

Technische Universität München
Lehrstuhl für Ernährung und Immunologie

Functional Characterization of Human Gut Microbiota in Inflammatory Bowel Disease Patients Using Gnotobiotic Humanized Mice

Amira Metwaly

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum
Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen
Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:

Prof. Dr. Wolfgang Liebl

Prüfer der

1. Prof. Dr. Dirk Haller

Dissertation:

2. Prof. Dr. Siegfried Scherer

3. Prof. Dr. Barbara Stecher

Die Dissertation wurde am 05.06.2019 bei der Technischen Universität
München eingereicht und durch die Fakultät Wissenschaftszentrum
Weihenstephan für Ernährung, Landnutzung und Umwelt am 29.10.2019
angenommen.

“Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less.”

Marie Curie

Abstract

Dysbiosis and metabolic alterations of the gut microbiome have been described in inflammatory bowel diseases (IBD). The aim of this work is to identify functional signatures associated with disease outcome or response to therapy in two longitudinal cohorts of patients with IBD, and to mechanistically characterize their pathogenic potential using a gnotobiotic humanized mouse model and an integrative multi-omics approach. We studied 68 IBD patients during 5-year treatment with autologous hematopoietic stem cell transplantation (35 CD patients) and during the first year of biological therapy (19 CD and 14 UC), enrolled in Spain and France. Fecal samples and tissue biopsies were collected both at baseline and at different time points during follow-up. To characterize changes in gut microbiome and metabolome, we performed 16S rRNA gene sequencing, global 16S predicted metagenomes, shotgun metagenomic sequencing on a subset of fecal samples and untargeted metabolomics. To address the functional impact of microbial dysbiosis, we established a humanized IBD mouse model by colonizing germfree (GF) *Il-10*^{-/-} mice with selected fecal samples from CD patients with different disease outcome.

Our data showed that temporal fluctuations in gut microbiota and metabolite profiles reflected the patient-related variations and the differences in disease activity. Fecal microbiome of active patients was enriched in microbial taxa involved in sulfur metabolism such as *Escherichia Shigella* and *Fusobacterium* and a high proportion of sulfate reducing bacteria, including *Desulfovibrio* and *Campylobacter*. Fecal metabolic profiling confirmed an enrichment of sulfated compounds, including bile acids, polyphenols and biogenic amines. Predicted metagenomes from 16S revealed enrichment of functional genes associated with sulfate and ion transport system metabolism, in IBD patients with active disease. In contrast, increased abundance of several basic biosynthetic processes correlated with a state of remission. Transplantation of microbiota from patients with active or inactive disease status was reproducibly sufficient to recreate disease phenotype in recipient *Il-10*^{-/-} GF mice. Interestingly, humanized mice reflected the dysbiotic features of their respective human donors and inflammation was driven by a variety of individual community profiles. Using a machine-learning algorithm, we identified a panel of 10 taxa that discriminates humanized mice by inflammatory status, where a microbial signature characterized by an overabundance of *Bacteroides fragilis* and *Desulfovibrio* classified humanized mice by inflammation with high accuracy. In accordance with the signature identified in humans, enrichment of sulfated metabolites was indicative for inflamed phenotype in mice, together with an abundance of genes mapping to sulfate metabolism and Type II, IV and VI secretion systems. Integration of microbiota and metabolite profiles from human and mice improved the predictive modelling of disease outcome significantly

and identified a network of functionally relevant bacteria-metabolite interactions linked to disease activity in CD. This work provides evidence that functional profiling is superior to taxonomic profiling for identifying predictive signatures in IBD. Despite the heterogeneity of CD patients gut microbiome at the taxonomic and metabolite levels, shared functional signatures correlate with disease severity. Multi-omics data integration improved the clinical outcome prediction and identified a signature involving sulfur metabolism and detoxification to be relevant in disease outcome.

ZUSAMMENFASSUNG

Eine Dysbiose und metabolische Veränderungen des Darmmikrobioms sind an entzündlichen Darmerkrankungen (IBD) beteiligt. Ziel dieser Arbeit ist es, funktionelle Signaturen zu identifizieren, die mit dem Krankheitsergebnis oder dem Ansprechen auf die Therapie in zwei Längskohorten von Patienten mit IBD verbunden sind, und ihr pathogenes Potenzial mit Hilfe eines gnotobiotischen humanisierten Mausmodells und eines integrativen Multiomik-Ansatzes mechanisch zu charakterisieren. Wir untersuchten 68 IBD-Patienten während der 5-jährigen Behandlung mit autologer hämatopoetischer Stammzelltransplantation (35 CD-Patienten (Morbus Crohn) und während des ersten Jahres der biologischen Therapie (19 CD und 14 UC (Ulcerative Colitis)), die in Spanien und Frankreich rekrutiert wurden. Kotproben und Gewebebiopsien wurden sowohl zu Studienbeginn als auch zu verschiedenen Zeitpunkten während der Nachsorge entnommen. Um Veränderungen im Darmmikrobiom und Metabolom zu charakterisieren, führten wir eine 16S rRNA-Gensequenzierung, globale 16S prognostizierte Metagenome, Shotgun Metagenomics und ungezielte Metabolomik an einer Teilmenge von Proben durch. Um die funktionellen Auswirkungen der mikrobiellen Dysbiose zu untersuchen, haben wir ein humanisiertes IBD-Mausmodell etabliert, indem wir keimfreie (GF) *Il-10^{-/-}* Mäuse mit ausgewählten Stuhlproben von CD-Patienten mit unterschiedlichem Krankheitsbild kolonisieren.

Unsere Daten zeigten, dass die zeitlichen Schwankungen der Darmmikrobiota- und Metabolitenprofile die patientenbezogenen Variationen und die Unterschiede in der Krankheitsaktivität widerspiegeln. Das fäkale Mikrobiom von Patienten mit aktiver Krankheit war mit mikrobiellen Taxa, die am Schwefelstoffwechsel beteiligt sind, wie *Escherichia Shigella* und *Fusobacterium* und einem hohen Anteil sulfatreduzierender Bakterien wie *Desulfovibrio* und *Campylobacter* angereichert. Die Profilerstellung des Fäkalstoffwechsels bestätigte eine erhöhte Häufigkeit von sulfatierten Verbindungen, einschließlich Gallensäuren, Polyphenolen und biogenen Aminen. Prognostizierte Metagenome aus 16S rRNA Daten, zeigten eine Anreicherung funktioneller Gene, die mit dem Stoffwechsel von Sulfat und Ionentransportsystemen verbunden sind, bei IBD-Patienten mit aktiver Krankheit. Im Gegensatz dazu korrelierte die zunehmende Häufigkeit mehrerer grundlegender biosynthetischer Prozesse mit einem Zustand der Remission. Die Transplantation von Mikrobiota von Patienten mit aktiver oder inaktiver Erkrankung war ausreichend reproduzierbar, um den Krankheitsphänotyp bei Empfänger-*Il-10^{-/-}* GF-Mäusen nachzubilden. Interessanterweise spiegelten humanisierte Mäuse die dysbiotischen Eigenschaften ihrer jeweiligen menschlichen Spender wider, und die Entzündung wurde durch eine Vielzahl von individuellen Gemeinschaftsprofilen angetrieben. Mit Hilfe eines maschinell lernenden Algorithmus identifizierten

wir ein Panel von 10 Taxa, welches humanisierte Mäuse nach Entzündungsstatus diskriminiert, wobei eine mikrobielle Signatur durch einen Überfluss an *Bacteroides fragilis* und *Desulfovibrio* humanisierte Mäuse nach Entzündung mit hoher Genauigkeit klassifiziert. In Übereinstimmung mit der beim Menschen identifizierten Signatur war die Anreicherung von sulfatierten Metaboliten ein Indikator für einen entzündeten Phänotyp bei Mäusen, zusammen mit einer Fülle von Genen, die auf das Sekretionssystem Typ II, IV und VI und die Glykolysepfade abgebildet werden.

Diese Arbeit hat gezeigt, dass funktionales Profiling dem taxonomischen Profiling überlegen ist, um prädiktive Signaturen in IBD zu identifizieren. Trotz der Heterogenität der Patientenproben auf taxonomischer und metabolischer Ebene rekapitulierte die Stuhltransplantation in keimfreie Empfängermäuse die Krankheitsaktivität. Wir identifizierten eine funktionelle Signatur von mikrobiellen Taxa, die am Schwefelstoffwechsel und an der Entgiftung beteiligt sind.

TABLE OF CONTENTS

1 Introduction 1

 1.1 *The healthy gut microbiome*..... 1

 1.1.1 Host-microbe mutualism in the human gut 1

 1.1.2 Taxonomic diversity of the healthy gut microbiota 2

 1.1.3 Functional redundancy and stability of the gut microbiome..... 3

 1.2 *Gut microbiota as a metabolic organ* 4

 1.3 *Microbial signatures in healthy gut microbiota* 6

 1.4. *Gut microbiota and inflammatory bowel diseases* 6

 1.4.1 Inflammatory bowel diseases (IBD) 7

 1.4.2 Microbial shifts in inflammatory bowel diseases 11

 1.5 *Longitudinal IBD cohorts for mechanistic studies of disease pathogenesis* 12

 1.6. *Mouse models as a translational tool to study human IBD*..... 14

 1.6.1 Interleukin-10 knockout mice: a mouse model for human IBD 14

 1.6.2 Gnotobiotic humanized mice: opportunities and challenges 15

 1.7. *Mining microbial signatures: structure and function* 19

 1.7.1 16S rRNA Gene Sequencing for community structure analysis 19

 1.7.2 Whole genome shotgun (WGS) sequencing for functional characterization 20

 1.7.3 Metabolomics - linking microbial diversity with metabolic activity..... 20

 1.7.4 Machine learning: recognizing patterns and connecting the dots..... 21

2. Aims and scope of dissertation..... 22

3. Materials and methods 24

 3.1 *Human study cohorts - cohort description and data collection* 24

 3.1.1 Hematopoietic stem cell transplantation (HSCT) cohort..... 24

 3.1.2 REMIND post-surgical cohort 26

 3.1.3 Biotherapy Cohort 28

 3.2 *Processing of human fecal samples* 30

 3.3 *Studies in gnotobiotic mice* 30

 3.3.1 Housing conditions..... 30

 3.3.2 Colonization of GF mice with human microbiota 30

 3.3.3 Bacterial enumeration in mouse fecal samples 31

 3.4 *Characterization of bacterial community's structure and function* 31

 3.4.1 Metagenomic DNA extraction from fecal samples..... 31

 3.4.2 Metagenomic DNA extraction from mucosa-associated bacteria (tissue biopsies)..... 31

 3.4.3 High-throughput 16S ribosomal RNA (rRNA) gene sequencing and microbiome profiling 32

 3.4.4. Whole-genome shotgun Sequencing..... 32

 3.4.5 Metabolomics 34

 3.5 *Phenotypic characterization of inflammation in humanized mice* 36

 3.5.1. *Histological scoring* 36

 3.5.2. Gene expression analysis 36

 3.6 *Statistical analysis*..... 37

4. Results	38
4.1 <i>Characterization of fecal microbial signatures in a cohort of adult CD patients</i>	38
4.1.1 HSCT- study cohort and patients' subgroups	38
4.1.2 IBD patients show temporal fluctuations and interpersonal variations through the course of disease	39
4.1.3 Stratification of IBD patients by disease activity shows separation of microbial profiles, with a substantial overlap between subgroups	42
4.1.4 Further stratification supported the small but significant differences according to disease activity	44
4.1.5 Effect of co-administration of biological and immunosuppressive therapy on shifts of gut microbiota in IBD patients post- HSCT	46
4.1.6 Microbial community diversity and richness at baseline could not predict response to HSCT therapy	48
4.1.7 Dynamic fluctuations of microbial community composition overtime in HSCT-treated CD patients	50
4.1.8 CD patients cluster into enterotypes characterized by the dominance of signature bacterial taxa	51
4.1.9 Unsupervised clustering emphasized the heterogeneity of gut microbiota structure in CD patients	51
4.2. <i>Mucosal-associated dysbiosis in CD patients treated with HSCT</i>	53
4.2.1. Characterization of colonic mucosa-associated dysbiosis	53
4.2.2. Characterization of ileal mucosa-associated dysbiosis	57
4.3 <i>Selection of patients for transfer experiments in germfree mice</i>	60
4.4 <i>Functional characterization of microbial signatures in IBD using gnotobiotic humanized mice</i>	62
4.4.1 Morphometric and immunologic changes in humanized mice	63
4.4.2. Humanized mice reflect the dysbiotic features of their respective donors	64
4.4.3 Alteration of bile acid metabolism in mice associated with dysbiotic human microbiota	65
4.4.4. Predicted metabolic functions of the disease-relevant microbial communities	67
4.4.5 Human Donor as well as humanized mice show individual functional signatures under inflammatory conditions	70
4.4.6 Untargeted metabolomics approach to characterize the changes in metabolite profiles of IBD patients with active or inactive disease post-HSCT	72
4.4.7 Significant enrichment of sulfated compounds is associated with inflammation in humanized mice	73
4.4.8 Integrative multi-omics signature improves the predictive modelling of disease outcome	74
4.5. <i>Validation of humanized mice in IBD research</i>	75
4.5.1 Transfer efficiency of human microbiota into germ-free mice.....	75
4.5.2 Established host-selected microbiota is efficiently transmissible to offspring	78
4.6 <i>Characterization of fecal microbial signatures in a cohort of biotherapy-treated adult CD and UC patients</i>	79
4.6.1 Crohn's disease and ulcerative colitis patients show unique fecal microbiota profiles.....	79
4.6.2 IBD patients show individualized response to biological therapy	80
4.6.3 Baseline composition and indication for response to biotherapy.....	81
4.6.4 Microbial shifts in response to biotherapy treatment	81
4.7 <i>Characterization of fecal microbial signatures in a cohort of post-operative recurrence cohort</i>	83
5. Discussion	85
5.1 <i>Identification of microbial signatures associated with response to HSCT therapy</i>	85
5.2 <i>Mucosa-associated microbiota in HSCT-treated CD patients</i>	88

5.3 *A Gnotobiotic mouse model to study the functional impact of microbial dysbiosis in IBD* 90

5.4 *Multi-omics approach unravels functional alterations of the gut microbiome* 92

Abbreviations **ciii**

References **cvi**

Acknowledgments **cxxi**

Publications and Presentations **cxxiii**

List of cited publications **cxxvii**

Curriculum Vitae **cxxviii**

1 Introduction

1.1 The healthy gut microbiome

1.1.1 Host-microbe mutualism in the human gut

The human gastrointestinal (GI) tract is colonized by a complex and dynamic array of microorganisms collectively termed as the gut microbiota. Gut microbiota comprises bacteria, archaea, bacteriophages, viruses, and fungi (Lozupone et al., 2013). Bacteria stands as the most dominant and diverse group of microorganisms in the human gut, with an outstanding number of unique species, around 1000 species (Herington et al., 2012). The number of bacteria gradually increases along the small intestine, ranging from 10^4 colony-forming units (CFU) per gram of luminal content in the jejunum to 10^7 CFU in the ileum. In the colon, anaerobic bacteria are predominant and bacterial total counts reach around 10^{12} CFU per gram of luminal content (Backhed et al., 2005) (Figure 1).

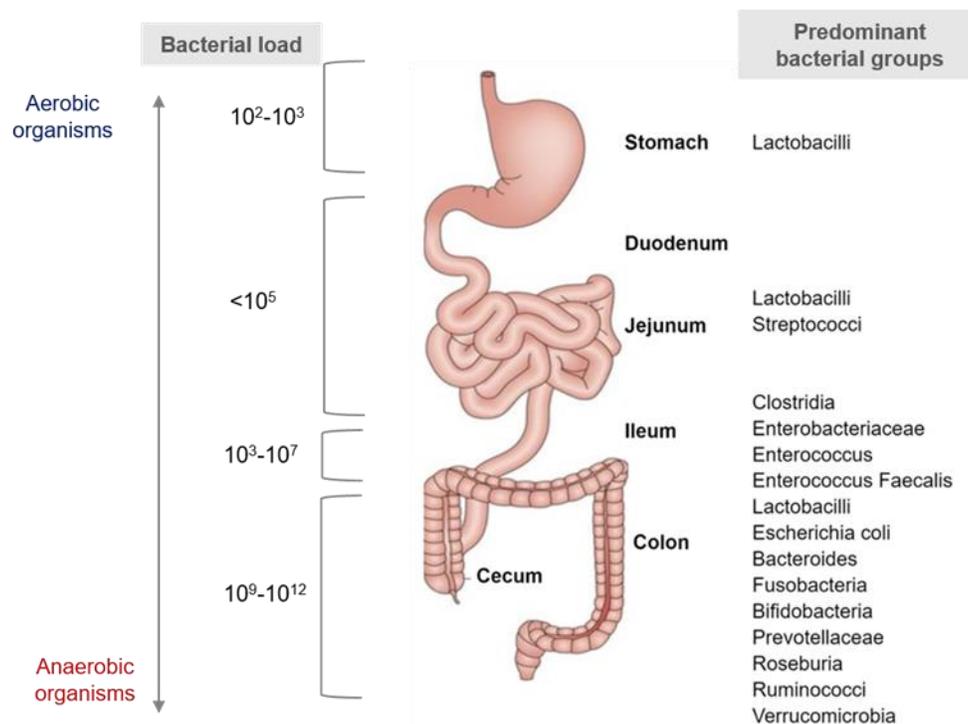


Figure 1 The Human Gastrointestinal Tract

A gradient of oxygen levels, bacterial load, and specific predominant bacterial groups exist along the different compartments. Acidic conditions inferred by the entry of bile acid and digestive enzymes characterize the proximal small intestine. The predominant bacterial groups colonizing the small intestine and colon reflect physiological differences along the length of the gut.

Over the past six million years, humans and microbes have co-evolved. Through this co-evolution, they have developed a mutual relationship. Gut microbes play an indispensable role in the maintenance of the gut homeostasis through various physiological functions such as maintaining the gut integrity (Natividad and Verdu, 2013), regulating host mucosal immunity maturation and development (Gensollen et al., 2016), producing vitamins and extracting energy and nutrients from the diet (den Besten et al., 2013). Additionally, they defend the host against pathogens through the production of mucus and the release of antimicrobial peptides.

1.1.2 Taxonomic diversity of the healthy gut microbiota

The formation of the gut microbiome starts during birth with exposure to bacterial communities of the mother upon passage through the birth canal, or of the mother's skin if delivered via cesarean section. Overtime, the diversity of the gut microbes expands tremendously during adolescence to reach an inconceivable number of microorganisms – up to 100 trillion bacterial cells (Bäckhed et al., 2015). Over the last decades, the vast advancements in culture-independent techniques and anaerobic culturing tools led to huge progress in the identification of gut bacterial species. An estimated number of over 1000 human gut bacterial species have been identified. The healthy gut microbiome is dominated by two major phyla – Firmicutes and Bacteroidetes (constituting 95% of the adult gut communities). Proteobacteria, Actinobacteria and Verrucomicrobia generally do not exceed 3% of the composition in healthy adults (Segata et al., 2012). Despite the temporal and inter-individual variation of gut microbiota structure, the community composition of the human gut is quite stable, with the dominant phyla conserved in healthy adults. Large-scale longitudinal studies have shown that samples collected from the one subject overtime are more similar to one another compared to those collected from other subjects, suggesting that each individual has his own relatively stable and unique gut microbial community (Caporaso et al., 2011; Gootenberg and Turnbaugh, 2011; Jakobsson et al., 2010; Turnbaugh et al., 2009) . In this context, large-scale population studies based on dense collection of longitudinal sampling is essential to understand the factors contributing to the stability of core functional microbiome. To better understand the stability of microbiome structure and function overtime, Mehta and colleagues characterized the fecal microbiome collected from 308 healthy adult men (Mehta et al., 2018). Four samples were collected from each donor, 2 were collected 48-72 hours apart, and the remaining two were collected 6-months later to assess the dynamics of gut microbiome structure overtime. Analysis of metagenomics and taxonomic profiles in this cohort confirmed that samples collected from the same person showed lower variation than samples collected from different persons. Interestingly, metatranscriptomes profiles showed comparable variation between and within individuals. From microbial ecology

perspective, this stability is explained by the ability of a healthy microbiota to resist perturbations which could possibly result from a pathogenic infection, change of dietary habits or prolonged use of medications, and to revert it back to a healthy state afterwards. These features of a healthy microbiota have been termed *resistance* and *resilience*, respectively (Lozupone et al., 2012).

1.1.3 Functional redundancy and stability of the gut microbiome

Functional fingerprinting by shotgun metagenomics allows high-resolution genomic analysis of the total microbial community DNA, including that of cultured and uncultured members and linking them to known functional genes. Over the past decade, large-scale projects, such as the Metagenomes of the Human Intestinal Tract (MetaHIT) and the Human Microbiome Project (HMP) intended for characterizing the phylogenetic diversity and the functional potential of the human intestinal microbiome and their relevance to health and disease. In MetaHIT project, metagenomics shotgun sequencing was performed on fecal samples from a cohort of 124 European adults (Qin et al., 2010). Two years later, the HMP published the results of 16S rRNA gene sequencing of microbial samples collected from different body habitats of 242 American healthy adults, as well as the results of metagenomics sequencing of a subset of 139 subjects (Consortium, 2013). Collectively, these studies have shown that while taxonomic composition varies largely between individuals at all body sites, the abundance of metabolic pathways or a functional core microbiome is relatively stable and consistent in healthy humans (Consortium, 2013; Qin et al., 2010; Turnbaugh and Gordon, 2009; Turnbaugh et al., 2009). This functional core involves a combination of microbial genes responsible for encoding metabolic, regulatory and physiological functions that are necessary to maintain gut homeostasis. These functions could be performed by functionally-redundant or exchangeable organisms in different individuals (Martiny et al., 2015; Xu et al., 2014). Functional redundancy is an established concept in ecology which explains how different species can perform similar functions needed to maintain homeostasis of the ecosystem (Lozupone et al., 2012; Mahowald et al., 2009). Examples of functional redundancy in the human gut microbiome exist in dietary carbohydrates processing (Ferrer et al., 2013), bile acids metabolism (Labbe et al., 2014), and the promotion of interleukin secretion (Morgan et al., 2015; Sokol et al., 2008). This “functional redundancy”, together with the “resistance” and “resilience” define the key features of a healthy gut microbiome (Bäckhed et al., 2012).

Recently, the development of high-fidelity, high throughput techniques, including metatranscriptomic, metaproteomic, and metabolomic analyses allowed deep characterization of the functional capacities and functional activity of gut microbiome. Data generated using functional omes showed greater sensitivity to perturbations compared to the taxonomic or metagenomics analyses,

emphasizing that taxonomic analysis alone is not sufficient to capture the changes microbiome undergo in IBD. A major contribution to this area was performed by, the Inflammatory Bowel Disease Multi'omics Database (IBDMDB) project. A number of 1,785 stool samples and 651 intestinal biopsies were collected from 132 individuals overtime for one year.(Lloyd-Price et al., 2019). Integrated multi-omics data showed that disease activity correlated with increased temporal variation, taxonomic, functional, and biochemical shifts. The study finally identified an interacting set of microbial, chemical and host factors that contribute to disease. Further, by comparing baseline samples to samples collected during inactive state, the study defined a dysbiosis index proposing a set of features common to inflammatory response in patients.

1.2 Gut microbiota as a metabolic organ

Gut microbiota plays a central role in regulating and maintaining various functions of host physiology. Through products of microbial metabolism, gut microbiota can communicate with cells throughout the body and regulate their activity. Many of these microbial metabolites, such as short-chain fatty acids and bile acids are taken up from the gut lumen to blood stream where they are transported to body organs and act as signaling molecules (**Figure 2**). The fundamental role of gut microbiota in regulating host metabolic functions was best studied by examining the phenotype of germ-free mice (GF). In this aspect, a number of studies investigated the role of microbiota in obesity, where the influence of microbiota on energy uptake and expenditure through regulation of metabolic pathways was established (Ding et al., 2004; Manchester et al., 2007). In this work, we focused on bile acids due to their connection to gut microbiota and IBD.

Bile acids are endogenous molecules that are synthesized from cholesterol in the liver. The primary bile acids produced in humans are cholic acid (CA) and chenodeoxycholic acid (CDCA), while in mice, the primary bile acids are CA and murocholic acids (MCAs). Primary bile acids get immediately conjugated predominantly to the amino acids glycine or to a lesser extent taurine (in humans) and exclusively to taurine (in mice) upon synthesis in the liver (Russell, 2003). Conjugated bile acids are transported then to the gallbladder, together with bilirubin, phospholipids and cholesterol (Trauner and Boyer, 2018). After food ingestion, the bile acids are secreted to the duodenum, where they are “deconjugated” through metabolizing reactions performed by gut microbiota. As they pass through the gastrointestinal tract, primary bile acids are further chemically modified by microbial enzymatic reactions in the colon into secondary bile acids. The major secondary bile acids are deoxycholic acid (DCA) and lithocholic acid (LCA) in humans, in addition to alpha, beta and omega-muricholic acid (α , β and ω MCA) in mice. Enzymes required for these chemical modifications are expressed by gut microbes. For example, members belonging to genera *Lactobacillus* and *Streptococcus* are known

to express bile salt hydrolase (BSH), and enzyme required for the initial step of de-conjugation. While members of *Clostridium* species (cluster XI and XVI) are known to express 7 α -dehydroxylation enzyme, that is for example required for the synthesis of essential secondary bile acids DCA and LCA (Ridlon et al., 2006). Investigation and annotation of these microbial modifications is an area of great interest, where links between gut microbial dysbiosis and disrupted metabolic functions were established in diseases, such as obesity and inflammatory bowel disease.

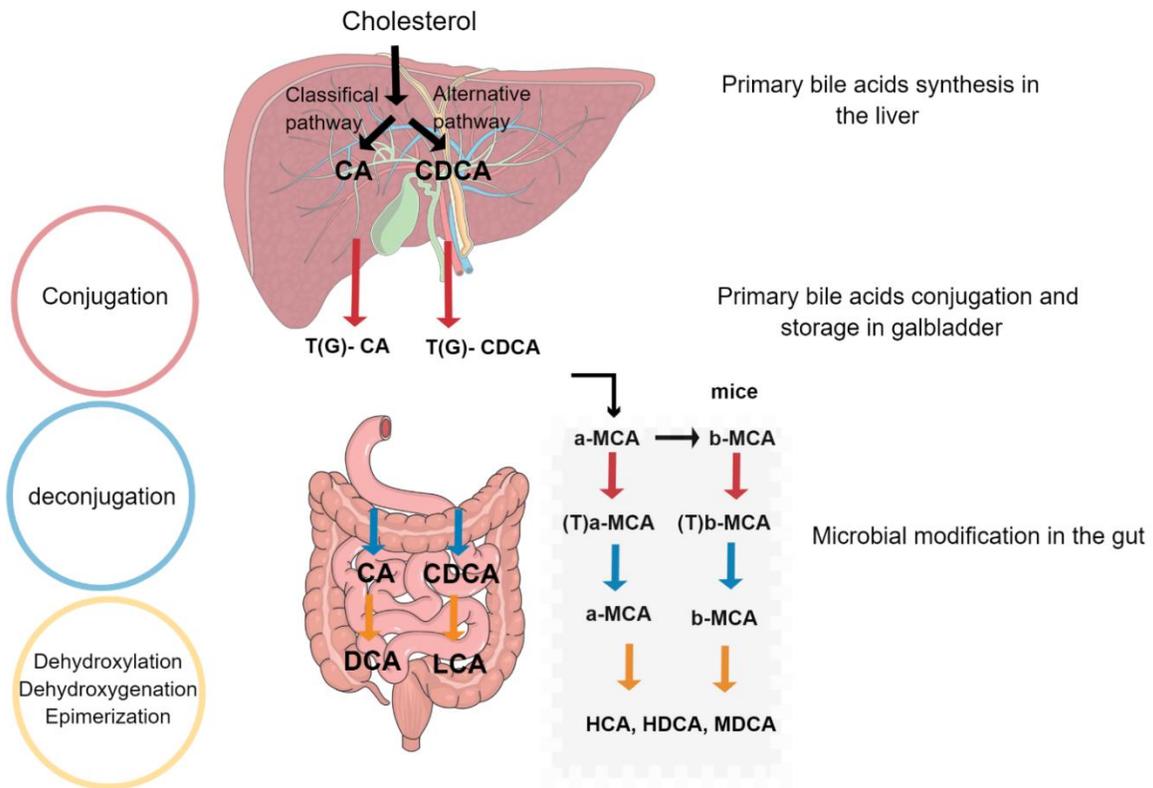


Figure 2 Bile acid metabolism by gut microbiota.

1.3 Microbial signatures in healthy gut microbiota

Identification of microbial signatures associated with a healthy status has been of great interest in clinical settings. In 2011, Arumugam and colleagues have analyzed 22 fecal shotgun metagenomes of individuals from 4 countries and showed that individuals' gut microbiota can be classified into "enterotypes" based on the abundance of signature bacterial taxa. These enterotypes are the *Bacteroides* type, the *Prevotella* type and the *Ruminococcus* type (Arumugam et al., 2011). Yet, in a following study, Wu and colleagues analyzed the microbial communities in fecal samples from 98 individuals who were undergoing dietary interventions. The microbial communities from fecal samples clustered into two enterotypes, *Bacteroides* and *Prevotella*. The enterotypes strongly correlated with long-term dietary habits, where a diet rich in protein and animal fat correlated strong with a *Bacteroides* enterotype, whereas a carbohydrates-rich diet correlated with *Prevotella*-dominated enterotype. The stability of community composition was changed 24 hours after initiating a high-fat/low-fiber or low-fat/high-fiber diet, but the enterotype class remained stable during the 10-day study, suggesting that the enterotypes clusters are associated with long-term dietary habits (Wu et al., 2012). Microbiota profiling of healthy adults from the US (Wu et al., 2012), healthy infants, children, and adults from Malawi, Venezuelan Amerindians, and the US (Yatsunencko et al., 2012), and of 250 healthy adults from the US (Consortium, 2013), failed to reproduce the same stratification, suggesting that despite a possible association of disease activity with enterotype cluster, they are not always reproducible and significant. In 2014, Dan Knights and colleagues performed a meta-analysis including both single-individual longitudinal samples and cross-sectional samples. In this work, they showed that a healthy gut microbiome can span much of the total variation of healthy human microbiome over a one-year time period, suggesting that enterotypes are continuous rather than discrete (Knights et al., 2014). Most recently, the same group revisited the notion of enterotypes and performed a refined meta-analysis on gut microbiota metagenomics datasets from three large projects (HMP, MetaHIT and Chinese II diabetes study) (Consortium, 2013; Qin et al., 2012). They concluded that enterotypes may be still relevant in clinical settings by stratifying disease associations, treatments or dietary interventions (Costea et al., 2018).

1.4. Gut microbiota and inflammatory bowel diseases

The imbalance of gut microbiota composition rather than specific putative pathogens can result in host-microbial mutual relationship disruption leading to the pathogenesis of several diseases, for example IBD. This compositional shift is termed as dysbiosis (Chamaillard and Radulovic, 2016).

1.4.1 Inflammatory bowel diseases (IBD)

1.4.1.1 Definition and clinical features

Inflammatory bowel diseases (IBD) describes a group of chronic and recurrent inflammatory conditions that affect the gastrointestinal tract and comprise Crohn's disease (CD) and ulcerative colitis (UC). While CD is characterized by transmural inflammation, UC is limited to the mucosal layer of the colon. The transmural phenotype of CD may additionally involve fibrosis and strictures, which are not common in UC. CD may affect any part along the entire intestinal tract; however, the most commonly affected area is the ileum and proximal colon. On the other hand, UC involves the rectum and could continue to involve parts of the colon (Silverberg et al., 2005).

1.4.1.2 Epidemiology, incidence and risk factors

Over the past decades, IBD became a global emerging disease with an increasing incidence and prevalence in most countries around the world. In Europe, the incidence rate reached 24.3 per 100,000 person-years for UC and 12.7 per 100,000 person-years for CD. In north America the incidence rates reached 19.2 cases per 100,000 person-years for UC and 20.2 for CD (Molodecky et al., 2012) . These numbers are much lower in Asia and the Middle east, however in some newly industrialized countries in Africa, Asia and south America, the incidence of IBD has been arising (Ng, 2017). While UC and CD are more common in individuals with Jewish ancestry and in Caucasians, it is less prevalent in Afro-Americans and Hispanic populations (Acheson, 1960; Roth et al., 1989; Sonnenberg et al., 1991). IBD is often diagnosed between the ages of 15 and 40 years, however some studies suggest a bimodal age distribution with a second peak between 50 and 80 years (Andres Ekbohm, Charles Helmck and Hans-Olov, 1991). Some reports have shown a slight difference in IBD incidence by gender with a female predominance in CD and male predominance in UC (Munkholm P1, Langholz E, Nielsen OH, Kreiner S, 1992; Shivashankar, 2017). Interestingly, immigration studies have shown an increased risk of developing IBD in individuals migrating from countries with low incidence rates to countries with higher incidence, particularly in the first-generation children (Bernstein and Shanahan, 2008a; Mikhailov and Furner, 2009), suggesting the importance of environmental risk factors in IBD pathogenesis.

Risk factors include genetic susceptibility, where a number of studies confirmed a high degree of familial aggregation and disease concordance in CD (Peeters et al., 1996). Along the same line, several large-scale population studies have identified a number of risk loci (Anderson et al., 2011; Ellinghaus et al., 2016; Julià et al., 2014; Kenny et al., 2012; Liu et al., 2015; Yang et al., 2014), with the most recent analysis describing 241 susceptibility loci to IBD (Katrina M de Lange, 2017).

However, the fact that genetic factors can explain only a small proportion of the disease pathogenesis indicates that gut microbiota and environmental risk factors interact with genetic susceptibility and lead to disease pathogenesis. Smoking, hygiene hypothesis, diet and stress are some of these factors that have been implicated in IBD incidence (BM, 1989). The hygiene hypothesis proposes that the lack of exposure to enteric pathogen along with improved hygienic conditions in early life led to the rise of immunologic disorders (Bernstein and Shanahan, 2008b; Shanahan and Bernstein, 2009). Children who are less exposed to enteric organisms are at a greater risk of developing exaggerated immune response to new antigens later in life (Gent et al., 1994). A number of hygiene-related factors including *Helicobacter pylori* infection, family size, urban upbringing, domestic hygiene and exposure to pets or farm animals have been addressed in IBD studies (Bernstein et al., 2006; Koloski et al., 2008; Lashner and Loftus, 2006). Additionally, stress factors can alter immune function (Söderholm et al., 2002) and has been suggested to be involved in the pathogenesis of IBD. Studies in animal models demonstrated that chronic stress can aggravate IBD by stimulating intestinal mucosa damage and impairing intestinal barrier function (Bernstein et al., 2006; Söderholm et al., 2002). Moreover, there has been a correlation between westernized diet (high fats, low fibers) and increased risk of IBD (Amre et al., 2007). This put together, IBD is believed to involve exaggerated response of the host immune system to the gut commensal microbiota in genetically predisposed hosts, highlighting the critical role of gut microbiota dysfunction in disease etiology.

