bacteria to the community functional potential, masking bacterial-bacterial and host-bacterial interactions (Rolig et al., 2015). These limitations emphasize the need for defined model systems to associate microbial taxa with specific properties, so that the function of complex communities to be predicted based solely on their structure. Colonization of gnotobiotic animals with certain bacterial species in combination with computational modeling paves the way for deciphering the functional capacity of individual bacteria in the context of a multi-species community (Rolig et al., 2015), although the limitation of this is that the germ-free animal is not normal and may have deficiencies in, for example, immune function.

1.4 Inflammatory bowel disease (IBD)

IBD is a group of disorders characterized by chronic inflammation of the GI tract with intermittent relapsing-remitting phases (Saleh and Elson, 2011). IBD comprises CD and UC. The incidence and prevalence of IBD are constantly rising and there are approximately 240000 individuals suffering from IBD in the UK with a prevalence rate of 243 or 145 per 100000 individuals for UC and CD respectively (Molodecky et al., 2012; Mowat et al., 2011). CD may affect any part of the GI tract, but the most common regions include the terminal ileum and the colon (Khor et al., 2011; Saleh and Elson, 2011). It is characterized by transmural inflammation (affecting all layers of the bowel wall), discontinuous lesions, ulceration, and granuloma formation (Bouma and Strober, 2003; Strober et al., 2007). Complications of CD include strictures obstructing the bowel and fistulae connecting bowel segments to other parts of the gut, skin or internal organs. UC is distinguished by inflammation restricted to the colon and complications such as those present in CD are absent. In fact, inflammatory lesions are confined to the mucosa and submucosa with cryptitis and crypt abscesses in UC (Bouma and Strober, 2003; Khor et al., 2011; Strober et al., 2007).

Intestinal homeostasis is perturbed in chronic inflammatory conditions such as IBD (Fava and Danese, 2011; Sartor, 2006, 2008). Under steady state conditions, a stringent control of bacterial load as well as non-pathological immune responses at a basal level against intestinal microbiota is necessary for the maintenance of homeostasis (Biswas et al., 2012; Fava and Danese, 2011). Diverse mechanisms encompassing epithelial barrier function, innate immune recognition and immune system regulation cooperate to sustain homeostasis (Maloy and Powrie, 2011). However, dysregulation of these pathways in combination with an altered microbial community composition (dysbiosis) predisposes to pathologic inflammation found in IBD (Fava and Danese, 2011; Sartor, 2006, 2008). IBD is the result of an over-exuberant immune response generated against commensal microbiota in genetically susceptible individuals (Saleh and Elson, 2011).

Impairment of epithelial integrity leads to continuous exposure of bacterial antigens to the mucosal immune system precipitating aberrant inflammatory responses (Sartor, 2006). Increased intestinal permeability has been observed in IBD patients, and it predicts relapse in patients with inactive disease (Khor et al., 2011; Wyatt et al., 1993). Until recently, the role of gut permeability in the pathogenesis of IBD such as UC and CD has been a conundrum. A breach in the epithelial barrier could be the initiating event or the result of preceding molecular changes. Studies in *mdr1a*$^{-/-}$ mice that constitute a model of spontaneous colitis support the notion that a breach in the intestine follows early changes in the expression profile of IECs (Collett et al., 2008). Therefore, perturbation of the epithelial barrier seems to follow an already predetermined route that will lead to IBD.

The exact aetiology of IBD remains unclear, but many factors that contribute to its pathogenesis have been identified (Sartor, 2006, 2008; Xavier and Podolsky, 2007). IBD is a multifaceted disease where genetics, host immune system and environmental factors such as gut microbiota interact, leading to a chronic inflammatory condition (Sartor, 2006). Genome wide association studies (GWAS) and meta-analyses of these studies have identified 163 susceptibility loci implicated in IBD pathogenesis (Jostins et al., 2012; Khor et al., 2011; Saleh and Elson, 2011). Interestingly, 110 out of 163 loci are shared between CD and UC, suggesting common disease mechanisms (Jostins et al., 2012). Furthermore, these genetic studies illustrate that the pathways involved in IBD are crucial for intestinal homeostasis, for example barrier function, autophagy, endoplasmic reticulum (ER) stress, microbial defence, innate and adaptive immune regulation (Khor et al., 2011; Knights et al., 2013).

Polymorphisms in NOD2 were the first to be associated with CD susceptibility (Hugot et al., 2001; Ogura et al., 2001). CD patients carrying NOD2 genetic variants show reduced α-defensin expression suggesting that Paneth cell dysfunction predisposes to IBD (Wehkamp et al., 2005). Accumulating evidence indicated that additional IBD risk alleles including the transcriptional factor TCF4 and the autophagy related 16-like 1 (ATG16L1) were linked with impaired AMP production or secretion of AMP granules in Paneth cells respectively (Bevins and Salzman, 2011; Cadwell et al., 2008). Autophagy is a survival process based on the breakdown of cellular constituents in response to diverse stimuli (such as starvation, infection, ER stress, protein misfolding or aggregation) (Biswas et al., 2012). A meta-analysis of GWAS points to genes associated with host defence as susceptibility risk factors for IBD (Jostins et al., 2012). Patients carrying the ATG16L1 risk allele have increased ER stress in Paneth cells, emphasizing the key role of this specialized cell lineage in CD pathogenesis (Deuring et al., 2014).

Genetic studies highlight the importance of host-microbe interactions in driving IBD pathogenesis. However, GWAS account for 23% and 16% of the genetic risk in CD and UC respectively, indicating that additional factors play a rather decisive role (Khor et al., 2011). In addition, not all individuals carrying IBD-related risk alleles will develop disease, showing incomplete prevalence of IBD and a rather decisive effect of the environment. Furthermore, diversion of faecal stream induced remission and mucosa repair in the excluded gut segment, while treatment with enteric-coated antibiotics ameliorated inflammation in active UC (Casellas et al., 1998; D'Haens et al., 1998). Importantly, experimental colitis does not develop when mice are raised in germ-free conditions (Khor et al., 2011; Maloy and Powrie, 2011). All these findings indicate that gut microbiota is a crucial factor contributing to IBD pathogenesis.

1.4.1 Gut microbiota in IBD

Studies in IBD patients and experimental animal models have consistently shown alterations in gut microbiota composition, so-called dysbiosis (Huttenhower et al., 2014; Manichanh et al., 2012; Rajilic-Stojanovic et al., 2013). IBD is also associated with reduced bacterial diversity (Huttenhower et al., 2014; Manichanh et al., 2006; Ott et al., 2004; Wlodarska et al., 2015a), which has mainly been attributed to the Firmicutes phylum (Kang et al., 2010;

Manichanh et al., 2006). Gammaproteobacteria and specifically the Enterobacteriaceae family are significantly enriched in patients with IBD and mouse models, whereas several taxa within the Firmicutes phylum are markedly reduced (Berry et al., 2012; Huttenhower et al., 2014; Kostic et al., 2014; Lupp et al., 2007; Manichanh et al., 2012; Martinez-Medina et al., 2006). Members of the Bacteroidetes phylum have also been found reduced in some IBD studies (Frank et al., 2007; Lepage et al., 2011; Nones et al., 2009; Willing et al., 2010).

The levels of *Escherichia coli,* member of the Enterobacteriaceae family, are elevated in faecal samples of UC patients (Sokol et al., 2006) and some of the *Escherichia coli* isolates display properties of AIEC (Manichanh et al., 2012). AIEC strains are able to adhere and invade IECs and replicate within macrophages (Kostic et al., 2014; Manichanh et al., 2012). Transient infection with AIEC instigated chronic colitis in mice lacking the flagellin receptor TLR5. AIEC colonization induced alterations in microbial community composition including a reduction in bacterial diversity and enhancement of the pro-inflammatory potential of the microbiota (Chassaing et al., 2014). The relative abundance of *Escherichia coli* is also increased in CD and this enrichment is more profound in mucosal samples compared with faecal samples (Kostic et al., 2014; Manichanh et al., 2012). The numbers of AEIC are higher in mucosal biopsy specimens from CD patients and particularly in those patients with predominant ileal involvement compared with healthy individuals (Martinez-Medina et al., 2009). Furthermore, patients having the ATG16L1 CD-associated susceptibility gene show a significant increase in the presence of AIEC in their intestinal mucosa (Deuring et al., 2014). Therefore, in addition to changes observed in faeces, similar trends in microbial composition have been observed in mucus residing bacteria (Martinez-Medina et al., 2009; Mylonaki et al., 2005; Swidsinski et al., 2002; Swidsinski et al., 2005). There is a greater density of bacteria on the colonic mucus layer of IBD patients and this is positively correlated with the severity of disease (Swidsinski et al., 2002). Another bacterial group able to invade the intestinal epithelium, *Fusobacterium* species, has also been found at increased abundance in the inflamed colonic mucosa of active UC (Ohkusa et al., 2002; Strauss et al., 2011) and in faeces during UC remission (Rajilic-Stojanovic et al., 2013). Notably, mice given rectal enemas with supernatants of *Fusobacterium varium* culture developed colonic mucosa erosion (Ohkusa, 2003).

The Firmicutes groups Clostridia clades IV and XIVa and particularly *Faecalibacterium prausnitzii* are underrepresented in IBD (Frank et al., 2007; Rajilic-Stojanovic et al., 2013;

Sokol et al., 2009; Varela et al., 2013; Vermeiren et al., 2012; Willing et al., 2010). *Faecalibacterium prausnitzii*, a major member of *Firmicutes* phylum, induces the secretion of anti-inflammatory cytokines, while it reduces pro-inflammatory cytokine synthesis (Sokol et al., 2008). Oral administration of live *Faecalibacterium prausnitzii* bacteria or their metabolites ameliorated the severity of TNBS-induced colitis (Sokol et al., 2008). It has been found that *Faecalibacterium prausnitzii* is dramatically reduced in ileal CD patients, whereas *Escherichia coli* is concomitantly increased (Willing et al., 2009). Interestingly, recovery of *Faecalibacterium prausnitzii* numbers at levels similar to healthy controls after relapse is associated with the maintenance of remission in UC (Varela et al., 2013). The genus *Faecalibacterium* belongs to butyrate-producing bacteria, which are generally reduced in IBD (Frank et al., 2007; Kang et al., 2010; Machiels et al., 2014). Other SCFA producing bacteria that were reduced in IBD include *Roseburia, Phascolarctobacterium*, Leuconostocaceae and *Eubacterium rectale* (Machiels et al., 2014; Morgan et al., 2012; Rajilic-Stojanovic et al., 2013).

As studies in individuals with IBD have examined the gut microbiota composition after the establishment of disease and clinical management (Kang et al., 2010; Lepage et al., 2011; Manichanh et al., 2006; Martinez et al., 2008; Papa et al., 2012; Willing et al., 2010), a study on newly diagnosed CD before treatment has also been performed (Gevers et al., 2014). This is the largest IBD-related microbiome study to date and has shown that treatment-naïve CD patients at the time of disease diagnosis have an imbalance in their microbial community network taken from tissue biopsies, whereas faecal microbial communities from these CD patients resemble those of healthy individuals (Gevers et al., 2014). The dysbiosis pattern is characterized by increased abundance of Enterobacteriaceae, Pasteurellaceae, Veillonellaceae and Fusobacteriaceae, and decreased abundance of Erysipelotrichales, Bacteroidales and Clostridiales (Gevers et al., 2014).

IBD is associated with microbial dysfunction, which in combination with a misbalanced microbial composition describe dysbiosis (Davenport et al., 2014; Gevers et al., 2014; Ilott et al., 2016; Morgan et al., 2012; Rooks et al., 2014; Schwab et al., 2014). Microbial gene functions related to oxidative stress resistance and nutrient transport are increased, whereas gene pathways implicated in basic biosynthetic processes such as amino acid biosynthesis and carbohydrate metabolism are decreased (Gevers et al., 2014; Ilott et al., 2016; Morgan et al., 2012; Rooks et al., 2014). These alterations in microbial function indicate that bacteria

dominating in IBD can withstand oxidative stress and also have a reduced ability to synthesize their own nutrients, but rather utilize the readily available metabolites released during inflammation (Kostic et al., 2014). Furthermore, pathways contributing to bacterial pathogenesis including cell motility and cell signalling are enriched (Morgan et al., 2012; Rooks et al., 2014), indicating a shift towards a microbiome with enhanced inflammatory potential.

1.4.2 Metabolites in IBD

Studies in humans based on metabolite profiling were able to discriminate IBD from healthy state (Jansson et al., 2009; Le Gall et al., 2011; Marchesi et al., 2007), CD from UC (Marchesi et al., 2007) as well as active disease from remission phase (Dawiskiba et al., 2014). Marchesi et al were the first to differentiate IBD patients from healthy individuals and stratify IBD patients either to CD or UC using ^{1}H NMR on faecal water extracts (Marchesi et al., 2007). Bacterial metabolites such as SCFAs, methylamine and trimethylamine were absent in faecal water extracts from IBD patients, whereas amounts of amino acids were elevated suggesting malabsorption from the compromised epithelial barrier. Of note, methylamine and trimethylamine were derived from the degradation of dietary compounds such as choline and carnitine by gut bacteria (Marchesi et al., 2007). A reduction in SCFAs levels has been linked with an altered microbiota balance and presumably a decrease in butyrate-producing bacteria (Sartor, 2008). In fact, depletion of butyrate and acetate was associated with reduced abundance of *Clostridium coccoides* and *Clostridium leptum* in IBD patients (Marchesi et al., 2007).

Metabolic profiling can discriminate patients according to the area of the gut being involved. A study of identical twins, which were discordant or concordant for CD, separated CD patients from healthy controls and subdivided CD cases to ileal or colonic (Jansson et al., 2009). Differences were found in metabolites affecting the pathways of amino acid metabolism, bile acid and fatty acid biosynthesis, and arachidonic acid -a precursor of inflammatory prostaglandin and leukotriene- synthesis (Jansson et al., 2009). Amino acids were elevated in the faeces of CD patients, distinguishing them from healthy controls and from the different CD subtypes (Jansson et al., 2009). Interestingly, the levels of bile acids in the faeces were also increased, indicating underlying inflammation in CD individuals despite

the fact that these patients were in remission (Jansson et al., 2009). Inflammation hinders bile acid absorption in the ileum, increasing their amount in the lumen (Jansson et al., 2009).

In another study, ^{1}H NMR analysis of faecal extracts differentiated UC from healthy subjects (Le Gall et al., 2011). The increased levels of taurine and cadaverine, which is the product of lysine degradation by decarboxylation reactions, distinguished UC from healthy condition, but the levels of SCFAs and amino acids were decreased yet not significant. High levels of a cadaverine metabolite, 5-aminovaleric acid, have been reported in a metabolic study of urine samples in the IL-10$^{-/-}$ mouse model of colitis (Lin et al., 2010). Moreover, disease activity has been linked with the levels of specific metabolites in patients with IBD; the levels of medium chain fatty acid hexanoate were inversely correlated with disease activity in CD whereas the benzenoid compound styrene, a product of protein fermentation, was positively correlated with the status of UC (De Preter et al., 2015).

Several metabolomic studies based on urine samples were able to discriminate IBD patients from healthy controls (Dawiskiba et al., 2014; Schicho et al., 2012; Stephens et al., 2013; Williams et al., 2010; Williams et al., 2009). A common finding in these studies is the lower levels of hippurate in IBD, suggesting its potential use as a biomarker. Hippurate or N-benzoglycine is a metabolite generated by bacterial fermentation of dietary aromatic compounds such as polyphenols, purines or aromatic amino acids to benzoic acid, which is subsequently conjugated to glycine in the liver and the kidneys (De Preter and Verbeke, 2013; Lees et al., 2013). Low hippurate levels may reflect microbial dysbiosis, as hippurate levels correlate with the presence of clostridia in the gut (Li et al., 2008). Remarkably, urine metabolite analysis was not so effective in discriminating UC from CD cohorts with the exception of only one study (Williams et al., 2009). This discrepancy may be attributed to the fact that IBD is a complex disease with varying phenotypes.

Williams et al were the first to perform a ^{1}H NMR based profiling on serum samples and identified distinct metabolic phenotypes between CD and UC, as well as between IBD and healthy individuals (Williams et al., 2012). Differences were detected in lipid and choline metabolism and encompassed lipoproteins, choline, the acute phase N-acetylated glycoprotein and amino acids (Williams et al., 2012). Another ^{1}H NMR metabolic profiling study of serum also implicated decreased lipid metabolism and altered amino acids (Zhang et al., 2013). Metabolic changes can also affect energy metabolism as evidenced by the

decreased levels of tricarboxylic acid (TCA) cycle intermediates and of metabolites involved in energy metabolism in serum and plasma of active UC and CD (Schicho et al., 2012). These alterations in combination with the increased levels of creatine that has a role in energy supply indicate higher energy requirements and rapid utilization of metabolites that feed energy producing pathways during intestinal inflammation (Schicho et al., 2010; Schicho et al., 2012).

Studies analyzing serum/plasma and colonic mucosa biopsies indicate that IBD impacts on amino acid metabolism (Balasubramanian et al., 2009; Bjerrum et al., 2010; Ooi et al., 2011; Schicho et al., 2012; Sharma et al., 2010; Williams et al., 2012; Yoshida et al., 2012). The levels of several amino acids were decreased in colonic mucosa from IBD patients compared with controls. As the levels of amino acids were increased in faecal extracts (Jansson et al., 2009; Marchesi et al., 2007), this may result from defects in metabolite absorption in the intestinal mucosa due to inflammation. Another possible explanation for these observed low amino acid levels is the increased catabolic reactions in IBD to satisfy enhanced energy requirements for mucosa repair (De Preter and Verbeke, 2013; Williams et al., 2012). For example, the levels of the amino acid glutamine decline in IBD (Bjerrum et al., 2010; Shiomi et al., 2011). Glutamine is an important energy source for colonocytes in addition to butyrate and accounts for approximately 30% of their energy needs (De Preter and Verbeke, 2013). In a DSS model of colitis, glutamine was reduced in serum and colonic tissue and its supplementation ameliorated colitis symptoms (Shiomi et al., 2011).

1.5 Animal models in IBD research

Experimental animal models, although not fully representative of human IBD, provide a firm basis to understand the mechanisms of IBD pathology and decipher pathways important for intestinal homeostasis. There are four main types of IBD models: spontaneous, genetically engineered, chemical, and inducible models by immune cell transfer (adoptive transfer) (Saleh and Elson, 2011). In this thesis, we will focus on the *mdr1a*$^{-/-}$ model of colitis.

1.5.1 The *mdr1a*$^{-/-}$ model of colitis

The genetically engineered *mdr1a*$^{-/-}$ model is a spontaneous model of colitis with an intact immune system (Panwala et al., 1998; Schinkel et al., 1994). As a mucosal immune system defect is not involved in the onset of disease, the *mdr1a*$^{-/-}$ model enables the study of epithelial barrier function and its relationship with bacteria (Ho et al., 2003). The *Mdr1a* locus encodes for P-glycoprotein (Pgp) (alternatively called ABCb1 transporter) and it is involved in the cellular export of xenobiotics (Ho et al., 2003; Panwala et al., 1998), protecting the intestinal epithelium against foreign substances. Pgp is a member of the ABC transporters superfamily, which is the largest family of membrane transport proteins (Borst and Elferink, 2002). It is an ATP-driven pump that depends on ATP hydrolysis for the transport of small molecules across the plasma membrane (Borst and Elferink, 2002). The *Mdr1a* gene is expressed in IECs, a subset of lymphocytes, hematopoietic cells and cells at the blood-brain barrier (Chaudhary et al., 1992; Panwala et al., 1998; van de Ven et al., 2009).

A series of evidence outlines the importance of the gut microbiota in *mdr1a*$^{-/-}$ mice. *Mdr1a*$^{-/-}$ mice had an altered microbiota composition compared with wild-type mice (Nones et al., 2009). Administration of antibiotics in *mdr1a*$^{-/-}$ mice led to the reversal of the disease phenotype and resolution of active inflammation (Panwala et al., 1998). In addition, *Helicobacter bilis* infection accelerated gut inflammation, whereas *Helicobacter hepaticus* had the opposite effect (Maggio-Price et al., 2002). Notably, experimental infection with *Helicobacter* species induces inflammation in other models of colitis including the *IL-10*$^{-/-}$ model (Burich et al., 2001).

Mdr1a$^{-/-}$ mice are characterized by a slow progression of colitis when raised in specific pathogen-free conditions (Ho et al., 2003; Panwala et al., 1998). They develop intestinal inflammation from 12 weeks of age onwards and the severity of symptoms reaches a peak at 25 weeks of age, when disease is stabilized (Nones et al., 2009; Resta-Lenert et al., 2005). It is noteworthy that changes in gene expression levels precede the onset of mucosal inflammation (Collett et al., 2008). Genes implicated in bacterial recognition were up-regulated, whereas genes associated with the maintenance of mucosal tolerance including the antimicrobial Reg3γ were down-regulated (Collett et al., 2008). These data suggest that early

molecular events impact on the response to intestinal bacteria creating a hypersensitive environment, followed by gut permeability and inflammation.

Human genetic studies indicate that certain polymorphisms in the MDR1A gene are linked with increased susceptibility to UC (Annese et al., 2006; Ho et al., 2006). The polymorphism C3435T of the MDR1A gene is associated with diminished expression of Pgp in the intestine. Thus, individuals carrying the 3435TT genotype are at a higher risk for development of UC (Schwab et al., 2003). Furthermore, expression of MDR1A gene is reduced in the inflamed mucosa of individuals suffering from UC (Blokzijl et al., 2007; Ho et al., 2005).

Collectively, Pgp has a crucial role in the control of commensal microbiota-intestinal epithelium interaction. Building on the Nones' study (Nones et al., 2009), the current research is focused on finding a link between early changes in the expression of important genes and alterations in microbial composition in *mdr1a*$^{-/-}$ mice.

1.6 Hypotheses, aims and objectives of the thesis

Host-microbe interactions are disrupted in IBD, leading to an overacting pathologic immune response to the gut microbiota. Studies in IBD patients and experimental animal models have consistently shown alterations in gut microbiota composition, referred to as dysbiosis. However, these studies have mainly focused on microbial changes in faecal samples taken after the onset of inflammation and IBD establishment. There is a lack of knowledge regarding whether a shift in microbial composition is present before the emergence of IBD. The gut microbiota occupies diverse niches across the colon including both the gut lumen and the mucus layer overlying the intestinal epithelium. Thus, very little is known about mucus-resident microbiota in the routine microbiota analyses in IBD. We hypothesized that alterations in gut microbiota composition, most likely in the mucus, precede the initiation of intestinal inflammation. Except for composition, microbial gene expression and subsequently metabolic pathways are reportedly perturbed in established IBD. Gut microbiota performs functions that impact on host metabolism and in turn is influenced by host epithelial responses. Therefore, it is reasonable to suggest that microbial functions exhibit some dysregulation before the onset of IBD and this is also reflected in endogenous metabolic networks. Epithelial cells through the secretion of various mediators shape gut microbiota.

We hypothesized that AMPs may have a role in regulating microbiota composition in IBD.

The global aim of the project is to characterize the microbial communities found in faeces and mucus at stages preceding the onset of intestinal inflammation and at disease manifestation. Moreover, this thesis aims to identify potential metabolic changes that could strongly influence the host-microbiota crosstalk over the course of IBD development. Finally, the role of AMPs in controlling microbiota configuration is considered. To address the aforementioned aims, three key objectives have been set:

1. To investigate microbial structure in mucus and faeces at stages preceding and during inflammation onset.

This objective was addressed using the *mdr1a*$^{-/-}$ spontaneous model of colitis (Chapter 2), which has an intact immune system. Overall microbial community composition was characterized using DGGE and Illumina high-throughput next generation sequencing, followed by advanced computational analysis. The total bacterial load and abundance of specific bacterial groups were also quantified and compared with wild-type littermate controls. Microbiota localization was assessed using fluorescent in situ hybridization (FISH) and correlated with severity of colitis. Other additional parameters that were assessed include different histological parameters.

2. To investigate microbial functional potential and endogenous metabolism at pre- and post- inflammation time points.

This objective was addressed in Chapter 3, using results of microbiota composition as described in Chapter 2. The predicted microbial gene encoding potential in mucus and faeces was calculated using developed algorithms based on microbial species composition. Microbial gene families were annotated to gene functional pathways and analysed for differences in their constituents. Urinary metabolite analysis using LC-MS was performed to analyze metabolites of host and bacterial origin, followed by advanced computational analysis. Predictive models correlating changes in metabolite profiles with parameters such as genotype and disease were also built.

3. To investigate the role of AMPs in determining gut microbiota composition.

This objective was addressed in Chapter 4 by examining the expression patterns of AMPs

that are present in the colon, before and during colitis onset. The numbers of IgA-producing plasma cells were also assessed.

References

(2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207-214.

Abraham, C., and Medzhitov, R. (2011). Interactions between the host innate immune system and microbes in inflammatory bowel disease. Gastroenterology *140*, 1729-1737.

Abreu, M.T. (2010). Toll-like receptor signalling in the intestinal epithelium: how bacterial recognition shapes intestinal function. Nature reviews. Immunology *10*, 131-144.

Abreu, M.T., Fukata, M., and Arditi, M. (2005). TLR signaling in the gut in health and disease. J Immunol *174*, 4453-4460.

Agus, A., Denizot, J., Thevenot, J., Martinez-Medina, M., Massier, S., Sauvanet, P., Bernalier-Donadille, A., Denis, S., Hofman, P., Bonnet, R.*, et al.* (2016). Western diet induces a shift in microbiota composition enhancing susceptibility to Adherent-Invasive E. coli infection and intestinal inflammation. Scientific reports *6*, 19032.

Ahmad, M.S., Krishnan, S., Ramakrishna, B.S., Mathan, M., Pulimood, A.B., and Murthy, S.N. (2000). Butyrate and glucose metabolism by colonocytes in experimental colitis in mice. Gut *46*, 493-499.

Ahmed, S., Macfarlane, G.T., Fite, A., McBain, A.J., Gilbert, P., and Macfarlane, S. (2007). Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. Applied and environmental microbiology *73*, 7435-7442.

Annese, V., Valvano, M.R., Palmieri, O., Latiano, A., Bossa, F., and Andriulli, A. (2006). Multidrug resistance 1 gene in inflammatory bowel disease: a meta-analysis. World journal of gastroenterology : WJG *12*, 3636-3644.

Arnison, P.G., Bibb, M.J., Bierbaum, G., Bowers, A.A., Bugni, T.S., Bulaj, G., Camarero, J.A., Campopiano, D.J., Challis, G.L., Clardy, J.*, et al.* (2013). Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. Natural product reports *30*, 108-160.

Arpaia, N., Campbell, C., Fan, X., Dikiy, S., van der Veeken, J., deRoos, P., Liu, H., Cross, J.R., Pfeffer, K., Coffer, P.J., and Rudensky, A.Y. (2013). Metabolites produced by commensal bacteria promote peripheral regulatory T-cell generation. Nature *504*, 451-455.

Artis, D., Wang, M.L., Keilbaugh, S.A., He, W., Brenes, M., Swain, G.P., Knight, P.A., Donaldson, D.D., Lazar, M.A., Miller, H.R.*, et al.* (2004). RELMbeta/FIZZ2 is a goblet cell-specific immune-effector molecule in the gastrointestinal tract. Proceedings of the National Academy of Sciences of the United States of America *101*, 13596-13600.

Atarashi, K., Nishimura, J., Shima, T., Umesaki, Y., Yamamoto, M., Onoue, M., Yagita, H., Ishii, N., Evans, R., Honda, K., and Takeda, K. (2008). ATP drives lamina propria T(H)17 cell differentiation. Nature *455*, 808-812.

Atarashi, K., Tanoue, T., Oshima, K., Suda, W., Nagano, Y., Nishikawa, H., Fukuda, S., Saito, T., Narushima, S., Hase, K., *et al.* (2013). Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature *500*, 232-236.

Atarashi, K., Tanoue, T., Shima, T., Imaoka, A., Kuwahara, T., Momose, Y., Cheng, G., Yamasaki, S., Saito, T., Ohba, Y., *et al.* (2011). Induction of colonic regulatory T cells by indigenous Clostridium species. Science *331*, 337-341.

Atuma, C., Strugala, V., Allen, A., and Holm, L. (2001). The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. American journal of physiology. Gastrointestinal and liver physiology *280*, G922-929.

Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., and Gordon, J.I. (2005). Host-bacterial mutualism in the human intestine. Science *307*, 1915-1920.

Backhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., Li, Y., Xia, Y., Xie, H., Zhong, H., *et al.* (2015). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. Cell host & microbe *17*, 690-703.

Balasubramanian, K., Kumar, S., Singh, R.R., Sharma, U., Ahuja, V., Makharia, G.K., and Jagannathan, N.R. (2009). Metabolism of the colonic mucosa in patients with inflammatory bowel diseases: an in vitro proton magnetic resonance spectroscopy study. Magnetic resonance imaging *27*, 79-86.

Bals, R., and Wilson, J.M. (2003). Cathelicidins--a family of multifunctional antimicrobial peptides. Cellular and molecular life sciences : CMLS *60*, 711-720.

Baumler, A.J., Bergstrom, K.S.B., Morampudi, V., Chan, J.M., Bhinder, G., Lau, J., Yang, H., Ma, C., Huang, T., Ryz, N., *et al.* (2015). Goblet Cell Derived RELM-β Recruits CD4+ T Cells during Infectious Colitis to Promote Protective Intestinal Epithelial Cell Proliferation. PLoS pathogens *11*, e1005108.

Bergstrom, K.S., Kissoon-Singh, V., Gibson, D.L., Ma, C., Montero, M., Sham, H.P., Ryz, N., Huang, T., Velcich, A., Finlay, B.B., *et al.* (2010). Muc2 protects against lethal infectious colitis by disassociating pathogenic and commensal bacteria from the colonic mucosa. PLoS pathogens *6*, e1000902.

Berndt, B.E., Zhang, M., Owyang, S.Y., Cole, T.S., Wang, T.W., Luther, J., Veniaminova, N.A., Merchant, J.L., Chen, C.C., Huffnagle, G.B., and Kao, J.Y. (2012). Butyrate increases IL-23 production by stimulated dendritic cells. American journal of physiology. Gastrointestinal and liver physiology *303*, G1384-1392.

Berry, D., Schwab, C., Milinovich, G., Reichert, J., Ben Mahfoudh, K., Decker, T., Engel, M., Hai, B., Hainzl, E., Heider, S., *et al.* (2012). Phylotype-level 16S rRNA analysis reveals new bacterial indicators of health state in acute murine colitis. The ISME journal *6*, 2091-2106.

Bevins, C.L., and Salzman, N.H. (2011). Paneth cells, antimicrobial peptides and maintenance of intestinal homeostasis. Nature reviews. Microbiology *9*, 356-368.

Birchenough, G.M., Johansson, M.E., Gustafsson, J.K., Bergstrom, J.H., and Hansson, G.C. (2015). New developments in goblet cell mucus secretion and function. Mucosal immunology *8*, 712-719.

Biswas, A., Petnicki-Ocwieja, T., and Kobayashi, K.S. (2012). Nod2: a key regulator linking microbiota to intestinal mucosal immunity. J Mol Med (Berl) *90*, 15-24.

Bjerrum, J.T., Nielsen, O.H., Hao, F., Tang, H., Nicholson, J.K., Wang, Y., and Olsen, J. (2010). Metabonomics in ulcerative colitis: diagnostics, biomarker identification, and insight into the pathophysiology. Journal of proteome research *9*, 954-962.

Blokzijl, H., Vander Borght, S., Bok, L.I., Libbrecht, L., Geuken, M., van den Heuvel, F.A., Dijkstra, G., Roskams, T.A., Moshage, H., Jansen, P.L., and Faber, K.N. (2007). Decreased P-glycoprotein (P-gp/MDR1) expression in inflamed human intestinal epithelium is independent of PXR protein levels. Inflammatory bowel diseases *13*, 710-720.

Boirivant, M., Amendola, A., Butera, A., Sanchez, M., Xu, L., Marinaro, M., Kitani, A., Di Giacinto, C., Strober, W., and Fuss, I.J. (2008). A transient breach in the epithelial barrier leads to regulatory T-cell generation and resistance to experimental colitis. Gastroenterology *135*, 1612-1623 e1615.

Borst, P., and Elferink, R.O. (2002). Mammalian ABC transporters in health and disease. Annual review of biochemistry *71*, 537-592.

Bouma, G., and Strober, W. (2003). The immunological and genetic basis of inflammatory bowel disease. Nature reviews. Immunology *3*, 521-533.

Bouskra, D., Brezillon, C., Berard, M., Werts, C., Varona, R., Boneca, I.G., and Eberl, G. (2008). Lymphoid tissue genesis induced by commensals through NOD1 regulates intestinal homeostasis. Nature *456*, 507-510.

Bowcutt, R., Forman, R., Glymenaki, M., Carding, S.R., Else, K.J., and Cruickshank, S.M. (2014). Heterogeneity across the murine small and large intestine. World journal of gastroenterology : WJG *20*, 15216-15232.

Brandl, K., Plitas, G., Schnabl, B., DeMatteo, R.P., and Pamer, E.G. (2007). MyD88-mediated signals induce the bactericidal lectin RegIII gamma and protect mice against intestinal Listeria monocytogenes infection. The Journal of experimental medicine *204*, 1891-1900.

Brestoff, J.R., and Artis, D. (2013). Commensal bacteria at the interface of host metabolism and the immune system. Nature immunology *14*, 676-684.

Brown, E.M., Sadarangani, M., and Finlay, B.B. (2013). The role of the immune system in governing host-microbe interactions in the intestine. Nature immunology *14*, 660-667.

Burger-van Paassen, N., Loonen, L.M., Witte-Bouma, J., Korteland-van Male, A.M., de Bruijn, A.C., van der Sluis, M., Lu, P., Van Goudoever, J.B., Wells, J.M., Dekker, J.*, et al.* (2012). Mucin Muc2 deficiency and weaning influences the expression of the innate defense genes Reg3beta, Reg3gamma and angiogenin-4. PloS one *7*, e38798.

Burich, A., Hershberg, R., Waggie, K., Zeng, W., Brabb, T., Westrich, G., Viney, J.L., and Maggio-Price, L. (2001). Helicobacter-induced inflammatory bowel disease in IL-10- and T cell-deficient mice. American journal of physiology. Gastrointestinal and liver physiology *281*, G764-778.

Burkitt, D.P. (1973). Epidemiology of large bowel disease: the role of fibre. The Proceedings of the Nutrition Society *32*, 145-149.

Cadwell, K., Liu, J.Y., Brown, S.L., Miyoshi, H., Loh, J., Lennerz, J.K., Kishi, C., Kc, W., Carrero, J.A., Hunt, S.*, et al.* (2008). A key role for autophagy and the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells. Nature *456*, 259-263.

Cao, S., Su, X., Zeng, B., Yan, H., Huang, Y., Wang, E., Yun, H., Zhang, Y., Liu, F., Li, W.*, et al.* (2016). The Gut Epithelial Receptor LRRC19 Promotes the Recruitment of Immune Cells and Gut Inflammation. Cell Reports *14*, 695-707.

Carmody, R.N., Gerber, G.K., Luevano, J.M., Jr., Gatti, D.M., Somes, L., Svenson, K.L., and Turnbaugh, P.J. (2015). Diet dominates host genotype in shaping the murine gut microbiota. Cell host & microbe *17*, 72-84.

Casellas, F., Borruel, N., Papo, M., Guarner, F., Antolin, M., Videla, S., and Malagelada, J.R. (1998). Antiinflammatory effects of enterically coated amoxicillin-clavulanic acid in active ulcerative colitis. Inflammatory bowel diseases *4*, 1-5.

Cash, H.L., Whitham, C.V., Behrendt, C.L., and Hooper, L.V. (2006). Symbiotic bacteria direct expression of an intestinal bactericidal lectin. Science *313*, 1126-1130.

Cerf-Bensussan, N., and Gaboriau-Routhiau, V. (2010). The immune system and the gut microbiota: friends or foes? Nature reviews. Immunology *10*, 735-744.

Chassaing, B., Koren, O., Carvalho, F.A., Ley, R.E., and Gewirtz, A.T. (2014). AIEC pathobiont instigates chronic colitis in susceptible hosts by altering microbiota composition. Gut *63*, 1069-1080.

Chaudhary, P.M., Mechetner, E.B., and Roninson, I.B. (1992). Expression and activity of the multidrug resistance P-glycoprotein in human peripheral blood lymphocytes. Blood *80*, 2735-2739.

Cho, I., Yamanishi, S., Cox, L., Methe, B.A., Zavadil, J., Li, K., Gao, Z., Mahana, D., Raju, K., Teitler, I.*, et al.* (2012). Antibiotics in early life alter the murine colonic microbiome and adiposity. Nature *488*, 621-626.

Christa, L., Carnot, F., Simon, M.T., Levavasseur, F., Stinnakre, M.G., Lasserre, C., Thepot, D., Clement, B., Devinoy, E., and Brechot, C. (1996). HIP/PAP is an adhesive protein expressed in hepatocarcinoma, normal Paneth, and pancreatic cells. The American journal of physiology *271*, G993-1002.

Chu, H., Pazgier, M., Jung, G., Nuccio, S.P., Castillo, P.A., de Jong, M.F., Winter, M.G., Winter, S.E., Wehkamp, J., Shen, B.*, et al.* (2012). Human alpha-defensin 6 promotes mucosal innate immunity through self-assembled peptide nanonets. Science *337*, 477-481.

Claesson, M.J., Jeffery, I.B., Conde, S., Power, S.E., O'Connor, E.M., Cusack, S., Harris, H.M., Coakley, M., Lakshminarayanan, B., O'Sullivan, O.*, et al.* (2012). Gut microbiota composition correlates with diet and health in the elderly. Nature *488*, 178-184.

Claus, S.P., Ellero, S.L., Berger, B., Krause, L., Bruttin, A., Molina, J., Paris, A., Want, E.J., de Waziers, I., Cloarec, O.*, et al.* (2011). Colonization-induced host-gut microbial metabolic interaction. mBio *2*, e00271-00210.

Cobo, E.R., Kissoon-Singh, V., Moreau, F., and Chadee, K. (2015). Colonic MUC2 mucin regulates the expression and antimicrobial activity of beta-defensin 2. Mucosal immunology *8*, 1360-1372.

Coburn, P.S., and Gilmore, M.S. (2003). The *Enterococcus faecalis* cytolysin: a novel toxin active against eukaryotic and prokaryotic cells. Cellular microbiology *5*, 661-669.

Collett, A., Higgs, N.B., Gironella, M., Zeef, L.A., Hayes, A., Salmo, E., Haboubi, N., Iovanna, J.L., Carlson, G.L., and Warhurst, G. (2008). Early molecular and functional changes in colonic epithelium that precede increased gut permeability during colitis development in mdr1a(-/-) mice. Inflammatory bowel diseases *14*, 620-631.

Cox, M.A. (2009). Short-chain fatty acids act as antiinflammatory mediators by regulating prostaglandin E2and cytokines. World Journal of Gastroenterology *15*, 5549.

Crabtree, B., Holloway, D.E., Baker, M.D., Acharya, K.R., and Subramanian, V. (2007). Biological and structural features of murine angiogenin-4, an angiogenic protein. Biochemistry *46*, 2431-2443.

Crosnier, C., Stamataki, D., and Lewis, J. (2006). Organizing cell renewal in the intestine: stem cells, signals and combinatorial control. Nature reviews. Genetics *7*, 349-359.

D'Haens, G.R., Geboes, K., Peeters, M., Baert, F., Penninckx, F., and Rutgeerts, P. (1998). Early lesions of recurrent Crohn's disease caused by infusion of intestinal contents in excluded ileum. Gastroenterology *114*, 262-267.

Dabard, J., Bridonneau, C., Phillipe, C., Anglade, P., Molle, D., Nardi, M., Ladire, M., Girardin, H., Marcille, F., Gomez, A., and Fons, M. (2001). Ruminococcin A, a new lantibiotic produced by a Ruminococcus gnavus strain isolated from human feces. Applied and environmental microbiology *67*, 4111-4118.

Davenport, M., Poles, J., Leung, J.M., Wolff, M.J., Abidi, W.M., Ullman, T., Mayer, L., Cho, I., and Loke, P. (2014). Metabolic alterations to the mucosal microbiota in inflammatory bowel disease. Inflammatory bowel diseases *20*, 723-731.

David, L.A., Maurice, C.F., Carmody, R.N., Gootenberg, D.B., Button, J.E., Wolfe, B.E., Ling, A.V., Devlin, A.S., Varma, Y., Fischbach, M.A.*, et al.* (2013). Diet rapidly and reproducibly alters the human gut microbiome. Nature.

Dawiskiba, T., Deja, S., Mulak, A., Zabek, A., Jawien, E., Pawelka, D., Banasik, M., Mastalerz-Migas, A., Balcerzak, W., Kaliszewski, K.*, et al.* (2014). Serum and urine metabolomic fingerprinting in diagnostics of inflammatory bowel diseases. World journal of gastroenterology : WJG *20*, 163-174.

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J.B., Massart, S., Collini, S., Pieraccini, G., and Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proceedings of the National Academy of Sciences of the United States of America *107*, 14691-14696.

De Preter, V., Arijs, I., Windey, K., Vanhove, W., Vermeire, S., Schuit, F., Rutgeerts, P., and Verbeke, K. (2012). Impaired butyrate oxidation in ulcerative colitis is due to decreased butyrate uptake and a defect in the oxidation pathway. Inflammatory bowel diseases *18*, 1127-1136.

De Preter, V., Machiels, K., Joossens, M., Arijs, I., Matthys, C., Vermeire, S., Rutgeerts, P., and Verbeke, K. (2015). Faecal metabolite profiling identifies medium-chain fatty acids as discriminating compounds in IBD. Gut *64*, 447-458.

De Preter, V., and Verbeke, K. (2013). Metabolomics as a diagnostic tool in gastroenterology. World journal of gastrointestinal pharmacology and therapeutics *4*, 97-107.

Dessein, R., Gironella, M., Vignal, C., Peyrin-Biroulet, L., Sokol, H., Secher, T., Lacas-Gervais, S., Gratadoux, J.J., Lafont, F., Dagorn, J.C.*, et al.* (2009). Toll-like receptor 2 is critical for induction of Reg3 expression and intestinal clearance of Yersinia pseudotuberculosis. Gut *58*, 771-776.

Dethlefsen, L., Huse, S., Sogin, M.L., and Relman, D.A. (2008). The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PLoS biology *6*, e280.

Deuring, J.J., Fuhler, G.M., Konstantinov, S.R., Peppelenbosch, M.P., Kuipers, E.J., de Haar, C., and van der Woude, C.J. (2014). Genomic ATG16L1 risk allele-restricted Paneth cell ER stress in quiescent Crohn's disease. Gut *63*, 1081-1091.

Devkota, S., Wang, Y., Musch, M.W., Leone, V., Fehlner-Peach, H., Nadimpalli, A., Antonopoulos, D.A., Jabri, B., and Chang, E.B. (2012). Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in Il10-/- mice. Nature *487*, 104-108.

Dicksved, J., Halfvarson, J., Rosenquist, M., Jarnerot, G., Tysk, C., Apajalahti, J., Engstrand, L., and Jansson, J.K. (2008). Molecular analysis of the gut microbiota of identical twins with Crohn's disease. The ISME journal *2*, 716-727.

Dieckgraefe, B.K., Crimmins, D.L., Landt, V., Houchen, C., Anant, S., Porche-Sorbet, R., and Ladenson, J.H. (2002). Expression of the regenerating gene family in inflammatory bowel disease mucosa: Reg Ialpha upregulation, processing, and antiapoptotic activity. Journal of investigative medicine : the official publication of the American Federation for Clinical Research *50*, 421-434.

Diehl, G.E., Longman, R.S., Zhang, J.X., Breart, B., Galan, C., Cuesta, A., Schwab, S.R., and Littman, D.R. (2013). Microbiota restricts trafficking of bacteria to mesenteric lymph nodes by CX(3)CR1(hi) cells. Nature.

Dominguez-Bello, M.G., Costello, E.K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., and Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proceedings of the National Academy of Sciences of the United States of America *107*, 11971-11975.

Donia, M.S., Cimermancic, P., Schulze, C.J., Wieland Brown, L.C., Martin, J., Mitreva, M., Clardy, J., Linington, R.G., and Fischbach, M.A. (2014). A systematic analysis of biosynthetic gene clusters in the human microbiome reveals a common family of antibiotics. Cell *158*, 1402-1414.

Donia, M.S., and Fischbach, M.A. (2015). Human microbiota. Small molecules from the human microbiota. Science *349*, 1254766.

Donohoe, D.R., Garge, N., Zhang, X., Sun, W., O'Connell, T.M., Bunger, M.K., and Bultman, S.J. (2011). The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. Cell metabolism *13*, 517-526.

Dorrestein, P.C., Mazmanian, S.K., and Knight, R. (2014). Finding the missing links among metabolites, microbes, and the host. Immunity *40*, 824-832.

Duboc, H., Rajca, S., Rainteau, D., Benarous, D., Maubert, M.A., Quervain, E., Thomas, G., Barbu, V., Humbert, L., Despras, G., *et al.* (2013). Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. Gut *62*, 531-539.

Duncan, S.H., Holtrop, G., Lobley, G.E., Calder, A.G., Stewart, C.S., and Flint, H.J. (2004). Contribution of acetate to butyrate formation by human faecal bacteria. The British journal of nutrition *91*, 915-923.

Duquesne, S., Destoumieux-Garzon, D., Peduzzi, J., and Rebuffat, S. (2007). Microcins, gene-encoded antibacterial peptides from enterobacteria. Natural product reports *24*, 708-734.

Earle, K.A., Billings, G., Sigal, M., Lichtman, J.S., Hansson, G.C., Elias, J.E., Amieva, M.R., Huang, K.C., and Sonnenburg, J.L. (2015). Quantitative Imaging of Gut Microbiota Spatial Organization. Cell host & microbe *18*, 478-488.

Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E., and Relman, D.A. (2005). Diversity of the human intestinal microbial flora. Science *308*, 1635-1638.

Ellis, D.I., Dunn, W.B., Griffin, J.L., Allwood, J.W., and Goodacre, R. (2007). Metabolic fingerprinting as a diagnostic tool. Pharmacogenomics *8*, 1243-1266.

Ermund, A., Schutte, A., Johansson, M.E., Gustafsson, J.K., and Hansson, G.C. (2013). Studies of mucus in mouse stomach, small intestine, and colon. I. Gastrointestinal mucus layers have different properties depending on location as well as over the Peyer's patches. American journal of physiology. Gastrointestinal and liver physiology *305*, G341-347.

Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., Clemente, J.C., Knight, R., Heath, A.C., Leibel, R.L., *et al.* (2013). The long-term stability of the human gut microbiota. Science *341*, 1237439.

Fava, F., and Danese, S. (2011). Intestinal microbiota in inflammatory bowel disease: friend of foe? World journal of gastroenterology : WJG *17*, 557-566.

Fiorucci, S., Mencarelli, A., Palladino, G., and Cipriani, S. (2009). Bile-acid-activated receptors: targeting TGR5 and farnesoid-X-receptor in lipid and glucose disorders. Trends in pharmacological sciences *30*, 570-580.

Forman, R.A., deSchoolmeester, M.L., Hurst, R.J., Wright, S.H., Pemberton, A.D., and Else, K.J. (2012). The goblet cell is the cellular source of the anti-microbial angiogenin 4 in the large intestine post Trichuris muris infection. PloS one *7*, e42248.

Frank, D.N., and Pace, N.R. (2008). Gastrointestinal microbiology enters the metagenomics era. Current opinion in gastroenterology *24*, 4-10.

Frank, D.N., St Amand, A.L., Feldman, R.A., Boedeker, E.C., Harpaz, N., and Pace, N.R. (2007). Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proceedings of the National Academy of Sciences of the United States of America *104*, 13780-13785.

Friswell, M.K., Gika, H., Stratford, I.J., Theodoridis, G., Telfer, B., Wilson, I.D., and McBain, A.J. (2010). Site and strain-specific variation in gut microbiota profiles and metabolism in experimental mice. PloS one *5*, e8584.

Fritz, J.H., Rojas, O.L., Simard, N., McCarthy, D.D., Hapfelmeier, S., Rubino, S., Robertson, S.J., Larijani, M., Gosselin, J., Ivanov, II*, et al.* (2012). Acquisition of a multifunctional IgA+ plasma cell phenotype in the gut. Nature *481*, 199-203.

Fu, J., Wei, B., Wen, T., Johansson, M.E., Liu, X., Bradford, E., Thomsson, K.A., McGee, S., Mansour, L., Tong, M.*, et al.* (2011). Loss of intestinal core 1-derived O-glycans causes spontaneous colitis in mice. The Journal of clinical investigation *121*, 1657-1666.

Furusawa, Y., Obata, Y., Fukuda, S., Endo, T.A., Nakato, G., Takahashi, D., Nakanishi, Y., Uetake, C., Kato, K., Kato, T.*, et al.* (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature *504*, 446-450.

Gallo, R.L., and Hooper, L.V. (2012). Epithelial antimicrobial defence of the skin and intestine. Nature reviews. Immunology *12*, 503-516.

Ganz, T. (2003). Defensins: antimicrobial peptides of innate immunity. Nature reviews. Immunology *3*, 710-720.

Geuking, M.B., Cahenzli, J., Lawson, M.A., Ng, D.C., Slack, E., Hapfelmeier, S., McCoy, K.D., and Macpherson, A.J. (2011). Intestinal bacterial colonization induces mutualistic regulatory T cell responses. Immunity *34*, 794-806.

Gevers, D., Kugathasan, S., Denson, L.A., Vazquez-Baeza, Y., Van Treuren, W., Ren, B., Schwager, E., Knights, D., Song, S.J., Yassour, M.*, et al.* (2014). The treatment-naive microbiome in new-onset Crohn's disease. Cell host & microbe *15*, 382-392.

Gill, S.R., Pop, M., Deboy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., Gordon, J.I., Relman, D.A., Fraser-Liggett, C.M., and Nelson, K.E. (2006). Metagenomic analysis of the human distal gut microbiome. Science *312*, 1355-1359.

Gromski, P.S., Muhamadali, H., Ellis, D.I., Xu, Y., Correa, E., Turner, M.L., and Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis--a marriage of convenience or a shotgun wedding. Analytica chimica acta *879*, 10-23.

Gutzeit, C., Magri, G., and Cerutti, A. (2014). Intestinal IgA production and its role in host-microbe interaction. Immunological reviews *260*, 76-85.

Hamer, H.M., Jonkers, D., Venema, K., Vanhoutvin, S., Troost, F.J., and Brummer, R.J. (2008). Review article: the role of butyrate on colonic function. Alimentary pharmacology & therapeutics *27*, 104-119.

Hansen, J., Gulati, A., and Sartor, R.B. (2010). The role of mucosal immunity and host genetics in defining intestinal commensal bacteria. Current opinion in gastroenterology *26*, 564-571.

Harris, D.C. (2010). Quantitative chemical analysis, 8th edn (New York: W.H. Freeman and Co.).

He, W., Wang, M.L., Jiang, H.Q., Steppan, C.M., Shin, M.E., Thurnheer, M.C., Cebra, J.J., Lazar, M.A., and Wu, G.D. (2003). Bacterial colonization leads to the colonic secretion of RELMbeta/FIZZ2, a novel goblet cell-specific protein. Gastroenterology *125*, 1388-1397.

Henderson, P., van Limbergen, J.E., Schwarze, J., and Wilson, D.C. (2011). Function of the intestinal epithelium and its dysregulation in inflammatory bowel disease. Inflammatory bowel diseases *17*, 382-395.

Ho, G.T., Moodie, F.M., and Satsangi, J. (2003). Multidrug resistance 1 gene (P-glycoprotein 170): an important determinant in gastrointestinal disease? Gut *52*, 759-766.

Ho, G.T., Nimmo, E.R., Tenesa, A., Fennell, J., Drummond, H., Mowat, C., Arnott, I.D., and Satsangi, J. (2005). Allelic variations of the multidrug resistance gene determine susceptibility and disease behavior in ulcerative colitis. Gastroenterology *128*, 288-296.

Ho, G.T., Soranzo, N., Nimmo, E.R., Tenesa, A., Goldstein, D.B., and Satsangi, J. (2006). ABCB1/MDR1 gene determines susceptibility and phenotype in ulcerative colitis: discrimination of critical variants using a gene-wide haplotype tagging approach. Human molecular genetics *15*, 797-805.

Hogan, S.P., Seidu, L., Blanchard, C., Groschwitz, K., Mishra, A., Karow, M.L., Ahrens, R., Artis, D., Murphy, A.J., Valenzuela, D.M.*, et al.* (2006). Resistin-like molecule beta regulates innate colonic function: barrier integrity and inflammation susceptibility. The Journal of allergy and clinical immunology *118*, 257-268.

Holmen Larsson, J.M., Thomsson, K.A., Rodriguez-Pineiro, A.M., Karlsson, H., and Hansson, G.C. (2013). Studies of mucus in mouse stomach, small intestine, and colon. III. Gastrointestinal Muc5ac and Muc2 mucin O-glycan patterns reveal a regiospecific distribution. American journal of physiology. Gastrointestinal and liver physiology *305*, G357-363.

Hooper, L.V. (2004). Bacterial contributions to mammalian gut development. Trends in microbiology *12*, 129-134.

Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the microbiota and the immune system. Science *336*, 1268-1273.

Hooper, L.V., Stappenbeck, T.S., Hong, C.V., and Gordon, J.I. (2003). Angiogenins: a new class of microbicidal proteins involved in innate immunity. Nature immunology *4*, 269-273.

Huang, N., Katz, J.P., Martin, D.R., and Wu, G.D. (1997). Inhibition of IL-8 gene expression in Caco-2 cells by compounds which induce histone hyperacetylation. Cytokine *9*, 27-36.

Hugot, J.P., Chamaillard, M., Zouali, H., Lesage, S., Cezard, J.P., Belaiche, J., Almer, S., Tysk, C., O'Morain, C.A., Gassull, M.*, et al.* (2001). Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. Nature *411*, 599-603.

Huttenhower, C., Kostic, A.D., and Xavier, R.J. (2014). Inflammatory bowel disease as a model for translating the microbiome. Immunity *40*, 843-854.

Iliev, I.D., Mileti, E., Matteoli, G., Chieppa, M., and Rescigno, M. (2009). Intestinal epithelial cells promote colitis-protective regulatory T-cell differentiation through dendritic cell conditioning. Mucosal immunology *2*, 340-350.

Ilott, N.E., Bollrath, J., Danne, C., Schiering, C., Shale, M., Adelmann, K., Krausgruber, T., Heger, A., Sims, D., and Powrie, F. (2016). Defining the microbial transcriptional response to colitis through integrated host and microbiome profiling. The ISME journal.

Ivanov, II, Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., Wei, D., Goldfarb, K.C., Santee, C.A., Lynch, S.V.*, et al.* (2009). Induction of intestinal Th17 cells by segmented filamentous bacteria. Cell *139*, 485-498.

Jakobsson, H.E., Abrahamsson, T.R., Jenmalm, M.C., Harris, K., Quince, C., Jernberg, C., Bjorksten, B., Engstrand, L., and Andersson, A.F. (2014a). Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section. Gut *63*, 559-566.

Jakobsson, H.E., Rodriguez-Pineiro, A.M., Schutte, A., Ermund, A., Boysen, P., Bemark, M., Sommer, F., Backhed, F., Hansson, G.C., and Johansson, M.E. (2014b). The composition of the gut microbiota shapes the colon mucus barrier. EMBO reports.

Jansson, J., Willing, B., Lucio, M., Fekete, A., Dicksved, J., Halfvarson, J., Tysk, C., and Schmitt-Kopplin, P. (2009). Metabolomics reveals metabolic biomarkers of Crohn's disease. PloS one *4*, e6386.

Jernberg, C., Lofmark, S., Edlund, C., and Jansson, J.K. (2007). Long-term ecological impacts of antibiotic administration on the human intestinal microbiota. The ISME journal *1*, 56-66.

Johansson, M.E., Ambort, D., Pelaseyed, T., Schutte, A., Gustafsson, J.K., Ermund, A., Subramani, D.B., Holmen-Larsson, J.M., Thomsson, K.A., Bergstrom, J.H.*, et al.* (2011a). Composition and functional role of the mucus layers in the intestine. Cellular and molecular life sciences : CMLS *68*, 3635-3641.

Johansson, M.E., Gustafsson, J.K., Holmen-Larsson, J., Jabbar, K.S., Xia, L., Xu, H., Ghishan, F.K., Carvalho, F.A., Gewirtz, A.T., Sjovall, H., and Hansson, G.C. (2014). Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. Gut *63*, 281-291.

Johansson, M.E., Gustafsson, J.K., Sjoberg, K.E., Petersson, J., Holm, L., Sjovall, H., and Hansson, G.C. (2010). Bacteria penetrate the inner mucus layer before inflammation in the dextran sulfate colitis model. PloS one *5*, e12238.

Johansson, M.E., and Hansson, G.C. (2011). Microbiology. Keeping bacteria at a distance. Science *334*, 182-183.

Johansson, M.E., Jakobsson, H.E., Holmen-Larsson, J., Schutte, A., Ermund, A., Rodriguez-Pineiro, A.M., Arike, L., Wising, C., Svensson, F., Backhed, F., and Hansson, G.C. (2015). Normalization of Host Intestinal Mucus Layers Requires Long-Term Microbial Colonization. Cell host & microbe *18*, 582-592.

Johansson, M.E., Larsson, J.M., and Hansson, G.C. (2011b). The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. Proceedings of the National Academy of Sciences of the United States of America *108 Suppl 1*, 4659-4665.

Johansson, M.E., Phillipson, M., Petersson, J., Velcich, A., Holm, L., and Hansson, G.C. (2008). The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. Proceedings of the National Academy of Sciences of the United States of America *105*, 15064-15069.

Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., Sharma, Y., Anderson, C.A.*, et al.* (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature *491*, 119-124.

Juge, N. (2012). Microbial adhesins to gastrointestinal mucus. Trends in microbiology *20*, 30-39.

Kagan, B.L., Selsted, M.E., Ganz, T., and Lehrer, R.I. (1990). Antimicrobial defensin peptides form voltage-dependent ion-permeable channels in planar lipid bilayer membranes. Proceedings of the National Academy of Sciences of the United States of America *87*, 210-214.

Kamada, N., and Nunez, G. (2013). Role of the gut microbiota in the development and function of lymphoid cells. J Immunol *190*, 1389-1395.

Kamada, N., Seo, S.U., Chen, G.Y., and Nunez, G. (2013). Role of the gut microbiota in immunity and inflammatory disease. Nature reviews. Immunology *13*, 321-335.

Kang, S., Denman, S.E., Morrison, M., Yu, Z., Dore, J., Leclerc, M., and McSweeney, C.S. (2010). Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray. Inflammatory bowel diseases *16*, 2034-2042.

Khor, B., Gardet, A., and Xavier, R.J. (2011). Genetics and pathogenesis of inflammatory bowel disease. Nature *474*, 307-317.

Knights, D., Lassen, K.G., and Xavier, R.J. (2013). Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. Gut *62*, 1505-1510.

Kobayashi, K.S., Chamaillard, M., Ogura, Y., Henegariu, O., Inohara, N., Nunez, G., and Flavell, R.A. (2005). Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. Science *307*, 731-734.

Koropatkin, N.M., Cameron, E.A., and Martens, E.C. (2012). How glycan metabolism shapes the human gut microbiota. Nature reviews. Microbiology *10*, 323-335.

Kostic, A.D., Gevers, D., Siljander, H., Vatanen, T., Hyotylainen, T., Hamalainen, A.M., Peet, A., Tillmann, V., Poho, P., Mattila, I.*, et al.* (2015). The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. Cell host & microbe *17*, 260-273.

Kostic, A.D., Xavier, R.J., and Gevers, D. (2014). The microbiome in inflammatory bowel disease: current status and the future ahead. Gastroenterology *146*, 1489-1499.

Krimi, R.B., Kotelevets, L., Dubuquoy, L., Plaisancie, P., Walker, F., Lehy, T., Desreumaux, P., Van Seuningen, I., Chastre, E., Forgue-Lafitte, M.E., and Marie, J.C. (2008). Resistin-like molecule beta regulates intestinal mucous secretion and curtails TNBS-induced colitis in mice. Inflammatory bowel diseases *14*, 931-941.

Kuczynski, J., Lauber, C.L., Walters, W.A., Parfrey, L.W., Clemente, J.C., Gevers, D., and Knight, R. (2012). Experimental and analytical tools for studying the human microbiome. Nature reviews. Genetics *13*, 47-58.

Lathrop, S.K., Bloom, S.M., Rao, S.M., Nutsch, K., Lio, C.W., Santacruz, N., Peterson, D.A., Stappenbeck, T.S., and Hsieh, C.S. (2011). Peripheral education of the immune system by colonic commensal microbiota. Nature *478*, 250-254.

Le Gall, G., Noor, S.O., Ridgway, K., Scovell, L., Jamieson, C., Johnson, I.T., Colquhoun, I.J., Kemsley, E.K., and Narbad, A. (2011). Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. Journal of proteome research *10*, 4208-4218.