1.4.1.3 Therapeutic options for IBD treatment

IBD is characterized by episodes of remission and relapse, where a subclinical inflammation persists leading to progressive tissue damage overtime. Thus, the main goals in treating IBD are to relieve symptoms, suppress inflammation and to achieve remission and endoscopic healing of mucosal lesions (Rogler et al., 2013), while improving the patient's quality of life (Hommes et al., 2012). Traditional treatment regimen involves a sequential "step-up" approach, where the choice of the treatment intensifies as the disease becomes more severe. Currently, there are five basic categories of medications used, namely anti-inflammatory drugs, immune-suppressants, biologics, antibiotics, and drugs for symptomatic relief (Triantafillidis et al., 2011). Clinicians typically initiate the treatment by 5-aminosalicylic acid (5-ASA) agents and antibiotics. This can be followed by corticosteroids, immune-modulators, and biologics. However, surgery is still indicated for approximately 40% of patients with CD and 30% of patients with UC at some point during the course of disease (Peyrin-Biroulet and Lémann, 2011). Some patients remain refractory to all current therapies and also not eligible for surgery due to the extensive disease that involves a large portion of the small intestine

(Hwang, 2008). Here, stem-cell therapy stands as a potential alternative to induce remission, possibly by erasing immune responses against microbes (Heslop et al., 2015).

However, in the recent years, treatment goals shifted from clinical (symptomatic) remission to endoscopic mucosal healing, and prevention of irreversible tissue damage. Besides, several reports have shown that a “top-down” approach, where more clinically-effective drugs (biologics and immunosuppressant) are initiated earlier, can alter the course of disease progression, and prevent irreversible complications (Hommes et al., 2012; Im et al., 2018). That said, in the era of personalized medicine, an individualized treatment approach based on tight and timely monitoring is believed to yield better outcomes and help clinicians decide which patients would benefit from an early use of more aggressive drugs. In the following section, treatment options discussed in this work are reviewed.

Hematopoietic stem cell transplantation (HSCT)

In the 1950s, HSCT was developed to treat the bone marrow failure that occurred in leukemia patients after being exposed to whole body radiation (Li and Sykes, 2012; Majhail et al., 2015). Two forms of this procedure were developed (allogenic or autologous HSCT). In allogenic HSCT, the patient’s bone marrow is eradicated and is replaced with stem cells derived from a suitable donor. On the contrary, autologous HSCT involves the replacement of the patient’s bone marrow with his own stem cells that are cryopreserved and subsequently used for transplantation. The procedure of HSCT involves “mobilization” of stem cells by a combination of chemotherapy and growth factors leading to the release of the stem cells from stem cell compartment into the blood stream, and consequently collected for storage in a process called “leukapheresis”. The recipient is prepared for transplantation in a process known as “conditioning”, where the bone marrow is ablated by total body irradiation and chemotherapy (Snowden et al., 2012). Following conditioning, the stem cells are infused back into the blood stream for further establishment and reconstitution of stem cell niche **(Figure 3)**.

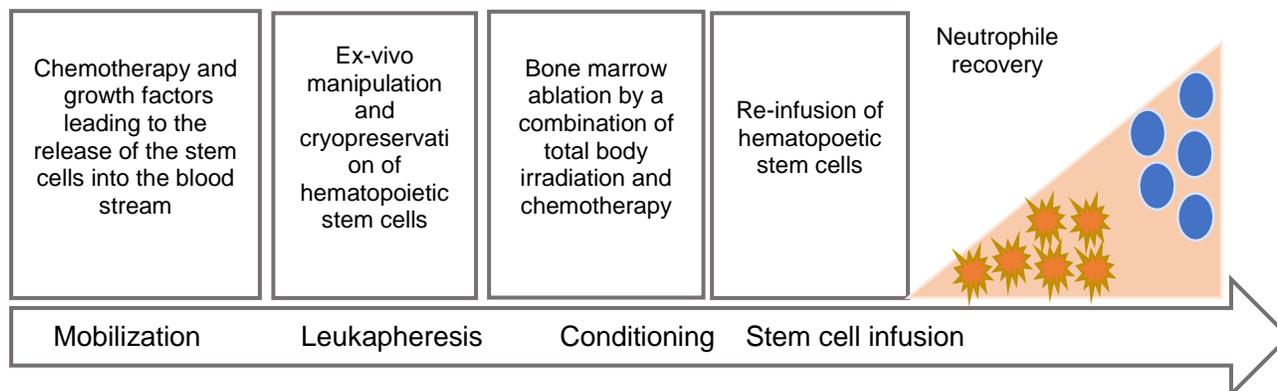


Figure 3 The process of autologous Hematopoietic stem cell transplantation

Allogenic HSCT proved to be effective in idiopathic adult IBD, with the possibility of replacing diseased hematopoietic cells with an HLA-matched, yet a genetically different one originating from a healthy donor. In the first clinical trial of this kind, 5 out of 7 patients improved after allogenic HSCT, of whom 3 patients maintained a long-term remission for 6-10 years (Lopez-Cubero et al., 1998). However, allogenic HSCT is high-risk procedure that is associated with 20% mortality rate (Keating et al., 2013) and therefore its use is reserved to patients with known monogenic diseases; for example patients with mutations in interleukin-10 [IL-10] or IL-10 receptor [IL-10R]) (Glocker and Kotlarz, 2009). Autologous HSCT (AHSCT) was therefore developed to avoid graft versus host disease and the high mortality rates of allogeneic transplantation. The mechanism of action is not yet clear, however it is believed that AHSCT works through the elimination of activated lymphocytes and the induction of *de novo* regeneration of the T-cell repertoire (Hawkey et al., 2015; Muraro et al., 2014). Since then, autologous HSCT has been used for the treatment of chronic refractory CD patients. The largest clinical study included 24 CD patients, 18 of whom were followed up for 5 years post the transplant (Burt et al., 2011). After the transplant, all patients achieved clinical remission, however most patients relapsed during the 5-years follow-up period. More recently, an international multi-centric trial (ASTIC) was carried out to compare the efficacy of AHSCT and conventional therapy in the treatment of refractory CD (Hawkey et al., 2015). The Autologous Stem Cell Transplantation International Crohn's Disease (ASTIC) trial is an open label, phase III trial that involved 11 transplant units in 7 countries. A total of 45 patients (age: 18-50) were recruited in this trial. AHSCT was indicated for these patients since they were refractory to all conventional therapies and were not amendable to surgery. The first results showed that more CD patients were able to stop all immunosuppressive therapy and maintained drug-free clinical remission compared to control group. Further assessment showed that the transplanted patients were still free of disease upon endoscopic examination during the first year-post HSCT.

Biological therapy

Biologics refer to natural or biologically modified compounds that are used for the treatment of diseases. Tumor necrosis factor alpha (TNF-alpha) is known to act as a pro-inflammatory cytokine that promotes CD pathogenesis. While non-biological conventional therapies could provide symptomatic relief to IBD patients, anti-TNF agents could considerably change the disease course, improve patients' quality of life and achieve mucosal healing (Rutgeerts et al., 2005). It may take up to 8 weeks after treatment initiation, however some patients could experience more immediate improvement. Examples of anti-TNF medications include: Adalimumab (Humira®), Certolizumab pegol (Cimzia®), Golimumab (Simponi®), Infliximab (Remicade®), Infliximab-dyyb (Inflectra®).

1.4.2 Microbial shifts in inflammatory bowel diseases

Cross-sectional studies on CD and UC patients showed significant alteration of microbial community's structure throughout the course of disease. The most consistent changes reported included an overall reduction of microbial community richness and diversity, accompanied by an inversed Firmicutes/Bacteroidetes ration and a remarkable increase in Gammaproteobacteria (Sokol and Seksik, 2010; Gevers *et al.*, 2014). The relative abundance of certain bacterial groups that are known to be beneficial was found to be reduced in IBD patients. For example, reduction of *Faecalibacterium prausnitzii*, *Roseburia hominis*, and *Bifidobacterium adolescentis* were consistently reported in IBD patients (Joossens et al., 2011; Maubert et al., 2014; Sokol et al., 2008). On the other hand, an increase of potential pathobionts of the phylum *Proteobacteria* was described in IBD patients (N.Barnichab, J.Denizotab, 2013). Interestingly, in a recent study assessment of strain-level diversity showed an increased strain diversity of potentially pathogenic species and, conversely reduced strain diversity in bacterial species known to be beneficial in stool samples from patients with IBD or IBS compared with controls (Vila et al., 2019). These changes eventually perturbates the host-microbial mutualism leading to functional dysbiosis and predisposition to IBD (Ianiro and Hansen, 2018). This imbalance (dysbiosis) of gut microbiota has been described in several cross-sectional studies which investigated patients' cohorts with already established disease. It is thus unclear whether dysbiosis is a crucial driving event or is simply a consequence of inflammation. Studies on early-onset or newly diagnosed patients could give us more insights into the relevance of dysbiotic intestinal microbiota in disease initiation. The first study of this kind was the "Bacteria in Inflammatory bowel disease in Scottish Children Undergoing Investigation before Treatment" (BISCUIT) study. This study investigated the changes of mucosa-associated bacterial composition in newly diagnosed children (13 CD and 12 UC patients). Recruited patients

were not affected by additional co-morbidities and were treatment naïve to IBD (have not received systematic antibiotics or steroids at least 3 months before sample collection), which eliminated a major confounding factor for microbial changes (Hansen et al., 2013). CD patients showed to have significantly reduced diversity compared to UC patients and healthy controls. In contrast to previous reports, *F. prausnitzii* was significantly more abundant in CD patients (Fujimoto et al., 2013; Gevers et al., 2014; Lopez-siles et al., 2012; Sokol et al., 2008).

Later in 2014, Gevers and colleagues published results from the largest cohort microbiome study on new-onset IBD, where they studied the changes in microbial composition in a multicenter cohort of new-onset CD patients, and including subjects of various disease phenotypes, localization and severity. They demonstrated that there is a considerable correlation between the inflammatory status and increased levels of *Enterobacteriaceae*, *Pasteurellaceae*, *Veillonellaceae*, and *Fusobacteriaceae*, and on the contrary with decreased levels of Erysipelotrichales, Bacteroidales, and Clostridiales. Interestingly, stratifying patients by antibiotic use, showed that antibiotics administration increases the microbial imbalance associated with CD. Additionally, comparing the microbial signatures between gut contents from ileum, rectum and fecal samples, showed that the rectal mucosa-associated microbiome gives the strongest correlation with disease outcome and stands as a better option for use to improve early diagnosis of CD (Gevers et al., 2014).

Taken together, these studies highlight the importance of confounding factors interfering with studying human cohorts with established disease. Following up IBD patients overtime in well-established longitudinal cohort provides an alternative to circumvent deviations arising from concurrent therapies, such as antibiotics and surgery (Halfvarson et al., 2017).

1.5 Longitudinal IBD cohorts for mechanistic studies of disease pathogenesis

The clinical course of CD can vary widely from mild to severe, making it difficult to predict disease progression based on the patient's status at initial diagnosis. Identifying factors that could influence the course of disease would help to improve prognosis and help clinicians make early decisions on follow-up and treatment options (INGER CAMILLA SOLBERG, MORTEN H. VATN, OLE HØIE, NJAAL STRAY, JOSTEIN SAUAR, JØRGEN JAHNSEN, BJØRN MOUM, IDAR LYGREN, 2007). In the microbiota field, prospective longitudinal cohorts are the best means to capture the changes that precede or follow disease onset, and to link these compositional shifts with mechanisms of disease pathogenesis. Few reports addressed this question on rather smaller cohorts (Martinez C, Antolin M, Santos J, Torrejon A, Casellas F, Borrueal N, Guarner F, 2008; Wills et al., 2014; Young et al., 2013), until Halfvarson and colleagues published their work on a prospective cohort of 128 IBD patients and 9 healthy controls in 2017 (Halfvarson et al., 2017). In this study, they collected fecal

samples every 3 months for a period of two years and looked into the dynamics of microbial changes overtime in health and disease. They demonstrated that microbiota of IBD patients fluctuates more dramatically than healthy subjects, based on deviation from a baseline they identified and termed as “healthy plane”. The distance to the “healthy plane” varied overtime in IBD patients but did not necessarily correlate with disease activity. However, a correlation was found between microbial shift and administration of drugs during treatment. These findings highlight the importance of potential confounding factors that could affect the reproducibility of results from cross-sectional studies. In-depth analysis of these temporal bacterial community dynamics in health and disease could help to identify “microbial signatures” as potential biomarkers and therapeutic targets.

In 2017, Pascal and colleagues characterized a longitudinal cohort of 178 participants from four European countries; comprising 40 unrelated healthy controls, 34 CD patients, 33 UC patients and 36 and 35 healthy controls related to patients with CD and UC, respectively (Pascal et al., 2017). Microbial dysbiosis, as measured by community richness, diversity and stability was more significant in CD compared to UC patients. Based on the 16S microbiota profiling, they proposed a microbial signature for CD based on the presence/absence or differentially-abundance of 8 bacterial taxa. Using these features, they built a model for prediction and validated the results against an internal test-set as well as external cohorts. When tested against all cohorts, the signature achieved high accuracy for the classification of CD patients versus healthy controls, patients with anorexia, IBS and UC (Pascal et al., 2017).

To determine whether gut microbiome may predict responses to IBD therapy, Ananthakrishnan and colleagues designed a prospective longitudinal study on anti-integrin treated IBD patients. Fecal samples were collected at baseline, 14, 30 and 54 weeks after initiation of therapy. Interestingly, patients achieving remission 14 weeks after treatment appeared to have significantly higher community richness and diversity, and an increased relative abundance of Burkholderiales and *Roseburia inulinivorans* at baseline. At the functional level, they identified several associations with microbial function including branched chain amino acid synthesis pathways that were enriched at baseline in patients achieving remission after therapy. In conclusion, the authors identified 2 taxa and 17 pathways that were predictive for the response to therapy, however, a specific mechanistic link to disease pathogenesis was not established and clinically relevant changes in metabolite concentrations were not confirmed (Ananthakrishnan, 2017).

To address the question of dynamic fluctuations of in microbiota compositions according to disease activity, Yilmaz and colleagues investigated the correlation of microbiota composition to disease activity based on clinical parameters or fecal calprotectin levels. Interestingly, very few taxa showed

similar trends among patients and data showed a huge inter-individual variations among patients, where microbial profiles showed to cluster based on individual-specific signature rather than disease activity status (Yilmaz et al., 2019). This put together, a prospective longitudinal cohort design is required to dissect changes preceding or coinciding with changes of disease course. This will lead to a transition from correlation to causation and give insights into the mechanistic role of microbiota, with an ultimate goal of improved early diagnosis and prediction of response to therapy. Now with the emerging multi-omics technologies, the integration of microbiome, metagenomics, metatranscriptomics and metabolite data with clinical data would provide a comprehensive approach to investigate and understand the mechanisms of IBD pathogenesis.

1.6. Mouse models as a translational tool to study human IBD

To address causality in disease pathogenesis, mouse models stand as an excellent tool to disentangle complex host-microbe interactions and to allow translation of complex human diseases. Numerous mouse models have been established to study IBD. These include chemical-induced colitis models, adoptive transfer models, models of spontaneous colitis, genetically engineered or transgenic mouse models, and lastly gnotobiotic humanized mice (Becker et al., 2015). While no single model reflects the complexity of human IBD, each model provides insights into one aspect of the disease. Here, genetically modified and gnotobiotic mouse models used in the context of the present work are reviewed.

1.6.1 Interleukin-10 knockout mice: a mouse model for human IBD

This model is one of the earliest mouse models for intestinal inflammation and remains to be one of the most commonly used models to study human IBD. The IL-10 knockout mouse model was generated in 1993 by Kuehn et al., by integrating a targeted mutation disrupting the IL-10 gene by replacing a 500 base-pair fragment of exon 1 with a termination codon and neo expression cassette and by introduction of termination codon into exon 3 (*Il10^{tm1Cgn}, IL10^{-/-}*) (Kühn et al., 1993). The clinical relevance of this models is based on the identification of a link between genetic polymorphisms at the IL-10 locus and an increased risk of both UC and CD (Franke et al., 2008, 2010). Additionally, a familial form of early onset CD has been linked to homozygous mutation in IL-10 receptor (Glocker et al., 2009). IL-10 knockout mice develop spontaneous inflammation of the colon, when housed under specific pathogen-free conditions (SPF). There, the inflammatory phenotype reflects the histological findings of human IBD (Bleich et al., 2004; Kühn et al., 1993). This model is characterized by a discontinuous patchy inflammation phenotype accompanied by immune cell infiltration in the lamina propria and submucosa, together with epithelia hyperplasia,

crypt abscesses, ulcers, mucin depletion and thickening of the intestinal wall (Berg et al., 1996). In 1998, Sellon and colleagues showed that IL-10 deficient mice did not develop colitis under germ-free conditions (Sellon et al., 1998). Later on, Sydora and colleagues demonstrated that both adult and neonatal germ-free IL-10 deficient mice developed colitis when colonized with complex microbial communities, suggesting that lack of IL-10 leads to loss of tolerance towards bacterial antigens independent of the time at which microbial colonization takes place (Sydora BC, Tavernini MM, Wessler A, Jewell LD, 2003). Additionally, results from antibiotics studies have demonstrated that antibiotic treatment prevents colitis development in IL-10 knockout mice, as well as in other mouse models of IBD. Interestingly, under germ-free conditions, IL-10 knockout mice responded differently depending on the bacterial species. For example, members of the *Lactobacillus* and *Bifidobacterium* species ameliorated colitis (McCarthy et al., 2003; Schultz et al., 2002), while mono-associations with *E. faecalis* (Balish and Warner, 2002; Steck et al., 2011), *E-coli* (Kim et al., 2005) or *Mycobacterium avium* (Naser et al., 2014) led to colonic inflammation. Similarly, dual association with *E. faecalis* and *E-coli* could drive colitis (Kim et al., 2007). Additionally, colonization of GF IL-10 knockout mice with a simplified human microbiota consortium (SIHUMI) of 7 human IBD-derived bacterial strains induced colitis (Eun et al., 2014). Collectively, this data confirms the impact of microbial community composition in driving inflammation and the usefulness of IL-10 deficient mice as a model to study human IBD.

1.6.2 Gnotobiotic humanized mice: opportunities and challenges

Gnotobiotic humanized mice are germ-free (GF) mice that are colonized with human gut microbiota of different levels of complexity (complex, minimal consortia or a single strain, all of human origin). Over the past two decades, several reports have been published on the use of gnotobiotic humanized mice in modelling multiple human diseases (**Table 1**), including obesity, type 1 diabetes, malnourishment, asthma, and IBD (Arrieta et al., 2016a; Blanton, 2016; Kostic et al., 2016; Lepage P, Häsler R, Spehlmann ME, Rehman A, Zvirbliene A, Begun A, Ott S, Kupcinskis L, Doré J, Raedler A, 2011; Turnbaugh et al., 2006). In the field of IBD, Nagao-Kitamoto and colleagues colonized GF IL-10 deficient with gut microbiota collected from CD and UC patients and healthy donors. They showed that humanized mice reflected the dysbiotic features; in terms of community richness and diversity, as well as the disease phenotype (inflammation status) of their respective human donors (Nagao-kitamoto et al., 2016). On the other hand, colonization of GF mice with gut microbiota from healthy subjects was significantly less inflammatory.

Although most humanization studies in literature proved to be successful in recapitulating human diseases, in-depth analysis of microbiota transfer efficiency of human microbiota to mice, and hence

validation of the model applicability was lacking. The published data suggests that human microbiota established in ex-GF humanized mice differs from the original human donor microbiota at least at lower taxonomic levels. Over the years, several validation studies were performed to address these questions (**Table 2**). One of the earliest validation studies was published by Wong and colleagues (1996). In their work, they colonized GF mice with complex microbiota for a period of 6 months. Over the colonization period, several predominant aerobic and anaerobic bacterial taxa persisted at stable numbers in the intestinal tracts of mice. However, *Bacillus species* and both aerobic and anaerobic *Lactobacillus species* disappeared within 7 days after inoculation (Wong et al., 1996). In a later study by Kibe and colleagues, they demonstrated that although humanized mice reflect the composition of respective human gut microbiota, there are differences between dominant bacterial populations. The authors concluded that humanized mice harbor a limited subset of bacterial taxa from the human donor that were preferentially selected by unknown microbe-host interactions (Kibe et al., 2005). In 2006, Rawls and colleagues conducted a series of intriguing experiments where they performed reciprocal transplantations of complex microbiota into GF zebrafish and mouse recipients (Rawls et al., 2006). Gut microbiota profiling showed that generally, gut communities resemble the community of origin at phyla level. However, these taxa changes in proportions to resemble the microbiota composition of the recipient host. Similar observations were made by Chung and colleagues (2012) who colonized GF mice with microbiota of human or mouse origins (Chung et al., 2012). They similarly demonstrated that gut microbial communities resembled its original inoculum, however bacterial community composition differed significantly at species level, in particular within the phylum *Firmicutes*. This taken together, humanized mice stands as an invaluable model in studying human microbiome, however, like any other model, they appear to have evolutionary and ecological limitations that need to be accounted for when findings are extrapolated to human disease (Arrieta et al., 2016b).

Table 1 Use of gnotobiotic humanized mice in modelling human disease

Disease	Colonization/Experiment	Human Disease recapitulation	Features of Transplanted Microbiota	Ref
Diet-induced Obesity (2009)	human fecal microbiota diet switch from low-fat, plant polysaccharide-rich diet to a high-fat, high-sugar diet	Humanized mice fed the Western diet have increased adiposity Transmissibility of phenotype by microbiota transplantation	Proportion of transferred bacterial taxa changed 33 out of 42 genus-level human donor taxa could be established at an abundance level >0.02% of humanized mice	(Turnbaugh et al., 2009)
Obesity (2013)	human microbiota from four twin pairs discordant for obesity, or with culture collections from an obese (ob) or lean (Ln) co-twin. mice were fed chow diet or one of two diets (high or in low in saturated fats)	Obesity and metabolic phenotype was transmissible to humanized mice Co-housing of (ob) and (Ln) rescued mice from the phenotype.	Unweighted UniFrac-based comparisons differentiated between mice harboring all lean versus all obese co-twins	(Ridaura et al., 2013)
Asthma (2016)	human microbiota from asthmatic children	Four bacterial taxa ; <i>Faecalibacterium</i> , <i>Lachnospira</i> , <i>Rothia</i> , and <i>Veillonella</i> (FLVR) ameliorated asthma in an OVA-challenged mouse model	Taxa-specific shifts in abundance at the family, genus, and species' level	(Arrieta et al., 2016a)
Childhood malnourishment (2016)	human microbiota from 6- and 18-month-old children with varying degrees of undernutrition mice fed a representative Malawian diet.	Transmission of impaired growth, altered bone morphology, and metabolic abnormalities in muscle, liver, and brain	Bacterial taxa harbored by recipient mice correlated with age-discriminatory bacterial taxa deduced from application of machine learning methods on Malawian infants during first 3 postnatal years	(Blanton, 2016)
Inflammatory Bowel Diseases (2016)	human microbiota from CD and UC patients	CD microbiota induced more severe colitis than healthy control microbiota, IBD-associated microbiota induced pro-inflammatory gene expression profile that resembles signatures found in CD patients	Microbial composition of donor patients' microbiota was not completely reproduced in humanized mice Dysbiotic features (reduced richness and diversity) recapitulated those in human donors	(Nagao-kitamoto et al., 2016)

Table 2 Validation studies on transplanted microbiota

Purpose	Experiment	Findings	Ref
Testing how host gut habitat shapes microbial structure	Reciprocal transplantations of mice and zebrafish gut microbiota	Proportions of transplanted bacterial taxa change to resemble the microbial community composition of the recipient host	(Rawls et al., 2006)
Validation of gut immune maturation in humanized mice	Colonization with mouse microbiota (MMb) or human microbiota (HMb) Determination of small intestinal immune maturation	HMb mouse intestines had low levels of CD4 ⁺ and CD8 ⁺ T cells, few proliferating T cells, few dendritic cells, and low antimicrobial peptide expression (all characteristics of GF mice) MMb and HMb Mouse Gut Microbiotas are similar in major bacterial phyla abundance with differences at the OTU level Loss of genus-level specially within the Phylum Firmicutes	(Chung et al., 2012)
Evaluation of microbiota transfer efficiency	Colonization with human microbiota GF recipient rats, GF recipient mice and antibiotic-treated specific pathogen-free (SPF) recipient mice Evaluation of transfer efficacy	Majority of abundant human bacterial phlotypes were established in rat, only a subset established in mice . Mice enriched for Bacteroidetes (<i>Bacteroides spp.</i>) Members of <i>Clostridia cluster IV</i> are poorly established in mice 8 abundant taxa belonging to Clostridia were well represented in rat, but completely excluded in mouse 86 phlotypes were detected in all animals, but not captured in human donor (rare-function) Firmicutes/Bacteroidetes ratio of human donor maintained in rats, but not in mice	(Wos-Oxley et al., 2012)

1.7. Mining microbial signatures: *structure and function*

The vast advancement of high-throughput technologies to study the gut microbiomes at various levels revolutionized our understanding of microbiota and its role in health and disease. Methods for studying gut microbiota can be classified as culture-dependent or culture-independent techniques. Combined use of both approaches allows the study of complex microbial communities' composition and functional capacities. While **culture-dependent techniques** allow the study of live bacteria and provide pure cultures of isolates for further mechanistic studies in animal models, they provide little information about community dynamics, community members' interactions and ignores the uncultivable microbes, which in reality make up most of the gut community diversity (Wang et al., 2017). As high-throughput **culture-independent** sequencing technologies emerged, the possibility to study complex microbial communities improved significantly, allowing for more detailed and rapid analysis of microbial communities. The two main approaches for analyzing the microbiome are 16S ribosomal RNA (rRNA) gene amplicons sequencing and shotgun metagenomics. While 16S rRNA gene sequencing relies on sequencing of variable regions in the universally conserved bacterial ribosomal genes, deeper resolution of community composition can be achieved by performing whole genome shotgun sequencing (WGS). 16S rRNA is commonly used to characterize the composition of complex bacterial communities, on the other hand, WGS provides insights into the functional potential of the microbial community, by identifying enriched genes and linking them to known functions. Additionally, the rapid advancement of technologies like nuclear magnetic resonance (NMR) and mass spectrometry (MS) has enabled the study of microbial metabolite profiles. In this section, technologies used to study gut microbiota in this work are reviewed.

1.7.1 16S rRNA Gene Sequencing for community structure analysis

The 16S rRNA gene is a highly conserved component of the transcriptional machinery. It contains 9 "hypervariable regions" that exhibit considerable sequence variability among different bacterial species and thus can be used to identify different species in a complex bacterial community (Van de Peer, 1996). These hypervariable regions are flanked by conserved regions in bacteria, which allows PCR amplification of target sequences using universal primers (Baker et al., 2003; Lu et al., 2000; McCabe et al., 1999). In this way, 16S amplicons reflecting the composition of the community organisms are generated. Reads from 16S rRNA gene sequencing are clustered according to their similarity, and Operational Taxonomic Units (OTU) are defined based on the similarity threshold (usually 97% similarity). Thus, OTUs are representative sequences of "clusters" comprising related sequences and can be handled as proxies for microbial "species". Taxonomic binning is achieved by comparing OTUs against databases, such as The Ribosomal

Database Project (RDP) (Maidak et al., 1997), GreenGenes (DeSantis et al., 2006), and SILVA (Quast et al., 2013) for identification and classification of the organisms in the microbiome. OTU counts are subsequently summarized in a table of relative abundances for each organism in each sample. For downstream analysis, several pipelines have been developed for analysis of 16S rRNA gene sequence data (Caporaso et al., 2010; Schloss et al., 2009). In this work, we use the tool developed in our group, Integrated Microbial Next Generation Sequencing (IMNGS) (Lagkouvardos et al., 2016).

1.7.2 Whole genome shotgun (WGS) sequencing for functional characterization

Whole Genome Shotgun (WGS) sequencing has the advantage of profiling both the taxonomic composition and the functional genetic capacity of a complex microbial community. The functional potential of communities is inferred by comparing sequencing reads against databases, such as KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway database (Kanehisa et al., 2004) and COG (Clusters of Orthologous Groups of proteins) database functional categories (Tatusov et al., 2000). Besides, the main advantage is that it can provide higher resolution of microbial communities by characterizing bacteria to their strain level, something that is not feasible by performing 16S rRNA gene sequencing. Collectively, metagenomics provides comprehensive knowledge of the entire microbial gene repertoire and its function.

1.7.3 Metabolomics - linking microbial diversity with metabolic activity

Metabolomics has been applied extensively in the study of gut microbiota to understand the relationship between gut microbiota and the metabolic status of the host (Yan et al., 2016). Technologies such as mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy are used to characterize and quantify low-molecular weight molecules (metabolites) in various biological samples (Goodacre et al., 2004). The integration of metagenomics and metabolomics provides important information on how microbial communities function in the human gut, how they impact the host metabolic phenotypes and provide a fingerprint of microbiota functional capacities. Metabolite measurements showed an association between metabolite profiles and host phenotype, and thus stand a promising tool for improved diagnosis and prognosis of metabolic diseases, as well as for prediction of response to therapies. Two types of measurements are available: untargeted or targeted analysis. **Untargeted analysis** offers the opportunity for novel target discovery, and usually involves the comparison of the metabolome between the control and treatment groups, to find differences between their metabolite profiles that could be attributed to specific biological functions. On the other hand, **targeted**

metabolomics measure defined groups of chemically characterized and biochemically annotated metabolites using internal standards and the use of quantitative measurements.

1.7.4 Machine learning: recognizing patterns and connecting the dots

IBD is a complex disease, driven by shifts in bacterial community composition, rather than single pathobionts. Hence, identification of bacterial taxa predictive of IBD would allow the use of bacterial signatures as biomarkers for improved prognosis and personalized treatment regimens. In this context, machine learning approaches are useful for identifying patterns and revealing microbe-microbe relationships in complex systems such as the human gut microbiome. Generally, traditional statistical testing is employed to identify differential abundance or prevalence of bacterial groups associated with different disease phenotypes, however these methods are limited and possibly overlook important intertwining relationships and correlations between important features of the ecosystem. In contrast, machine learning algorithms recognize specific patterns and features that can be used for prediction. The process of supervised learning involves training the algorithm on a set of pre-categorized data points, and building a model that can predict the correct categorical class of future novel input data based on what the algorithm has previously learnt. The model then determines a set of features that best discriminate between different categories. A good model is one that can be applied to external datasets and yield accurate predictions. Hence, a model that is too general will fail to capture differences or make distinctions between groups (underfitting), while a model that is too specific would fail to extrapolate to data points outside the training set (external dataset) (overfitting). Training the machine learning model on more input data leads to better and more accurate predictions (Knights et al., 2011a). In context of microbial signatures discovery, supervised classification requires training data, where each training data point has values for a number of features (for example, relative abundance of taxa) and an associated qualitative dependent variable giving the correct classification (for example, disease status or response to therapy). The model would be trained on a set of communities with known categorized phenotype or class and identifies a set of discriminative variables as “microbial signatures” which can be then used to predict phenotypes of other bacterial communities in future experiments (Knights et al., 2011b). Identification of a set of disease-associated discriminative taxa would accurately classify patients from healthy controls or responders from non-responders. However, it is not necessarily enough to determine whether these changes are a cause of consequence of disease. Validation in external cohorts and the analysis of large prospective cohorts could help overcome this challenge by following up the patients before the onset of disease.

2. Aims and scope of dissertation

Substantial effort has been dedicated to assessing the use of microbial signatures as a diagnostic biomarker tool for inflammatory bowel disease (IBD). Despite a growing body of evidence suggesting a causative role of the intestinal microbiota in disease pathogenesis, the impact of bacteria on disease progression, disease phenotypes and the risk to relapse after initial response to therapy is still unclear. In clinical practice, pharmacological treatment of the disease relies on suppression of the exaggerated immune system through the blockade of inflammatory processes by immune-suppressive drugs and biological therapy (mostly antibodies targeting the cytokine TNF). However, a significant proportion of patients fail to improve or maintain long-term remission. For these patients, surgical resection or palliative treatment is the last resort. Accordingly, there is a great need to optimize response through monitoring of the changes in gut microbiome patients undergo overtime and identifying patients who are most likely to respond to treatments, before initiation of therapy.

The aim of the present work was to investigate the functional role of intestinal microbiota in driving disease pathogenesis, and to ultimately identify microbial signatures linked to change of disease state or response to therapy, with a potential of being used as biomarker signatures in the clinical settings. Towards this aim, we adopted an integrative multi-omics approach, combining 16S rRNA microbial profiling, metagenomics, metabolomics, with humanized gnotobiotic mice.

This dissertation consists of four parts. The first part dealt with the characterization of a longitudinal cohort of Crohn's disease (CD) patients who underwent autologous hematopoietic stem cell transplantation (HSCT). Patients were monitored for up to 5-years upon HSCT therapy, and microbial profiling was performed on stool samples collected at baseline (n=15), during remission (n=35) or during relapse (n=20). Patients were classified as responders or non-responders based on clinical and endoscopic assessment one year after HSCT initiation. In the second part, a humanized IBD model was established, by colonizing germfree *Il-10^{-/-}* mice with selected fecal samples from CD patients to address the functional impact of microbial profiles in IBD. Changes in gut microbiome and metabolome in human donors and humanized mice were assessed at multiple levels, including taxonomic composition (16S rRNA gene sequencing), functional potential (shotgun metagenomics and predicted metagenomes), as well functional activity (targeted and untargeted metabolite profiling). In the third part, gut microbiome in two additional cohorts of IBD was characterized. Patients were either treated with biological therapy (n=33; 19 CD and 14 UC), or with surgery (n=13). Stool samples were collected at both baseline and at different time points during follow-up. In the last part of this work, the adequacy of humanized

mouse model in IBD research was addressed. We validated the transfer efficiency of human complex microbiota into germ-free mice, assessed different colonization protocols and the stability of humanized mice in terms of inter-generational microbiota transfer efficiency.

3. Materials and methods

3.1 Human study cohorts - cohort description and data collection

3.1.1 Hematopoietic stem cell transplantation (HSCT) cohort

This longitudinal study comprised 35 refractory Crohn's disease patients who were recruited in Barcelona (Department of Gastroenterology, IDIBAPS, Hospital Clínic, CIBERehd, Barcelona, Prof. Dr. Julian Panes and Dr. Azucena Salas). Patients were treated by autologous hematopoietic stem cell transplantation (AHSCT) as last line-therapy. Mucosal biopsies from involved areas, as well as, fecal samples were collected longitudinally overtime starting from (t=0) before the HSCT and at different time points during the 5-year follow-up period. In order to address the question of which microbial factors are associated with inducing remission or response in these patients, we formulated our question around two main points: (i) identification of microbial changes associated with response or failure after HSCT and (ii) identification of microbial changes associated with inducing long term remission or otherwise relapse (**Figure 4**).

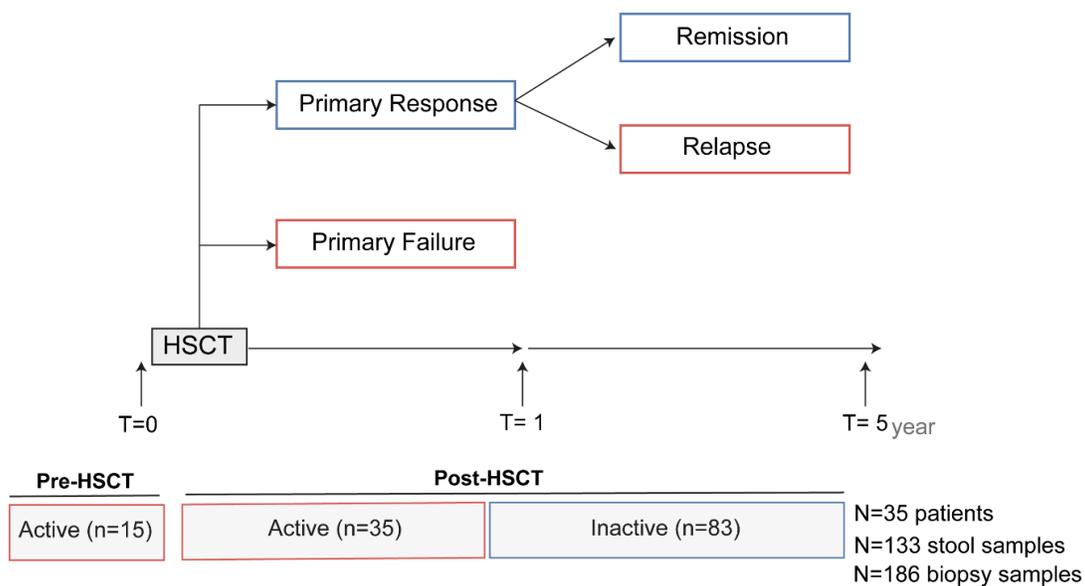


Figure 4 Hematopoietic Stem Cell Transplantation Study Design

Fecal samples were collected from CD patients who underwent HSCT both at baseline and at different time points during 5-year follow-up

CD patients presented different phenotypes, disease location and disease activity. The patients were sub-classified based on whether they are receiving treatment or not and based on whether they are smokers, quitters or never-smoked (**Table 3**).