Lees, H.J., Swann, J.R., Wilson, I.D., Nicholson, J.K., and Holmes, E. (2013). Hippurate: the natural history of a mammalian-microbial cometabolite. Journal of proteome research *12*, 1527-1546.

Lehotzky, R.E., Partch, C.L., Mukherjee, S., Cash, H.L., Goldman, W.E., Gardner, K.H., and Hooper, L.V. (2010). Molecular basis for peptidoglycan recognition by a bactericidal lectin. Proceedings of the National Academy of Sciences of the United States of America *107*, 7722-7727.

Lepage, P., Hasler, R., Spehlmann, M.E., Rehman, A., Zvirbliene, A., Begun, A., Ott, S., Kupcinskas, L., Dore, J., Raedler, A., and Schreiber, S. (2011). Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. Gastroenterology *141*, 227-236.

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., and Gordon, J.I. (2008). Evolution of mammals and their gut microbes. Science *320*, 1647-1651.

Ley, R.E., Turnbaugh, P.J., Klein, S., and Gordon, J.I. (2006). Microbial ecology: human gut microbes associated with obesity. Nature *444*, 1022-1023.

Li, M., Wang, B., Zhang, M., Rantalainen, M., Wang, S., Zhou, H., Zhang, Y., Shen, J., Pang, X., Zhang, M.*, et al.* (2008). Symbiotic gut microbes modulate human metabolic phenotypes. Proceedings of the National Academy of Sciences of the United States of America *105*, 2117-2122.

Lin, H.M., Barnett, M.P., Roy, N.C., Joyce, N.I., Zhu, S., Armstrong, K., Helsby, N.A., Ferguson, L.R., and Rowan, D.D. (2010). Metabolomic analysis identifies inflammatory and noninflammatory metabolic effects of genetic modification in a mouse model of Crohn's disease. Journal of proteome research *9*, 1965-1975.

Lin, H.M., Helsby, N.A., Rowan, D.D., and Ferguson, L.R. (2011). Using metabolomic analysis to understand inflammatory bowel diseases. Inflammatory bowel diseases *17*, 1021-1029.

Liu, L., Li, L., Min, J., Wang, J., Wu, H., Zeng, Y., Chen, S., and Chu, Z. (2012). Butyrate interferes with the differentiation and function of human monocyte-derived dendritic cells. Cellular immunology *277*, 66-73.

Loonen, L.M., Stolte, E.H., Jaklofsky, M.T., Meijerink, M., Dekker, J., van Baarlen, P., and Wells, J.M. (2013). REG3gamma-deficient mice have altered mucus distribution and increased mucosal inflammatory responses to the microbiota and enteric pathogens in the ileum. Mucosal immunology.

Lupp, C., Robertson, M.L., Wickham, M.E., Sekirov, I., Champion, O.L., Gaynor, E.C., and Finlay, B.B. (2007). Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. Cell host & microbe *2*, 119-129.

Mabbott, N.A., Donaldson, D.S., Ohno, H., Williams, I.R., and Mahajan, A. (2013). Microfold (M) cells: important immunosurveillance posts in the intestinal epithelium. Mucosal immunology *6*, 666-677.

Machiels, K., Joossens, M., Sabino, J., De Preter, V., Arijs, I., Eeckhaut, V., Ballet, V., Claes, K., Van Immerseel, F., Verbeke, K.*, et al.* (2014). A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. Gut *63*, 1275-1283.

Macpherson, A.J., and Uhr, T. (2004). Induction of protective IgA by intestinal dendritic cells carrying commensal bacteria. Science *303*, 1662-1665.

Madsen, K.L., Malfair, D., Gray, D., Doyle, J.S., Jewell, L.D., and Fedorak, R.N. (1999). Interleukin-10 gene-deficient mice develop a primary intestinal permeability defect in response to enteric microflora. Inflammatory bowel diseases *5*, 262-270.

Maggio-Price, L., Shows, D., Waggie, K., Burich, A., Zeng, W., Escobar, S., Morrissey, P., and Viney, J.L. (2002). Helicobacter bilis infection accelerates and H. hepaticus infection delays the development of colitis in multiple drug resistance-deficient (mdr1a-/-) mice. The American journal of pathology *160*, 739-751.

Maloy, K.J., and Powrie, F. (2011). Intestinal homeostasis and its breakdown in inflammatory bowel disease. Nature *474*, 298-306.

Manichanh, C., Borruel, N., Casellas, F., and Guarner, F. (2012). The gut microbiota in IBD. Nature reviews. Gastroenterology & hepatology *9*, 599-608.

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., Nalin, R., Jarrin, C., Chardon, P., Marteau, P.*, et al.* (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut *55*, 205-211.

Marchesi, J.R., Holmes, E., Khan, F., Kochhar, S., Scanlan, P., Shanahan, F., Wilson, I.D., and Wang, Y. (2007). Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. Journal of proteome research *6*, 546-551.

Mariat, D., Firmesse, O., Levenez, F., Guimaraes, V., Sokol, H., Dore, J., Corthier, G., and Furet, J.P. (2009). The Firmicutes/Bacteroidetes ratio of the human microbiota changes with age. BMC microbiology *9*, 123.

Martin, F.P., Dumas, M.E., Wang, Y., Legido-Quigley, C., Yap, I.K., Tang, H., Zirah, S., Murphy, G.M., Cloarec, O., Lindon, J.C.*, et al.* (2007). A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model. Molecular systems biology *3*, 112.

Martinez, C., Antolin, M., Santos, J., Torrejon, A., Casellas, F., Borruel, N., Guarner, F., and Malagelada, J.R. (2008). Unstable composition of the fecal microbiota in ulcerative colitis during clinical remission. The American journal of gastroenterology *103*, 643-648.

Martinez-Augustin, O., and Sanchez de Medina, F. (2008). Intestinal bile acid physiology and pathophysiology. World journal of gastroenterology : WJG *14*, 5630-5640.

Martinez-Medina, M., Aldeguer, X., Gonzalez-Huix, F., Acero, D., and Garcia-Gil, L.J. (2006). Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. Inflammatory bowel diseases *12*, 1136-1145.

Martinez-Medina, M., Aldeguer, X., Lopez-Siles, M., Gonzalez-Huix, F., Lopez-Oliu, C., Dahbi, G., Blanco, J.E., Blanco, J., Garcia-Gil, L.J., and Darfeuille-Michaud, A. (2009). Molecular diversity of Escherichia coli in the human gut: new ecological evidence supporting the role of adherent-invasive E. coli (AIEC) in Crohn's disease. Inflammatory bowel diseases *15*, 872-882.

Maslowski, K.M., Vieira, A.T., Ng, A., Kranich, J., Sierro, F., Yu, D., Schilter, H.C., Rolph, M.S., Mackay, F., Artis, D.*, et al.* (2009). Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. Nature *461*, 1282-1286.

Mazmanian, S.K., Round, J.L., and Kasper, D.L. (2008). A microbial symbiosis factor prevents intestinal inflammatory disease. Nature *453*, 620-625.

McAuley, J.L., Linden, S.K., Png, C.W., King, R.M., Pennington, H.L., Gendler, S.J., Florin, T.H., Hill, G.R., Korolik, V., and McGuckin, M.A. (2007). MUC1 cell surface mucin is a critical element of the mucosal barrier to infection. The Journal of clinical investigation *117*, 2313-2324.

McDole, J.R., Wheeler, L.W., McDonald, K.G., Wang, B., Konjufca, V., Knoop, K.A., Newberry, R.D., and Miller, M.J. (2012). Goblet cells deliver luminal antigen to CD103+ dendritic cells in the small intestine. Nature *483*, 345-349.

McGuckin, M.A., Linden, S.K., Sutton, P., and Florin, T.H. (2011). Mucin dynamics and enteric pathogens. Nature reviews. Microbiology *9*, 265-278.

McVay, L.D., Keilbaugh, S.A., Wong, T.M., Kierstein, S., Shin, M.E., Lehrke, M., Lefterova, M.I., Shifflett, D.E., Barnes, S.L., Cominelli, F.*, et al.* (2006). Absence of

bacterially induced RELMbeta reduces injury in the dextran sodium sulfate model of colitis. The Journal of clinical investigation *116*, 2914-2923.

Meyer-Hoffert, U., Hornef, M.W., Henriques-Normark, B., Axelsson, L.G., Midtvedt, T., Putsep, K., and Andersson, M. (2008). Secreted enteric antimicrobial activity localises to the mucus surface layer. Gut *57*, 764-771.

Micenkova, L., Staudova, B., Bosak, J., Mikalova, L., Littnerova, S., Vrba, M., Sevcikova, A., Woznicova, V., and Smajs, D. (2014). Bacteriocin-encoding genes and ExPEC virulence determinants are associated in human fecal Escherichia coli strains. BMC microbiology *14*, 109.

Michalek, R.D., Gerriets, V.A., Jacobs, S.R., Macintyre, A.N., MacIver, N.J., Mason, E.F., Sullivan, S.A., Nichols, A.G., and Rathmell, J.C. (2011). Cutting edge: distinct glycolytic and lipid oxidative metabolic programs are essential for effector and regulatory CD4+ T cell subsets. J Immunol *186*, 3299-3303.

Miki, T., Holst, O., and Hardt, W.D. (2012). The bactericidal activity of the C-type lectin RegIIIbeta against Gram-negative bacteria involves binding to lipid A. The Journal of biological chemistry *287*, 34844-34855.

Mizoguchi, E., Xavier, R.J., Reinecker, H.C., Uchino, H., Bhan, A.K., Podolsky, D.K., and Mizoguchi, A. (2003). Colonic epithelial functional phenotype varies with type and phase of experimental colitis. Gastroenterology *125*, 148-161.

Molodecky, N.A., Soon, I.S., Rabi, D.M., Ghali, W.A., Ferris, M., Chernoff, G., Benchimol, E.I., Panaccione, R., Ghosh, S., Barkema, H.W., and Kaplan, G.G. (2012). Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology *142*, 46-54 e42; quiz e30.

Mora, J.R., Iwata, M., Eksteen, B., Song, S.Y., Junt, T., Senman, B., Otipoby, K.L., Yokota, A., Takeuchi, H., Ricciardi-Castagnoli, P.*, et al.* (2006). Generation of gut-homing IgA-secreting B cells by intestinal dendritic cells. Science *314*, 1157-1160.

Mora, J.R., and von Andrian, U.H. (2008). Differentiation and homing of IgA-secreting cells. Mucosal immunology *1*, 96-109.

Morampudi, V., Dalwadi, U., Bhinder, G., Sham, H.P., Gill, S.K., Chan, J., Bergstrom, K.S., Huang, T., Ma, C., Jacobson, K.*, et al.* (2016). The goblet cell-derived mediator RELM-beta drives spontaneous colitis in Muc2-deficient mice by promoting commensal microbial dysbiosis. Mucosal immunology.

Morgan, X.C., and Huttenhower, C. (2014). Meta'omic analytic techniques for studying the intestinal microbiome. Gastroenterology *146*, 1437-1448 e1431.

Morgan, X.C., Tickle, T.L., Sokol, H., Gevers, D., Devaney, K.L., Ward, D.V., Reyes, J.A., Shah, S.A., LeLeiko, N., Snapper, S.B.*, et al.* (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome biology *13*, R79.

Mortensen, P.B., and Clausen, M.R. (1996). Short-chain fatty acids in the human colon: relation to gastrointestinal health and disease. Scandinavian journal of gastroenterology. Supplement *216*, 132-148.

Mowat, C., Cole, A., Windsor, A., Ahmad, T., Arnott, I., Driscoll, R., Mitton, S., Orchard, T., Rutter, M., Younge, L.*, et al.* (2011). Guidelines for the management of inflammatory bowel disease in adults. Gut *60*, 571-607.

Mukherjee, S., Zheng, H., Derebe, M.G., Callenberg, K.M., Partch, C.L., Rollins, D., Propheter, D.C., Rizo, J., Grabe, M., Jiang, Q.X., and Hooper, L.V. (2014). Antibacterial membrane attack by a pore-forming intestinal C-type lectin. Nature *505*, 103-107.

Muyzer, G., de Waal, E.C., and Uitterlinden, A.G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Applied and environmental microbiology *59*, 695-700.

Mylonaki, M., Rayment, N.B., Rampton, D.S., Hudspith, B.N., and Brostoff, J. (2005). Molecular characterization of rectal mucosa-associated bacterial flora in inflammatory bowel disease. Inflammatory bowel diseases *11*, 481-487.

Nair, M.G., Guild, K.J., Du, Y., Zaph, C., Yancopoulos, G.D., Valenzuela, D.M., Murphy, A., Stevens, S., Karow, M., and Artis, D. (2008). Goblet cell-derived resistin-like molecule beta augments CD4+ T cell production of IFN-gamma and infection-induced intestinal inflammation. J Immunol *181*, 4709-4715.

Nicholson, J.K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., and Pettersson, S. (2012). Host-gut microbiota metabolic interactions. Science *336*, 1262-1267.

Nicholson, J.K., Lindon, J.C., and Holmes, E. (1999). 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica; the fate of foreign compounds in biological systems *29*, 1181-1189.

Nielsen, D.S., Moller, P.L., Rosenfeldt, V., Paerregaard, A., Michaelsen, K.F., and Jakobsen, M. (2003). Case study of the distribution of mucosa-associated Bifidobacterium species, Lactobacillus species, and other lactic acid bacteria in the human colon. Applied and environmental microbiology *69*, 7545-7548.

Niess, J.H., Brand, S., Gu, X., Landsman, L., Jung, S., McCormick, B.A., Vyas, J.M., Boes, M., Ploegh, H.L., Fox, J.G.*, et al.* (2005). CX3CR1-mediated dendritic cell access to the intestinal lumen and bacterial clearance. Science *307*, 254-258.

Nijmeijer, R.M., Gadaleta, R.M., van Mil, S.W., van Bodegraven, A.A., Crusius, J.B., Dijkstra, G., Hommes, D.W., de Jong, D.J., Stokkers, P.C., Verspaget, H.W.*, et al.* (2011). Farnesoid X receptor (FXR) activation and FXR genetic variation in inflammatory bowel disease. PloS one *6*, e23745.

Nones, K., Knoch, B., Dommels, Y.E., Paturi, G., Butts, C., McNabb, W.C., and Roy, N.C. (2009). Multidrug resistance gene deficient (mdr1a-/-) mice have an altered caecal microbiota that precedes the onset of intestinal inflammation. Journal of applied microbiology *107*, 557-566.

Noor, Z., Burgess, S.L., Watanabe, K., and Petri, W.A., Jr. (2016). Interleukin-25 Mediated Induction of Angiogenin-4 Is Interleukin-13 Dependent. PloS one *11*, e0153572.

Nunes-Alves, C. (2014). Microbiome: Commensally sourced antibiotics. Nature reviews. Microbiology *12*, 726.

O'Neil, D.A., Porter, E.M., Elewaut, D., Anderson, G.M., Eckmann, L., Ganz, T., and Kagnoff, M.F. (1999). Expression and regulation of the human beta-defensins hBD-1 and hBD-2 in intestinal epithelium. J Immunol *163*, 6718-6724.

Ogawa, H., Fukushima, K., Naito, H., Funayama, Y., Unno, M., Takahashi, K., Kitayama, T., Matsuno, S., Ohtani, H., Takasawa, S.*, et al.* (2003). Increased expression of HIP/PAP and regenerating gene III in human inflammatory bowel disease and a murine bacterial reconstitution model. Inflammatory bowel diseases *9*, 162-170.

Ogura, Y., Bonen, D.K., Inohara, N., Nicolae, D.L., Chen, F.F., Ramos, R., Britton, H., Moran, T., Karaliuskas, R., Duerr, R.H.*, et al.* (2001). A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature *411*, 603-606.

Ohkusa, T. (2003). Induction of experimental ulcerative colitis by Fusobacterium varium isolated from colonic mucosa of patients with ulcerative colitis. Gut *52*, 79-83.

Ohkusa, T., Sato, N., Ogihara, T., Morita, K., Ogawa, M., and Okayasu, I. (2002). Fusobacterium varium localized in the colonic mucosa of patients with ulcerative colitis stimulates species-specific antibody. Journal of gastroenterology and hepatology *17*, 849-853.

Ooi, M., Nishiumi, S., Yoshie, T., Shiomi, Y., Kohashi, M., Fukunaga, K., Nakamura, S., Matsumoto, T., Hatano, N., Shinohara, M.*, et al.* (2011). GC/MS-based profiling of amino acids and TCA cycle-related molecules in ulcerative colitis. Inflammation research : official journal of the European Histamine Research Society ... [et al.] *60*, 831-840.

Ott, S.J., Musfeldt, M., Wenderoth, D.F., Hampe, J., Brant, O., Folsch, U.R., Timmis, K.N., and Schreiber, S. (2004). Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. Gut *53*, 685-693.

Pabst, O., Herbrand, H., Friedrichsen, M., Velaga, S., Dorsch, M., Berhardt, G., Worbs, T., Macpherson, A.J., and Forster, R. (2006). Adaptation of solitary intestinal lymphoid tissue in response to microbiota and chemokine receptor CCR7 signaling. J Immunol *177*, 6824-6832.

Palmer, C., Bik, E.M., DiGiulio, D.B., Relman, D.A., and Brown, P.O. (2007). Development of the human infant intestinal microbiota. PLoS biology *5*, e177.

Panwala, C.M., Jones, J.C., and Viney, J.L. (1998). A novel model of inflammatory bowel disease: mice deficient for the multiple drug resistance gene, mdr1a, spontaneously develop colitis. J Immunol *161*, 5733-5744.

Papa, E., Docktor, M., Smillie, C., Weber, S., Preheim, S.P., Gevers, D., Giannoukos, G., Ciulla, D., Tabbaa, D., Ingram, J.*, et al.* (2012). Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. PloS one *7*, e39242.

Pearce, E.L., Walsh, M.C., Cejas, P.J., Harms, G.M., Shen, H., Wang, L.S., Jones, R.G., and Choi, Y. (2009). Enhancing CD8 T-cell memory by modulating fatty acid metabolism. Nature *460*, 103-107.

Pelaseyed, T., Bergstrom, J.H., Gustafsson, J.K., Ermund, A., Birchenough, G.M., Schutte, A., van der Post, S., Svensson, F., Rodriguez-Pineiro, A.M., Nystrom, E.E.*, et al.* (2014). The mucus and mucins of the goblet cells and enterocytes provide the first defense line of the gastrointestinal tract and interact with the immune system. Immunological reviews *260*, 8-20.

Penders, J., Vink, C., Driessen, C., London, N., Thijs, C., and Stobberingh, E.E. (2005). Quantification of Bifidobacterium spp., Escherichia coli and Clostridium difficile in faecal samples of breast-fed and formula-fed infants by real-time PCR. FEMS microbiology letters *243*, 141-147.

Peterson, D.A., Frank, D.N., Pace, N.R., and Gordon, J.I. (2008). Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. Cell host & microbe *3*, 417-427.

Peterson, D.A., McNulty, N.P., Guruge, J.L., and Gordon, J.I. (2007). IgA response to symbiotic bacteria as a mediator of gut homeostasis. Cell host & microbe *2*, 328-339.

Peterson, L.W., and Artis, D. (2014). Intestinal epithelial cells: regulators of barrier function and immune homeostasis. Nature reviews. Immunology *14*, 141-153.

Petnicki-Ocwieja, T., Hrncir, T., Liu, Y.J., Biswas, A., Hudcovic, T., Tlaskalova-Hogenova, H., and Kobayashi, K.S. (2009). Nod2 is required for the regulation of commensal microbiota in the intestine. Proceedings of the National Academy of Sciences of the United States of America *106*, 15813-15818.

Putsep, K., Axelsson, L.G., Boman, A., Midtvedt, T., Normark, S., Boman, H.G., and Andersson, M. (2000). Germ-free and colonized mice generate the same products from enteric prodefensins. The Journal of biological chemistry *275*, 40478-40482.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T.*, et al.* (2010). A human gut microbial gene catalogue established by metagenomic sequencing. Nature *464*, 59-65.

Rajilic-Stojanovic, M., Shanahan, F., Guarner, F., and de Vos, W.M. (2013). Phylogenetic analysis of dysbiosis in ulcerative colitis during remission. Inflammatory bowel diseases *19*, 481-488.

Ramanan, D., Tang, M.S., Bowcutt, R., Loke, P., and Cadwell, K. (2014). Bacterial sensor Nod2 prevents inflammation of the small intestine by restricting the expansion of the commensal Bacteroides vulgatus. Immunity *41*, 311-324.

Rawls, J.F., Mahowald, M.A., Ley, R.E., and Gordon, J.I. (2006). Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. Cell *127*, 423-433.

Rescigno, M. (2011). The intestinal epithelial barrier in the control of homeostasis and immunity. Trends in immunology *32*, 256-264.

Rescigno, M., Urbano, M., Valzasina, B., Francolini, M., Rotta, G., Bonasio, R., Granucci, F., Kraehenbuhl, J.P., and Ricciardi-Castagnoli, P. (2001). Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. Nature immunology *2*, 361-367.

Resta-Lenert, S., Smitham, J., and Barrett, K.E. (2005). Epithelial dysfunction associated with the development of colitis in conventionally housed mdr1a-/- mice. American journal of physiology. Gastrointestinal and liver physiology *289*, G153-162.

Ridlon, J.M., Kang, D.J., and Hylemon, P.B. (2006). Bile salt biotransformations by human intestinal bacteria. Journal of lipid research *47*, 241-259.

Rodriguez-Pineiro, A.M., Bergstrom, J.H., Ermund, A., Gustafsson, J.K., Schutte, A., Johansson, M.E., and Hansson, G.C. (2013). Studies of mucus in mouse stomach, small intestine, and colon. II. Gastrointestinal mucus proteome reveals Muc2 and Muc5ac accompanied by a set of core proteins. American journal of physiology. Gastrointestinal and liver physiology *305*, G348-356.

Rolig, A.S., Parthasarathy, R., Burns, A.R., Bohannan, B.J., and Guillemin, K. (2015). Individual Members of the Microbiota Disproportionately Modulate Host Innate Immune Responses. Cell host & microbe *18*, 613-620.

Rooks, M.G., Veiga, P., Wardwell-Scott, L.H., Tickle, T., Segata, N., Michaud, M., Gallini, C.A., Beal, C., van Hylckama-Vlieg, J.E., Ballal, S.A., *et al.* (2014). Gut microbiome composition and function in experimental colitis during active disease and treatment-induced remission. The ISME journal *8*, 1403-1417.

Rossini, V., Zhurina, D., Radulovic, K., Manta, C., Walther, P., Riedel, C.U., and Niess, J.H. (2014). CX3CR1(+) cells facilitate the activation of CD4 T cells in the colonic lamina propria during antigen-driven colitis. Mucosal immunology *7*, 533-548.

Round, J.L., Lee, S.M., Li, J., Tran, G., Jabri, B., Chatila, T.A., and Mazmanian, S.K. (2011). The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. Science *332*, 974-977.

Round, J.L., and Mazmanian, S.K. (2010). Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. Proceedings of the National Academy of Sciences of the United States of America *107*, 12204-12209.

Saleh, M., and Elson, C.O. (2011). Experimental inflammatory bowel disease: insights into the host-microbiota dialog. Immunity *34*, 293-302.

Sanos, S.L., Bui, V.L., Mortha, A., Oberle, K., Heners, C., Johner, C., and Diefenbach, A. (2009). RORgammat and commensal microflora are required for the differentiation of mucosal interleukin 22-producing NKp46+ cells. Nature immunology *10*, 83-91.

Sartor, R.B. (2006). Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. Nature clinical practice. Gastroenterology & hepatology *3*, 390-407.

Sartor, R.B. (2008). Microbial influences in inflammatory bowel diseases. Gastroenterology *134*, 577-594.

Sass, V., Schneider, T., Wilmes, M., Korner, C., Tossi, A., Novikova, N., Shamova, O., and Sahl, H.G. (2010). Human -Defensin 3 Inhibits Cell Wall Biosynthesis in Staphylococci. Infection and immunity *78*, 2793-2800.

Schicho, R., Nazyrova, A., Shaykhutdinov, R., Duggan, G., Vogel, H.J., and Storr, M. (2010). Quantitative metabolomic profiling of serum and urine in DSS-induced ulcerative colitis of mice by (1)H NMR spectroscopy. Journal of proteome research *9*, 6265-6273.

Schicho, R., Shaykhutdinov, R., Ngo, J., Nazyrova, A., Schneider, C., Panaccione, R., Kaplan, G.G., Vogel, H.J., and Storr, M. (2012). Quantitative metabolomic profiling of serum, plasma, and urine by (1)H NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. Journal of proteome research *11*, 3344-3357.

Schinkel, A.H., Smit, J.J., van Tellingen, O., Beijnen, J.H., Wagenaar, E., van Deemter, L., Mol, C.A., van der Valk, M.A., Robanus-Maandag, E.C., te Riele, H.P., and et al. (1994). Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. Cell *77*, 491-502.

Schwab, C., Berry, D., Rauch, I., Rennisch, I., Ramesmayer, J., Hainzl, E., Heider, S., Decker, T., Kenner, L., Muller, M.*, et al.* (2014). Longitudinal study of murine microbiota activity and interactions with the host during acute inflammation and recovery. The ISME journal *8*, 1101-1114.

Schwab, M., Schaeffeler, E., Marx, C., Fromm, M.F., Kaskas, B., Metzler, J., Stange, E., Herfarth, H., Schoelmerich, J., Gregor, M.*, et al.* (2003). Association between the C3435T MDR1 gene polymorphism and susceptibility for ulcerative colitis. Gastroenterology *124*, 26-33.

Shanahan, M.T., Carroll, I.M., Grossniklaus, E., White, A., von Furstenberg, R.J., Barner, R., Fodor, A.A., Henning, S.J., Sartor, R.B., and Gulati, A.S. (2014). Mouse Paneth cell antimicrobial function is independent of Nod2. Gut *63*, 903-910.

Sharma, U., Singh, R.R., Ahuja, V., Makharia, G.K., and Jagannathan, N.R. (2010). Similarity in the metabolic profile in macroscopically involved and un-involved colonic mucosa in patients with inflammatory bowel disease: an in vitro proton ((1)H) MR spectroscopy study. Magnetic resonance imaging *28*, 1022-1029.

Shi, L.Z., Wang, R., Huang, G., Vogel, P., Neale, G., Green, D.R., and Chi, H. (2011). HIF1alpha-dependent glycolytic pathway orchestrates a metabolic checkpoint for the differentiation of TH17 and Treg cells. The Journal of experimental medicine *208*, 1367-1376.

Shiomi, Y., Nishiumi, S., Ooi, M., Hatano, N., Shinohara, M., Yoshie, T., Kondo, Y., Furumatsu, K., Shiomi, H., Kutsumi, H.*, et al.* (2011). GCMS-based metabolomic study in mice with colitis induced by dextran sulfate sodium. Inflammatory bowel diseases *17*, 2261-2274.

Sina, C., Gavrilova, O., Forster, M., Till, A., Derer, S., Hildebrand, F., Raabe, B., Chalaris, A., Scheller, J., Rehmann, A.*, et al.* (2009). G protein-coupled receptor 43 is essential for neutrophil recruitment during intestinal inflammation. J Immunol *183*, 7514-7522.

Smith, P.M., Howitt, M.R., Panikov, N., Michaud, M., Gallini, C.A., Bohlooly, Y.M., Glickman, J.N., and Garrett, W.S. (2013). The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. Science *341*, 569-573.

Sokol, H., Lepage, P., Seksik, P., Dore, J., and Marteau, P. (2006). Temperature gradient gel electrophoresis of fecal 16S rRNA reveals active Escherichia coli in the microbiota of patients with ulcerative colitis. Journal of clinical microbiology *44*, 3172-3177.

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermudez-Humaran, L.G., Gratadoux, J.J., Blugeon, S., Bridonneau, C., Furet, J.P., Corthier, G.*, et al.* (2008). Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proceedings of the National Academy of Sciences of the United States of America *105*, 16731-16736.

Sokol, H., Seksik, P., Furet, J.P., Firmesse, O., Nion-Larmurier, I., Beaugerie, L., Cosnes, J., Corthier, G., Marteau, P., and Dore, J. (2009). Low counts of Faecalibacterium prausnitzii in colitis microbiota. Inflammatory bowel diseases *15*, 1183-1189.

Sommer, F., Adam, N., Johansson, M.E., Xia, L., Hansson, G.C., and Backhed, F. (2014). Altered mucus glycosylation in core 1 O-glycan-deficient mice affects microbiota composition and intestinal architecture. PloS one *9*, e85254.

Sommer, F., and Backhed, F. (2013). The gut microbiota--masters of host development and physiology. Nature reviews. Microbiology *11*, 227-238.

Sonnenburg, E.D., and Sonnenburg, J.L. (2014). Starving our microbial self: the deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. Cell metabolism *20*, 779-786.

Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.H., Westover, B.P., Weatherford, J., Buhler, J.D., and Gordon, J.I. (2005). Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science *307*, 1955-1959.

Spor, A., Koren, O., and Ley, R. (2011). Unravelling the effects of the environment and host genotype on the gut microbiome. Nature reviews. Microbiology *9*, 279-290.

Stappenbeck, T.S., Hooper, L.V., and Gordon, J.I. (2002). Developmental regulation of intestinal angiogenesis by indigenous microbes via Paneth cells. Proceedings of the National Academy of Sciences of the United States of America *99*, 15451-15455.

Steenwinckel, V., Louahed, J., Lemaire, M.M., Sommereyns, C., Warnier, G., McKenzie, A., Brombacher, F., Van Snick, J., and Renauld, J.C. (2009). IL-9 promotes IL-13-dependent paneth cell hyperplasia and up-regulation of innate immunity mediators in intestinal mucosa. J Immunol *182*, 4737-4743.

Stephens, N.S., Siffledeen, J., Su, X., Murdoch, T.B., Fedorak, R.N., and Slupsky, C.M. (2013). Urinary NMR metabolomic profiles discriminate inflammatory bowel disease from healthy. Journal of Crohn's & colitis *7*, e42-48.

Steppan, C.M., Brown, E.J., Wright, C.M., Bhat, S., Banerjee, R.R., Dai, C.Y., Enders, G.H., Silberg, D.G., Wen, X., Wu, G.D., and Lazar, M.A. (2001). A family of tissue-specific resistin-like molecules. Proceedings of the National Academy of Sciences of the United States of America *98*, 502-506.

Stewart, J.A., Chadwick, V.S., and Murray, A. (2005). Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children. Journal of medical microbiology *54*, 1239-1242.

Strauss, J., Kaplan, G.G., Beck, P.L., Rioux, K., Panaccione, R., Devinney, R., Lynch, T., and Allen-Vercoe, E. (2011). Invasive potential of gut mucosa-derived Fusobacterium nucleatum positively correlates with IBD status of the host. Inflammatory bowel diseases *17*, 1971-1978.

Strober, W., Fuss, I., and Mannon, P. (2007). The fundamental basis of inflammatory bowel disease. The Journal of clinical investigation *117*, 514-521.

Su, L., Shen, L., Clayburgh, D.R., Nalle, S.C., Sullivan, E.A., Meddings, J.B., Abraham, C., and Turner, J.R. (2009). Targeted epithelial tight junction dysfunction causes immune activation and contributes to development of experimental colitis. Gastroenterology *136*, 551-563.

Suau, A., Bonnet, R., Sutren, M., Godon, J.J., Gibson, G.R., Collins, M.D., and Dore, J. (1999). Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Applied and environmental microbiology *65*, 4799-4807.

Suzuki, K., Maruya, M., Kawamoto, S., Sitnik, K., Kitamura, H., Agace, W.W., and Fagarasan, S. (2010). The sensing of environmental stimuli by follicular dendritic cells promotes immunoglobulin A generation in the gut. Immunity *33*, 71-83.

Suzuki, K., Meek, B., Doi, Y., Muramatsu, M., Chiba, T., Honjo, T., and Fagarasan, S. (2004). Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. Proceedings of the National Academy of Sciences of the United States of America *101*, 1981-1986.

Swann, J.R., Want, E.J., Geier, F.M., Spagou, K., Wilson, I.D., Sidaway, J.E., Nicholson, J.K., and Holmes, E. (2011). Systemic gut microbial modulation of bile acid metabolism in host tissue compartments. Proceedings of the National Academy of Sciences of the United States of America *108 Suppl 1*, 4523-4530.

Swidsinski, A., Ladhoff, A., Pernthaler, A., Swidsinski, S., Loening-Baucke, V., Ortner, M., Weber, J., Hoffmann, U., Schreiber, S., Dietel, M., and Lochs, H. (2002). Mucosal flora in inflammatory bowel disease. Gastroenterology *122*, 44-54.

Swidsinski, A., Weber, J., Loening-Baucke, V., Hale, L.P., and Lochs, H. (2005). Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. Journal of clinical microbiology *43*, 3380-3389.

Tait Wojno, E.D., and Artis, D. (2012). Innate lymphoid cells: balancing immunity, inflammation, and tissue repair in the intestine. Cell host & microbe *12*, 445-457.

Topping, D.L., and Clifton, P.M. (2001). Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. Physiological reviews *81*, 1031-1064.

Trompette, A., Gollwitzer, E.S., Yadava, K., Sichelstiel, A.K., Sprenger, N., Ngom-Bru, C., Blanchard, C., Junt, T., Nicod, L.P., Harris, N.L., and Marsland, B.J. (2014). Gut microbiota

metabolism of dietary fiber influences allergic airway disease and hematopoiesis. Nature medicine *20*, 159-166.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P*., et al.* (2009a). A core gut microbiome in obese and lean twins. Nature *457*, 480-484.

Turnbaugh, P.J., Ridaura, V.K., Faith, J.J., Rey, F.E., Knight, R., and Gordon, J.I. (2009b). The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Science translational medicine *1*, 6ra14.

Turner, J.R. (2009). Intestinal mucosal barrier function in health and disease. Nature reviews. Immunology *9*, 799-809.

Ubeda, C., and Pamer, E.G. (2012). Antibiotics, microbiota, and immune defense. Trends in immunology *33*, 459-466.

Vaishnava, S., Behrendt, C.L., Ismail, A.S., Eckmann, L., and Hooper, L.V. (2008). Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface. Proceedings of the National Academy of Sciences of the United States of America *105*, 20858-20863.

Vaishnava, S., Yamamoto, M., Severson, K.M., Ruhn, K.A., Yu, X., Koren, O., Ley, R., Wakeland, E.K., and Hooper, L.V. (2011). The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. Science *334*, 255-258.

van Ampting, M.T.J., Rodenburg, W., Vink, C., Kramer, E., Schonewille, A.J., Keijer, J., van der Meer, R., and Bovee-Oudenhoven, I.M.J. (2009). Ileal Mucosal and Fecal Pancreatitis Associated Protein Levels Reflect Severity of Salmonella Infection in Rats. Digestive diseases and sciences *54*, 2588-2597.

van de Ven, R., Scheffer, G.L., Scheper, R.J., and de Gruijl, T.D. (2009). The ABC of dendritic cell development and function. Trends in immunology *30*, 421-429.

Van der Sluis, M., De Koning, B.A., De Bruijn, A.C., Velcich, A., Meijerink, J.P., Van Goudoever, J.B., Buller, H.A., Dekker, J., Van Seuningen, I., Renes, I.B., and Einerhand, A.W. (2006). Muc2-deficient mice spontaneously develop colitis, indicating that MUC2 is critical for colonic protection. Gastroenterology *131*, 117-129.

van Es, J.H., Jay, P., Gregorieff, A., van Gijn, M.E., Jonkheer, S., Hatzis, P., Thiele, A., van den Born, M., Begthel, H., Brabletz, T*., et al.* (2005). Wnt signalling induces maturation of Paneth cells in intestinal crypts. Nature cell biology *7*, 381-386.

Varela, E., Manichanh, C., Gallart, M., Torrejon, A., Borruel, N., Casellas, F., Guarner, F., and Antolin, M. (2013). Colonisation by Faecalibacterium prausnitzii and maintenance of clinical remission in patients with ulcerative colitis. Alimentary pharmacology & therapeutics *38*, 151-161.

Vavassori, P., Mencarelli, A., Renga, B., Distrutti, E., and Fiorucci, S. (2009). The bile acid receptor FXR is a modulator of intestinal innate immunity. J Immunol *183*, 6251-6261.

Veldhoen, M., and Brucklacher-Waldert, V. (2012). Dietary influences on intestinal immunity. Nature reviews. Immunology *12*, 696-708.

Vermeiren, J., Van den Abbeele, P., Laukens, D., Vigsnaes, L.K., De Vos, M., Boon, N., and Van de Wiele, T. (2012). Decreased colonization of fecal Clostridium coccoides/Eubacterium rectale species from ulcerative colitis patients in an in vitro dynamic gut model with mucin environment. FEMS microbiology ecology *79*, 685-696.

Vora, P., Youdim, A., Thomas, L.S., Fukata, M., Tesfay, S.Y., Lukasek, K., Michelsen, K.S., Wada, A., Hirayama, T., Arditi, M., and Abreu, M.T. (2004). Beta-defensin-2 expression is regulated by TLR signaling in intestinal epithelial cells. J Immunol *173*, 5398-5405.

Walker, C.R., Hautefort, I., Dalton, J.E., Overweg, K., Egan, C.E., Bongaerts, R.J., Newton, D.J., Cruickshank, S.M., Andrew, E.M., and Carding, S.R. (2013). Intestinal intraepithelial lymphocyte-enterocyte crosstalk regulates production of bactericidal angiogenin 4 by Paneth cells upon microbial challenge. PloS one *8*, e84553.

Wang, L., Fouts, D.E., Starkel, P., Hartmann, P., Chen, P., Llorente, C., DePew, J., Moncera, K., Ho, S.B., Brenner, D.A.*, et al.* (2016). Intestinal REG3 Lectins Protect against Alcoholic Steatohepatitis by Reducing Mucosa-Associated Microbiota and Preventing Bacterial Translocation. Cell host & microbe *19*, 227-239.

Wang, X., Heazlewood, S.P., Krause, D.O., and Florin, T.H. (2003). Molecular characterization of the microbial species that colonize human ileal and colonic mucosa by using 16S rDNA sequence analysis. Journal of applied microbiology *95*, 508-520.

Wang, Y.D., Chen, W.D., Yu, D., Forman, B.M., and Huang, W. (2011). The G-protein-coupled bile acid receptor, Gpbar1 (TGR5), negatively regulates hepatic inflammatory response through antagonizing nuclear factor kappa light-chain enhancer of activated B cells (NF-kappaB) in mice. Hepatology *54*, 1421-1432.

Wehkamp, J., Harder, J., Weichenthal, M., Mueller, O., Herrlinger, K.R., Fellermann, K., Schroeder, J.M., and Stange, E.F. (2003). Inducible and constitutive beta-defensins are differentially expressed in Crohn's disease and ulcerative colitis. Inflammatory bowel diseases *9*, 215-223.

Wehkamp, J., Schmid, M., Fellermann, K., and Stange, E.F. (2005). Defensin deficiency, intestinal microbes, and the clinical phenotypes of Crohn's disease. Journal of leukocyte biology *77*, 460-465.

Wei, M., Shinkura, R., Doi, Y., Maruya, M., Fagarasan, S., and Honjo, T. (2011). Mice carrying a knock-in mutation of Aicda resulting in a defect in somatic hypermutation have impaired gut homeostasis and compromised mucosal defense. Nature immunology *12*, 264-270.

Williams, H.R., Cox, I.J., Walker, D.G., Cobbold, J.F., Taylor-Robinson, S.D., Marshall, S.E., and Orchard, T.R. (2010). Differences in gut microbial metabolism are responsible for reduced hippurate synthesis in Crohn's disease. BMC gastroenterology *10*, 108.

Williams, H.R., Cox, I.J., Walker, D.G., North, B.V., Patel, V.M., Marshall, S.E., Jewell, D.P., Ghosh, S., Thomas, H.J., Teare, J.P.*, et al.* (2009). Characterization of inflammatory

bowel disease with urinary metabolic profiling. The American journal of gastroenterology *104*, 1435-1444.

Williams, H.R., Willsmore, J.D., Cox, I.J., Walker, D.G., Cobbold, J.F., Taylor-Robinson, S.D., and Orchard, T.R. (2012). Serum metabolic profiling in inflammatory bowel disease. Digestive diseases and sciences *57*, 2157-2165.

Willing, B., Halfvarson, J., Dicksved, J., Rosenquist, M., Jarnerot, G., Engstrand, L., Tysk, C., and Jansson, J.K. (2009). Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. Inflammatory bowel diseases *15*, 653-660.

Willing, B.P., Dicksved, J., Halfvarson, J., Andersson, A.F., Lucio, M., Zheng, Z., Jarnerot, G., Tysk, C., Jansson, J.K., and Engstrand, L. (2010). A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. Gastroenterology *139*, 1844-1854 e1841.

Wilson, B.A., Salyers, A.A., White, D.D., and Winkler, M.E. (2011). Bacterial pathogenesis: a molecular approach, 3rd edition edn (Washington DC: Americal Society for Microbiology).

Wlodarska, M., Kostic, A.D., and Xavier, R.J. (2015a). An integrative view of microbiome-host interactions in inflammatory bowel diseases. Cell host & microbe *17*, 577-591.

Wlodarska, M., Willing, B.P., Bravo, D.M., and Finlay, B.B. (2015b). Phytonutrient diet supplementation promotes beneficial Clostridia species and intestinal mucus secretion resulting in protection against enteric infection. Scientific reports *5*, 9253.

Wyatt, J., Vogelsang, H., Hubl, W., Waldhoer, T., and Lochs, H. (1993). Intestinal permeability and the prediction of relapse in Crohn's disease. Lancet *341*, 1437-1439.

Xavier, R.J., and Podolsky, D.K. (2007). Unravelling the pathogenesis of inflammatory bowel disease. Nature *448*, 427-434.

Yang, D., Chertov, O., Bykovskaia, S.N., Chen, Q., Buffo, M.J., Shogan, J., Anderson, M., Schroder, J.M., Wang, J.M., Howard, O.M., and Oppenheim, J.J. (1999). Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. Science *286*, 525-528.

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P*., et al.* (2012). Human gut microbiome viewed across age and geography. Nature *486*, 222-227.

Yoshida, M., Hatano, N., Nishiumi, S., Irino, Y., Izumi, Y., Takenawa, T., and Azuma, T. (2012). Diagnosis of gastroenterological diseases by metabolome analysis using gas chromatography-mass spectrometry. Journal of gastroenterology *47*, 9-20.

Zelante, T., Iannitti, R.G., Cunha, C., De Luca, A., Giovannini, G., Pieraccini, G., Zecchi, R., D'Angelo, C., Massi-Benedetti, C., Fallarino, F*., et al.* (2013). Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. Immunity *39*, 372-385.

Zhang, Y., Lin, L., Xu, Y., Lin, Y., Jin, Y., and Zheng, C. (2013). 1H NMR-based spectroscopy detects metabolic alterations in serum of patients with early-stage ulcerative colitis. Biochemical and biophysical research communications *433*, 547-551.

Zheng, Y., Valdez, P.A., Danilenko, D.M., Hu, Y., Sa, S.M., Gong, Q., Abbas, A.R., Modrusan, Z., Ghilardi, N., de Sauvage, F.J., and Ouyang, W. (2008). Interleukin-22 mediates early host defense against attaching and effacing bacterial pathogens. Nature medicine *14*, 282-289.

Zoetendal, E.G., Vaughan, E.E., and de Vos, W.M. (2006). A microbial world within us. Molecular microbiology *59*, 1639-1650.

Chapter 2

Compositional changes in the gut mucus microbiota precede the onset of colitic inflammation

M. Glymenaki[1], A. Brass[2], G. Warhurst[3], A.J. McBain[4], K.J. Else[1]†, S.M. Cruickshank[1]†*

[1]Faculty of Life Sciences, University of Manchester, Manchester, UK

[2]Bio-Health Informatics Group, School of Computer Science, University of Manchester, Manchester, UK

[3]Infection, Injury and Inflammation Research Group, Division of Medicine and Neurosciences, University of Manchester and Salford Royal Hospitals NHS Trust, Salford, UK

[4]Manchester Pharmacy School, University of Manchester, Manchester, UK

†These authors contributed equally to the study.

***Correspondence to:** Sheena Cruickshank, Faculty of Life Sciences, University of Manchester, Manchester, UK

sheena.cruickshank@manchester.ac.uk

Manuscript under review in Inflammatory Bowel Diseases

Author contributions

Maria Glymenaki: Designed and performed experiments, analysed data and wrote the manuscript.

Andy Brass: Supervised analysis of sequencing data, wrote the model in Python and edited the manuscript.

Geoffrey Warhurst: Designed and supervised the project, critically reviewed and edited the manuscript.

Andrew McBain: Designed and supervised the project, critically reviewed and edited the manuscript.

Kathryn Else: Designed and supervised the project, critically reviewed and edited the manuscript.

Sheena Cruickshank: Designed and supervised the project, critically reviewed and edited the manuscript.

2.1 Abstract

Inflammatory bowel disease (IBD) is associated with an inappropriate immune response to the gut microbiota. Notably, patients with IBD reportedly have alterations in faecal microbiota. However, the colonic microbiota occupies both the gut lumen and the mucus covering the epithelium. Thus, information about mucus-resident microbiota fails to be conveyed in the routine microbiota analyses of stool samples. Further, studies analyzing microbiota in IBD have mainly focused on stool samples taken after onset of inflammation. In contrast, here we investigate both temporal and spatial changes in colonic microbiota communities preceding the onset of colitis. We studied mucus and stool microbiota using a spontaneous model of colitis, the *mdr1a*$^{-/-}$ mouse, and their respective wild-type littermate controls in a time series mode. Using this approach we show for the first time that microbial dysbiosis with reduced abundance of Clostridiales was evident in the mucus but not stools, of colitis-prone mice *mdr1a*$^{-/-}$ mice 12 weeks before the onset of detectable inflammation. This altered microbial composition was coupled with a significantly thinner mucus layer. Upon emergence of inflammation, dysbiosis was evident in stool and at this time point, the spatial segregation between microbiota and host tissue was also disrupted, correlated with worsened inflammation. Our results reveal that microbial dysbiosis is detectable prior to changes in the stools. Importantly, dysbiosis in the mucus layer preceded development of colitis. Thus, our data reveal the importance of mucus sampling for understanding the underlying aetiology of IBD and fundamental processes underlying disease progression.

2.2Introduction

The gastrointestinal tract (GIT) microbiota is a complex and diverse microbial ecosystem (Backhed et al., 2005). The GIT microbiota performs essential functions for host physiology such as nutrient metabolism, inhibition of pathogen colonization, and education and development of the immune system (Hooper et al., 2012; Kamada et al., 2013). Despite the beneficial aspects of the microbiota, its abundance and proximity to the host poses a continuous challenge for the mucosal immune system. Thereby, the immune system has evolved mechanisms to tolerate the gut microbiota (Brown et al., 2013; Hooper et al., 2012) and anatomically contain it in the lumen and the outer mucus layer in order to maintain homeostasis (Johansson et al., 2008). A dense inner mucus layer that is considered sterile and is enriched in antimicrobial products mediates physical segregation of microbes from the intestinal epithelium (Johansson et al., 2011; Vaishnava et al., 2011).

Inflammatory bowel disease (IBD), comprising Crohn's disease (CD) and ulcerative colitis (UC), is thought to result from dysregulated immune responses to the microbiota (Khor et al., 2011; Maloy and Powrie, 2011) and has been associated with decreased bacterial diversity (particularly the Firmicutes phylum) and an altered balance of residential microbiota, so-called dysbiosis (Huttenhower et al., 2014; Manichanh et al., 2012; Rajilic-Stojanovic et al., 2013). Dysbiosis in IBD is varied and includes structural changes in the gut microbiota and impairments in microbial metabolic function (Morgan et al., 2012). Compositional or functional dysbiosis might compromise immune homeostasis, resulting in excessive inflammation against gut microbiota and vice versa, immune dysfunction might drive shifts in microbial structure establishing dysbiosis (Huttenhower et al., 2014; Kostic et al., 2014). Studies in individuals with IBD have examined the gut microbiota composition either during disease onset before treatment (Gevers et al., 2014) or after establishment of disease and clinical management (Kang et al., 2010; Lepage et al., 2011; Manichanh et al., 2006; Martinez et al., 2008; Papa et al., 2012; Willing et al., 2010). Furthermore, the majority of human studies have focused on stool samples, which do not fully replicate the complexity of the microbiota communities *in situ* (Kang et al., 2010; Manichanh et al., 2006; Papa et al., 2012). Bacterial composition differs depending on the niche that microbiota inhabit, for example the small intestine versus the large intestine or mucus layer overlying the epithelium

versus the gut lumen (Momozawa et al., 2011). Stool samples cannot capture these differences and it may be that more subtle changes in microbial communities undetected in stool are evident in specific niches in the gut prior to inflammation. We sought to address the composition and location of microbiota over time focusing on time points before the onset of any detectable inflammation and at onset of inflammation. Many mouse models of IBD rely on alerting the immune system (either by impairing it or activating it), or adding chemical or infectious agents that drive UC. Whilst the models have provided valuable information, they may not be the most appropriate models to understand the local changes in bacteria over time, as changes may be attributed to the trigger and not the environment. For this reason we used the *mdr1a*$^{-/-}$ spontaneous model of UC, which has an intact immune system (Panwala et al., 1998), and littermate controls in order to control for the environment. Our data revealed that changes in mucus but not stool bacteria preceded the onset of inflammation and highlight the importance of sampling the mucus to better understand disease mechanisms.

2.3 Materials and methods

2.3.1 Maintenance of animals

Mdr1a$^{-/-}$ mice (FVB.129P2-Abcb1atm1Bor N7) (Schinkel et al., 1994) and control FVB mice purchased from Taconic Farms (Albany, NY, USA) were crossbred to generate F2 littermates and maintained under specific, pathogen-free (SPF) conditions at the University of Manchester. Co-housed littermate *mdr1a*$^{-/-}$ and wild-type (WT) males were used throughout to ensure shared microbiota. All experimental procedures in mice were performed in accordance with the regulations issued by the Home Office under amended ASPA, 2012.

2.3.2 Isolation of bacterial genomic DNA

Stool samples were collected and frozen immediately in dry ice. Colon tissue was opened up longitudinally and washed in sterile PBS to remove luminal contents. Mucus was scraped for the collection of mucus resident bacteria. DNA extraction from stool and mucus samples was performed using the QIAamp® DNA Stool Mini Kit (Qiagen, Manchester, UK).

2.3.3 Denaturing gradient gel electrophoresis (DGGE) analysis

16S rRNA gene was amplified using the universal primer pairs P3_GC-341F and P2_518R (Supplementary Table 1), purified using the MinELute PCR Purification kit (Qiagen) and loaded onto a DGGE gel having a 30%-70% denaturing gradient. Gel images were analyzed using Bionumerics software (Applied Maths, TX, USA). A similarity matrix between sample fingerprints was extracted and principal component analysis (PCA) was performed using MATLAB (MathWorks, MA, USA). Non-metric multidimensional scaling (NMDS) using Bray-Curtis dissimilarity distances and statistical analysis (PERMANOVA; adonis function) was performed using the vegan package in R (http://cran.r-project.org).

2.3.4 Quantification of bacterial load and Gram-positive/Gram-negative bacteria numbers

Bacterial load in faeces and mucus layer was measured by qPCR, using universal 16S rRNA gene primers (forward 5'-ACTCCTACGGGAGGCAGCAGT -3', reverse 5'-ATTACCGCGGCTGCTGGC -3') and the Power SYBR Green Master Mix (Applied Biosystems, supplied by Fischer Scientific, Loughborough, UK) as previously described (Vaishnava et al., 2011). Gram-positive and Gram-negative bacteria numbers were quantified using Taqman® Gene Expression Master Mix (Applied Biosystems) with specific probes (Supplementary Table 1) as previously reported (Wu et al., 2008). Bacterial abundances were determined using standard curves constructed by a Topo-TA vector plasmid containing a cloned region (179 bp) of the 16S rRNA gene. Bacterial measurements in stool and mucus samples were normalized to the initial amount of each sample. Bacterial measurements correspond to the number of gene copies and not actual bacterial numbers or colony forming units.

2.3.5 16S rRNA gene sequencing analysis

16S amplicon sequencing targeting the V3 and V4 variable regions of the 16S rRNA (Supplementary Table 1) was performed on the Illumina MiSeq platform according to manufacturer's guidelines. The data were further processed using QIIME v.1.9.0 (Caporaso et al., 2010). Multivariate association with linear models (MaAsLin) analysis was performed using default parameters (http://huttenhower.sph.harvard.edu/maaslin) (Morgan et al., 2012).

2.3.6 Histology and colitis score

Distal colon was stained with haematoxylin and eosin (H&E) for histological assessment of tissue integrity and with alcian blue dye for goblet cell visualization. For enumeration of crypt lengths and goblet cells, the average number from a minimum of 60 crypts/slide from at least three slides was calculated per mouse. Slide pictures were taken using the slide scanner (Pannoramic 250 Flash), and analyzed using Panoramic Viewer software (3DHISTECH) and ImageJ (http://rsb.info.nih.gov/ij). All tissue was blinded prior to analysis. The sum of scores

for crypt length elongation (score 0-4), goblet cell depletion (score 0-4), muscle wall thickness (score 0-4), inflammatory cell infiltration (score 0-4) and destruction of architecture (score 0 or 3-4) was calculated (maximum 20) (Supplementary Table 2).

2.3.7 Fluorescence in situ hybridization (FISH) and mucin MUC2 immunostaining

Distal colon tissue was fixed in methanol-Carnoy's solution to preserve mucus integrity (Johansson et al., 2008). FISH staining was performed using the universal bacterial probe EUB338 (5'–Cy3-GCTGCCTCCCGTAGGAGT-3'), followed by immunostaining with a rabbit polyclonal MUC2 antibody and goat anti-rabbit Alexa-Fluor® 488 antibody (Life technologies, Paisley, UK). A nonspecific probe directed at fungus *Cryptococcus* was used as a control (5'-Cy3-CCAGCCCTTATCCACCGA-3'). Sections were mounted in ProLong® antifade mounting medium containing 4',6-diamidino-2-phenylindole (DAPI) (Life technologies) and analyzed using ImageJ. A minimum of 3 different distal colon sections was used for each sample. The thickness of the inner mucus layer was quantified by measuring the distance between epithelium and the start of outer mucus layer. Mucus penetration by bacteria was also evaluated. Gut location score was assigned values ranging from 0 to 4, thus: 0= bacteria in the lumen and outer mucus layer; 1= bacteria in the inner mucus layer; 2= bacteria in contact with the epithelium; 3= bacteria in the crypts and 4= bacteria in the lamina propria. All slides were scored blindly.

2.3.8 Statistical analysis

Statistical analysis was performed using GraphPad Prism 6 (GraphPad software, CA, USA). Normally distributed data were analyzed by one-way ANOVA with Tukey's posttest or unpaired t test as appropriate to the number of comparisons being made. Data that did not exhibit a normal distribution were analyzed using the nonparametric Kruskal-Wallis test with Dunn's posttest or Mann-Whitney test. Correlation analysis was performed using the Pearson correlation coefficient. $P<0.05$ was considered as statistically significant (*$P<0.05$, **$P<0.01$, ***$P<0.001$, ****$P<0.0001$).

2.4Results

2.4.1 Microbiota are altered in the mucus prior to the onset of inflammation

We monitored colitis progression in *mdr1a*$^{-/-}$ mice from 4 weeks until 25 weeks (Supplementary Figure 1A) and observed a slow progression of disease incidence in agreement with previous reports in *mdr1a*$^{-/-}$ mice reared under specific, pathogen-free (SPF) conditions (Collett et al., 2008; Dommels et al., 2007). We selected an early time point at 6 weeks to represent a disease-free post-weaning state and a later time point, at 18 weeks, when signs of inflammation begin to appear (Figure 1A). Histological examination of 6-week-old *mdr1a*$^{-/-}$ and wild-type (WT) mice revealed normal gut morphology with no signs of inflammation (Supplementary Figure 1B), in agreement with previous findings from our group based on gene expression analysis (Collett et al., 2008). By 18-weeks however, a subset of *mdr1a*$^{-/-}$ mice exhibited histological evidence of inflammation characteristic of progression to active disease (Supplementary Figure 1C). By 25 weeks most of the older *mdr1a*$^{-/-}$ mice had very inflamed guts with either moderate to severe disease (Supplementary Figure 1A). *Mdr1a*$^{-/-}$ mice had similar crypt length, muscle wall thickness and goblet cell numbers relative to their WT counterparts at both time points with the exception of those 18-week-old mice with indications of colitis (Supplementary Figure 1D - G).

To identify changes in microbiota composition that could promote a disturbed microbiota-epithelium crosstalk, we investigated the microbial consortium in both stools and mucus. Principal component analysis (PCA) of microbial fingerprints showed a shift in the overall community composition between young WT and *mdr1a*$^{-/-}$ mice that was limited to the mucus at 6 weeks (Figure 1B). In contrast, the microbial profiles in stool samples were very similar, indicating that changes in microbial composition start locally at specific niches. Non-metric multidimensional scaling (NMDS) analysis of these data using Bray-Curtis distances also revealed a diversification of mucus microbial communities between control and *mdr1a*$^{-/-}$ mice at 6 weeks (PERMANOVA: $R2=0.128$, $P<0.05$ for mucus; no difference in group variance).

We examined whether the altered microbial profile is associated with overall differences in the balance of Gram-positive and Gram-negative bacteria. Many pathobionts that belong to Gram-negative bacteria such as Enterobacteriaceae have increased abundance in IBD stool samples (Manichanh et al., 2012). We found similar numbers of Gram-positive and Gram-negative bacteria between WT and knockout mice in the mucus and stools at 6 weeks (Figure

1C and Supplementary Figure 2A); although mucus Gram-positive bacteria were more variable (F test of variance, $P<0.05$). The numbers of Gram-positive bacteria present in the stools were significantly higher than Gram-negative bacteria in both WT and knockout mice ($P<0.001$) (Supplementary Figure 2A).

Despite the different mucus microbial profile at 6 weeks, the total bacterial load, as assessed by qPCR of total 16S rRNA gene copy number, remained the same between young WT and *mdr1a*$^{-/-}$ animals at that site (Figure 1D). We also observed no significant difference in stool bacterial loads (Supplementary Figure 2B). Therefore, bacterial numbers were equivalent in the mucus, but the taxonomic composition of bacterial communities was different (Figure 1E). Interestingly, the relative abundance of Clostridiales, which has previously been associated with the maintenance of tolerance to the microbiota (Atarashi et al., 2011) was increased in WT controls, whereas the S24-7 family, part of Bacteroidetes phylum, exhibited reduced levels; this microbial pattern observed in the mucus was not reproduced in the stool samples (Supplementary Figure 2C). Microbial diversity was not affected by genotype either in mucus or faeces (Supplementary Figure 2D).

Principal coordinate analysis (PCoA) of the weighted UniFrac distance showed distinct clustering of mucus and faecal microbial assemblies (Supplementary Figure 3A) as shown in previous reports (Gevers et al., 2014; Momozawa et al., 2011; Morgan et al., 2012). 16S rRNA gene survey analysis also revealed differences in microbial composition (Supplementary Figure 3B) and diversity (Supplementary Figure 3C). Firmicutes were more abundant in the mucus, whereas Bacteroidetes dominated the lumen (Supplementary Figure 3B). The levels of those two phyla could distinguish sampling location using a predictor model in Python (file:///Users/maria/Downloads/metagenome.html).