Table 3 Characteristics of patients initiating Hematopoietic Stem Cell Transplantation (HSCT) therapy

Gender F / M, n (%)	22 (71) / 9 (29)	CDAI mean (min-max)	284.7 (123.4-455.7)
Age at inclusion mean, yr (min-max)	28.9 (16.5 – 49.3)	Extraintestinal manifestations, n (%)	23 (74)
Duration of disease, yr mean (min-max)	9.9 (1.8 – 25.7)	Perianal disease, n (%)	16 (52)
Location (Montreal), n (%)		Active at inclusion	6 (19)
L1 (ileal)	2 (7)	Smoking history, n (%)	
L2 (colonic)	8 (26)	Smoker	3 (10)
L3 (ileocolonic)	14 (45)	Non-smoker	18 (58)
L1 + L4 (ileal + upper)	1 (3)	Previous smoker	10 (32)
L3 + L4 (ileocolonic + upper)	6 (19)	Previous therapies, n (%)	
Activity location before HSCT, n (%)		Infliximab, Adalimumab	31 (100)
L1 + L4 (ileal + upper)	7 (23)	Thiopurines	30 (97)
L2 (colonic)	15 (48)	Methotrexate	28 (90)
L3 (ileocolonic)	9 (29)	Other immunomodulators	17 (55)
Behaviour, n (%)		Other biologics	16 (52)
Inflammatory	21 (68)	Study medication	3 (10)
Stricturing	3 (10)	Previous CD surgeries, n (%)	15 (48)
Penetrating	7 (22)	Ostomy at inclusion	6 (19)
		Body Mass Index, mean (min-max)	22.7 (14-39.7)

3.1.2 REMIND post-surgical cohort

We studied a subset of this cohort which was recruited in Paris (Hospital San Louis, Prof.Dr. Matthieu Allez and Dr. Lionel Le-Bouhris). The REMIND group performed a multi-center prospective study in 9 French university medical centers. Adult CD patients undergoing ileocolonic resection were included in the analysis. Exclusion criteria included pregnancy and patients with dysplasia on surgical specimen. Demographic data including gender, age, smoking and status at surgery were included. Additionally, clinical data including disease phenotype, disease activity indices, medical history were recorded. In this cohort of patients, recurrence manifests as inflammation in the neo-terminal ileum and at the anastomosis in 88% of cases. Fecal samples were collected at baseline before the surgery and 6-months post-surgery. Based on the endoscopic scoring, 8 patients achieved remission, 5 patients were in recurrence (n=3, substantial recurrence and n=2 in advanced recurrence (**Figure 5**)).

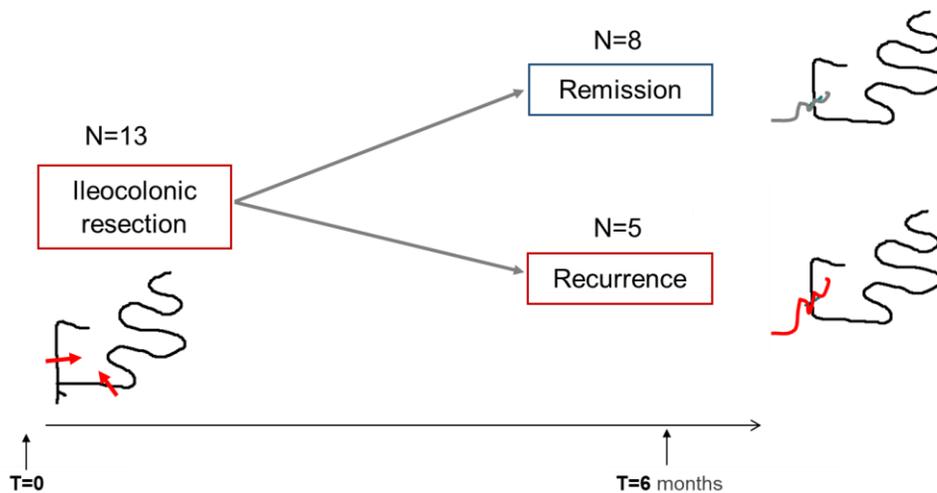


Figure 5 REMIND Post-surgical cohort Study Design

Stool samples were collected from patients at (T0) at the ileocolonic resection surgery and at (T=6 months) post- surgery. Thirteen paired samples are included in the analysis; of each 8 patients went in remission, and 5 recurred (n=3; substantial recurrence, n=2; advanced recurrence).

Rutgeerts score (RS) was used for the assessment of post-surgical recurrence. RS scoring system encompasses five categories (i,0–i,4) (**Table 4**). for the endoscopic examination based on the severity of lesions at the anastomosis and neo-terminal ileum region after ileocolonic resection (Rutgeerts et al., 1990). Clinical studies showed a significant correlation of RS score with the endoscopic recurrence after surgery, and thus it is widely used in clinical settings. A score of i2 or i3 and i4 correspond to substantial recurrence or advanced recurrence, respectively, while a score of i0 and i1 correspond to endoscopic remission.

Table 4 Rutgeerts Scoring System for assessment of post-surgical recurrence in CD patients

Rutgeerts Grade	Decoding
i0	Post-surgery endoscopic remission
i1	Post-surgery endoscopic remission
i2	Post-surgery substantial recurrence
i3	Post-surgery advanced recurrence
i4	Post-surgery advanced recurrence

3.1.3 Biotherapy Cohort

This longitudinal study comprised 33 patients with IBD (19 CD and 14 UC) who are treated with anti-TNF and other biotherapy drugs and recruited in Paris (Hospital San Louis, Prof. Dr. Matthieu Allez and Dr. Lionel Le-Bouhris). The mean HBI and Mayo score at baseline were (10.2 and 7.2) respectively with a mean CRP of (16.02 mg/L). At week 14, 6 CD patients and 1 UC patient met the criteria of clinical remission. Fecal samples were collected at baseline, at week 14 and at week 52 (1Y) post- initiation of biotherapy. Patients who did not receive antibiotics at least 3-months before initiation of biological therapy are included in the analysis. Biological therapy included anti-TNF (Adalimumab, Infliximab or other biologics (Vedolizumab and Ustekinumab) (**Figure 6**). Patients gender administered drugs and disease activity scores are enlisted in (**Table 5**).

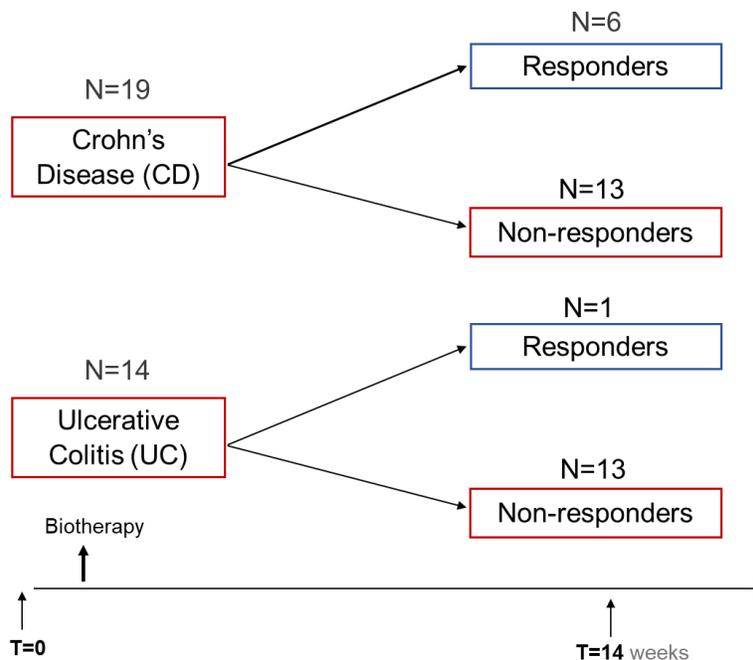


Figure 6 Biotherapy Study Design

Stool samples were collected from patients both at the initiation of therapy (T=0) and 14 weeks later at (T=14 weeks) post. Thirty-three patients were included (19 CD and 14 UC)

Table 5 Characteristics of the included patients initiating biotherapy treatment

		CD	UC
Gender (F/M)		(9 /10)	(7/6)
Disease Location	ileum	5	NA
	Colon	1	1
	rectum	5	6
	Sigmoid	7	7
	anastomosis	1	NA
Treatment	Adalimumab	4	1
	Humira	3	1
	Humira Imurel	1	
	Humira Metho	1	
	Inflectra	1	1
	Inflectra Imurel	1	
	Infliximab	3	5
	Remicade	1	
	Ustekinumab	2	
	Vedolizumab	2	3
	Simponi		2
	Golimumab	1	
	CDEIS/UCEIS	14 ±9	4±2
	HBI/Mayo	8 ±6	6±3

3.2 Processing of human fecal samples

Frozen fecal samples were transferred to a UV-sterilized biosafety hood and pulverized to be homogeneous under liquid nitrogen using sterile mortar and pestle, as described previously (Ridaura et al., 2013). To prevent cross-contamination of samples, individual autoclaved sets (one mortar and one pestle each) were used to process each sample and UV-sterilization was carried out before and after the processing of each stool sample. Aliquoted samples were stored at -80°C until use.

3.3 Studies in gnotobiotic mice

Mouse experiments and the treatment protocols were approved by the local institution in charge (Regierung von Oberbayern, approval number (55.2-1-54-2532-133-2014)). All animals were housed in the germ-free (GF) mouse facility at the Technical University of Munich (School of Life Sciences Weihenstephan).

3.3.1 Housing conditions

Animals were housed in gnotobiotic isolators (plastic flexible film) and maintained at (12 h light/dark cycles at 24-26°C) and fed a standard autoclaved chow diet (Ssniff, Soest, Germany, M-Z V1124-300 for GF-animals) *ad libitum* and were sacrificed by CO₂. Germfree status was checked by cultivation of feces in Wilkins-Chalgren Anaerobe (WCA) broth (OXOID, UK) and by Gram staining of fecal suspensions before start of the experiments and at sampling time. Presence of mold was controlled by a mold-trap.

3.3.2 Colonization of GF mice with human microbiota

Colonization of germ-free mice was performed by diluting 400 mg frozen aliquot of the pulverized human stool sample (homogenized in final concentration of 20% sterile glycerol and frozen at -80°C) in 2 mL of reduced PBS (PBS supplemented with 0.05% L-cysteine-HCl) in an anaerobic Coy chamber (atmosphere, 75% N₂, 20% CO₂, 5% H₂) and then vortexed at room temperature for 5 minutes. The fecal suspension was allowed to settle by gravity for 5 minutes to exclude residual particulate matter, afterwards, the clear supernatant was transferred under anaerobic conditions into an anaerobic crimped tube (Hungate) which is then transferred to the gnotobiotic facility. Each germ-free recipient mouse received 100 µL of the suspension via oral gavage using 20 Gauge gavage needle (Fine Science Tools). Colonized mice were housed in group-specific isolators reserved to mice colonized with the same human microbiota. Mice were sacrificed 1 week or 4 weeks after colonization.

3.3.3 Bacterial enumeration in mouse fecal samples

Bacterial enumeration was determined by performing serial dilutions of fecal suspension. Fecal samples were diluted 10-fold in filter-sterilized reduced PBS and homogenized by vortexing using sterile glass beads. Serial dilutions were plated on WCA media under aerobic and anaerobic conditions. Colony forming units (CFU) were counted 24 hours after incubation at 37°C in anaerobic coy chamber and number of CFU per gram of feces/content was calculated.

3.4 Characterization of bacterial community's structure and function

3.4.1 Metagenomic DNA extraction from fecal samples

DNA was extracted from frozen mouse colon content or pulverized human fecal samples by bead-beating followed by a modified version of the protocol by Godon et al (Godon et al., 1997). Briefly, a volume of 600 µL DNA stabilizing solution (Stratec Biomedical) was added to the frozen fecal aliquots in 2-ml screw-cap polypropylene microcentrifuge tube containing sterile 500 mg Silica beads (0.1-mm-diameter; BioSpec Products) and kept on ice. After the addition of 250 µL of 4 M guanidine thiocyanate - 0.1 M Tris (pH 7.5) and 500 µL of 5% N-lauroyl sarcosine - 0.1 M phosphate buffer (pH 8.0), fecal suspensions were vortexed briefly and incubated at 70°C for 1 hour with constant shaking. The mixture was mechanically disrupted by bead beating using a FastPrep®-24 bead beater (MP Biomedicals) supplied with a 24 x 2 mL cooling adaptor 3 times each for 40 seconds at a speed of 6.5 m/ sec. Afterwards, an amount of 15 mg of Polyvinylpolypyrrolidone (PVPP, Sigma Aldrich) was added as polyphenol adsorbent, and suspension was centrifuged for 3 minutes at 15.000×g at 4 °C. Supernatant was recovered in a new 2-mL tube, and further centrifuged for 3 minutes at 15.000 g at 4 °C. To remove bacterial RNA, a volume of 2µl RNase (10mg/ml) was added to 500 µl clear supernatant and incubated at 37°C for 30 minutes with constant shaking. Finally, the genomic DNA was purified using the NucleoSpin® gDNA clean-up kit (Macherey Nagel) following the manufacturer's instructions. Concentration and purity of the extracted DNA was determined using the NanoDrop® Spectrophotometer ND-1000 (Thermo Scientific), and samples were stored at 4 °C during library preparation and at -20 °C for long-term storage.

3.4.2 Metagenomic DNA extraction from mucosa-associated bacteria (tissue biopsies)

DNA was isolated from biopsies using the NucleoSpin® Tissue kit (Macherey-Nagel). Briefly, tissue biopsies were resuspended in 180 µL sterile filtered lysis buffer (20 mM Tris/HCl; 2 mM EDTA; 1% Triton X-100 (pH 8) and supplemented with 20 mg/mL lysozyme freshly before use. Tissue suspensions were incubated in a thermoshaker (37 °C, 30 min, 950 rpm). Afterwards, Proteinase K (10mg/mL) was added and the suspension was vortexed vigorously and incubated

in a thermoshaker (56°C, 1-3 h, 950 rpm) until complete lysis of tissue pieces. Further purification steps were performed using the NucleoSpin® Tissue kit following the manufacturer's instructions.

3.4.3 High-throughput 16S ribosomal RNA (rRNA) gene sequencing and microbiome profiling

16S rRNA high-throughput gene sequencing and downstream analysis were performed as previously described (Lagkouvardos and Kl, 2015). In brief, the V3-V4 regions of the 16S rRNA gene was amplified (25 cycles for fecal samples, 15x15 cycles for biopsies) using previously described two-step protocol (Berry et al. 2011) using forward and reverse primers 341F-785R (Klindworth et al., 2013). Amplicons were purified using the AMPure XP system (Beckmann). Afterwards, sequencing was carried out with pooled samples in paired-end modus (PE275) using a MiSeq system (Illumina, Inc.) according to the manufacturer's instructions and 25% (v/v) PhiX standard library. Data were analyzed as described previously (Lagkouvardos et al., 2016). Raw reads were processed using IMNGS pipeline based on the UPARSE approach (Edgar, 2013). Sequences were demultiplexed, trimmed to the first base with a quality score <3 and then paired. Sequences with less than 300 and more than 600 nucleotides and paired reads with an expected error >3 were excluded from the analysis. Remaining reads were trimmed by 5 nucleotides on each end to avoid GC bias and nonrandom base composition. The presence of chimeras was tested using UCHIME (Edgar et al., 2011). Clustering of Operational taxonomic units (OTUs) was done at 97% sequence similarity. Taxonomic binning was assigned at 80% confidence level using the RDP classifier (Wang et al., 2007) and compared to that of the SILVA ribosomal RNA gene database project (Quast et al., 2013). Analysis was performed using R-package Rhea (Lagkouvardos et al., 2017). Briefly, rarefaction curves were used to assess sequencing depth and to eliminate low quality reads. OTUs counts were normalized and percentage relative abundance was computed. Beta-diversity analysis was used to assess the diversity between groups based on generalized UniFrac distances. Alpha diversity within species was calculated based on species richness and Shannon effective number of species.

3.4.4. Whole-genome shotgun Sequencing

DNA concentrations were measured using Quant-It™ PicoGreen® dsDNA Assay Kit (ThermoFisher Scientific, MA, USA) and a spectrofluorometer (SpectraMax Gemini EM microplate reader Molecular Devices, LLC, USA). DNA purity check was assessed spectrophotometrically (Nano Drop 1000, ThermoFisher Scientific, USA). In total 200 ng of DNA per sample was sheared using an E220 Focused-ultrasonicator (Covaris® Inc., MA, USA) targeting 300-400 bp fragments following Covaris's instructions. Metagenomic libraries were constructed using NEBNext® Ultra II™ DNA Library Prep Kit for Illumina®. Dual indexing was done using the kit NEBNext® Multiplex

Oligos for Illumina® (Dual index primers set 1, New England BioLabs, UK). Purification and size selection were performed based on Agencourt® AMPure® XP (Beckman-Coulter, MA, USA). Libraries inserts ranged between 400 and 500 bp were evaluated using a Fragment Analyzer™ (Advanced Analytical, IA, USA) using the DNF-474 High Sensitivity NGS Kit (Agilent, Waldbronn, Germany). One sample with sterile water was used as a control for the metagenomics library preparation and sequencing. Libraries quantification were performed using Quant-It™ PicoGreen® dsDNA Assay Kit. Libraries were diluted to 12 pM and 1% PhiX control DNA (Illumina, CA, USA) was spiked in and the libraries were sequenced on an Illumina HiSeq 2500 (Illumina, CA, USA) using the Rapid run paired-end mode (2 × 250 bp).

Quality control

The raw metagenome samples (with an average of ~ 19 million read pairs per sample) were processed using Trimmomatic version 0.36. First the adapter sequences were removed, retained reads with target length of at least 90 bp and strictness parameter 0.4 (MAXINFO:30:0.4) were further processed to obtain good quality of ~ 15 million read pairs in average per sample. Reads coding for ribosomal genes were removed from the samples using sortmeRNA version 2.1b by mapping the quality-controlled samples against SILVA database version 132 to obtain non-rRNA good quality read-pairs of ~15 million read-pairs per sample in average. Reads related to the human (~ 12 million read-pairs per sample in average) and mouse genome (~ 11 million read-pairs per sample in average) were further removed from the samples by mapping against their genomes using hisat2. Finally, PhiX contaminant reads were also removed human (~ 11 million read pairs per sample in average).

Taxonomic profiling

Pre-processed reads from samples from the same treatments were pooled and assembled using megahit version 1.1.3-0. Assembled contigs were used for taxonomic annotation using CAT database, available on (<https://github.com/dutilh/CAT>). This tool internally uses prodigal v2.6.3 for gene prediction and DIAMOND v 0.9.14 for the alignment against the non-redundant (nr) protein database. Annotated contigs were used as reference and mapped against corresponding samples from each treatment to the reference contigs using bbmap. The taxonomic profile was used as a measure for relative abundance. LEfSe analysis was performed from this relative taxonomic abundance profile using <http://huttenhower.sph.harvard.edu/lefse/>. Those taxa with a Kruskal-Wallis p-value <5% and LDA score with at least 100x fold change (log10 fold change of 2) were considered as potential taxonomic biomarkers.

Functional profiling

Gene-annotation was performed for the assembled contigs from each treatment using prodigal v2.6.3. In order to obtain KEGG annotation and to reconstruct KEGG Genes and thereby the functionally associated gene sets named KEGG Modules, the amino acid sequences were used in the KEGG internal annotation tool named GHOSTKOALA. The fraction of KEGG Modules present in a treatment was obtained using R package 'metQy'. This allowed a characterization of the functional capabilities of the microbial communities based on the complete or incomplete presence of functional units. The annotated gene coding sequences obtained from prodigal were used as reference to map for the post-processed reads from the corresponding samples using bbmap to obtain the relative abundance of the KEGG Modules. The relative abundance of KEGG Modules of samples from different treatments was obtained by LEfSe analysis. KEGG Modules with a Kruskal-Wallis p-value <5% and LDA score with at least 10x fold change (log₁₀ fold change of 1) were considered as potential functional biomarkers (<http://huttenhower.sph.harvard.edu/lefse/>).

3.4.5 Metabolomics

Targeted and Untargeted metabolomics measurement and data analysis was performed in collaboration with Sinah Schmidt (Chair of Food Chemistry and Molecular Sensory Science, Technical University of Munich) and Andreas Dunkel (Leibniz-Institute for Food Systems Biology at the Technical University of Munich), respectively.

Sample preparation

Mouse colon content (20 mg) was mixed with 1 mL methanol-based Dehydrocholic acid extraction solvent (1.3 µmol/L) as an internal standard in a 2 mL bead beater tube (CKMix 2 mL, Bertin Technologies, Montigny-le-Bretonneux, France) filled with ceramic beads (1.4 mm and 2.8 mm ceramic beads i.d.). The mixture was homogenized by bead beating using a bead beater (Precellys Evolution, Bertin Technologies) supplied with a Cryolys cooling module (Bertin Technologies, cooled with liquid nitrogen) 3 times each for 20 seconds with 15 seconds breaks in between, at a speed of 8000 rpm. Afterwards, the suspension was centrifuged (10 min, 6000 rpm, 10°C), using an Eppendorf Centrifuge 5415R (Eppendorf, Hamburg, Germany). Finally, the 100 µL clear supernatant was mixed with 20 µL internal standard solution (c = 7 µmol/L) and injected into the LC-MS/MS system for targeted and the LC-TOF-MS system for untargeted analysis. Preparation of human fecal extracts was achieved in similar manner. Human fecal samples (100 mg) were mixed with 5 mL extraction solvent in a 15 mL bead beater tube (CKMix50 15 mL, Bertin Technologies, Montigny-le-Bretonneux, France), filled with ceramic beads (2.8 mm and 5.0 mm

ceramic beads i.d.). The subsequent steps were performed as described above for the mouse colon content samples.

Liquid Chromatography-Triple Quadrupole Mass Spectrometry (LC-MS/MS)

For quantitation a QTRAP **6500** mass spectrometer (Sciex, Darmstadt, Germany) was used in negative electrospray ionization (ESI) mode in combination with Multiple reaction monitoring (MRM) for detection and quantification of bile acids. For detection of the target ions, an ion spray voltage of -4500 V and the following ion source parameters were used: curtain gas (35 psi), temperature (450°C), gas 1 (55 psi), gas 2 (65 psi) and entrance potential (-10 V). The MS parameters and LC conditions were optimized using commercially available standards of endogenous bile acids and deuterated bile acids, for the simultaneous and unequivocal quantification of selected 34 analytes. For separation of the analytes, a Nexera X2 UHPLC (Shimadzu Europa GmbH, Duisburg, Germany) was used. The system consists of two LC pump systems 30AD, a DGU-20A5 degasser, a SIL-30AC auto-sampler, a CTO-30A column oven and a CBM-20A controller, and equipped with a 100 × 2.1 mm, 100 Å, 1.7 µm, Kinetex C18 column (Phenomenex, Aschaffenburg, Germany). Chromatography was performed with a constant flow rate of 0.35 mL/min using a mobile phase consisted of water (eluent A) and acetonitrile/methanol (3/1, v/v, eluent B), both containing 10 mM ammonium acetate and 0.1% formic acid. The gradient elution started with 32% B for 1.5 min, increased in 4.5 min to 50% B, in 2 min to 60% B, in 1 min to 62% B, increased in 2 min to 80% B, held for 0.5 min, increased in 0.5 min to 100% B; held 2 min isocratically at 100% B, decreased in 0.5 min to the initial ratio of 32% B, followed by 2 min of re-equilibration. The injection volume for all samples was 1 µL, the column oven temperature was set to 40°C, and the auto-sampler was kept at 10°C. Data acquisition and instrumental control were performed with Analyst 1.6.2 software (Sciex, Darmstadt, Germany).

Liquid Chromatography-Time of Flight-Mass Spectrometry (LC-TOF-MS)

For untargeted LC-MS analysis, an Nexera X2 UHPLC system (Shimadzu, Duisburg, Germany) consisting of two LC-30AD pumps, a SIL-30AC auto sampler, a CTO-30A column oven and a CBM-20A system controller was connected to 6600 TripleTof instrument (Sciex, Darmstadt, Germany) equipped with an IonDrive ion source (Sciex) operating in positive and negative electrospray mode. After each fifth sample the instruments calibration was verified and corrected using ESI Positive or ESI Negative Calibration solution (Sciex) and a Calibrant Delivery System (Sciex). UHPLC separation was performed on a 100 × 2.1 mm, 100 Å, 1.7 µm, Kinetex C18 column (Phenomenex, Aschaffenburg, Germany) using water (mobile phase A) and acetonitrile (mobile phase B) with 0.1% formic acid each and the following gradient program: 0 min 36% B, 2 min 36% B, 3.5 min 80% B, 5 min 100% B, 7 min 100% B, 8 min 36% B, 12 min 36% B. The total flow of the chromatography was set to 0.25 mL/min and separation was performed at 40°C. The mass

spectrometer was operated in the SWATH mode with a series of 19 consecutive experiments per 1.05 sec measurement cycle. After starting with a high-resolution scan of the intact precursor ions from 50 to 1000 m/z for 200 ms, fragment ions were generated by means of collision-induced fragmentation subsequently for precursor ions within 18 separate windows ranging from 50 to 600 m/z (window width 30 Da each, 1 Da overlap), the resulting fragment spectra were recorded in the high sensitivity mode (50 ms acquisition per window). Ion spray voltage was set at -4500 V in negative and 5500 V in positive mode and the following source parameters were applied: curtain gas 35 psi, gas 1 55 psi, gas 2 65 psi, temperature 500°C. Declustering potential was set to 80 V for all experiments while the collision energy was 10 V for precursor ion scans and 35 V including 25 V collision energy spread for the fragmentation in the individual SWATH windows.

3.5 Phenotypic characterization of inflammation in humanized mice

3.5.1. Histological scoring

Cecal Swiss-roll tissues were fixed in 4% formalin for 24 hours. Afterwards, they were dehydrated, embedded in paraffin and cut in 5 µm sections followed by hematoxylin and eosin (H&E) staining. Scoring of H&E stained tissue sections was done blindly by single observer through evaluating lamina propria mononuclear cell infiltration, crypt hyperplasia, goblet cell depletion and architectural distortion resulting in a score spanning from 0 to 12; as described previously (Erben et al., 2014) Images were acquired using Digital microscope M8 (PreciPoint GmbH).

3.5.2. Gene expression analysis

Total RNA was isolated from the site of inflammation (cecum tissue). Tissues were stored in RNAlater (Sigma Aldrich) at -80 °C until processing Tissue samples were thawed down on ice and mixed with RA1 buffer (Macherey Nagel) and β-Mercapthoethanol (Sigma Aldrich). Tissue pieces (ca. 30 mg) were opened longitudinally and disrupted by means of mechanical lysis with a sterile plastic pastille. For homogenization, tissue suspensions were passed through a 0.9 mm syringe needle and were loaded on QIAshredder columns (QIAGEN). Isolation of RNA was performed following manufacturer's instructions (NucleoSpin® RNAII kit; Macherey-Nagel). Quantity and purity of the RNA was measured using NanoDrop® (Thermo Fisher Scientific). A total of 1 µg RNA in 13 µl PCR-grade H₂O using random hexamers and M-MLV RT Point Mutant Synthesis System (Promega) was used for reverse transcription (Table). qPCR was performed using the LightCycler® 480 from Roche with 10 µM UPL-Probe, 20 µM forward and reverse primer, ca. 125 ng/ µl cDNA and 2x Probe or Master Mix. The settings for qPCR are displayed in **(Table 6)**. Primers and probes used were: *tnf* (5'-tgccatgtctcagcctcttc 3'; 5'-gaggccatttgggaacttct-3'; Probe #49); *gapdh* (5'-tccactcatggcaaattcaa-3'; 5' ttgatgttagtgggtctcg-3'; Probe #9). *IFNγ* (5'-ggaggaactggcaaaaggat-3';5'-ttcaagacttcaaagagtctgagg-3', Probe#21); *IL6* (5'-

gctaccaaactggatataatcagga 3', 5'-ccaggtagctatgggtactccagaa-3', Probe#6). Calculation of the relative expression of mRNA was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) and normalized for the expression of GAPDH. Data were expressed as fold change against Wildtype germ-free control matching control.

Table 6 Protocol for qPCR gene expression analysis

cDNA synthesis Components	Volume per sample [μ l]
Random-Hexamer [200ng/ μ l]	1
5 min. 70°C	
5 min. 4°C	
5x 1. Strang Puffer (Promega)	5 μ l
PCR grade H ₂ O	3,1 μ l
rRNasin® (Promega)	0.65
dNTP Mix 10mM	1,25 μ l
M-MLV (200U/ μ l) (Promega)	1 μ l
10 min. 25°C	
50 min. 48°C	

3.6 Statistical analysis

Statistical analysis was performed with GraphPad Prism (version 7.00; GraphPad Software, San Diego, CA). For comparison between two groups, Student's two-tailed unpaired t-test was used. For comparison between more than two groups one-way ANOVA followed by pairwise comparison testing (Bonferroni post-hoc test). $P < 0.05$ was considered significant. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$. Data is presented as mean \pm SD.

4. Results

4.1 Characterization of fecal microbial signatures in a cohort of adult CD patients

4.1.1 HSCT- study cohort and patients' subgroups

This longitudinal study comprised 35 Crohn's disease patients who were unresponsive or intolerant to conventional medications and not amenable for surgery (Age=28.9 years; Disease duration= 9.9 years; CDAI=284.7; Colonic involvement =77%). Enrolled patients underwent autologous hematopoietic stem cell transplantation (AHSCT) and after treatment they were closely followed up. Fecal samples were collected longitudinally overtime starting from (T=0) before the HSCT and every 6 months up to 5 years post-HSCT. To assess intra- and inter-individual variations in microbial community composition, we performed sequencing of 16S rRNA gene amplicons generated from V3/V4 variable regions of fecal (n=133) and biopsy mucosa-associated bacterial DNA (n=186) (**Figure 7**).

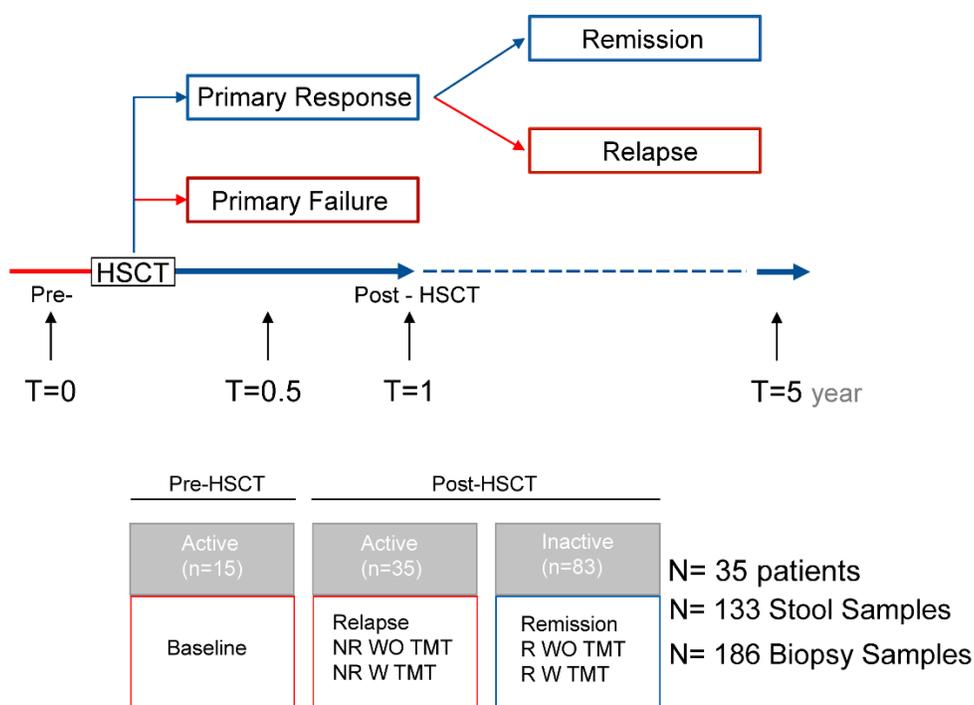


Figure 7 HSCT cohort and patients' groups

Stool and tissue biopsy samples were collected before and after HSCT

4.1.2 IBD patients show temporal fluctuations and interpersonal variations through the course of disease

To characterize changes in microbial community structure between IBD patients at different stages during the course of disease and healthy subjects, we performed a global analysis of all subgroups and stratified patients by disease phenotype into patients at baseline (pre-HSCT), IBD patients during an active and an inactive state of disease (post-HSCT); (n=15, n=35, n=83 respectively). As a control, we included two groups of healthy donors. Control I comprises a group of healthy subjects from Barcelona (n= 8). Control II comprises fecal samples collected longitudinally over 21 consecutive days from 2 healthy subjects in Munich (n=53). The results showed significant separation of microbial profiles between samples collected from IBD patients and healthy subjects (**Figure 8A**). A significant separation of microbial profiles was observed between samples collected from patients during remission, and those collected during active state of the disease, at baseline or during relapse or non-response (p-value=0.00875, p-value=0.0088, respectively). On the other hand, samples collected during active state of the disease (baseline, non-response or relapse) showed overlapping microbial profiles (p-value=0.152). Fecal samples collected from patients during remission showed clear separation from healthy controls (p-value=0.001). Alpha diversity analysis showed that in patients with IBD, the richness of fecal microbiome is reduced compared to healthy subjects. While in clinical and endoscopic remission, remitting patients do not show complete recovery to healthy microbiome state “*normbiosis*”, and they rather show an intermediate level of community richness and diversity between patients with active disease and healthy subjects (**Figure 8B**).

Looking into intra and interpersonal changes of microbial communities in IBD patients, longitudinal samples collected from each patient showed to cluster together. Fluctuations of microbial profiles proved to vary during the course of disease (**Figure 8C**). We used a 10-fold cross validation classifier to predict different disease categories within the human cohort, and generated ROC curves for comparison. A ROC curve or Receiver operating curve is used to show the trade-off between the true positive rate and the false positive rate. The area under the ROC curve is a measure of the model accuracy. An area under the curve (AUC) of 1.0 implies perfect accuracy. The results showed that a model could classify the patients with IBD based on disease activity and based on the clinical outcome post-HSCT with an AUC=0.79 and AUC=0.83, respectively (**Figure 8D**).