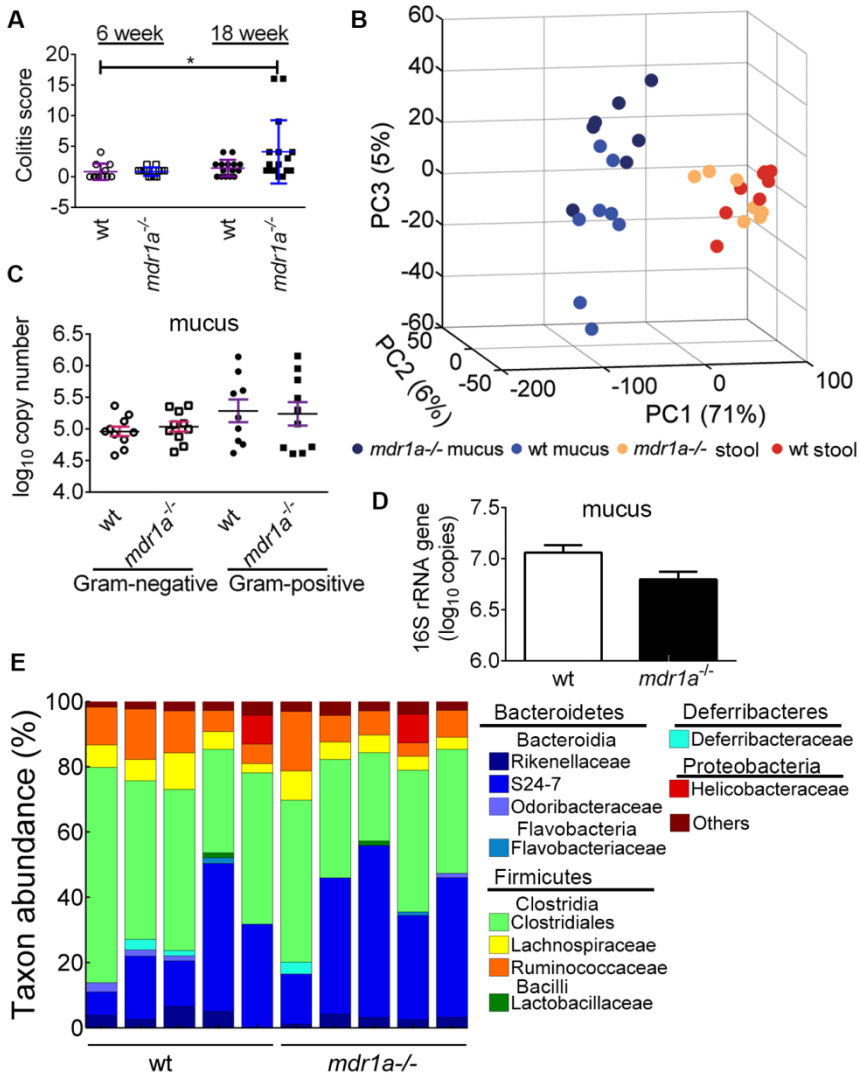


Figure 1. Differences in *mdr1a*$^{-/-}$ microbiota composition in the mucus are evident before onset of inflammation. Stool and mucus microbiota were analyzed by Illumina sequencing from WT and *mdr1a*$^{-/-}$ mice at 6 weeks. **(A)** Histopathology scoring of distal colon tissue samples obtained from 6-week and 18-week WT and *mdr1a*$^{-/-}$ mice (n=11-16). **(B)** PCA plot of fingerprinting data of faecal and mucus-associated microbiota from WT and *mdr1a*$^{-/-}$ mice (n=7-8/group), showing differential clustering of mucus microbiota from WT and *mdr1a*$^{-/-}$ mice but no clear clustering in stool bacteria (PERMANOVA; $P<0.05$ for mucus and non-significant for stool). **(C)** Balance of Gram-positive and Gram-negative bacteria in mucus microbial communities. n=10 mice per genotype **(D)** Bacterial load from the mucus was quantified by 16S rRNA gene copy number. n=5 mice per genotype; data are from 3 littermate groups. **(E)** Summary of the relative abundance of mucus microbiota at the family level (n=5/group). In case sequences could not be taxonomically assigned to the family level, the closest relative was given. Clostridiales corresponds to order level. The category 'others' represents bacterial families with abundance below 0.1%. Data shown as means ± SD. *, $P < 0.05$ by Kruskal-Wallis test with Dunn's multiple comparisons test **(A)**, one-way ANOVA with Tukey's multiple comparisons test **(C)** and Mann-Whitney test **(D)**.

2.4.2 Microbiota differences start in the mucus and progress to the stools over time, concordant with the onset of inflammation

Our data revealed changes in the mucus bacteria prior to any indicators of inflammation at 6 weeks. We hypothesized that such changes in microbiota composition start in the mucus and are later evident in the stools. Indeed, a differential clustering of global microbial profiles between 18-week-old WT and *mdr1a*$^{-/-}$ mice exhibiting indicators of inflammation was evident in both the mucus and stool (Figure 2A). NMDS analysis of these data using Bray-Curtis distances also revealed separation of mucus microbial communities between control and *mdr1a*$^{-/-}$ mice (PERMANOVA: R2=0.293, P=0.015 for mucus and R2=0.192, P=0.021 for stools; no difference in group variance).

The abundance of Gram-positive was much higher only in *mdr1a*$^{-/-}$ mice compared to the levels of Gram-negative bacteria in the mucus at 18 weeks (Figure 2B). On the contrary, levels of Gram-positive bacteria in the stools were increased in both WT and *mdr1a*$^{-/-}$ mice compared with Gram-negative bacteria levels (Supplementary Figure 4A). Notably *mdr1a*$^{-/-}$ mice exhibiting signs of inflammation showed a significant increase in the numbers of Gram-negative bacteria in the stools ($P<0.05$, Supplementary Figure 4B). Increased Gram-negative bacteria numbers in the stools but not the mucus were positively correlated with inflammation ($r=0.775$, $r^2=0.601$, $P<0.01$, Figure 2C). Thus, there was an association between *mdr1a*$^{-/-}$ mice that presented with a higher colitis score and elevated levels of Gram-negative bacteria.

Overall 18-week-old *mdr1a*$^{-/-}$ mice harbored similar numbers of mucus and stool bacteria (Figure 2D and Supplementary Figure 4C). Nonetheless, they showed an altered abundance of bacterial groups (Figure 2E and Supplementary Figure 4D). Members of the mucus microbial communities that belong to *Bacteroides* and the mucus-degrading *Akkermansia* were increased in 18-week-old *mdr1a*$^{-/-}$ mice, whereas levels of S24-7 group were reduced. We also noted an increase in Clostridiales in older *mdr1a*$^{-/-}$ mice. *Mdr1a*$^{-/-}$ mice that exhibited severe inflammation had a distinctive microbial profile compared with healthy age-matched knockout mice (Figure 3A, Supplementary Figure 4D and Supplementary Figure 5). There was evidence that the microbial profiles of the mucus and stools were more similar to each other in mice showing overt inflammation than those with little or no inflammation.

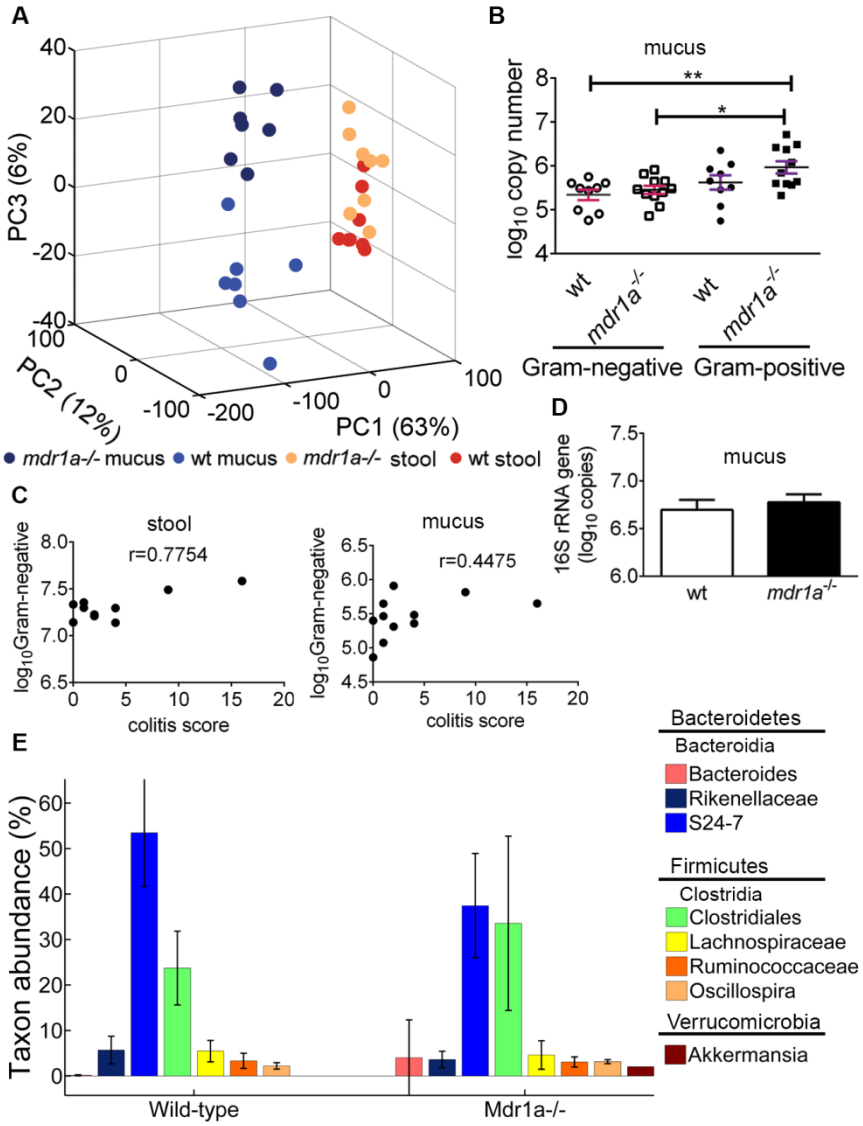


Figure 2. Microbial configuration in the mucus differs between WT and *mdr1a*$^{-/-}$ mice on disease onset. (A) PCA plot of stool and mucus samples based on fingerprinting data at 18 weeks of age from WT and *mdr1a*$^{-/-}$ mice (n=7-8/group) showing diversification of both mucus and lumen microbial communities between WT and *mdr1a*$^{-/-}$ mice (PERMANOVA: R2=0.293, P=0.015 for mucus and R2=0.192, P=0.021 for stools). *Mdr1a*$^{-/-}$ mice had no or a moderate inflammation. **(B)** Balance of Gram-positive and Gram-negative bacteria in the mucus microbial populations. n=10 mice/genotype. **(C)** Correlation of Gram-negative bacterial numbers with inflammation scoring in stool and mucus. **(D)** Quantification of mucus bacteria based on 16S rRNA gene copy number. n=5 mice per genotype from 5 littermate groups. **(E)** Distribution of bacterial taxa at the genus level in mucus samples (n=5/group) at 0.02% cutoff. Rikenellaceae, S24-7, Lachnospiraceae, Ruminococceae correspond to family level. Clostridiales corresponds to order. Data shown are means ± SEM. *p<0.05; **p<0.01 as determined by one-way ANOVA with Tukey's multiple comparisons test **(B)** and Mann-Whitney test **(D)**. P<0.05 in stool samples in **(C)** by Pearson correlation analysis.

To gain insight into the effect of *mdr1a* gene deficiency into driving inflammation, we looked at the microbial structure in WT and *mdr1a*$^{-/-}$ mice. Interestingly, compared with WT, *mdr1a*$^{-/-}$ mice showed similar alpha diversity both in their mucus and stools (Figure 3B and C). To determine a correlation between the gut microbial composition and status of disease, we used the multivariate association with linear models (MaAsLin) analysis (Morgan et al., 2012). Progression to colitis associated with an increase in the abundance of S24-7 (Figure 3D), whereas Clostridiales tended to decline in frequency (Figure 3E). Our data suggest that specific bacterial taxa are altered in the mucus in colitis-prone animals before disease onset and that these differences are later also apparent in the stools after inflammation has developed.

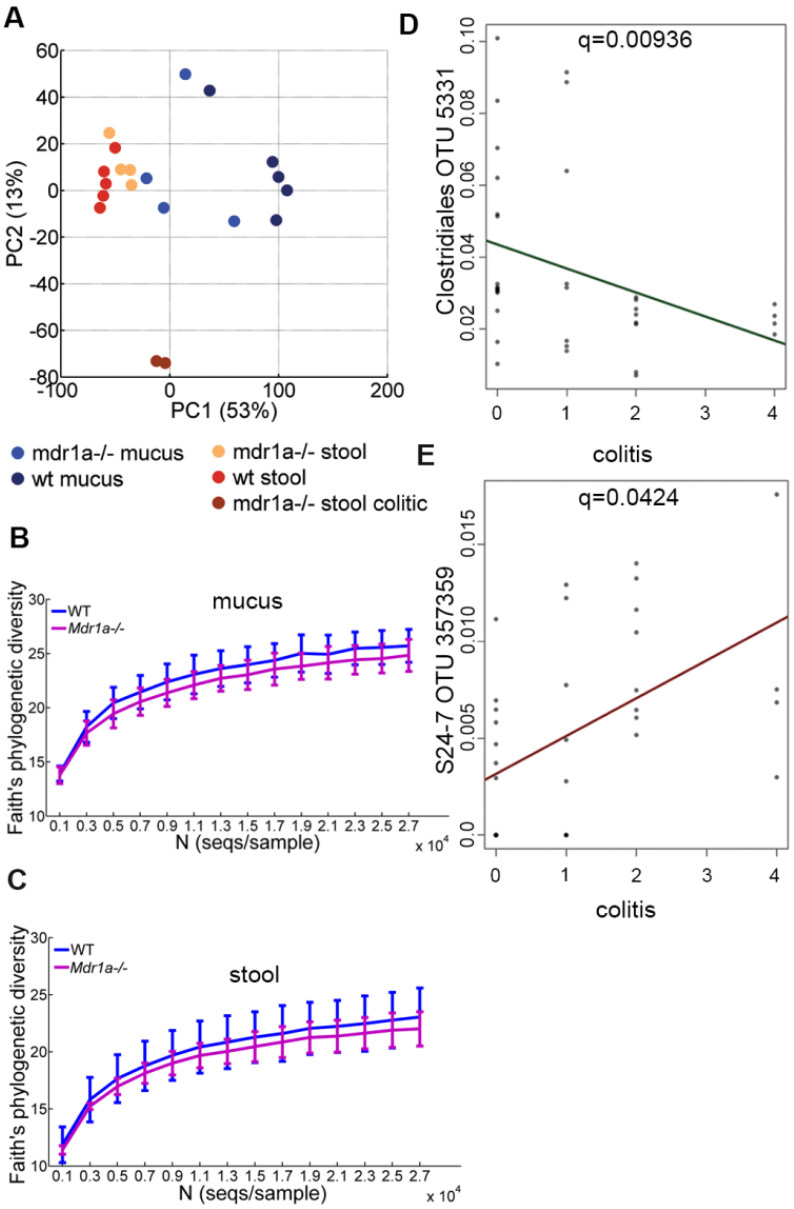


Figure 3. Disrupted bacterial relationships in *mdr1a*$^{-/-}$ animals progressing to disease. **(A)** PCA plot of stool and mucus samples based on fingerprinting data from WT and *mdr1a*$^{-/-}$ mice at 18 weeks (n=4-6/group) showing differential clustering of stool microbial communities in *mdr1a*$^{-/-}$ animals with no inflammation compared with age-matched *mdr1a*$^{-/-}$ with inflammation. **(B)** Alpha diversity rarefaction plots of phylogenetic diversity in WT and *mdr1a*$^{-/-}$ mice in mucus and **(C)** stools at 18 weeks (n=5/group, means± SD). **(D and E)** Significant community shifts were associated with progression to inflammation in *mdr1a*$^{-/-}$ mice compared with healthy WT and disease-free *mdr1a*$^{-/-}$ mice. Analysis and associated FDR-corrected p values were computed using MaAsLin.

2.4.3 Altered location of microbiota in colitis-prone mice

To identify whether microbial localization is implicated in an altered host-bacterial relationship, we examined the spatial distribution of microbiota in the context of colitis development by fluorescence *in situ* hybridization (FISH) using a universal 16S rRNA gene probe (Supplementary Figure 6A - C). We observed that bacteria were mostly located at the outer mucus layer in 6-week old WT and *mdr1a*$^{-/-}$ mice; however, few bacteria were occasionally found in the inner mucus layer (Figure 4A and B). The spatial separation between microbiota and intestinal mucosa was maintained in 18-week WT mice. In contrast, microbiota were frequently observed to have penetrated the inner mucus barrier and contacted the intestinal epithelial surface in age-matched cohoused *mdr1a*$^{-/-}$ mice ($P<0.01$, Figure 4A and C). In cases of severe inflammation, bacteria also colonized the intestinal crypts and penetrated the lamina propria.

Immunostaining of MUC2 revealed that the thickness of the inner mucus was decreased in *mdr1a*$^{-/-}$ mice compared with age-matched controls at 6 weeks ($P<0.05$, Figure 4D). Moreover, as *mdr1a*$^{-/-}$ mice aged (18 weeks), the mucus layer diminished significantly ($P<0.05$, Figure 4E). Our findings reveal an association between bacterial penetration, as evidenced in higher gut location score, and the grade of inflammation ($r=0.69$, $r^2=0.476$, $P<0.05$, Figure 4F). A reduction in mucus thickness was negatively correlated with the grade of inflammation ($r=-0.754$, $r^2=0.569$, $P<0.05$, Figure 4G). The decrease in mucus layer may provide the opportunity for colonization by commensal bacteria that are normally excluded from direct epithelial contact. These results suggest that the spatial relationship between microbiota and host intestinal tissue is disrupted before the onset of colitis symptoms, creating a hypersensitive environment that challenges the immune system.

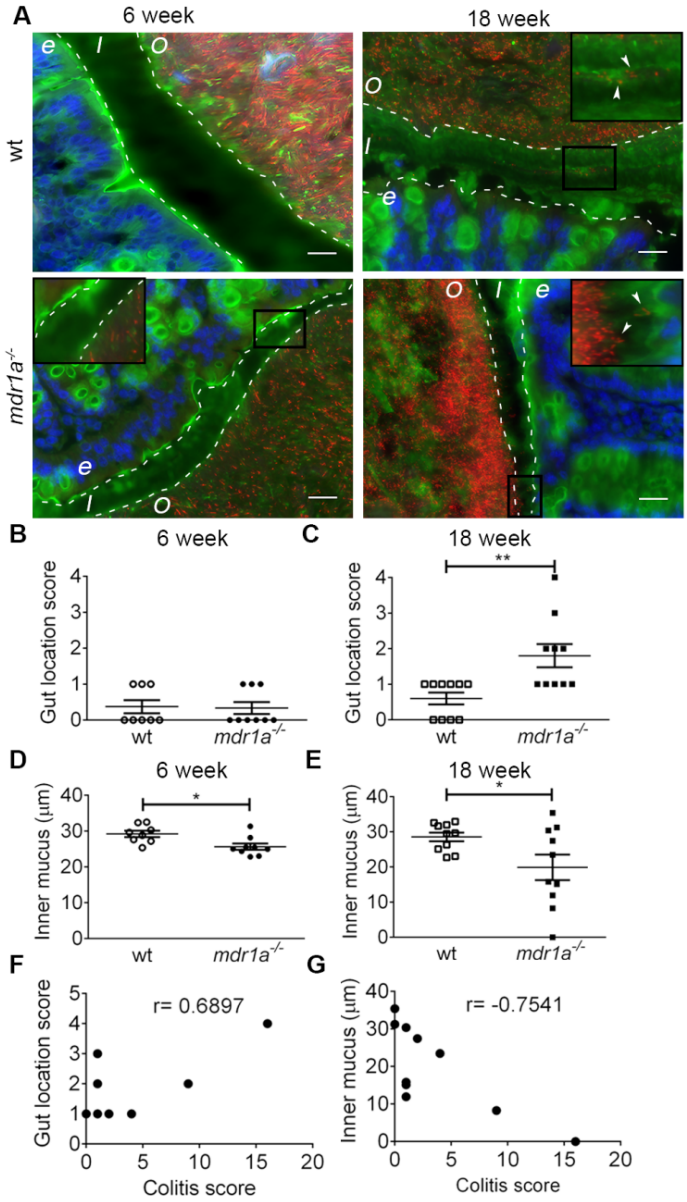


Figure 4. Altered gut microbial location and mucus thickness precedes inflammation in colitis-prone *mdr1a*$^{-/-}$ mice. (A) Representative images of microbiota localization in distal colon tissue by FISH at two time points, 6 weeks and 18 weeks. Bacteria were stained with the universal eubacterial 16S rRNA gene probe-EUB338 (red) and mucus was identified with an anti-MUC2 specific antiserum (green). Tissue was counterstained with DAPI (blue). Dotted lines indicate the inner mucus layer (I) that separates the intestinal epithelium (e) from the outer mucus layer (O). Arrowheads point to bacteria in the inner mucus layer and in contact with the epithelium. Sections are representative of > 8 littermate groups. Scale bars=20µm. **(B)** Scoring of bacterial location in distal colon tissue samples derived from 6- and **(C)** 18-week old WT and *mdr1a*$^{-/-}$ mice (n=8-10). **(D)** Thickness of the inner mucus layer in WT and *mdr1a*$^{-/-}$ mice at 6-weeks and **(E)** 18-weeks. **(F)** Correlation of bacterial location scoring and **(G)** inner mucus thickness with disease status in *mdr1a*$^{-/-}$ mice at 18 weeks. A higher colitis score was associated with decreased mucus thickness and increased bacterial penetrance. Data shown are means ± SEM. *, P<0.05; **, P<0.01 as determined by unpaired

t-test in **(B)** and **(D)**, unpaired t-test with Welch's correction in **(E)** and Mann-Whitney test **(C)**. $P<0.05$ in **(F, G)** by Pearson correlation analysis.

2.5Discussion

The composition of the gut microbiota is known to be disturbed in chronic intestinal inflammatory conditions (Huttenhower et al., 2014; Kostic et al., 2014; Manichanh et al., 2012). Here we show for the first time that compositional changes in the mucus microbiota can be detected long before the onset of inflammation (12 weeks before inflammation was evident) in colitis-prone mice. Despite the fact that the total bacterial load was similar, *mdr1a*$^{-/-}$ animals had reduced proportions of Clostridiales and higher abundance of S24-7 group in their mucus communities. Clostridiales have been associated with the development of Tregs (Atarashi et al., 2013; Atarashi et al., 2011; Furusawa et al., 2013) and it has been reported that *mdr1a*$^{-/-}$ animals have reduced frequency of Foxp3^{+} Tregs in the lamina propria (Tanner et al., 2013). Therefore, the reduction in mucus Clostridiales in *mdr1a*$^{-/-}$ animals could have altered the mucus-microbiota/immune-dialogue and as a result impaired the induction of Tregs. Interestingly, a previous study that correlated high Immunoglobulin A (IgA) binding of bacteria with colitogenic potential implicated the S24-7 as commensal bacteria with high IgA coating (Palm et al., 2014). Notably, our data reveals an increased abundance of the S24-7 in the mucus of colitis-prone mice. Thus, early changes in microbiota may be favoring an environment that is more colitogenic and less tolerogenic.

In accordance with our findings a previous study (Nones et al., 2009) showed differences in the caecal microbiota of *mdr1a*$^{-/-}$ animals compared with controls at around 12 weeks of age. Inflammation was not assessed in this study, but rather inferred to be absent, based on an earlier study from the same group (Dommels et al., 2007); however, it is important to state that other groups have reported that 12 weeks is associated with active inflammation (Collett et al., 2008; Resta-Lenert et al., 2005). Changes such as expression of interferon response genes, levels of inflammation and disease onset can be variable in the *mdr1a*$^{-/-}$ model, thus it is critical to understand what the level of inflammation or inflammatory markers are in the same mice from which samples are collected for microbiota analysis (Collett et al., 2008). Importantly, our study identifies changes in colon mucus bacteria at a very early time point when there is no evidence of either histological inflammation or expression of any molecular indicators of inflammation (Collett et al., 2008). In addition, it is important to control for the environment in microbiota studies, as there is substantial microbial variation between cages, location and litters (Friswell et al., 2010); we have achieved this by using littermate controls in co-housing conditions.

Spatial segregation of microbiota is crucial for the maintenance of homeostasis (Johansson et al., 2011). *Mdr1a*$^{-/-}$ animals had a thinner inner mucus barrier at an early age, indicating a compromised mucus barrier before the emergence of any indicators of inflammation. A marked reduction of the mucus thickness was evident during the development of inflammation in agreement with previous reports on active inflammation in murine UC models (Perez-Bosque et al., 2015) or IBD patients (Fyderek et al., 2009; Larsson et al., 2011). Notwithstanding, spatial segregation of microbiota was maintained in *mdr1a*$^{-/-}$ mice at a young age (i.e. 6 weeks) as previously described (Collett et al., 2008). Bacteria have been shown to penetrate the inner mucus layer and reach the epithelium in murine models of colitis and patients with active UC (Johansson et al., 2014). Akin to this work, breach of the mucus barrier in *mdr1a*$^{-/-}$ animals was most frequently observed during the onset of inflammation. Indeed, *mdr1a*$^{-/-}$ mice with a higher penetration of invading bacteria also had worsened gut pathology. Bacteria initially colonized the epithelium and in severe inflammation invaded the host tissue. Previous studies also suggested a correlation between severity of inflammation and bacterial penetration in the inner colonic mucus (Johansson et al., 2014; Perez-Bosque et al., 2015), but to our knowledge our study is the first to show reduced mucus barrier thickness and thus altered mucus properties prior to disease onset. Recent findings indicate that bacteria themselves can influence mucus barrier properties (Jakobsson et al., 2014) and since mucus bacterial composition is altered in *mdr1a*$^{-/-}$ mice prior to disease, it may contribute to that impaired mucus barrier.

The shift in microbial community composition that was initially identified in the mucus of colitis-prone mice progresses to stools over time, concurrently with the development of inflammation. The abundance of Gram-negative bacteria was elevated in the stools of colitis-prone animals and showed a positive correlation with the severity of inflammation. The increase in Gram-negative bacteria is a sign of microbial imbalance (Huttenhower et al., 2014; Peterson et al., 2008). We found that the levels of *Bacteroides*, *Akkermansia* and Clostridiales increased on disease onset in the intestinal mucus in *mdr1a*$^{-/-}$ mice, whereas the S24-7 group declined in abundance. *Akkermania* sp. are mucin-degrading bacteria residing in mucus (Derrien et al., 2008). Studies in human UC samples from tissue biopsies and faeces have shown reduced numbers of *Akkermansia* sp. (Png et al., 2010; Rajilic-Stojanovic et al., 2013); however, these studies report an increase in other mucolytic bacteria (Hickey et al., 2015). The increase in *Akkermansia* abundance together with an increased proportion of

Bacteroides, which produce mucus-degrading sulfatases (Hickey et al., 2015), suggests an excess of liberated mucus glycans, facilitating the expansion of harmful bacteria that forage on these readily provided glycans. The differences in communities at 6 and 18 weeks reveal the importance of sampling over time to better unpick potential drivers and initiators of inflammation and/or homeostasis. The distinct microbial pattern we observed differentiated not only disease-prone animals from healthy controls, but also disease-prone animals exhibiting varying grades of colitis. Severe inflammation was associated with a different microbial profile compared with low-grade inflammation.

Overall, our data indicate that differences in microbiota composition start in the mucus before the onset of inflammation and then later extend to stools in parallel with an altered mucus barrier. On disease onset, loss of spatial compartmentalization of bacteria could further trigger the immune system and fuel the inflammatory cascade and vice versa. Our findings provide a framework for understanding the temporal and spatial variation of microbial communities in the context of IBD and highlight the importance of sampling mucus-associated microbiota.

Acknowledgements

We thank L. Hooper for the Topo-TA plasmid vector, R.K. Grencis and D. Thornton for the MUC2-antiserum and the anti-rabbit Alexa-Fluor® 488 antibody. M. Constantinou provided extensive technical support in Matlab. G. Singh and C. Knight provided technical assistance for data analysis. We also thank European Bioinformatics Institute (EBI) and the Qiime community forum for help in sequencing analysis. Sequences of the bacterial 16S rRNA genes obtained from faecal and mucus bacteria have been deposited to the EBI (study accession number PRJEB6905). This work was supported by a UK Biotechnology and Biological Sciences Research Council (BBSRC) studentship for MG awarded to SC, KE, AJM and GW and a grant from the European Crohn's and Colitis Organization (ECCO) awarded to SC. The authors declare no conflicts of interest.

References

Atarashi, K., Tanoue, T., Oshima, K., Suda, W., Nagano, Y., Nishikawa, H., Fukuda, S., Saito, T., Narushima, S., Hase, K., *et al.* (2013). Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature *500*, 232-236.

Atarashi, K., Tanoue, T., Shima, T., Imaoka, A., Kuwahara, T., Momose, Y., Cheng, G., Yamasaki, S., Saito, T., Ohba, Y., *et al.* (2011). Induction of colonic regulatory T cells by indigenous Clostridium species. Science *331*, 337-341.

Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., and Gordon, J.I. (2005). Host-bacterial mutualism in the human intestine. Science *307*, 1915-1920.

Brown, E.M., Sadarangani, M., and Finlay, B.B. (2013). The role of the immune system in governing host-microbe interactions in the intestine. Nature immunology *14*, 660-667.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. Nature methods *7*, 335-336.

Collett, A., Higgs, N.B., Gironella, M., Zeef, L.A., Hayes, A., Salmo, E., Haboubi, N., Iovanna, J.L., Carlson, G.L., and Warhurst, G. (2008). Early molecular and functional changes in colonic epithelium that precede increased gut permeability during colitis development in mdr1a(-/-) mice. Inflammatory bowel diseases *14*, 620-631.

Derrien, M., Collado, M.C., Ben-Amor, K., Salminen, S., and de Vos, W.M. (2008). The Mucin degrader Akkermansia muciniphila is an abundant resident of the human intestinal tract. Applied and environmental microbiology *74*, 1646-1648.

Dommels, Y.E., Butts, C.A., Zhu, S., Davy, M., Martell, S., Hedderley, D., Barnett, M.P., McNabb, W.C., and Roy, N.C. (2007). Characterization of intestinal inflammation and identification of related gene expression changes in mdr1a(-/-) mice. Genes & nutrition *2*, 209-223.

Friswell, M.K., Gika, H., Stratford, I.J., Theodoridis, G., Telfer, B., Wilson, I.D., and McBain, A.J. (2010). Site and strain-specific variation in gut microbiota profiles and metabolism in experimental mice. PloS one *5*, e8584.

Furusawa, Y., Obata, Y., Fukuda, S., Endo, T.A., Nakato, G., Takahashi, D., Nakanishi, Y., Uetake, C., Kato, K., Kato, T., *et al.* (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature *504*, 446-450.

Fyderek, K., Strus, M., Kowalska-Duplaga, K., Gosiewski, T., Wedrychowicz, A., Jedynak-Wasowicz, U., Sladek, M., Pieczarkowski, S., Adamski, P., Kochan, P., and Heczko, P.B. (2009). Mucosal bacterial microflora and mucus layer thickness in adolescents with inflammatory bowel disease. World journal of gastroenterology : WJG *15*, 5287-5294.

Gevers, D., Kugathasan, S., Denson, L.A., Vazquez-Baeza, Y., Van Treuren, W., Ren, B., Schwager, E., Knights, D., Song, S.J., Yassour, M., *et al.* (2014). The treatment-naive microbiome in new-onset Crohn's disease. Cell host & microbe *15*, 382-392.

Hickey, C.A., Kuhn, K.A., Donermeyer, D.L., Porter, N.T., Jin, C., Cameron, E.A., Jung, H., Kaiko, G.E., Wegorzewska, M., Malvin, N.P., *et al.* (2015). Colitogenic *Bacteroides thetaiotaomicron* Antigens Access Host Immune Cells in a Sulfatase-Dependent Manner via Outer Membrane Vesicles. Cell host & microbe *17*, 672-680.

Hooper, L.V., Littman, D.R., and Macpherson, A.J. (2012). Interactions between the microbiota and the immune system. Science *336*, 1268-1273.

Huttenhower, C., Kostic, A.D., and Xavier, R.J. (2014). Inflammatory bowel disease as a model for translating the microbiome. Immunity *40*, 843-854.

Jakobsson, H.E., Rodriguez-Pineiro, A.M., Schutte, A., Ermund, A., Boysen, P., Bemark, M., Sommer, F., Backhed, F., Hansson, G.C., and Johansson, M.E. (2014). The composition of the gut microbiota shapes the colon mucus barrier. EMBO reports.

Johansson, M.E., Gustafsson, J.K., Holmen-Larsson, J., Jabbar, K.S., Xia, L., Xu, H., Ghishan, F.K., Carvalho, F.A., Gewirtz, A.T., Sjovall, H., and Hansson, G.C. (2014). Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. Gut *63*, 281-291.

Johansson, M.E., Larsson, J.M., and Hansson, G.C. (2011). The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. Proceedings of the National Academy of Sciences of the United States of America *108 Suppl 1*, 4659-4665.

Johansson, M.E., Phillipson, M., Petersson, J., Velcich, A., Holm, L., and Hansson, G.C. (2008). The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. Proceedings of the National Academy of Sciences of the United States of America *105*, 15064-15069.

Kamada, N., Seo, S.U., Chen, G.Y., and Nunez, G. (2013). Role of the gut microbiota in immunity and inflammatory disease. Nature reviews. Immunology *13*, 321-335.

Kang, S., Denman, S.E., Morrison, M., Yu, Z., Dore, J., Leclerc, M., and McSweeney, C.S. (2010). Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray. Inflammatory bowel diseases *16*, 2034-2042.

Khor, B., Gardet, A., and Xavier, R.J. (2011). Genetics and pathogenesis of inflammatory bowel disease. Nature *474*, 307-317.

Kostic, A.D., Xavier, R.J., and Gevers, D. (2014). The microbiome in inflammatory bowel disease: current status and the future ahead. Gastroenterology *146*, 1489-1499.

Larsson, J.M., Karlsson, H., Crespo, J.G., Johansson, M.E., Eklund, L., Sjovall, H., and Hansson, G.C. (2011). Altered O-glycosylation profile of MUC2 mucin occurs in active ulcerative colitis and is associated with increased inflammation. Inflammatory bowel diseases *17*, 2299-2307.

Lepage, P., Hasler, R., Spehlmann, M.E., Rehman, A., Zvirbliene, A., Begun, A., Ott, S., Kupcinskas, L., Dore, J., Raedler, A., and Schreiber, S. (2011). Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. Gastroenterology *141*, 227-236.

Maloy, K.J., and Powrie, F. (2011). Intestinal homeostasis and its breakdown in inflammatory bowel disease. Nature *474*, 298-306.

Manichanh, C., Borruel, N., Casellas, F., and Guarner, F. (2012). The gut microbiota in IBD. Nature reviews. Gastroenterology & hepatology *9*, 599-608.

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., Nalin, R., Jarrin, C., Chardon, P., Marteau, P.*, et al.* (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut *55*, 205-211.

Martinez, C., Antolin, M., Santos, J., Torrejon, A., Casellas, F., Borruel, N., Guarner, F., and Malagelada, J.R. (2008). Unstable composition of the fecal microbiota in ulcerative colitis during clinical remission. The American journal of gastroenterology *103*, 643-648.

Momozawa, Y., Deffontaine, V., Louis, E., and Medrano, J.F. (2011). Characterization of bacteria in biopsies of colon and stools by high throughput sequencing of the V2 region of bacterial 16S rRNA gene in human. PloS one *6*, e16952.

Morgan, X.C., Tickle, T.L., Sokol, H., Gevers, D., Devaney, K.L., Ward, D.V., Reyes, J.A., Shah, S.A., LeLeiko, N., Snapper, S.B.*, et al.* (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome biology *13*, R79.

Nones, K., Knoch, B., Dommels, Y.E., Paturi, G., Butts, C., McNabb, W.C., and Roy, N.C. (2009). Multidrug resistance gene deficient (mdr1a-/-) mice have an altered caecal microbiota that precedes the onset of intestinal inflammation. Journal of applied microbiology *107*, 557-566.

Palm, N.W., de Zoete, M.R., Cullen, T.W., Barry, N.A., Stefanowski, J., Hao, L., Degnan, P.H., Hu, J., Peter, I., Zhang, W.*, et al.* (2014). Immunoglobulin A coating identifies colitogenic bacteria in inflammatory bowel disease. Cell *158*, 1000-1010.

Panwala, C.M., Jones, J.C., and Viney, J.L. (1998). A novel model of inflammatory bowel disease: mice deficient for the multiple drug resistance gene, mdr1a, spontaneously develop colitis. J Immunol *161*, 5733-5744.

Papa, E., Docktor, M., Smillie, C., Weber, S., Preheim, S.P., Gevers, D., Giannoukos, G., Ciulla, D., Tabbaa, D., Ingram, J.*, et al.* (2012). Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. PloS one *7*, e39242.

Perez-Bosque, A., Miro, L., Maijo, M., Polo, J., Campbell, J., Russell, L., Crenshaw, J., Weaver, E., and Moreto, M. (2015). Dietary intervention with serum-derived bovine immunoglobulins protects barrier function in a mouse model of colitis. American journal of physiology. Gastrointestinal and liver physiology *308*, G1012-1018.

Peterson, D.A., Frank, D.N., Pace, N.R., and Gordon, J.I. (2008). Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. Cell host & microbe *3*, 417-427.

Png, C.W., Linden, S.K., Gilshenan, K.S., Zoetendal, E.G., McSweeney, C.S., Sly, L.I., McGuckin, M.A., and Florin, T.H. (2010). Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. The American journal of gastroenterology *105*, 2420-2428.

Rajilic-Stojanovic, M., Shanahan, F., Guarner, F., and de Vos, W.M. (2013). Phylogenetic analysis of dysbiosis in ulcerative colitis during remission. Inflammatory bowel diseases *19*, 481-488.

Resta-Lenert, S., Smitham, J., and Barrett, K.E. (2005). Epithelial dysfunction associated with the development of colitis in conventionally housed mdr1a-/- mice. American journal of physiology. Gastrointestinal and liver physiology *289*, G153-162.

Schinkel, A.H., Smit, J.J., van Tellingen, O., Beijnen, J.H., Wagenaar, E., van Deemter, L., Mol, C.A., van der Valk, M.A., Robanus-Maandag, E.C., te Riele, H.P., and et al. (1994). Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. Cell *77*, 491-502.

Tanner, S.M., Staley, E.M., and Lorenz, R.G. (2013). Altered generation of induced regulatory T cells in the FVB.mdr1a-/- mouse model of colitis. Mucosal immunology *6*, 309-323.

Vaishnava, S., Yamamoto, M., Severson, K.M., Ruhn, K.A., Yu, X., Koren, O., Ley, R., Wakeland, E.K., and Hooper, L.V. (2011). The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. Science *334*, 255-258.

Willing, B.P., Dicksved, J., Halfvarson, J., Andersson, A.F., Lucio, M., Zheng, Z., Jarnerot, G., Tysk, C., Jansson, J.K., and Engstrand, L. (2010). A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. Gastroenterology *139*, 1844-1854 e1841.

Wu, Y.D., Chen, L.H., Wu, X.J., Shang, S.Q., Lou, J.T., Du, L.Z., and Zhao, Z.Y. (2008). Gram stain-specific-probe-based real-time PCR for diagnosis and discrimination of bacterial neonatal sepsis. Journal of clinical microbiology *46*, 2613-2619.

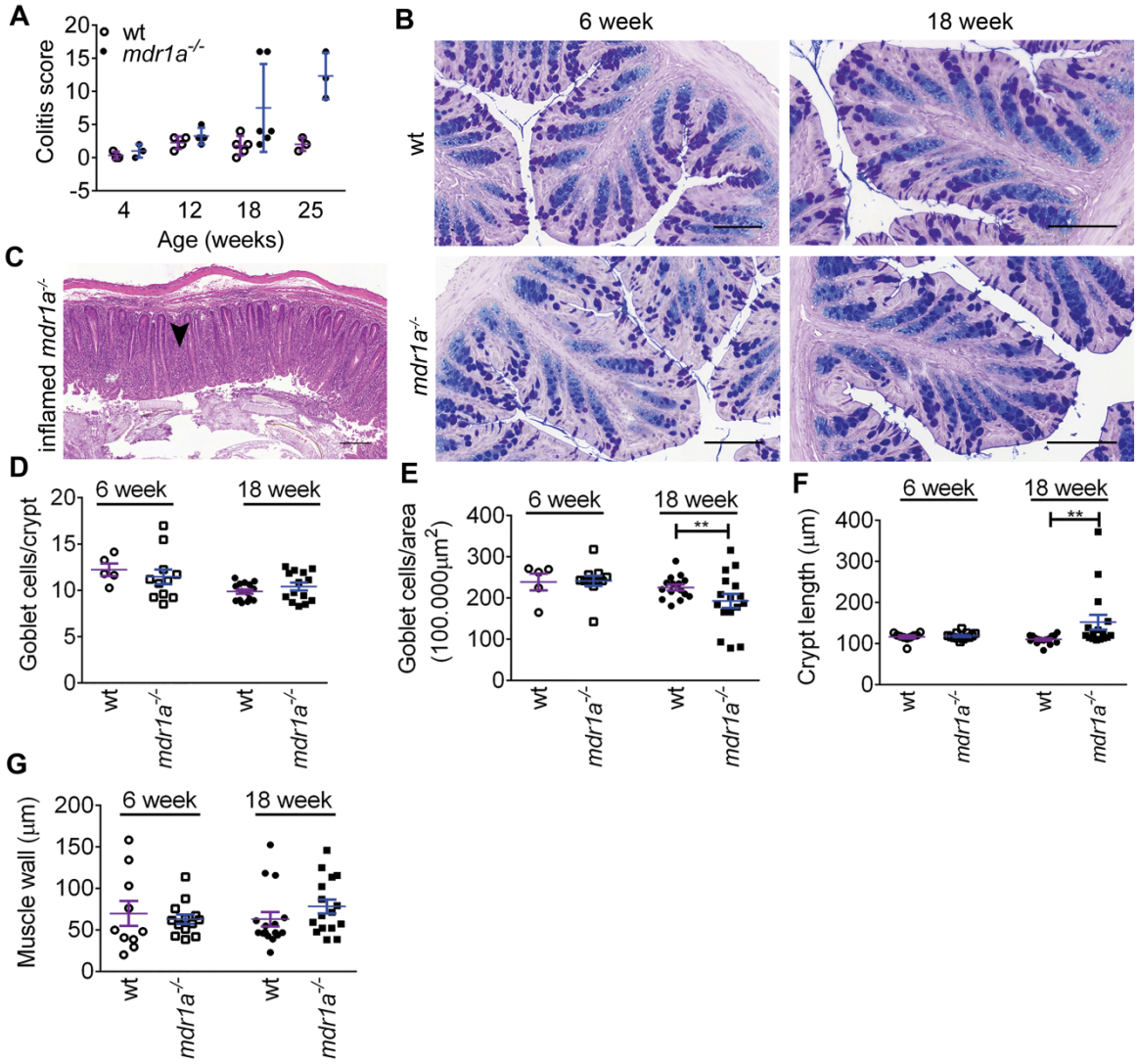
Supplementary materials

Supplementary Table 1.Bacterial 16S rRNA primers.

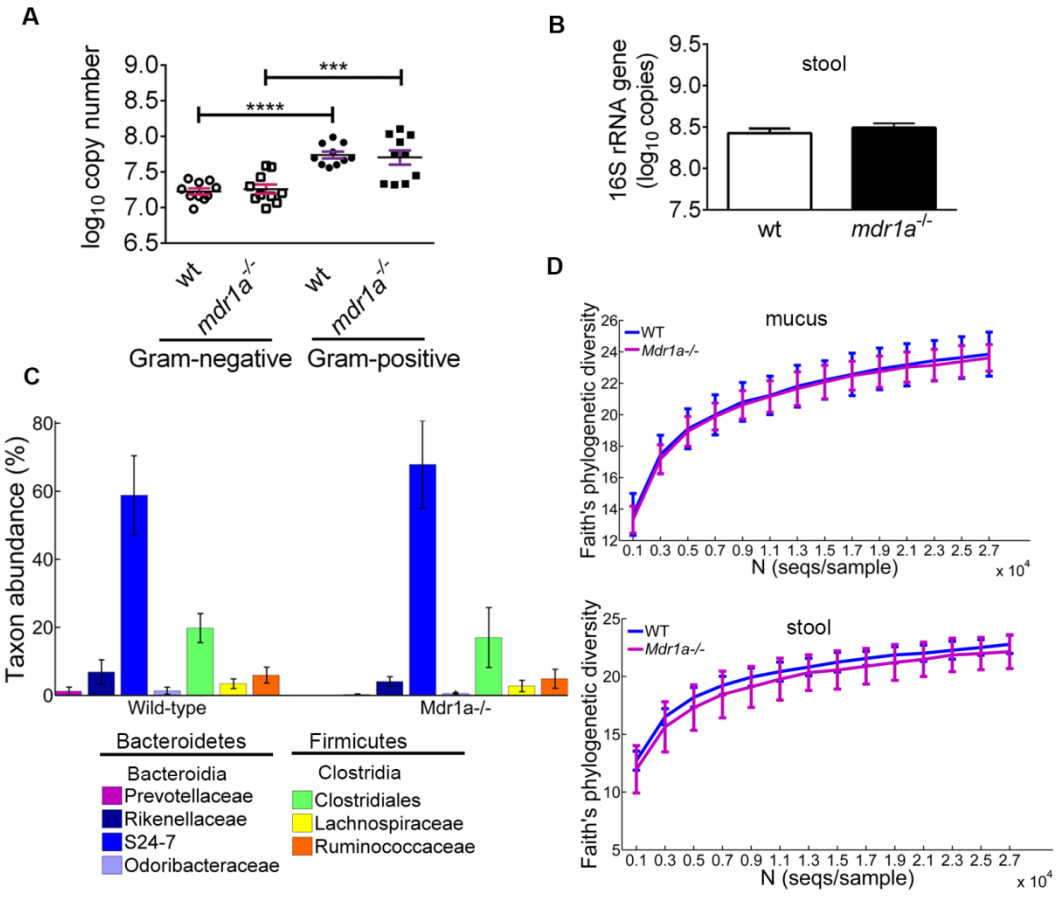
16S rRNA gene	P3_GC-341F	5’-CGC CCG CCG CGC GCG GCG GGC GGG GCG GGG GCA CGG GGG GCC TAC GGG AGG CAG CAG -3’
	P2_518R	5’-ATT ACC GCG GCT GCT GG-3’
Eubacterial (universal)	F	5’-ACTCCTACGGGAGGCAGCAGT -3’
	R	5’-ATTACCGCGGCTGCTGGC -3’
16S rRNA gene sequencing (containing adapters)	F	5’-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3’
	R	5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3'
Gram-positive/Gram-negative	P967F	5’-CAACGCGAAGAACCTTACC-3’
	P1194R	5’-ACGTCATCCCCACCTTCC-3’
	Gram-positive probe	5’-FAM-ACGACAACCATGCACCACCTG-TAMRA-3’
	Gram-negative probe	5’-HEX-ACGACAGCCATGCAGCACCT-TAMRA-3’

Supplementary Table 2. Colitis scoring table.

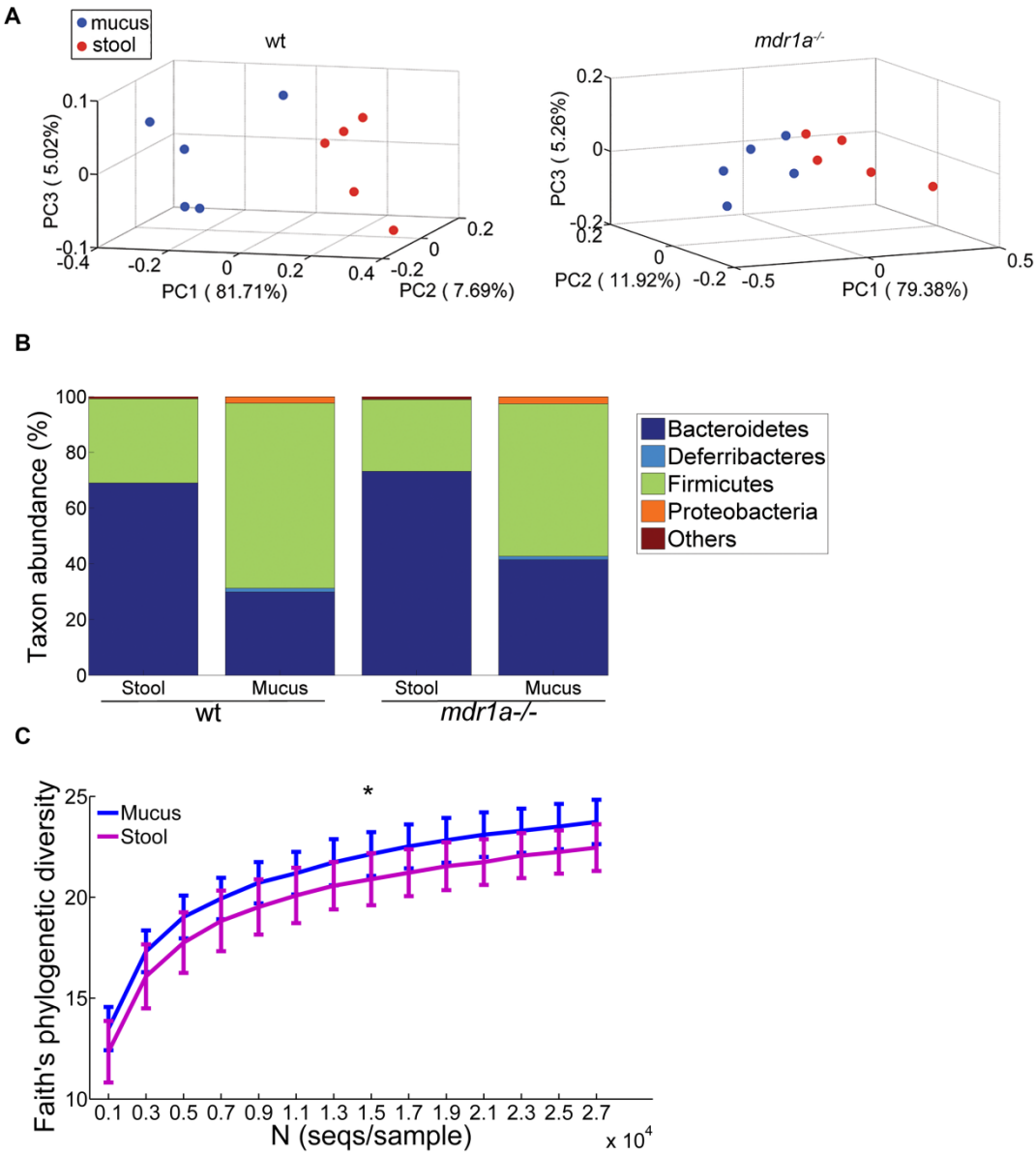
Score	Crypt length elongation	Goblet cell depletion	Muscle wall thickness	Inflammatory cell infiltration	Destruction of architecture
0	Normal crypt length	Normal goblet cell numbers	Normal thickness	No or occasional inflammatory cells in the lamina propria	Normal architecture
1	Up to 25% increase	Up to 25% decrease	Up to 25% increase	Increased number of inflammatory cells in the lamina propria	-
2	25%-50%	25%-50%	25%-50%	Higher numbers of inflammatory cells in the lamina propria	-
3	50%-75%	50%-75%	50%-75%	Confluent inflammatory cells in the lamina propria	Crypt deformity
4	>75%	>75%	>75%	Confluent Inflammatory cells in the lamina propria extending to the submucosa	Gross structural changes including ulcers



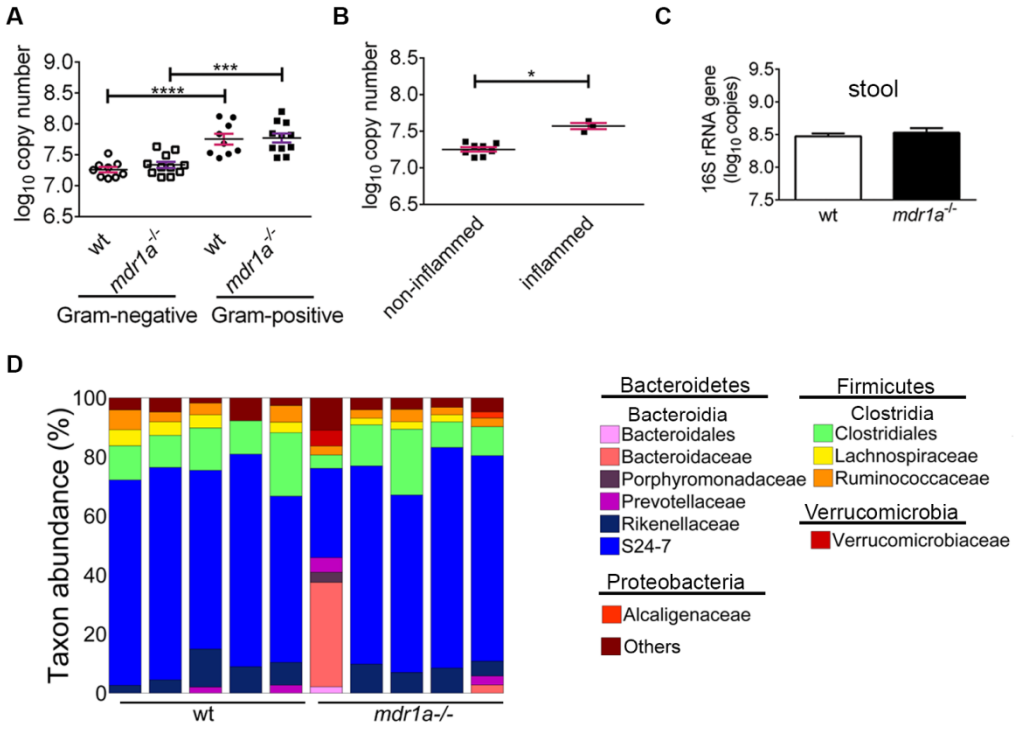
Supplemental Figure 1. Gut morphology in WT and *mdr1a*$^{-/-}$ mice. (A) Histopathology scoring of distal colon tissue samples obtained from WT and *mdr1a*$^{-/-}$ mice at 4, 12, 18 and 25 weeks. N=3-6 mice per group. Data shown are means ± SD. **(B)** Representative goblet cell stained sections from WT and *mdr1a*$^{-/-}$ mice at 6 and 18 weeks. Scale bar =100µm. **(C)** Representative H&E section from the distal colon of an inflamed *mdr1a*$^{-/-}$ mouse (18 weeks). Arrow indicates inflammatory cell infiltration. Scale bar=200µm. **(D)** Goblet cell numbers per crypt and **(E)** per area, **(F)** crypt length and **(G)** muscle wall thickness of WT and *mdr1a*$^{-/-}$ mice at 6 and 18 weeks (n=6-12). Data shown are means ± SEM. **, $P < 0.01$ by Kruskal-Wallis test with Dunn's multiple comparisons test **(A, D –G)**. **, $P < 0.01$ by unpaired t test with Welch's correction at 18-week time point **(E)**.



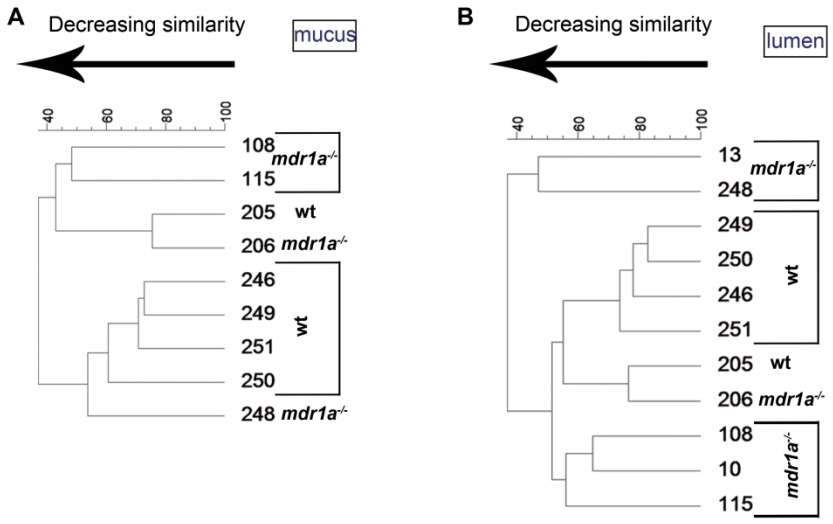
Supplemental Figure 2. Faecal microbial configuration is similar in WT and *mdr1a*$^{-/-}$ animals at 6 weeks. (A) Balance of Gram-positive and Gram-negative bacteria in faecal microbial communities (n=10 mice per genotype). **(B)** Bacterial load from stool samples was quantified based on 16S rRNA gene copy number. N=5 mice per genotype; data are from 3 littermate groups. **(C)** Bar graph showing the relative abundance of faecal microbiota at the family level (n=5/group). Where sequences could not be taxonomically assigned to the family level, the closest relative was given. Clostridiales corresponds to order level. The category 'others' represents bacterial families with abundance below 0.1%. **(D)** Rarefaction curves of phylogenetic diversity between WT and *mdr1a*$^{-/-}$ in mucus and stools (n=5/group). Data shown are means ± SEM in **(A-B)** and means ± SD in **(C-D)**. ***, $p < 0.001$; ****, $p < 0.0001$ as determined by one-way ANOVA with Tukey's multiple comparisons test **(A)**, Mann-Whitney test **(B)**, Wilcoxon rank-sum test **(C)** and Monte Carlo permutations followed by FDR correction **(D)**.



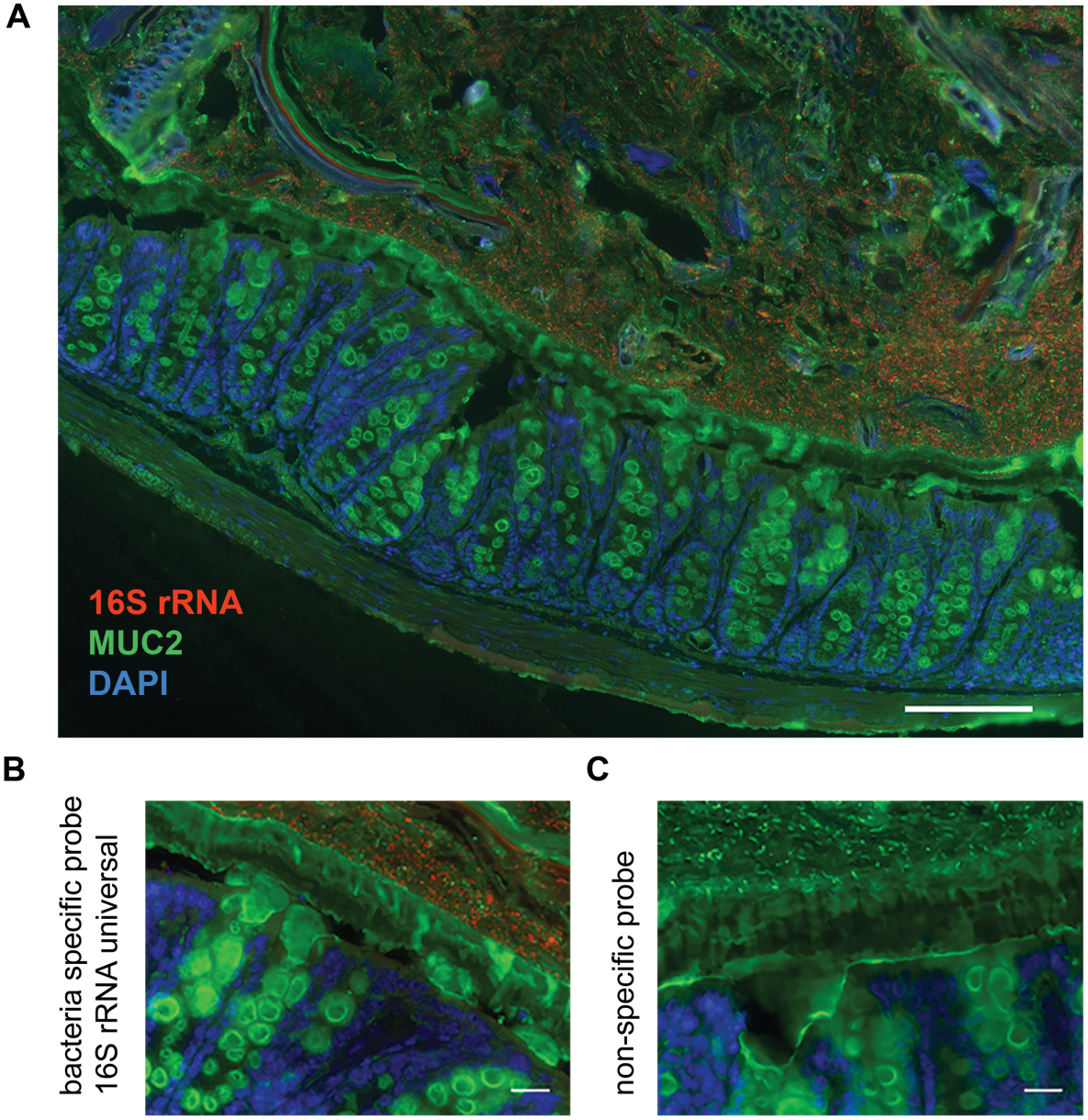
Supplemental Figure 3. Microbial community composition diversifies mucus from faecal matter. (A) Principal Coordinate Analysis (PCoA) plot of the weighted UniFrac distance comparing microbial assemblages in mucus and stool between *mdr1a*$^{-/-}$ and WT littermates at 6 weeks showed differential clustering of mucus versus faecal bacteria. Significance of segregation was tested with PERMANOVA based on 1000 permutations, P=0.029 for WT and P=0.017 for *mdr1a*$^{-/-}$. **(B)** Mean relative abundance of the main bacterial taxa at the phylum level in WT and *mdr1a*$^{-/-}$ animals at 6 weeks (n=5/group). The category 'others' represents low abundance bacterial classes (< 0.1%). **(C)** Rarefaction curves using Faith's phylogenetic measure of alpha diversity in mucus and stool samples from both WT and *mdr1a*$^{-/-}$ at 6 weeks (n=5/group). Data shown are means ± SD. *, P< 0.05 as determined by Monte Carlo permutations followed by FDR correction **(C)**.