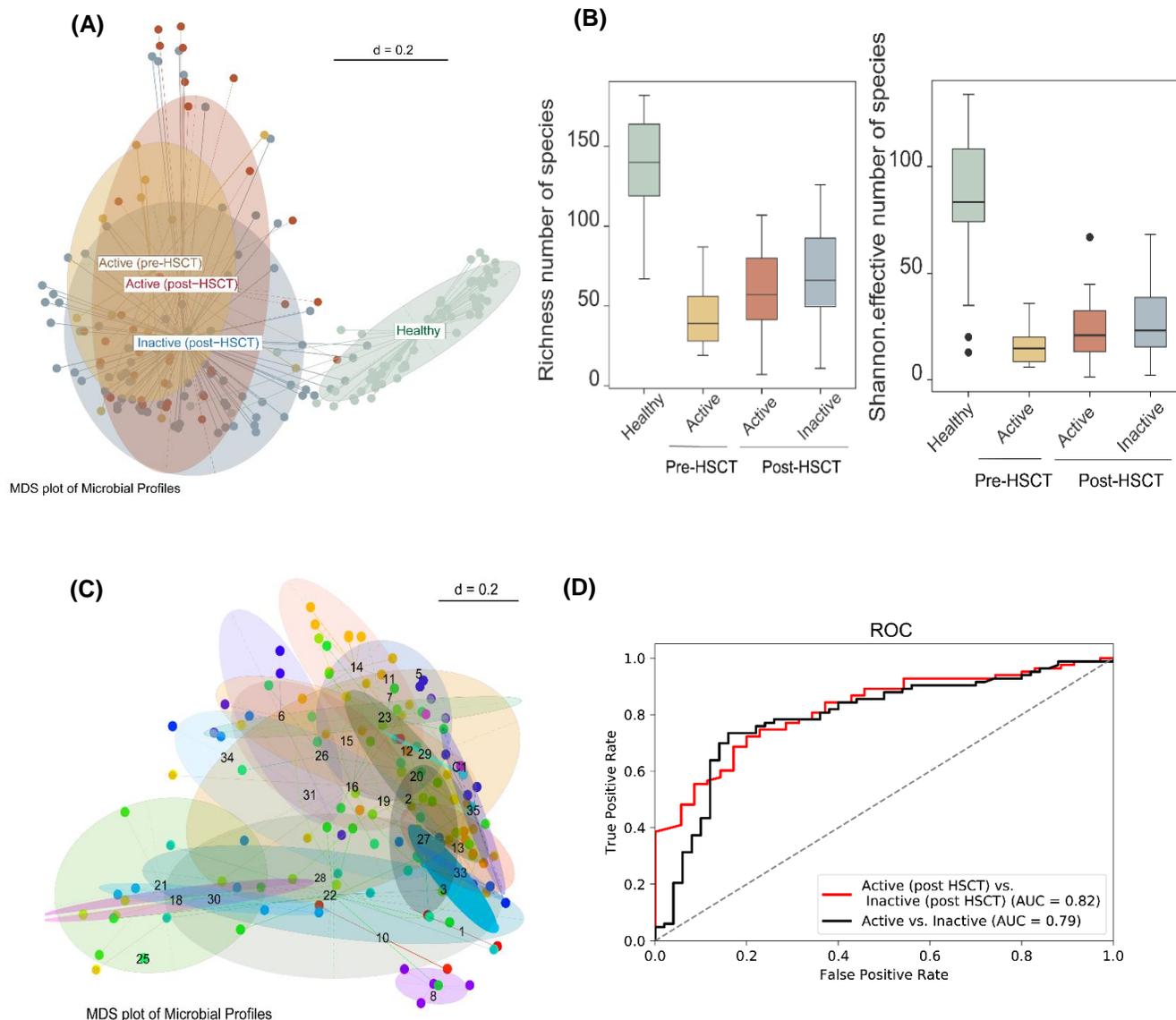


Figure 8 IBD patients show fluctuations through the course of disease

(A) MDS plot of bacterial composition of stool samples collected at different time points and with different clinical outcomes in HSCT-treated IBD patients ($n = 133$ samples, $n = 35$ patients). Each circle represents an individual microbial composition and each color represents disease activity status **(B)** Community richness in IBD patients compared to healthy subjects **(C)** MDS plot of bacterial composition of stool samples collected from IBD patients. Each circle represents an individual microbial composition and each color represents samples collected from the same patient longitudinally over time. **(D)** ROC curve based on 10-fold classifier.

To assess temporal variations in community composition within and between healthy donors and patients with IBD, we computed a matrix of generalized UniFrac distances for all pairwise comparisons of all 194 fecal samples included in the analysis. Generalized UniFrac distance measure has the advantage of detecting compositional changes of both rare and abundant taxa, in contrast to the classical and commonly used Weighted or Unweighted UniFrac distances. To quantify the fluctuations overtime, we compared the UniFrac distances within and between two healthy donors, as well as within selected patients with IBD. The results revealed that healthy donors maintained a relatively stable microbial community structure, where the distance between communities from the same individual was less than the distance between communities originating from two individuals (**Figure 9A**) On the contrary, patients with IBD showed to have vast fluctuations, where the distances between two microbial communities varied significantly (**Figure 9B**). This was particularly more pronounced in patients where there was a change of disease status, for example patient 16 who relapsed at (week 116 post-HSCT).

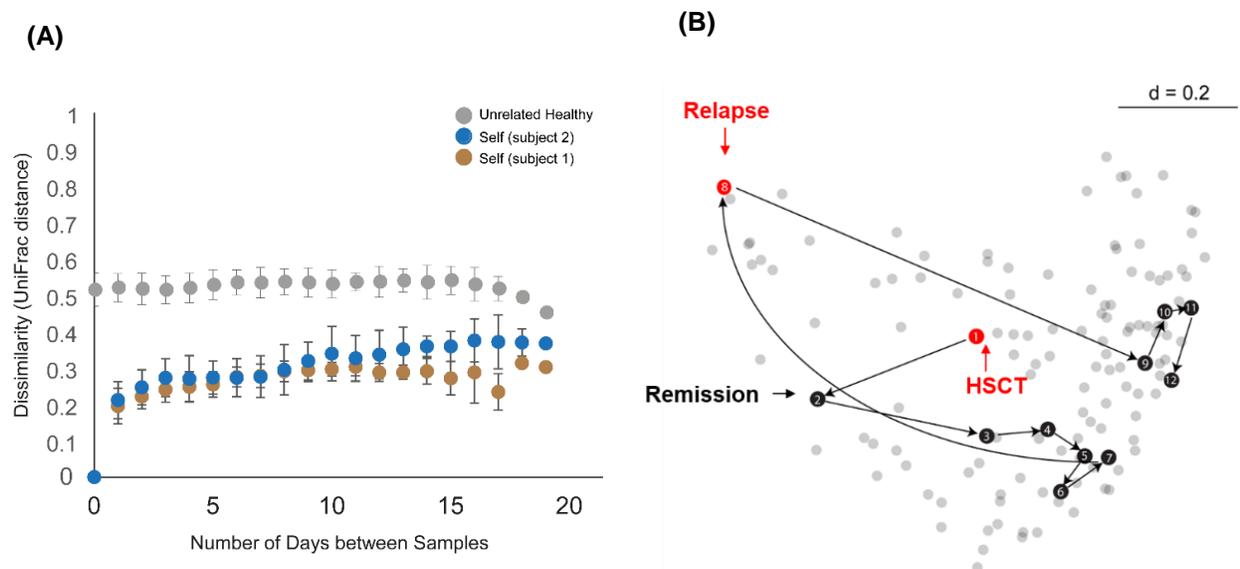


Figure 9 Time course of intra- and interpersonal variations in fecal bacterial communities

(A) Generalized UniFrac distances, from pairwise comparisons within and between all samples obtained from two health individuals given individual plotted as mean values \pm SD., and **(B)** from an IBD patient longitudinally collected samples

4.1.3 Stratification of IBD patients by disease activity shows separation of microbial profiles, with a substantial overlap between subgroups

order to look at specific comparisons between patients' subgroups, we stratified patients by disease activity and looked at microbial profiles and community diversity measures, as well as differentially abundant taxonomic groups. Active patients (n= 50) included patients at inclusion in the study (baseline), non-remission (with or without treatment) and patients who relapsed after initial response. On the other hand, inactive patients (n=83) included patients who responded to HSCT therapy, patients who stayed on long-term remission, or those who went into remission after gaining responsiveness to biologics therapy. Microbiota profiles showed substantial variation between active and inactive patients. Nevertheless, and despite considerable overlap, beta-diversity analysis identified significant differences between patients in active and inactive disease state (**Figure 10A**). In addition, community richness was significantly reduced in active patients (**Figure 10B**).

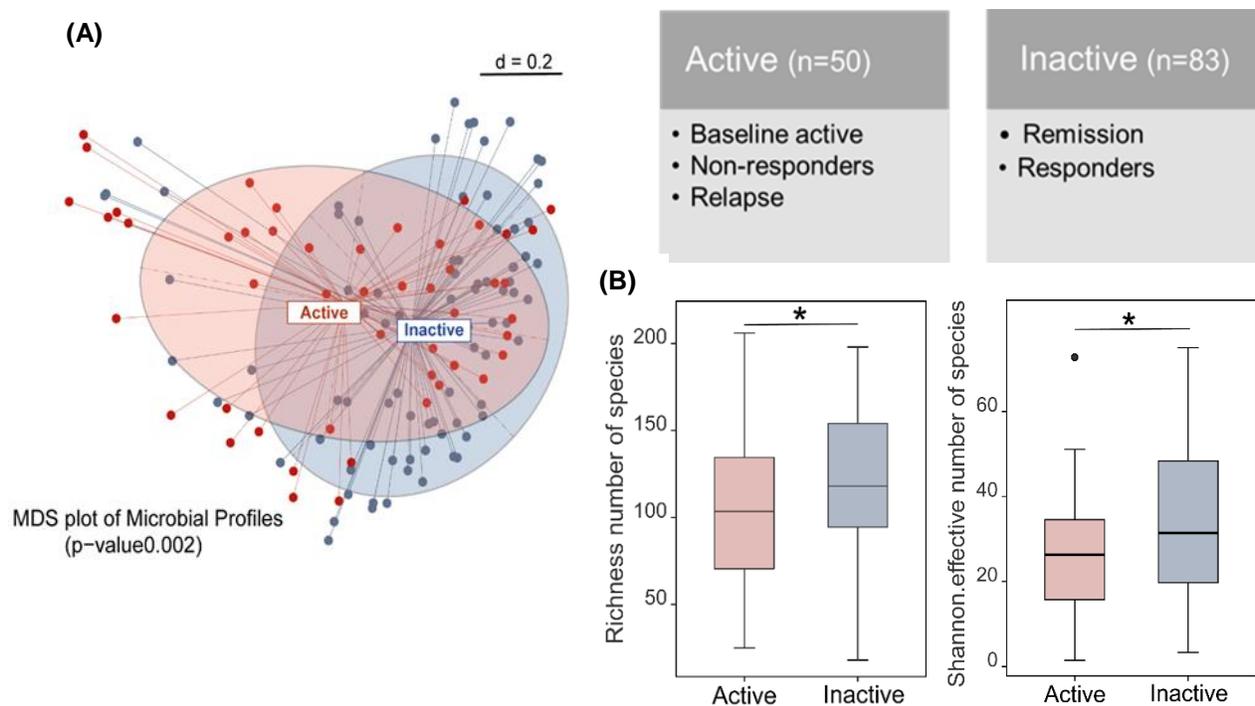


Figure 10 Stratification of IBD patients based on disease activity

(A) Multidimensional scaling (MDS) analysis displaying significant phylogenetic separation between fecal microbial communities of (active) or (inactive) CD patients undergoing HSCT **(B)** Alpha-diversity analysis showing reduced community richness and diversity in active patients

Patients with active disease had a significant increase of Firmicutes and Proteobacteria represented in a significantly increased abundance of families *Lachnospiraceae* and Gammaproteobacteria respectively. A panel of 11 OTUs that were significantly different between active and inactive patients were identified. Patients with active disease had a significant increase in relative abundance of OTUs within *Escherichia-Shigella*, *Parasutterella* and *Faecalibacterium* and significant increase in prevalence of OTUs within *Blautia*, *Escherichia-Shigella*, *Clostridium XIVa*. On the other hand, patients with inactive disease had a significant increase of the phylum *Bacteroidetes* represented in an increased prevalence of *Porphyromonadaceae* and *Rikenellaceae*. At species level, patients at inactive state of the disease had increased prevalence of OTUs within *Parabacteroides*, *Bacteroides* and *Dialister* (**Figure 11**).

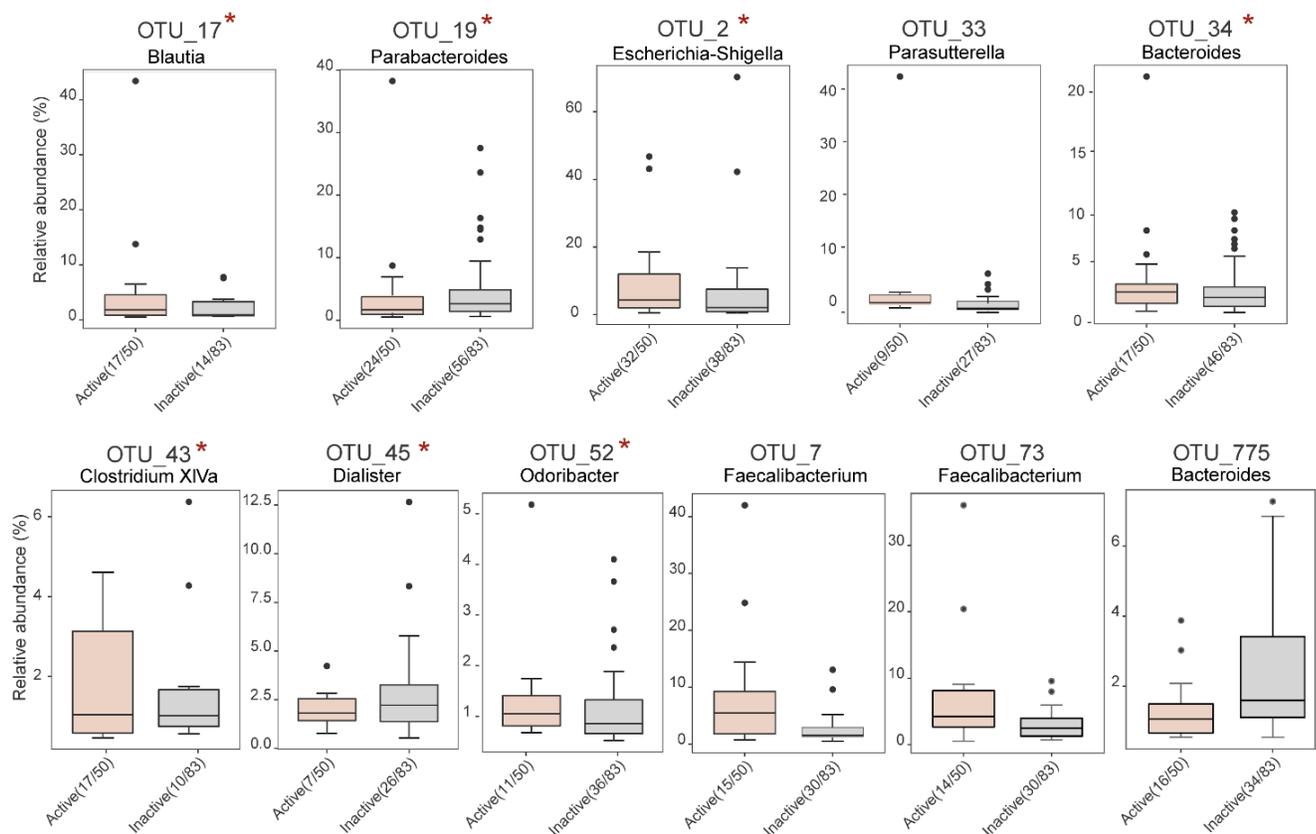


Figure 11 Differential abundant taxa in IBD patients with active disease

4.1.4 Further stratification supported the small but significant differences according to disease activity

Further stratification of patients into active (pre-HSCT), active (post-HSCT) and inactive (post-HSCT) supported the small but significant differences in beta-diversity and community richness according to disease activity. Samples from patients who responded to HSCT or those who remained in long-term remission after HSCT showed significantly increased richness ($p= 0.0225$) and diversity ($p= 0.0399$) in comparison with samples collected at baseline before HSCT (**Figure 12B**). Beta-diversity analysis showed significant separation between inactive patients (post-HSCT) and both active patients at baseline ($p=0.0075$) and active patients post-HSCT ($p=0.0075$) (**Figure 12A**).

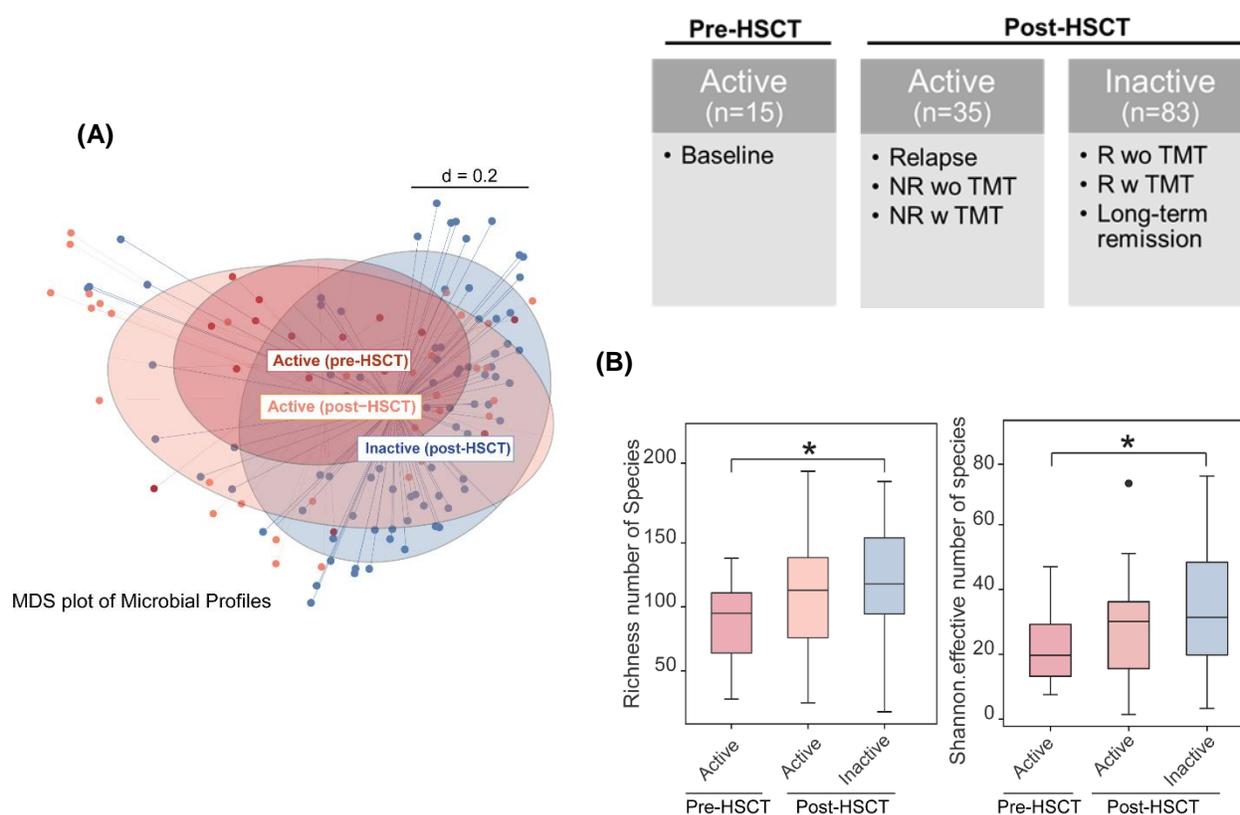


Figure 12 Stratification of IBD patients Microbiota Composition by timepoint

(A) MDS plot of bacterial composition of stool samples collected at different time points stratified based on time before and after HSCT and disease activity status (B) Species richness and diversity in the two groups.

In this comparison, patients who responded to HSCT therapy or remained in long-term remission showed reduced relative abundances of *Escherichia-Shigella* sp.(p=0.0087) sp. and *Ruminococcus gnavus* sp. (p=0.0414). They also showed reduced prevalence of *Megasphaera* sp. (p=0.0318) and *Clostridium* sp. (p=0.0207) and an increased prevalence of *Subdoligranulum* sp. (p=0.0256), *Oscillibacter* sp. (p=0.0256) (**Figure 13**).

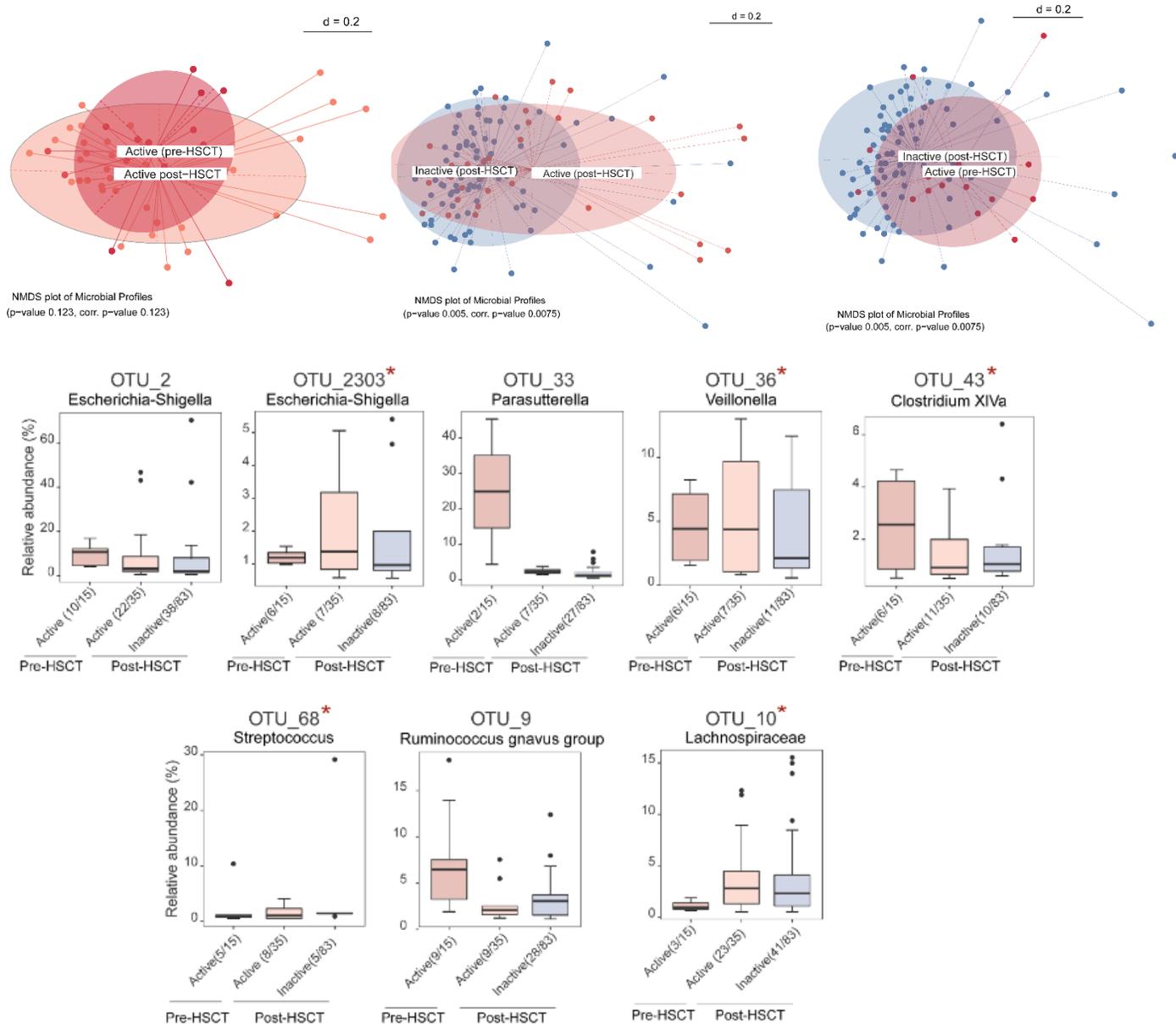


Figure 13 Pairwise comparisons and differentially abundant taxa between the subgroups

4.1.5 Effect of co-administration of biological and immunosuppressive therapy on shifts of gut microbiota in IBD patients post- HSCT

In order to investigate the effect of co-administered drugs on shifts in gut microbial composition, we further stratified samples post- HSCT by the co-administration of additional drugs. This resulted in 5 groups that include patients who never responded to the treatment (NR W TMT), those who regained responsiveness to previous therapies they were refractory to before HSCT (R W TMT), samples from patients who are active but still not on treatment (NR WO TMT) and patients who responded to HSCT and remained disease-free, without co-administration of other drugs (R WO TMT). Interestingly, while there was a significant separation between “disease-free” microbial communities coming from patients in remission; with and without co-administered drugs ($P=0.0267$), there was almost an overlap of the disease-associated microbiota from non-responders with or without treatments and relapsed patients (**Figure 14A,B**). This was additionally showed by pairwise-comparisons of subgroups (**Figure 15**).

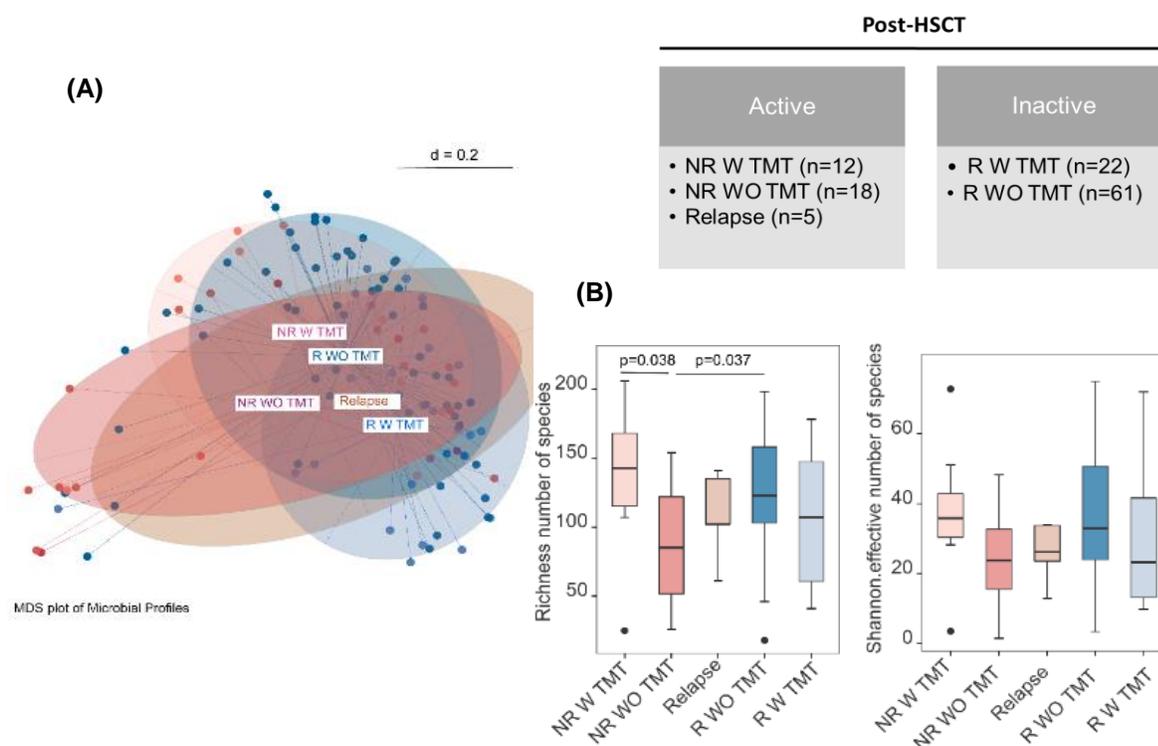


Figure 14 Effect of Co-administration of biological and Immunosuppressive therapy on Shifts of Gut Microbiota in IBD patients post- HSCT

(A) MDS plot of bacterial composition of stool samples collected at different time points stratified based on disease activity status and co-administration of drugs (B) Species richness and diversity in the subgroups. Statistical test: Wilcoxon test (relative abundances) – Fischer’s test (Prevalence).

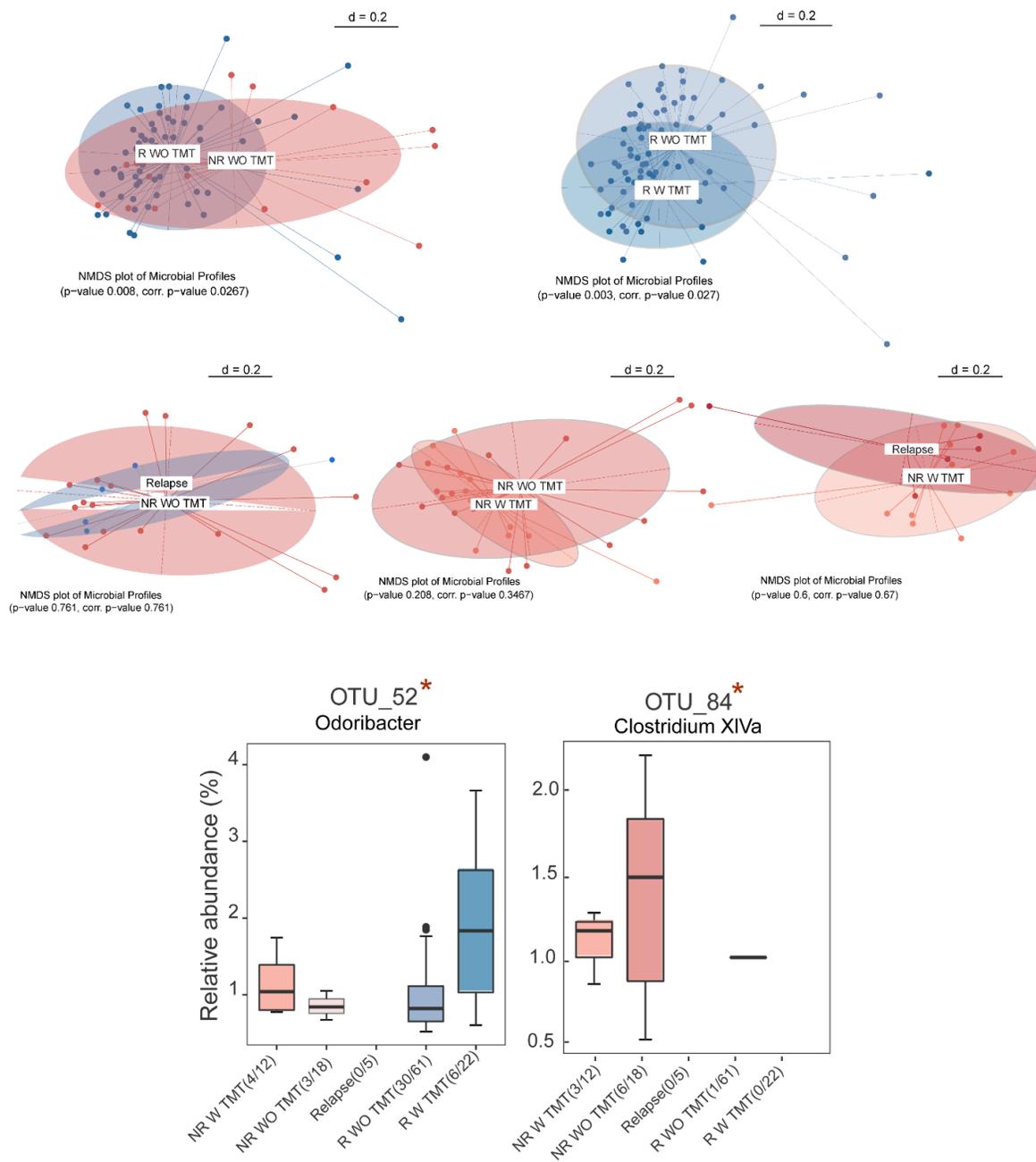


Figure 15 Pairwise comparisons and differentially abundant taxa in subgroups of IBD patients stratified based on administration of additional drugs

4.1.6 Microbial community diversity and richness at baseline could not predict response to HSCT therapy

In an attempt to make predictions of response to therapy based on microbial profiles of patients at baseline, we picked a subset of patients. We were challenged by the fact that we have only 12 patients; where we have stool samples collected at baseline and having follow-up samples. This subset comprised 6 responders and 6 non-responders. A responder is defined as CD patient who reached endoscopic, and clinical remission at Week 52 (1Y) post-HSCT (**Figure 16A**). However, we did not have stool samples collected at 52-weeks' time point from all patients, therefore we used samples collected at 26 weeks post-HSCT for this analysis, instead.

The results showed that there was no significant difference in diversity measures between responder and non-responders at baseline (**Figure 16B**), showing that patients' fecal microbiota community richness and diversity are not sufficient to predict the response to therapy 6-months post-HSCT. Beta-diversity analysis showed no specific clustering of samples from responders or non-responders at baseline (**Figure 16C**). We next looked at variations of bacterial communities before and after HSCT. Quantifying UniFrac distances, as a measure of phylogenetic relatedness between the two groups showed that the effect of treatment did not follow a consistent pattern. In other words, changes were individual-specific and the shifts between samples before and after treatment ranged from short to extreme in both groups, suggesting that fecal microbiota from IBD patients are differentially responsive or sensitive to HSCT treatment, possibly due to the presence or absence of specific bacterial groups (**Figure 16D**).

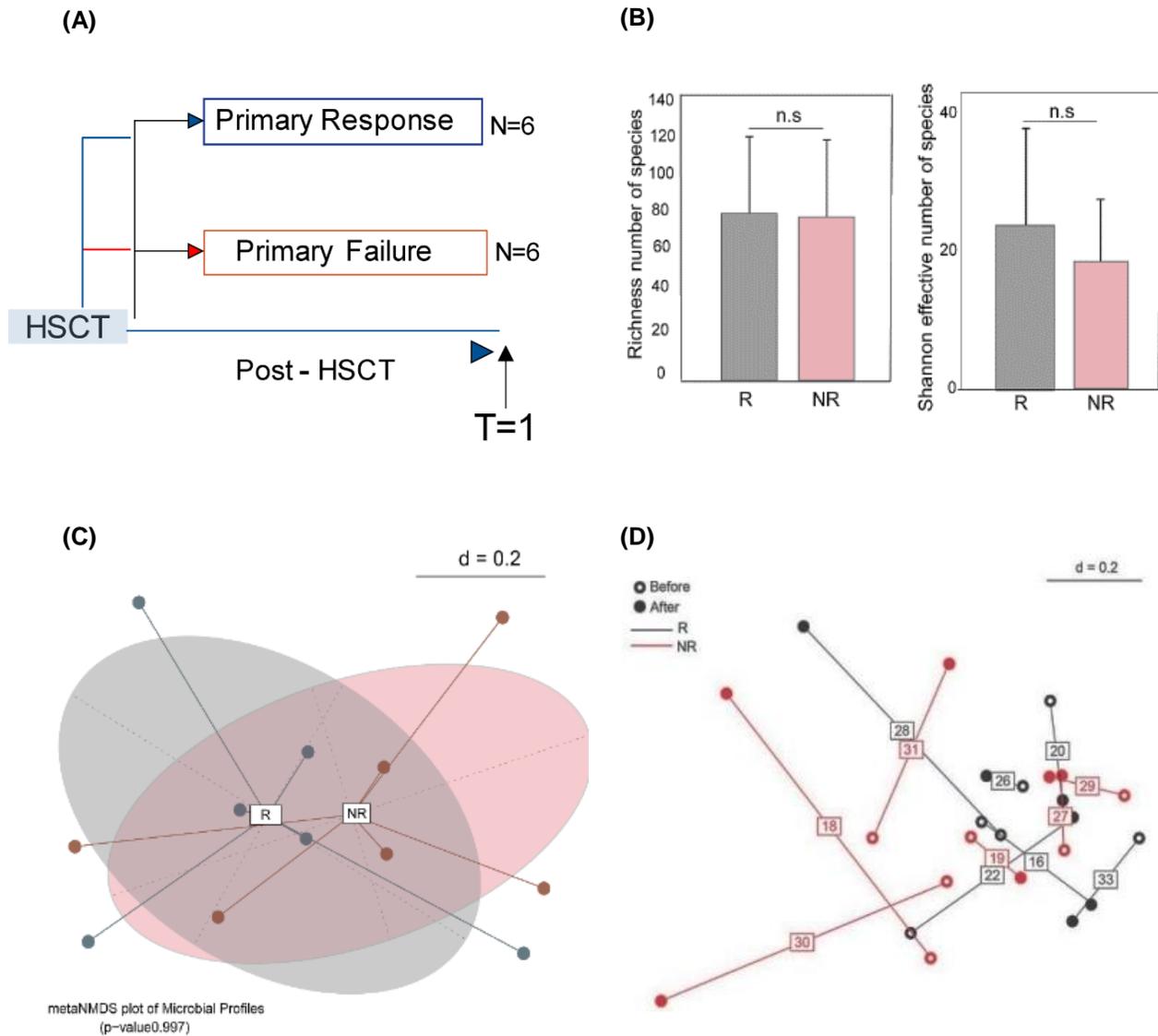


Figure 16 Microbial community diversity and richness at baseline could not predict response to HSCT therapy

(A) Prediction of response to therapy (B) Species richness and diversity in the responder and non-responders at baseline. (C) NMDS plot of bacterial composition of CD patients stool samples collected at baseline of patients who are either in remission (n=6) or at relapse (n=6) one-year post-HSCT (D) UniFrac distances in responders and non-responders before and after HSCT (6 months post-HSCT).

4.1.7 Dynamic fluctuations of microbial community composition overtime in HSCT-treated CD patients

To characterize the microbial composition fluctuations in HSCT-treated patients, we looked at microbial community composition of fecal samples collected overtime. We included a subset of sample collected at baseline (n=14), at 6 months (n=17) and at 1-year post-HSCT (n=14). Beta-diversity analysis showed that there was no clear separation between microbial community profiles overtime (**Figure 17A**). Next, we compared community richness and diversity overtime, between responders and non-responder. The results showed that there was a steady increase of community richness in both groups, however, it did not reach significance (**Figure 17B**). While the effective number of species in the responders group showed a tendency to increase overtime, it reached a plateau in the non-responders 26 weeks post-HSCT (**Figure 17C**). One OTU belonging to the IBD-associated genus *Escherichia-Shigella* showed reduced abundance in responders (**Figure 17D**).

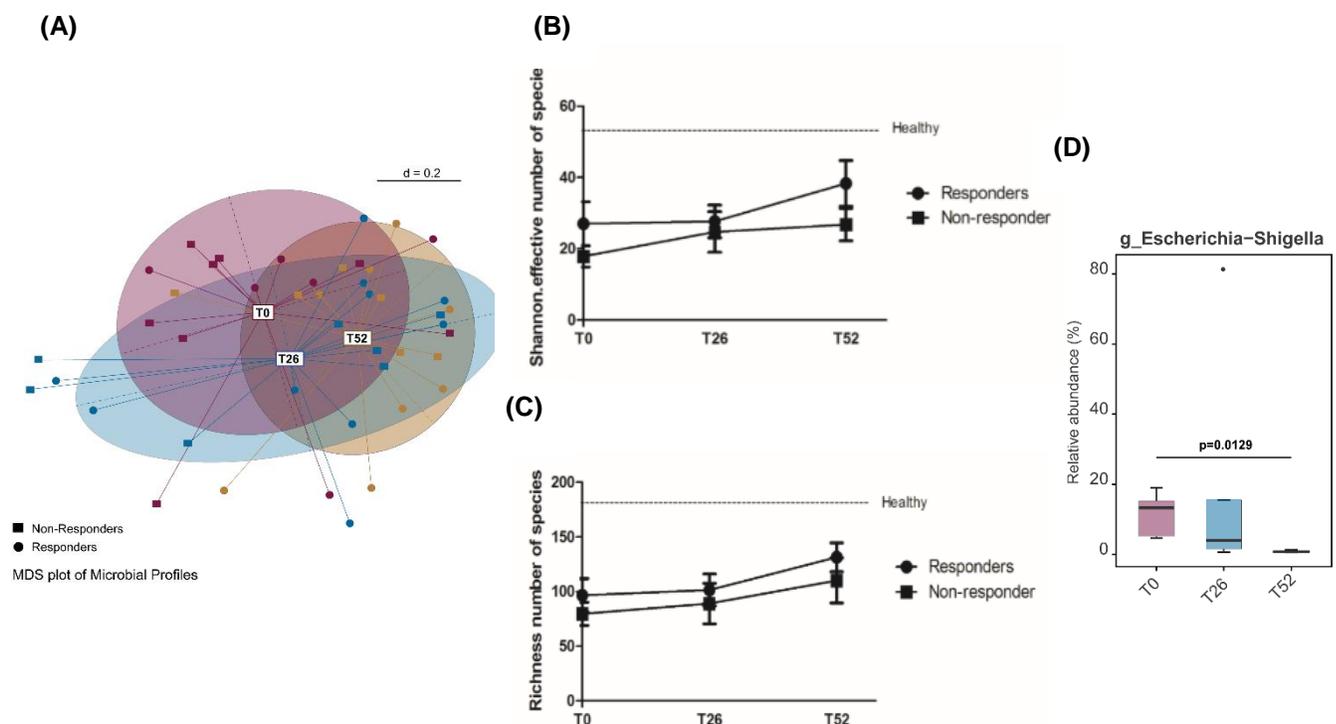


Figure 17 Dynamic fluctuations of microbial community composition overtime in HSCT-treated CD patients

(A) MDS plot of responders and non-responder's microbial composition overtime **(B)** Richness and **(C)** diversity over the course of disease in IBD patients followed overtime after HSCT **(D)** differential abundant taxon in responders.

4.1.8 CD patients cluster into enterotypes characterized by the dominance of signature bacterial taxa

We then looked into the microbiota composition of all patients' samples independently from their clinical characteristics using an unsupervised clustering approach (*k*-means clustering algorithm and Calinski–Harabasz index). Unsupervised beta-diversity analysis showed that all samples from IBD patients formed three robust clusters representing three distinct microbial community structures; that were defined as clusters A, B and C based on the taxonomic composition of samples belonging to each of these 3 clusters. While samples belonging to cluster A showed to have microbiota profile driven by *Bacteroidetes*, cluster B was dominant in *Ruminococcus*, and samples belonging to cluster C were particularly enriched in Firmicutes and gamma or delta Proteobacteria (**Figure 18A and C**). Interestingly, these 3 clusters were associated with different levels of species richness and diversity (**Figure 18B**).

4.1.9 Unsupervised clustering emphasized the heterogeneity of gut microbiota structure in CD patients

To look for microbial signatures possibly associated with patient's disease status, we mapped patients' samples to each of these three clusters and the disease status of patients' samples belonging to each cluster. This analysis showed that patients with active state of disease are distributed amongst these three distinct clusters, with a trend towards clustering of inflamed patients in cluster C which did not reach statistical significance. These findings emphasize the heterogeneity of IBD patients gut microbiome at the taxonomic level. Interestingly, patients proved to show fluctuations in microbial profiles through the course of disease, especially during the transition from non-diseased to a diseased state. The transitions of microbial profiles (distinct clusters) through the course of disease is illustrated in patient 16 as an example in (**Figure 18D**). Samples no.6 represents a sample collected when the patient relapsed after long-term remission 2.5Y post-HSCT.

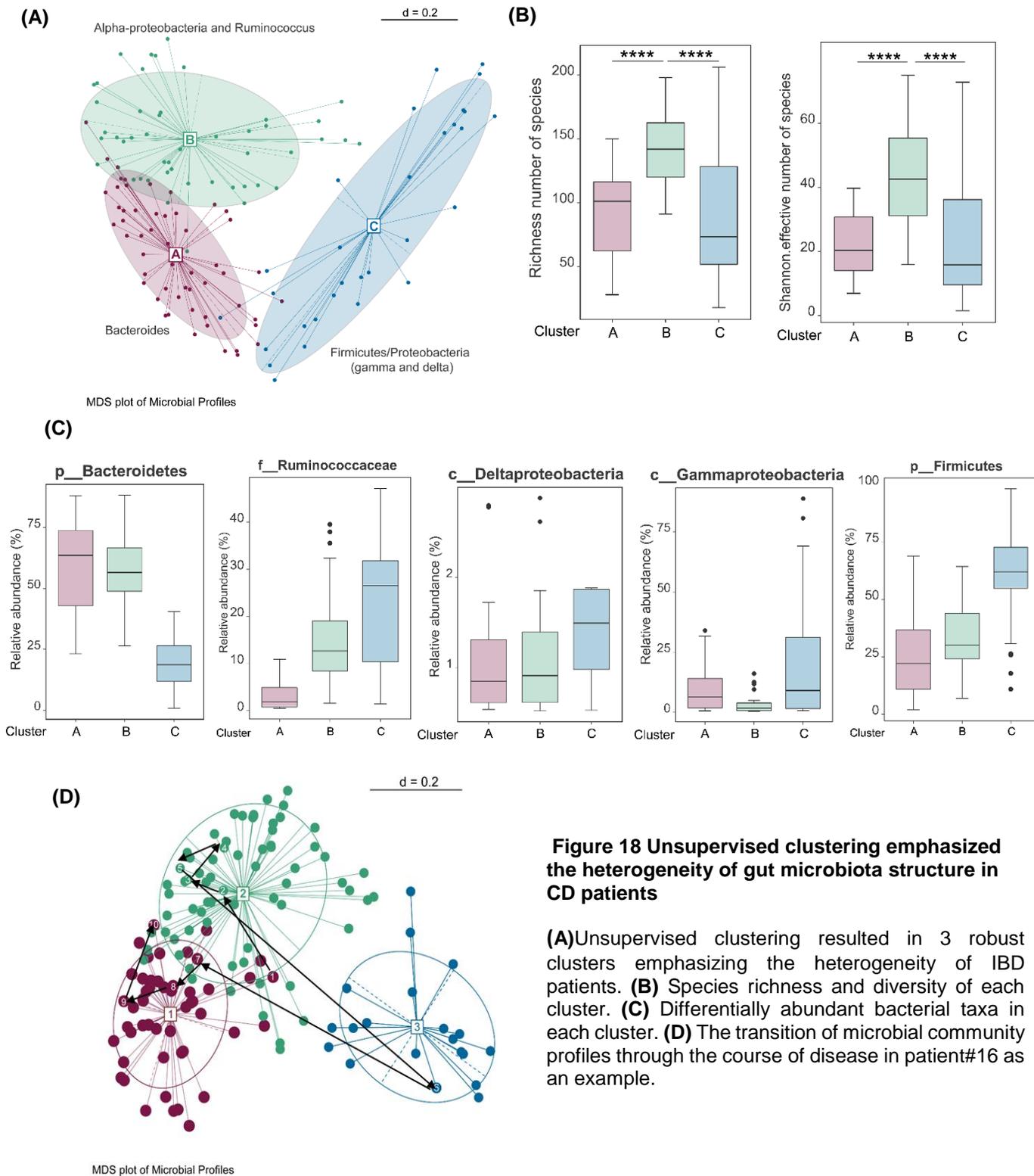


Figure 18 Unsupervised clustering emphasized the heterogeneity of gut microbiota structure in CD patients

(A)Unsupervised clustering resulted in 3 robust clusters emphasizing the heterogeneity of IBD patients. **(B)** Species richness and diversity of each cluster. **(C)** Differentially abundant bacterial taxa in each cluster. **(D)** The transition of microbial community profiles through the course of disease in patient#16 as an example.

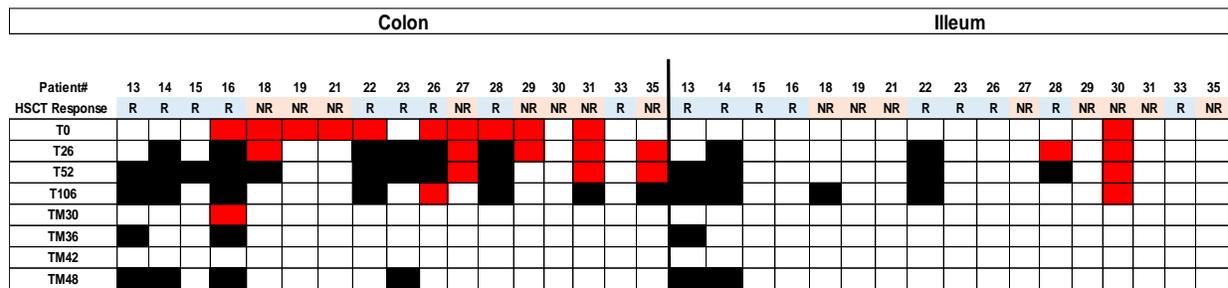
4.2. Mucosal-associated dysbiosis in CD patients treated with HSCT

CD patients included in this study presented with different disease phenotype; ileal, colonic or ileocolonic. In order to characterize specific imbalances in the mucosa-associated microbiota of patients with different disease localizations, we looked into the microbial profiles of patients presenting with colonic or ileal phenotypes separately. Patients' samples included in the analysis are summarized in **(Figure 19A)**. Clinical characteristics of patients with colonic or ileal involvement are depicted in **(Figure 19B)**. Simple Endoscopic Score for Crohn's Disease (SES-CD) was used to describe disease severity or response to therapy globally or in the involved areas only.

4.2.1. Characterization of colonic mucosa-associated dysbiosis

First, we performed a global analysis of all subgroups and stratified patients by disease activity and looked at microbial profiles and diversity measures, as well as differentially abundant taxonomic groups. Active patients included patients at inclusion in the study (baseline), non-responders (with or without treatment) and patients who relapsed after initial response. On the other hand, inactive patients included patients who responded to HSCT therapy, patients who stayed on long-term remission, or those who went into remission after gaining responsiveness to biologics therapy. Activity is determined based on clinical and endoscopic assessment. Beta-diversity analysis of colonic tissue biopsies collected during active or inactive state of disease showed a significant separation of microbial profiles **(Figure 20A)**. Alpha-diversity analysis showed reduced microbial community richness in patients with active disease **(Figure 20B)**. In order to look at bacterial taxa that are differentially abundant between active and inactive colonic mucosa-associated microbiota, we constructed a heatmap of the most abundant OTUs represented as an average of relative abundances of all samples belonging to the same group. Specific bacterial taxa showed to be more abundant or prevalent in inflammation. Tissue biopsies from patients in active disease showed to have higher relative abundance of *Escherichia-Shigella*, *Halomonas*, *Streptococcus* and *Schewanella*. On the contrary, they showed reduced relative abundance of members belonging to the genera *Fecalibacterium*, *Blautia*, *Akkermansia* and family *Lachnospiraceae* **(Figure 20C)**.

(A)



(B)

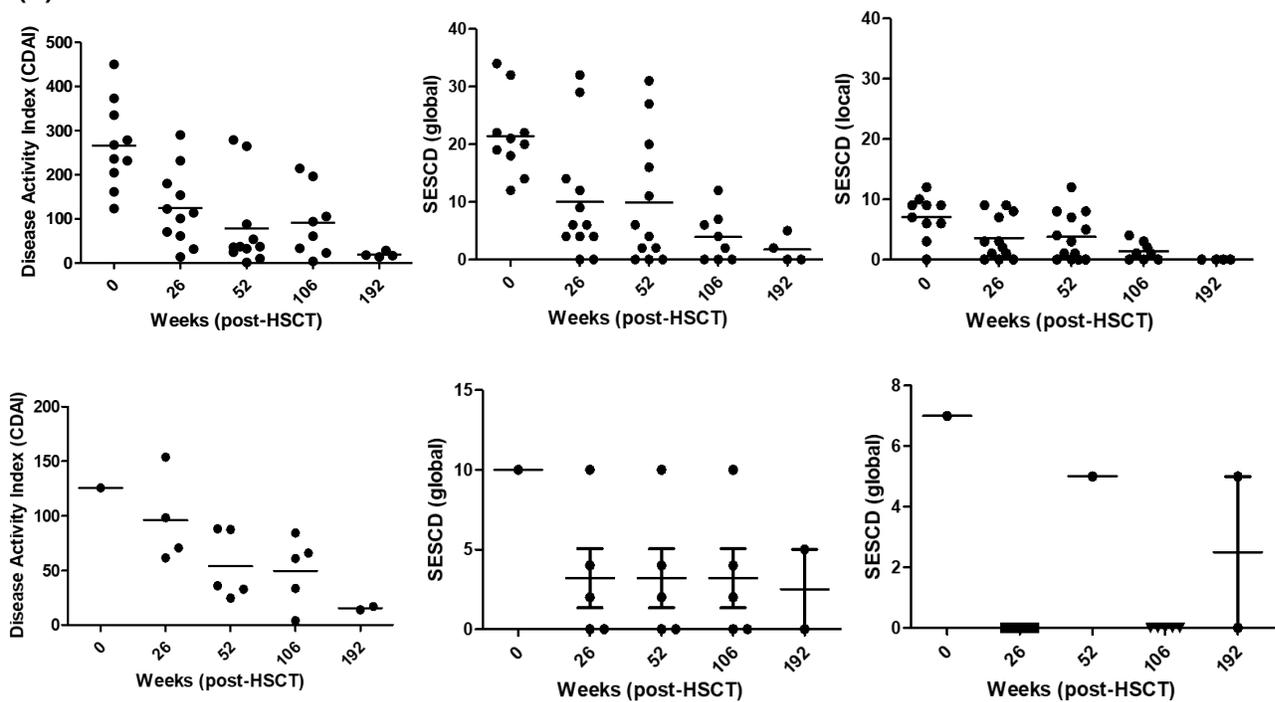


Figure 19 Mucosal-associated microbiota samples included in the analysis

(A) Summary of patients' samples included in the analysis. Patients present with ileal, colonic or ileocolonic phenotype. Red blocks represent samples collected from the involved tissue during an active state of the disease, while black blocks represent samples collected from the involved tissues during an inactive state of the disease. (B) Clinical characteristics of patients included in the analysis.

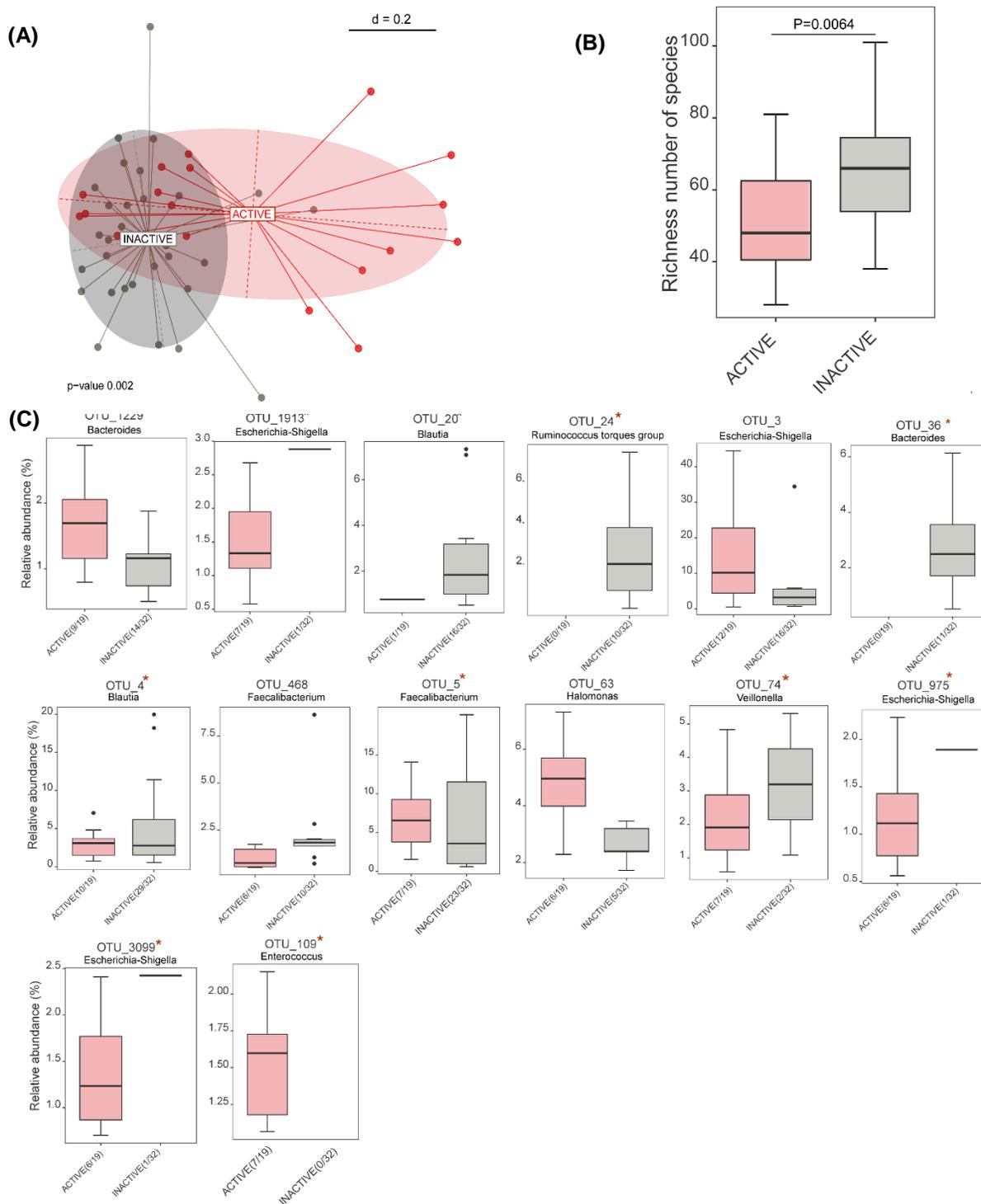


Figure 20 Characterization of colonic mucosa-associated dysbiosis

(A) MDS plot of microbial community composition across "involved" colon samples **(B)** Microbial community richness in mucosa-associated microbial community in colon tissues during active and inactive states of the disease. **(C)** Differential abundant taxa in colon mucosal biopsies of active and inactive CD patients undergoing HSCT * Phylotypes significantly more prevalent.

To assess temporal and inter-individual variations in community composition in mucosa-associated microbiota, we looked at beta-diversity analysis using Bray-Curtis distance metric. The results revealed that although patients with IBD showed to have vast fluctuations, where the distances between two microbial communities varied significantly during the course of disease, samples originating from the same patient seemed to cluster together in most cases (**Figure 21A**). Mucosa-associated microbiota at baseline showed to vary significantly between active patients. Overtime, more patients responded to HSCT therapy and the variation among microbial communities seemed to decrease moving to uninfamed status (**Figure 21B**). Due to the considerable inter-individual variation between patients, we looked into the temporal alterations of microbial community composition in individual patients, overtime. While each patient showed different dynamics through the course of disease, some bacterial groups showed to be relevant in most cases These include *Escherichia-Shigella*, *Halomonas*, *Shewanella*, *Ruminococcus gnavus* and *Streptococcus* (**S 1, S 2, S 3.**).

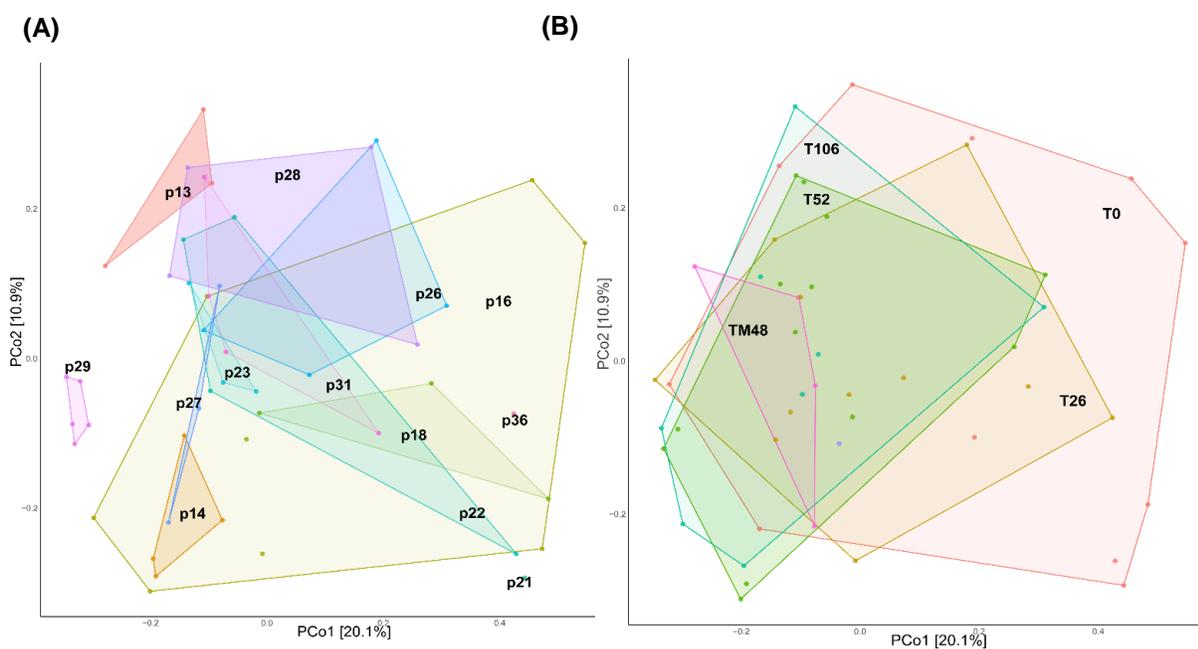


Figure 21 Temporal and Inter-Individual Variations in colonic-mucosa associated microbiota

Principal coordinate analysis (PCoA) profile of microbial community composition across “involved” colon samples using Bray-Curtis distance metric. The percentage of variation explained by PC1 and PC2 are indicated in the axes **(A)** Samples grouped per patient **(B)** Samples grouped per time-point before or after HSCT.

4.2.2. Characterization of ileal mucosa-associated dysbiosis

We applied the same analysis on the ileal mucosa-associated microbiota in patients with ileal or ileocolonic phenotype. In this case, we were limited by a lower number of samples (n= 6 patients, 19 samples). Active patients included patients at inclusion in the study (baseline), non-responders (with or without treatment) and patients who relapsed after initial response. Inactive patients included patients who responded to HSCT therapy, patients who stayed on long-term remission, or those who went into remission after gaining responsiveness to biologics therapy. Activity is determined based on clinical and endoscopic assessment. Beta-diversity comparison of ileal tissue biopsies collected during active or inactive state of disease showed a significant separation of microbial profiles (stronger than in case of colonic tissue) (**Figure 22A**). Alpha-diversity analysis showed significantly reduced microbial community richness in active compared to inactive patients (**Figure 22B**). In comparison to colonic tissue biopsies, the number of bacterial species is much lower in ileal mucosa.

We then looked at bacterial taxa that are differentially abundant between active and inactive ileal mucosa-associated microbiota. Interestingly, we observed a remarkable increase of *Bifidobacterium* and the *Lactobacillus* group in active patients compared to inactive patients. Additionally, an increase of *Veillonella*, *Escherichia-Shigella* and *Acidominococcus*, *Anaerostipes* and *Klebsiella* was also observed. Compared to inactive patients, members belonging to *Faecalibacterium* genus were sharply decreased in the ileal biopsy specimens of the active CD patients (0% versus 6.5%). The same trend was observed in *Fusobacterium* and *Blautia* genera (**Figure 22C**).

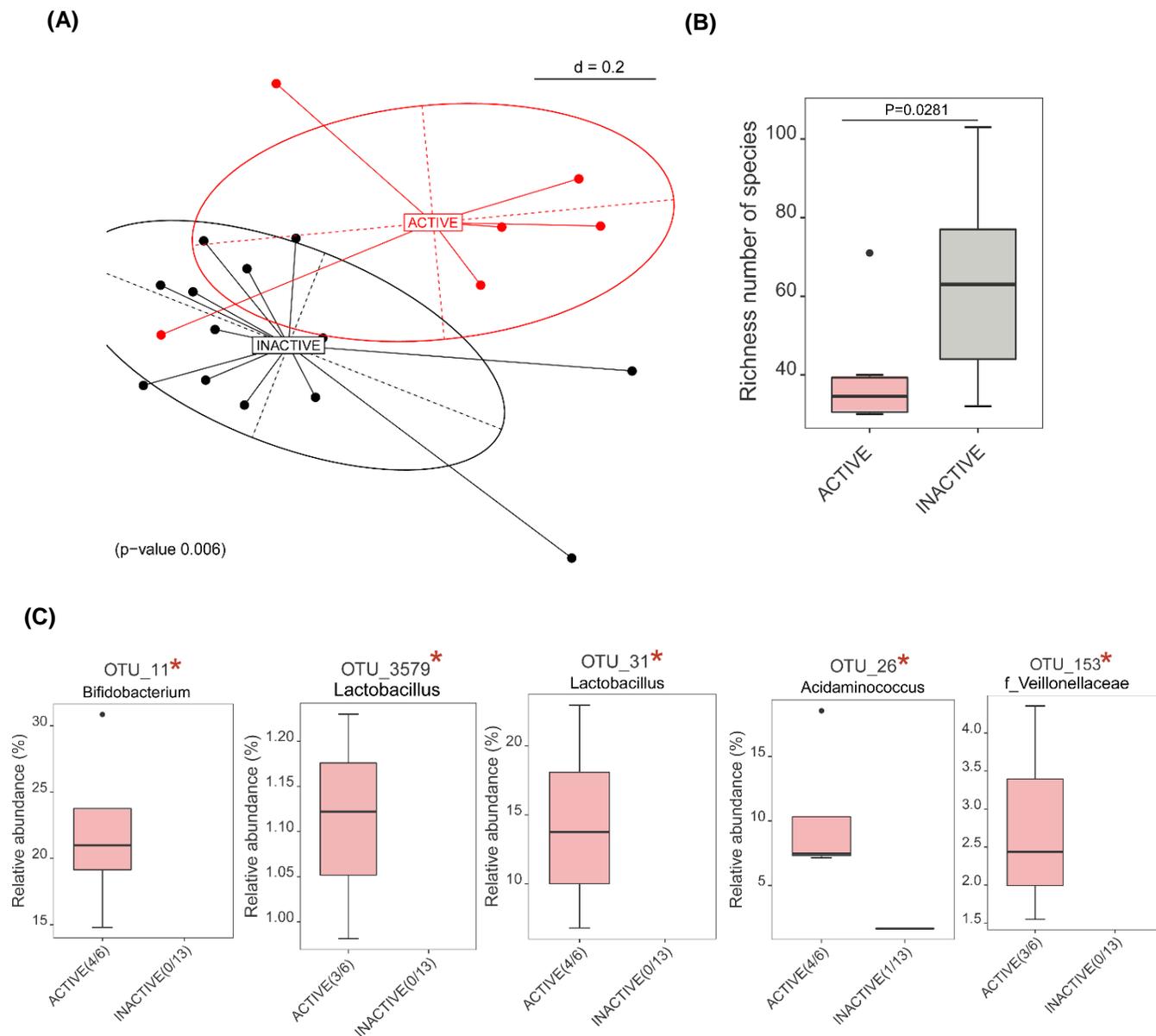


Figure 22 Characterization of ileal mucosa-associated dysbiosis

(A) MDS plot of microbial community composition across “involved” ileal biopsy samples **(B)** Microbial community richness and diversity in mucosa-associated microbial community in ileal tissues during active and inactive states of the disease. **(C)** Differential abundant taxa in ileal mucosal biopsies of active and inactive CD patients undergoing HSCT.

Similarly, to assess temporal and inter-individual variations in community composition in mucosa-associated microbiota, we looked at beta-diversity analysis using Bray-Curtis distance metric. The results showed that samples originating from the same patient seemed to cluster together (**Figure 23A**). However, the number of samples was too low to make further conclusions on the temporal variations (**Figure 23B**). We then looked into the temporal alterations of microbial community composition in individual patients, overtime. While each patient showed different dynamics through the course of disease, some bacterial groups showed to be relevant in most cases (**S 4, S 5, S 6**). These include (as described earlier), *Lactobacillus*, *Bifidobacterium*, *Escherichia-Shigella* and *Acidaminococcus*. It is noteworthy that these findings are limited by the very low number of samples collected during an active state of the disease (n=5; 4 of which originate from the same patient).

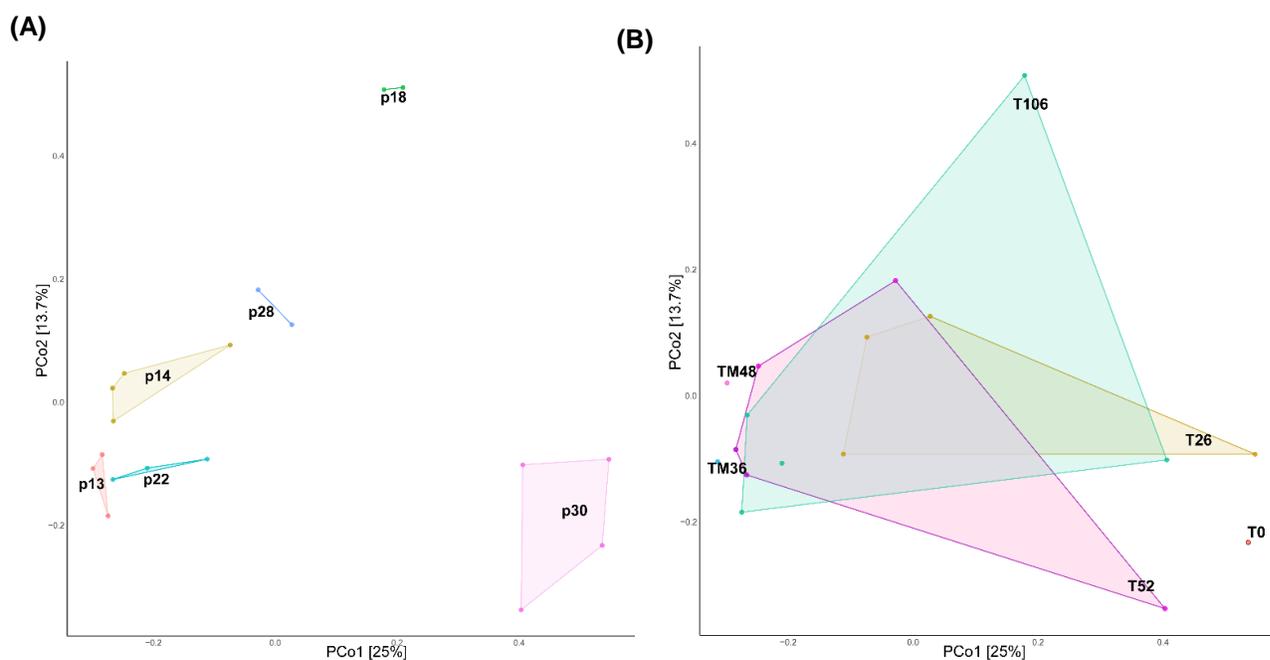


Figure 23 Temporal and Inter-Individual Variations in ileal mucosa-associated microbiota

Principal coordinate analysis (PCoA) profile of microbial community composition across “involved” ileal samples using Bray-Curtis distance metric. The percentage of variation explained by PC1 and PC2 are indicated in the axes **(A)** Samples grouped per patient **(B)** Samples grouped per time-point before or after HSCT.