Supplemental Figure 4. Differences in microbial composition in the stools of WT and *mdr1a*$^{-/-}$ mice start to appear on disease onset. (A) Balance of Gram-positive and Gram-negative bacteria in the faecal microbial populations (n=10 mice/genotype) at 18 weeks. **(B)** Altered balance of Gram-negative bacteria in the stools of inflamed *mdr1a*$^{-/-}$ mice compared with disease-free *mdr1a*$^{-/-}$ (n=3-7 mice) at 18 weeks. **(C)** Quantification of faecal bacteria based on 16S rRNA gene copy number at 18 weeks. N=5 mice/genotype from 5 littermate groups. **(D)** Distribution of bacterial taxa at the family level in faeces at 18 weeks (n=5/genotype). Where sequences could not be taxonomically assigned to the family level, the closest relative was given. The category 'others' represents bacterial families with abundance below 0.2%. Bacteroidales and Clostridiales correspond to order. The first and the last column of taxon abundance in the *mdr1a*$^{-/-}$ group derive from animals that have progressed to colitis (severe inflammation) compared to the other animals in the *mdr1a*$^{-/-}$ group. Data shown are means ± SEM. *, $p < 0.05$; ***, $p < 0.001$; ****, $p < 0.0001$ as determined by one-way ANOVA with Tukey's multiple comparisons test **(A)** and Mann-Whitney test **(B, C)**.



Supplemental Figure 5. Cluster analysis of microbial fingerprints at 18 weeks. The mucus **(A)** and faecal **(B)** bacteria from WT and *mdr1a*$^{-/-}$ mice were analysed. *Mdr1a*$^{-/-}$ mice displayed moderate and severe inflammation. Mice exhibiting severe inflammation (mice id: 248 and 13) formed a separate cluster compared with other age-matched *mdr1a*$^{-/-}$ mice. The percent of similarity between clusters is shown.



Supplemental Figure 6. Spatial separation between microbiota and the intestinal epithelium in the distal colon. Distal colon sections were hybridized with a universal 16S rRNA gene probe-EUB338 (red); mucus (green) was stained with an anti-MUC2 specific antiserum and sections were counterstained with DAPI (blue). Samples taken from a WT animal are shown at 10x magnification **(A)** and 40X magnification **(B)**. **(C)** FISH analysis of distal colon tissue using a non-specific probe (negative control) that recognizes the fungus *Cryptococcus* (red). Scale bar= 100 μm in **(A)** and 20 μm in **(B)** and **(C)**.

Chapter 3

Stability in metabolite and metagenome profiles before the onset of colitis-type inflammation

M. Glymenaki[1], A. Barnes[2], S. O' Hagan[3], G. Warhurst[4], A.J. McBain[1], I.D. Wilson[5], D.B. Kell[3], K.J. Else[1], S.M. Cruickshank[1]*

[1]Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[2]Shimadzu Corporation, Manchester, UK

[3]School of Chemistry, University of Manchester, Manchester, UK; Manchester Institute of Biotechnology, University of Manchester, Manchester, UK

[4]Infection, Injury and Inflammation Research Group, Division of Medicine and Neurosciences, University of Manchester and Salford Royal Hospitals NHS Trust, Salford, UK

[5]Division of Computational and Systems Medicine, Department of Surgery and Cancer, Imperial College London, London, UK

***Correspondence to:** Sheena Cruickshank, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

sheena.cruickshank@manchester.ac.uk

Author contributions

Maria Glymenaki: Designed the study, collected data, performed experiments, analyzed data and wrote the manuscript.

Alan Barnes: Performed metabolomics experiments, analyzed data, wrote LC-MS methodology and analysis, and edited and critically reviewed the manuscript.

Stephen O' Hagan: Implemented advanced metabolite software analysis and wrote relevant methodology, edited and critically reviewed the manuscript.

Geoffrey Warhurst: Conceived and designed the study, supervised the study, edited and critically reviewed the manuscript.

Andrew McBain: Conceived and designed the study, supervised the study, edited and critically reviewed the manuscript.

Ian Wilson: Conceived and designed the study, supervised the study, edited and critically reviewed the manuscript.

Doug Kell: Supervised the study, edited and critically reviewed the manuscript.

Kathryn Else: Conceived and designed the study, supervised the study, edited and critically reviewed the manuscript.

Sheena Cruickshank: Conceived and designed the study, supervised the study, edited and critically reviewed the manuscript.

3.1 Abstract

Inflammatory bowel disease (IBD) is associated with alterations in microbial composition and metabolism, but it is unclear whether this is correlative or causative, as current studies have mainly focused on changes after the onset of disease. We have previously shown differences in gut microbiota composition preceding colitis-induced inflammation. Changes were initially confined to the mucus layer and were evident in stools only after colitis development. In the present study we aimed to investigate whether differences in predicted microbial gene content and endogenous metabolite profiles accompanied such changes in bacterial composition. We examined the functional potential of mucus and stool microbial communities in the *mdr1a*$^{-/-}$ mouse model of colitis and their respective wild-type littermate controls using PICRUSt on 16S rRNA sequencing data. Our findings indicate that despite changes in microbial communities, microbial functional pathways are stable prior to the development of mucosal inflammation. In support of these findings, targeted and untargeted analysis in urine samples showed that metabolite profiles are unaffected by intestinal inflammation. Previously published metabolic markers of IBD showed no difference between *mdr1a*$^{-/-}$ mice and controls at stages preceding and during inflammation. However, metabolic profiles discriminated colitis-prone *mdr1a*$^{-/-}$ genotype from wild-type. Our results indicate resilience of the metabolic network regarding inflammation-related changes and suggest that genotype-differentiating metabolites could have value in predicting risk of IBD with a potential clinical use in at least a subset of patients carrying MDR1A polymorphisms.

3.2Introduction

Inflammatory bowel disease (IBD), which includes Crohn's disease (CD) and ulcerative colitis (UC), is associated with an overreacting immune response and alterations in gut microbial communities referred to as dysbiosis (Khor et al., 2011; Saleh and Elson, 2011). Dysbiosis in IBD is characterized by decreased bacterial diversity and an imbalanced microbial composition (Huttenhower et al., 2014; Manichanh et al., 2012; Sartor, 2008). Enterobacteriaceae are enriched, whereas clades IV and XIVa Clostridia and members of Bacteroidetes are reduced during IBD development (Gevers et al., 2014; Huttenhower et al., 2014; Manichanh et al., 2012).

Perturbations in microbial gene content abundance and expression occur as a consequence of the underlying IBD-associated dysbiosis and have been reported in IBD and experimental models of colitis (Davenport et al., 2014; Ilott et al., 2016; Morgan et al., 2012; Rooks et al., 2014; Schwab et al., 2014). Microbial gene functions related to oxidative stress resistance and nutrient transport were increased in colitis at the expense of basic biosynthetic processes such as amino acid biosynthesis, thus indicating alterations in energy metabolism within the intestinal microbiota during IBD (Ilott et al., 2016; Morgan et al., 2012; Rooks et al., 2014). To date, studies on the gene functional profile of gut microbial communities have focused on changes during active inflammation or remission and thus they may be occurring as a consequence of inflammation, while there is little information on potential changes preceding inflammation that may be causative.

Changes in the taxonomic composition of microbial species or their activities may impact on the metabolic processes in the colon, leading to an altered metabolite profile (De Preter and Verbeke, 2013; Le Gall et al., 2011; Li et al., 2008; Williams et al., 2009). Metabolite profiling studies using a range of biofluids such as faecal water, urine or serum have been used to differentiate IBD patients from healthy individuals (Le Gall et al., 2011; Marchesi et al., 2007; Schicho et al., 2012; Williams et al., 2009; Williams et al., 2012). In particular, urinary metabolites reflect endogenous metabolites produced by host metabolism as well as metabolic products of bacterial metabolism and host-bacterial co-metabolism such as hippurate (De Preter and Verbeke, 2013; Lees et al., 2013). Hence, the analysis of

metabolites in urine offers a means by which systemic changes seen before and during IBD can be investigated.

We have previously shown that changes in gut microbiota profiles in the mucus but not in faeces precede onset of inflammation in colitis-prone mice (Glymenaki et al., 2016). We utilized the *mdr1a*$^{-/-}$ spontaneous model of colitis, because it has an intact immune system (Panwala et al., 1998; Resta-Lenert et al., 2005) and polymorphisms of this gene are linked with increased susceptibility to UC in humans (Annese et al., 2006; Schwab et al., 2003). Based on our previous findings, we hypothesized that shifts in mucus microbial communities may correlate with changed function and altered metabolite profiles. To assess the functional profile of microbial communities in both mucus and faeces, we performed an *in silico* analysis of 16S rRNA gene sequencing data coupled with reference genomes. As a subset of the encoded functions of microbial communities is expressed at any given time, we further employed a metabonomic approach using urine samples with the aim to detect potential metabolite changes that strongly influence the host-microbiota crosstalk prior to inflammation. Our findings indicate stability of microbial gene coding potential and endogenous metabolites prior to the development of mucosal inflammation and suggest resilience of metabolism before disease outbreak.

3.3 Materials and methods

3.3.1 Maintenance of animals

Mdr1a$^{-/-}$ mice (FVB.129P2-*Abcb1a*^{tm1Bor} N7) (Schinkel et al., 1994) and control FVB mice (Taconic Farms, NY, USA) were crossbred to generate F1 heterozygotes, which in turn were crossbred to generate F2 littermate controls. Mice were given autoclaved standard chow and sterile acidified water (pH=3.2) *ad libitum*. *Mdr1a*$^{-/-}$ and wild-type (WT) control males were maintained under co-housing conditions to ensure shared microbiota. All animals were kept under specific, pathogen-free (SPF) conditions at the University of Manchester and experiments were performed according to the regulations issued by the Home Office under amended ASPA, 2012.

3.3.2 Histology and colitis scoring

Distal colon tissue was fixed, paraffin-embedded and stained with haematoxylin and eosin or alcian blue dye for colitis scoring as previously described (Glymenaki et al., 2016). In brief, the sum of scores for crypt length elongation (score 0-4), goblet cell depletion (score 0-4), muscle wall thickness (score 0-4), inflammatory cell infiltration (score 0-4) and destruction of architecture (score 0 or 3-4) was calculated.

3.3.3 Isolation of bacterial genomic DNA

Bacterial genomic DNA was isolated from faecal and mucus samples as previously reported (Glymenaki et al., 2016) and extracted using the QIAamp® DNA Stool Mini Kit (Qiagen, Manchester, UK) with an additional bead beating step.

3.3.4 16S rRNA gene sequencing analysis

The V3 and V4 variable regions of the 16S rRNA gene were PCR amplified for sequencing on the Illumina MiSeq platform according to manufacturer's guidelines as previously

described (Glymenaki et al., 2016). Sequences were submitted to European Bioinformatics Institute (EBI) for quality filtering (Hunter et al., 2014) and were further processed using the Quantitative Insights Into Microbial Ecology (QIIME) pipeline v.1.9.0 (Caporaso et al., 2010). They were assigned to operational taxonomic units (OTUs) using a closed-reference OTU picking strategy (Edgar, 2010) and taxonomically classified using the Greengenes database filtered at 97% identity (McDonald et al., 2012; Wang et al., 2007).

PICRUSt (phylogenetic investigation of communities by reconstruction of unobserved states) was then applied on the Greengenes picked OTU table to generate metagenomic data and derive KEGG (Kyoto Encyclopaedia of Genes and Genomes) Orthology gene abundance data (Langille et al., 2013). KEGG Orthology gene family abundances were summarized at a higher hierarchical level at pathway-level categories for easier biological interpretation. Non-microbial categories such as 'Organismal Systems' and 'Human Diseases' were excluded from further analysis. Beta diversity of rarefied KEGG pathway data was calculated using the Bray-Curtis distance metric and visualized using Principal Coordinate Analysis (PCoA) in Matlab (MathWorks, MA, USA). KEGG pathway abundance data between groups were compared using group_significance.py in QIIME (Caporaso et al., 2010). Metagenomic data were also analysed using STatistical Analysis of Metagenomic Profiles (STAMP) software (Parks and Beiko, 2010). To examine PICRUSt's predictive accuracy, the weighted nearest sequenced taxon index (NSTI) values were calculated (Supplementary Table 1).

3.3.5 Urine sample collection and preparation

Urine samples were collected from mice at designated time points in clean autoclaved cages or by injection in the bladder during culling and stored at -80˚C (n=18 WT 6 weeks, n=18 *mdr1a*$^{-/-}$ 6 weeks, n=17 WT 18 weeks, n=12 *mdr1a*$^{-/-}$ 18 weeks). Urine preparation was performed according to a previously published method (Wang et al., 2009). In brief, samples were thawed at room temperature and centrifuged at 10,000g for 5 min. Then, 10μL urine was diluted with 40μL water (or by 1:4 v:v in samples of less than 10μL). 5 μL of each diluted sample was injected onto to the analytical column. Quality control (QC) samples were prepared by pooling aliquots of 10μL of each sample.

3.3.6 Ultra high performance liquid chromatography- mass spectrometry (UHPLC-MS) metabolite analysis

5µL aliquots of each sample (maintained at 4°C in the autosampler) were injected for separation by UHPLC, using gradient elution, with a Nexera LC system (Shimadzu Corporation, Kyoto, Japan) on to the analytical column, Acquity HSS T3 1.8um C18 (100 x 2.1mm) at a flow rate of 0.4 mL/min, with the column maintained at 40 °C. The chromatographic system used a binary solvent system according to a previously published method (Andersen et al., 2013), delivered as a gradient of solvent A (water containing 0.1% formic acid) and solvent B (acetonitrile containing 0.1% formic acid). The gradient conditions were: 2% B (2 min), to 35% B (12 min), to 100% (18 min) held to 23.5 min, re-equilibration time was 5 min. Samples were analysed by LC-MS using a quadrupole ion-trap time-of-flight mass spectrometer (LCMS-IT-TOF, Shimadzu Corporation, Kyoto, Japan) equipped with an electrospray source in both positive and negative mode (polarity switching time of 100 msecs). The mass range measured was m/z 60-1250 in MS mode. Mass calibration was carried out using a trifluoroacetic acid sodium solution (2.5 mmol/L) from 50 to 1000 Da. Other instrument parameters included: ion source temperature of 250 °C, heated capillary temperature of 230 °C, electrospray voltage of 4.5 kV, electrospray nebulization gas flow was 1.5 L/minute, detector voltage 1.7 kV. Data acquisition and processing used software LCMS Solution (version 3.8, Shimadzu Corporation, Kyoto, Japan).

3.3.7 LC-MS data analysis and processing

Profiling Solution software (version 1.1, Shimadzu Corporation, Kyoto, Japan) was used to create an aligned data array of retention time, m/z and intensity data for both positive and negative ion data as previously described (Loftus et al., 2011). Sample data acquisition was performed in four acquisition events to provide optimum sensitivity in the lower mass range whilst not saturating in the higher mass range: positive mode m/z 60-200, positive mode m/z 140-1250, negative mode m/z 60-200, negative mode m/z 140-1250. Data was combined into a single data array containing 3129 ions in which no data point was excluded. Profiling Solution software generated an aligned data array of both positive and negative ion MS data, which was subsequently exported to SIMCA-P (version 14, Umetrics, MKS Instruments Inc., Sweden) for principal components analysis (PCA). Following noise reduction thresholding, a

data array was processed using SIMCA-P and scaled to unit variance (the base weight is computed as 1/sdj, with sdj is the standard deviation of variable j computed around the mean). No variable was excluded in this analysis. Metabolite features were statistically tested for their quantitative significance by considering the reproducibility of the ion signal in the pooled QC sample. Profiling Solution processing parameters included: 15mDa ion bin m/z tolerance, 0.2 min ion bin RT tolerance, noise threshold 100000. Pooled QA/QC parameters: 80% ions required from all QC samples, better than 30% relative standard deviation (RSD) peak area precision, better than 5% RSD retention time precision.

3.3.8 Multivariate statistics on LC-MS data

LC-MS data of urine samples were subjected to multivariate statistical analysis using KNIME (Berthold et al., 2008; Mazanetz et al., 2012; O'Hagan and Kell, 2015) and R R (http://cran.r-project.org). Pre-processing involved removing QC and "singleton" data, followed by application of a correlation filter for removal of correlated features (threshold = 0.98) and Z - scores normalization ($Z = (x - \mu)/\sigma$). In addition to PCA, multivariate regression was applied, as it correlates independent variables in matrix X (i.e. metabolomics data) to corresponding dependent variables in matrix Y (i.e. groups, classes) (Barker and Rayens, 2003). Partial least squares (PLS) regression and random forests (RF) regression was used to construct predictive regression models for better discrimination of sample groups (Breiman, 2001; O'Hagan and Kell, 2015). PLS- linear discriminant analysis (PLS-LDA), a specific form of PLS regression was used in the case of dependent categorical variables. The performance of PLS-regression model was tested using cross-validation with the 'leave one out' method. All data were used for training in the model, which potentially does not rule out potential over-fitting of the data.

3.3.9 Feature permutation method

To identify mass ions, also referred to as features, that contribute to differential classification of sample groups, mass ions were permutated and classification cross entropy was calculated. An RF classifier was used in a simplified method as follows. First, the mean cross entropy and sigma for 100 repeats of the unpermuted RF data were calculated and then each feature column was permuted and the cross entropy was recalculated. Repeats on each permutation

were performed in a loop to derive the mean cross entropy. The difference between the mean permutated result and mean unpermuted result was measured and if the difference was found to be greater than 1.96 * sigma, it was considered significant with a P value equal to 0.95. The Storey multiple comparison correction method was further applied (Storey and Tibshirani, 2003). When a feature is permuted, the cross entropy will increase if it is significant meaning that classification will get worse.

When we have relatively few samples and noisy data, machine-learning methods can often pick out noise as features. To deal with this issue, re-pre-processing the data offers a way of systematic error removal. Thus, data were re-processed with a 0.3min tolerance in retention time alignment and a noise thresholding set to 1,000,000 was used. Nevertheless, this approach still generated a small number of noise ions, so a second stage analysis was applied that generated the so-called Chromatogram Matrix in which generic peak integration parameters were applied to all chromatographic peaks from ions identified in the spectral matrix. PCA and regression analysis of this data also led to similar conclusions as the initial analysis before re-processing.

3.3.10 Metabolite identification

To identify biologically significant components, high mass accuracy MS and MSn fragment ion information was used to identify the most likely candidate formula (mass accuracy of LCMS-IT-TOF typically better than 5ppm). Authentic chemical standards were also purchased for metabolite identification. Endogenous creatine and histidine were used as internal standard compounds for data normalisation. Data acquisition and processing used software LCMS Solution (version 3.8). Analysis of pooled QC data of known endogenous metabolites showed acceptable precision with better than 30% RSD (peak area ratio) for most of the targeted metabolites with exception to leucine and isoleucine that are challenging to separate by reversed phase chromatography (Supplementary Table 2).

3.3.11 Statistical analysis

Statistical analysis was performed using R and GraphPad Prism 6 (GraphPad software, CA, USA). The vegan package in R was also used for carrying adonis test. Normally distributed data were analysed by unpaired t test. Data that did not exhibit a normal distribution were analysed using Mann-Whitney test or the nonparametric Kruskal-Wallis test with Dunn's posttest as appropriate to the number of comparisons being made. $P<0.05$ was considered as statistically significant (*$P<0.05$, **$P<0.01$). All P-values were corrected for multiple hypothesis testing using the Benjamini- Hochberg false discovery rate (FDR) method unless otherwise stated (Benjamini and Hochberg, 1995).

3.4Results

3.4.1 Gut microbial functional profiles remain stable prior to inflammation

We have previously reported that alterations in microbial composition are found in mucus but not in the stools of *mdr1a*$^{-/-}$ mice in contrast to WT controls at 6 weeks of age, approximately 12 weeks before signs of overt intestinal inflammation are present (Glymenaki et al., 2016). However, by the later time point of 18 weeks of age, signs of inflammation were apparent and changes in microbial composition were detected in both the mucus and stools (Glymenaki et al., 2016). To investigate the microbial functional potential in mucus and stools prior and during the development of colitis, we applied PICRUSt on 16S rRNA gene amplicon sequencing data (Langille et al., 2013). PICRUSt infers the approximate gene content of detected phylotypes. PICRUSt-predicted gene families represented by KEGG Orthology groups were binned into KEGG metabolic pathways to assess the similarity of the functional state of the microbiome. PCoA analysis using Bray-Curtis distance demonstrated that samples were separated mainly according to their location (i.e. mucus or stool), which accounted for 93.57% of the overall variation between samples (Adonis test; R2=0.52, P=0.001) (Figure 1). Samples were further stratified according to their sampling location (i.e. mucus or stool) to examine possible differentiation across age without the confounding effect of location (Supplementary Figure 1A-B). However, we observed no clear clustering of samples depending on age. In a similar mode, genotype was not able to separate samples obtained at 6 and 18 weeks from mucus and stools (Supplementary Figure 2A-D).

The relative abundance of KEGG metabolic pathways was similar between WT and *mdr1a*$^{-/-}$ mice before the onset of intestinal inflammation irrespective of microbial habitat, mucus or stools (Supplementary Figure 3A-B). Functional pathways' abundance remained unmodified even when signs of inflammation started (Supplementary Figure 4A-B). Thus, functional categories implicated in metabolism, genetic information processing, environmental information processing and cellular processes presented a steady pattern across time independent of genotype. Notably, a great proportion of the microbial functions were poorly characterized or unclassified. Differences in KEGG metabolic pathways were only observed between mucus and stool microbial communities (Supplementary Figure 5). Collectively, these data suggest that microbial functional potential is remarkably stable prior to inflammation in the gut.

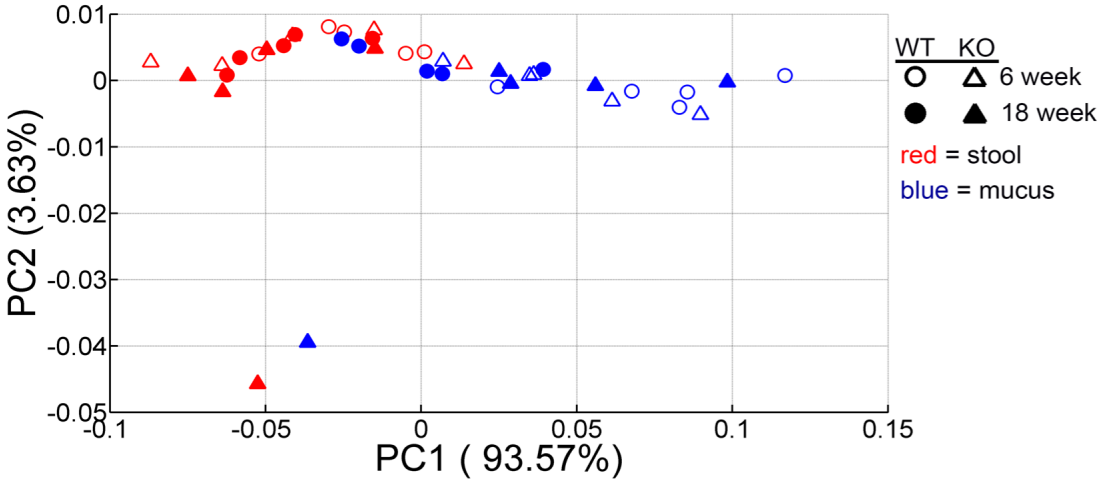


Figure 1. Stability of the microbial functional potential prior to the development of colitis. Microbial genes were inferred by PICRUSt from 16S rRNA gene sequences and assigned to functional pathways as organized in KEGG database. Principal Coordinate Analysis (PCoA) plot of Bray-Curtis distance comparing microbial functional profiles between *mdr1a*$^{-/-}$ and WT littermates at 6 and 18 weeks showed clustering of samples according to sampling location (i.e. mucus and stools). Bray-Curtis distances were calculated based on KEGG pathway abundance values. WT mice are shown in circles and knockout (KO) mice in triangles; open symbols correspond to 6 weeks whereas filled ones to 18 weeks. Mucus is depicted in blue and stool in red.

3.4.2 Targeted metabolomics reveals no difference in identified IBD marker metabolites

Dysbiosis in IBD is reportedly associated with disruption of host-microbiota dialogue with an impact on host immune system and metabolism resulting in loss of homeostasis (Kamada et al., 2013). To examine putative effects of previously reported altered microbial composition on the metabolite profiles of *mdr1a*$^{-/-}$ mice versus WT controls, we performed LC-MS analysis of urinary samples collected at 6 and 18 weeks. We selected urine, as it provides a pool of endogenous host metabolites that also reflect bacterial metabolism (De Preter and Verbeke, 2013; Lees et al., 2013). Previously published endogenous metabolites known to significantly change in human and murine IBD studies were targeted in order to detect if there were similar differences in *mdr1a*$^{-/-}$ mice before and during the onset of inflammation. Hippurate has been considered a significant marker metabolite in human studies, as it is known to decrease in UC and CD patients (Schicho et al., 2012; Williams et al., 2009).

To account for sample-to-sample differences, endogenous creatine or histidine was used as an internal standard compound for all metabolites. No significant differences in urinary metabolites including hippurate could be identified between 6 and 18-week WT and *mdr1a*$^{-/-}$ mice with the exception of arginine, which was reduced over time in both *mdr1a*$^{-/-}$ ($P=0.0012$ when compared with 18-week old WT) and WT mice ($P=0.0113$ when compared with WT at 18 weeks) (Figure 2). Further to this, levels of previously reported influential metabolites were comparable between WT and *mdr1a*$^{-/-}$ mice at 18 weeks, when signs of inflammation began to appear (Supplementary Figure 6).

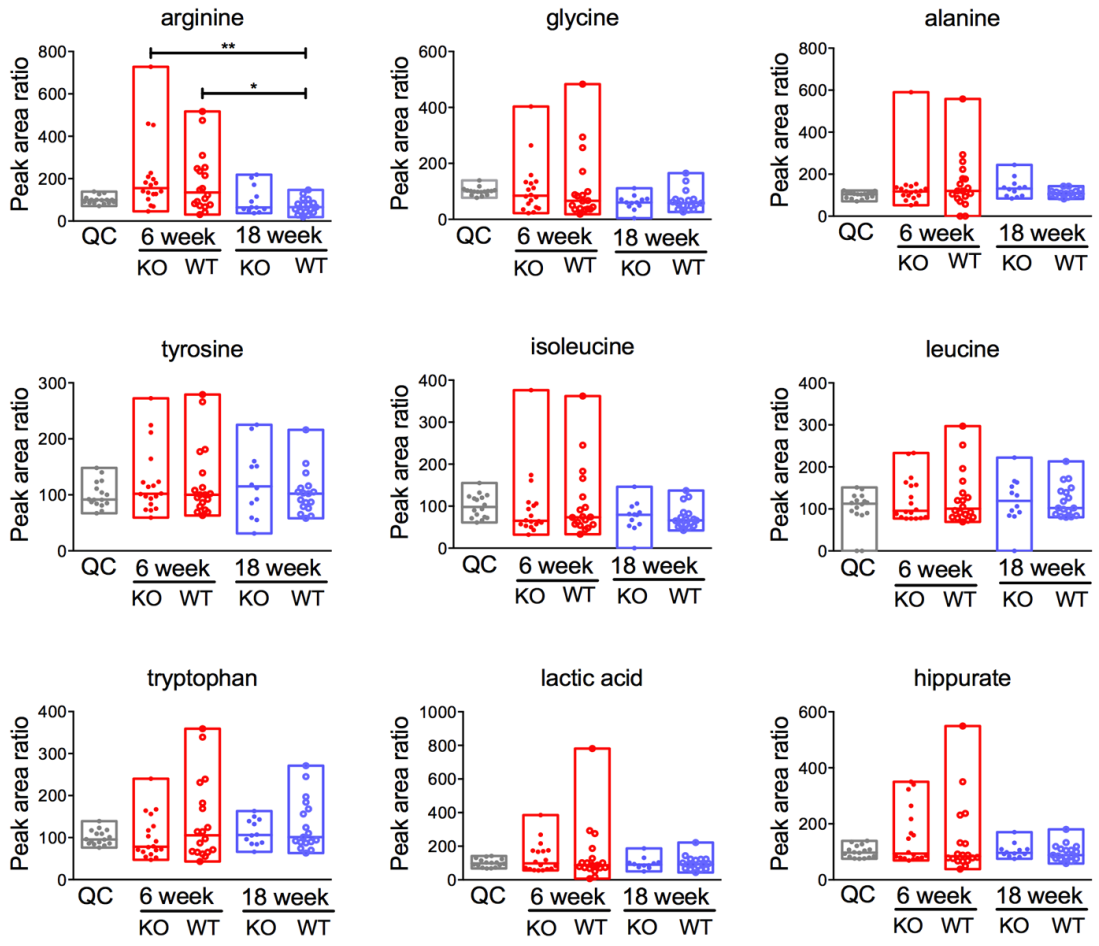


Figure 2. IBD marker metabolite levels in urinary samples from WT and *mdr1a*$^{-/-}$ mice at 6 and 18 weeks. No differences were identified in the levels of metabolites when comparing samples of the same genotype across time or between WT and KO samples at 6 weeks or 18 weeks. Creatine was used as an internal control. The median is shown as a line and bars capture the minimum and maximum. *, P<0.05; **, P<0.01 as determined by Kruskal-Wallis test with Dunn's multiple comparisons test.

3.4.3 Metabolite profiling discriminates colitis-prone genotype prior to the onset of inflammation

Since targeted analysis looks only at a small fraction of the urinary metabolite profile, we also performed an untargeted analysis to examine a wider pool of metabolites. PCA analysis was first carried out on LC-MS mass ion data to provide an overview of the variations between WT and *mdr1a*$^{-/-}$ mice at 6 and 18 weeks (Figure 3). PCA revealed a pattern of variation associated with age reflected by differential clustering of 6-week and 18-week samples, which explained 41.7% of the variation.

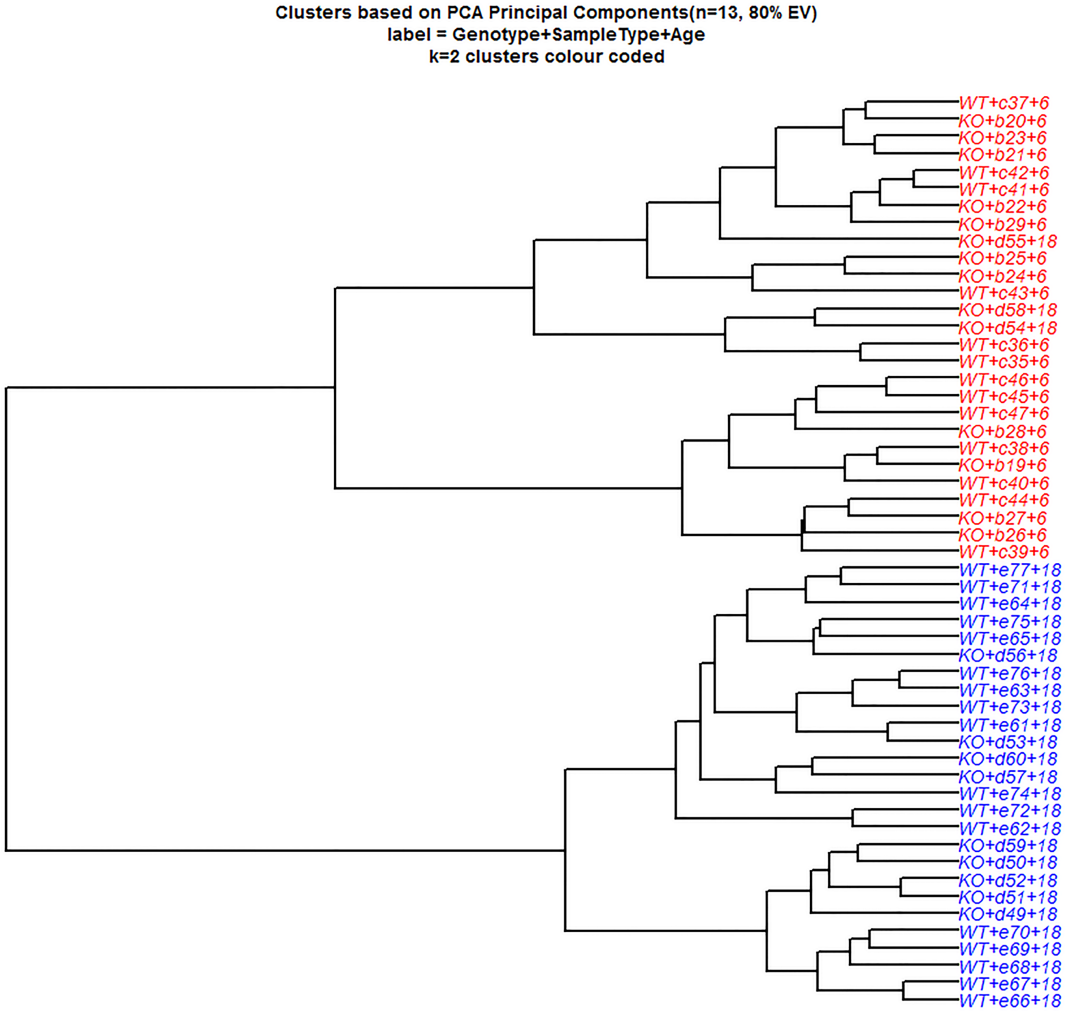


Figure 3. Stability of urinary metabolite profiles in colitis-prone *mdr1a*$^{-/-}$ animals during onset of inflammation. Principal Components Analysis (PCA) of mass ions measured by LC-MS for urine samples from *mdr1a*$^{-/-}$ mice and WT littermate controls at 6 and 18 weeks of age. Clustering of PCA data separated samples according to age, which accounted for 41.7% of the total variance. Data are shown in a dendrogram.

PCA is an unsupervised method for assessing variation among samples that ignores the correlation between LC-MS mass ions and sample characteristics such as age, genotype or colitis scoring. For this reason, we applied PLS-LDA that is a supervised classification method and considers the correlation of LC-MS variables to the class membership of samples (for example WT or *mdr1a*$^{-/-}$ in the case of genotype) to maximize their separation. PLS-LDA analysis on urinary metabolite profiles exhibited a better separation of groups based on genotype compared with PCA (Figure 4). LC-MS profiles from WT and *mdr1a*$^{-/-}$ mice were differentially clustered with an accuracy of 80% in predictive regression models (Table 1). To discover the metabolites, which contributed to the discrimination of urinary metabolic profiles between WT and knockout (KO) animals, we performed permutational analysis. The mean cross entropy changes on feature permutation found differences based on genotype; however, study limitations prevent assigning metabolite biomarkers from this data (Supplementary Figure 7 and 8).

PLS regression and RF regression were used to build models in order to identify changes in metabolite profiles that could predict varying scores of colitis (Supplementary Figure 9). We found that metabolic variations were not correlated with different colitis scores, as they deviated from the identity line. These data suggest that metabolic changes are independent of ongoing inflammation, but rather they are related to genotype. As *mdr1a*$^{-/-}$ mice are prone to colitis development as they age (Panwala et al., 1998; Resta-Lenert et al., 2005), differences observed in metabolites not affiliated with inflammation could indicate a predictive risk of IBD.

Table 1. Prediction statistics for the PLS-LDA[a] obtained from LC-MS data.

row ID	TP[b]	FP[c]	TN[d]	FN[e]	Recall	Precision	Sensitivity	Specificity
Mdr1a$^{-/-}$	8	1	9	4	0.67	0.89	0.67	0.90
WT	9	4	8	1	0.90	0.69	0.90	0.67

[a] 5 latent variables were used for the PLS-LDA model, [b] True Positives, [c] False Positives, [d] True Negatives, [e] False Negatives. F-measure for *mdr1a*$^{-/-}$ is 0.76 and for WT is 0.78.Overall accuracy is 0.77 and Cohen's kappa 0.55. All LC-MS mass data were used for the PLS-LDA model. No sample was excluded from analysis.

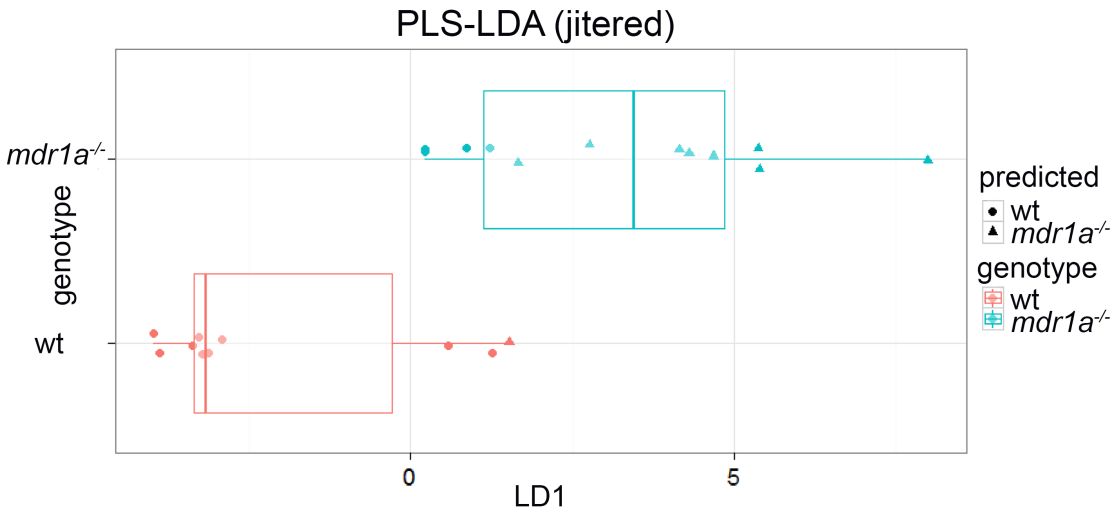


Figure 4. Partial least squares-linear discriminant analysis (PLS-DA) discriminates urinary metabolite profiles from *mdr1a*$^{-/-}$ and WT mice based on genotype. Boxplot of genotype predictions for WT and *mdr1a*$^{-/-}$ mice, where the central rectangle spans the first quartile to the third quartile with median shown as a line; whiskers above and below the box represent the maximum and minimum respectively. True *mdr1a*$^{-/-}$ are in blue, predicted *mdr1a*$^{-/-}$ are triangles; true WT are in red, predicted WT are in circles. Therefore, blue circles are false negatives and red triangles are false positives.

3.5Discussion

IBD is associated with an altered microbiota composition, which may also be translated to an altered metabolic activity in the gut (Kostic et al., 2014; Morgan et al., 2012). Here we show that microbial functional pathways remain stable before the onset of intestinal inflammation in colitis-prone mice. In our previous findings we observed changes in the overall taxonomic composition in mucus of *mdr1a*$^{-/-}$ mice 12 weeks prior to inflammation, and in both mucus and faeces during colitis development at 18 weeks (Glymenaki et al., 2016). Therefore, our results indicate that changes in predicted microbial functional gene potential do not accompany altered bacterial composition. Interestingly, previous studies have demonstrated that different assemblages of microbial communities may converge to similar metabolic functions in the gut (Kostic et al., 2015; Turnbaugh et al., 2009). Differences in the clustering of microbial gene pathways were only observed between mucus and stool microbial communities, which are known to differ in microbial composition (Momozawa et al., 2011; Morgan et al., 2012).

Our data indicate a stable pattern in the functional profiling of microbial communities in *mdr1a*$^{-/-}$ mice relative to their WT counterparts at time points before and during the start of inflammation. Previous functional profiling studies of the gut microbiome have mainly focused on established disease (Davenport et al., 2014; Ilott et al., 2016; Morgan et al., 2012; Rooks et al., 2014). Microbial gene families involved in nutrient transport and oxidative stress resistance were increased during active colitis (Ilott et al., 2016; Morgan et al., 2012; Rooks et al., 2014), whereas energy metabolism and amino acid biosynthesis pathways were reduced (Morgan et al., 2012; Rooks et al., 2014). Additionally, microbial functions contributing to bacterial pathogenesis were enriched (Morgan et al., 2012; Rooks et al., 2014). These functional changes suggest an adaptation of multiple microbes to accommodate the environmental stressors present in the inflamed gut. A recent study of the colonic microbiota from healthy individuals carrying *FUT2* gene polymorphism, a CD risk allele, revealed alterations at both the compositional and functional level of the gut microbiome, which were accompanied by sub-clinical inflammation in the intestinal mucosa (Tong et al., 2014).

We inferred gene content of the gut microbiome using PICRUSt, which gives an accurate but approximate prediction based on reference microbial genomes (Langille et al., 2013).

However, potential bias may be introduced from 16S rRNA reads that failed to map to reference OTUs and the lack of a sufficient number of reference genomes. In our study, 93.56% of 16S rRNA sequences mapped to reference OTUs, but NSTI scores were high (>0.15) indicating low availability of closely related reference genomes, which is also illustrated in the proportions of poorly characterized and unclassified gene functions. Thus, we further analyzed metabolites to complement microbial functional profiling observations.

To shed light on putative metabolic changes, we performed a metabolite analysis in urine samples. Targeted metabolomic approach showed that the levels of metabolites previously reported to change in human and murine IBD studies were similar between WT and *mdr1a*$^{-/-}$ mice at both 6 and 18 weeks with the exception of arginine, which significantly decreased with age. PCA analysis of untargeted metabolites also showed an age-dependent variation. These findings are consistent with previous studies that have reported age-related effects on metabolome (Lin et al., 2009; Murdoch et al., 2008). Hippurate, which is consistently reduced in UC and CD (Schicho et al., 2012; Williams et al., 2009), displayed comparable levels among groups. Predictive models were also not able to correlate metabolite profiles to colitis scoring. Taken together, our results suggest that the overall metabolic state of the animal is not affected, pertaining to a state of homeostasis, until the onset of inflammation.

Clinical studies of IBD in humans that exhibit disease symptoms and have established disease have shown changes in metabolite profiles (Le Gall et al., 2011; Marchesi et al., 2007; Schicho et al., 2012; Williams et al., 2009; Williams et al., 2012). Further to this, studies in IL-10$^{-/-}$ mice as well as dextran sulphate sodium (DSS)-colitis induced mice have shown that differences in metabolite profiles were more profound once inflammation has progressed and colitis has developed (Lin et al., 2010; Lin et al., 2009; Martin et al., 2009; Murdoch et al., 2008; Schicho et al., 2010). Therefore, metabolic changes observed in these studies follow the development of intestinal inflammation and as a consequence the disruption of host-microbiota homeostasis (De Preter and Verbeke, 2013; Dorrestein et al., 2014). In our study, we looked at stages before the development of overt mucosal inflammation and found no metabolic changes.

Our results show that metabolite profiles classified differentially according to genotype. Metabolic composition differed between WT and m*dr1a*$^{-/-}$ mice, which progress to colitis with aging (Panwala et al., 1998; Resta-Lenert et al., 2005). In support of genotype-related

metabolic changes, a recent study on dietary supplementation has shown that *mdr1a*$^{-/-}$ mice fed a control diet had lower levels of short chain fatty acids (SCFAs) and higher levels of lactic and succinic acid in their faeces (Paturi et al., 2012). Of note, butyrate is a SCFA that contributes to the maintenance of immune tolerance (Furusawa et al., 2013; Smith et al., 2013).

Collectively, our findings demonstrate that the activity of microbial communities and urinary metabolites are remarkably stable in *mdr1a*$^{-/-}$ mice in the face of ensuing gut inflammation despite changes in mucus microbial communities. The biochemistry of the host and its associated gut microbiota seems to remain unaffected at stages preceding full disease manifestation possibly due to function redundancy in microbial and host metabolic pathways or underlying subtle changes with a minimal effect. However, metabolite profiles differ depending on genotype, indicating alterations in the metabolic network unrelated to inflammation. The fact that we did not observe changes in metabolites associated with gut inflammation suggests that these genotype-affiliated metabolites could constitute a predictive risk of IBD and warrant further investigation.

Acknowledgements

We thank N. Loftus, A. Brass and C. Knight for critical discussions on the study. M. Constantinou provided technical support in Matlab. This work was supported by a UK Biotechnology and Biological Sciences Research Council (BBSRC) studentship for MG awarded to SC, KE, AM and GW and a grant from the European Crohn's and Colitis Organization (ECCO) awarded to SC. The authors declare no conflicts of interest.

References

Andersen, M.-B.S., Reinbach, H.C., Rinnan, Å., Barri, T., Mithril, C., and Dragsted, L.O. (2013). Discovery of exposure markers in urine for Brassica-containing meals served with different protein sources by UPLC-qTOF-MS untargeted metabolomics. Metabolomics : Official journal of the Metabolomic Society *9*, 984-997.

Annese, V., Valvano, M.R., Palmieri, O., Latiano, A., Bossa, F., and Andriulli, A. (2006). Multidrug resistance 1 gene in inflammatory bowel disease: a meta-analysis. World journal of gastroenterology : WJG *12*, 3636-3644.

Barker, M., and Rayens, W. (2003). Partial least squares for discrimination. Journal of Chemometrics *17*.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Statist. Soc. *57*, 289-300.

Berthold, M.R., Cebron, N., Dill, F., Gabriel, T.R., Kötter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K., and Wiswedel, B. (2008). KNIME: The Konstanz Information Miner. In Data Analysis, Machine Learning and Applications: Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation e.V., Albert-Ludwigs-Universität Freiburg, March 7–9, 2007, C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 319-326.

Breiman, L. (2001). Random Forests. Machine Learning *45*, 5-32.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I.*, et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. Nature methods *7*, 335-336.

Davenport, M., Poles, J., Leung, J.M., Wolff, M.J., Abidi, W.M., Ullman, T., Mayer, L., Cho, I., and Loke, P. (2014). Metabolic alterations to the mucosal microbiota in inflammatory bowel disease. Inflammatory bowel diseases *20*, 723-731.

De Preter, V., and Verbeke, K. (2013). Metabolomics as a diagnostic tool in gastroenterology. World journal of gastrointestinal pharmacology and therapeutics *4*, 97-107.

Dorrestein, P.C., Mazmanian, S.K., and Knight, R. (2014). Finding the missing links among metabolites, microbes, and the host. Immunity *40*, 824-832.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics *26*, 2460-2461.

Furusawa, Y., Obata, Y., Fukuda, S., Endo, T.A., Nakato, G., Takahashi, D., Nakanishi, Y., Uetake, C., Kato, K., Kato, T.*, et al.* (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature *504*, 446-450.

Gevers, D., Kugathasan, S., Denson, L.A., Vazquez-Baeza, Y., Van Treuren, W., Ren, B., Schwager, E., Knights, D., Song, S.J., Yassour, M.*, et al.* (2014). The treatment-naive microbiome in new-onset Crohn's disease. Cell host & microbe *15*, 382-392.

Glymenaki, M., Brass, A., Warhurst, G., McBain, A.J., Else, K.J., and Cruickshank, S.M. (2016). Differences in gut mucus microbiota precede the onset of colitis-induced inflammation. Submitted.

Hunter, S., Corbett, M., Denise, H., Fraser, M., Gonzalez-Beltran, A., Hunter, C., Jones, P., Leinonen, R., McAnulla, C., Maguire, E.*, et al.* (2014). EBI metagenomics--a new resource for the analysis and archiving of metagenomic data. Nucleic acids research *42*, D600-606.

Huttenhower, C., Kostic, A.D., and Xavier, R.J. (2014). Inflammatory bowel disease as a model for translating the microbiome. Immunity *40*, 843-854.

Ilott, N.E., Bollrath, J., Danne, C., Schiering, C., Shale, M., Adelmann, K., Krausgruber, T., Heger, A., Sims, D., and Powrie, F. (2016). Defining the microbial transcriptional response to colitis through integrated host and microbiome profiling. The ISME journal.

Kamada, N., Seo, S.U., Chen, G.Y., and Nunez, G. (2013). Role of the gut microbiota in immunity and inflammatory disease. Nature reviews. Immunology *13*, 321-335.

Khor, B., Gardet, A., and Xavier, R.J. (2011). Genetics and pathogenesis of inflammatory bowel disease. Nature *474*, 307-317.

Kostic, A.D., Gevers, D., Siljander, H., Vatanen, T., Hyotylainen, T., Hamalainen, A.M., Peet, A., Tillmann, V., Poho, P., Mattila, I.*, et al.* (2015). The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. Cell host & microbe *17*, 260-273.

Kostic, A.D., Xavier, R.J., and Gevers, D. (2014). The microbiome in inflammatory bowel disease: current status and the future ahead. Gastroenterology *146*, 1489-1499.

Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R.*, et al.* (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature biotechnology *31*, 814-821.

Le Gall, G., Noor, S.O., Ridgway, K., Scovell, L., Jamieson, C., Johnson, I.T., Colquhoun, I.J., Kemsley, E.K., and Narbad, A. (2011). Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. Journal of proteome research *10*, 4208-4218.

Lees, H.J., Swann, J.R., Wilson, I.D., Nicholson, J.K., and Holmes, E. (2013). Hippurate: the natural history of a mammalian-microbial cometabolite. Journal of proteome research *12*, 1527-1546.

Li, M., Wang, B., Zhang, M., Rantalainen, M., Wang, S., Zhou, H., Zhang, Y., Shen, J., Pang, X., Zhang, M.*, et al.* (2008). Symbiotic gut microbes modulate human metabolic phenotypes. Proceedings of the National Academy of Sciences of the United States of America *105*, 2117-2122.

Lin, H.M., Barnett, M.P., Roy, N.C., Joyce, N.I., Zhu, S., Armstrong, K., Helsby, N.A., Ferguson, L.R., and Rowan, D.D. (2010). Metabolomic analysis identifies inflammatory and noninflammatory metabolic effects of genetic modification in a mouse model of Crohn's disease. Journal of proteome research *9*, 1965-1975.

Lin, H.M., Edmunds, S.I., Helsby, N.A., Ferguson, L.R., and Rowan, D.D. (2009). Nontargeted urinary metabolite profiling of a mouse model of Crohn's disease. Journal of proteome research *8*, 2045-2057.

Loftus, N., Barnes, A., Ashton, S., Michopoulos, F., Theodoridis, G., Wilson, I., Ji, C., and Kaplowitz, N. (2011). Metabonomic Investigation of Liver Profiles of Nonpolar Metabolites Obtained from Alcohol-Dosed Rats and Mice Using High Mass Accuracy MSn Analysis. Journal of Proteome Research *10*, 705-713.

Manichanh, C., Borruel, N., Casellas, F., and Guarner, F. (2012). The gut microbiota in IBD. Nature reviews. Gastroenterology & hepatology *9*, 599-608.

Marchesi, J.R., Holmes, E., Khan, F., Kochhar, S., Scanlan, P., Shanahan, F., Wilson, I.D., and Wang, Y. (2007). Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. Journal of proteome research *6*, 546-551.

Martin, F.P., Rezzi, S., Philippe, D., Tornier, L., Messlik, A., Holzlwimmer, G., Baur, P., Quintanilla-Fend, L., Loh, G., Blaut, M.*, et al.* (2009). Metabolic assessment of gradual development of moderate experimental colitis in IL-10 deficient mice. Journal of proteome research *8*, 2376-2387.

Mazanetz, M.P., Marmon, R.J., Reisser, C.B., and Morao, I. (2012). Drug discovery applications for KNIME: an open source data mining platform. Current topics in medicinal chemistry *12*, 1965-1979.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The ISME journal *6*, 610-618.

Momozawa, Y., Deffontaine, V., Louis, E., and Medrano, J.F. (2011). Characterization of bacteria in biopsies of colon and stools by high throughput sequencing of the V2 region of bacterial 16S rRNA gene in human. PloS one *6*, e16952.

Morgan, X.C., Tickle, T.L., Sokol, H., Gevers, D., Devaney, K.L., Ward, D.V., Reyes, J.A., Shah, S.A., LeLeiko, N., Snapper, S.B.*, et al.* (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome biology *13*, R79.

Murdoch, T.B., Fu, H., MacFarlane, S., Sydora, B.C., Fedorak, R.N., and Slupsky, C.M. (2008). Urinary metabolic profiles of inflammatory bowel disease in interleukin-10 gene-deficient mice. Analytical chemistry *80*, 5524-5531.

O'Hagan, S., and Kell, D.B. (2015). Software review: the KNIME workflow environment and its applications in genetic programming and machine learning. Genetic Programming and Evolvable Machines *16*, 387-391.

Panwala, C.M., Jones, J.C., and Viney, J.L. (1998). A novel model of inflammatory bowel disease: mice deficient for the multiple drug resistance gene, mdr1a, spontaneously develop colitis. J Immunol *161*, 5733-5744.

Parks, D.H., and Beiko, R.G. (2010). Identifying biologically relevant differences between metagenomic communities. Bioinformatics *26*, 715-721.

Paturi, G., Mandimika, T., Butts, C.A., Zhu, S., Roy, N.C., McNabb, W.C., and Ansell, J. (2012). Influence of dietary blueberry and broccoli on cecal microbiota activity and colon morphology in mdr1a(-/-) mice, a model of inflammatory bowel diseases. Nutrition *28*, 324-330.

Resta-Lenert, S., Smitham, J., and Barrett, K.E. (2005). Epithelial dysfunction associated with the development of colitis in conventionally housed mdr1a-/- mice. American journal of physiology. Gastrointestinal and liver physiology *289*, G153-162.

Rooks, M.G., Veiga, P., Wardwell-Scott, L.H., Tickle, T., Segata, N., Michaud, M., Gallini, C.A., Beal, C., van Hylckama-Vlieg, J.E., Ballal, S.A., *et al.* (2014). Gut microbiome composition and function in experimental colitis during active disease and treatment-induced remission. The ISME journal *8*, 1403-1417.

Saleh, M., and Elson, C.O. (2011). Experimental inflammatory bowel disease: insights into the host-microbiota dialog. Immunity *34*, 293-302.

Sartor, R.B. (2008). Microbial influences in inflammatory bowel diseases. Gastroenterology *134*, 577-594.

Schicho, R., Nazyrova, A., Shaykhutdinov, R., Duggan, G., Vogel, H.J., and Storr, M. (2010). Quantitative metabolomic profiling of serum and urine in DSS-induced ulcerative colitis of mice by (1)H NMR spectroscopy. Journal of proteome research *9*, 6265-6273.

Schicho, R., Shaykhutdinov, R., Ngo, J., Nazyrova, A., Schneider, C., Panaccione, R., Kaplan, G.G., Vogel, H.J., and Storr, M. (2012). Quantitative metabolomic profiling of serum, plasma, and urine by (1)H NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. Journal of proteome research *11*, 3344-3357.

Schinkel, A.H., Smit, J.J., van Tellingen, O., Beijnen, J.H., Wagenaar, E., van Deemter, L., Mol, C.A., van der Valk, M.A., Robanus-Maandag, E.C., te Riele, H.P., and et al. (1994). Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. Cell *77*, 491-502.

Schwab, C., Berry, D., Rauch, I., Rennisch, I., Ramesmayer, J., Hainzl, E., Heider, S., Decker, T., Kenner, L., Muller, M., *et al.* (2014). Longitudinal study of murine microbiota activity and interactions with the host during acute inflammation and recovery. The ISME journal *8*, 1101-1114.

Schwab, M., Schaeffeler, E., Marx, C., Fromm, M.F., Kaskas, B., Metzler, J., Stange, E., Herfarth, H., Schoelmerich, J., Gregor, M., *et al.* (2003). Association between the C3435T MDR1 gene polymorphism and susceptibility for ulcerative colitis. Gastroenterology *124*, 26-33.

Smith, P.M., Howitt, M.R., Panikov, N., Michaud, M., Gallini, C.A., Bohlooly, Y.M., Glickman, J.N., and Garrett, W.S. (2013). The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. Science *341*, 569-573.

Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences of the United States of America *100*, 9440-9445.

Tong, M., McHardy, I., Ruegger, P., Goudarzi, M., Kashyap, P.C., Haritunians, T., Li, X., Graeber, T.G., Schwager, E., Huttenhower, C.*, et al.* (2014). Reprograming of gut microbiome energy metabolism by the FUT2 Crohn's disease risk polymorphism. The ISME journal *8*, 2193-2206.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P.*, et al.* (2009). A core gut microbiome in obese and lean twins. Nature *457*, 480-484.

Wang, J., Reijmers, T., Chen, L., Van Der Heijden, R., Wang, M., Peng, S., Hankemeier, T., Xu, G., and Van Der Greef, J. (2009). Systems toxicology study of doxorubicin on rats using ultra performance liquid chromatography coupled with mass spectrometry based metabolomics. Metabolomics : Official journal of the Metabolomic Society *5*, 407-418.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology *73*, 5261-5267.

Williams, H.R., Cox, I.J., Walker, D.G., North, B.V., Patel, V.M., Marshall, S.E., Jewell, D.P., Ghosh, S., Thomas, H.J., Teare, J.P.*, et al.* (2009). Characterization of inflammatory bowel disease with urinary metabolic profiling. The American journal of gastroenterology *104*, 1435-1444.

Williams, H.R., Willsmore, J.D., Cox, I.J., Walker, D.G., Cobbold, J.F., Taylor-Robinson, S.D., and Orchard, T.R. (2012). Serum metabolic profiling in inflammatory bowel disease. Digestive diseases and sciences *57*, 2157-2165.

Supplementary materials

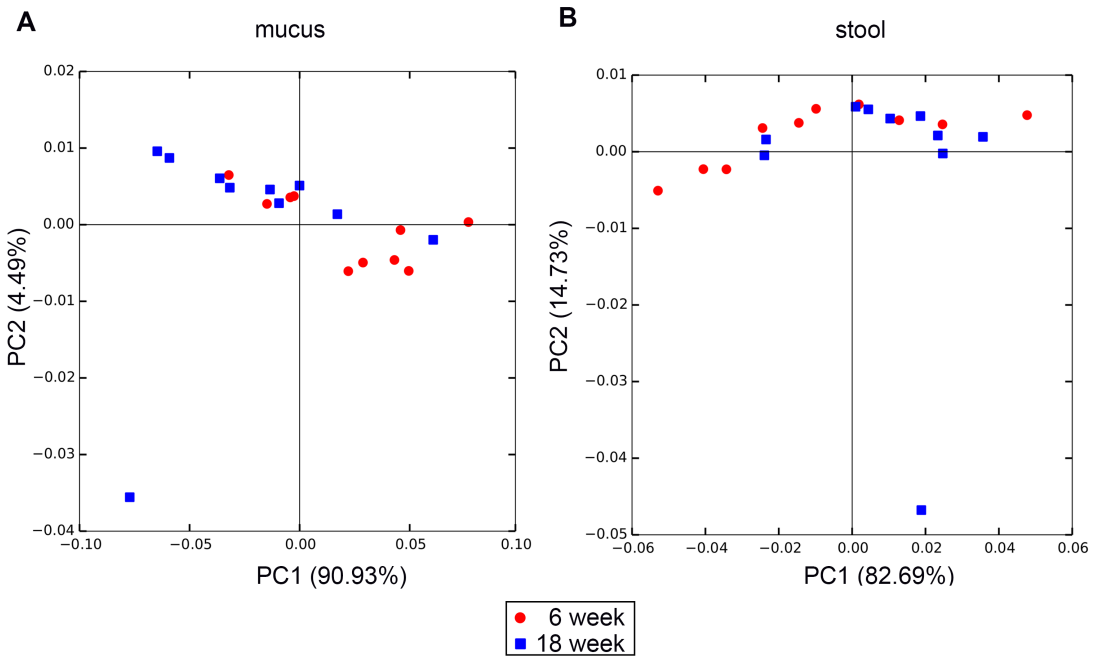
Supplementary Table 1. NSTI values to evaluate PICRUSt accuracy.