4.3 Selection of patients for transfer experiments in germfree mice

In order to address the functional impact of microbial communities, we developed a humanized IBD model by colonizing germfree *Il-10^{-/-}* mice with fecal samples from CD patients. From the HSCT-treated IBD cohort, we selected 3-paired samples from CD patients representing different disease activities and community clusters (**Figure 24B**). These 3 patients represented 3 different disease-course scenarios (**Figure 24A**). As shown patient 16 responded to HSCT and maintained in remission for 2Y, before he relapsed 6 months later (2.5Y post-HSCT). She was treated then with anti-TNF therapy and maintained a long-term remission. Patient 28 is also a responder who maintained remission for 2Y post-HSCT. On the other hand, patient 27 did not respond to HSCT. Stool samples were collected from the three donors and were used to orally gavage GF IL-10 deficient and WT mice. We analyzed donor stool samples by 16S rRNA gene sequencing to characterize the shifts in microbial communities in terms of taxonomic composition and diversity indices (**Figure 25**).

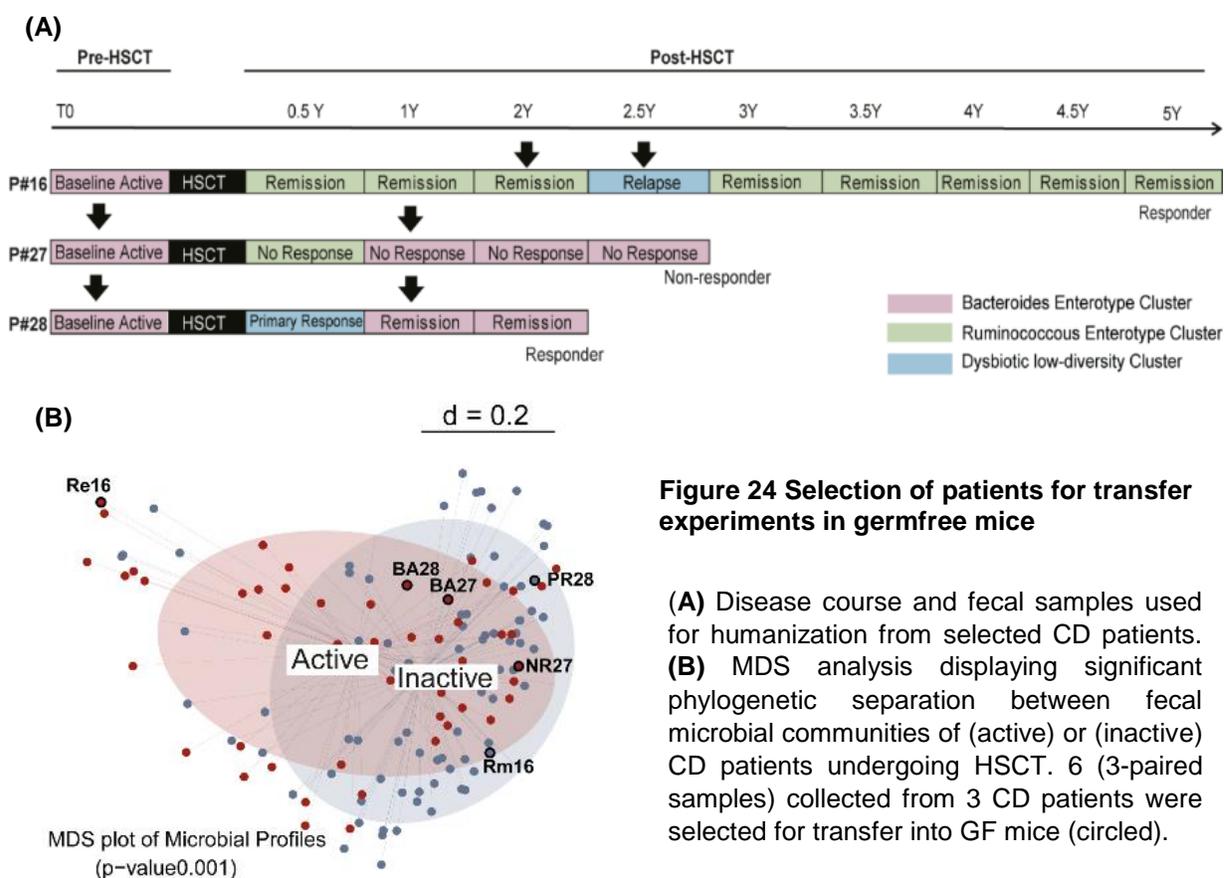


Figure 24 Selection of patients for transfer experiments in germfree mice

(A) Disease course and fecal samples used for humanization from selected CD patients. (B) MDS analysis displaying significant phylogenetic separation between fecal microbial communities of (active) or (inactive) CD patients undergoing HSCT. 6 (3-paired samples) collected from 3 CD patients were selected for transfer into GF mice (circled).

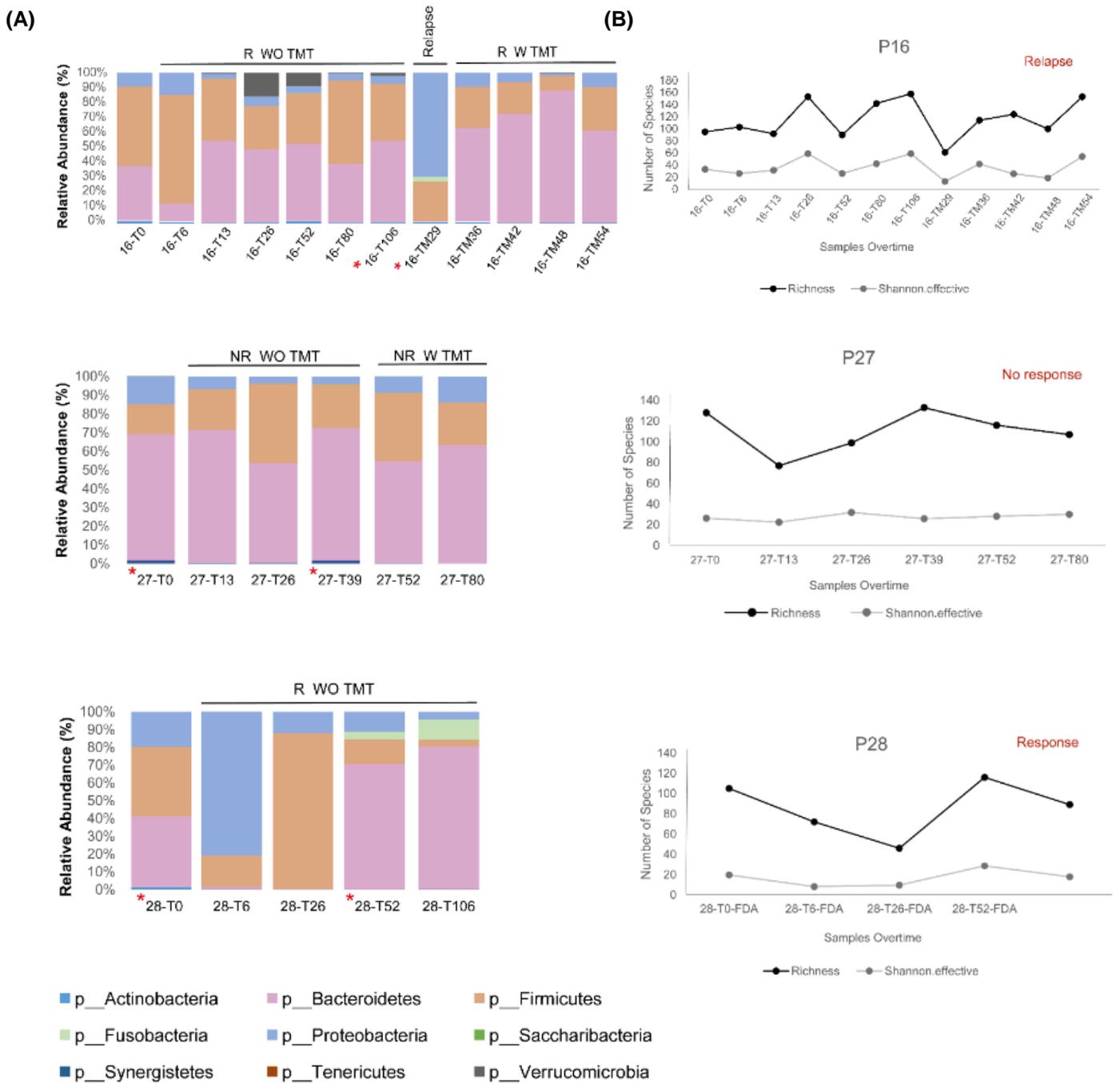


Figure 25 Patient’s clinical course and microbial composition of fecal samples collected before and over-time after HSCT

(A) Patient’s clinical course and microbial composition of fecal samples collected before and over-time after HSCT. Two fecal samples were selected for transplantation in GF mice **(B)** Richness and Diversity of fecal samples over-time.

4.4 Functional characterization of microbial signatures in IBD using gnotobiotic humanized mice

We colonized germfree wild-type (WT) and *Il10*^{-/-} mice at the age of 8 weeks with CD patient-derived microbiota for 4 weeks using 3-paired samples from 3 CD patients (**Figure 26A,B**). Humanized *Il10*^{-/-} mice developed inflammation in the cecum tissue as shown by histopathological evaluation and recapitulated the disease phenotype of their respective human donors (**Figure 26C**). Furthermore, our data showed that mice colonized with fecal microbiota from patients at baseline (pre-HSCT) developed milder inflammation compared to mice colonized with microbiota from patients with active state of disease (post-HSCT). In contrast, *Il10*^{-/-} mice colonized with microbiota from patients at initial clinical response or at long-term remission remained disease-free (**Figure 26D**). Wild-type mice remained disease-free (not shown).

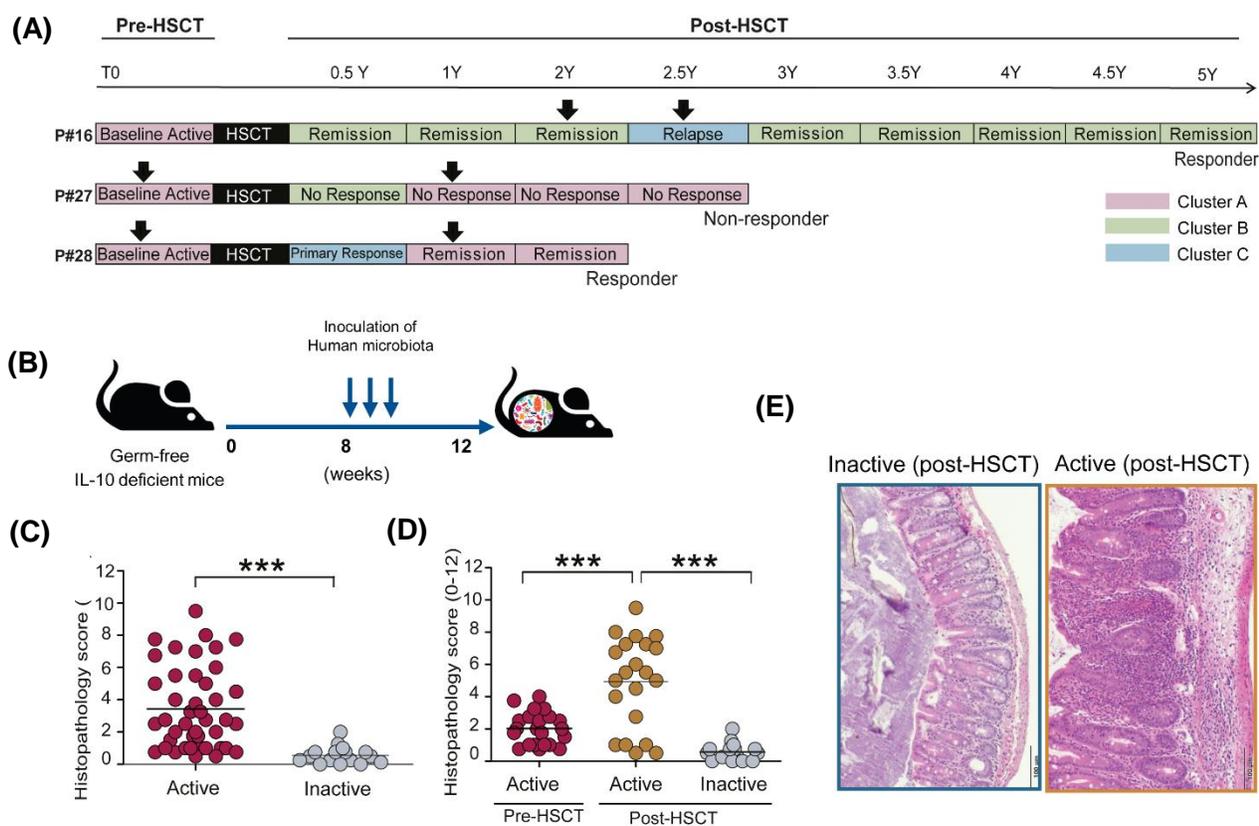


Figure 26 Functional characterization of microbial signatures in IBD using gnotobiotic humanized mice

(A) CD patient samples selected for transfer into GF mice (indicated by arrowheads). **(B)** Experimental setup. GF WT and *Il10*^{-/-} mice were colonized at the age of 8 weeks with selected CD patient-derived microbiota for 4 weeks **(C)** Colonization of germfree *Il10*^{-/-} mice with fecal microbiota, stratified by disease activity of human donors, and **(D)** sub-stratification by time before or after HSCT **(E)** Representative HE staining of cecum tissue in GF *Il10*^{-/-} mice, mice colonized with microbiota from patients with active or inactive state of disease. * $p \leq 0.05$, ** $p < 0.01$, **** $p < 0.001$. One-way ANOVA followed by *Bonferroni* post-hoc comparisons test.

4.4.1 Morphometric and immunologic changes in humanized mice

Development of inflammation was evidenced by gross morphological changes indicative of immune activation, including cecal atrophy and significantly increased MLN and spleen to body weight ratios (**Figure 27A**). Gene expression analysis of total cecal tissue showed significantly increased interferon gamma (IFN- γ) and interleukin-6 (IL-6) expression in humanized mice colonized with fecal microbiota from CD patients during active disease. A trend towards increased expression of tumor necrosis factor alpha (TNF- α) did not reach statistical significance (**Figure 27B and C**). Pearson's correlation coefficient was used to assess a possible linear relationship between levels of inflammation and expression of inflammatory cytokines (**Figure 27D**).

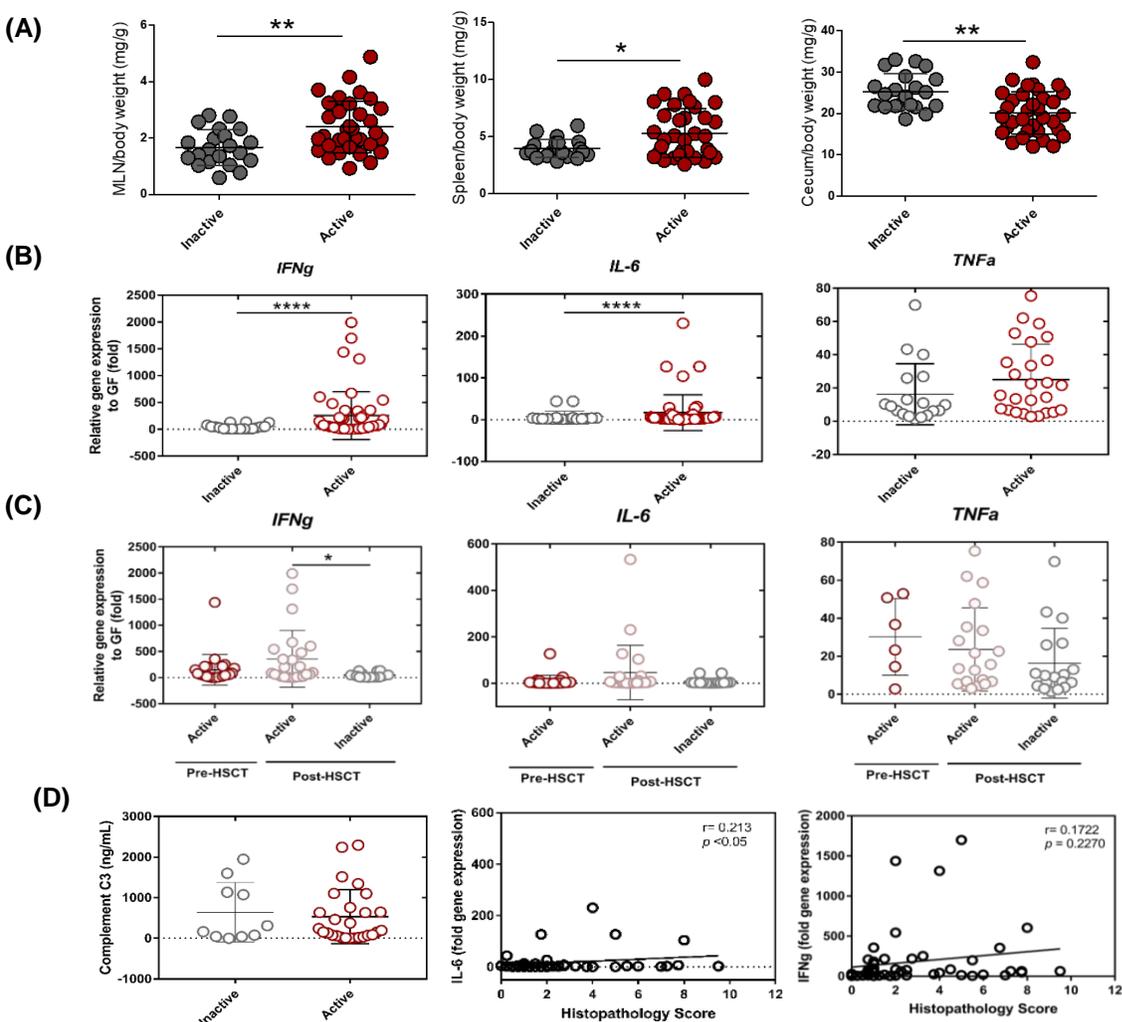


Figure 27 Morphometric and immunologic changes in humanized mice

(A) MLN, Spleen and cecum/body weight ratios. **(B, C)** Gene expression of pro-inflammatory cytokines. **(D)** C3 levels and Pearson's correlation between inflammation and gene expression. Gene expression values correspond to fold change values of $\Delta\Delta CT$. Using Student's t-test. * $p \leq 0.05$, ** $p < 0.01$, **** $p < 0.001$.

4.4.2. Humanized mice reflect the dysbiotic features of their respective donors

16S rRNA gene sequencing analysis showed that humanized mice reflected the microbial community composition and diversity levels of their respective donors. Beta-diversity analysis of microbial profiles clearly demonstrates that the community profiles of the recipient mice (with halos) mimics the microbiota composition of the corresponding donor microbial profiles (with arrows) in 6 distinct community clusters (**Figure 28A**). In addition, humanized mice and corresponding human donors showed similar phyla composition (**Figure 28B**). Alpha diversity analysis showed that humanized mice followed the patterns of the human donors in terms of community richness and diversity (**Figure 28C and D**).

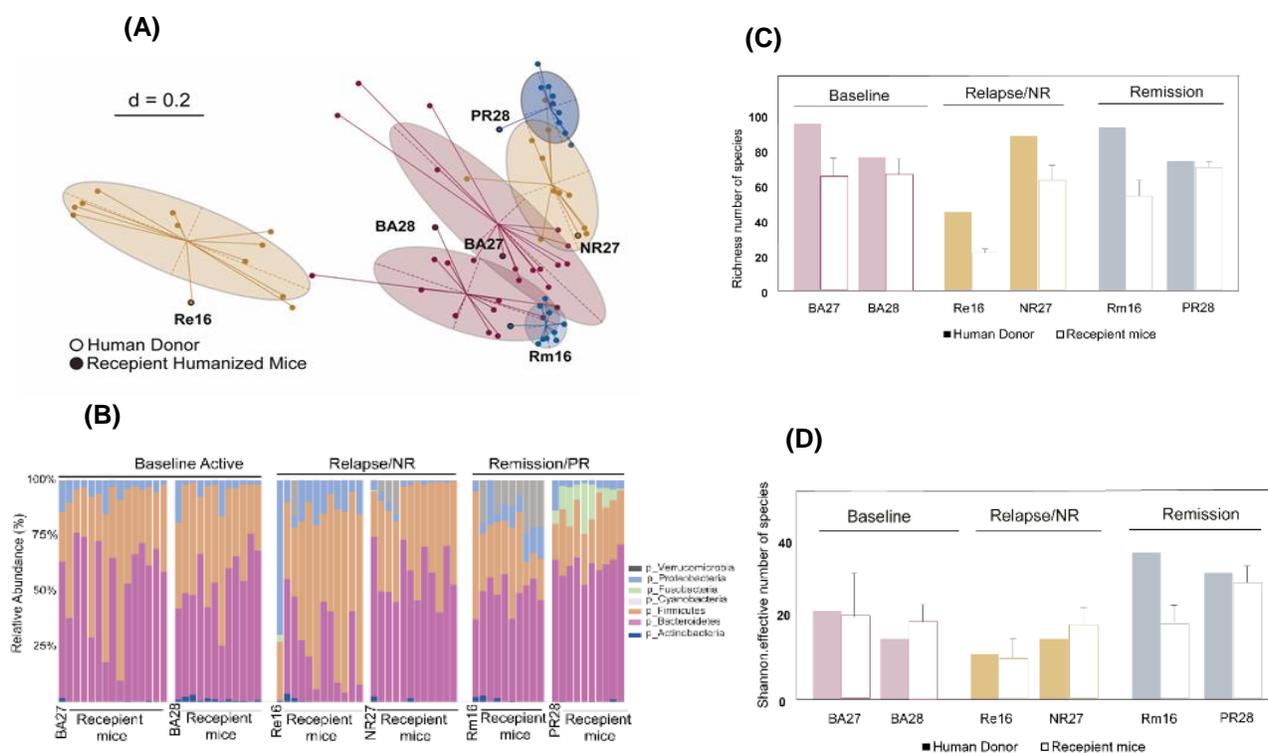


Figure 28 Humanized mice reflect the dysbiotic features of their respective donors

(A) MDS of microbial profiles showing that humanized mice reflect the individual features of their respective human donors. **(B)** Taxonomic composition of human donors and respective humanized mice showing similarities at the phylum level. **(C)** Richness and **(D)** diversity of human donors and respective humanized mice. Abbr.: BA, Baseline active; PR, Primary responder; NR, non-responder; Rm, Remission; Re, Relapse.

4.4.3 Alteration of bile acid metabolism in mice associated with dysbiotic human microbiota

To characterize the changes in metabolomics profiles, we performed targeted metabolomics to measure the concentrations of 19 primary, secondary as well as conjugated bile acids in human donors and corresponding humanized mice. Principal component analysis (PCA) emphasized the difference between metabolic profiles in humans and mice due to host-specificity in bile acid metabolism and conjugation (**Figure 29A**). The inter-individual variation in the metabolic profiles of the human donors reflected the variation previously seen in microbiota profiling. Additionally, the metabolome of humanized mice originating from the same human donor clustered together. Interestingly, humanized mice colonized with microbiota originating from CD patients with active disease (baseline, non-responders or relapse) appeared to have more restricted metabolic profiles compared to humanized mice associated with microbiota from CD patients in remission (**Figure 29B**).

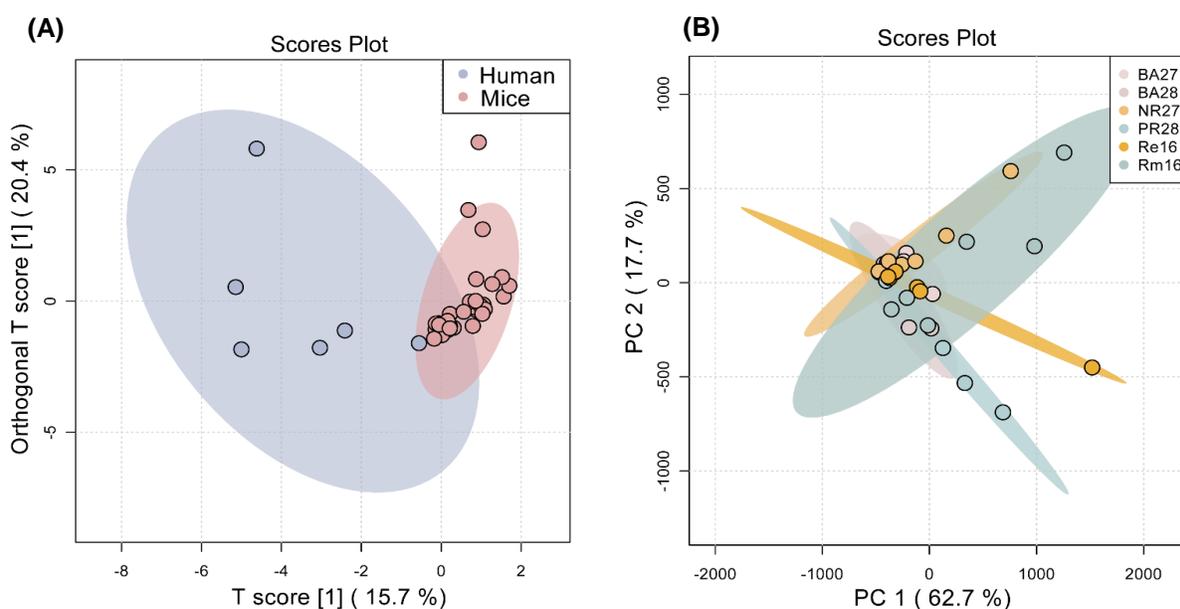


Figure 29 Alteration of bile acid metabolism in mice associated with dysbiotic human microbiota

(A) OrthoPLSDA (OPLSDA) plot of bile acid profiles in human donors and humanized mice **(B)** Score plot of OPLS-DA. The different points in the score plots refer to individual mice, and the samples are color-coded according to the CD donor disease activity. Abbr.: BA, Baseline active; PR, Primary responder; NR, non-responder; Rm, Remission; Re, Relapse.

We next examined bile acid profiles in humanized mice. The results showed clear separation between **inflamed (Active pre- and post-HSCT)** and non-inflamed (**inactive post-HSCT**) mice as shown by the PCoA (**Figure 30A**). Correlation and fold changes analysis identified Cholic acid 7-sulphate, lithocholic acid and its derivatives to be correlating with disease-free state in mice. In contrast, tauro-conjugated bile acids and glycocholic acid correlated with inflammation (**Figure 30B, C and D**).

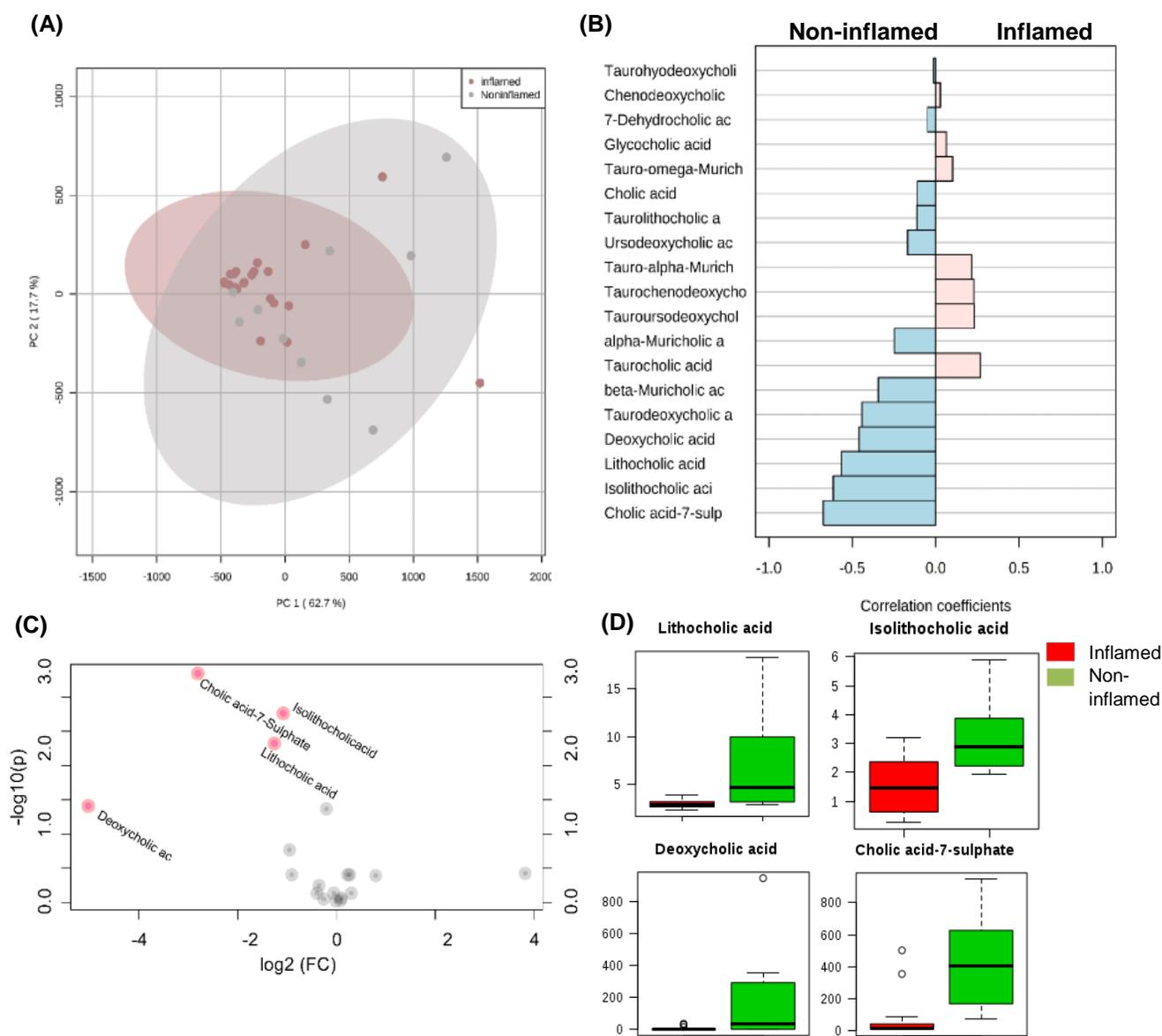


Figure 30 Identification of discriminative bile acid repertoire in humanized mice

(A) PCoA plot of bile acid profiles in human donors and humanized mice **(B)** Correlation plot of bile acid concentrations and inflammation **(C)** Volcano plot showing fold changes of different bile acids in humanized mice (categorized by inflammation status) **(D)** Top enriched bile acid

4.4.4. Predicted metabolic functions of the disease-relevant microbial communities

To predict the putative functions of the microbial communities relevant to disease activity in HSCT-treated IBD patients and similarly in humanized mice, we used PICRUSt to infer the functional content of the microbiota based on closed-reference OTU picking. PLS discriminative analysis (PLSDA) showed the separation between IBD patients in remission and those in relapse, based on the predicted bacterial functional pathways (**Figure 31A**). We compared 182 KEEG orthologs. Identifying the top 25 differentially abundant functional genes between IBD patients with active and inactive state of the disease (at FDR of 0.05) showed a switch towards increased Sulfur metabolism, RIG-I-like receptor signaling pathway, indicative for pattern recognition receptors responsible for detecting viral pathogens and generating innate immune responses, ABC transporters and purine metabolism pathways which were shown to be involved in colitis exacerbation (**Figure 31C**). Looking at functional modules differentially abundant between patients in remission or in relapse emphasized the enrichment of metabolic pathways involved in sulfur transport system and other ions (e.g Manganese and Nickel) transport systems. On the other hands, patients in remission showed an enrichment of functional pathways involved in basic biosynthesis processes (**Figure 31B**).

We applied the same analysis on the humanized mice. PLSDA showed better separation of the humanized mice based on the inflammatory status (**Figure 32A**). Non-inflamed humanized mice showed an enrichment of functional pathways involved in sugar transport systems, multiple amino acids biosynthesis and sulfate or glutathione transport system. On the other hand, inflamed humanized mice showed an enrichment of functional pathways involved in toxin transport, sulfur reduction, LPS biosynthesis and Histidine degradation (**Figure 32B**). Interestingly, identifying the top 25 differentially abundant KEEG orthologs between inflamed and non-inflamed humanized mice (at FDR of 0.05) emphasized the individualized signatures in humanized mice receiving different human microbiota. In **Figure 32C**, mice colonized with the same microbiota are underlined.

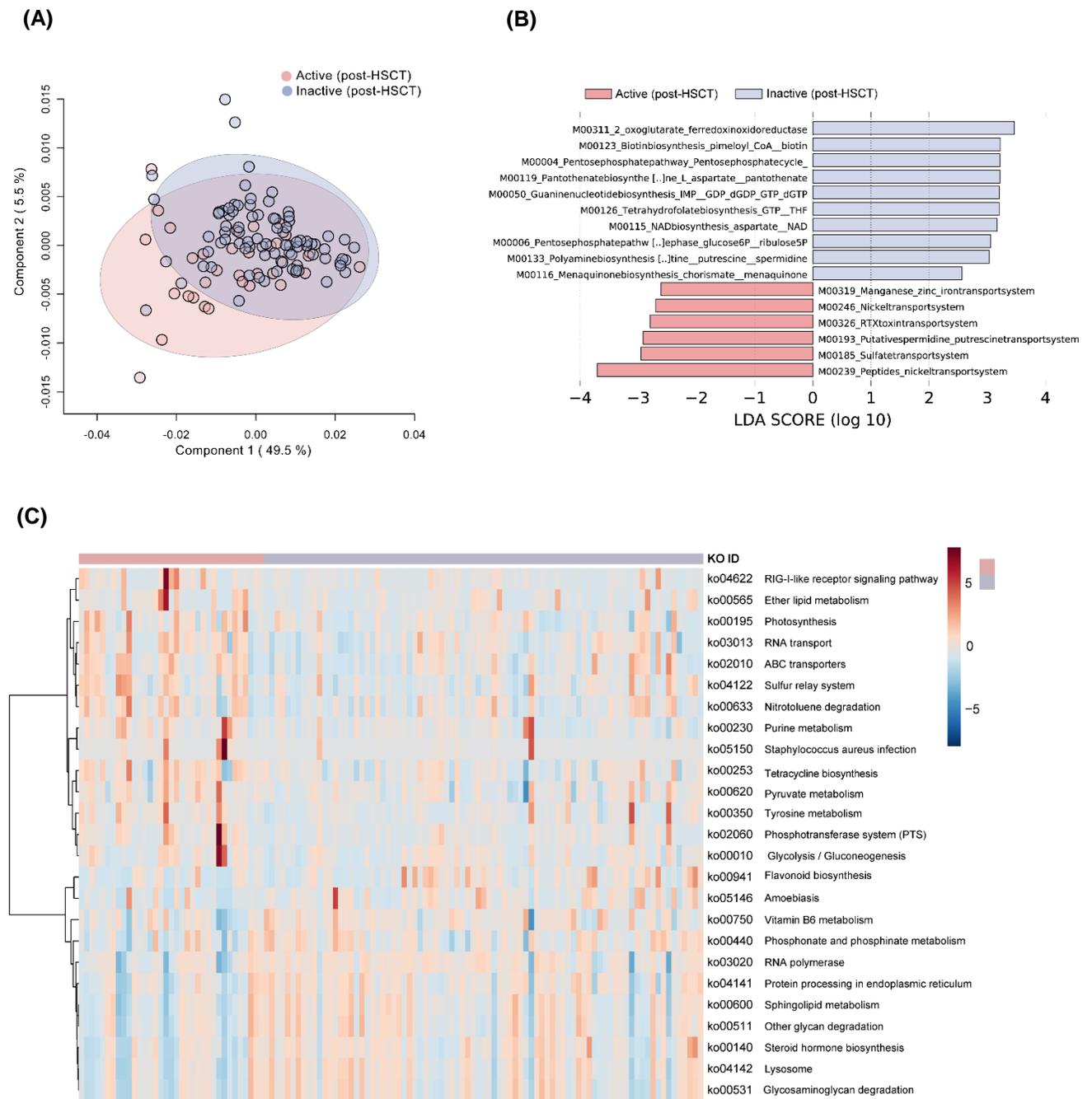


Figure 31 Functional pathways enriched in the fecal microbiota of IBD patients under inflammatory conditions

(A) PLS-DA score plot of predicted functional pathways in IBD patients with and without active disease and **(C)** IBD patients in remission or relapse. **(B)** Supervised analysis with LefSe showing functional modules differentially enriched under remission or relapse. **(C)** Heatmap showing relative abundance of the top 25 differentially abundant KEEGs under inflammatory conditions.

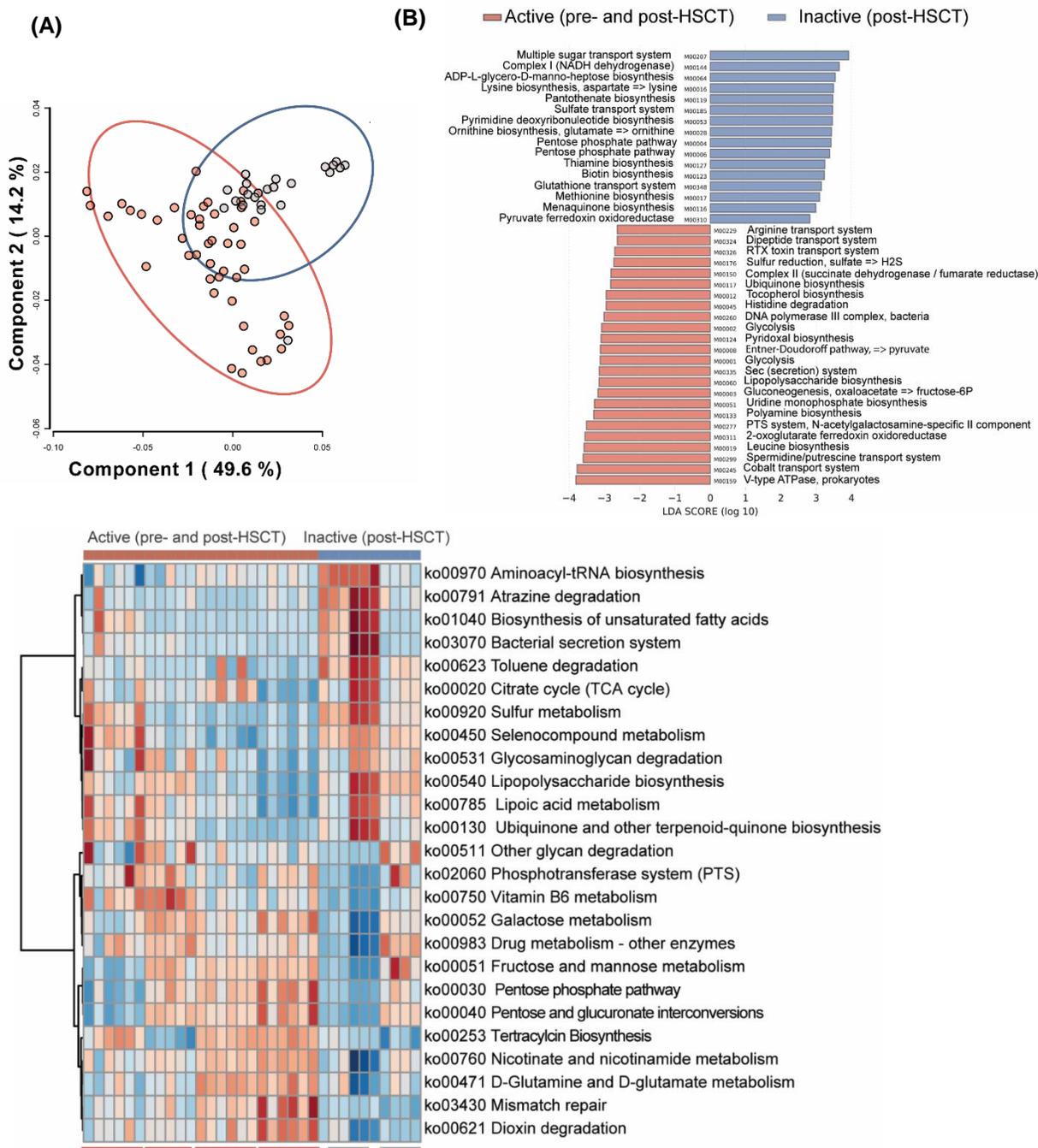


Figure 32 Functional pathways enriched in the fecal microbiota of humanized mice under inflammatory conditions

(A) PLS-DA score plot of predicted functional pathways in inflamed and non-inflamed humanized mice (B) functional modules differentially enriched under inflammation (C) Heatmap showing relative abundance of the top 25 differentially abundant KEGGs under inflammatory conditions

4.4.5 Human Donor as well as humanized mice show individual functional signatures under inflammatory conditions

To examine the differences between microbiomes of the selected human donors under inflamed or non-inflamed conditions, we performed metagenomics shotgun sequencing on the 6 microbial communities collected from the 3 human donors and 3 mice/group of respective humanized mice. PCA analysis of gene abundance data showed marked inter-individual variability in the gene content of fecal microbiome from the 3 human donors (**Figure 33A**). Interestingly, fecal samples collected from patients under inflammatory conditions or under non-inflammatory conditions did not necessarily cluster together, emphasizing the observation that individualized functional signatures are capable of driving inflammation. The number of differentially abundant gene orthologs was most evident in P16 (Remission vs. Relapse), where distinct functional genes were enriched in the inflamed condition. Less number of genes were differentially abundant in P28 (active baseline vs. initial response). As expected, very low number of genes were differentially abundant in P27 (active baseline vs. non-response), where both microbial communities represent an inflamed situation. This emphasized the intermediate inflammatory potential of fecal samples at baseline, where the highest separation could be characterized in remission versus relapse conditions (**Figure 33B**). Looking at the overlap between the inflammation-associated microbiota did not yield a considerable number of shared genes (**Figure 33C**).

A similar analysis on gut microbiome gene content of humanized mice yielded a better separation, as shown by the PCA analysis (**Figure 33D**). Humanized mice colonized with the same human donor microbiota seemed to cluster together. However, no clear separation based on the inflammatory outcome was evident. We performed PLS discriminative analysis to maximize the variation between inflamed and non-inflamed mice and constructed a heatmap of the top 25 differential genes (**Figure 33E**). Interestingly, humanized mice colonized with the same human microbiota shared a functional signature that is unique and distinct from other humanized mice, even if they share the same disease outcome, inflamed (active pre-and post-HSCT) or non-inflamed (inactive post-HSCT).

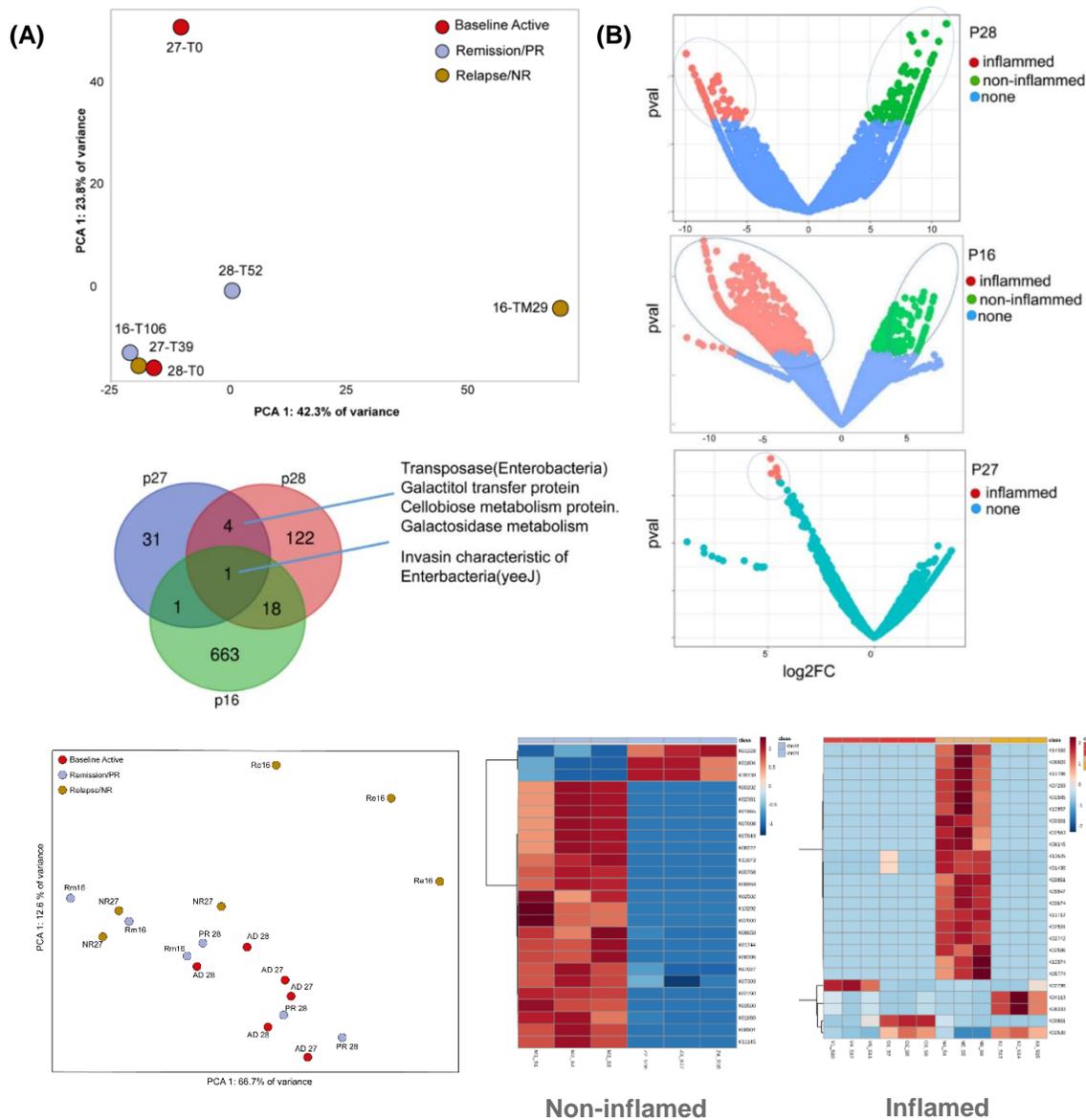


Figure 33 Comparison of metagenomic profiling of human donors and humanized mice under inflammation

(A) PCA plot of 6 human donors grouped by disease state **(B)** Volcano plots of fecal microbiota from 3 IBD patients (human donors). Each volcano plot shows the estimated fold changes (x-axis) versus the $-\log_{10}$ p-values (y-axis) for differential abundance of KEEGs between each two samples/patient. **(C)** Venn Diagram showing the overlap of KEEGs among samples collected during inflammation from the 3 IBD patients. **(D)** PCA of functional genes in selected humanized mice under inflammatory conditions **(E)** Heatmap of the top 25 differentially abundant KEEGs in mice.

4.4.6 Untargeted metabolomics approach to characterize the changes in metabolite profiles of IBD patients with active or inactive disease post-HSCT

We characterized fecal metabolome changes of both IBD patients and humanized mice using liquid chromatography-mass spectrometry (LC-MS). We used both reverse-phase ultrahigh performance liquid chromatography (RP-UHPLC) and hydrophilic interaction ultrahigh performance liquid chromatography (HILIC-UHPLC) techniques to cover both polar and nonpolar metabolites. An unsupervised principal component analysis (PCA) was performed to understand the relationships among the data matrix. Then, a supervised partial least-squares discriminate analysis (PLS-DA) was performed. PLS-DA model separated patients with active or inactive state of the disease (**Figure 34A**). Identification of differential metabolites features showed higher number of compounds linked to inflammation (**Figure 34B**). Metabolite identification using structural ontologies showed trend towards regulation of sulfur metabolism pathways, however enrichment analysis did not reach statistical significance. Integration of microbiota and metabolite data showed improved separation of IBD patients based on disease outcome post-HSCT (**Figure 34C**).

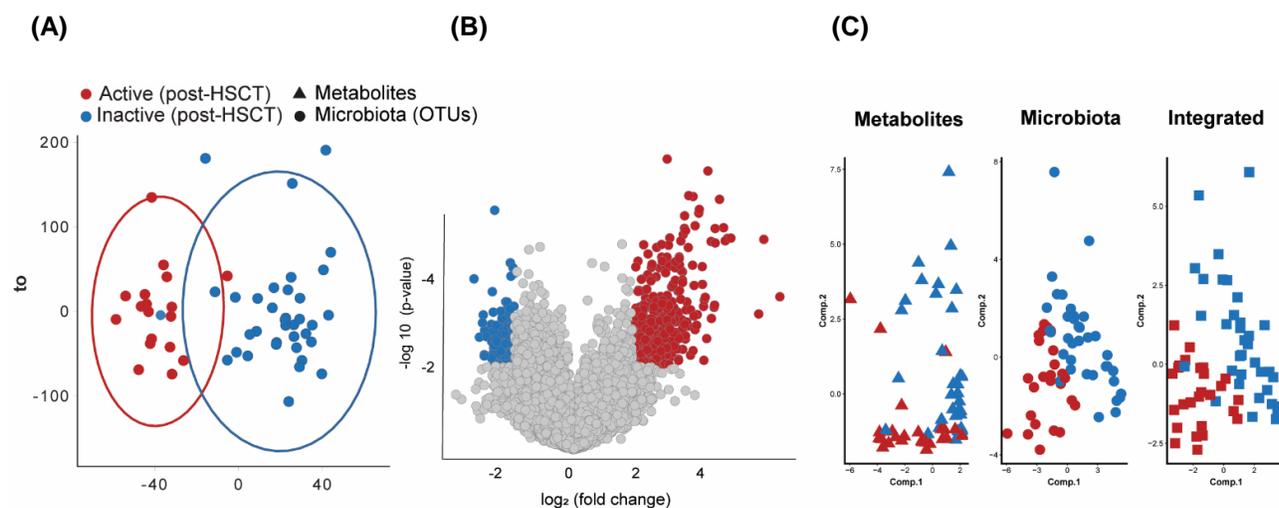


Figure 34 Alterations of fecal microbiota metabolite profiles in IBD patients

(A) PLS-DA score plot of metabolite profiles **(B)** Volcano plot of significantly differential metabolites (at cutoff fold change $\log_2 = 2$). **(C)** Separation of IBD patients according to disease activity post-HSCT based on metabolite features, microbiota and on the integrated dataset.

4.4.7 Significant enrichment of sulfated compounds is associated with inflammation in humanized mice

To visualize the metabolic changes under inflammation, we performed a PCA analysis and we studied the metabolite differences between inflamed and non-inflamed humanized mice as shown in **(Figure 35A)**. A supervised PLSDA plot showed clear separation of humanized mice based on inflammatory outcome **(Figure 35B)**. Of the 28606 features captured, a number of 672 features were differentially abundant in inflamed mice, while 580 features were differentially abundant in non-inflamed mice as shown by the volcano plot **(Figure 35C)**. Interestingly, in accordance to what we observed in IBD patients; regulation of sulfur metabolism seemed to be linked to inflammation in humanized mice. Additionally, humanized mice seemed to cluster together based on their original human donor, emphasizing the relevance of individualized metabolite signatures to disease pathogenesis **(Figure 35D)**.

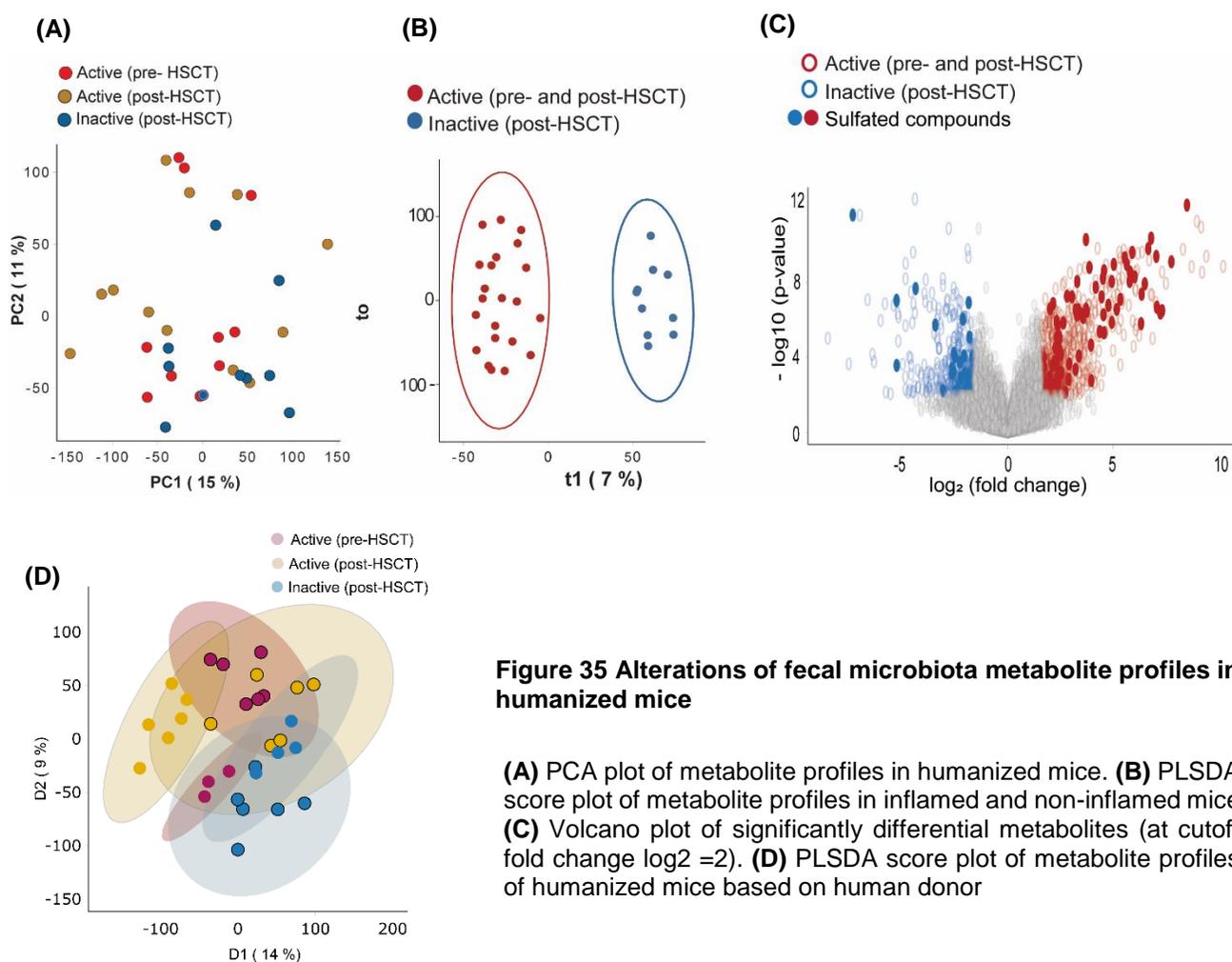


Figure 35 Alterations of fecal microbiota metabolite profiles in humanized mice

(A) PCA plot of metabolite profiles in humanized mice. **(B)** PLSDA score plot of metabolite profiles in inflamed and non-inflamed mice **(C)** Volcano plot of significantly differential metabolites (at cutoff fold change log₂ =2). **(D)** PLSDA score plot of metabolite profiles of humanized mice based on human donor

4.4.8 Integrative multi-omics signature improves the predictive modelling of disease outcome

To evaluate if the integration of both microbial and metabolite signals could be used to classify patients based on disease activity, we combined the datasets from IBD patients and humanized mice (**Figure 36A**). The separation of both datasets based on inflammation (disease activity) is illustrated in the PCA (**Figure 36B**). Interacting microbiome and metabolite features were used to train RF classifier (**Figure 36C**). This integration of metabolite and microbial species produced a significant improvement in classification accuracy relative to microbial features alone (**Figure 36D**).

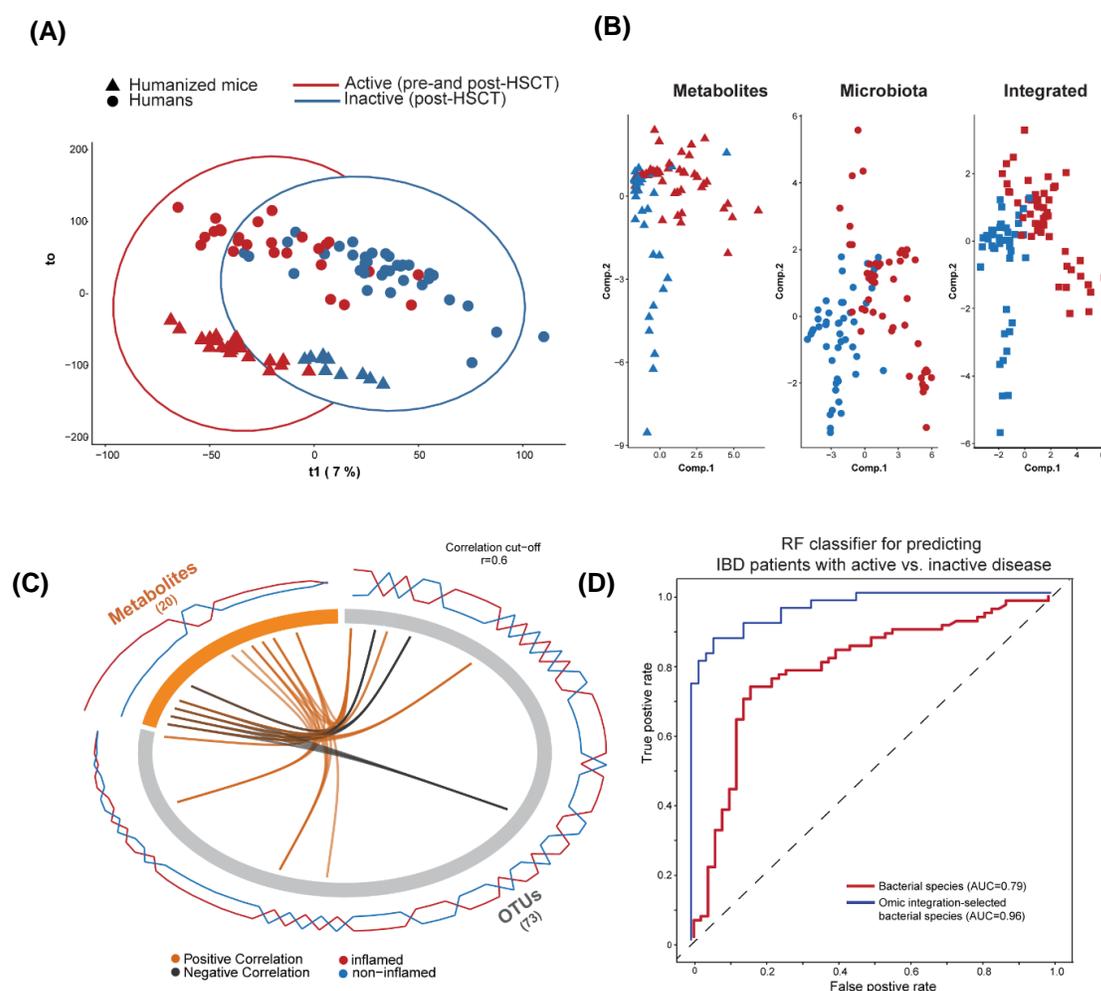


Figure 36 Integrative multi-omics signature improves the predictive modelling of disease outcome

(A) PLSDA plot of integrated omics data from IBD patients and humanized mice **(B)** Separation of patients and mice samples based on metabolite, microbiota and the integrated datasets **(C)** Circos plot shows the positive (negative) correlation ($r > 0.6$) between selected features as indicated by the orange (black) links between and within the different omics data **(D)** ROC curve showing the predictability of the omics-integration selected OTUs compared to that based on taxonomic data only.

4.5. Validation of humanized mice in IBD research

4.5.1 Transfer efficiency of human microbiota into germ-free mice

To address the efficiency of fecal microbiota transfer from human to mouse, we performed fecal transplantations of specific-pathogen-free (SPF) and healthy or dysbiotic human microbiota, respectively, into GF wild-type recipient mice. Expectedly, SPF mouse microbiota was retrieved in high efficiency (around 80%) but transfer of human microbiota into GF mouse was only partially successful. Percentage relative abundance of OTUs in original human donors' inoculum and those established in recipient mice after 4-weeks of colonization were calculated. Heatmaps demonstrate that the transfer of human microbiota is less efficient compared to the transfer of mouse SPF-derived microbiota (**Figure 37A**) Interestingly, humanized mice select a specific repertoire of OTUs to grow in relatively high abundance (**Figure 37B**).

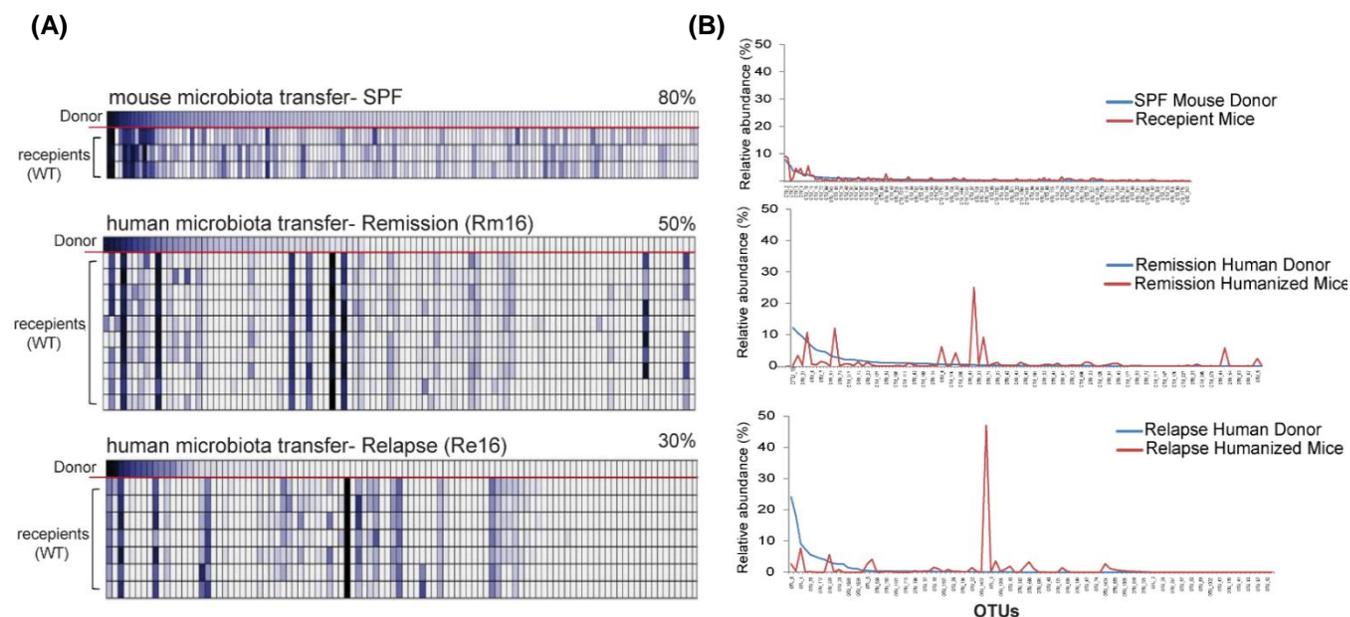


Figure 37 Transfer efficiency of human microbiota in comparison with mouse microbiota

(A) Comparison of the transfer efficiency of mouse and human donor microbiota into GF recipient mice; heatmaps showing comparisons of percentage relative abundance of OTUs in SPF mouse donor and GF-recipient wildtype mice, human IBD donor in remission and or relapse and respective humanized wildtype mice. Percentage transfer efficiency is indicated **(B)** Histograms showing the relative abundance of OTUs in original donor microbiota (SPF mouse donor, or human IBD donor in remission or relapse).

Based on the transfer of 6 different microbial communities from 3 patients with IBD into GF mice, humanized recipient mice reflected the microbial dysbiotic features of their original human donors. The transfer efficiency varied in the different experiments and ranged from 25%-80%, as shown by the UniFrac distances as proxy for dissimilarity (**Figure 38**) Reduced richness and diversity of the microbial ecosystem was observed in mice that were associated with the fecal microbiota from an IBD patient in relapse when compared to mice that received fecal microbiota transplantation with samples from the same patient in remission. In contrast to SPF-colonized mice, OTU-based analysis revealed that humanized mice are composed of very small number of abundant and diverse OTUs and a large number of OTUs that are present at low abundance (**Table 7**). Depending on donor's condition, the transfer of bacteria from human donor to recipient mice was particularly efficient for Bacteroidetes families, including *Bacteroidaceae*, *Rikenellaceae* and *Porphyromonadaceae*, and *Lachnospiraceae* from the phylum Firmicutes. In contrast, colonization of bacteria from the *Ruminococcaceae*, *Streptococcaceae* families of the Firmicutes phylum and *Enterobacteriaceae* from the phylum Proteobacteria was apparently insufficient. (**Table 8**).

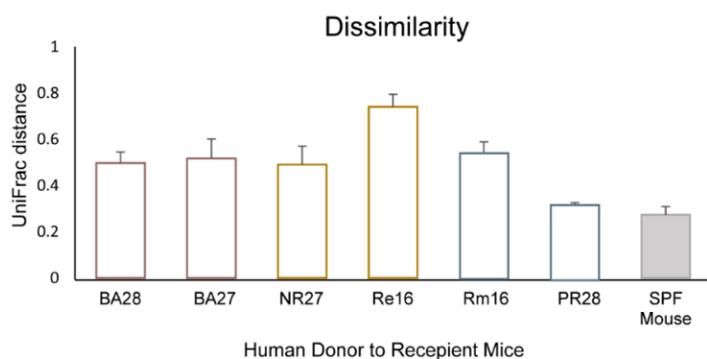


Figure 38 Dissimilarity of humanized mice and respective human donors

Generalized UniFrac distance between each human donor and the respective humanized mice, as a proxy for dissimilarity or transfer efficiency.

Table 7 Changes of bacterial taxa in humanization

	Number of Most Abundant OTUs (85% of whole community)	Number of all OTUs (Dominant and Rare)	Richness	Shannon.effective
Donor SPF Mouse	70	158	158	74.99
Receptient Mouse	66	153	141.67 ± 6.51	59.29 ± 2.12
Human Relapse	11	66	66	13.82
Humanized Mice (Re)	13	88	52.57±8.26	9.90 ± 2.09
Human Remission	23	78	78	27.4
Humanized Mice (Rm)	15	89	56.9 ± 5.2	14.4 ± 3

Table 8 Core and rare bacterial phylotypes in human/mouse donors and respective recipient mice

Family	Human Relapse [%]	(ReMb) Humanized Mice [%]	Human Remission [%]	(RmMb) Humanized mice [%]
f__Bacteroidaceae	0.14	45-65	22	12-40
f__Porphyromonadaceae	0.08	0-0.03	8	4-10
f__Rhodospirillaceae	0	0	0.30	3-20
f__Rikenellaceae	0	0	3	0.8-3.5
f__Lachnospiraceae	0.01	0.05-3.2	28	21-51
f__Ruminococcaceae	0.03	0.015-0.06	32	5-12
f__Enterobacteriaceae	65	5-12	0.09	0-0.06
f__Streptococcaceae	9.40	0	0	0

4.5.2 Established host-selected microbiota is efficiently transmissible to offspring

To investigate intergenerational microbiota, transfer efficiency and stability overtime, we began the experiment with germ-free WT129Sv 1 male and 2 female mice. Mice were inoculated with human fecal material at the age for 8 weeks, for 4-weeks to allow transfer and establishment of human fecal microbiota. Two weeks after inoculation, mice were cohoused for breeding. Parents (F0, n=2) and offspring (F1, n=9) were sampled at the age of 18 and 12 weeks respectively (Figure 39A). Microbiota community structure and OTU transfer efficiency were assessed by 16S rRNA gene sequencing. Data showed that after initial colonization of mice (F0) with human microbiota, F1 generation of mice maintained the engrafted human microbiota. A number of taxa were not efficiently transferred from human donor to recipient mice in the first generation (UniFrac distance $d=0.5$). However, once established in the first generation, the microbiota is transmissible to the next generation with higher efficiency (UniFrac distance=0.2) (Figure 39B, C).

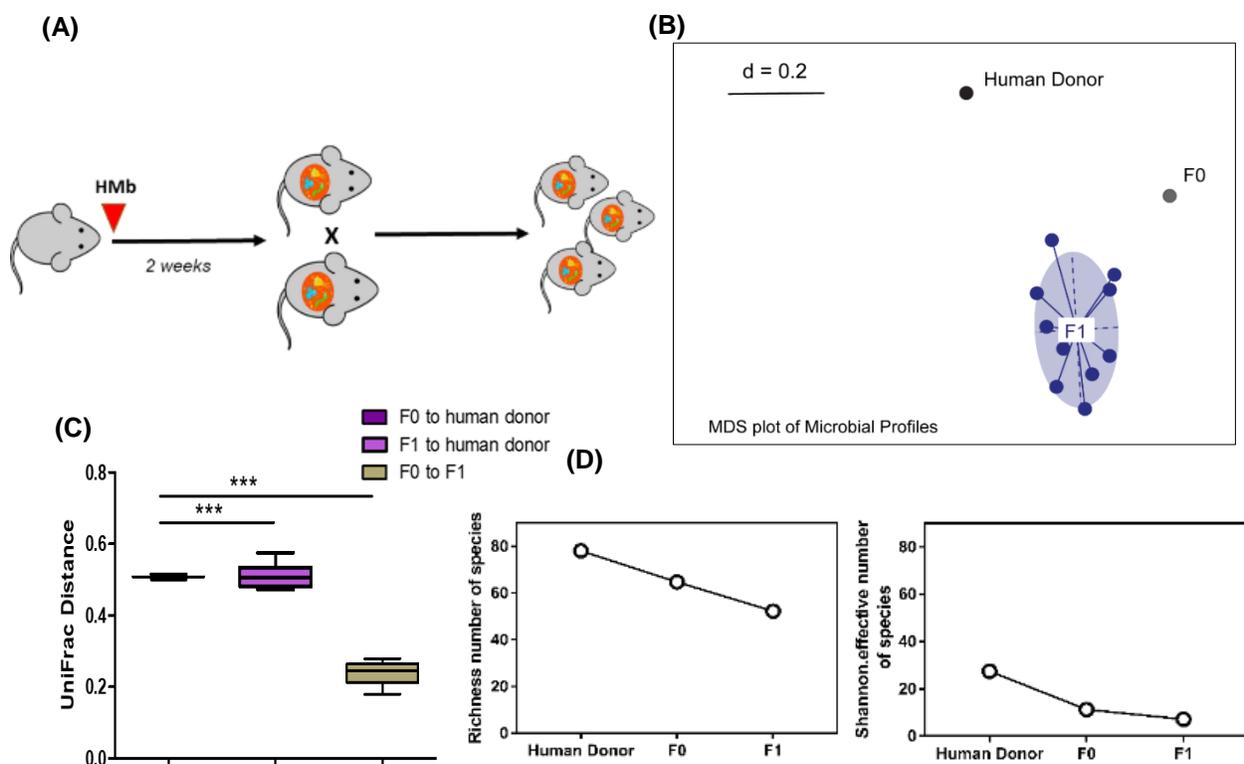


Figure 39 Intergenerational Microbiota Transfer Efficiency and stability

(A) Experimental setup **(B)** NMDS plot showing shifts in microbial profiles **(C)** Uni-Frac distances between Human donor, humanized mice (F0) and the offspring (F1) **(D)** Richness and Shannon effective counts drop from original human donor to the humanized mice.