Group	Mean	Standard deviation (SD)	Age (weeks)	Genotype	Location
Group A	0.290	0.023	6	*mdr1a*$^{-/-}$	Stool
Group B	0.267	0.021	6	wt	Stool
Group C	0.229	0.028	6	*mdr1a*$^{-/-}$	Mucus
Group D	0.203	0.024	6	wt	Mucus
Group E	0.262	0.045	18	*mdr1a*$^{-/-}$	Stool
Group F	0.285	0.019	18	wt	Stool
Group G	0.222	0.026	18	*mdr1a*$^{-/-}$	Mucus
Group H	0.260	0.027	18	wt	Mucus

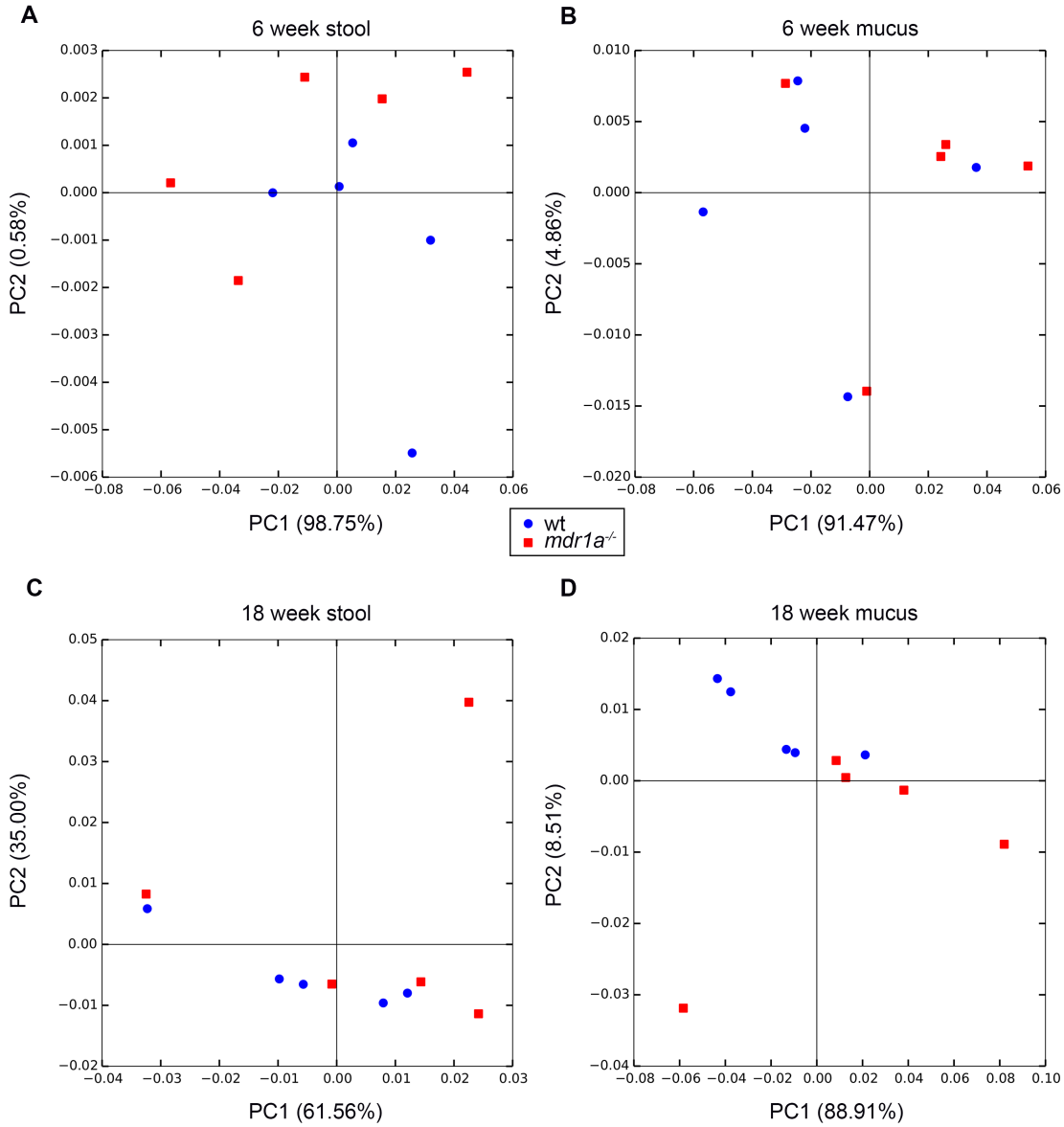
Supplementary Table 2. Previously published significant endogenous metabolites in human IBD and in murine models. These metabolites were detected in our analysis and confirmed by authentic standard analysis.

Metabolite	[M+H]+	RT[a] (min)	QC[b] % RSD[c]	Formula	Sample type that has been identified
Arginine	175.1190	0.55	19%	C6H14N4O2	Urine
Glycine	76.0393	0.65	14%	C2H5NO2	Urine
Alanine	90.0550	0.69	15%	C3H7NO2	Urine
Tyrosine	182.0812	1.48	24%	C9H11NO3	Faecal water
Isoleucine	132.1019	1.72	28%	C6H13NO2	Faecal water
Leucine	132.1019	1.87	43%	C6H13NO2	Faecal water
Tryptophan	205.0972	5.35	18%	C11H12N2O2	Urine
Lactic acid	91.0390	6.54	24%	C3H6O3	Faecal water
Hippurate	180.0655	6.61	21%	C9H9NO3	Urine
Creatinine	114.0662	0.67	9%	C4H7N3O	Urine
Mannitol	183.0863	0.60	26%	C6H14O6	Urine
Carnitine	162.1125	0.64	9%	C7H15NO3	Urine
Valine	118.0863	0.64	15%	C5H11NO2	Faecal water
Glucose	181.0707	0.66	31%	C6H12O6	Faecal water, Urine
Allantoin	159.0513	0.65	15%	C4H6N4O3	Urine
Trigonelline	138.055	0.67	10%	C7H7NO2	Urine
Acetoacetate	103.039	0.76	20%	C4H6O3	Urine
Glycylproline	173.0921	0.85	41%	C7H12N2O3	Urine
Asparginine	133.0608	0.83	33%	C4H8N2O3	
Methionine	150.0583	1.07	15%	C5H11NO2S	Urine
Hypoxanthine	137.0458	1.19	29%	C5H4N4O	Urine
Glutamine	147.0764	1.27	26%	C5H10N2O3	Urine
Proline	116.0706	1.45	26%	C5H9NO2	Serum
Phenylalanine	166.0863	3.65	52%	C9H11NO2	Urine
Xylose	151.0601	4.11	37%	C5H10O5	Urine
Succinate	119.0339	6.91	20%	C4H6O4	Urine
Aspartic acid	134.0302	7.06	25%	C4H7NO4	Faecal water
Lactose	343.1235	7.35	50%	C12H22O11	Urine
GPCho(16:0/0:0)	496.3398	16.34	12%	C24H50NO7P	
GPCho(18:0/0:0)	524.3711	17.29	11%	C26H54NO7P	

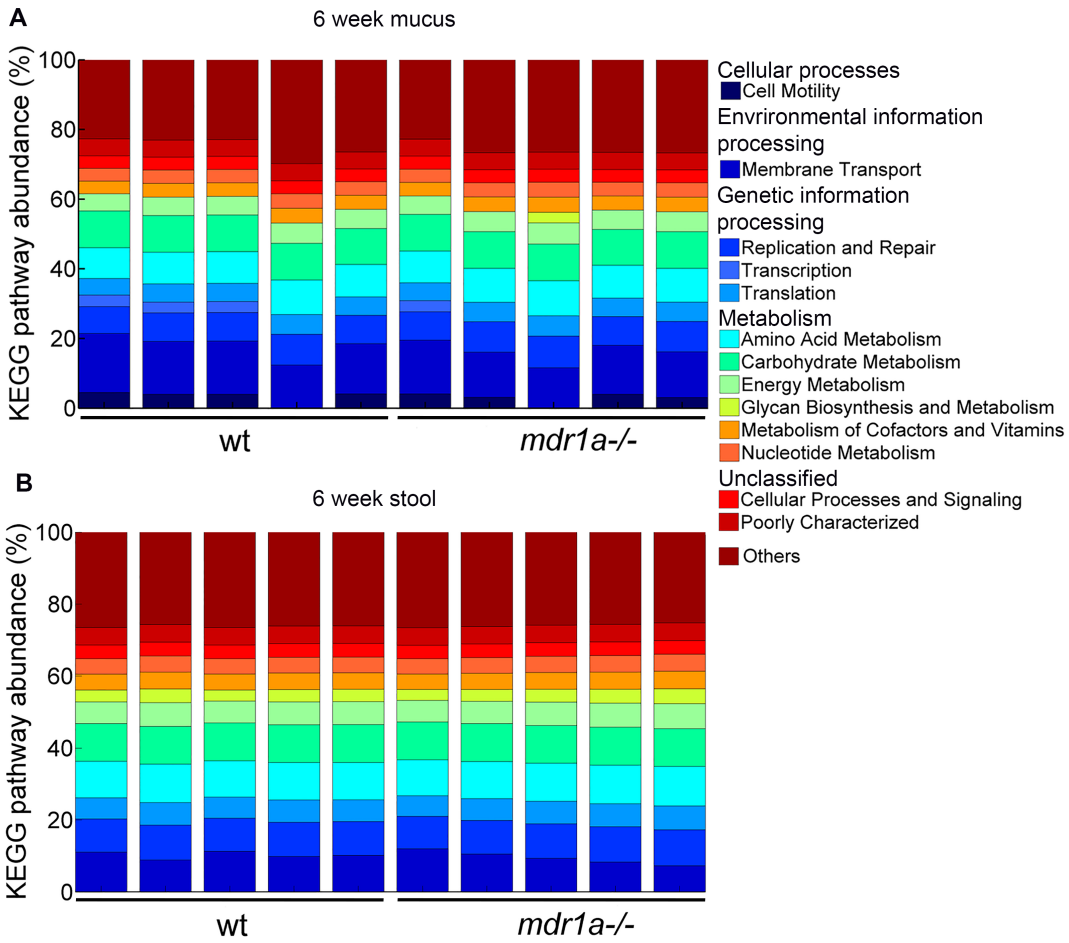
[a] Retention time (RT), [b] Quality control (QC), [c] Relative standard deviation (RSD) calculated by dividing the standard deviation by the mean in the current study.



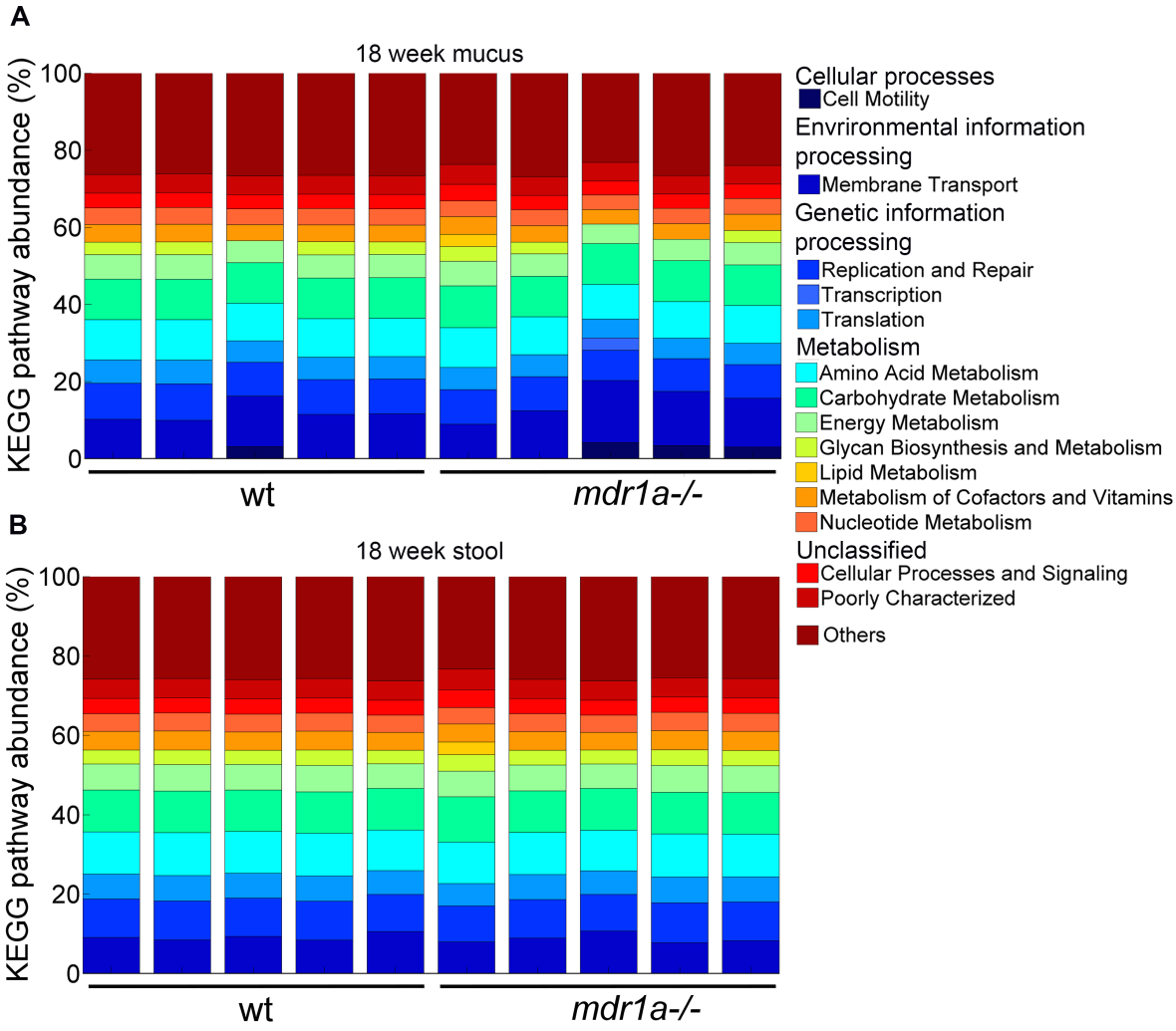
Supplemental Figure 1. The effect of age on microbial gene functional patterns. PCoA plot of **(A)** mucus and **(B)** stool samples based on Bray-Curtis distance of KEGG metabolic pathways. Age had no effect in segregating groups in separate clusters.



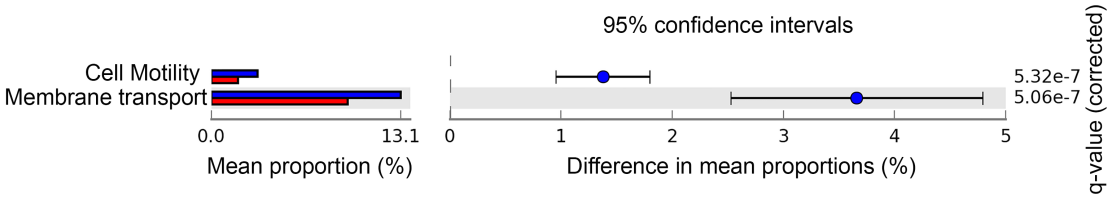
Supplemental Figure 2. Impact of genotype on microbial gene functional patterns. PCoA plots using Bray-Curtis distance metric revealed no clustering based on genotype in stool microbial communities at **(A)** 6 or **(C)** 18 weeks or in mucus-associated bacteria at **(B)** 6 weeks or **(D)** 18 weeks.



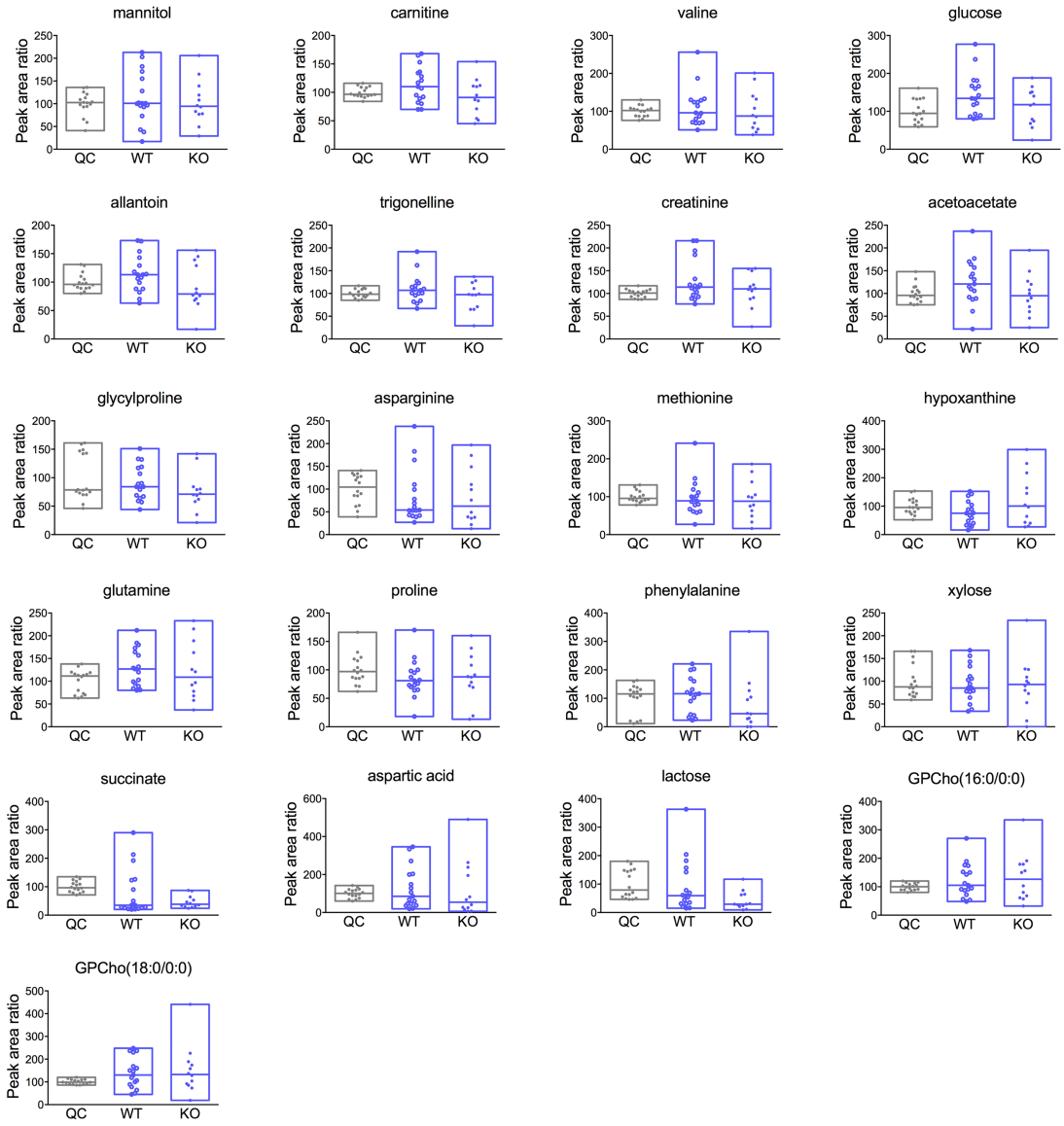
Supplemental Figure 3. Similarity of the microbial functional potential in WT and *mdr1a*$^{-/-}$ mice before the onset of inflammation. Relative abundance of KEGG metabolic pathways in **(A)** mucus and **(B)** stool microbial communities at 6 weeks. The category 'others' represents KEGG pathways with abundance below 0.3%.



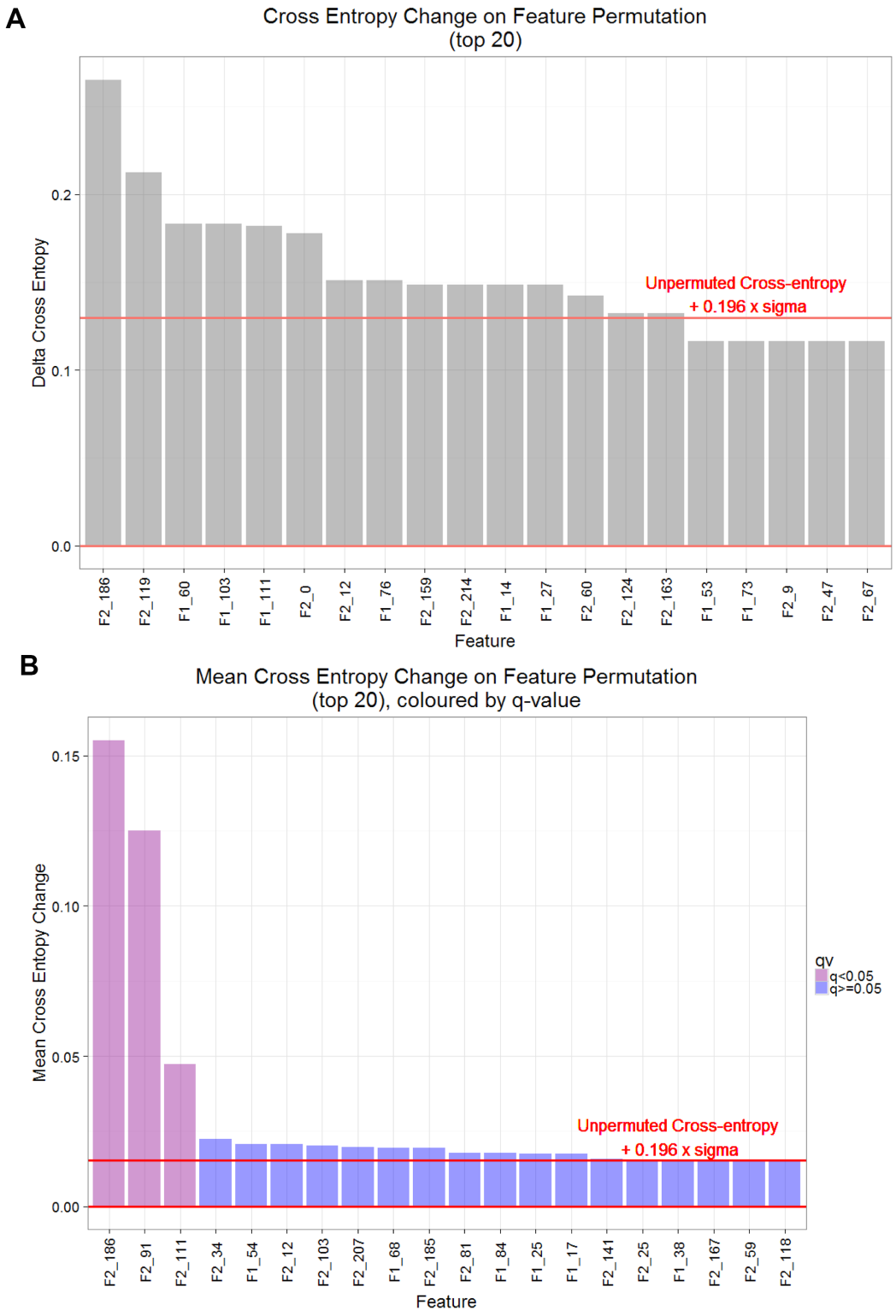
Supplemental Figure 4. Resilience of the microbial functional potential in WT and colitis prone *mdr1a*$^{-/-}$ mice during colitis onset. Relative abundance of KEGG metabolic pathways in **(A)** mucus and **(B)** stool microbial communities at 18 weeks. The category 'others' represents KEGG pathways with abundance below 0.3%.



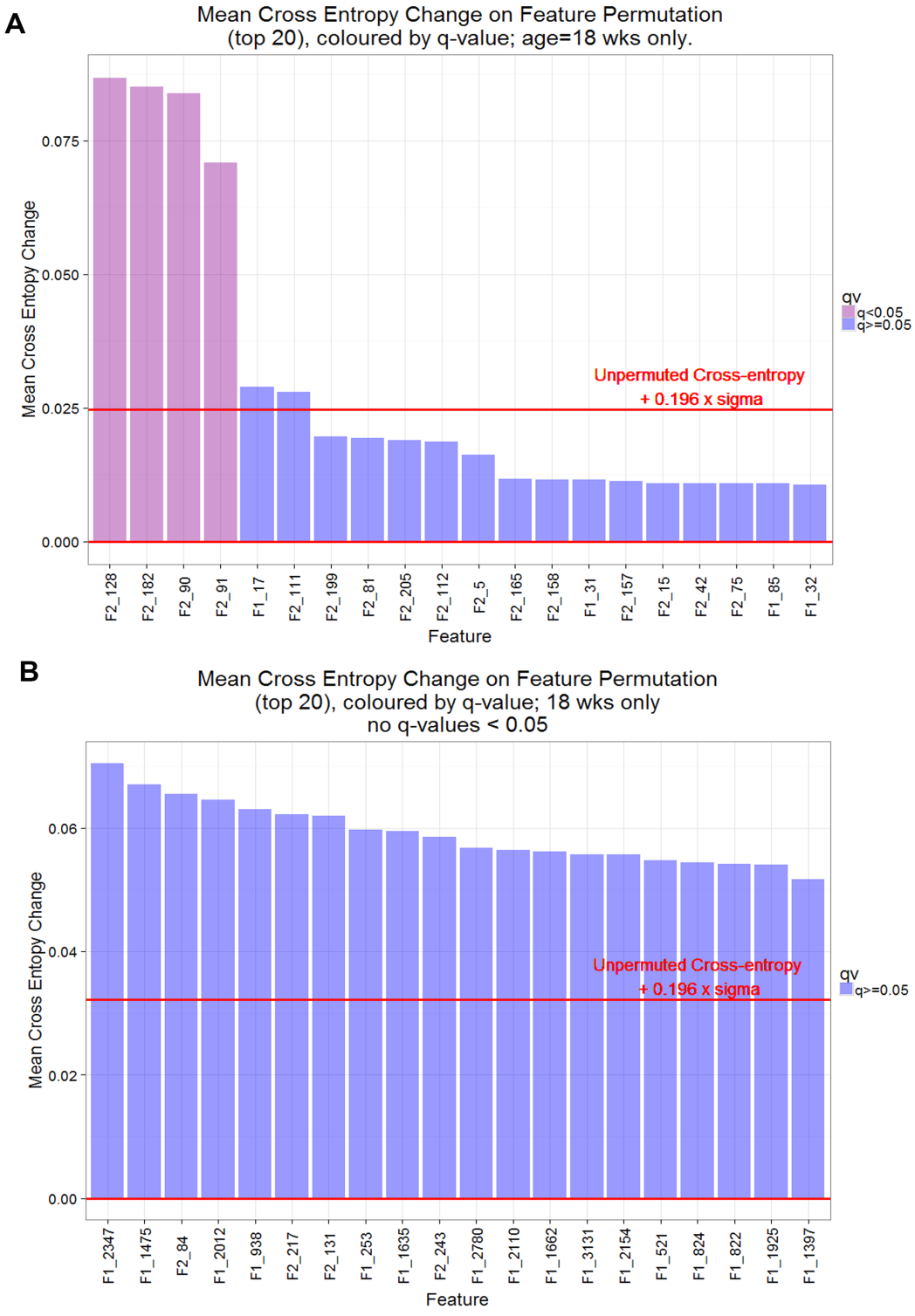
Supplemental Figure 5. Differences in KEGG pathways from mucus and stool microbial communities. The pathways at level 2 subsystems are shown. Pathways overrepresented in the mucus (blue) or stools (red) are indicated. Corrected p-values were calculated using Benjamini–Hochberg false discovery rate (FDR). Effect size measures (difference between proportions) and their 95% confidence intervals are shown.



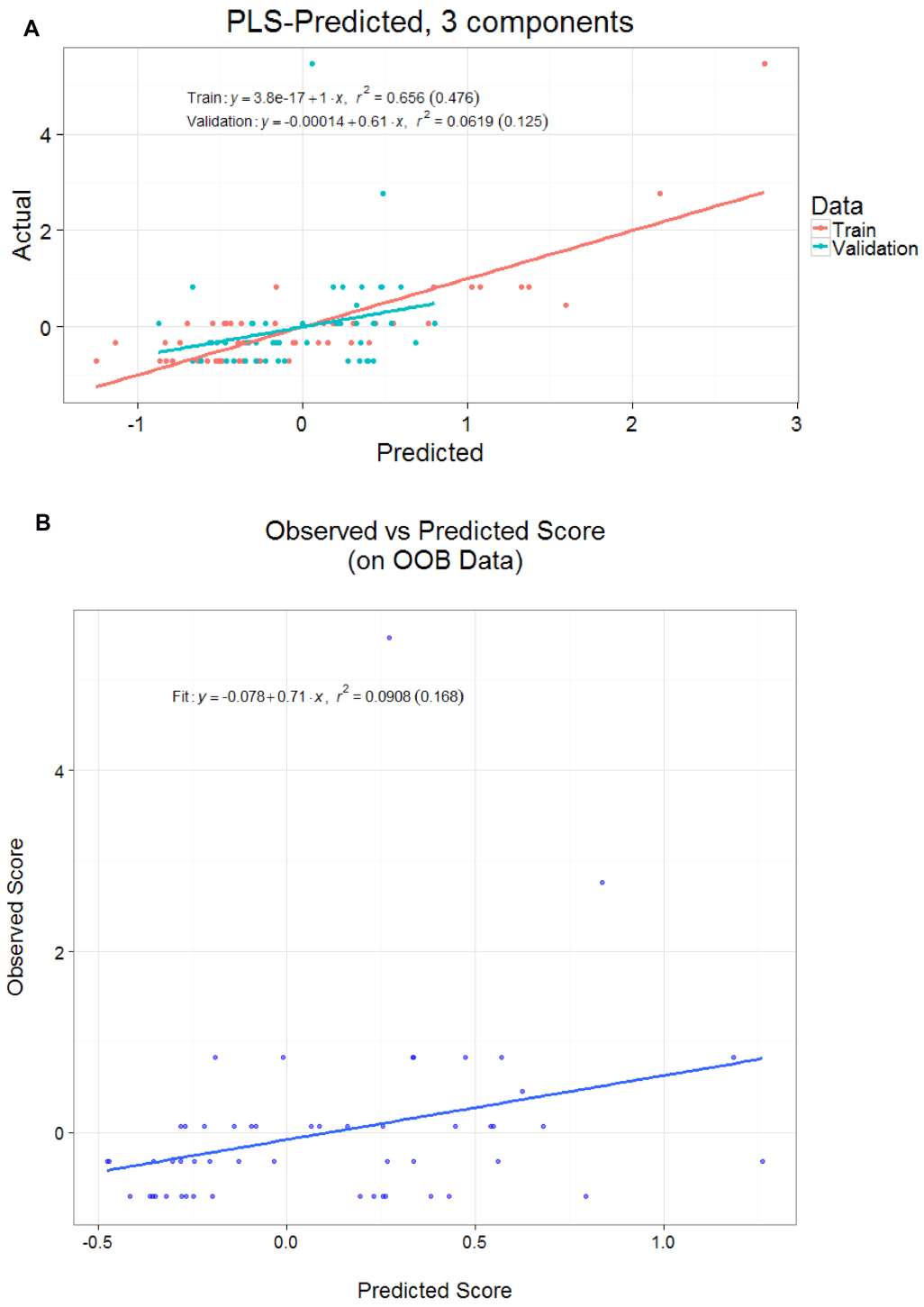
Supplemental Figure 6. Levels of known IBD reported marker metabolites in urinary samples from WT and *mdr1a*$^{-/-}$ mice at 18 weeks. No differences were identified in the levels of metabolites in WT and KO samples during onset of signs of inflammation at 18 weeks. Creatine was used as an internal control in these calculations. The median is shown as a line and bars capture the minimum and maximum. Unpaired t-test or Mann Whitney test were applied for comparison between WT and KO samples.



Supplemental Figure 7. Feature permutation for the identification of discriminatory mass ions responsible for differential classification based on genotype. (A) Features were permuted using mass ion data from all samples as input, and cross entropy was calculated using a random forests (RF) classifier. **(B)** Features with a difference between mean cross entropy of permuted and unpermuted data greater than 1.96 * sigma were regarded as significant. The Storey multiple correction method was applied.



Supplemental Figure 8. Feature permutation for the detection of mass ions contributing to separation of metabolite profiles of WT and *mdr1a*$^{-/-}$ mice. Features were permuted using mass ion data from 18-week old samples only as input and cross entropy was calculated using a RF classifier. Permutation results are shown from 18-week samples **(A)** of the initial dataset and **(B)** of the re-processed dataset to correct for noise errors. Features with a difference between mean cross entropy of permuted and unpermuted data greater than 1.96 * sigma were regarded as significant. The Storey multiple correction method was applied.



Supplemental Figure 9. Changes in metabolite profiles are not related to intestinal inflammation. The PLS and RF regression plots show that the predicted colitis scores based on the pattern of urinary metabolites does not correlate with the actual colitis scores. **(A)** PLS plot; actual versus predicted score for training data and leave-one-out cross validation data with linear regression fit; also showing fit equations with squared Pearson correlation coefficient and (squared Spearman correlation coefficient). **(B)** Random Forest plot, colitis score predictions using "Out-of-Bag" data with linear fit, fit equation and squared correlation coefficients. The low R2 values for both techniques indicate poor predictive performance.

Chapter 4

Characterization of antimicrobial protein expression prior to colitis development in mdr1a knockout mice

M. Glymenaki[1], G. Warhurst[2], A.J. McBain[1], K.J. Else[1]†, S.M. Cruickshank[1]†*

[1]Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

[2]Infection, Injury and Inflammation Research Group, Division of Medicine and Neurosciences, University of Manchester and Salford Royal Hospitals NHS Trust, Salford, UK

†These authors contributed equally to the study.

***Correspondence to:** Sheena Cruickshank, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK

sheena.cruickshank@manchester.ac.uk

Author contributions

Maria Glymenaki: Designed and performed experiments, analysed data and wrote the manuscript.

Geoffrey Warhurst: Designed and supervised the project.

Andrew McBain: Designed and supervised the project.

Kathryn Else: Designed and supervised the project, critically reviewed and edited the manuscript.

Sheena Cruickshank: Designed and supervised the project, critically reviewed and edited the manuscript.

4.1 Abstract

Although expression of antimicrobial proteins (AMPs) is altered in inflammatory bowel disease (IBD), the precise role of antimicrobial factors in shaping gut microbiota composition during IBD progression remains unclear. Previous work using the *mdr1a*$^{-/-}$ model of spontaneous colitis, showed a dramatic reduction of the antimicrobials regenerating islet-derived protein 3 γ (Reg3γ) and the related pancreatitis-associated protein (PAP) in the colon before the start of inflammation. Furthermore, we have observed alterations in mucus microbial communities prior to the onset of inflammation in *mdr1a*$^{-/-}$ mice. To assess whether alterations in antimicrobial expression accounted for differences in mucus bacterial composition, we examined the expression patterns of a range of antimicrobials including Reg3γ, β-defensin 1, angiogenin 4 (Ang4) and resistin-like molecule beta (Relm-β), and IgA. Our results revealed that, despite changes in mucus microbial communities, *mdr1a*$^{-/-}$ mice 12 weeks prior to colonic inflammation and wild-type littermate controls had similar expression of antimicrobial effector molecules. Levels of AMPs were also not different between controls and *mdr1a*$^{-/-}$ mice during the development of colitis, with the exception of β-defensin 1 and IgA that were increased in cases of severe inflammation. Our results suggest that other factors involved in epithelial-microbiota interactions contribute to alterations in mucus microbiota composition prior to the onset of colitis.

4.2Introduction

Antimicrobial proteins (AMPs), produced by specialized epithelial cells, are enriched in the mucus layer and contribute to host defence by limiting bacterial access to the epithelial surface (Meyer-Hoffert et al., 2008; Vaishnava et al., 2011). AMPs comprise representatives of several major protein families including defensins, cathelicidins, C-type lectins (such as regenerating islet-derived protein 3 (Reg3) family) and ribonucleases (RNases such as angiogenin 4 (Ang4)) (Bevins and Salzman, 2011; Gallo and Hooper, 2012). There are differences in the types of AMPs produced between different regions of the gut, indicating their variable sources of origin including Paneth cells, enterocytes or goblet cells (Bowcutt et al., 2014). Secretory IgA also contributes to innate defence mechanisms by restricting bacterial access to the intestinal epithelium (Peterson and Artis, 2014) and preventing aberrant bacterial expansion (Suzuki et al., 2004; Wei et al., 2011).

Microbial signals are necessary to induce the expression of some AMPs, whereas others, mostly α-defensins, show a constitutive pattern of expression (Bevins and Salzman, 2011; Gallo and Hooper, 2012). Activation of Toll-like receptor (TLR) signaling in response to recognition of conserved microbial motifs drives the expression of AMPs, which control bacterial numbers at the mucosal surface (Brandl et al., 2007; Vaishnava et al., 2008; Vaishnava et al., 2011; Vora et al., 2004). Therefore, expression of AMPs is regulated by the gut microbiota and in turn AMPs shape gut microbiota composition to maintain immune homeostasis (Salzman et al., 2010; Vaishnava et al., 2011). Disruption of homeostasis in inflammatory bowel disease (IBD) is associated with an altered gut microbiota composition (Kostic et al., 2014). IBD patients and experimental animal models show increased expression of AMPs including β-defensins, Reg3γ and resistin-like molecule beta (Relm-β) in the colonic mucosa during inflammation (Fahlgren et al., 2004; Hogan et al., 2006; Mizoguchi et al., 2003; Ogawa et al., 2003; Ramanan et al., 2014; Wehkamp et al., 2003). Given that AMPs can regulate microbial communities, altered AMP expression may function as a mechanism to counteract changed microbial composition and increased colonization.

Previous studies using the *mdr1a*$^{-/-}$ model of spontaneous colitis have shown decreased expression of genes implicated in the maintenance of intestinal homeostasis before the onset of colonic inflammation (Collett et al., 2008; Dommels et al., 2007). Transcriptome analysis

using whole colon identified significant changes as early as 4 weeks of age (Collett et al., 2008). Genes implicated in antigen processing and bacterial recognition were upregulated, whereas the antimicrobial Reg3γ and its related anti-inflammatory factor pancreatitis-associated protein (PAP) were strongly downregulated in *mdr1a*$^{-/-}$ mice (Collett et al., 2008; Dommels et al., 2007). Similar changes in gene expression were also observed in colonic epithelial cells (CECs) from disease-free 12-week old (Dommels et al., 2007) and 4-week old *mdr1a*$^{-/-}$ mice (Collett et al., 2008), indicating that these changes are localized in the epithelium (Collett et al., 2008; Dommels et al., 2007). Furthermore, we have reported alterations in gut microbiota composition in the mucus but not in faeces in *mdr1a*$^{-/-}$ mice 12 weeks prior to any histological signs of inflammation (Glymenaki et al., 2016). However, it is not known what causes the changes in mucus microbiota before the onset of colitis. We hypothesized that shifts in mucus microbial communities may be explained in part by altered epithelial response and in particular AMPs expression. It is crucial to control for the environment in microbiota studies (Friswell et al., 2010) and we have achieved this by using co-housed littermate controls. In contrast, the previous two studies on colonic gene expression in *mdr1a*$^{-/-}$ mice used non-littermates, which were also individually housed (Collett et al., 2008; Dommels et al., 2007), thus potential confounding effects of the environment cannot be overruled. To address these issues, we examined the colonic expression of AMPs and IgA at the molecular and protein level using co-housed littermate controls at stages preceding and during colitis onset.

4.3 Materials and methods

4.3.1 Maintenance of animals

Mdr1a$^{-/-}$ mice (FVB.129P2-Abcb1atm1Bor N7) (Schinkel et al., 1994) and control FVB mice were purchased from Taconic Farms (Albany, NY, USA) and crossbred to generate F1 heterozygotes. F1 mice were further crossbred to generate F2 littermate controls, which were used for experiments. All animals were maintained under specific, pathogen-free (SPF) conditions at the University of Manchester. Mice were given autoclaved standard chow and sterile acidified water (pH=3.2) *ad libitum*. Co-housed littermate *mdr1a*$^{-/-}$ and wild-type (wt) males were used to ensure shared microbiota. All experimental procedures in mice were performed in accordance with the regulations issued by the Home Office under amended ASPA, 2012.

4.3.2 RNA extraction and cDNA synthesis

RNA extraction from proximal colon tissue samples was performed using TRIsure (Bioline, London, UK) in accordance with the manufacturer's instructions. In brief, tissue was homogenized using a FastPrep -24 Instrument (MP Biochemicals, OH, USA), phases were separated using chloroform (Sigma-Aldrich, Dorset, UK) and RNA was precipitated in isopropanol (Sigma-Aldrich). Resulting RNA was resuspended in nuclease-free water and its concentration and integrity was measured in a Nanodrop ND-1000 Spectrophotometer (Labtech International, Uckfield, UK). RNA was stored at -80°C until further use. 2µg of RNA was treated with RNase-free DNase (Promega, Southampton, UK) and converted to cDNA using the Bioscript kit (Bioline) according to manufacturer's guidelines and stored at -20°C.

4.3.3 Gene expression analysis by quantitative PCR (qPCR)

qPCR was performed using the LightCycler 480 Probes master (Roche Diagnostics, West Sussex, UK) on the LightCycler 480 Instrument (Roche). A mouse-specific Universal

ProbeLibrary Set (Roche) containing pre-validated, dual-labelled RT-PCR probes was used. Each PCR reaction contained 10µl of LightCycler 480 Probes master, 500nM of each primer pair, 200nM of each probe, 1µl template and nuclease-free H2O up to 20µl. Primer sequences for Relm-β, Reg3γ, β-defensin 1, Ang4 and IL-22 and two housekeeping genes (β-actin and villin) are shown in Supplementary Table 1. Gene expression was normalized based on the two housekeeping genes and a control sample (6-week WT) used as calibrator in the Roche Applied Science software. A three-step cycling reaction was performed using the following conditions: 5 min at 95°C for 1 cycle and 10 sec at 95°C, 20 sec at 60°C and 1 sec at 72°C for 40 cycles.

4.3.4 Histology

Distal colon snips taken at autopsy were placed in fixative solution containing 8% (v/v) formaldehyde, 0.9% (w/v) sodium chloride, 2% (v/v) glacial acetic acid and 0.5% (w/v) cetrimide in dH_2O, processed and paraffin-embedded. Paraffin-embedded tissue was sectioned at 5-µm thickness. Frozen histological sections were also prepared by OCT embedded distal colon tissue. 6-µm sections were cut, placed onto Superfrost plus microscope slides (Thermo Scientific, Paisley, UK) and kept at -80˚C until further use.

4.3.5 Immunohistochemistry

Sections embedded in paraffin were stained for the presence of antimicrobial peptides Relm-β, Ang4 and Reg3γ as previously described (Forman et al., 2012), whereas frozen OCT sections were used for IgA staining (Forman et al., 2016). Frozen tissue sections were fixed in 4% PFA (Sigma-Aldrich) on ice for 10 min. For both types of immunohistochemical staining, endogenous peroxidase activity was quenched by incubation with 1.5U/ml glucose oxidase type II-S: from *Aspergillus niger* (Sigma-Aldrich) in the presence of 1.8 mg/ml D-glycose (Sigma-Aldrich) and 0.064 mg/ml sodium azide (Sigma-Aldrich) for 20 min at 37˚C. Antigen retrieval was performed only in wax sections by incubation with pepsin digest-all^{TM}3 (Invitrogen, Paisley, UK) for 10 min at 37˚C. Non-specific binding was blocked by incubation with 7% goat serum (Thermo Scientific) for Relm-β and Reg3γ, 1.5% donkey serum (Sigma-Aldrich) for Ang4 and 7% rat serum (AbD Serotec, Kidlington, UK) for IgA.

The avidin/biotin kit (Thermo Fischer Scientific) was used to block endogeous avidin/biotin binding sites according to manufacturer's guidelines. Sections were incubated with purified polyclonal antibody to Relm-β (Abcam, Cambridge, UK), purified polyclonal Reg3γ antibody (Aviva Systems Biology, CA, USA) (5 μg/ml in PBS), sheep anti-mouse polyclonal Ang4 (12.5 μg/ml in 1.5% donkey serum) or biotin-conjugated rat anti-mouse IgA (BD-Biosciences, Oxford, UK) (5μg/ml in PBS). Following that the slides were incubated with the secondary biotinylated goat anti-rabbit IgG $F(ab')_2$ (Santa Cruz Biotechology, TX, USA) at 1 μg/ml in the case of Relm-β and Reg3γ or $F(ab')_2$ donkey anti-sheep biotin (Stratech Scientific, Suffolk, UK) at 0.85 μg/ml in the case of Ang4. Then, ABC kit (avidin-biotin complex) (Vector laboratories, Peterborough, UK) was added and colour development was monitored under a microscope after treatment with the 3, 3'-diaminobenzidine (DAB) peroxidase substrate kit (Vector laboratories). Sections were counterstained with HaemQS (Vector laboratories) for 1 min and mounted in aqueous mounting medium (AbD Serotec). For IgA staining 0.1% (w/v) saponin in PBS was used in all washing steps.

Paraffin-embedded sections were stained for IL-22 as follows. Slides were incubated in a microwave oven for 5 min in citrate (pH=6) for antigen retrieval. Endogenous peroxidase was quenched in 0.3% (v/v) H_2O_2 from 30% Hydrogen peroxide solution (Sigma-Aldrich) for 20 min, followed by blocking with the reagent Tyramide Signal Amplification (TSA™) Plus Cyanine 3 (Perkin Elmer, Llantrisant, UK) for 30 min. Slides were incubated with the rabbit polyclonal IL-22 (Abcam) (10 μg/ml in 0.05% BSA (w/v)) followed by a biotinylated goat anti-rabbit antibody (Dako, Cambridgeshire, UK) (8.2 μg/ml). After streptavidin-POD conjugate horseradish peroxidase (HRP) (Roche Diagnostics), DAB (Vector laboratories) was added to the slides and colour development was monitored under microscope.

Images of the sections were obtained using a SPOT Insight QE camera and SPOT software (Diagnostic instruments, MI, USA) and analysed using ImageJ (http://rsb.info.nih.gov/ij). All slides were blinded prior to analysis and the numbers of Relm-β, Ang4 and Reg3γ positive cells per section were counted in a minimum of 20 crypts from three sections per mouse. The number of IgA and IL-22 positive cells was quantified per field of view. Staining with relevant isotype controls was performed to confirm there was no non-specific staining (Supplementary Figure 1).

Statistical analysis

Statistical analysis was performed using GraphPad Prism 6 (GraphPad software, CA, USA). Data that did not exhibit a normal distribution were analyzed using the nonparametric Kruskal-Wallis test with Dunn's post-test. Correlation analysis was performed using the Pearson correlation coefficient. $P<0.05$ was considered as statistically significant (*$P<0.05$, **$P<0.01$, ***$P<0.001$).

4.4Results

4.4.1 Colitis-prone *mdr1a*$^{-/-}$ and wt mice have similar levels of AMPs in their colon

To identify potential changes in AMPs that could explain observed differences in mucus microbial communities in *mdr1a*$^{-/-}$ mice prior to mucosal inflammation (Glymenaki et al., 2016), we examined the levels of expression of AMPs, known to be found in the large intestine, including Reg3-γ, Relm-β, Ang4 and β-defensin 1 (Bowcutt et al., 2014). We selected two time points: one at 6 weeks, when no histological signs of inflammation are present, and one at 18 weeks, when inflammation started, as shown in our previous study of colitis progression in *mdr1a*$^{-/-}$ mice (Glymenaki et al., 2016). Gene expression of Relm-β in the proximal colon was similar between wt and *mdr1a*$^{-/-}$ mice at 6 weeks or 18 weeks of age (Figure 1A). Despite the fact that Relm-β expression was initially low at 6 weeks of age in both groups of mice, it significantly increased nine-fold in wt animals ($P<0.05$) and five-fold in *mdr1a*$^{-/-}$ mice (non-significant) at 18 weeks of age. Therefore, the level of Relm-β trancription was elevated over time. In contrast to the mRNA transcript data, the number of Relm-β positive cells per crypt in the distal colon decreased over time in both wt and *mdr1a*$^{-/-}$ mice ($P<0.05$) (Figure 1B). Immunohistochemical analysis showed that the numbers of Relm-β positive cells were comparable between wt and *mdr1a*$^{-/-}$ mice at both 6 and 18 weeks of age (Figure 1C).

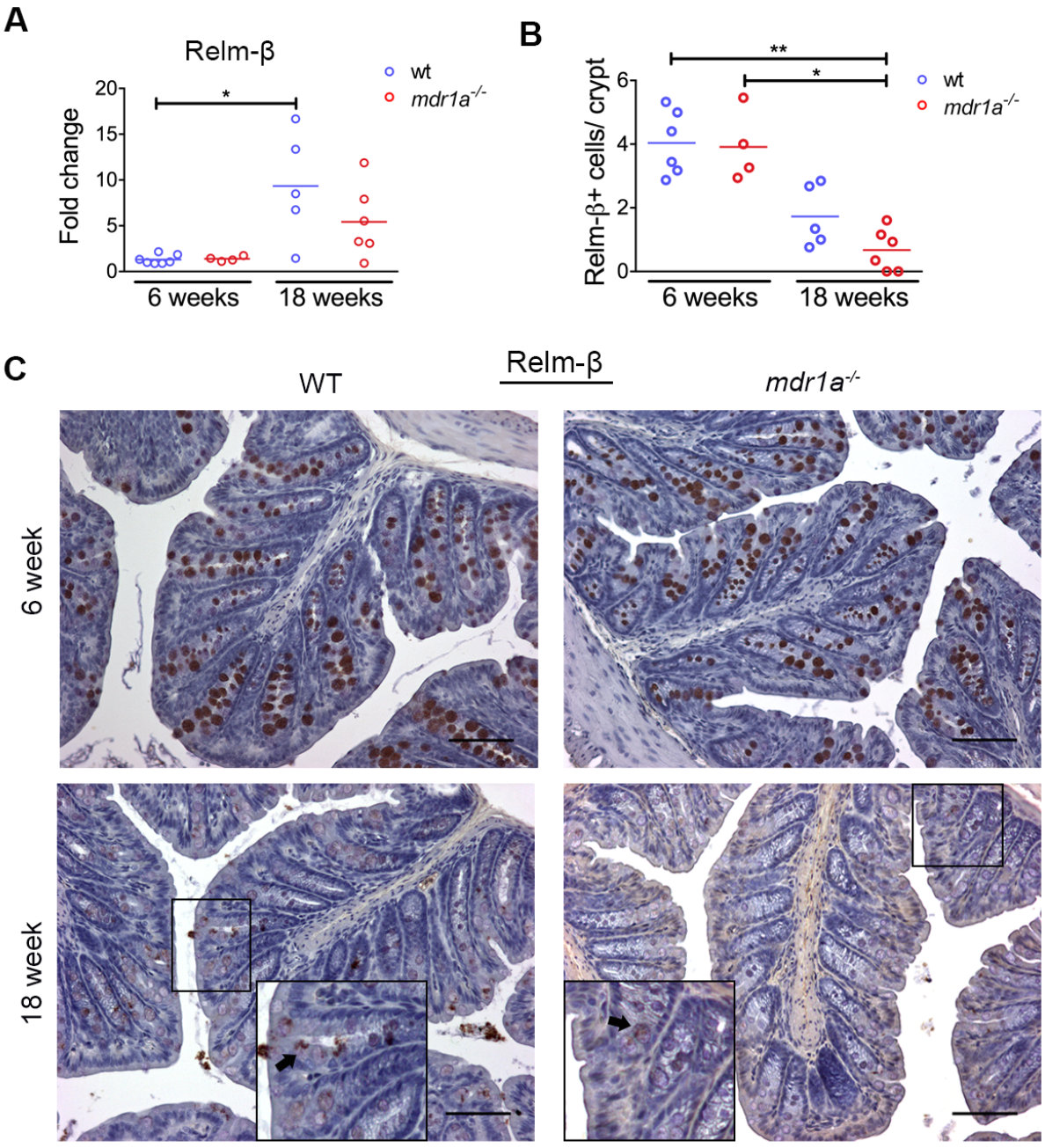


Figure 1. **Expression levels of Relm-β in gut tissue from *mdr1a*$^{-/-}$ and wt control mice at 6 and 18 weeks of age. (A)** qPCR analysis of Relm-β gene expression in proximal colon tissue. Fold change in Relm-β expression after normalization to two housekeeping genes and to a 6-week old wt mouse. N=4-7. **(B)** Quantification of Relm-β staining is shown as numbers of Relm-β positive cells per crypt in the distal colon. N=4-6. **(C)** Representative images of Relm-β immunohistochemistry in the distal colon tissue. The mean is shown in all data. * P<0.05, ** P<0.01 as determined by Kruskal-Wallis test with Dunn's multiple comparisons test in **(A and B)**. Scale bar= 100µm. Arrows in inset images indicate positive cells.

Analysis of Ang4 gene expression levels in the proximal colon showed low expression for both wt and *mdr1a*$^{-/-}$ mice at 6 and 18 weeks of age except for 18-week old wt mice, where a spread in the levels of Ang4 gene expression was seen (Figure 2A). These results of Ang4 were also reflected at the protein level as indicated by immunohistochemistry. Low levels of Ang4 protein were observed in both genotypes at 6 and 18 weeks of age (Figure 2B). The only exception was 18-week old wt animals, where there was high variability in the number of Ang4 positive cells per crypt. Ang4 positive cells in the distal colon were goblet cells as previously reported (Forman et al., 2012) (Figure 2C).

As we have shown similar expression of Relm-β and Ang4 between wt and *mdr1a*$^{-/-}$ mice at 6 and 18 weeks of age, we further investigated their expression at two additional time points. We selected 4 weeks of age to represent a post-weaning time and 12 weeks of age, because it is an intermediate time point before colitis onset at 18 weeks of age. At 4 weeks of age, wt controls expressed higher levels of Relm-β ($P<0.05$) than in *mdr1a*$^{-/-}$ mice and there was also a trend for higher Ang4 expression in wt mice (Supplementary Figure 2A-D). In contrast, there was no significant difference between the groups at 12 weeks of age. Relm-β and Ang4 positive cells were also counted per area in addition to counts per crypt, as colitis is associated with crypt elongation which may make the numbers of cells expressing AMPs elevated due to longer crypts (Supplementary Figure 2E-F). However, irrespective of whether cells were counted per crypt or per area, the numbers of Relm-β and Ang4 positive cells per area showed the same trends.

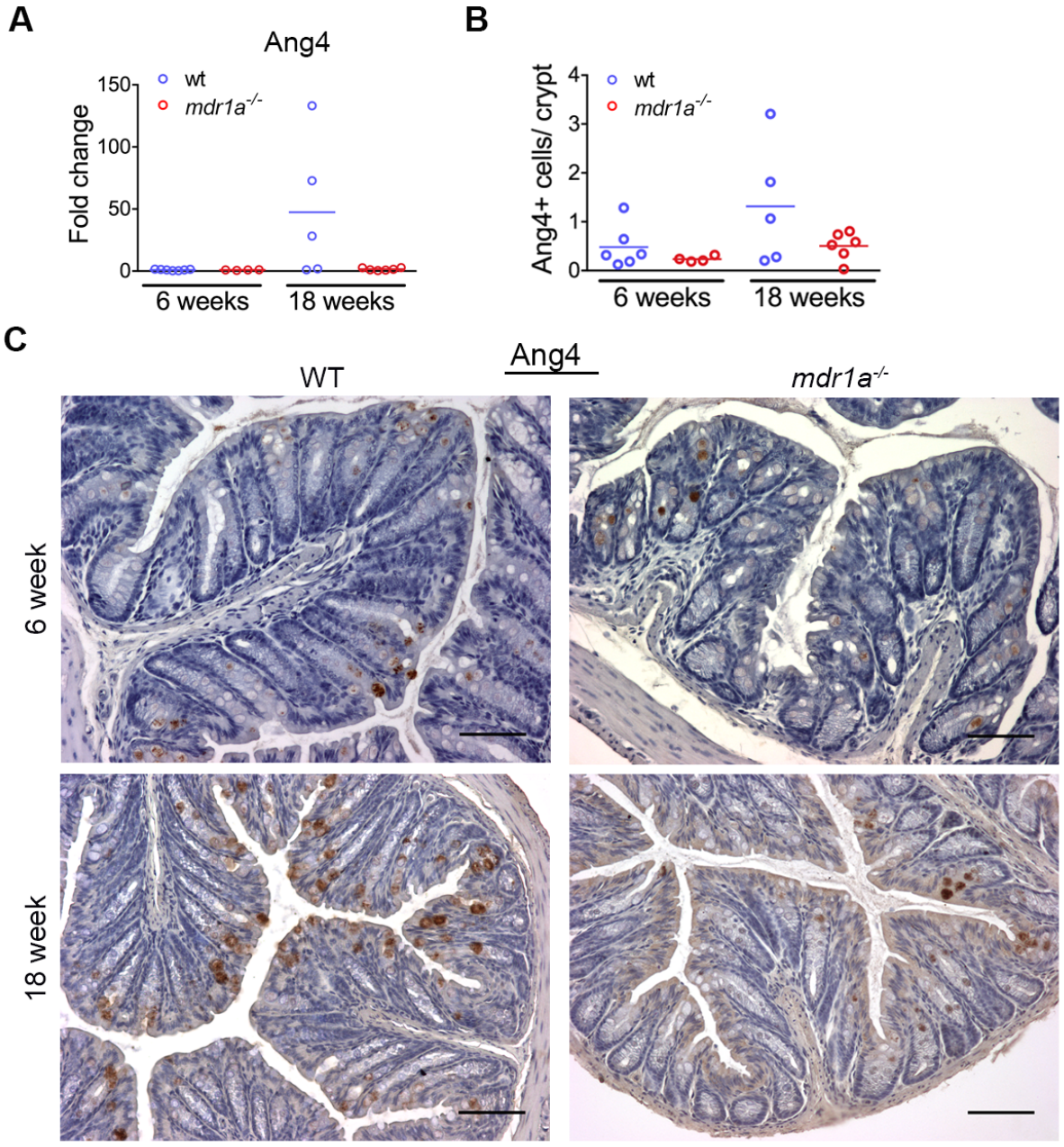


Figure 2. **Expression levels of Ang4 in gut tissue from *mdr1a*$^{-/-}$ and wt mice at 6 and 18 weeks of age. (A)** qPCR analysis of Ang4 gene expression in proximal colon tissue. Fold change in Ang4 expression after normalization to two housekeeping genes and to a 6-week old wt mouse. N= 4-7. **(B)** Quantification of Ang4 staining is shown as numbers of Ang4 positive cells per crypt in the distal colon. N=4-6. **(C)** Representative Ang4 immunohistochemistry sections from the distal colon. Data are depicted as dots with mean. Kruskal-Wallis test with Dunn's multiple comparisons test in **(A and B)**. Scale bar= 100μm.

Previous studies (Collett et al., 2008; Dommels et al., 2007) have shown that expression of Reg3γ is strongly downregulated in *mdr1a*$^{-/-}$ mice before the development of colitis. However, our qPCR analysis demonstrated that although Reg3γ was present at low levels in *mdr1a*$^{-/-}$ mice at 6 weeks of age, the levels of Reg3γ expression were comparable between *mdr1a*$^{-/-}$ and wt littermate controls (Figure 3A). Immunohistochemical staining for Reg3γ showed no positive cells in the distal colon of both *mdr1a*$^{-/-}$ and wt mice (data not shown). Another AMP, β-defensin 1, was almost undetectable in both wt and *mdr1a*$^{-/-}$ mice at both time points (Figure 3B). Although it remained at low levels in wt mice, there was variable expression in colitis-prone *mdr1a*$^{-/-}$ mice at 18 weeks of age.

As Reg3γ production can be driven by IL-22 (Sanos et al., 2009), we also examined the levels of IL-22 cytokine expression by qPCR. Levels of IL-22 were generally low at both time points, with a few outliers (Figure 3C). Elevated levels in 18-week old *mdr1a*$^{-/-}$ mice were consistent with the exacerbated colitis in those individuals. Staining with IL-22 did not show any positive cells in the distal colon in accordance with the low levels of IL-22 transcripts (data not shown). We found no correlation between IL-22 and Reg3γ gene expression levels (Figure 3D), suggesting that expression of Reg3γ was not fully regulated by IL-22. Collectively, our findings indicate that although we observed age-dependent changes in AMPs expression, differences in the levels of AMPs cannot explain the altered microbial composition in *mdr1a*$^{-/-}$ mice compared to wt littermate controls.

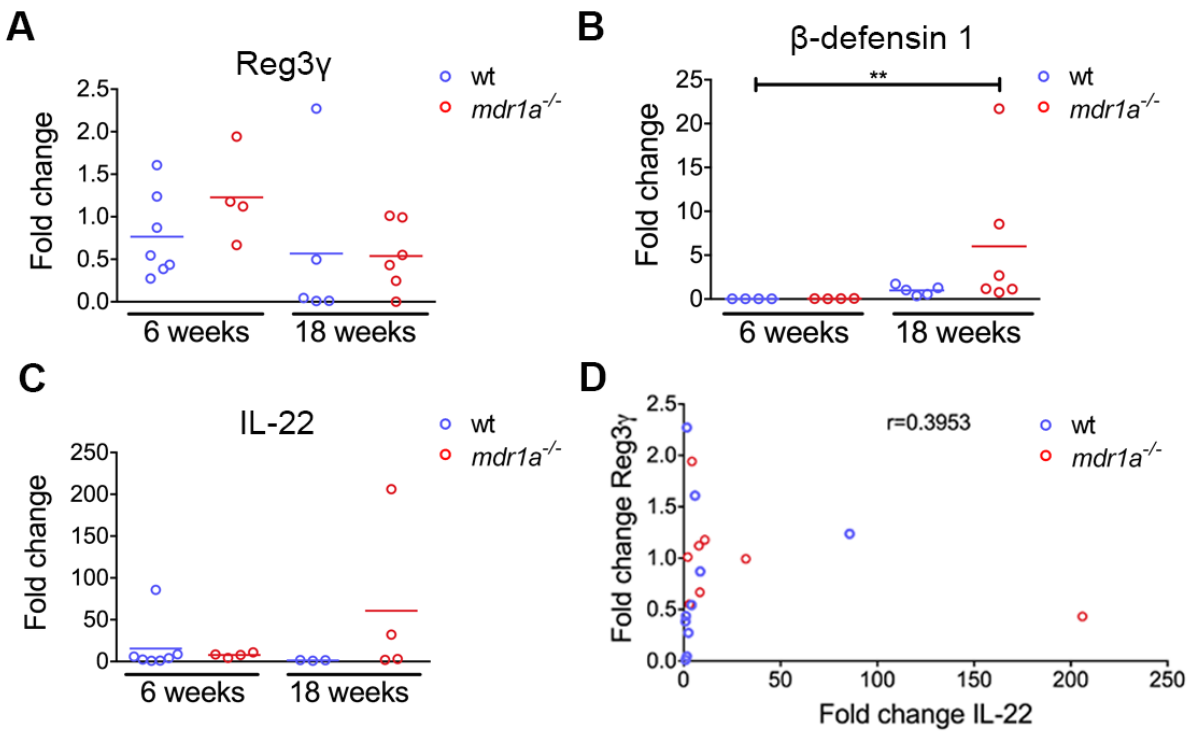


Figure 3. Expression levels of Reg3γ, β-defensin, and IL-22. mRNA expression of **(A)** Reg3γ (n=4-7), **(B)** β-defensin 1 (n=4-6) and **(C)** IL-22 (n=3-7) was assessed by qPCR in proximal colon tissue samples from *mdr1a*$^{-/-}$ and wt controls at 6 and 18 weeks of age. Fold changes in gene expression after normalization to two housekeeping genes and to a 6-week old wt mouse. The data shown as dot plots with mean. **(D)** Correlation analysis of Reg3γ and IL-22 gene expression. All data were used for correlation analysis. ** P<0.01 by Kruskal-Wallis test with Dunn's multiple comparisons test in **(A, B and C)**. Pearson correlation analysis in **(D)**.

4.4.2 IgA expression in *mdr1a*$^{-/-}$ and wt mice

Secretory IgA produced by plasma cells contributes to epithelial barrier integrity facilitating microbial clearance, and thus regulates microbial presence (Henderson et al., 2011). IgA positive cells in distal colon tissue sections were quantified by immunohistochemistry (Figure 4A). The number of IgA positive cells was comparable between wt and *mdr1a*$^{-/-}$ mice at 6 weeks of age (Figure 4B). At 18 weeks, elevated numbers of IgA-producing plasma cells in *mdr1a*$^{-/-}$ mice were consistent with the exacerbated colitis in those individuals.

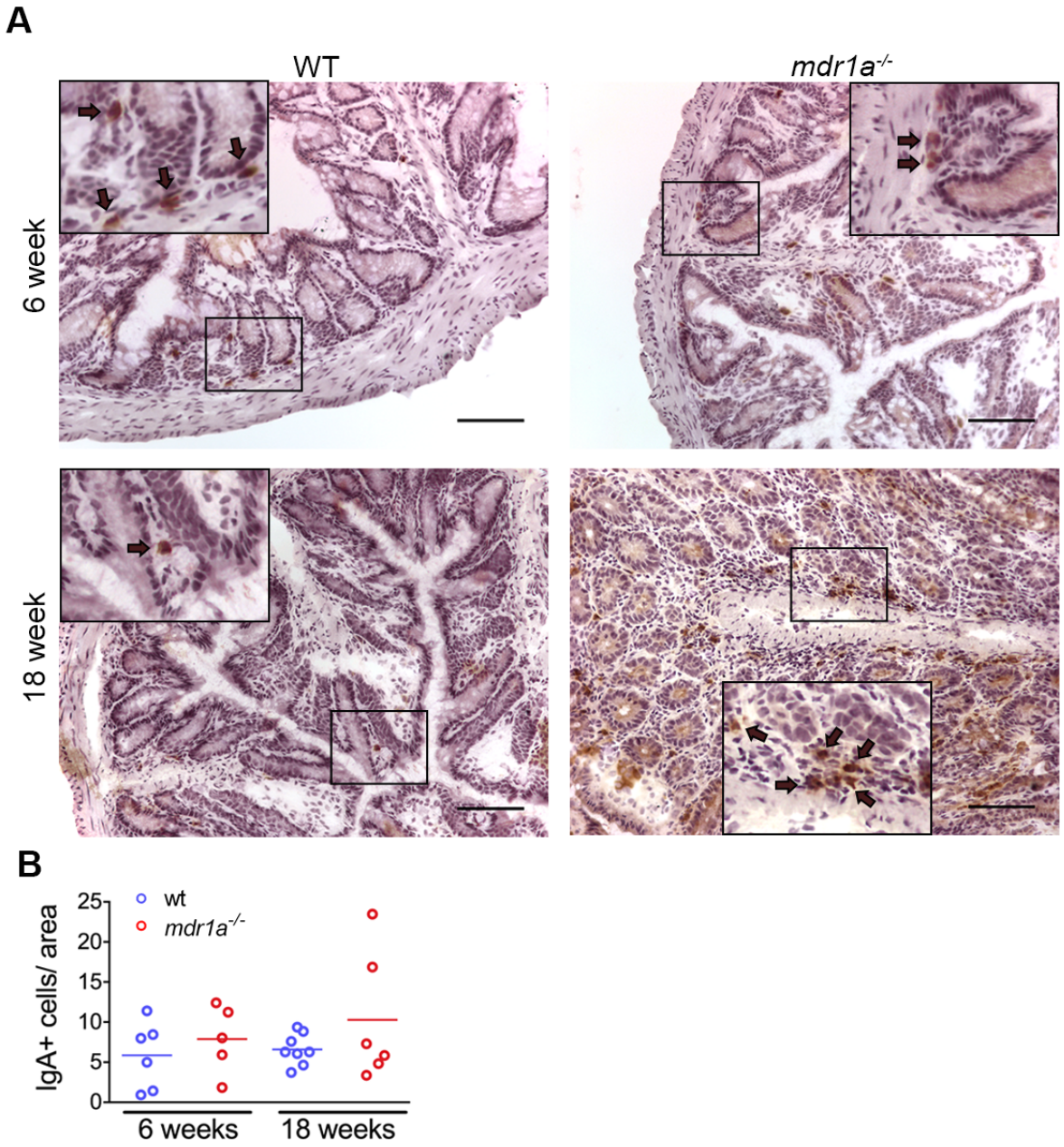


Figure 4. Expression of IgA in *mdr1a*$^{-/-}$ and wt mice. (A) Representative images from IgA immunohistochemical staining in distal colon in *mdr1a*$^{-/-}$ and wt mice at 6 and 18 weeks of age. Scale bar= 100µm. Arrows in inset images indicate positive cells. **(B)** Quantification of IgA positive cells per field of view. N=5-8. Data shown as plots with mean. Kruskal-Wallis test with Dunn's multiple comparisons test.

4.5Discussion

Our findings show similar patterns of AMP expression between wt and *mdr1a*$^{-/-}$ mice 12 weeks before the onset of intestinal inflammation. Although we have previously observed changes in mucus microbial composition between wt and *mdr1a*$^{-/-}$ mice 12 weeks prior to colitis onset (Glymenaki et al., 2016), herein we demonstrate that the colonic expression of Relm-β, Reg3γ, Ang4 and β-defensin 1 is not altered. In contrast, previous studies on *mdr1a*$^{-/-}$ mice have demonstrated a dramatic reduction of the antimicrobials Reg3γ and PAP before the onset of colonic inflammation (Collett et al., 2008; Dommels et al., 2007).

Differences in animal husbandry and maintenance between these studies and our own could account for the apparent discrepancies in the findings especially in the case of Reg3γ. Animals were not bred to generate littermate controls and were housed in individual cages (Collett et al., 2008; Dommels et al., 2007); however, in our study we used littermate controls maintained under co-housing conditions. Importantly, a previous study on *Nod2*$^{-/-}$ mice has pinpointed the effect of litter and housing conditions on the expression profile of AMPs (Shanahan et al., 2014). Age is another important factor that could have influenced the disparate results between the previous studies and our own. Previous studies have used 4-week (Collett et al., 2008) and 12-week old (Dommels et al., 2007) *mdr1a*$^{-/-}$ mice, whereas we used mice at 6 weeks of age. Notably, we found differences in Relm-β and Ang4 expression between wt and *mdr1a*$^{-/-}$ mice once they have been weaned at 4 weeks of age, suggesting that as these mice get older the differences in AMP levels drop off. Moreover, we analyzed mRNA transcript levels of AMPs in proximal colon, whereas the other studies have used whole colon and CECs (Collett et al., 2008; Dommels et al., 2007). Therefore, it could be possible that regional differences in AMPs may have influenced our results. For instance, the AMPs Reg3γ, Reg3β and Ang4 have shown a heterogeneous pattern of expression between the proximal and distal colon, where they were expressed in the proximal but were almost absent in the distal colon (Burger-van Paassen et al., 2012). Relm-β exhibited an opposite pattern of expression with highest levels in the distal colon and caecum, and low expression in the proximal colon (He et al., 2003).

We found no significant differences in AMPs expression between control and *mdr1a*$^{-/-}$ mice at onset of inflammation, however there was high individual variability within groups. Expression of AMPs showed differences at stages before and during colitis onset, indicating

an age-dependent effect on AMPs expression in agreement with a previous study (Burger-van Paassen et al., 2012). Increased β-defensin 1 and IL-22 expression and IgA-producing B cell numbers were only observed in mice exhibiting severe inflammation. Studies in patients with established IBD and in animal models of chemically induced colitis or infectious colitis have shown upregulation of β-defensin 2, Reg3γ, PAP or Relm-β in inflamed colonic epithelial tissue (Dieckgraefe et al., 2002; Hogan et al., 2006; Mizoguchi et al., 2003; Nair et al., 2008; O'Neil et al., 1999; Ogawa et al., 2003). Increased AMPs expression may acts as a mechanism to counteract increased bacterial colonization and intestinal epithelium penetration.

Overall, our data indicate similar expression of antimicrobial effector molecules prior to colonic inflammation, despite changes in mucus microbial communities. Therefore, other factors that remain to be identified contribute to alterations in mucus microbiota composition. Since we have observed reduced mucus thickness in addition to altered mucus microbiota composition before the onset of intestinal inflammation (Glymenaki et al., 2016), it is highly likely that mucus properties such as glycosylation are changed and thus have an effect on mucus microbiota. In fact, mucus glycosylation can affect microbial composition and thereby host physiology and susceptibility to inflammation (Goto et al., 2014; Sommer et al., 2014). Future studies on mucus properties in *mdr1a*$^{-/-}$ mice are warranted.

Acknowledgements

We thank P. Davies and M. Little for technical assistance. This work was supported by a UK Biotechnology and Biological Sciences Research Council (BBSRC) studentship for MG awarded to SC, KE, AM and GW. The authors declare no conflicts of interest.

References

Bevins, C.L., and Salzman, N.H. (2011). Paneth cells, antimicrobial peptides and maintenance of intestinal homeostasis. Nature reviews. Microbiology *9*, 356-368.

Bowcutt, R., Forman, R., Glymenaki, M., Carding, S.R., Else, K.J., and Cruickshank, S.M. (2014). Heterogeneity across the murine small and large intestine. World journal of gastroenterology : WJG *20*, 15216-15232.

Brandl, K., Plitas, G., Schnabl, B., DeMatteo, R.P., and Pamer, E.G. (2007). MyD88-mediated signals induce the bactericidal lectin RegIII gamma and protect mice against intestinal Listeria monocytogenes infection. The Journal of experimental medicine *204*, 1891-1900.

Burger-van Paassen, N., Loonen, L.M., Witte-Bouma, J., Korteland-van Male, A.M., de Bruijn, A.C., van der Sluis, M., Lu, P., Van Goudoever, J.B., Wells, J.M., Dekker, J.*, et al.* (2012). Mucin Muc2 deficiency and weaning influences the expression of the innate defense genes Reg3beta, Reg3gamma and angiogenin-4. PloS one *7*, e38798.

Collett, A., Higgs, N.B., Gironella, M., Zeef, L.A., Hayes, A., Salmo, E., Haboubi, N., Iovanna, J.L., Carlson, G.L., and Warhurst, G. (2008). Early molecular and functional changes in colonic epithelium that precede increased gut permeability during colitis development in mdr1a(-/-) mice. Inflammatory bowel diseases *14*, 620-631.

Dieckgraefe, B.K., Crimmins, D.L., Landt, V., Houchen, C., Anant, S., Porche-Sorbet, R., and Ladenson, J.H. (2002). Expression of the regenerating gene family in inflammatory bowel disease mucosa: Reg Ialpha upregulation, processing, and antiapoptotic activity. Journal of investigative medicine : the official publication of the American Federation for Clinical Research *50*, 421-434.

Dommels, Y.E., Butts, C.A., Zhu, S., Davy, M., Martell, S., Hedderley, D., Barnett, M.P., McNabb, W.C., and Roy, N.C. (2007). Characterization of intestinal inflammation and identification of related gene expression changes in mdr1a(-/-) mice. Genes & nutrition *2*, 209-223.

Fahlgren, A., Hammarstrom, S., Danielsson, A., and Hammarstrom, M.L. (2004). beta-Defensin-3 and -4 in intestinal epithelial cells display increased mRNA expression in ulcerative colitis. Clinical and experimental immunology *137*, 379-385.

Forman, R., Bramhall, M., Logunova, L., Svensson-Frej, M., Cruickshank, S.M., and Else, K.J. (2016). Eosinophils may play regionally disparate roles in influencing IgA(+) plasma cell numbers during large and small intestinal inflammation. BMC immunology *17*, 12.

Forman, R.A., deSchoolmeester, M.L., Hurst, R.J., Wright, S.H., Pemberton, A.D., and Else, K.J. (2012). The goblet cell is the cellular source of the anti-microbial angiogenin 4 in the large intestine post Trichuris muris infection. PloS one *7*, e42248.

Friswell, M.K., Gika, H., Stratford, I.J., Theodoridis, G., Telfer, B., Wilson, I.D., and McBain, A.J. (2010). Site and strain-specific variation in gut microbiota profiles and metabolism in experimental mice. PloS one *5*, e8584.

Gallo, R.L., and Hooper, L.V. (2012). Epithelial antimicrobial defence of the skin and intestine. Nature reviews. Immunology *12*, 503-516.

Glymenaki, M., Brass, A., Warhurst, G., McBain, A.J., Else, K.J., and Cruickshank, S.M. (2016). Differences in gut mucus microbiota precede the onset of colitis-induced inflammation. Submitted.

Goto, Y., Obata, T., Kunisawa, J., Sato, S., Ivanov, II, Lamichhane, A., Takeyama, N., Kamioka, M., Sakamoto, M., Matsuki, T.*, et al.* (2014). Innate lymphoid cells regulate intestinal epithelial cell glycosylation. Science *345*, 1254009.

He, W., Wang, M.L., Jiang, H.Q., Steppan, C.M., Shin, M.E., Thurnheer, M.C., Cebra, J.J., Lazar, M.A., and Wu, G.D. (2003). Bacterial colonization leads to the colonic secretion of RELMbeta/FIZZ2, a novel goblet cell-specific protein. Gastroenterology *125*, 1388-1397.

Henderson, P., van Limbergen, J.E., Schwarze, J., and Wilson, D.C. (2011). Function of the intestinal epithelium and its dysregulation in inflammatory bowel disease. Inflammatory bowel diseases *17*, 382-395.

Hogan, S.P., Seidu, L., Blanchard, C., Groschwitz, K., Mishra, A., Karow, M.L., Ahrens, R., Artis, D., Murphy, A.J., Valenzuela, D.M.*, et al.* (2006). Resistin-like molecule beta regulates innate colonic function: barrier integrity and inflammation susceptibility. The Journal of allergy and clinical immunology *118*, 257-268.

Kostic, A.D., Xavier, R.J., and Gevers, D. (2014). The microbiome in inflammatory bowel disease: current status and the future ahead. Gastroenterology *146*, 1489-1499.

Meyer-Hoffert, U., Hornef, M.W., Henriques-Normark, B., Axelsson, L.G., Midtvedt, T., Putsep, K., and Andersson, M. (2008). Secreted enteric antimicrobial activity localises to the mucus surface layer. Gut *57*, 764-771.

Mizoguchi, E., Xavier, R.J., Reinecker, H.C., Uchino, H., Bhan, A.K., Podolsky, D.K., and Mizoguchi, A. (2003). Colonic epithelial functional phenotype varies with type and phase of experimental colitis. Gastroenterology *125*, 148-161.

Nair, M.G., Guild, K.J., Du, Y., Zaph, C., Yancopoulos, G.D., Valenzuela, D.M., Murphy, A., Stevens, S., Karow, M., and Artis, D. (2008). Goblet cell-derived resistin-like molecule beta augments CD4+ T cell production of IFN-gamma and infection-induced intestinal inflammation. J Immunol *181*, 4709-4715.

O'Neil, D.A., Porter, E.M., Elewaut, D., Anderson, G.M., Eckmann, L., Ganz, T., and Kagnoff, M.F. (1999). Expression and regulation of the human beta-defensins hBD-1 and hBD-2 in intestinal epithelium. J Immunol *163*, 6718-6724.

Ogawa, H., Fukushima, K., Naito, H., Funayama, Y., Unno, M., Takahashi, K., Kitayama, T., Matsuno, S., Ohtani, H., Takasawa, S.*, et al.* (2003). Increased expression of HIP/PAP and regenerating gene III in human inflammatory bowel disease and a murine bacterial reconstitution model. Inflammatory bowel diseases *9*, 162-170.

Peterson, L.W., and Artis, D. (2014). Intestinal epithelial cells: regulators of barrier function and immune homeostasis. Nature reviews. Immunology *14*, 141-153.

Ramanan, D., Tang, M.S., Bowcutt, R., Loke, P., and Cadwell, K. (2014). Bacterial sensor Nod2 prevents inflammation of the small intestine by restricting the expansion of the commensal Bacteroides vulgatus. Immunity *41*, 311-324.

Salzman, N.H., Hung, K., Haribhai, D., Chu, H., Karlsson-Sjoberg, J., Amir, E., Teggatz, P., Barman, M., Hayward, M., Eastwood, D*., et al.* (2010). Enteric defensins are essential regulators of intestinal microbial ecology. Nature immunology *11*, 76-83.

Sanos, S.L., Bui, V.L., Mortha, A., Oberle, K., Heners, C., Johner, C., and Diefenbach, A. (2009). RORgammat and commensal microflora are required for the differentiation of mucosal interleukin 22-producing NKp46+ cells. Nature immunology *10*, 83-91.

Schinkel, A.H., Smit, J.J., van Tellingen, O., Beijnen, J.H., Wagenaar, E., van Deemter, L., Mol, C.A., van der Valk, M.A., Robanus-Maandag, E.C., te Riele, H.P., and et al. (1994). Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. Cell *77*, 491-502.

Shanahan, M.T., Carroll, I.M., Grossniklaus, E., White, A., von Furstenberg, R.J., Barner, R., Fodor, A.A., Henning, S.J., Sartor, R.B., and Gulati, A.S. (2014). Mouse Paneth cell antimicrobial function is independent of Nod2. Gut *63*, 903-910.

Sommer, F., Adam, N., Johansson, M.E., Xia, L., Hansson, G.C., and Backhed, F. (2014). Altered mucus glycosylation in core 1 O-glycan-deficient mice affects microbiota composition and intestinal architecture. PloS one *9*, e85254.

Suzuki, K., Meek, B., Doi, Y., Muramatsu, M., Chiba, T., Honjo, T., and Fagarasan, S. (2004). Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. Proceedings of the National Academy of Sciences of the United States of America *101*, 1981-1986.

Vaishnava, S., Behrendt, C.L., Ismail, A.S., Eckmann, L., and Hooper, L.V. (2008). Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface. Proceedings of the National Academy of Sciences of the United States of America *105*, 20858-20863.

Vaishnava, S., Yamamoto, M., Severson, K.M., Ruhn, K.A., Yu, X., Koren, O., Ley, R., Wakeland, E.K., and Hooper, L.V. (2011). The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. Science *334*, 255-258.

Vora, P., Youdim, A., Thomas, L.S., Fukata, M., Tesfay, S.Y., Lukasek, K., Michelsen, K.S., Wada, A., Hirayama, T., Arditi, M., and Abreu, M.T. (2004). Beta-defensin-2 expression is regulated by TLR signaling in intestinal epithelial cells. J Immunol *173*, 5398-5405.

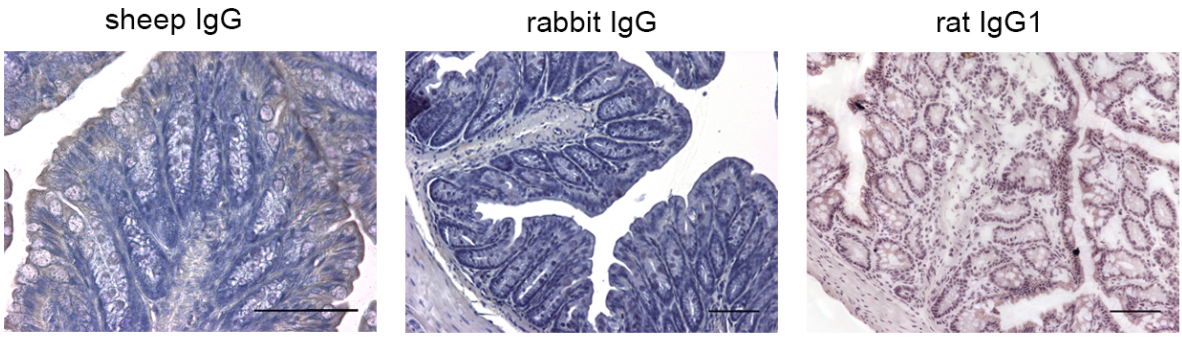
Wehkamp, J., Harder, J., Weichenthal, M., Mueller, O., Herrlinger, K.R., Fellermann, K., Schroeder, J.M., and Stange, E.F. (2003). Inducible and constitutive beta-defensins are differentially expressed in Crohn's disease and ulcerative colitis. Inflammatory bowel diseases *9*, 215-223.

Wei, M., Shinkura, R., Doi, Y., Maruya, M., Fagarasan, S., and Honjo, T. (2011). Mice carrying a knock-in mutation of Aicda resulting in a defect in somatic hypermutation have impaired gut homeostasis and compromised mucosal defense. Nature immunology *12*, 264-270.

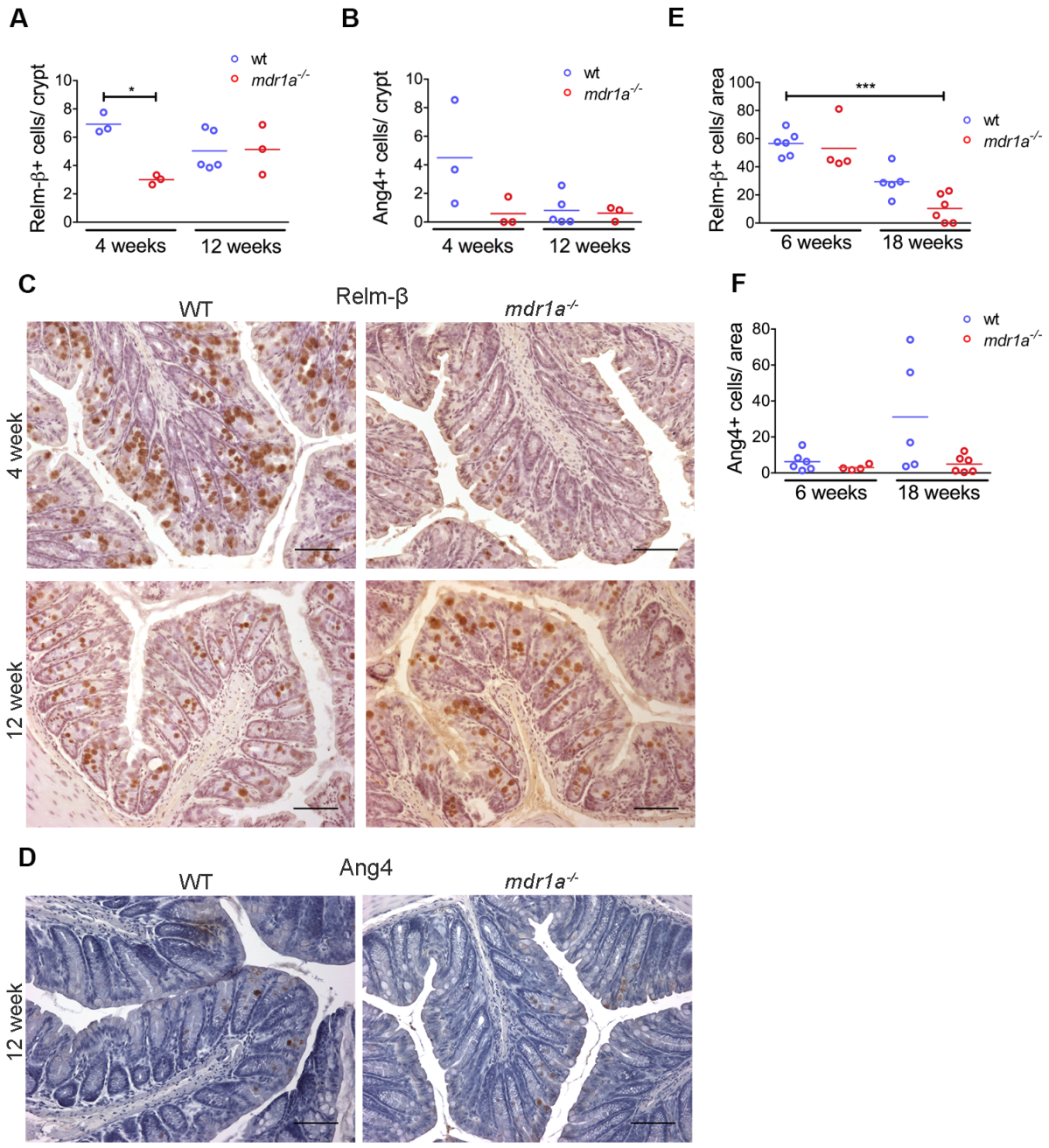
Supplementary materials

Supplementary Table 1. List of primers and probes used for qPCR.

Gene	Primer sequence	Universal Probe Library Set
β-defensin 1	Left: 5'-CCTGGCTGCCACCACTAT-3' Right: 5'-CTTGTGAGAATGCCAACACCT-3'	Probe 2
Relm-β	Left: 5'-GGAAGCTCTCAGTCGTCAAGA-3' Right: 5'-GCACATCCAGTGACAACCAT-3'	Probe 18
Reg3γ	Left: 5'-GCTTCCCCGTATAACCATCA-3' Right: 5'-GCATCTTTCTTGGCAACTTCA-3'	Probe 51
Ang4	Left: 5'-CCCCAGTTGGAGGAAAGC-3' Right: 5'-CGTAGGAATTTTTCGTACCTTTCA-3'	Probe 106
IL-22	Left: 5'-TGACGACCAGAACATCCAGA-3' Right: 5'-AATCGCCTTGATCTCTCCAC-3'	Probe 94
β-actin	Left: 5'- CTAAGGCCAACCGTGAAAAG-3' Right: 5'- ACCAGAGGCATACAGGGACA-3'	Probe 64
villin	Left: 5'-ATCTCCCTGAGGGTGTGGA-3' Right: 5'- AGTGAAGTCTTCGGTGGACAG-3'	Probe 4



Supplemental Figure 1. Isotype control staining. Representative images showing isotype control staining in distal colon. Scale bar represents =100μm.



Supplemental Figure 2. Differential expression of Relm-β and Ang4. Quantification of Relm-β **(A)** and Ang4 **(B)** staining is shown as numbers of Relm-β or Ang4 positive cells per crypt in the distal colon at 4 and 12 weeks of age. N=3-5. Representative Relm-β **(C)** and Ang4 **(D)** immunohistochemistry sections from the distal colon. Scale bar =100µm. Numbers of Relm-β **(E)** and Ang4 **(F)** positive cells per 100µm^2 area in the distal colon of wt and *mdr1a*$^{-/-}$ mice at 6 and 18 weeks of age. N=4-6. Data shown as dot plots with mean **(A, B, E and F).** * P<0.05, *** P<0.001 as determined by Kruskal-Wallis test with Dunn's multiple comparisons test.

Chapter 5

Summary discussion

5.1 Inflammatory bowel disease: the role of microbiome

Inflammatory bowel disease (IBD) including Crohn's disease (CD) and ulcerative colitis (UC) is a chronic relapsing inflammatory disorder affecting the gastrointestinal (GI) tract with an increasing prevalence in the developed world (Khor et al., 2011; Saleh and Elson, 2011). Environmental factors associated with a modern lifestyle have been proposed to be responsible for the rise in IBD as exemplified in the hygiene hypothesis (Strachan, 1989). Decreased exposure to infectious bacterial agents due to enhanced sanitation and widespread antibiotics usage, as well as alterations in eating habits characterized by a high-fat/ low-fibre diet modulate the gut microbiota with an effect on immune homeostasis (Burkitt, 1973; Sonnenburg and Sonnenburg, 2014). IBD is a state of dysregulated host-microbiota interactions, wherein an exaggerated immune response directed at the commensal bacteria is generated in genetically susceptible individuals (Saleh and Elson, 2011).

Disruption of host-microbe interactions has shaped the landscape of IBD pathophysiology (Wlodarska et al., 2015). Genetic studies have shown that polymorphisms in regulatory pathways essential for intestinal homeostasis such as epithelial barrier function, host defence, immune activation and tolerance act as susceptibility risks for IBD development (Jostins et al., 2012; Khor et al., 2011; Maloy and Powrie, 2011). Furthermore, despite the existence of IBD genetic variants in engineered mouse models, disease does not develop when these mice are raised in germ-free conditions (Khor et al., 2011; Manichanh et al., 2012). IBD patients and experimental animal models have consistently shown alterations in microbiota composition, so-called dysbiosis (Huttenhower et al., 2014; Manichanh et al., 2012; Rajilic-Stojanovic et al., 2013). Collectively, these data indicate that the gut microbiota plays an essential role in IBD pathology.

A key unresolved issue is whether changes in microbiota composition are a secondary effect of the chronic intestinal inflammation or precede the development of IBD and contribute to its initiation and progression or both. To address some of these issues, this thesis has used the *mdr1a*$^{-/-}$ spontaneous model of UC (Panwala et al., 1998) in order to examine whether alterations in microbiota structure precede the onset of intestinal inflammation. Since most of the studies on microbial composition in IBD have focused on microbial changes in faeces, we have sampled mucus in addition to faeces to clarify whether both mucus and faecal microbial

communities or either one of them differ before the manifestation of signs of inflammation. We have investigated microbiota localisation, function and epithelial cell responses at stages before and during onset of inflammation to gain a more holistic view of the microbial community organization and regulation. Also, we have addressed potential changes in metabolite profiles to better understand how the host-microbiota crosstalk is affected during the course of colitis progression. This work could indicate pathways or targets for the design of new disease therapeutics and could advance the discovery of potential disease biomarkers in IBD.

In Chapter 2, we defined the gut microbiota composition in *mdr1a*$^{-/-}$ mice and their wild-type littermate controls in both faeces and mucus before and during colitis onset. Notably, our findings indicated that alterations in gut microbiota composition preceded the onset of intestinal inflammation and these changes were observed only in the mucus layer but not in faeces. Previously published studies in IBD research have also reported changes in microbiota structure, but in established disease. Furthermore, our results showed alterations in microbial composition in both mucus and faeces at onset of inflammation. These data suggest that microbial changes start in the mucus layer and then proceed to the lumen in concomitance with inflammation progression and colitis development. Moreover, we characterized microbiota localisation in the mucus layer and observed a reduction in the thickness of the inner colonic mucus layer in m*dr1a*$^{-/-}$ mice prior to inflammation and increased microbial penetration into the inner mucus layer during colitis onset.

In Chapter 3, we examined whether alterations in microbial gene functional capacity and urinary metabolite profiles followed changes in microbial taxonomic composition. The microbial gene content was similar between *mdr1a*$^{-/-}$ and wild-type mice supporting a stability of the gut microbiome at the face of ensuing gut inflammation. Furthermore, our data showed that metabolites previously reported to change in IBD were similar between *mdr1a*$^{-/-}$ and wild-type mice at stages preceding and during inflammation. Also, disease scores could not be predicted based on variations in metabolite profiles within *mdr1a*$^{-/-}$ mice, indicating that changes in metabolite profiles were unrelated to inflammation.

In Chapter 4, we investigated whether changes in the expression of antimicrobial proteins (AMPs), which are present in the large intestine, could account for observed differences in microbial composition. The findings were similar levels of expression of the antimicrobials

regenerating islet-derived protein 3 γ (Reg3γ), angiogenin 4 (Ang4), β-defensin 1 and resistin-like molecule beta (Relm-β) between *mdr1a*$^{-/-}$ and wild-type mice, suggesting that these antimicrobials do not contribute to the altered bacterial composition. We hypothesized that other mechanisms drive alterations in microbiota composition especially in the mucus prior to the onset of colitis.

5.2 Does the gut microbiota change before the onset of IBD?

IBD is associated with alterations in gut microbiota composition and reduced bacterial diversity (Huttenhower et al., 2014; Wlodarska et al., 2015). Enterobacteriaceae and particularly *Escherichia coli* are enriched in patients with IBD and mouse models, whereas Clostridia clades IV and XIVa including *Roseburia* and *Faecalibacterium prausnitzii*, and members of the Bacteroidetes phylum are markedly reduced (Huttenhower et al., 2014; Kostic et al., 2014; Lupp et al., 2007; Manichanh et al., 2012). Studies in individuals with IBD have examined the gut microbiota composition either during disease onset before treatment (Gevers et al., 2014) or after establishment of disease and clinical management (Kang et al., 2010; Lepage et al., 2011; Manichanh et al., 2006; Martinez et al., 2008; Papa et al., 2012; Willing et al., 2010). Furthermore, most of these studies have largely examined microbial composition in the faeces, which does not necessarily reflect microbial configuration in local niches such as the mucus layer (Kang et al., 2010; Manichanh et al., 2006; Papa et al., 2012). To identify whether the gut microbiota changes before the onset of colitis, we studied microbial composition in both mucus and faeces 12 weeks prior to the start of inflammation in *mdr1a*$^{-/-}$ and wild-type littermate controls. Strikingly, we found that alterations in gut microbiota composition preceded the onset of intestinal inflammation and these changes were confined to the mucus layer.

In agreement with our findings, the largest IBD-related microbiome study to date has shown that treatment-naïve CD patients at the time of disease diagnosis had an imbalance in their microbial community network taken from tissue biopsies akin to our mucus sampling, whereas faecal microbial communities from these CD patients resembled those of healthy individuals (Gevers et al., 2014). Therefore, mucus sampling is critical for better understanding the pathways that govern microbe-host interaction. A possible explanation for the manifestation of differences in the mucus layer but not in faecal samples could be that the

mucus-associated bacteria are closer to the intestinal epithelium and the underlying immune system, so they are likely to have the greatest functional impact on host health and conversely most likely to be affected by the epithelial and immune compartments. Notably, identification of mucus microbial differences in the treatment-naïve CD cohort was only performed at the time of diagnosis and thus concurrent with the manifestation of disease symptoms (Gevers et al., 2014); however, in this study we were able to observe alterations prior to the onset of colitis.

Microbial dysbiosis is also associated with susceptibility to other conditions. Evidence derived from skin inflammatory disorders including atopic dermatitis (also known as eczema) and psoriasis implicates microbial communities in disease pathogenesis (Gao et al., 2008; Kong et al., 2012; Williams and Gallo, 2015). For instance, atopic dermatitis (AD) is characterized by a marked reduction in skin bacterial diversity and overgrowth of *Staphylococcus* species and in particular *Staphylococcus aureus (S. aureus)* during disease flares (Kong et al., 2012). Dysbiosis involving *S. aureus* and *Corynebacterium* species has also been observed in primary immunodeficiencies that manifest an eczematous dermatitis nearly identical to AD, further supporting a role for dysbiosis in skin inflammation (Kubo et al., 2012; Oh et al., 2013). However, it has not been clarified whether changes in skin microbiota composition are the result of inflammation or the cause of it. A recent elegant study performed in *Adam17*$^{fl/fl}$*Sox9*-^{Cre} mice that lack the protease a disintegrin and metalloproteinase 17 (ADAM17) in SRY (Sex Determining Region Y)-Box 9 (Sox9)-expressing tissues including epidermis revealed that dysbiosis indeed drives eczematous inflammation (Kobayashi et al., 2015). These mice developed eczematous skin and dysbiosis that recapitulated human AD and eczematous dermatitis. Dysbiotic flora including *S. aureus* and *Corynebacterium bovis* contributed differentially to skin inflammation. *S. aureus,* which emerged prior to the onset of inflammation, induced eczema formation, whereas *Corynebacterium bovis* emerged later and potently drove T helper 2 (T_H2) cell responses (Kobayashi et al., 2015). These findings demonstrate that a shift in microbiota composition, in which *S. aureus* and *Corynebacterium* species were major constituents, is a pathological factor for eczematous dermatitis development.

Given that we observed alterations in mucus microbial communities in *mdr1a*$^{-/-}$ mice before the start of intestinal inflammation, we could speculate that these microbial changes drive inflammation and have a causative role as was shown in eczematous dermatitis (Kobayashi et

al., 2015). In support of the role of microbial alterations in colitis pathophysiology, studies in IBD have reported changes in microbial composition in sites out of the boundaries of the inflammatory lesions, suggesting that they are not a secondary effect of mucosal inflammation (Strober et al., 2007; Swidsinski et al., 2005). To confirm this hypothesis, further experiments on germ-free mice colonized with specific bacterial members as well as a longitudinal analysis of sequential sampling of mucus and faecal microbial communities from the same mice until colitis development would be required.

Skin microbiota, although different from gut microbiota in density, diversity, habitation niches and effect on the immune system, also regulates innate and adaptive immunity (Belkaid and Segre, 2014). So, despite the differences between the microbial populations inhabiting skin and gut, both of them share the capacity to modulate the immune response. Additionally, the incidence of atopic dermatitis has doubled over time but with no clear cause (Asher et al., 2006), resembling the constant rise of IBD. *Mdr1a*$^{-/-}$ mice have an impaired epithelial barrier function due to the lack of P-glycoprotein, a xenobiotic transporter, (Panwala et al., 1998; Schinkel et al., 1997), while *Adam17*$^{fl/fl}$*Sox9*-^{Cre} mice have defective epidermal growth factor receptor (EGFR) signalling (Kobayashi et al., 2015). Interestingly AD is characterized by a disruption in stratum corneum barrier due to mutations in the structural protein filaggrin (Irvine et al., 2011). Hence, both *mdr1a*$^{-/-}$ and *Adam17*$^{fl/fl}$*Sox9*-^{Cre} mice are characterized by epithelial barrier dysfunction. All these observations including the immunomodulatory capacity of both skin and gut microbiota, the increasing rise and prevalence of AD similar to IBD and the implication of the epithelial barrier in both IBD and AD add support to the hypothesis that the gut microbiota is highly likely to be involved in the initiation and progression of IBD in a manner similar to inflammatory skin disorders.

Alterations in the mucus microbiota encompassed reduced representation of Clostridiales and enhanced abundance of the S24-7 group that belongs to Bacteroidales in *mdr1a*$^{-/-}$ animals relative to wild-type controls. Clostridia have been associated with the development of regulatory T cells (Tregs) (Atarashi et al., 2013; Atarashi et al., 2011; Furusawa et al., 2013) and *mdr1a*$^{-/-}$ animals have decreased levels of forkhead box P3 (Foxp3^{+}) Tregs in the lamina propria (Tanner et al., 2013). Therefore, the reduction in mucus Clostridiales in *mdr1a*$^{-/-}$ animals could have impaired the generation of Tregs, resulting in loss of immunoregulatory signals and breakdown of immune homeostasis, which could in turn drive inflammation. We have not assessed the frequency of the Treg compartment in the present study, but analysis of

immune cells and expression of anti-inflammatory/ pro-inflammatory markers in the same tissue samples from which mucus and faecal microbial communities were derived shall be performed in future analyses.

Our results showed no significant differences in the numbers of Gram-positive and Gram-negative bacteria between *mdr1a*$^{-/-}$ and wild-type mice. A possible explanation is that the Gram-positive probe used was designed using bacteria affiliated with the Bacilli (Wu et al., 2008) and thus Clostridia including taxa Lachnospiraceae and Ruminococcaceae were less well targeted. Also, the Gram-negative probe did not detect all the members of the Bacteroidetes phylum such as Bacteroidaceae and Prevotellaceae. A more targeted analysis using probes for each major taxonomic group could provide further insight on the balance of certain bacterial taxa before the onset of inflammation.

To examine whether changes in microbial communities preceded onset of inflammation in other experimental models of colitis except from *mdr1a*$^{-/-}$, we collected samples from interleukin-10-deficient (IL-10$^{-/-}$) mice infected with a high-dose of *Trichuris muris* at day 16 of infection. IL-10$^{-/-}$ mice are more susceptible to chronic inflammation in response to infection due to their inability to mount an effective T_H2 response (Schopf et al., 2002). We selected day 16 of infection, because day 21 is associated with peak of cytokine responses and is the time when worm expulsion begins. We observed differences in microbial communities between IL-10$^{-/-}$ and wild-type mice before the cytokine peak (Appendix). However, a caveat of this study is that only faecal samples were collected, as mice were harvested for another experiment at a later time point. Furthermore, wild-type mice used in the study were bought in and were not littermate controls with *IL-10*$^{-/-}$ mice. Therefore, a repeat of this study with a better design is needed; samples should be taken from both mucus and faecal microbial communities and mice used should be littermate controls. Selection of an appropriate mouse model of colitis to confirm our findings is complicated by the fact that IBD is a heterogeneous group of disorders regarding disease location, behaviour and severity (Saleh and Elson, 2011; Strober et al., 2007). Furthermore, murine models of experimental colitis rely on disrupting the precarious immune system balance (either by driving over-activation or impairing tolerance), or adding chemical or infectious agents that drive IBD. Thus, some of the experimental models have already a dysfunctional immune system or altered microbial compartment, making it harder to identify and isolate the role of dysbiosis in driving inflammation. The mdr1a model of colitis offers the advantage of being an

epithelial barrier model with an intact immune system (Panwala et al., 1998) and hence enabling focus on the crosstalk between bacteria and mucosa. Other spontaneous models of colitis that could be useful in studying potential shifts in microbiota composition before disease onset include models with defects in mucus glycans such as core 1-derived O-glycans (Fu et al., 2011) or in immune regulation (Saleh and Elson, 2011), however, the kind of gene depletion involved should be considered when interpreting results, as it may have impacted on immune function or barrier function.

5.3 Does dysbiosis causes IBD in genetically susceptible individuals?

In our study we reported alterations in microbiota structure in *mdr1a*$^{-/-}$ mice before the development of any signs of inflammation. The *mdr1a*$^{-/-}$ is a genetically engineered model of colitis that lacks completely the *Mdr1a* gene, so our findings suggest that dysbiosis drives IBD in the context of a genetically susceptible setting. A pending question is whether microbiota changes cause inflammation or further stimuli are required. Studies in animal models of colitis have demonstrated that microbiota shifts predispose to colitis but require additional insults –genetic, immunologic or chemical- for disease manifestation (Huttenhower et al., 2014; Strober, 2013). The first study showing that gut microbiota can drive spontaneous colitis *de novo* in normal hosts came from mice lacking recombination-activating gene 2 (RAG2) and the transcriptional factor T-bet, also termed TRUC mice due to immunodeficiency (TR) and resemblance to human UC (UC), which transmitted colitis to co-housed wild-type controls (Garrett et al., 2010). TRUC mice possessed a colitogenic microbiota characterized by *Klebsiella pneumoniae* and *Proteus mirabilis,* which was responsible for colitis transmission. Thus, can these colitogenic bacteria transmit colitis without the presence of an underlying genetic predisposition to IBD? Administration of Tregs ameliorated colitis in TRUC mice without affecting the levels of colitogenic bacteria, indicating that these bacteria are not intrinsically colitogenic but rather require an abnormal immune microenvironment to exert their function (Garrett et al., 2010). Mice deficient in Nod-like receptor pyrin domain-containing protein 6 (NLRP6) although having an altered microbiota enriched in Prevotellaceae and TM7 did not cause spontaneous colitis in co-housed wild-type mice, although co-housing made wild-type mice more susceptibility to dextran sulfate sodium (DSS)- induced colitis (Elinav et al., 2011). Furthermore, mice

lacking nucleotide oligomerization domain 2 (NOD2), a risk gene for CD, did not develop spontaneous colitis despite alterations in microbiota composition; however they were more susceptible to severe colitis upon DSS challenge (Couturier-Maillard et al., 2013). In addition, the altered microbiota in *Nod2*$^{-/-}$ caused increased susceptibility to induced colitis in co-housed wild-type mice (Couturier-Maillard et al., 2013). These data indicate that in addition to an altered microbiota composition and genetic susceptibility sometimes an exogenous trigger is required to cause pathologic inflammation. Therefore, our observations in agreement with these studies suggest that if there is an underlying predisposition to IBD and a balanced gut microbiota, they prevent tipping the balance towards disease, but a dysbiosis tips the balance towards full-blown inflammation.

Interestingly, early antibiotic exposure is a risk factor for IBD development (Hviid et al., 2011; Shaw et al., 2010). Antibiotic treatment amplified dysbiosis in CD compared to patients not receiving antibiotics upon diagnosis (Gevers et al., 2014). As gut microbiota influences immune system development and function, early changes in microbiota composition for example due to antibiotics could impact on immune response and lead to immune-mediated pathologies later in life. Notably, healthy unaffected relatives of patients suffering from CD, who have an increased risk of developing CD, have a dysbiotic gut microbial community, further supporting a role for dysbiosis in preceding or predisposing to IBD (Hedin et al., 2016; Joossens et al., 2011).

5.4 Is microbiota localisation affected by the altered mucus microbial composition?

Spatial segregation of microbiota from direct contact with the intestinal epithelium is essential for the maintenance of immune homeostasis (Johansson et al., 2011b; Vaishnava et al., 2011). The mucus layer provides a physical barrier excluding microbes, but it also functions as a niche for microbial colonisation, as it is a source rich in mucin glycans (Johansson and Hansson, 2011; Sonnenburg et al., 2005). The mucus in the colon consists of an inner stratified firm layer that is considered devoid of bacteria and a thicker outer loose layer that commensal bacteria may reside (Johansson et al., 2008). Herein we report that *mdr1a*$^{-/-}$ mice had reduced thickness of the inner mucus layer, indicating a compromised

mucus barrier before the emergence of intestinal inflammation. Notably, reduced mucus thickness was not associated with increased microbial penetration before colitis onset. Bacteria have been shown to penetrate the inner mucus layer and reach the epithelium in experimental models of colitis and patients with active UC (Johansson et al., 2014). Mice treated with the chemical colitogen DSS exhibited increased bacterial penetration in the inner colonic mucus layer within 12 hours of treatment, long before any infiltration of inflammatory cells and signs of colitis (Johansson et al., 2010). Therefore, findings in DSS-induced colitis along with our observations in *mdr1a*$^{-/-}$ mice suggest that altered mucus properties may be an initial event triggering inflammation and thus predisposing to colitis.

Given that bacteria can influence mucus barrier properties such as thickness and permeability (Jakobsson et al., 2014; Petersson et al., 2011) and since mucus bacterial composition is altered in *mdr1a*$^{-/-}$ mice prior to colitis, we could speculate that changes in microbial community contribute to that impaired mucus barrier. Alterations in mucus glycan composition could also have impacted on mucin cross-linking affecting mucus structure. The major component of mucus, mucin MUC2, is heavily O-glycosylated and glycans contribute to about 80% of its mass and thus play a rather important role in determining mucus properties (Johansson et al., 2011a). Apart from mucins, mucus organization relies on additional proteins that stabilize the mucus by additional cross-linking (Johansson et al., 2009). Mucus organization further depends on external factors in the surrounding milieu such as calcium levels and pH (Ambort et al., 2012). As *mdr1a*$^{-/-}$ lack the xenobiotics transporter P-glycoprotein in intestinal epithelial cells (IECs) (Panwala et al., 1998), this could also change metabolites present in the local environment with a potential effect on mucus. Another possible explanation relies on alterations in mucus distribution incurred by changes in AMP. Mice deficient in Reg3γ had impaired mucus distribution, which was absent in certain parts along the small intestine (Loonen et al., 2013). Although our findings showed no differences in the expression levels of Reg3γ, this does not exclude the possibility that altered expression of other AMPs may contribute to a defective mucus barrier in *mdr1a*$^{-/-}$ mice.

Further work on altered microbial localisation is needed to understand the mechanisms behind mucus barrier compromise at a time before the onset of inflammation. Although we did not report any differences in the number of goblet cells between *mdr1a*$^{-/-}$ and wild-type mice before colitis onset, quantitative analysis of MUC2 mRNA levels is necessary in order to confirm whether alterations in mucus secretion are accountable for the reduced mucus

thickness or other mechanisms are implicated. Determining mucus glycosylation profiles as well as local metabolite niches will provide insight in pathways whose dysregulation leads to IBD pathogenesis.

5.5 Gut microbiota changes: onset of inflammation

Herein the data showed alterations in microbiota composition both in the mucus and faeces at onset of inflammation. A study performed on newly diagnosed CD patients prior to treatment reported perturbations in mucosa-associated microbial communities but not in faecal samples (Gevers et al., 2014). A limited number of bacterial taxa were associated with CD in faeces, indicating that dysbiosis has not reached the lumen yet (Gevers et al., 2014). This study supports our observation that dysbiosis starting at a local niche expands to the luminal community in parallel with advancing inflammation. It would be interesting to address the mechanisms of the system dynamics that govern microbe-microbe interactions and drive progression of dysbiosis in *mdr1a*$^{-/-}$ mice in future studies.

Our findings demonstrated that the numbers of Gram-negative bacteria were increased in faecal samples from *mdr1a*$^{-/-}$ mice and were also correlated with the severity of inflammation. Further to this, we noted an increase in mucolytic bacteria such as *Akkermansia*. It is possible that inflammation may have driven these microbial changes. Conditions prevailing in the inflamed gut provide a growth advantage to species that can adapt to environmental and nutritional changes associated with inflammation. Enhanced oxidative respiration due to elevated nitrate and thiosulfate levels in combination with a higher concentration of oxygen in the luminal microenvironment of the inflamed gut due to higher blood flow favours the growth of oxygen-tolerant organisms such as Enterobacteriaceae, whereas it impedes obligate anaerobic bacteria such as Clostridia and Bacteroidia (Baumler and Sperandio, 2016; Gevers et al., 2014; Zeng et al., 2016). The abundance of anaerobic bacteria has indeed been shown to be decreased in faecal samples from patients with established IBD (Papa et al., 2012). Although the abundance of aerotolerant and obligate anaerobic bacteria has not been assessed in this study, it would be interesting to examine the abundance of certain taxonomic groups such as Enterobacteriaceae. The increase in mucus-degrading bacteria could also provide a source of readily available mucus glycans, facilitating the expansion of potentially harmful bacteria. The expansion of *Escherichia coli* in the inflamed colon of DSS-treated mice was

favoured by the elevated levels of mucin-derived sialic acid sugars (Huang et al., 2015). Increase in glycans is also exploited by the infectious pathogens *Salmonella enterica* serovar Typhimurium and *Clostridium difficile* (Ng et al., 2013).

Studies on infection with GI pathogens such as *Clostridium difficile*, *Citrobacter rodentium*, and *Salmonella* support a role for dysbiosis in driving and amplifying inflammation (Antharam et al., 2013; Baumler and Sperandio, 2016; Koenigsknecht et al., 2015; Zeng et al., 2016). Antibiotic treatment disrupted the gut microbiota composition affecting metabolite levels of bile acids, glucose, free fatty acids, dipeptides and sugar alcohols; this altered metabolite repertoire was exploited by *Clostridium difficile* to enhance its growth and establish its colonisation, while eliciting inflammation (Modi et al., 2014; Theriot et al., 2014). Porphyromonadaceae, Lachnospiraceae, *Lactobacillus*, *Alistipes* and *Turicibacter* confer resistance to *Clostridium difficile*, whereas *Escherichia* and *Streptococcus* are associated with increased susceptibility (Baumler and Sperandio, 2016; Buffie et al., 2015; Schubert et al., 2015). Given the ability of pathogenic bacteria to flourish by invoking intestinal inflammation, it would be interesting to suppose that during inflammation in IBD opportunistic bacteria not only find the chance to thrive but also further contribute to inflammation and thus aggravate dysbiosis, culminating in an endless feed-forward loop. Indeed, a longitudinal study on the mucosa-associated microbiota in UC patients demonstrated variations in community composition over time, which corresponded to disease severity (Fite et al., 2013).

Changes in gut microbiota composition in *mdr1a*$^{-/-}$ mice occurred concurrently with loss of spatial segregation at onset of inflammation. Inner mucus thickness was markedly reduced and bacteria penetrated the inner mucus layer and even contacted the epithelium in agreement with previous studies on active inflammation in murine models of colitis (Johansson et al., 2014; Perez-Bosque et al., 2015) and IBD patients (Fyderek et al., 2009; Johansson et al., 2014; Larsson et al., 2011). Disruption of separation of microbiota with the host epithelium provides an additional trigger for launching a pathologic inflammatory response and sustaining it in colitis.

5.6 Host-microbiota crosstalk: do metabolite profiles accompany gut microbiota changes?

The gut microbiota performs functions that contribute to host metabolism and affect physiological processes at the intestinal interface (Dorrestein et al., 2014; Ursell et al., 2014). Microbes metabolize nutrients and xenobiotics producing a variety of metabolites including amino acids, bile acids, lipids, vitamins and short chain fatty acids (SCFAs), while other products comprise lipopolysaccharide (LPS), peptidoglycan and bacteriocins (Brestoff and Artis, 2013; Donia and Fischbach, 2015). Having shown differences in mucus microbiota composition before colitis onset, we investigated whether gene functional potential was also altered in these microbial communities. Notably, we observed similarities in the predicted microbial gene encoding functions between *mdr1a*$^{-/-}$ and wild-type control mice prior to inflammation despite changes in taxonomic composition. The only differentiation in microbial gene functional pathways reported was between the mucus compartment and faeces, which differ grossly in their microbial composition (Momozawa et al., 2011; Morgan et al., 2012). Previous studies have also shown that microbial communities that differ in their composition may converge to similar metabolic pathways (2012; Kostic et al., 2015; Turnbaugh et al., 2009).

Studies in IBD and experimental models of colitis have reported alterations in microbial functional profiles, which in concordance with a misbalanced microbial composition characterise dysbiosis (Davenport et al., 2014; Ilott et al., 2016; Morgan et al., 2012; Rooks et al., 2014; Schwab et al., 2014). Microbial gene functions related to oxidative stress resistance and nutrient transport are increased in colitis, whereas gene pathways implicated in basic biosynthetic processes such as amino acid biosynthesis are decreased, indicating a shift in energy metabolism during IBD establishment (Ilott et al., 2016; Morgan et al., 2012; Rooks et al., 2014). Furthermore, there is enrichment in pathways contributing to bacterial pathogenesis including cell motility and cell signalling (Morgan et al., 2012; Rooks et al., 2014). These functional changes of the gut microbiome denote an adaptation of multiple members of the microbial community to withstand environmental stressors such as nutrient deprivation and increased oxidative stress dominating in the inflamed gut. Based on these findings, we could speculate that changes in microbial gene functional potential occur as a

consequence of inflammation, further supporting our observations of microbial functional stability in *mdr1a*$^{-/-}$ mice before the onset of intestinal inflammation.

Functional profiling studies of the gut microbiome have mainly focused on stages of established IBD (Davenport et al., 2014; Ilott et al., 2016; Morgan et al., 2012; Rooks et al., 2014). A study of the colonic microbiota from healthy individuals carrying fucosyltransferase 2 (FUT2) gene polymorphism, a risk factor for CD, revealed alterations at both the compositional and functional level (Tong et al., 2014). Microbial metabolic pathways associated with energy metabolism such as amino acid biosynthesis and carbohydrate metabolism were disrupted. This pattern of functional changes was also recapitulated in humanized *FUT2*$^{-/-}$ mice (Tong et al., 2014). So, do changes in microbial functional profiles precede inflammation in addition to being incurred by it? Structural and functional perturbations of the gut microbiome in patients carrying FUT2 polymorphism were associated with sub-clinical inflammation in the intestinal mucosa (Tong et al., 2014). As the affected microbial functional pathways in subjects having FUT2 gene variants were related to energy metabolism and since these identical pathways have been found altered in IBD, it is highly likely that underlying inflammation may have triggered these changes in FUT2 individuals predisposing them to CD.

A longitudinal study performed on infants at risk of developing type 1 diabetes (T1D) followed them until clinical appearance of disease providing a clearer picture on the time onset of microbial functional differences. Analysis of the human infant gut microbiome showed that T1D was preceded by a 25% decrease in species diversity at a time window after the presence of autoantibodies in blood circulation but before clinical diagnosis (Kostic et al., 2015). During this time window, there was a spike in the relative abundance of a few bacterial groups including *Blautia*, Rikenellaceae*, Ruminococcus* and *Streptococcus*, which contain pathobiont species. Conversely, there was a drop in the abundance of bacterial groups commonly depleted in inflammation such as Lachnospiraceae and Veillonellaceae. The altered microbial composition preceding T1D onset was associated with changes in gut bacterial gene content and faecal metabolite profile, both of which had features of a pro-inflammatory environment (Kostic et al., 2015). More specifically, microbial gene functions shifted from the synthesis of nutrients such as amino acids to transportation of readily available sugars in the inflamed tissue. Therefore, microbiota changes happening before the diagnosis of T1D may promote a metabolic environment in the gut that enhances

inflammation, leading ultimately to a clinical disease outcome. Whether changes in microbial gene functions occur at a certain time window in IBD in a similar manner with T1D remains to be identified. When we examined microbial gene content at onset of inflammation, we found that microbial functional pathways remained stable in *mdr1a*$^{-/-}$ mice. It is possible that alterations in microbiota composition have not had an impact on gene functional profile yet. Furthermore, as a subset of *mdr1a*$^{-/-}$ mice suffered from severe colitis at this time as evidenced by the variability in colitis scores, another possible explanation is that inflammation was not so profound to impact microbial gene function in all *mdr1a*$^{-/-}$ mice at this point. Altogether these data support the resilience of the gut microbiome and that establishment of inflammation is required to drive changes in microbial gene function. A longitudinal analysis of microbial communities in *mdr1a*$^{-/-}$ mice would be useful to examine potential differences in microbial gene content along with colitis progression.

We inferred the approximate gene content of gut microbial communities using an *in silico* analysis based on microbial species composition (Langille et al., 2013). Nevertheless, this type of analysis does not provide information about the expression of encoded microbial gene functions. To gain insight into how expressed microbial gene functions impact on host-microbiota relationship before and during inflammation onset, we performed a metabolomics analysis on urinary samples. Metabolomics is particularly useful for detecting subtle changes in metabolic pathways, before phenotypic changes appear, because it integrates information about gene regulation, post-translational modification and pathway interaction (Yoshida et al., 2012). We examined urinary metabolites in particular, as they reflect endogenous metabolites produced by host metabolism and host-bacterial co-metabolism, as well as products of bacterial metabolism (De Preter and Verbeke, 2013; Lees et al., 2013).

Analysis of urinary metabolite profiles revealed resilience of the metabolic network to colitis-induced inflammation. Metabolic markers of IBD previously reported to change in human and experimental animal studies showed no difference between *mdr1a*$^{-/-}$ and control mice at stages preceding and during inflammation. Importantly, the levels of hippurate, a metabolite consistently reduced in IBD studies (Schicho et al., 2012; Williams et al., 2009), remained unchanged in *mdr1a*$^{-/-}$ mice compared with wild-type. Hippurate levels have been correlated with the presence of clostridia in the gut (Li et al., 2008) and since clostridia species are reduced in IBD (Sokol et al., 2008), low hippurate levels may reflect microbial dysbiosis (Williams et al., 2010). Although we observed a reduction in clostridia in mucus in *mdr1a*$^{-/-}$

mice prior to inflammation onset, this change was not accompanied by a concomitant decrease in hippurate. It is possible that microbiota changes in a local niche are unable to affect metabolism in such an extent that can be captured by urine metabolomics. Therefore, metabolite analysis in other bodily fluids such as faeces or the mucus could provide an alternative in order to discern potential subtle metabolite differences.

In further support of our findings of metabolic stability in *mdr1a*$^{-/-}$ mice at the face of ensuing gut inflammation, it was also showed that metabolite profiles were not correlated with colitis development. To date, studies on metabolite profiles in IBD have focused on patients with established disease and have demonstrated alterations compared to healthy individuals (Le Gall et al., 2011; Marchesi et al., 2007; Schicho et al., 2012; Williams et al., 2009; Williams et al., 2012). Levels of SCFAs, secondary bile acids, tricarboxylic acid cycle intermediates, which are implicated in energy metabolism, are reduced in IBD, whereas amounts of amino acids are elevated suggesting malabsorption due to compromised epithelial barrier (Duboc et al., 2013; Le Gall et al., 2011; Marchesi et al., 2007; Schicho et al., 2012). These changes in metabolites accommodate to altered energy metabolism and enhanced energy requirements during intestinal inflammation (Dawiskiba et al., 2014; Schicho et al., 2010; Schicho et al., 2012). Moreover, studies in IL-10$^{-/-}$ mice and DSS-treated mice have shown that metabolite differences were more pronounced, as these mice progressed to inflammation and exhibited symptoms of colitis (Lin et al., 2010; Lin et al., 2009; Martin et al., 2009; Murdoch et al., 2008; Schicho et al., 2010). These findings further support that differences in metabolite profiles during IBD are secondary to inflammation following disruption of the host-microbiota homeostatic relationship.

Our results indicated that metabolite profiles discriminated mice according to their genotype. Metabolic composition differed between wild-type controls and colitis-prone m*dr1a*$^{-/-}$ mice. Given that observed metabolic alterations were unaffected by inflammation, but rather they were associated with a genotype that causes spontaneous colitis, we could hypothesize that such genotype-differentiating metabolites have a value in predicting risk of IBD in at least a subset of patients carrying MDR1A polymorphisms. Identification of the nature of such discriminating metabolites is thus really important and should be addressed in future studies. Importantly, *mdr1a*$^{-/-}$ mice fed a control diet had lower levels of SCFAs and higher levels of succinic acid in their faeces compared with *mdr1a*$^{-/-}$ mice on a control diet supplemented with blueberries or broccoli (Paturi et al., 2012). It would be tempting to speculate that SCFAs,

which are known to be involved in the maintenance of immune tolerance (Furusawa et al., 2013; Smith et al., 2013), are indeed the genotype-discriminating metabolites, whose alteration ultimately leads to chronic inflammation. Of note, a reduction in SCFAs levels has been linked with an altered microbiota composition and especially a decrease in butyrate-producing bacteria such as clostridia species (Sartor, 2008).

5.7 Epithelial response: do IECs regulate the altered microbial composition?

Previous studies in *mdr1a*$^{-/-}$ mice have shown decreased expression of genes involved in the maintenance of mucosal integrity at stages before the onset of colonic inflammation (Collett et al., 2008; Dommels et al., 2007). Transcriptome analysis using whole colon identified significant changes in genes as early as 4 weeks of age (Collett et al., 2008). Genes implicated in bacterial recognition and antigen presentation were upregulated, whereas genes with a role in homeostasis were downregulated. Some of these genes that displayed a markedly reduced expression were Reg3γ and the related pancreatitis associated protein (PAP) (Collett et al., 2008; Dommels et al., 2007). Similar transcriptomic changes were also observed in colonic epithelial cells (CECs), indicating their epithelial origin (Collett et al., 2008; Dommels et al., 2007). Given that Reg3γ expression controls the composition of mucosa-associated gut microbiota (Cash et al., 2006; Vaishnava et al., 2011) and since we have reported changes in mucus microbiota prior to colitis, we hypothesized that changes in AMP expression could account for observed differences in microbial composition. Strikingly, we found similar levels of expression of the antimicrobials Reg3γ, Ang4, β-defensin 1 and Relm-β between *mdr1a*$^{-/-}$ and wild-type mice. Interestingly, a study on activation-induced cytidine deaminase knockout mice (*AID*$^{-/-}$) demonstrated that the expression of AMPs was not affected by the aberrant expansion of segmented filamentous bacteria (SFB), which established a state of dysbiosis (Suzuki et al., 2004).

A possible explanation for the discrepancy between previous studies on *mdr1a*$^{-/-}$ mice and our own is that we used co-housed littermate controls, whereas the other studies used discrete colonies of mice (non-littermates), which were individually housed (Collett et al., 2008; Dommels et al., 2007). It is crucial to control for the environment in microbiota studies, as there is substantial microbial variation between cages, location and litters (Friswell et al.,

2010); thus results obtained may be confounded by the environment were it not properly controlled. Furthermore, differences in diet could have contributed to the disparity between previous studies on *mdr1a*$^{-/-}$ mice and our own. We used acidified filtered water and irradiated food for the maintenance of *mdr1a*$^{-/-}$ and wild-type mice, whereas filtered water and standard chow (Collett et al., 2008) or AIN-76A powdered diet prepared in-house and water (Dommels et al., 2007), in which it is not specified if it was filtered or acidified, were used in the other studies. It is also not stated in the previous work whether FVB control mice were given the same diet as *mdr1a*$^{-/-}$ mice, so it cannot be assumed they had the same diets. Moreover, we looked on AMPs gene expression in proximal colon whereas other studies have used whole colon and isolated CECs. Therefore, it could be possible that regional differences in AMPs may have influenced our observed results. For example, a study on *Trichuris muris* infection has shown differences in the expression profiles between proximal CECs, where the worm is located, and distal CECs, which are infection clear (Cruickshank et al., 2009).

Our data showed increased β-defensin 1 expression and immunoglobulin A (IgA)-producing B-cells in *mdr1a*$^{-/-}$ mice exhibiting severe inflammation. Studies in patients suffering from IBD and in experimental models of colitis have shown upregulation of AMPs such as β-defensin 2, Reg3γ/ PAP and Relm-β in inflamed colonic epithelial tissue (Dieckgraefe et al., 2002; Hogan et al., 2006; Mizoguchi et al., 2003; Nair et al., 2008; O'Neil et al., 1999; Ogawa et al., 2003). Increased AMPs expression may act as a mechanism to counteract increased bacterial colonization and intestinal epithelium penetration. Recently, it has been shown that gut microbiota have encoded mechanisms enabling them to resist elevated levels of inflammation-associated AMPs and as a result to persist in the adverse conditions of an inflammatory environment (Cullen et al., 2015).

5.8 Conclusions

Using the *mdr1a*$^{-/-}$ spontaneous model of colitis, our data revealed that alterations in microbiota composition started in the mucus before the onset of intestinal inflammation (Figure 5.1). Microbiota differences later extended to the faecal microbial community in parallel with colitis development. Except for microbial composition, localisation of microbiota was also affected showing a compromised mucus barrier prior to disease and loss

of spatial segregation on disease onset, thus leading to further induction and perpetuation of chronic intestinal pathology. Despite changes in taxonomic composition, the microbial functional potential was stable at the face of ensuing gut inflammation. Supporting this, urinary metabolite analysis showed that changes in metabolite profiles were unrelated to inflammation, indicating resilience of the metabolic network. Notably, metabolite profiles could distinguish the colitis-prone genotype suggesting their potential usage as a predictive risk for IBD. Observed changes in mucus microbiota composition could not be attributed to altered expression of antimicrobials Reg3γ, Ang4, β-defensin 1 and Relm-β or increased numbers of IgA-producing plasma cells. Therefore, other factors that remain to be identified such as mucus glycans or epithelial metabolic products could contribute to alterations in microbiota composition.

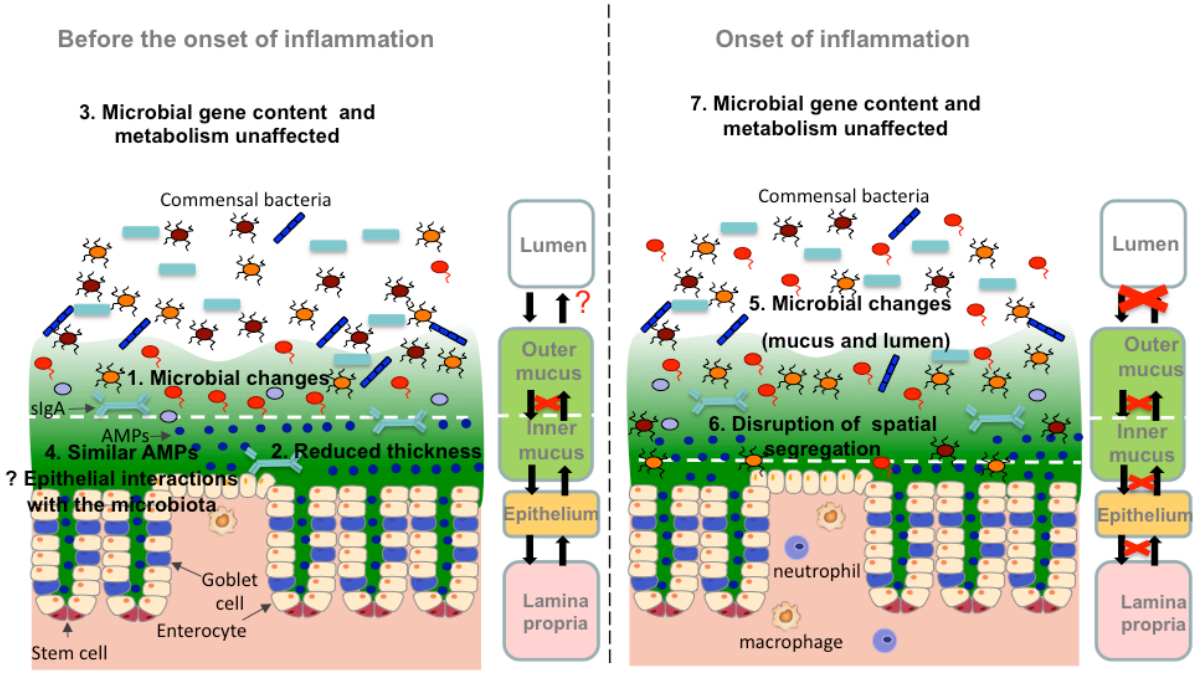


Figure 5.1. Schematic showing possible mechanisms involved at stages preceding colitis and during its onset. Diagram depicts a cartoon illustrating the details of the cells and structures in the colon with a simple schematic in which the lamina propria is in pink, the epithelium in orange, the mucus in green and the lumen in white and arrows outline the potential communication channels between the host and microbiota **(a)** before onset of inflammation and **(b)** during inflammation. **(a)** Before the onset of intestinal inflammation: 1) Alterations in microbiota composition start in the mucus in the *mdr1a*$^{-/-}$ spontaneous model of colitis and 2) this is coupled with reduced inner mucus thickness (dotted line). Both of these changes may disrupt the interplay of mucus microbial communities with the intestinal epithelium and possibly with other microbial constituents in the lumen as indicated in the schematic. 3) However, microbial gene content and urinary metabolites are not affected. 4) Epithelial interactions may shape the mucus microbiota and mucus thickness. Our data excluded antimicrobial proteins (AMPs) and IgA, but other possibilities are altered mucus glycosylation and epithelial metabolic products. **(b)** Onset of inflammation: 5) Changes in microbiota composition are found in mucus and lumen, while 6) spatial segregation of microbiota is disrupted. 7) The microbial functional potential and urinary metabolite profiles remain stable.

5.9 Future work

This thesis has addressed whether changes in gut microbial composition precede intestinal inflammation and how microbial location, epithelial response and host-bacterial metabolism change over the course of colitis development. However, there are still many issues that need to be resolved and questions to be answered.

The data have shown that changes in mucus resident microbiota precede the onset of intestinal inflammation and this was associated with reduced Clostridiales and increased S24-7. Alterations in microbiota composition were found in both mucus and faeces upon emergence of intestinal inflammation. The abundance of Gram-negative bacteria and mucus-degrading bacteria were increased on disease onset in colitis-prone mice. To further validate these findings and reinforce our observations, there are a number of steps that need to be undertaken. First, qPCR analysis of specific bacterial groups including Clostridia, S24-7 and *Akkermansia* needs to be performed, because although we did DGGE on a larger cohort of samples to confirm broad differences, we only sampled a relatively small group in our next-generation sequencing experiments but with qPCR a larger cohort of samples can be used. In addition, 16S rRNA sequencing gives information about the relative abundance of these groups, but not their absolute abundance. In contrast, qPCR can provide semi-quantitative data and thus the absolute numbers of these bacterial groups can be assessed. Furthermore, the Gram-positive/ Gram-negative qPCR could not target all Gram-positive and Gram-negative bacteria respectively. The Gram-positive probe used was designed using bacteria affiliated with the Bacilli (Wu et al., 2008) and thus clostridia including taxa Lachnospiraceae and Ruminococcaceae were less well targeted. Also, the Gram-negative probe cannot detect all the members of the Bacteroidetes phylum such as Bacteroidaceae and Prevotellaceae. Therefore, the Gram-positive/ Gram-negative qPCR has omitted important taxonomic groups. An additional qPCR analysis for the major taxonomic groups that are not targeted by the generic Gram-positive/Gram-negative qPCR could provide further insight on the balance of certain bacterial taxa before and during colitis onset. Living in a highly populated environment the gut microbiota has developed an extensive dynamic network of interactions (Faust and Raes, 2012; Faust et al., 2012). Identification of microbial association networks and how they are disrupted before and during disease is an interesting issue that could provide valuable insight in how dysbiosis drives IBD. Characterization of microbial networks

and identification of potential phylogenetic relationships that govern them is being investigated as part of another research project ongoing in the laboratory. It would also be interesting to correlate the presence of specific bacterial members with the inflammation status of *mdr1a*$^{-/-}$ mice. We only assessed tissue histologically and monitored animals for gross pathological changes, but it is possible there are inflammatory changes in the tissue at the molecular or protein levels that precede the onset of evident inflammation. Thus, it is important to quantify the expression of key inflammatory cytokines such as tumor necrosis factor-alpha (TNF-α) and interferon-gamma (IFN-γ). qPCR analysis of inflammatory markers would confirm our findings based on histological assessment of inflammation.

Fluorescent in situ hybridization (FISH) was performed to visualize gut microbiota localisation in wild-type and *mdr1a*$^{-/-}$ mice. The findings were reduced mucus thickness before the onset of intestinal inflammation and loss of spatial segregation when colitis appeared. We assessed inner colonic mucus thickness and microbial permeability in the inner mucus using gut tissue fixed in Carnoy's solution. As mucus has high water content, processing of gut tissue such as fixation and sectioning introduces changes disrupting its native state. In particular, Carnoy's fixative induces shrinkage of mucus thickness and generation of spaces lacking mucus in proximity to the epithelium (Earle et al., 2015; Johansson et al., 2008). Studies performed *ex-vivo* using for example the charcoal method or fluorescent beads on colon explants would add more validity to our *in-vitro* observations based on FISH analysis of Carnoy's fixed gut tissue. Another experimental tool in visualization of bacteria that has recently been developed and tested in gut anaerobic bacteria involves the addition of fluorescent tags in bacterial cells that are later introduced into mice, allowing *in-vivo* imaging of microbe-microbe interactions (Geva-Zatorsky et al., 2015). Tracking of fluorescent bacteria in the gut of *mdr1a*$^{-/-}$ mice would enable us to study microbe-microbe and microbe-host interactions in real-time over the course of colitis using non-invasive imaging techniques.

It would be interesting to characterise mucus properties given the compromised mucus barrier and the altered mucus microbiota composition we observed before the onset of histological inflammation in *mdr1a*$^{-/-}$ mice. Although we did not report any differences in the number of goblet cells between *mdr1a*$^{-/-}$ and wild-type mice before colitis onset, qPCR analysis of MUC2 mRNA levels is required in order to confirm whether alterations in mucus secretion are accountable for the reduced mucus thickness or other mechanisms involving for example

mucus glycosylation are implicated. Analysis of mucus glycans could be performed by mass spectrometry (MS) and relative abundance of glycosyltrasferases could be assessed by proteomic analysis (Johansson et al., 2015). Since expression of AMPs was found to be similar between wild-type and *mdr1a*$^{-/-}$ mice, it is highly likely that other factors contribute to altered mucus microbial composition. Therefore, in addition to mucus glycans, it would be interesting to look at the immune and epithelial cells. Microarray analysis or RNA sequencing in IECs could provide a potential mechanistic link between the lack of P-glycoprotein transporter in *mdr1a*$^{-/-}$ mice and its effect on host-microbiota crosstalk, which leads to chronic inflammation.

The data have indicated a remarkable stability of the microbial functional gene potential and metabolic network during colitis development. We inferred the predicted microbial gene content based on microbial species composition using a developed algorithm (Langille et al., 2013). To confirm our observations, future studies could perform a metagenomics or metatranscriptomics analysis to quantify bacterial gene expression. Metabolite profiles were also analyzed in urine samples using liquid chromatography (LC)-MS to understand how host-microbiota relationship affects metabolism. Our data showed that changes in metabolites were unaffected by inflammation. As metabolites differ depending on the biological fluid and the chosen analytical method, it would be interesting to perform metabolite analysis on faecal samples or even in the intestinal mucosa and compare these results with urinary metabolite data. Due to LC-MS, we were unable to assess levels of SCFAs and future studies using a different technique such as gas chromatography coupled with MS could address this issue. Imaging MS, which allows the identification of proteins and small molecules such as bacterial or host metabolites locally in tissue sections (Rath et al., 2012), could also be applied to provide insight in the metabolic networks that characterize gut microbial communities.

In research, it is important observations in one animal model to be transferrable to other animal models and ultimately in humans. To confirm the potential pathogenic role of certain bacterial groups, experiments in germ-free mice would be required in order to provide a mechanistic link. Also, in future studies we could focus not only on bacteria, but also on other components of the gut microbiota such as fungi, because fungi have a large impact on immune homeostasis despite their small numbers and also affect the composition of the gut microbiota (Kumamoto, 2016). It would be interesting if observations made in the *mdr1a*$^{-/-}$

model could be translated in UC patients or even a subset of them carrying MDR1A polymorphisms. However, an issue for human studies is that patients undergo routines such as colonic lavage before they are biopsied, thus the samples being taken will have been flushed and cannot fully replicate the *in situ* bacterial communities (Harrell et al., 2012). This issue is addressed in the recently developed advanced method of mucus sampling, OriCol™ (http://originsciences.com/index.php/microbiome), which retains mucosa-residing microbial communities. Identification of bacteria associated with IBD would contribute to a better understanding of the underlying mechanisms associated with disease pathogenesis, allowing the design of better therapeutic strategies. They could also provide diagnostic biomarkers for early disease diagnosis and monitoring tools for disease surveillance.

References

(2012). Structure, function and diversity of the healthy human microbiome. Nature *486*, 207-214.

Ambort, D., Johansson, M.E., Gustafsson, J.K., Nilsson, H.E., Ermund, A., Johansson, B.R., Koeck, P.J., Hebert, H., and Hansson, G.C. (2012). Calcium and pH-dependent packing and release of the gel-forming MUC2 mucin. Proceedings of the National Academy of Sciences of the United States of America *109*, 5645-5650.

Antharam, V.C., Li, E.C., Ishmael, A., Sharma, A., Mai, V., Rand, K.H., and Wang, G.P. (2013). Intestinal dysbiosis and depletion of butyrogenic bacteria in Clostridium difficile infection and nosocomial diarrhea. Journal of clinical microbiology *51*, 2884-2892.

Asher, M.I., Montefort, S., Bjorksten, B., Lai, C.K., Strachan, D.P., Weiland, S.K., Williams, H., and Group, I.P.T.S. (2006). Worldwide time trends in the prevalence of symptoms of asthma, allergic rhinoconjunctivitis, and eczema in childhood: ISAAC Phases One and Three repeat multicountry cross-sectional surveys. Lancet *368*, 733-743.

Atarashi, K., Tanoue, T., Oshima, K., Suda, W., Nagano, Y., Nishikawa, H., Fukuda, S., Saito, T., Narushima, S., Hase, K.*, et al.* (2013). Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature *500*, 232-236.

Atarashi, K., Tanoue, T., Shima, T., Imaoka, A., Kuwahara, T., Momose, Y., Cheng, G., Yamasaki, S., Saito, T., Ohba, Y.*, et al.* (2011). Induction of colonic regulatory T cells by indigenous Clostridium species. Science *331*, 337-341.

Baumler, A.J., and Sperandio, V. (2016). Interactions between the microbiota and pathogenic bacteria in the gut. Nature *535*, 85-93.

Belkaid, Y., and Segre, J.A. (2014). Dialogue between skin microbiota and immunity. Science *346*, 954-959.

Brestoff, J.R., and Artis, D. (2013). Commensal bacteria at the interface of host metabolism and the immune system. Nature immunology *14*, 676-684.

Buffie, C.G., Bucci, V., Stein, R.R., McKenney, P.T., Ling, L., Gobourne, A., No, D., Liu, H., Kinnebrew, M., Viale, A.*, et al.* (2015). Precision microbiome reconstitution restores bile acid mediated resistance to Clostridium difficile. Nature *517*, 205-208.

Burkitt, D.P. (1973). Epidemiology of large bowel disease: the role of fibre. The Proceedings of the Nutrition Society *32*, 145-149.

Cash, H.L., Whitham, C.V., Behrendt, C.L., and Hooper, L.V. (2006). Symbiotic bacteria direct expression of an intestinal bactericidal lectin. Science *313*, 1126-1130.

Collett, A., Higgs, N.B., Gironella, M., Zeef, L.A., Hayes, A., Salmo, E., Haboubi, N., Iovanna, J.L., Carlson, G.L., and Warhurst, G. (2008). Early molecular and functional changes in colonic epithelium that precede increased gut permeability during colitis development in mdr1a(-/-) mice. Inflammatory bowel diseases *14*, 620-631.

Couturier-Maillard, A., Secher, T., Rehman, A., Normand, S., De Arcangelis, A., Haesler, R., Huot, L., Grandjean, T., Bressenot, A., Delanoye-Crespin, A., *et al.* (2013). NOD2-mediated dysbiosis predisposes mice to transmissible colitis and colorectal cancer. The Journal of clinical investigation *123*, 700-711.

Cruickshank, S.M., Deschoolmeester, M.L., Svensson, M., Howell, G., Bazakou, A., Logunova, L., Little, M.C., English, N., Mack, M., Grencis, R.K., *et al.* (2009). Rapid dendritic cell mobilization to the large intestinal epithelium is associated with resistance to Trichuris muris infection. J Immunol *182*, 3055-3062.

Cullen, T.W., Schofield, W.B., Barry, N.A., Putnam, E.E., Rundell, E.A., Trent, M.S., Degnan, P.H., Booth, C.J., Yu, H., and Goodman, A.L. (2015). Gut microbiota. Antimicrobial peptide resistance mediates resilience of prominent gut commensals during inflammation. Science *347*, 170-175.

Davenport, M., Poles, J., Leung, J.M., Wolff, M.J., Abidi, W.M., Ullman, T., Mayer, L., Cho, I., and Loke, P. (2014). Metabolic alterations to the mucosal microbiota in inflammatory bowel disease. Inflammatory bowel diseases *20*, 723-731.

Dawiskiba, T., Deja, S., Mulak, A., Zabek, A., Jawien, E., Pawelka, D., Banasik, M., Mastalerz-Migas, A., Balcerzak, W., Kaliszewski, K., *et al.* (2014). Serum and urine metabolomic fingerprinting in diagnostics of inflammatory bowel diseases. World journal of gastroenterology : WJG *20*, 163-174.

De Preter, V., and Verbeke, K. (2013). Metabolomics as a diagnostic tool in gastroenterology. World journal of gastrointestinal pharmacology and therapeutics *4*, 97-107.

Dieckgraefe, B.K., Crimmins, D.L., Landt, V., Houchen, C., Anant, S., Porche-Sorbet, R., and Ladenson, J.H. (2002). Expression of the regenerating gene family in inflammatory bowel disease mucosa: Reg Ialpha upregulation, processing, and antiapoptotic activity. Journal of investigative medicine : the official publication of the American Federation for Clinical Research *50*, 421-434.

Dommels, Y.E., Butts, C.A., Zhu, S., Davy, M., Martell, S., Hedderley, D., Barnett, M.P., McNabb, W.C., and Roy, N.C. (2007). Characterization of intestinal inflammation and identification of related gene expression changes in mdr1a(-/-) mice. Genes & nutrition *2*, 209-223.

Donia, M.S., and Fischbach, M.A. (2015). Human microbiota. Small molecules from the human microbiota. Science *349*, 1254766.

Dorrestein, P.C., Mazmanian, S.K., and Knight, R. (2014). Finding the missing links among metabolites, microbes, and the host. Immunity *40*, 824-832.

Duboc, H., Rajca, S., Rainteau, D., Benarous, D., Maubert, M.A., Quervain, E., Thomas, G., Barbu, V., Humbert, L., Despras, G., *et al.* (2013). Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. Gut *62*, 531-539.

Earle, K.A., Billings, G., Sigal, M., Lichtman, J.S., Hansson, G.C., Elias, J.E., Amieva, M.R., Huang, K.C., and Sonnenburg, J.L. (2015). Quantitative Imaging of Gut Microbiota Spatial Organization. Cell host & microbe *18*, 478-488.

Elinav, E., Strowig, T., Kau, A.L., Henao-Mejia, J., Thaiss, C.A., Booth, C.J., Peaper, D.R., Bertin, J., Eisenbarth, S.C., Gordon, J.I., and Flavell, R.A. (2011). NLRP6 inflammasome regulates colonic microbial ecology and risk for colitis. Cell *145*, 745-757.

Faust, K., and Raes, J. (2012). Microbial interactions: from networks to models. Nature reviews. Microbiology *10*, 538-550.

Faust, K., Sathirapongsasuti, J.F., Izard, J., Segata, N., Gevers, D., Raes, J., and Huttenhower, C. (2012). Microbial co-occurrence relationships in the human microbiome. PLoS computational biology *8*, e1002606.

Fite, A., Macfarlane, S., Furrie, E., Bahrami, B., Cummings, J.H., Steinke, D.T., and Macfarlane, G.T. (2013). Longitudinal analyses of gut mucosal microbiotas in ulcerative colitis in relation to patient age and disease severity and duration. Journal of clinical microbiology *51*, 849-856.

Friswell, M.K., Gika, H., Stratford, I.J., Theodoridis, G., Telfer, B., Wilson, I.D., and McBain, A.J. (2010). Site and strain-specific variation in gut microbiota profiles and metabolism in experimental mice. PloS one *5*, e8584.

Fu, J., Wei, B., Wen, T., Johansson, M.E., Liu, X., Bradford, E., Thomsson, K.A., McGee, S., Mansour, L., Tong, M.*, et al.* (2011). Loss of intestinal core 1-derived O-glycans causes spontaneous colitis in mice. The Journal of clinical investigation *121*, 1657-1666.

Furusawa, Y., Obata, Y., Fukuda, S., Endo, T.A., Nakato, G., Takahashi, D., Nakanishi, Y., Uetake, C., Kato, K., Kato, T.*, et al.* (2013). Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. Nature *504*, 446-450.

Fyderek, K., Strus, M., Kowalska-Duplaga, K., Gosiewski, T., Wedrychowicz, A., Jedynak-Wasowicz, U., Sladek, M., Pieczarkowski, S., Adamski, P., Kochan, P., and Heczko, P.B. (2009). Mucosal bacterial microflora and mucus layer thickness in adolescents with inflammatory bowel disease. World journal of gastroenterology : WJG *15*, 5287-5294.

Gao, Z., Tseng, C.H., Strober, B.E., Pei, Z., and Blaser, M.J. (2008). Substantial alterations of the cutaneous bacterial biota in psoriatic lesions. PloS one *3*, e2719.

Garrett, W.S., Gallini, C.A., Yatsunenko, T., Michaud, M., DuBois, A., Delaney, M.L., Punit, S., Karlsson, M., Bry, L., Glickman, J.N.*, et al.* (2010). Enterobacteriaceae act in concert with the gut microbiota to induce spontaneous and maternally transmitted colitis. Cell host & microbe *8*, 292-300.

Geva-Zatorsky, N., Alvarez, D., Hudak, J.E., Reading, N.C., Erturk-Hasdemir, D., Dasgupta, S., von Andrian, U.H., and Kasper, D.L. (2015). In vivo imaging and tracking of host-microbiota interactions via metabolic labeling of gut anaerobic bacteria. Nature medicine *21*, 1091-1100.

Gevers, D., Kugathasan, S., Denson, L.A., Vazquez-Baeza, Y., Van Treuren, W., Ren, B., Schwager, E., Knights, D., Song, S.J., Yassour, M.*, et al.* (2014). The treatment-naive microbiome in new-onset Crohn's disease. Cell host & microbe *15*, 382-392.

Harrell, L., Wang, Y., Antonopoulos, D., Young, V., Lichtenstein, L., Huang, Y., Hanauer, S., and Chang, E. (2012). Standard colonic lavage alters the natural state of mucosal-associated microbiota in the human colon. PloS one *7*, e32545.

Hedin, C., van der Gast, C.J., Rogers, G.B., Cuthbertson, L., McCartney, S., Stagg, A.J., Lindsay, J.O., and Whelan, K. (2016). Siblings of patients with Crohn's disease exhibit a biologically relevant dysbiosis in mucosal microbial metacommunities. Gut *65*, 944-953.

Hogan, S.P., Seidu, L., Blanchard, C., Groschwitz, K., Mishra, A., Karow, M.L., Ahrens, R., Artis, D., Murphy, A.J., Valenzuela, D.M.*, et al.* (2006). Resistin-like molecule beta regulates innate colonic function: barrier integrity and inflammation susceptibility. The Journal of allergy and clinical immunology *118*, 257-268.

Huang, Y.L., Chassard, C., Hausmann, M., von Itzstein, M., and Hennet, T. (2015). Sialic acid catabolism drives intestinal inflammation and microbial dysbiosis in mice. Nature communications *6*, 8141.

Huttenhower, C., Kostic, A.D., and Xavier, R.J. (2014). Inflammatory bowel disease as a model for translating the microbiome. Immunity *40*, 843-854.

Hviid, A., Svanstrom, H., and Frisch, M. (2011). Antibiotic use and inflammatory bowel diseases in childhood. Gut *60*, 49-54.

Ilott, N.E., Bollrath, J., Danne, C., Schiering, C., Shale, M., Adelmann, K., Krausgruber, T., Heger, A., Sims, D., and Powrie, F. (2016). Defining the microbial transcriptional response to colitis through integrated host and microbiome profiling. The ISME journal.

Irvine, A.D., McLean, W.H.I., and Leung, D.Y.M. (2011). Filaggrin Mutations Associated with Skin and Allergic Diseases. New England Journal of Medicine *365*, 1315-1327.

Jakobsson, H.E., Rodriguez-Pineiro, A.M., Schutte, A., Ermund, A., Boysen, P., Bemark, M., Sommer, F., Backhed, F., Hansson, G.C., and Johansson, M.E. (2014). The composition of the gut microbiota shapes the colon mucus barrier. EMBO reports.

Johansson, M.E., Ambort, D., Pelaseyed, T., Schutte, A., Gustafsson, J.K., Ermund, A., Subramani, D.B., Holmen-Larsson, J.M., Thomsson, K.A., Bergstrom, J.H.*, et al.* (2011a). Composition and functional role of the mucus layers in the intestine. Cellular and molecular life sciences : CMLS *68*, 3635-3641.

Johansson, M.E., Gustafsson, J.K., Holmen-Larsson, J., Jabbar, K.S., Xia, L., Xu, H., Ghishan, F.K., Carvalho, F.A., Gewirtz, A.T., Sjovall, H., and Hansson, G.C. (2014). Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and patients with ulcerative colitis. Gut *63*, 281-291.

Johansson, M.E., Gustafsson, J.K., Sjoberg, K.E., Petersson, J., Holm, L., Sjovall, H., and Hansson, G.C. (2010). Bacteria penetrate the inner mucus layer before inflammation in the dextran sulfate colitis model. PloS one *5*, e12238.

Johansson, M.E., and Hansson, G.C. (2011). Microbiology. Keeping bacteria at a distance. Science *334*, 182-183.

Johansson, M.E., Jakobsson, H.E., Holmen-Larsson, J., Schutte, A., Ermund, A., Rodriguez-Pineiro, A.M., Arike, L., Wising, C., Svensson, F., Backhed, F., and Hansson, G.C. (2015). Normalization of Host Intestinal Mucus Layers Requires Long-Term Microbial Colonization. Cell host & microbe *18*, 582-592.

Johansson, M.E., Larsson, J.M., and Hansson, G.C. (2011b). The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. Proceedings of the National Academy of Sciences of the United States of America *108 Suppl 1*, 4659-4665.

Johansson, M.E., Phillipson, M., Petersson, J., Velcich, A., Holm, L., and Hansson, G.C. (2008). The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. Proceedings of the National Academy of Sciences of the United States of America *105*, 15064-15069.

Johansson, M.E., Thomsson, K.A., and Hansson, G.C. (2009). Proteomic analyses of the two mucus layers of the colon barrier reveal that their main component, the Muc2 mucin, is strongly bound to the Fcgbp protein. Journal of proteome research *8*, 3549-3557.

Joossens, M., Huys, G., Cnockaert, M., De Preter, V., Verbeke, K., Rutgeerts, P., Vandamme, P., and Vermeire, S. (2011). Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut *60*, 631-637.

Jostins, L., Ripke, S., Weersma, R.K., Duerr, R.H., McGovern, D.P., Hui, K.Y., Lee, J.C., Schumm, L.P., Sharma, Y., Anderson, C.A.*, et al.* (2012). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature *491*, 119-124.

Kang, S., Denman, S.E., Morrison, M., Yu, Z., Dore, J., Leclerc, M., and McSweeney, C.S. (2010). Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray. Inflammatory bowel diseases *16*, 2034-2042.

Khor, B., Gardet, A., and Xavier, R.J. (2011). Genetics and pathogenesis of inflammatory bowel disease. Nature *474*, 307-317.

Kobayashi, T., Glatz, M., Horiuchi, K., Kawasaki, H., Akiyama, H., Kaplan, D.H., Kong, H.H., Amagai, M., and Nagao, K. (2015). Dysbiosis and Staphylococcus aureus Colonization Drives Inflammation in Atopic Dermatitis. Immunity *42*, 756-766.

Koenigsknecht, M.J., Theriot, C.M., Bergin, I.L., Schumacher, C.A., Schloss, P.D., and Young, V.B. (2015). Dynamics and establishment of Clostridium difficile infection in the murine gastrointestinal tract. Infection and immunity *83*, 934-941.

Kong, H.H., Oh, J., Deming, C., Conlan, S., Grice, E.A., Beatson, M.A., Nomicos, E., Polley, E.C., Komarow, H.D., Program, N.C.S.*, et al.* (2012). Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis. Genome research *22*, 850-859.

Kostic, A.D., Gevers, D., Siljander, H., Vatanen, T., Hyotylainen, T., Hamalainen, A.M., Peet, A., Tillmann, V., Poho, P., Mattila, I.*, et al.* (2015). The dynamics of the human infant gut microbiome in development and in progression toward type 1 diabetes. Cell host & microbe *17*, 260-273.

Kostic, A.D., Xavier, R.J., and Gevers, D. (2014). The microbiome in inflammatory bowel disease: current status and the future ahead. Gastroenterology *146*, 1489-1499.

Kubo, A., Nagao, K., and Amagai, M. (2012). Epidermal barrier dysfunction and cutaneous sensitization in atopic diseases. The Journal of clinical investigation *122*, 440-447.

Kumamoto, C.A. (2016). The Fungal Mycobiota: Small Numbers, Large Impacts. Cell host & microbe *19*, 750-751.

Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R.*, et al.* (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature biotechnology *31*, 814-821.

Larsson, J.M., Karlsson, H., Crespo, J.G., Johansson, M.E., Eklund, L., Sjovall, H., and Hansson, G.C. (2011). Altered O-glycosylation profile of MUC2 mucin occurs in active ulcerative colitis and is associated with increased inflammation. Inflammatory bowel diseases *17*, 2299-2307.

Le Gall, G., Noor, S.O., Ridgway, K., Scovell, L., Jamieson, C., Johnson, I.T., Colquhoun, I.J., Kemsley, E.K., and Narbad, A. (2011). Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. Journal of proteome research *10*, 4208-4218.

Lees, H.J., Swann, J.R., Wilson, I.D., Nicholson, J.K., and Holmes, E. (2013). Hippurate: the natural history of a mammalian-microbial cometabolite. Journal of proteome research *12*, 1527-1546.

Lepage, P., Hasler, R., Spehlmann, M.E., Rehman, A., Zvirbliene, A., Begun, A., Ott, S., Kupcinskas, L., Dore, J., Raedler, A., and Schreiber, S. (2011). Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. Gastroenterology *141*, 227-236.

Li, M., Wang, B., Zhang, M., Rantalainen, M., Wang, S., Zhou, H., Zhang, Y., Shen, J., Pang, X., Zhang, M.*, et al.* (2008). Symbiotic gut microbes modulate human metabolic phenotypes. Proceedings of the National Academy of Sciences of the United States of America *105*, 2117-2122.

Lin, H.M., Barnett, M.P., Roy, N.C., Joyce, N.I., Zhu, S., Armstrong, K., Helsby, N.A., Ferguson, L.R., and Rowan, D.D. (2010). Metabolomic analysis identifies inflammatory and noninflammatory metabolic effects of genetic modification in a mouse model of Crohn's disease. Journal of proteome research *9*, 1965-1975.

Lin, H.M., Edmunds, S.I., Helsby, N.A., Ferguson, L.R., and Rowan, D.D. (2009). Nontargeted urinary metabolite profiling of a mouse model of Crohn's disease. Journal of proteome research *8*, 2045-2057.

Loonen, L.M., Stolte, E.H., Jaklofsky, M.T., Meijerink, M., Dekker, J., van Baarlen, P., and Wells, J.M. (2013). REG3gamma-deficient mice have altered mucus distribution and increased mucosal inflammatory responses to the microbiota and enteric pathogens in the ileum. Mucosal immunology.

Lupp, C., Robertson, M.L., Wickham, M.E., Sekirov, I., Champion, O.L., Gaynor, E.C., and Finlay, B.B. (2007). Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. Cell host & microbe *2*, 119-129.

Maloy, K.J., and Powrie, F. (2011). Intestinal homeostasis and its breakdown in inflammatory bowel disease. Nature *474*, 298-306.

Manichanh, C., Borruel, N., Casellas, F., and Guarner, F. (2012). The gut microbiota in IBD. Nature reviews. Gastroenterology & hepatology *9*, 599-608.

Manichanh, C., Rigottier-Gois, L., Bonnaud, E., Gloux, K., Pelletier, E., Frangeul, L., Nalin, R., Jarrin, C., Chardon, P., Marteau, P.*, et al.* (2006). Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut *55*, 205-211.

Marchesi, J.R., Holmes, E., Khan, F., Kochhar, S., Scanlan, P., Shanahan, F., Wilson, I.D., and Wang, Y. (2007). Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. Journal of proteome research *6*, 546-551.

Martin, F.P., Rezzi, S., Philippe, D., Tornier, L., Messlik, A., Holzlwimmer, G., Baur, P., Quintanilla-Fend, L., Loh, G., Blaut, M.*, et al.* (2009). Metabolic assessment of gradual development of moderate experimental colitis in IL-10 deficient mice. Journal of proteome research *8*, 2376-2387.

Martinez, C., Antolin, M., Santos, J., Torrejon, A., Casellas, F., Borruel, N., Guarner, F., and Malagelada, J.R. (2008). Unstable composition of the fecal microbiota in ulcerative colitis during clinical remission. The American journal of gastroenterology *103*, 643-648.

Mizoguchi, E., Xavier, R.J., Reinecker, H.C., Uchino, H., Bhan, A.K., Podolsky, D.K., and Mizoguchi, A. (2003). Colonic epithelial functional phenotype varies with type and phase of experimental colitis. Gastroenterology *125*, 148-161.

Modi, S.R., Collins, J.J., and Relman, D.A. (2014). Antibiotics and the gut microbiota. The Journal of clinical investigation *124*, 4212-4218.

Momozawa, Y., Deffontaine, V., Louis, E., and Medrano, J.F. (2011). Characterization of bacteria in biopsies of colon and stools by high throughput sequencing of the V2 region of bacterial 16S rRNA gene in human. PloS one *6*, e16952.

Morgan, X.C., Tickle, T.L., Sokol, H., Gevers, D., Devaney, K.L., Ward, D.V., Reyes, J.A., Shah, S.A., LeLeiko, N., Snapper, S.B.*, et al.* (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome biology *13*, R79.

Murdoch, T.B., Fu, H., MacFarlane, S., Sydora, B.C., Fedorak, R.N., and Slupsky, C.M. (2008). Urinary metabolic profiles of inflammatory bowel disease in interleukin-10 gene-deficient mice. Analytical chemistry *80*, 5524-5531.

Nair, M.G., Guild, K.J., Du, Y., Zaph, C., Yancopoulos, G.D., Valenzuela, D.M., Murphy, A., Stevens, S., Karow, M., and Artis, D. (2008). Goblet cell-derived resistin-like molecule beta augments CD4+ T cell production of IFN-gamma and infection-induced intestinal inflammation. J Immunol *181*, 4709-4715.

Ng, K.M., Ferreyra, J.A., Higginbottom, S.K., Lynch, J.B., Kashyap, P.C., Gopinath, S., Naidu, N., Choudhury, B., Weimer, B.C., Monack, D.M., and Sonnenburg, J.L. (2013). Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. Nature *502*, 96-99.

O'Neil, D.A., Porter, E.M., Elewaut, D., Anderson, G.M., Eckmann, L., Ganz, T., and Kagnoff, M.F. (1999). Expression and regulation of the human beta-defensins hBD-1 and hBD-2 in intestinal epithelium. J Immunol *163*, 6718-6724.

Ogawa, H., Fukushima, K., Naito, H., Funayama, Y., Unno, M., Takahashi, K., Kitayama, T., Matsuno, S., Ohtani, H., Takasawa, S.*, et al.* (2003). Increased expression of HIP/PAP and regenerating gene III in human inflammatory bowel disease and a murine bacterial reconstitution model. Inflammatory bowel diseases *9*, 162-170.

Oh, J., Freeman, A.F., Program, N.C.S., Park, M., Sokolic, R., Candotti, F., Holland, S.M., Segre, J.A., and Kong, H.H. (2013). The altered landscape of the human skin microbiome in patients with primary immunodeficiencies. Genome research *23*, 2103-2114.

Panwala, C.M., Jones, J.C., and Viney, J.L. (1998). A novel model of inflammatory bowel disease: mice deficient for the multiple drug resistance gene, mdr1a, spontaneously develop colitis. J Immunol *161*, 5733-5744.

Papa, E., Docktor, M., Smillie, C., Weber, S., Preheim, S.P., Gevers, D., Giannoukos, G., Ciulla, D., Tabbaa, D., Ingram, J.*, et al.* (2012). Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. PloS one *7*, e39242.

Paturi, G., Mandimika, T., Butts, C.A., Zhu, S., Roy, N.C., McNabb, W.C., and Ansell, J. (2012). Influence of dietary blueberry and broccoli on cecal microbiota activity and colon morphology in mdr1a(-/-) mice, a model of inflammatory bowel diseases. Nutrition *28*, 324-330.

Perez-Bosque, A., Miro, L., Maijo, M., Polo, J., Campbell, J., Russell, L., Crenshaw, J., Weaver, E., and Moreto, M. (2015). Dietary intervention with serum-derived bovine immunoglobulins protects barrier function in a mouse model of colitis. American journal of physiology. Gastrointestinal and liver physiology *308*, G1012-1018.

Petersson, J., Schreiber, O., Hansson, G.C., Gendler, S.J., Velcich, A., Lundberg, J.O., Roos, S., Holm, L., and Phillipson, M. (2011). Importance and regulation of the colonic mucus barrier in a mouse model of colitis. American journal of physiology. Gastrointestinal and liver physiology *300*, G327-333.

Rajilic-Stojanovic, M., Shanahan, F., Guarner, F., and de Vos, W.M. (2013). Phylogenetic analysis of dysbiosis in ulcerative colitis during remission. Inflammatory bowel diseases *19*, 481-488.

Rath, C.M., Alexandrov, T., Higginbottom, S.K., Song, J., Milla, M.E., Fischbach, M.A., Sonnenburg, J.L., and Dorrestein, P.C. (2012). Molecular analysis of model gut microbiotas by imaging mass spectrometry and nanodesorption electrospray ionization reveals dietary metabolite transformations. Analytical chemistry *84*, 9259-9267.

Rooks, M.G., Veiga, P., Wardwell-Scott, L.H., Tickle, T., Segata, N., Michaud, M., Gallini, C.A., Beal, C., van Hylckama-Vlieg, J.E., Ballal, S.A.*, et al.* (2014). Gut microbiome

composition and function in experimental colitis during active disease and treatment-induced remission. The ISME journal *8*, 1403-1417.

Saleh, M., and Elson, C.O. (2011). Experimental inflammatory bowel disease: insights into the host-microbiota dialog. Immunity *34*, 293-302.

Sartor, R.B. (2008). Microbial influences in inflammatory bowel diseases. Gastroenterology *134*, 577-594.

Schicho, R., Nazyrova, A., Shaykhutdinov, R., Duggan, G., Vogel, H.J., and Storr, M. (2010). Quantitative metabolomic profiling of serum and urine in DSS-induced ulcerative colitis of mice by (1)H NMR spectroscopy. Journal of proteome research *9*, 6265-6273.

Schicho, R., Shaykhutdinov, R., Ngo, J., Nazyrova, A., Schneider, C., Panaccione, R., Kaplan, G.G., Vogel, H.J., and Storr, M. (2012). Quantitative metabolomic profiling of serum, plasma, and urine by (1)H NMR spectroscopy discriminates between patients with inflammatory bowel disease and healthy individuals. Journal of proteome research *11*, 3344-3357.

Schinkel, A.H., Mayer, U., Wagenaar, E., Mol, C.A., van Deemter, L., Smit, J.J., van der Valk, M.A., Voordouw, A.C., Spits, H., van Tellingen, O.*, et al.* (1997). Normal viability and altered pharmacokinetics in mice lacking mdr1-type (drug-transporting) P-glycoproteins. Proceedings of the National Academy of Sciences of the United States of America *94*, 4028-4033.

Schopf, L.R., Hoffmann, K.F., Cheever, A.W., Urban, J.F., Jr., and Wynn, T.A. (2002). IL-10 is critical for host resistance and survival during gastrointestinal helminth infection. J Immunol *168*, 2383-2392.

Schubert, A.M., Sinani, H., and Schloss, P.D. (2015). Antibiotic-Induced Alterations of the Murine Gut Microbiota and Subsequent Effects on Colonization Resistance againstClostridium difficile. mBio *6*, e00974-00915.

Schwab, C., Berry, D., Rauch, I., Rennisch, I., Ramesmayer, J., Hainzl, E., Heider, S., Decker, T., Kenner, L., Muller, M.*, et al.* (2014). Longitudinal study of murine microbiota activity and interactions with the host during acute inflammation and recovery. The ISME journal *8*, 1101-1114.

Shaw, S.Y., Blanchard, J.F., and Bernstein, C.N. (2010). Association between the use of antibiotics in the first year of life and pediatric inflammatory bowel disease. The American journal of gastroenterology *105*, 2687-2692.

Smith, P.M., Howitt, M.R., Panikov, N., Michaud, M., Gallini, C.A., Bohlooly, Y.M., Glickman, J.N., and Garrett, W.S. (2013). The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. Science *341*, 569-573.

Sokol, H., Lay, C., Seksik, P., and Tannock, G.W. (2008). Analysis of bacterial bowel communities of IBD patients: what has it revealed? Inflammatory bowel diseases *14*, 858-867.

Sonnenburg, E.D., and Sonnenburg, J.L. (2014). Starving our microbial self: the deleterious consequences of a diet deficient in microbiota-accessible carbohydrates. Cell metabolism *20*, 779-786.

Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.H., Westover, B.P., Weatherford, J., Buhler, J.D., and Gordon, J.I. (2005). Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science *307*, 1955-1959.

Strachan, D.P. (1989). Hay fever, hygiene, and household size. BMJ *299*, 1259-1260.

Strober, W. (2013). Impact of the gut microbiome on mucosal inflammation. Trends in immunology *34*, 423-430.

Strober, W., Fuss, I., and Mannon, P. (2007). The fundamental basis of inflammatory bowel disease. The Journal of clinical investigation *117*, 514-521.

Suzuki, K., Meek, B., Doi, Y., Muramatsu, M., Chiba, T., Honjo, T., and Fagarasan, S. (2004). Aberrant expansion of segmented filamentous bacteria in IgA-deficient gut. Proceedings of the National Academy of Sciences of the United States of America *101*, 1981-1986.

Swidsinski, A., Weber, J., Loening-Baucke, V., Hale, L.P., and Lochs, H. (2005). Spatial organization and composition of the mucosal flora in patients with inflammatory bowel disease. Journal of clinical microbiology *43*, 3380-3389.

Tanner, S.M., Staley, E.M., and Lorenz, R.G. (2013). Altered generation of induced regulatory T cells in the FVB.mdr1a-/- mouse model of colitis. Mucosal immunology *6*, 309-323.

Theriot, C.M., Koenigsknecht, M.J., Carlson, P.E., Jr., Hatton, G.E., Nelson, A.M., Li, B., Huffnagle, G.B., J, Z.L., and Young, V.B. (2014). Antibiotic-induced shifts in the mouse gut microbiome and metabolome increase susceptibility to Clostridium difficile infection. Nature communications *5*, 3114.

Tong, M., McHardy, I., Ruegger, P., Goudarzi, M., Kashyap, P.C., Haritunians, T., Li, X., Graeber, T.G., Schwager, E., Huttenhower, C.*, et al.* (2014). Reprograming of gut microbiome energy metabolism by the FUT2 Crohn's disease risk polymorphism. The ISME journal *8*, 2193-2206.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P.*, et al.* (2009). A core gut microbiome in obese and lean twins. Nature *457*, 480-484.

Ursell, L.K., Haiser, H.J., Van Treuren, W., Garg, N., Reddivari, L., Vanamala, J., Dorrestein, P.C., Turnbaugh, P.J., and Knight, R. (2014). The intestinal metabolome: an intersection between microbiota and host. Gastroenterology *146*, 1470-1476.

Vaishnava, S., Yamamoto, M., Severson, K.M., Ruhn, K.A., Yu, X., Koren, O., Ley, R., Wakeland, E.K., and Hooper, L.V. (2011). The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. Science *334*, 255-258.

Williams, H.R., Cox, I.J., Walker, D.G., Cobbold, J.F., Taylor-Robinson, S.D., Marshall, S.E., and Orchard, T.R. (2010). Differences in gut microbial metabolism are responsible for reduced hippurate synthesis in Crohn's disease. BMC gastroenterology *10*, 108.

Williams, H.R., Cox, I.J., Walker, D.G., North, B.V., Patel, V.M., Marshall, S.E., Jewell, D.P., Ghosh, S., Thomas, H.J., Teare, J.P.*, et al.* (2009). Characterization of inflammatory bowel disease with urinary metabolic profiling. The American journal of gastroenterology *104*, 1435-1444.

Williams, H.R., Willsmore, J.D., Cox, I.J., Walker, D.G., Cobbold, J.F., Taylor-Robinson, S.D., and Orchard, T.R. (2012). Serum metabolic profiling in inflammatory bowel disease. Digestive diseases and sciences *57*, 2157-2165.

Williams, M.R., and Gallo, R.L. (2015). The role of the skin microbiome in atopic dermatitis. Current allergy and asthma reports *15*, 65.

Willing, B.P., Dicksved, J., Halfvarson, J., Andersson, A.F., Lucio, M., Zheng, Z., Jarnerot, G., Tysk, C., Jansson, J.K., and Engstrand, L. (2010). A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. Gastroenterology *139*, 1844-1854 e1841.

Wlodarska, M., Kostic, A.D., and Xavier, R.J. (2015). An integrative view of microbiome-host interactions in inflammatory bowel diseases. Cell host & microbe *17*, 577-591.

Wu, Y.D., Chen, L.H., Wu, X.J., Shang, S.Q., Lou, J.T., Du, L.Z., and Zhao, Z.Y. (2008). Gram stain-specific-probe-based real-time PCR for diagnosis and discrimination of bacterial neonatal sepsis. Journal of clinical microbiology *46*, 2613-2619.

Yoshida, M., Hatano, N., Nishiumi, S., Irino, Y., Izumi, Y., Takenawa, T., and Azuma, T. (2012). Diagnosis of gastroenterological diseases by metabolome analysis using gas chromatography-mass spectrometry. Journal of gastroenterology *47*, 9-20.

Zeng, M.Y., Inohara, N., and Nunez, G. (2016). Mechanisms of inflammation-driven bacterial dysbiosis in the gut. Mucosal immunology.

Chapter 6

Supplementary materials and methods

6.1 Animal husbandry

6.1.1 Maintenance of animals

Mdr1a$^{-/-}$ mice (FVB.129P2-Abcb1atm1Bor N7) (Schinkel et al., 1994) and control FVB mice obtained from Taconic Farms (Albany, NY, USA) were crossbred to generate F1 heterozygotes. F1 mice were further crossbred to generate F2 littermate controls, which were used for experiments. Mice were given autoclaved standard chow and sterile acidified water (pH=3.2) *ad libitum*. Co-housed littermate *mdr1a*$^{-/-}$ and wild-type (WT) males were used throughout to ensure shared microbiota and samples were harvested at designated time points as specified. C57BL/6 (purchased from Harlan, Blackthorn, UK) and IL-10$^{-/-}$ male mice (Berg et al., 1996) raised on a C57BL/6 background (bred in house) were infected with 200 *Trichuris muris* eggs in 200µl of ultra-pure distilled water (Milli-Q® water purification system, Merck Millipore, Darmstadt, Germany) via oral gavage at 8 weeks of age (provided by Dr M. Bramhall). Mice purchased from external supplier were left to acclimatize before any experimental procedure. Stool samples were collected at day 16 of infection when mice were 10 weeks old. All mice were housed in groups of 3-5 animals in individual ventilated cages and maintained under constant 12-hour light dark cycle at 21-23˚C under specific, pathogen-free (SPF) conditions in the Biological Services Facility at the University of Manchester. All experimental procedures in mice were performed in accordance with the regulations issued by the Home Office under amended ASPA, 2012.

6.1.2 Genotyping of *mdr1a*$^{-/-}$ mice

Mouse ear punches were collected and tissue was lysed in the presence of 100mM Tris-HCl pH=8.5, 5mM EDTA pH=8.0, 200mM NaCl, 0.2% SDS and 10 mg/ml Proteinase K (Promega, Southampton, UK) by incubation at a thermomixer at 55˚C for 30 minutes. Tissue homogenate was centrifuged for removal of debris for 10 minutes at maximum speed and supernatants were collected. DNA was precipitated by addition of 300µl isopropanol followed by another centrifugation step at maximum speed (13,000g) for 3 minutes. The supernatant was discarded and the pelleted DNA was washed with 70% ethanol. After centrifugation for 1 minute and removal of the supernatant, the pellet was left to air-dry for

20 minutes at room temperature (RT). DNA was solubilised in 100µl DNase/RNase free H_2O (Promega) and stored at -20˚C.

For the identification of *mdr1a* gene, the primer pairs AS2 (5' – CTC CTC CAA GGT GCA TAG ACC – 3'), AW2 (5' – CCC AGC TCT TCA TCT AAC TAC CCT G -3') and AKO2 (5' – CTT CCC AGC CTC TGA GCC CAG -3') were used as previously described (Collett et al., 2008). 1µl of DNA was added in each PCR reaction, which contained at a final concentration 2mM MgCl2, 200µM of each dNTP, 0.5µM of each primer pair and 2U of FastStart Taq DNA polymerase (Roche Diagnostics, West Sussex, UK). The PCR conditions used were as follows: denaturation/ activation at 95˚C for 8 minutes followed by 30 cycles of denaturation at 94˚C for 45 seconds, annealing at 60˚C for 1 minute and elongation at 72˚C for 1 minute, and a final elongation step at 72˚C for 5 minutes.

6.2 Histology

6.2.1 Tissue processing

Ileum, proximal and distal colon snips were taken at autopsy and placed in fixative solution containing 8% (v/v) formaldehyde, 0.9% (w/v) sodium chloride, 2% (v/v) glacial acetic acid and 0.5% (w/v) cetrimide in dH_2O for 24 hours. Samples were then transferred into 70% (v/v) ethanol and processed using a micro-spin tissue processor STP 120 (Thermo Scientific, supplied by Thermo Fisher Scientific, Paisley, UK) through conditions of 70% (v/v) ethanol for 15 minutes, 90% (v/v) ethanol for 30 minutes, 95% (v/v) ethanol for 30 minutes, two changes of 100% (v/v) ethanol for 60 minutes, 100% (v/v) ethanol for 20 minutes, xylene for 15 minutes, two changes of xylene for 30 minutes (all at 40˚C) and twice in fibrowax pastillated wax (Beckon Dickinson, Oxford, UK) for 1 hour at 60˚C. The processed tissue samples were mounted in wax blocks using a Microm EC 350 embedding system (Thermo Scientific) and 5µm sections were cut using a Microm HM 325 rotary microtome (Thermo Scientific). The sections were placed on standard glass microscope slides (Thermo Scientific) after floating on a water bath and left to dry.

6.2.2 Haematoxylin and eosin (H&E) staining

Paraffin-embedded tissue sections were stained with haematoxylin and eosin for histological assessment of tissue integrity. Prior to staining, tissue sections were dewaxed in Citroclear (HD supplies, Aylesbury, UK) twice for 20 minutes each and taken to distilled water through decreasing concentrations of ethanol (100%, 90%, 70%, 50%) for 2 minutes each change. After rehydration, slides were stained with Harris' haematoxylin (Sigma-Aldrich, Dorset, UK) for 5 minutes, washed in running tap water for 1 minute and differentiated in acid alcohol (1% HCl in 70% alcohol) for 10 seconds. Following that the slides were washed in running tap water for 5 minutes before the final staining step with 1% eosin (Sigma-Aldrich) for 30 seconds. Sections were washed in tap water for 2 minutes and then dehydrated through increasing concentrations of ethanol (50%, 70%, 90%, 100%) for 30 seconds, cleared in Citroclear (HD supplies) twice for 2 minutes and mounted using DPX mounting medium (Thermo Scientific). Slide pictures were taken using the Pannoramic 250 Flash slide scanner (3DHISTECH, Budapest, Hungary) or an Axioskop upright microscope with a Coolsnap ES camera (Photometrics, Arizona, USA) through MetaVue Software (Molecular Devices, CA, USA). All pictures were analysed using Panoramic Viewer software (3DHISTECH) and/or ImageJ (http://rsb.info.nih.gov/ij). For enumeration of crypt lengths, the average number from a minimum of 20 crypts/section and three sections per mouse was calculated. All slides were measured in a blind, randomized order.

6.2.3 Goblet cell staining

To visualize goblet cells in gut sections, acid and neutral mucopolysaccharides were stained using alcian blue dye/periodic acid- Schiff's (PAS) as first described by Mowry (Mowry, 1956). Paraffin-embedded tissue sections were dewaxed in Citroclear (HD supplies) and rehydrated though a series of decreasing ethanol concentrations as described above (section 6.2.2). Sections were stained in 1% (w/v) alcian blue in 3% (v/v) acetic acid pH=2.5 for 5 minutes and then washed in distilled water. Following that they were oxidised in 1% (w/v) periodic acid (Sigma-Aldrich) for 5 minutes. Slides were rinsed in distilled water for 1 minute, washed in tap water for 5 minutes and again rinsed in distilled water for 1 minute. Sections were then treated with Schiff's reagent (Scientific Laboratory Supplies, Nottingham, UK) for 15 minutes and washed as in the previous step. Finally, they were counterstained

with Mayer's heamatoxylin (Sigma-Aldrich) for 1 minute and 'blued' in running tap water. Sections were dehydrated to 100% (v/v) ethanol, cleared and mounted using DPX mounting medium (Thermo Scientific) as above (section 6.2.2). Goblet cells appear as blue/magenta/purple clusters and the average number from a minimum of 20 crypts/section and three sections per mouse was calculated. All slides were measured in a blind, randomized order.

6.3 Molecular biology

6.3.1 Analysis of gut microbial communities based on 16S rRNA gene

Faecal samples were collected from *mdr1a*$^{-/-}$ and WT mice, and frozen immediately in dry ice. Distal colon tissue was excised, opened up longitudinally and washed in sterile phosphate-buffered saline (PBS) for the removal of luminal contents. Mucus was scraped for the collection of bacteria that inhabited the outer mucus and the inner adherent layer. Stool Lysis (ASL) buffer from the QIAamp® DNA Stool Mini Kit (Qiagen, Manchester, UK) was added to the mucus. Samples were kept frozen at -80˚ C until further analysis.

6.3.1.1 DNA extraction from faecal and mucus samples

DNA extraction from faecal and mucus samples was performed using the QIAamp® DNA Stool Mini Kit (Qiagen). In brief, samples were lysed by incubation with Stool Lysis (ASL) buffer at 95˚C for 45 minutes. Samples were also bead beaten for thorough homogenization. After lysis, DNA-damaging substances and other inhibitors released in the sample were adsorbed to InhibitEX matrix tablets. After incubating the suspension for 1 minute at RT, samples were centrifuged at full speed to pellet any inhibitors that had bound to InhibitEX matrix. The supernatant was incubated with proteinase K in the presence of Lysis (AL) buffer at 70˚C for 10 minutes for the digestion and degradation of proteins. Absolute ethanol (96-100%) (Sigma-Aldrich) was subsequently added and the lysate was loaded onto the QIAamp spin column. DNA was adsorbed onto the silica membrane after two centrifugation steps. DNA was then purified by washing with wash buffer 1 (AW1) and AW2 sequentially, and eluted in Elution (AE) buffer. DNA yield was determined by measuring the concentration of

the eluted DNA at 260nm in a Nanodrop ND-1000 Spectrophotometer (Labtech International, Uckfield, UK). Purity of DNA was also determined by comparing the A260/A280 ratio, which should be between 1.7-1.9. Extracted DNA was run in 1% agarose gel to further verify its quality and quantity. DNA concentration was adjusted at 100ng/µl for stool samples and 250ng/µl for mucus samples, and DNA was stored at -20˚C for further use.

6.3.1.2 Amplification of 16S rRNA gene

For the identification of different bacterial species, 16S rRNA gene was amplified using the primer pairs P3_GC-341F (5'-CGC CCG CCG CGC GCG GCG GGC GGG GCG GGG GCA CGG GGG GCC TAC GGG AGG CAG CAG -3') and P2_518R (5'-ATT ACC GCG GCT GCT GG-3'). 100ng of DNA from stool samples or 250ng of DNA from mucus samples was added in each PCR reaction, which contained at a final concentration 2mM $MgCl_2$, 200µM of each dNTP, 0.4µM of each primer pair and 2U of FastStart Taq DNA polymerase (Roche Diagnostics). Bovine Serum Albumin (BSA) (New England Biolabs, Herts, UK) at a final concentration of 0.1 µg/µl was also added to the PCR mixture. The PCR conditions used were as follows: denaturation/ activation at 95˚ for 5 minutes followed by 30 cycles of denaturation at 95˚ for 1 minute, annealing at 55˚ for 1 minute and elongation at 72˚ for 1 minute, and a final elongation step at 72˚ for 10 minutes.

Amplified DNA was purified using the MinELute PCR Purification kit (Qiagen). DNA was solubilised in the presence of PB buffer containing a pH indicator and withheld into a QIAquick column after centrifugation for 1 minute at maximum speed. The flow-through was discarded and the QIAquick column was washed with PE buffer followed by two centrifugations each for 1 minute at maximum speed. DNA was finally eluted in EB buffer (10mM Tris-Cl, pH=8.5). The concentration of DNA was determined by measurement at 260nm and was adjusted, so that the final concentration is 30ng/µl.

6.3.1.3 Denaturing gradient gel electrophoresis (DGGE)

A denaturing gradient acrylamide gel was prepared according to the technique first developed by Fischer and Lerman (Fischer and Lerman, 1983). Initially, two stock solutions either with

100% denaturant or without denaturant were prepared. For making the without denaturant solution, 40% acrylamide/bis-acrylamide solution (Fisher Scientific, supplied by Thermo Fisher Scientific, Loughborough, UK), 50x TAE (2M Tris base, 1M glacial acetic acid, 50mM EDTA pH=8) and dH_2O were used, whereas the 100% denaturant solution additionally included 40% deionized formamide (Sigma-Aldrich) and 7M urea (Sigma-Aldrich). Stock solutions were degased under vacuum and then low and high denaturant solutions (i.e. 30% and 70% respectively) were prepared from the stock solutions. Finally, 10% ammonium persulfate (Sigma-Aldrich) and *N*,*N*,*N*′,*N*′-Tetramethylethylenediamine (TEMED) (Sigma-Aldrich) were added to facilitate the polymerization. The denaturing gel gradient was formed using the gradient maker Hoefer SG-50 (Hoefer, MA, USA). 150ng of purified DNA per sample was loaded onto a DGGE gel having a 30%-70% denaturing gradient formed with urea and formamide. DNA was loaded in a 2x gel loading dye (0.25 ml 2% bromophenol blue, 0.25ml 2% xylene cyanol, 7ml 100% glycerol and 2.5ml H_2O). The DGGE gel was run at 60˚C for 16 hours at 63V in 1x TAE buffer using the DCODE Universal Mutation Detection System (Bio-Rad, Hurts, UK). Gels were stained with SYBR Gold nucleic acid stain (Invitrogen, supplied by Thermo Fisher Scientific) for 30 minutes. Gel bands were visualized under UV illumination and RAW images were captured using a Canon D60 DSLR camera.

6.3.1.4 Cluster analysis

Gel images were processed using Adobe Photoshop CS6 (Adobe, CA, USA) and were then analysed using Bionumerics software (Applied Maths, TX, USA). At first, lane boundaries were defined in order to correct any potential distortion of lanes and normalization values were set up to reduce the background noise. Reference bands were defined across the gel for proper alignment and all bands were automatically detected. Then, these fingerprint data were used to create pairwise comparisons between lanes. The Unweighted Pair Group Method with Arithmetic Mean (UPGMA) algorithm was used to compare matching profiles and similarity relationships were depicted in a dendrogram. UPGMA is an unsupervised hierarchical clustering algorithm that uses the mean similarity across all cluster data values (Loewenstein et al., 2008). A similarity matrix was also extracted and principal component analysis (PCA) was performed using MATLAB (MathWorks, MA, USA) to reveal any relationships among

clusters found in the dendrogram. Non-metric multidimensional scaling (NMDS) using Bray-Curtis dissimilarity distances and statistical analysis (PERMANOVA; adonis function) were performed using the vegan package in R (http://cran.r-project.org).

6.3.2 Quantification of bacterial load and Gram-positive/Gram-negative bacteria numbers

6.3.2.1 Transformation of JM109 competent cells

The pCR™2.1-TOPO® plasmid vector (Life technologies, acquired by Thermo Fischer Scientific, Loughborough, UK) containing a cloned region (179 bp) of the 16S rRNA gene (provided by Dr L. Hooper) was used to construct standard curves for bacteria quantification. 1µg of plasmid DNA was blotted in Whatman filter paper (Sigma-Aldrich) and eluted in 100µl EB buffer (10mM Tris-HCl, pH=8.5) (Qiagen). Eluted plasmid vector transformed JM109 competent cells (Stratagene, CA, USA) for plasmid purification. JM109 cells were aliquoted (100µl) and 0.8µl of β-mercaptoethanol was added to the cell aliquot and left for 10 minutes on ice, swirling gently every 2 minutes. 1µl of the pCR™2.1-TOPO® plasmid DNA was added and cells were incubated on ice for 30 minutes followed by a heat-pulse at 42°C for 45 seconds and an additional incubation on ice for 2 minutes. 0.9ml of preheated SOC medium were added to the cell aliquot and incubated at 37°C for 1 hour with shaking at 225-250rpm. Cells were concentrated by centrifugation at 150g for 10 minutes followed by resuspension in 150µl of SOC medium. The transformation mixture was plated on LB agar plates containing 50µg/ml kanamycin and plates were incubated at 37°C overnight for selection of positive colonies.

6.3.2.2 Purification of plasmid DNA

Positively selected colonies were inoculated in LB liquid broth containing 50µg/ml kanamycin and cultured overnight. Purification of plasmid DNA was performed using QIAprep® Miniprep kit (Qiagen) according to manufacturer's guidelines. Bacterial overnight cultures of positively selected colonies were harvested by centrifugation at 10000g for 3 minutes at RT. Bacterial cells were resuspended in buffer P1 and lysed with the addition of

buffer P2. Buffer N3 was added to the lysate, which was then centrifuged at full speed for 10 minutes. The supernatant was loaded onto the QIAprep spin column and plasmid DNA was adsorbed onto the QIAprep silica membrane after a centrifugation step at full speed. Plasmid DNA was purified by washing with buffer PE and eluted in EB buffer. Plasmid DNA yield was determined by measuring the concentration of the eluted DNA at 260nm in a Nanodrop ND-1000 Spectrophotometer (Labtech International). Plasmid DNA was stored at -20°C for further use.

6.3.2.3 Quantification of bacterial load

Standards for the calculation of bacterial copy number were based on the multiplication of the molecular weight of the plasmid (4079bp) with the coefficient 1.096×10^{-21} g/bp (method from Caroline Sommerville, Strathclyde University, Glasgow). Given the mass of a single copy of the plasmid by the previous equation and the initial concentration of purified plasmid DNA, the amount of added water in the initial sample was calculated, resulting in the generation of standards of final plasmid DNA concentration with each one corresponding to a different bacterial copy number. Different plasmid DNA concentrations reflecting different copy numbers were used to generate a standard curve to measure bacterial abundances. Bacterial load in the faecal and mucus samples was measured by qPCR, using universal 16S rRNA gene primers (forward 5'-ACTCCTACGGGAGGCAGCAGT -3' and reverse 5'-ATTACCGCGGCTGCTGGC -3') and the Power SYBR® Green Master Mix (Applied Biosystems, supplied by Thermo Fisher Scientific) as previously described (Vaishnava et al., 2011). Each qPCR reaction contained 1x SYBR Green PCR mastermix, 300nM of each primer pair, 1µl template and nuclease-free water (Promega) made up to 20µl. qPCR was carried out in a StepOnePlus™ Real-Time PCR system (Applied Biosystems) with the following cycling conditions: 10 minutes at 95°C for 1 cycle, and 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. Bacterial measurements in faecal and mucus samples were normalized to the initial amount of each sample. Bacterial measurements correspond to the number of gene copies and not actual bacterial numbers or colony forming units (CFUs).

6.3.2.4 Quantification of Gram-positive and Gram-negative bacteria numbers

Gram-positive and Gram-negative bacteria numbers in faecal and mucus samples were quantified using Taqman® Gene Expression Master Mix (Applied Biosystems) with eubacterial-specific 16S primers (forward 5'-CAACGCGAAGAACCTTACC-3', reverse 5'-ACGTCATCCCCACCTTCC-3') and specific probes (Gram-positive probe 5'-FAM-ACGACAACCATGCACCACCTG-TAMRA-3', Gram-negative probe 5'-HEX-ACGACAGCCATGCAGCACCT-TAMRA-3') as previously described (Wu et al., 2008). Each qPCR reaction contained 10µl Taqman® Gene Expression Master Mix, 2µl of each 16S primer pair (4µM), 1µl template, 2µl of Gram-positive or Gram-negative probe (1µM) and 3µl nuclease-free water (Promega) in a total volume of 20µl. Taqman qPCR was carried out in a StepOnePlus™ Real-Time PCR system (Applied Biosystems) with the following cycling conditions: an initial holding stage for 2 minutes at 50°C, then 10 minutes at 95°C and 40 cycles of 15 seconds at 95°C and 1 minute at 60°C. Bacterial measurements in faecal and mucus samples were normalized to the initial amount of each sample. Bacterial measurements correspond to the number of gene copies and not actual bacterial numbers or CFUs.

6.3.3 16S metagenomic sequencing library preparation

6.3.3.1 Amplicon PCR

Amplicon 16S PCR was performed using 16S rRNA primers targeting the V3 and V4 region with Illumina overhang adapter sequences attached (forward 5'-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG CCT ACG GGN GGC WGC AG-3', reverse 5'-GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GGA CTA CHV GGG TAT CTA ATC C-3'). Each PCR reaction contained at a final concentration 12.5ng of microbial genomic DNA, 250nM of each primer pair, 1x KAPA HiFi HotStart ReadyMix (KAPA Biosystems, MA, USA) and nuclease-free water (Promega) made up to 25µl. PCR conditions were as follows: 95ºC for 3 minutes followed by 25 cycles of 95ºC for 30 seconds, 55ºC for 30 seconds and 72ºC for 30 seconds, and a final extension step at 72ºC for 5 minutes. The size of the PCR product (~550bp) was verified by running an agarose gel.

6.3.3.2 PCR clean-up

16S PCR amplicons were purified from free primers and primer dimer species using Agencourt® AMPure® XP beads (Beckman Coulter Genomics, MA, USA). 20µl of AMPure® XP beads brought at RT were added to each PCR sample in a 96-well U-bottom microplate (Thermo Scientific) and incubated at RT for 5 minutes. The plate was placed on a magnetic stand (Life Technologies, acquired by Thermo Fisher Scientific) for 2 minutes until clearance of the supernatant, which was then removed and discarded. With the PCR plate on the magnetic stand, the beads were washed twice with freshly prepared 80% (v/v) ethanol (200µl) for 30 seconds and then left to air-dry for 10 minutes. The PCR plate was removed from the magnetic stand and 52.5µl of 10mM Tris-HCl pH=8.5 (Qiagen) was added to each sample well and left to incubate for 5 minutes. The plate was placed on the magnetic stand for 4 minutes and 50µl of the supernatant from each sample was retrieved and stored at -20ºC until further use. 16S rRNA gene amplicons underwent further preparations steps carried out by the Core Genomics Facility at the University of Manchester for subsequent analysis in the Illumina MiSeq system (Illumina, CA, USA).

6.3.4 Isolation of genomic DNA from bacterial cultures

Escherichia coli and *Staphyloccocus aureus* were cultured in Wilkins-Chalgren broth (Sigma-Aldrich) and Mueller Hinton broth (Sigma-Aldrich) until they reached mid-log phase. Serial dilutions of bacteria were plated into agar plates for calculation of CFUs. Bacterial suspension was further used for genomic DNA extraction. Faecal samples derived from WT mice were cultured under anaerobic and aerobic conditions and single colonies were selected and sub-cultured. DNA extraction and 16S rRNA PCR were performed as described in sections 6.3.1.1 and 6.3.1.2.

6.3.4.1 Isolation of genomic DNA from Gram-negative bacteria

Genomic DNA from a bacterial culture of *Escherichia coli* was isolated using the QIAamp® DNA Mini Kit (Qiagen) according to provided guidelines. Briefly, bacterial cells were lysed in the presence of Tissue Lysis (ATL) buffer and Proteinase K. After incubation at 56°C for

20 minutes, AL buffer was added followed by incubation at 70°C for 10 minutes. Absolute ethanol (96%-100%) was added and the mixture was applied to the QIAamp spin column. After two serial washes with AW1 and AW2 buffer, genomic DNA was eluted in AE buffer. The concentration and purity of DNA was determined by measurement at 260nm in a Nanodrop ND-1000 Spectrophotometer (Labtech International) and by comparing the A260/A280 ratio respectively. DNA was stored at -20°C.

6.3.4.2 Isolation of genomic DNA from Gram positive bacteria

Genomic DNA from a bacterial culture of *Staphylococcus aureus* was isolated using the QIAamp® DNA Mini Kit (Qiagen) according to manufacturer's guidelines. Briefly, bacteria were incubated in 200 µg/ml lysostaphin (Sigma-Aldrich) in 20mM Tris-HCl (pH=8.0), 2mM EDTA, 1.2% Triton® for 45 minutes at 37°C to lyse the multi-layered cell wall. Then, ATL buffer and Proteinase K were added and the mixture was incubated at 56°C for 30 minutes and then for a further 15 minutes at 95°C. Absolute ethanol (96%-100%) was added and the solution was applied to the QIAamp spin column. After two serial washes with AW1 and AW2 buffer, genomic DNA was eluted in AE buffer. The concentration of DNA was determined by measurement at 260nm in a Nanodrop ND-1000 Spectrophotometer (Labtech International). Purity of DNA was determined by the A260/A280 ratio. DNA was stored at -20°C.

6.3.5 Real-time quantitative PCR (RT-qPCR) in gut tissue

6.3.5.1 RNA extraction

Tissue samples from the proximal colon were collected from *mdr1a*$^{-/-}$ and WT mice and placed in TRIsure (Bioline, London, UK) filled Lysing Matrix D 2ml tubes (MP Biochemicals, OH, USA). Samples were immediately snap frozen in dry ice and kept at -80°C until further procession. The tissue was homogenised using a FastPrep®-24 Instrument (MP Biochemicals) twice at 6m/s for 30 seconds and total RNA was extracted according to the manufacturer's instructions for TRIsure. In brief, homogenised samples were transferred to new RNA-free eppendorf tubes and 200µl chloroform (Sigma-Aldrich) was added per 1ml

of TRIsure used for phase separation. The mixture was shaken vigorously for 15 seconds and then left to rest for 2-3 minutes at RT. Afterwards, the samples were centrifuged at 12000g for 15 minutes at 4°C and the upper RNA-containing aqueous phase was collected and 500µl chilled isopropanol (Sigma-Aldrich) was added to it. RNA was left to precipitate for 10 minutes at RT. The solution was centrifuged at 12000g for 15 minutes at 4°C, the supernatant was discarded and the pellet was washed with 1ml 75% (v/v) ethanol. Samples were vortexed and centrifuged at 7500g for 5 minutes at 4°C. Ethanol was aspirated and the pellet was left to air-dry. Finally, RNA was resuspended in 20µl nuclease-free water (Promega) and maintained at -80°C. The concentration of total RNA was measured by absorbance at 260 nm in a Nanodrop ND-1000 Spectrophotometer and the ratio A260/A280 was also calculated to determine the purity of RNA (1.7-2.1).

6.3.5.2 cDNA synthesis

Contaminating genomic DNA was removed by treatment with 2µl RNase-free DNase (Promega) in a mixture containing 3µl nuclease-free water (Promega), 1µl DNase 10x reaction buffer (Promega) and 4µl of RNA (500ng/µl) for 30 minutes at 37°C. 1µl DNase Stop Solution (Promega) was added to stop DNase treatment by incubating at 65°C for 10 minutes. Reverse transcription was carried out by adding 2µl 10µM oligo DT primers (Eurofins, Ebersberg, Germany) and 12µl nuclease-free water (Promega) in 11µl of DNase-treated RNA mixture and incubating at 70°C for 5 minutes. cDNA synthesis mix (8µl 5x reaction buffer (Bioline), 4µl 2.5mM dNTPs (Bioline), 2µl RNasin® ribonuclease inhibitor (40U/µl) (Promega) and 1µl Bioscript™ Reverse Transciptase (200U/ µl) (Bioline)) was added to the RNA mixture and incubated at 42°C for 50 minutes and at 85°C for 5minutes. The cDNA synthesis reaction was stored at -20°C.

6.3.5.3 Quantitative real time PCR

Quantitative real time PCR (RT-PCR) was performed using the LightCycler® 480 Probes master (Roche) on the LightCycler® 480 Instrument (Roche) according to the manufacturer's guidelines. A mouse-specific Universal ProbeLibrary Set (Roche) containing pre-validated, dual-labelled (labelled 5'-terminal with fluorescein/FAM and 3' –proximal with a dark

quencher dye) RT-PCR probes was used. Each PCR reaction contained 10µl of LightCycler® 480 Probes master (Roche), 1µl of each primer pair (10µM), 0.4µl of each probe (10µM), 6.6µl H_2O and 1µl template. Primer sequences for Relm-β, Reg3γ, β-defensin 1, Ang4 and IL-22 and two housekeeping genes (β-actin and villin) are shown in Table 1. Gene expression was analysed using an advanced analysis mode, the E-method (efficiency method), in the Roche Applied Science software. A three-step cycling reaction was performed using the following conditions: 5 minutes at 95°C for 1 cycle and 10 seconds at 95°C, 20 seconds at 60°C and 1 second at 72°C for 40 cycles.

Table 1. List of primers and probes used for RT-PCR.

Gene	Primer sequence	Universal Probe Library Set
β-defensin 1	Left: 5'-CCTGGCTGCCACCACTAT-3' Right: 5'-CTTGTGAGAATGCCAACACCT-3'	Probe 2
Relm-β	Left: 5'-GGAAGCTCTCAGTCGTCAAGA-3' Right: 5'-GCACATCCAGTGACAACCAT-3'	Probe 18
Reg3γ	Left: 5'-GCTTCCCCGTATAACCATCA-3' Right: 5'-GCATCTTTCTTGGCAACTTCA-3'	Probe 51
Ang4	Left: 5'-CCCCAGTTGGAGGAAAGC-3' Right: 5'-CGTAGGAATTTTTCGTACCTTTCA-3'	Probe 106
IL-22	Left: 5'-TGACGACCAGAACATCCAGA-3' Right: 5'-AATCGCCTTGATCTCTCCAC-3'	Probe 94
β-actin	Left: 5'- CTAAGGCCAACCGTGAAAAG-3' Right: 5'- ACCAGAGGCATACAGGGACA-3'	Probe 64
villin	Left: 5'-ATCTCCCTGAGGGTGTGGA-3' Right: 5'- AGTGAAGTCTTCGGTGGACAG-3'	Probe 4

6.4 Immunohistochemistry

6.4.1 Tissue processing

Proximal and distal colon samples taken at autopsy were either placed in 8% (v/v) formaldehyde, 0.9% (w/v) sodium chloride, 2% (v/v) glacial acetic acid and 0.5% (w/v) cetrimide in dH_2O for 24 hours, processed and sectioned as described (section 6.2.1) or

placed in optimal cutting temperature (OCT) compound (Thermo Scientific Raymond Lamb, supplied by Fischer Scientific), snap frozen in dry ice and stored at -80˚C. Sections of OCT embedded tissue were cut at 6µm thickness using a Leica CM 1100 microtome (Leica Biosystems, Milton Keynes, UK), placed onto Superfrost® plus microscope slides (Thermo Scientific) and kept at -80˚C until further use.

6.4.2 Relm-β, Ang4 and Reg3γ immunohistochemistry

Distal colon tissue sections embedded in paraffin were stained for the presence of antimicrobial peptides Relm-β, Ang4 and Reg3γ as previously described (Forman et al., 2012). All sections were dewaxed in Citroclear (HD supplies) and rehydrated as described in section 6.2.2. Slides were washed twice in PBS for 5 minutes after every step during the whole procedure. Endogenous peroxidase activity was quenched by incubation with 1.5U/ml glucose oxidase type II-S: from *Aspergillus niger* (Sigma-Aldrich) in the presence of 1.8 mg/ml D-glycose (Sigma-Aldrich) and 0.064 mg/ml sodium azide (Sigma-Aldrich) for 20 minutes at 37˚C. Antigen retrieval was performed by incubation with pepsin digest-all™3 (Invitrogen) for 10 minutes at 37˚C. Binding to non-specific sites was blocked by 1-hour incubation with 7% (v/v) goat serum (Thermo Scientific) in PBS for Relm-β and Reg3γ or 1.5% donkey serum (Sigma-Aldrich) in PBS for Ang4. The avidin/biotin kit (Life Technologies, acquired by Thermo Fischer Scientific) was used to block endogeous avidin/biotin binding sites according to manufacturer's guidelines. Sections were then incubated with purified polyclonal antibody to Relm-β (Abcam, Cambridge, UK) (1:2000 diluted in PBS) or 12.5 µg/ml sheep anti-mouse polyclonal Ang4 (diluted in 1.5% donkey serum) or 5 µg/ml purified polyclonal Reg3γ antibody (Aviva Systems Biology, CA, USA) (diluted in PBS) for 1 hour at RT. Following that the slides were incubated with the secondary biotinylated goat anti-rabbit IgG $F(ab')_2$ (Santa Cruz Biotechology, TX, USA) at 1 µg/ml in the case of Relm-β and Reg3γ or $F(ab')_2$ donkey anti-sheep biotin (Stratech Scientific, Suffolk, UK) at 0.85 µg/ml in the case of Ang4 for 1 hour. Then, ABC kit (avidin-biotin complex, Vector laboratories, Peterborough, UK) was added for 30 minutes and colour development was monitored under a microscope after treatment with the 3, 3'-diaminobenzidine (DAB) peroxidase substrate kit (Vector laboratories). Sections were counterstained with HaemQS (Vector laboratories) for 1 minute, washed in running tap water

and mounted in permanent aqueous mounting medium (AbD Serotec, Kidlington, UK). Images of the sections were obtained using a SPOT Insight QE camera and SPOT software (Diagnostic instruments, MI, USA) and analysed using ImageJ. For quantification of Relm-β, Ang4 and Reg3γ expression a minimum of 3 sections per mouse was used. The numbers of Relm-β, Ang4 and Reg3γ positive cells per 20 crypts per section were quantified in a blind randomized order. Staining with relevant isotype controls was performed to confirm there was no non-specific staining.

6.4.3 IgA staining

Immunohistochemical analysis for IgA was performed as follows. After air-drying, frozen tissue sections were fixed in 4% (w/v) paraformaldehyde (Sigma-Aldrich) on ice for 10 minutes and washed in 0.1% (w/v) saponin in PBS. Slides were washed twice in 0.1% (w/v) saponin in PBS after every treatment throughout the staining process. Endogenous peroxidase was quenched by incubation with 1.5U/ml glucose oxidase (Sigma-Aldrich) in 1x glucose sodium azide solution in PBS for 20 minutes at 37˚C (section 6.4.2). Non-specific binding was blocked by incubation with the avidin/biotin blocking kit (Life Technologies, acquired by Thermo Fischer Scientific) followed by blocking with 7% (v/v) rat serum (AbD Serotec) for 1 hour. Sections were then incubated with biotin-conjugated rat anti-mouse IgA antibody (BD-Biosciences, Oxford, UK) at 5µg/ml in PBS for 1 hour. After incubation with the ABC kit (avidin-biotin complex, Vector laboratories) for 30 minutes, peroxidase staining (DAB, Vector laboratories) was visualized. Sections were counterstained with HaemQS (Vector laboratories) for 1 minute, washed in running tap water and mounted in permanent aqueous mounting medium (AbD Serotec). The number of IgA positive cells per field of view was quantified in a blind randomized order. Staining with relevant isotype control was performed to confirm there was no non-specific staining.

6.4.4 IL-22 staining

Paraffin embedded sections were dewaxed and rehydrated for IL-22 staining as described in section 6.2.2. Slides were incubated in a microwave oven for 5 minutes in citrate (pH =6) and left at RT for 10 minutes for antigen retrieval. Endogenous peroxidase was quenched in 0.3%

(v/v) H_2O_2 from 30% (v/v) Hydrogen peroxide solution (Sigma-Aldrich) for 20 minutes at RT. Slides were washed in washing buffer (0.05% (w/v) BSA in 1x PBS) for 5 minutes on a shaking platform followed by blocking with the reagent Tyramide Signal Amplification (TSA™) Plus Cyanine 3 (Perkin Elmer, Llantrisant, UK) for 30 minutes at RT. Slides were washed twice in washing buffer and this was repeated after every step during the rest of the procedure. Slides were incubated with the rabbit polyclonal IL-22 (Abcam) (10 µg/ml) or with the isotype control polyclonal IgG (Abcam) for 1 hour at RT followed by a 30-minutes' incubation with the secondary biotinylated goat anti-rabbit antibody (Dako, Cambridgeshire, UK) (8.2 µg/ml). Slides were then incubated with streptavidin-POD conjugate horseradish peroxidase (HRP) (1:1000) (Roche Diagnostics) for 30 minutes at RT. After incubation with DAB (Vector laboratories) for 4 minutes and colour development, slides were counterstained, washed and dehydrated though alcohols (section 6.2.2) before mounting in DPX (Thermo Scientific). The number of IL-22 positive cells per field of view was quantified in a blind randomized order. Staining with relevant isotype controls was performed to confirm there was no non-specific staining.

6.4.5 Mucus preservation in histological sections

6.4.5.1 Tissue fixation to preserve intestinal mucus

Distal colon sections containing faecal material were placed in methanol-Carnoy's fixative solution (60% (v/v) dry methanol, 30% (v/v) chloroform, 10% (v/v) glacial acetic acid). Tissue was fixed for a minimum of 24 hours and a maximum of 1 week at RT as previously described (Johansson and Hansson, 2012). Then, fixed tissue samples were washed two times in dry methanol for 30 minutes each, followed by two washes in absolute ethanol for 20 minutes each. Tissues were further processed with a micro-spin tissue processor STP 120 (Thermo Scientific) and were incubated in ethanol for 15 minutes, twice in xylene for 20 minutes each, and twice in fibrowax pastillated wax (Beckon Dickinson) for 1 hour at 60ºC. The processed tissue samples were then embedded in paraffin using a Microm EC 350 embedding system (Thermo Scientific) and 5µm sections were cut using a Microm HM 325 rotary microtome (Thermo Scientific). The sections were placed on glass slides after floating on a water bath and left to dry.

6.4.5.2 Fluorescence in situ hybridization (FISH) and mucin MUC2 immunostaining

Sections were dewaxed by incubation at 60ºC for 10 minutes followed by two additional incubations in xylene substitute solution (Sigma-Aldrich) each for 10 minutes with the first solution pre-warmed to 60ºC. Afterwards, slides were brought to 99.5% ethanol twice for 5 minutes each and were left to air dry. Sections were then hybridized with 0.5µg of the universal bacterial probe EUB338 (5'–Cy3-GCTGCCTCCCGTAGGAGT-3') targeted against the 16S rRNA gene as previously described (Johansson and Hansson, 2012). 50µl of hybridization buffer (20mM Tris-HCl pH=7.4, 0.9M NaCl, 0.1% (w/v) SDS) pre-warmed to 46ºC was mixed with 0.5µg of EUB338 probe and added dropwise onto sections, which were placed on a humidified chamber and incubated at 46ºC for 2 hours. Slides were washed in pre-warmed (48ºC) FISH washing buffer (20mM Tris-HCl pH=7.4, 0.9M NaCl) for 15 minutes and then immediately immersed in cold water (4ºC) and washed three times in PBS for 2 minutes each. Sections were subjected to a reduction step in 10mM DL-Dithiothreitol (Sigma-Aldrich) (diluted in 0.1M Tris-HCl pH=8) for 30 minutes and an alkylation step in 25mM Iodoacetamide (Sigma-Aldrich) (diluted in 0.1M Tris-HCl pH=8) for 30 minutes. After washing three times in PBS for 2 minutes each, blocking solution (1:100 goat serum in PBS) was added and slides were incubated at 4ºC for 30 minutes. Sections were then incubated with a polyclonal Muc2 antibody raised in rabbit (provided by Dr R. Grencis and Dr D. Thornton) (1:200 dilution in blocking buffer) overnight at 4ºC. After washing three times in cold PBS for 2 minutes each, the secondary antibody goat anti-rabbit Alexa-Fluor® 488 (Life technologies, acquired by Thermo Fischer Scientific) (2.5 µg/ml in PBS) was added and the sections were incubated for 1 hour at 4ºC. Sections were washed three times in cold PBS for 2 minutes each and mounted in ProLong® antifade mounting medium containing 4',6-diamidino-2-phenylindole (DAPI) (Life technologies, acquired by Thermo Fischer Scientific). A nonspecific probe directed at fungus Cryptococcus was used as a control (5'-Cy3-CCAGCCCTTATCCACCGA-3'). Slides were visualized using an Olympus BX51 upright microscope and captured using a Coolsnap ES camera (Photometrics) through MetaVue Software (Molecular Devices) and analysed using ImageJ. Specific band pass filter sets for DAPI, FITC and Cy3 were used to prevent bleed-through from one channel to the next. A minimum of three different distal colon sections was used for each sample.

6.5 16S rRNA analysis of microbial communities

6.5.1 16S rRNA gene sequence processing and operational taxonomic unit (OTU) selection

16S amplicon sequencing targeting the V3 and V4 variable regions of the 16S rRNA gene was performed on the Illumina MiSeq platform (section 6.3.3) and generated paired-end reads of 300bp in each direction. Illumina reads were initially demultiplexed to remove adapter sequences and trim primers. Then, Illumina paired-end reads were merged together using SeqPrep (http://github.com/jstjohn/SeqPrep) and submitted to EMBL-EBI's metagenomics pipeline (Hunter et al., 2014). Reads were pre-processed to remove low-quality and uninformative reads before further analysis steps in the Quantitative Insights Into Microbial Ecology (QIIME) pipeline v.1.9.0 (Caporaso et al., 2010b). The quality-filtering process included removal of reads with low quality ends (i.e. ambiguous leading/trailing bases), removal of reads where the proportion of ambiguous bases is higher than 10% and removal of reads with length less than 300bp. Sequences corresponding to rRNA were selected using a filter for prokaryotic rRNA reads and were clustered to OTUs using a closed-reference OTU picking strategy in QIIME based on uclust (Edgar, 2010). OTU representative sequences were then taxonomically classified using the Greengenes reference database (version 13.8) filtered at 97% identity (McDonald et al., 2012; Wang et al., 2007). A resulting OTU table was generated giving the OTU abundances in each sample with taxonomic identification for each OTU.

Sequences were also processed using the vsearch analysis pipeline (https://github.com/torognes/vsearch). Demultiplexed sequences derived from Illumina pipeline initial processing were joined together and annotated to their respective sample ids, so that each read would indicate which sample it came from. Sequences were then quality filtered using the value 19 as phred quality threshold in QIIME (Bokulich et al., 2013). The following steps were performed in vsearch pipeline. Sequences were dereplicated and sorted by size to discard single reads. Clustering was performed using cluster_smallmem and 97% pairwise identity with maximum accepts and rejects values set at 8 and 64 respectively. Chimeras were detected using uchime denovo followed by uchime_ref using the gold database as a reference database (Edgar et al., 2011). Mapping to initial input sequences was

performed using global pairwise alignment (usearch_global) at 97% identity with top_hits_only option enabled. The resulting map.uc file was converted into a QIIME cluster file by running the script uc2otutab_mod.py (http://drive5.com/python/). Afterwards, taxonomy was assigned using the RDP (Ribosomal Database Project) classifier (Wang et al., 2007) and the Greengenes database (version 13.8) (McDonald et al., 2012) resulting in the generation of an OTU table. The OTU representative sequences were aligned to the Greengenes core set alignment using PyNAST (Caporaso et al., 2010a). A lane mask filter was applied to the aligned sequences to remove highly variable regions and retain the conserved positions of 16S rRNA gene for phylogenetic inference. A phylogenetic tree was built after alignment of representative sequences using FastTree (Price et al., 2009). Figure 6.1 provides an overview of QIIME analysis pipeline for 16S high-throughput sequencing data.

Singletons and OTUs with less than 10 sequences as total observation count were excluded from subsequent analysis. The resulting OTUs table was checked for contamination using SourceTracker, which identifies the potential source and proportion of contamination (Knights et al., 2011). OTU tables were summarized at different taxonomic levels (i.e. phylum, class, order, family and genus) showing relative abundance for each sample.

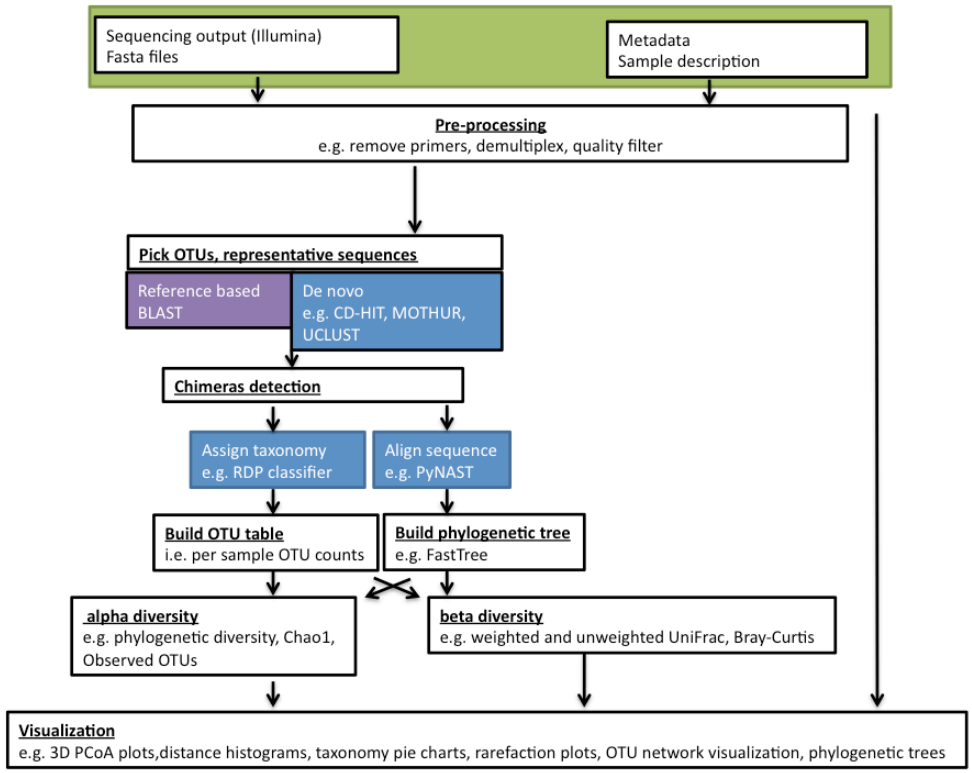


Figure 6.1. Overview of the QIIME analysis pipeline. Sequencing reads combined with metadata supplied by the user undergo an initial pre-processing stage to demultiplex the barcoded primers from the samples and to perform quality filtering. At the next stage, sequences are clustered into OTUs either using a reference database or de-novo based on sequence similarity. Data are 'denoised' to remove chimeric sequences. Then, taxonomy is assigned, sequences are aligned and a phylogenetic tree is built. At this stage, an OTU table showing the counts of each OTU in each sample is produced. The phylogenetic distance and the OTU table can be used to perform alpha- and beta-diversity measurements, which are combined with the metadata to produce graphics for easy interpretation of the data.

6.5.2 Microbial community compositional analysis

Alpha- and beta-diversity measures, reflecting within and between samples' diversity respectively, were calculated with QIIME (Caporaso et al., 2010b) and principal coordinate analysis (PCoA) plots were produced in Emperor (Vazquez-Baeza et al., 2013). Measures of alpha-diversity such as Chao1, observed OTUs, Faith's phylogenetic diversity and Shannon index were calculated based on rarefied OTU tables to account for variation in sequencing depth. The distance between microbial communities was calculated using weighted and unweighted UniFrac (Lozupone et al., 2006; Lozupone and Knight, 2005). UniFrac metric is a phylogenetic distance method for comparison of microbial communities based on the degree to which pairs of communities share a fraction of the branch length in a common phylogenetic tree (Lozupone and Knight, 2005). Weighted UniFrac weighs the phylogenetic differences between microbial communities according to the relative abundance of each taxon (quantitative measure), whereas unweighted UniFrac distances are dependent on the presence/absence of microbial community members (qualitative measure) (Lozupone et al., 2006; Lozupone and Knight, 2005). PCoA analysis was employed to assess microbial community variation among samples and to reveal potential clustering patterns. Alpha-diversities were compared using compare_alpha_diversity.py and beta diversities were compared using adonis in compare_categories.py. Differences in OTUs abundance were tested using both DESeq (Love et al., 2014; McMurdie and Holmes, 2014) and by rarefying microbial sequencing data.

To examine complex microbial communities with many confounding variables (covariates), multivariate statistical analysis was applied by regressing the relative abundance of each bacterial taxon on each one of the sample variables. Multivariate association with linear models (MaAsLin) analysis is a pipeline that uses multivariate association tests to identify significant associations of bacterial taxa with a specific sample variable without the confounding effect of other sample variables. MaAsLin, an additive general linear model with boosting, was performed using default parameters (http://huttenhower.sph.harvard.edu/maaslin) (Morgan et al., 2012). Sample id was used as a random effect and the following variables (genotype, age, location, colitis, litter, housing and total reads) were used to identify association with the abundance of each taxon. All p-values were corrected for multiple hypothesis testing using the Benjamini- Hochberg false discovery

rate (FDR) method (Benjamini and Hochberg, 1995), and significant association was considered below a FDR q-value threshold of 0.05.

Data presented in the study are derived from the EBI analysis. MaAsLin analysis was performed using vsearch data. Analysis via the EBI and vsearch pipeline concurred. Phylum abundance data were used to build a model to predict sample location (i.e. mucus or stool) and it can be found in the link file:///Users/maria/Downloads/metagenome.html.

6.5.3 Metagenome inference and functional pathway reconstruction

In order to predict the functional profiles of microbial communities from 16S sequence data, we used PICRUSt (phylogenetic investigation of communities by reconstruction of unobserved states) (Langille et al., 2013). PICRUSt is an ancestral state reconstruction algorithm, which is based on marker gene survey data such as 16S rRNA sequencing data and a database of known reference genomes to estimate with high accuracy the gene content of microbial communities. The resulting OTU table from closed reference OTU picking strategy in QIIME analysis was normalized, by dividing each OTU with its predicted 16S rRNA gene copy number to better reflect the true abundance of each OTU. Normalized OTU abundances were then multiplied by gene families' abundances inferred for each OTU using a reference database. The final output from this gene content prediction step was an annotated table with gene family counts for every sample, where gene family corresponds to KEGG (Kyoto Encyclopedia of Genes and Genomes) Orthology (KO) copy numbers (Kanehisa et al., 2012). Inferred KO gene abundances were summarized at a higher hierarchical level at pathway-level categories for easier biological interpretation of metagenomic data.

To examine PICRUSt's predictive accuracy on functional abundance, the weighted nearest sequenced taxon index (NSTI) values were calculated. NSTI values represent the average branch length that separates each OTU in a sample from its closest relative with a sequenced reference genome, weighted by the abundance of that OTU in the sample (Langille et al., 2013). Therefore, NSTI values summarize the extent to which OTUs in a sample are related to sequenced genomes. Low NSTI values indicate higher prediction accuracy. For example, an NSTI value of 0.03 means that the gene content of the average OTU in a given sample can be predicted from a reference genome that belongs to same species (97% sequence identity)

(Langille et al., 2013). Non-microbial categories such as 'Organismal Systems' and 'Human Diseases' were excluded from further analysis. Comparisons between groups based on KEGG pathway abundance were performed using the group_significance.py in QIIME (Caporaso et al., 2010b). Beta diversity of KO data was calculated using the Bray-Curtis distance metric and visualized using Principal Coordinate Analysis (PCoA) in QIIME using Emperor (Vazquez-Baeza et al., 2013). Data were further analyzed using STatistical Analysis of Metagenomic Profiles (STAMP) software (Parks and Beiko, 2010).

6.6 Statistical analysis

Statistical analysis was performed using R, Matlab and GraphPad Prism 6 (GraphPad software, CA, USA). Detailed analysis of the statistical tests applied for each dataset are given in the materials and methods section of each results' chapter.

6.6.1 Analysis of metabolites data

6.6.1.1 Multivariate statistics on liquid chromatography – mass spectrometry (LC-MS) data

LC-MS data of urine samples were subjected to multivariate statistical analysis using KNIME (Berthold et al., 2008; Mazanetz et al., 2012; O'Hagan and Kell, 2015) and R (http://cran.r-project.org). Principal components analysis (PCA) was performed to provide an overview of the samples' distribution and identify potential patterns of variation. Data pre-processing involved removing quality control (QC) samples and "singletons", followed by application of a correlation filter for removal of correlated features (threshold = 0.98) and Z -scores normalization ($Z = (x - \mu)/\sigma$). PCA calculates principal components, which are linear combinations of the initial variables (i.e. metabolites), explaining most of the variation within the dataset (Ramette, 2007). Score plots of PCA analysis were generated and each sample was represented in the new coordinate space. The corresponding loading plot for each principal component was also produced to indentify which mass ions contribute to patterns of variation as observed in the scores plot.

Multivariate regression was applied for data analysis, as it correlates independent variables in matrix X (i.e. metabolomics data) to corresponding dependent variables in matrix Y (i.e. groups, classes) (Barker and Rayens, 2003). This approach aims to maximize the covariance between X and Y matrices by finding a linear relation. Thus, partial least squares (PLS) regression was used to construct predictive regression models for better discrimination of sample groups (Barker and Rayens, 2003; Gromski et al., 2015). Y variables (i.e. sample groups) were predicted from the model based on a reduced number of factors (PLS components) (Gromski et al., 2015). The performance of each model was tested using cross-validation with the 'leave one out' method. All data were used for training in the model, which potentially does not rule out potential over-fitting of the data.

Random forests (RF) regression was further applied to build prediction models (O'Hagan and Kell, 2015). RF is a classification method, in which many decision trees are constructed using different sets of random variables and samples (Breiman, 2001; Cutler et al., 2007). An advantage of this method is that it is robust to over-fitting and no data transformation (such as standardization) is required prior to the analysis (Gromski et al., 2015). The original data is split in training and test sets using bootstrapping (with replacement), whereby training sets are useful for tree construction and test sets for calculation of prediction accuracy.

A specific form of PLS regression is PLS- linear discriminant analysis (PLS-LDA). PLS-LDA, a supervised classification method, relates LC-MS variables to the class membership of samples to maximize the separation of samples according to their classification. Therefore, PLS-DA handles dependent categorical variables compared with PLS regression that uses dependent continuous variables (Gromski et al., 2015). The PLS-LDA model chosen was the one that gave the lowest mean classification error rate for 20 "bootstrap samples". The misclassification matrix describes the number of correctly predicted samples, the specificity, sensitivity and the positive and negative predictive values of the model. Score plots were generated and mass ions responsible for differences between classes were searched for by inspection of regression vectors and variable importance in projection (VIP) scores. However, as the VIP threshold is hard to define, a way of settling this is by looking at separate validation data versus threshold. As the data are insufficient for this, a features permutation approach was followed.

6.6.1.2 Feature permutation

LC-MS peaks of permuted features using the whole dataset as input showed that that the three ion signals coming as significant were of low spectral intensity (Chapter 4, Supplementary Figure 7A-B). Therefore, confidence in mass accuracy was not sufficient to assign these mass ions to known metabolites. As 6-week animals had higher variation in the targeted metabolites than 18-week animals, which were more closely clustered, and since genotype appeared to be the main discriminating factor, subsequent permutation analysis was performed including only 18-week animals.

Permutation analysis based on RF classification of 18-week LC-MS samples identified four ions as discriminatory (Chapter 4, Supplementary Figure 8A). The first ion detected was however absent from the profiling data array due to data misalignment (the spectral ion matrix was binned with a 15mDa tolerance and 0.2min). As a consequence of data misalignment, data were re-processed with a 0.3min tolerance in retention time. Setting data alignment tolerances aims to enable alignment of chromatographically resolved ions from different data files; setting a tolerance too narrow can cause misalignment through systematic changes during batch acquisition, conversely setting tolerances too wide can cause incorrect ion binning due to isomers (particularly lipid species), which can be binned incorrectly with too high a tolerance. The same can also happen with mass tolerance misalignment however with good LC separation applied it is rare for retention time and mass tolerance issues to occur at the same time. In the case of this data, analysis of peak area data after re-processing revealed that retention time misalignment had occurred and the ion found as significant was in fact the same in all sample groups and was not significant. Following reprocessing the data with wider retention time tolerance may have slightly changed the PCA plots, regression and classification results, so they were reprocessed and recalculated but no change was found.

The spectral matrix data processing parameters were set to “de-isotope” the data array to avoid duplication of ions, however some isotopes can still remain in the matrix if isotope intensities are not sufficiently aligned. The statistical analysis, although very powerful, may have been finding features within LC-MS noise; thereby data were re-processed again applying a noise thresholding set to 1,000,000 (previously set to 100,000). Nevertheless, this approach still generated a small number of noise ions, so a second stage analysis was applied that generated a Chromatogram Matrix in which generic peak integration parameters were

applied to all peaks identified in the spectral matrix.

When we have relatively few samples and noisy data, machine learning methods can often pick out noise as features and as a result this warrants cautiousness about claims made for the contribution of certain ions. To deal with this issue, re-pre-processing data offers a way of systematic error removal. Another permutation technique including permutation of the target class a few hundred times was also used, as in that way the link between features and target class would break, allowing us to determine the likelihood of getting a good classification accuracy "by accident"; additionally it may give a similar insight into accidental feature ranking. A caveat in this method is that issues such as de-isotoping errors and mass binning errors would probably manifest as systematic errors and typical statistical methods such as permutation may not be of help.

PCA and regression analysis of the newly pre-processed data also led to similar conclusion as the initial analysis before re-processing. Classification accuracy of this data was similar to that found for the full data set, however feature importance was not flagged as significant when looking at q-values; the best q-value was very poor at 0.48 (Chapter 4, Supplementary Figure 8B). Since using the 18-week data only results in the number of cases being halved, it is possible that the power of the analysis was compromised.

References

Barker, M., and Rayens, W. (2003). Partial least squares for discrimination. Journal of Chemometrics *17*.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Statist. Soc. *57*, 289-300.

Berg, D.J., Davidson, N., Kuhn, R., Muller, W., Menon, S., Holland, G., Thompson-Snipes, L., Leach, M.W., and Rennick, D. (1996). Enterocolitis and colon cancer in interleukin-10-deficient mice are associated with aberrant cytokine production and CD4(+) TH1-like responses. The Journal of clinical investigation *98*, 1010-1020.

Berthold, M.R., Cebron, N., Dill, F., Gabriel, T.R., Kötter, T., Meinl, T., Ohl, P., Sieb, C., Thiel, K., and Wiswedel, B. (2008). KNIME: The Konstanz Information Miner. In Data Analysis, Machine Learning and Applications: Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation e.V., Albert-Ludwigs-Universität Freiburg, March 7–9, 2007, C. Preisach, H. Burkhardt, L. Schmidt-Thieme, and R. Decker, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 319-326.

Bokulich, N.A., Subramanian, S., Faith, J.J., Gevers, D., Gordon, J.I., Knight, R., Mills, D.A., and Caporaso, J.G. (2013). Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nature methods *10*, 57-59.

Breiman, L. (2001). Random Forests. Machine Learning *45*, 5-32.

Caporaso, J.G., Bittinger, K., Bushman, F.D., DeSantis, T.Z., Andersen, G.L., and Knight, R. (2010a). PyNAST: a flexible tool for aligning sequences to a template alignment. Bioinformatics *26*, 266-267.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I.*, et al.* (2010b). QIIME allows analysis of high-throughput community sequencing data. Nature methods *7*, 335-336.

Collett, A., Higgs, N.B., Gironella, M., Zeef, L.A., Hayes, A., Salmo, E., Haboubi, N., Iovanna, J.L., Carlson, G.L., and Warhurst, G. (2008). Early molecular and functional changes in colonic epithelium that precede increased gut permeability during colitis development in mdr1a(-/-) mice. Inflammatory bowel diseases *14*, 620-631.

Cutler, D.R., Edwards, T.C., Jr., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., and Lawler, J.J. (2007). Random forests for classification in ecology. Ecology *88*, 2783-2792.

Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics *26*, 2460-2461.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. Bioinformatics *27*, 2194-2200.

Fischer, S.G., and Lerman, L.S. (1983). DNA fragments differing by single base-pair substitutions are separated in denaturing gradient gels: correspondence with melting theory.

Proceedings of the National Academy of Sciences of the United States of America *80*, 1579-1583.

Forman, R.A., deSchoolmeester, M.L., Hurst, R.J., Wright, S.H., Pemberton, A.D., and Else, K.J. (2012). The goblet cell is the cellular source of the anti-microbial angiogenin 4 in the large intestine post Trichuris muris infection. PloS one *7*, e42248.

Gromski, P.S., Muhamadali, H., Ellis, D.I., Xu, Y., Correa, E., Turner, M.L., and Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis--a marriage of convenience or a shotgun wedding. Analytica chimica acta *879*, 10-23.

Hunter, S., Corbett, M., Denise, H., Fraser, M., Gonzalez-Beltran, A., Hunter, C., Jones, P., Leinonen, R., McAnulla, C., Maguire, E.*, et al.* (2014). EBI metagenomics--a new resource for the analysis and archiving of metagenomic data. Nucleic acids research *42*, D600-606.

Johansson, M.E., and Hansson, G.C. (2012). Preservation of mucus in histological sections, immunostaining of mucins in fixed tissue, and localization of bacteria with FISH. Methods in molecular biology *842*, 229-235.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. Nucleic acids research *40*, D109-114.

Knights, D., Kuczynski, J., Charlson, E.S., Zaneveld, J., Mozer, M.C., Collman, R.G., Bushman, F.D., Knight, R., and Kelley, S.T. (2011). Bayesian community-wide culture-independent microbial source tracking. Nature methods *8*, 761-763.

Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepile, D.E., Vega Thurber, R.L., Knight, R.*, et al.* (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nature biotechnology *31*, 814-821.

Loewenstein, Y., Portugaly, E., Fromer, M., and Linial, M. (2008). Efficient algorithms for accurate hierarchical clustering of huge datasets: tackling the entire protein space. Bioinformatics *24*, i41-49.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology *15*, 550.

Lozupone, C., Hamady, M., and Knight, R. (2006). UniFrac--an online tool for comparing microbial community diversity in a phylogenetic context. BMC bioinformatics *7*, 371.

Lozupone, C., and Knight, R. (2005). UniFrac: a new phylogenetic method for comparing microbial communities. Applied and environmental microbiology *71*, 8228-8235.

Mazanetz, M.P., Marmon, R.J., Reisser, C.B., and Morao, I. (2012). Drug discovery applications for KNIME: an open source data mining platform. Current topics in medicinal chemistry *12*, 1965-1979.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., and Hugenholtz, P. (2012). An improved Greengenes taxonomy

with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The ISME journal *6*, 610-618.

McMurdie, P.J., and Holmes, S. (2014). Waste not, want not: why rarefying microbiome data is inadmissible. PLoS computational biology *10*, e1003531.

Morgan, X.C., Tickle, T.L., Sokol, H., Gevers, D., Devaney, K.L., Ward, D.V., Reyes, J.A., Shah, S.A., LeLeiko, N., Snapper, S.B., *et al.* (2012). Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome biology *13*, R79.

Mowry, R.W. (1956). Alcian blue techniques for the histochemical study of acidic carbohydrates. J. Histochem. and Cytochem *4*, 403-407.

O'Hagan, S., and Kell, D.B. (2015). Software review: the KNIME workflow environment and its applications in genetic programming and machine learning. Genetic Programming and Evolvable Machines *16*, 387-391.

Parks, D.H., and Beiko, R.G. (2010). Identifying biologically relevant differences between metagenomic communities. Bioinformatics *26*, 715-721.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Molecular biology and evolution *26*, 1641-1650.

Ramette, A. (2007). Multivariate analyses in microbial ecology. FEMS microbiology ecology *62*, 142-160.

Schinkel, A.H., Smit, J.J., van Tellingen, O., Beijnen, J.H., Wagenaar, E., van Deemter, L., Mol, C.A., van der Valk, M.A., Robanus-Maandag, E.C., te Riele, H.P., and et al. (1994). Disruption of the mouse mdr1a P-glycoprotein gene leads to a deficiency in the blood-brain barrier and to increased sensitivity to drugs. Cell *77*, 491-502.

Vaishnava, S., Yamamoto, M., Severson, K.M., Ruhn, K.A., Yu, X., Koren, O., Ley, R., Wakeland, E.K., and Hooper, L.V. (2011). The antibacterial lectin RegIIIgamma promotes the spatial segregation of microbiota and host in the intestine. Science *334*, 255-258.

Vazquez-Baeza, Y., Pirrung, M., Gonzalez, A., and Knight, R. (2013). EMPeror: a tool for visualizing high-throughput microbial community data. GigaScience *2*, 16.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology *73*, 5261-5267.

Wu, Y.D., Chen, L.H., Wu, X.J., Shang, S.Q., Lou, J.T., Du, L.Z., and Zhao, Z.Y. (2008). Gram stain-specific-probe-based real-time PCR for diagnosis and discrimination of bacterial neonatal sepsis. Journal of clinical microbiology *46*, 2613-2619.

Appendix

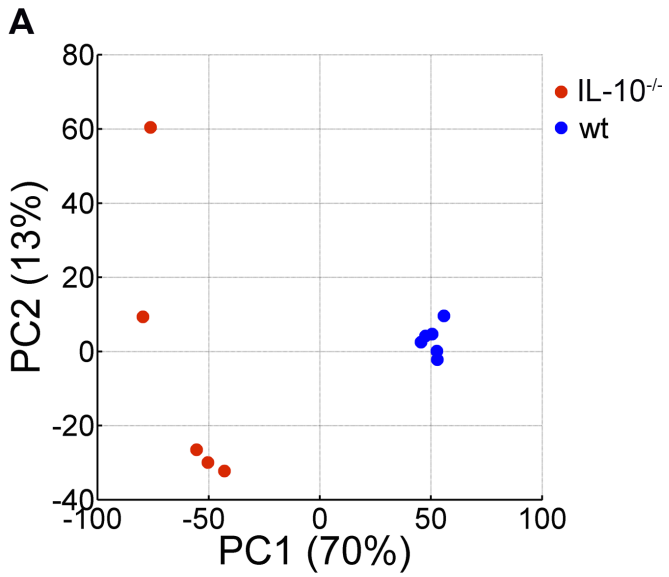


Figure 1. Principical component analysis (PCA) of feacal microbial communities derived from *IL-10*$^{-/-}$ and wt mice. Both groups of mice were infected with a high dose of *Trichuris muris* and faeces were collected at day 16 of infection. PCA based on fingerprint data reveals differential clustering of the two groups.

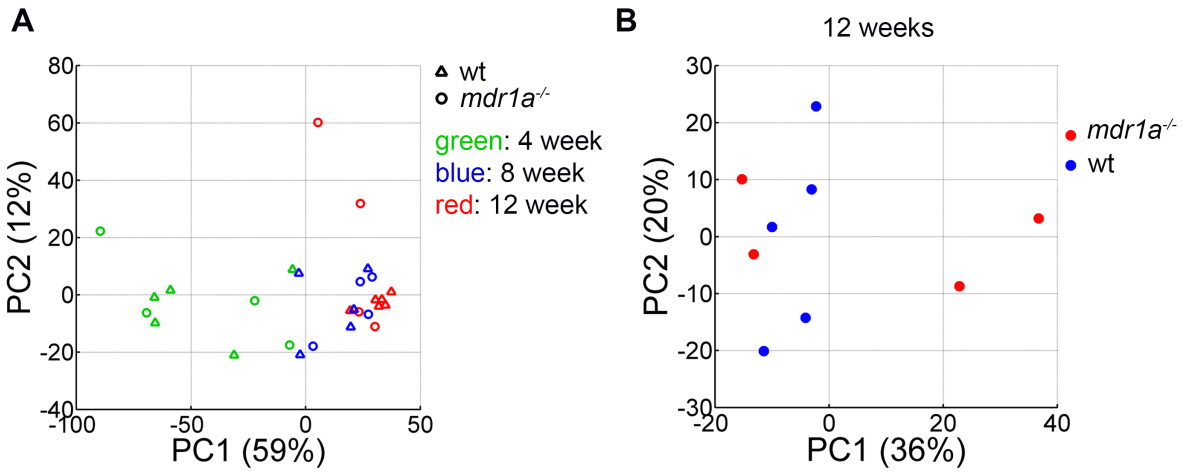


Figure 2. Microbial community composition differs after weaning. Faecal samples were collected from *mdr1a*$^{-/-}$ and wt mice at 4, 8 and 12 weeks of age. **(A)** Principal component analysis (PCA) of fingerprint data shows a change in microbial composition after the mice are weaned. (**A)** PCA of fingerprint data at 12 weeks indicates no clear clustering between faecal microbial communities from *mdr1a*$^{-/-}$ and wt mice.

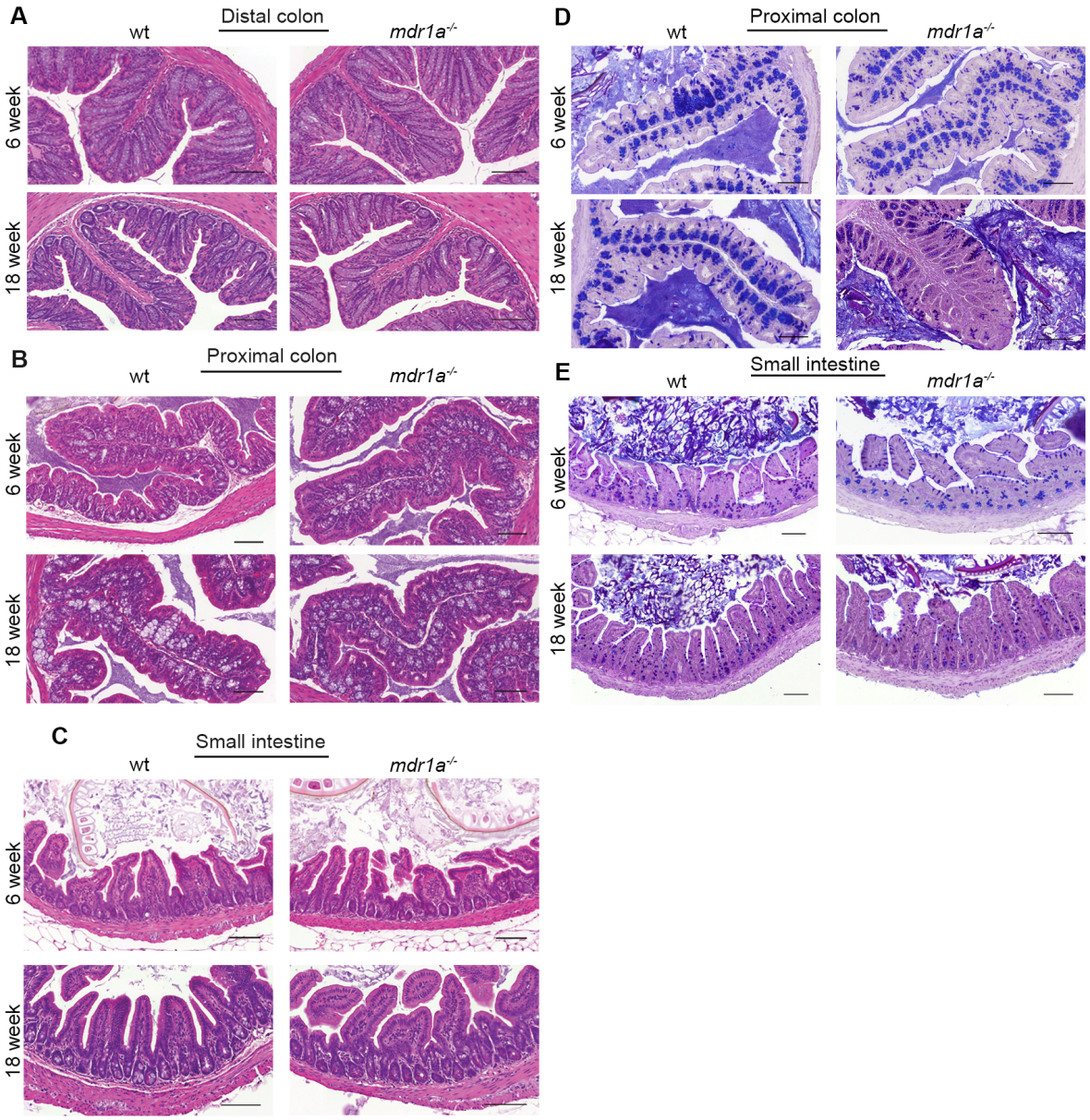


Figure 3. Gut morphology in WT and *mdr1a*$^{-/-}$ mice. Representative H&E sections from the **(A)** distal colon, **(B)** proximal colon and **(C)** small intestine at 6 and 18 weeks. Representative goblet cell stained sections from the **(D)** proximal colon and **(E)** small intestine at 6 and 18 weeks. Scale bar =100μm.