4.6 Characterization of fecal microbial signatures in a cohort of biotherapy-treated adult CD and UC patients

4.6.1 Crohn's disease and ulcerative colitis patients show unique fecal microbiota profiles

In order to get an overview of microbial community composition in CD and UC patients included in the study, we performed a global beta-diversity analysis of all samples collected from CD and UC patients. Fecal samples from CD (n=50) and UC (n=27) patients showed distinct microbial community composition (**Figure 40A**). Inflamed (38 CD) and non-inflamed (10 CD) patients showed to have significantly distinct microbiota profiles (**Figure 40B**). Similarly, CD patients showed trends towards reduced community richness and diversity, however it did not reach significance, possibly due the huge variation (**Figure 40C and D**). This couldn't be addressed in UC patients since only one UC patient responded to the treatment.

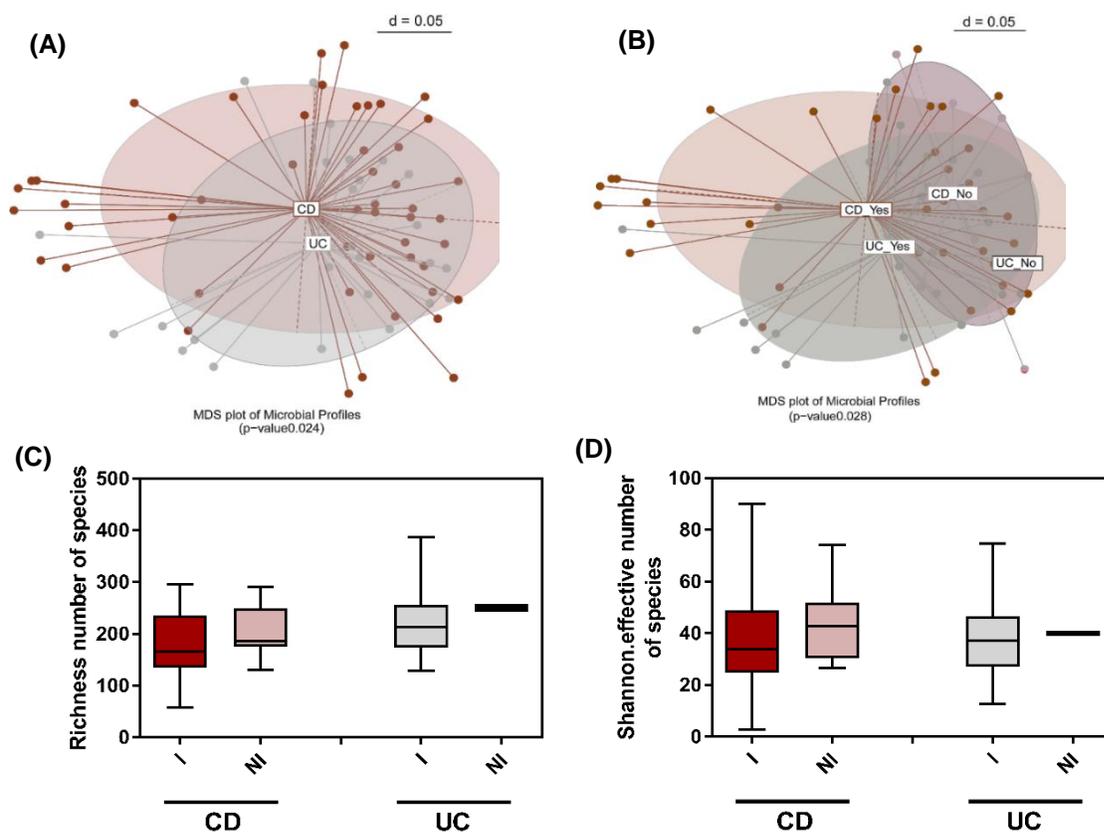


Figure 40 Higher Inter-individual variation in CD compared to UC patients

(A) Beta diversity analysis of microbial community composition across stool samples from CD and UC patients. **(B)** sub-stratified by inflammation (Inflamed; Yes) and (non-inflamed; No). **(C)** Community richness and **(D)** diversity in subgroups.

4.6.2 IBD patients show individualized response to biological therapy

We next looked at variations of bacterial communities before and after biological therapy. Quantifying generalized UniFrac distances, as a measure of dissimilarity between paired samples showed that the effect of treatment did not follow a consistent pattern. In other words, transitions were individual- specific and the microbial shifts between samples before and after treatment ranged from short to extreme in both groups (responders and non-responders), suggesting that fecal microbiome from CD patients are differentially responsive or sensitive to the biological therapy (**Figure 41A and B**). In both CD and UC groups, patients were classified into responders or non-responders. The assessment is based on endoscopic scoring as well as Harvey-Bradshaw index (HBI) for CD or Mayo score for UC at baseline, 14 weeks and 52 weeks' post-biological therapy.

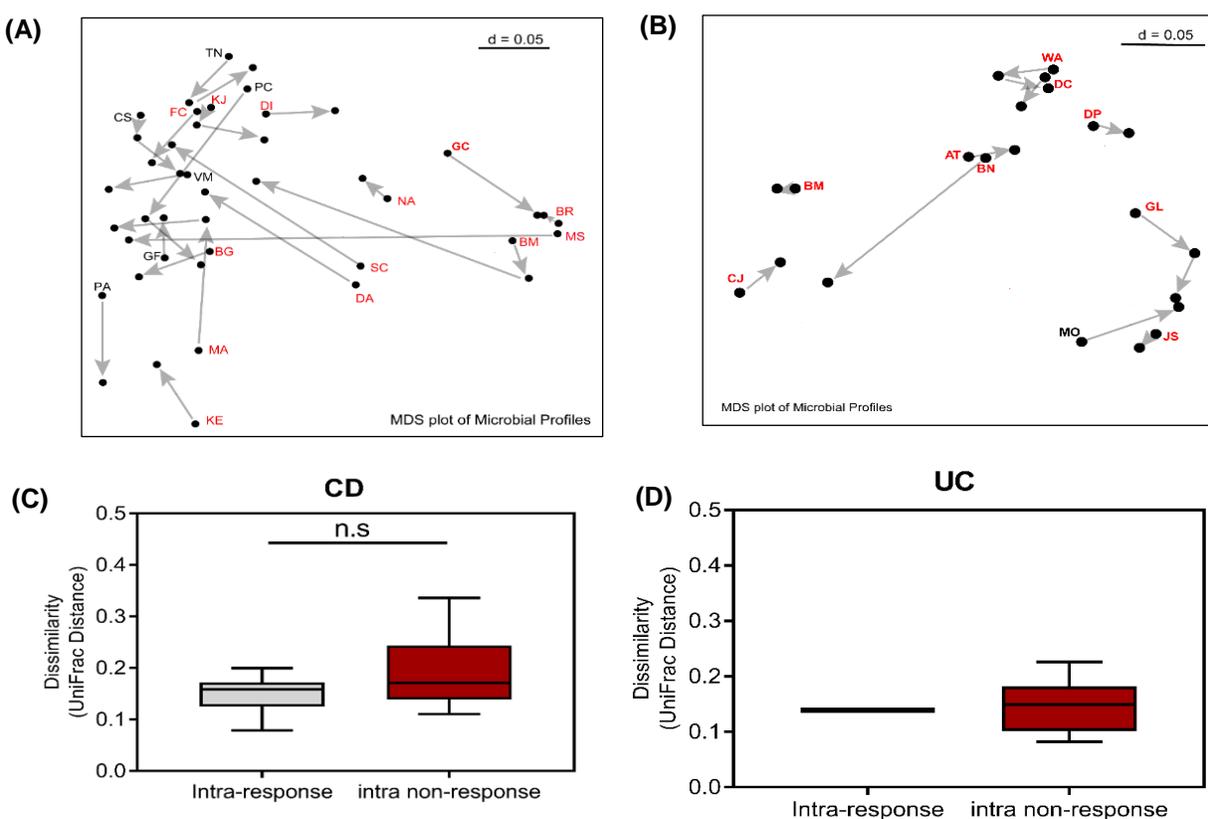


Figure 41 Effect of biological therapy on shifts in microbial community composition in IBD patients

(A, B) Transitions of microbial community profiles after initiation of therapy in CD and UC patients. The length of the connecting line corresponds to generalized UniFrac distance as a measure for dissimilarity plotted in **(C and D)**

4.6.3 Baseline composition and indication for response to biotherapy

Microbial community richness and diversity at baseline showed to be higher in CD who responded at week 14, however it did not reach statistical significance. Similar analysis could not be performed in UC patients because most of them did not respond to treatment (**Figure 42A**). Interestingly, beta-diversity analysis differentiated between CD patients who responded and those who did not at baseline (**Figure 42B**). *Escherichia Shigella* showed to be significantly more abundant at baseline among patients who did not respond at W14 (**Figure 42C**). Using taxonomic composition and diversity measures RF ML algorithm classified patients at baseline into responders and non-responders with an (AUC=0.564 and AUC=0.679), respectively (**Figure 42D**). Identification of classifying features showed an overabundance of *Odoribacter*, *Parasutterella*, *Flavinifactor* and *Alistipes* to be correlating with response to therapy at baseline (**Figure 42E**).

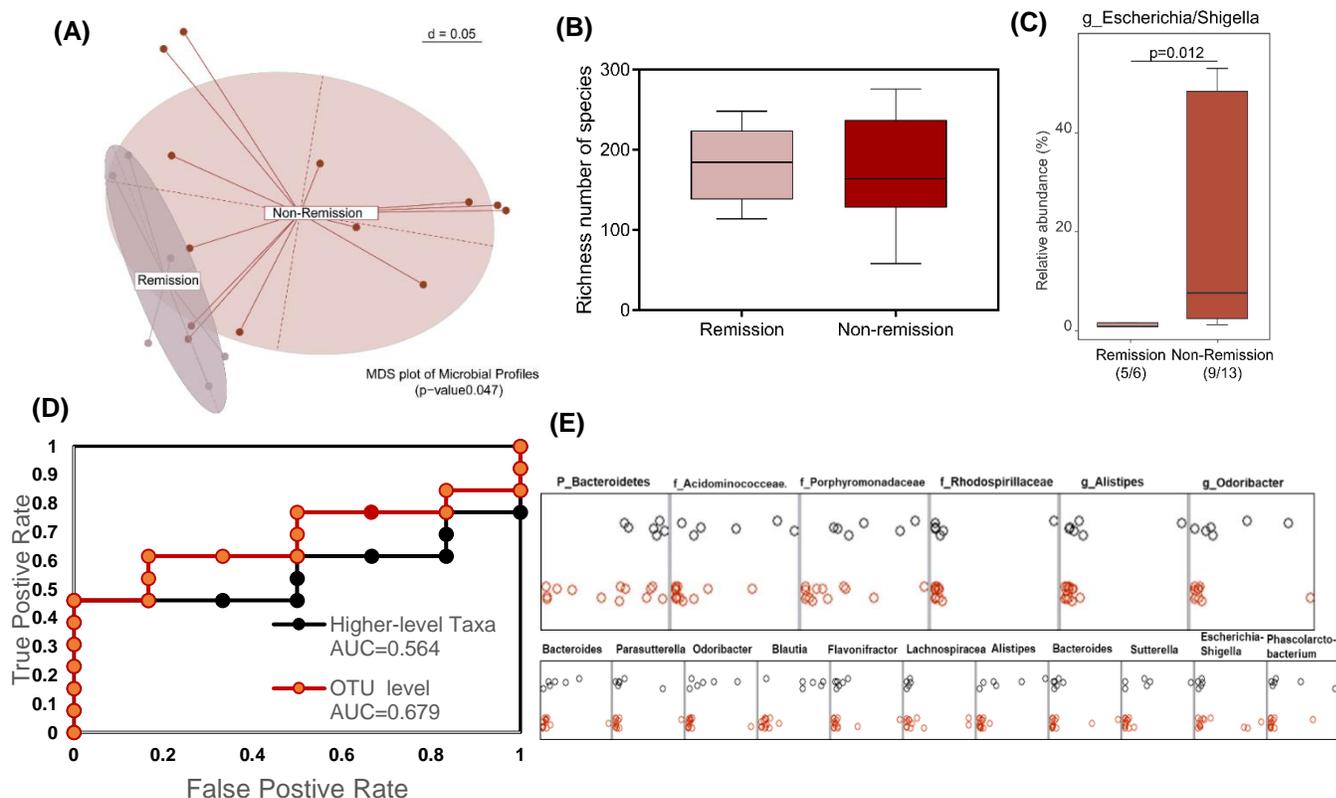


Figure 42 Baseline Composition and Indication for Response to biotherapy

(A) Beta-diversity analysis of microbial profiles at baseline **(B)** Community richness **(C)** Differentially abundant taxon between responders and non-responders at baseline **(D)** The training deployed a 5-fold cross validation scheme. A model incorporating taxonomy and community richness classified patients by 56% or 68%. **(E)** Classifying features for each model

We looked at changes in community richness and diversity in responders and non-responders 14 weeks after initiation of biotherapy treatment. CD patients who responded to biotherapy, showed higher levels of community richness and diversity; ranging from (150-300 species) and (25-80 species) respectively. In contrast, most of the non-responders did not show an improvement of diversity indices, and they generally remained at lower levels compared to responders (Richness= 50-300 species and Shannon. Effective= 10 – 60 species). Wider range of baseline diversity levels was noted in UC non-responder patients (**Figure 43**).

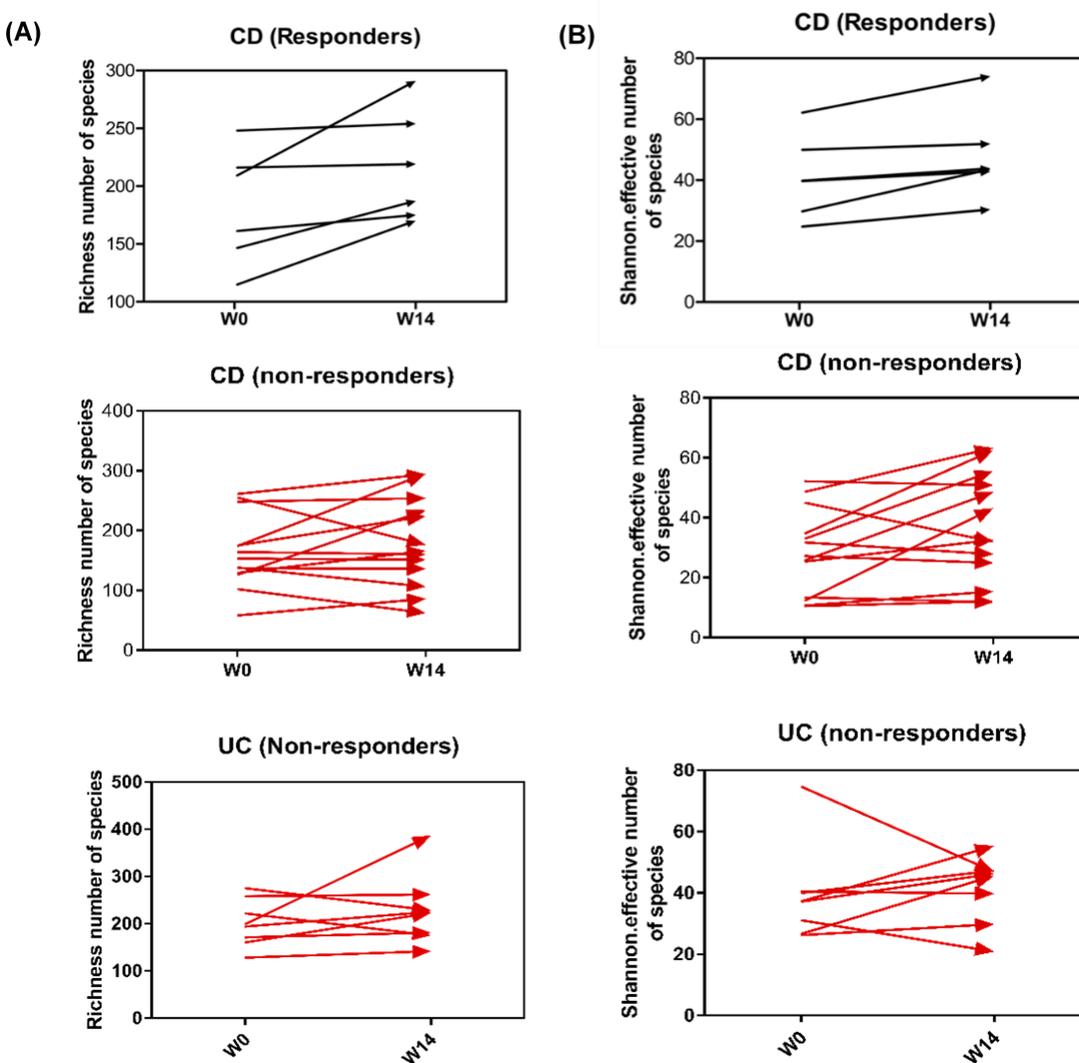


Figure 43 Changes in community structure at week 14 post-biotherapy

(A) Community richness and (B) diversity in responders and non-responders 14 weeks after initiation of biotherapy in CD and UC patients

4.7 Characterization of fecal microbial signatures in a cohort of post-operative recurrence cohort

In order to characterize the microbial dysbiosis associated with post-surgical resection, we analyzed 13 paired- samples collected from CD patients at time of surgery and 6 months post-ileocolonic resection using 16S rRNA sequencing of gene amplicons. To assess the dissimilarity of microbial profiles among the subgroups, we performed a Bray-Curtis based beta-diversity analysis. We stratified fecal samples into: **(1)** active at baseline – from patients on antibiotics or **(2)** From patients who stopped antibiotic treatment at least 3 months prior to the surgery, **(3)** samples collected from patients who maintained remission 6 months' post-surgery and **(4) and (5)**. those who relapsed substantially or severely 6 months after surgery. Interestingly, we observed that stool samples from inflamed patients at baseline separated significantly into two groups based on their antibiotic intake history (**Figure 44A**). The number of samples in the relapse groups were too low to make conclusions on shifts in microbial community composition. Alpha-diversity analysis showed a trend towards reduced community richness in patients who recurred compared to those who maintained remission, however it did not reach statistical significance (**Figure 44B**). As expected, patients on antibiotics treatment showed reduced community richness compared to those who stopped antibiotic treatment at least 3 months prior to surgery/ sample collection.

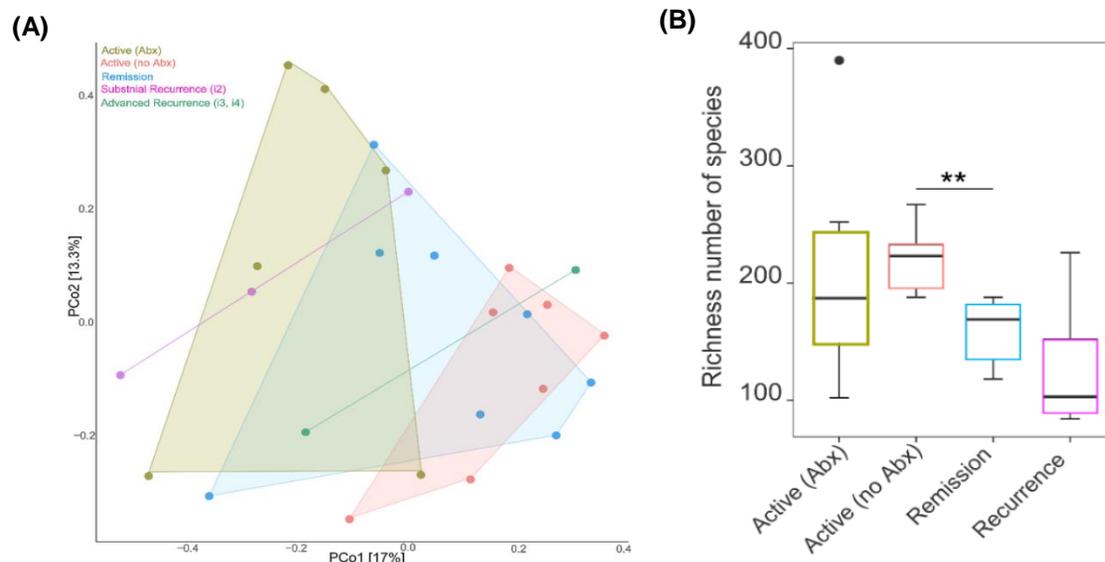


Figure 44 Characterization of post-surgical recurrence associated microbial dysbiosis

(A) Principal coordinate analysis (PCoA) profile of microbial community composition across stool samples using Bray-Curtis distance metric. The percentage of variation explained by PC1 and PC2 are indicated in the axes. **(B)** Microbial community richness at baseline at surgery, at remission or recurrence post-surgery.

Due to the small sample size, we looked at the individual variations overtime, at surgery and 6-months later to characterize the shifts in microbial community composition. Relapsing and non-relapsing patients showed individualized microbial shifts in response to surgery (**Figure 45**). In each case, the microbial community composition showed different shifts towards remission or relapse. However, in most of the relapsing patients, an increase in the relative abundance of members belonging to the genera *Escherichia-Shigella*, *Lachnoclostridium* and to some extent *Acidaminococcus* was observed (**Figure S12**).

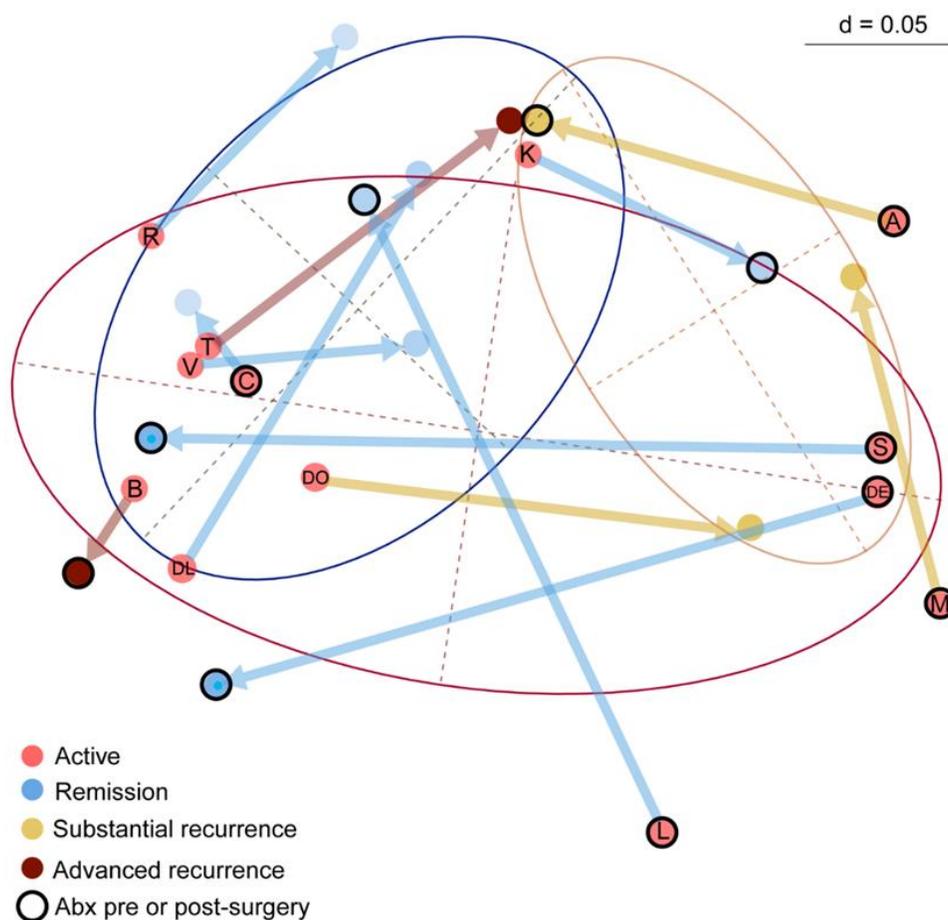


Figure 45 Microbial shifts pre- and post-surgery in remitters and non-remitters

NMDS plot showing microbial shifts in response to surgery in relapsing and non-relapsing patients

5. Discussion

The present work investigated the composition and community changes in the gut microbiome of three longitudinal IBD patient cohorts who were undergoing three different therapies: 1. Autologous Hematopoietic Stem Cell Transplantation (AH SCT); 2. biological therapy (anti-TNF and anti-integrin), and 3. surgical resection. We identified microbial signatures linked to disease outcome and response to therapy in two of the cohorts. To bridge the gap between microbiota composition and their actual function, we employed a multi-omics approach comprising metagenomics profiling, in addition to targeted and untargeted metabolomics profiling to assess the functional potential and functional activity of the microbial signatures we identified. To address host-microbe interactions of relevant microbial communities, we established a gnotobiotic humanized mouse model to assess host response to disease-relevant bacteria.

5.1 Identification of microbial signatures associated with response to HSCT therapy

While the involvement of the microbiota in IBD has been intensively studied, the data on the contribution of the intestinal microbiota to HSCT therapy outcome is lacking. In this unique longitudinal study, we characterized the composition and diversity of gut microbial communities in Crohn's disease patients undergoing autologous HSCT; making it one of the first reports on microbial changes driven by HSCT therapy in IBD. This single-center cohort from Barcelona included 35 patients with CD, who were unresponsive and refractory to all the conventional therapies. Six months after autologous HSCT, 70% of the recruited CD patients achieved drug-free clinical remission (López-García et al., 2017). Assessment of long-term remission at 5 years of follow up showed that 15% of the patients remained in drug-free remission (López-García et al., 2017). Although promising, HSCT is a major procedure with a relatively high mortality rate (2-10%) (Jauregui-Amezaga et al., 2016). Accordingly, there is a great need to optimize response through monitoring the changes patients undergo overtime and identifying patients who are most likely to respond to HSCT before initiation of therapy.

Characterizing changes in taxonomic composition of fecal bacterial communities of HSCT-treated CD patients as a function of disease activity status emphasized that during remission patients showed significantly different microbial community structure compared to patients during active state of disease; at baseline or when they relapsed after long-term remission. Patients at baseline or during relapse showed overlapping microbial communities characterized by reduced community richness and diversity. Altogether, fecal bacterial communities from CD patients

(baseline, during remission or relapse) showed clear separation from those of healthy subjects. Longitudinal sampling of CD patients over a 5-year period highlighted the fluctuations of microbial profiles during disease, where a change of disease status correlated with dramatic changes in microbial community structure. As a control, a repeated sampling of 2 healthy adults over a 4-weeks period showed that the temporal changes were minimal, with each individual maintaining a relatively stable community profiles overtime. Though in clinical and endoscopic remission, remitting patients do not show complete recovery to healthy microbiome diversity indices, and they rather show an intermediate level between active patients and healthy subjects. Previous studies on longitudinal cohorts of patients with IBD showed that fluctuations in the gut microbiome overtime underline the severity of inflammation. Compared to healthy subjects, IBD patients have more volatile dynamics of the gut microbiome (Halfvarson et al., 2017).

In our study, there was a considerable overlap of microbial communities amongst IBD patients with active disease and patients with inactive disease, additionally, an unsupervised beta-diversity analysis showed that IBD patients formed 3 distinct clusters, where patients with active state of disease showed heterogenous microbial community structure. Using the LEfSe algorithm, we identified specific genera that were enriched or depleted in the microbiota of IBD (baseline, active or inactive post-HSCT) versus healthy subjects. As expected, several of the enriched bacterial groups in IBD are known to be associated with aggravated CD-inflammation or cancer; including *Escherichia-Shigella* (Bringer et al., 2004; Rolhion and Darfeuille-michaud, 2018) and *Fusobacterium* (Strauss et al., 2011)(Kostic et al., 2012). Others include *Ruminococcus gnavus*, which has been implicated in a number of studies as a pathobiont. In this context, a recent FMT study on UC patients reported an increased rate of relapse in UC patients who received fecal microbiota from donors with high levels of *R.gnavus* (Fuentes et al., 2002). *Eisenbergiella* has been previously reported to be enriched in microbiota from colorectal cancer patients(Zhang et al., 2018). In addition, members belonging to the sulfate reducing *Desulfovibrio* genus showed to be enriched in IBD patients. These organisms produce hydrogen sulfide; a toxic metabolic end-product, and subsequently impair the functions of the intestinal epithelium through inhibition of butyrate oxidation (Loubinoux et al., 2002). While the butyrate producing *Clostridiales* are known to be generally beneficial to the host, *Clostridium XIVA* and *Romboutsia* showed to be enriched in IBD patients. Looking closely at the individual subgroups, they seem to be particularly prevalent in IBD patients who achieved remission post-HSCT. Prior literature showed that the species *Romboutsia ilealis*, is known to produce hydrogen and carbon dioxide together with acetate, formate and lactate that are utilized by *Desulfovibrio* for their growth and respiration and thus

contribute to colonic inflammation (Gerritsen et al., 2017). In addition, the results confirm that CD-associated dysbiosis is characterized by disappearance of multiple beneficial taxa that appear to be significantly enriched in healthy subjects. These beneficial organisms include *Prevotella*, multiple genera belonging to family *Ruminococcaceae*, *Faecalibacterium*, *Christensenellaceae*, *Eubacterium hallii*, *Coliinsella*, *Methanobrevibacter*, *Roseburia* and *Oscillibacter*. Previous reports described a concurrent decrease of *Roseburia* and *Ruminococcaceae* in IBD; where their role in maintaining host homeostasis is impaired. In healthy subjects, *Ruminococcaceae* consume hydrogen and produce acetate which is utilized by *Roseburia* to produce butyrate (Duncan et al., 2002),(Morgan et al., 2012). *Christensenellaceae*, *Methanobrevibacter*, *Oscillibacter* have been shown to be reduced in a cohort of IBD patients (Pascal et al., 2017). *E. hallii* is an anaerobic propionate producer gut commensal, known to have function in maintaining host homeostasis (Engels et al., 2016). Moreover, a previous study showed that administration of live *E. hallii* in obese and diabetic db/db mice increased energy metabolism and improved insulin sensitivity (Udayappan et al., 2016). Using machine learning algorithms, we showed that a combination of diversity indices and taxonomic composition at OTU species level could classify patients to patients with active or inactive diseases, and further classify patients based on clinical outcome after HSCT in this cohort with an AUC of 0.79 and 0.83, respectively. Obviously, there is a mounting body of evidence that gut dysbiosis is characteristic for patients with IBD. However, it is still not clear whether this dysbiosis is the cause or the consequence of the disease. With this in mind, we looked the patient's microbial profiles at baseline to investigate whether the microbiota trajectory before initiation of therapy could predict response at the primary endpoint. A trend towards a more diverse microbial composition at baseline was associated with clinical remission at the primary endpoint; however, this did not reach statistical significance. In addition, beta-diversity analysis of microbial profiles revealed a poor separation between remitters and non-remitters at baseline, and discriminative LEfSe analysis did not yield any differentially abundant taxa. One of the issues in analysing the data from patients at baseline is the small sample size (6 remitters and 6 non-remitters). Given the remarkable inter-individual variability known of human microbiota at the taxonomic level, it was challenging to make any predictions on response to HSCT therapy based on this small cohort.

An important aspect of this cohort is that the patients were intensively treated with biologics and immunosuppressive drugs before they underwent HSCT as last line therapy. This was manifested in the huge inter-individual variations among patients' microbial profiles at baseline. Indeed, this high inter-individual variation in microbiota composition introduced noise to our analysis, possibly

resulting into smaller correlation coefficients. Consequently, we excluded the fecal samples collected at baseline from the analysis and re-evaluated the results. The results showed that remitters and non-remitters showed a clear separation of microbial profiles after HSCT. Using LEfSe algorithm, we identified a list of taxa that discriminate between both groups. In addition to the ones identified before as markers discriminating IBD patients from healthy subjects, the results showed an enrichment of a number of organisms that have been described previously as pathobionts. These include *Enterococcus* (Zhou et al., 2016), *Megasphaera* (Dhiman, 2012), *Hemophilus*; a bacterium previously reported to positively correlate with CD (Gevers et al., 2014), *Granulicatella* (Hai-Qin et al., 2018) and *Campylobacter* (of Proteobacteria and shown to release cytolethal toxins, leading to cell cycle arrest, chromatin fragmentation and apoptosis (Smith et al., 2018) . Expectedly, remitters had increased levels of beneficial microbes including *Akkermansia*, *Barnesiella*, *Oscillibacter*, *Roseburia* and *Odoribacter*. *Akkermansia* is known to exert an anti-inflammatory action in the gut and enhances the intestinal barrier function. Furthermore, an increase in the abundance of *Akkermansia* was observed in cancer patients who responded to immunotherapy (Naito et al., 2018). *Barnesiella* one of the depleted taxa in relapsing CD patients was shown to have a protective effect. Previous research confirmed that microbiota dominated with *Barnesiella* was shown to clear vancomycin-resistant *Enterococcus* (VRE) ,a pathogen causing infections in immunocompromised patients undergoing allogeneic hematopoietic stem cell transplantation (Ubeda, 2013). Similarly, we observed a negative correlation of *Odoribacter* genus with inflammation. *Odoribacter* belongs to the *Porphyromonadaceae* family and known to be a producer of acetate, propionate, and butyrate. Hence, it may affect host inflammation via reduced SCFA availability(Markus et al., 2011).

5.2 Mucosa-associated microbiota in HSCT-treated CD patients

Most of the microbiota studies investigate fecal samples (representing the luminal microbial niche), clearly due to the ease of frequent collection. However, several recent studies have shown that the microbial compositions of the luminal microbiota (fecal samples) and the mucosa-associated microbiota (from biopsies) largely differ, suggesting that these two distinct microbial ecosystems perform different functions within the intestinal microbial niche (Ringel et al., 2015). For example, mucosa-associated microbiota modulate systemic immune responses (Lee and Hase, 2014) , microbial metabolites can penetrate the epithelial barrier, interact with immune cells, and subsequently exert functions essential for the host (Dorrestein et al., 2015).

To examine the microbiota community composition at the mucosal surfaces, we analyzed mucosal-associated microbiota from a subset of the HSCT-treated patients. Clearly, the microbial

composition of CD patients with ileal or ileocolonic involvement significantly differed. Hence, we looked at each group separately. The comparison of the mucosa-associated microbiota between active and inactive CD patients with colonic involvement confirmed the significant enrichment in *Escherichia-Shigella* and *Ruminococcus gnavus* in inflammation. Additionally, *Halomonas*, *Streptococcus* and *Shewanella* correlated with inflammation particularly in mucosa tissue with colonic involvement. In accordance with these results *Halomonas* and *Shewanella* were significantly more abundant in patients with rectal and distal cancers (Flemer et al., 2017). *Halomonas* is known to have halophilic properties (Quesada et al., 2018; Ruiz-Ruiz et al., 2011), while *Shewanella* has sodium permeation ability (Wang et al., 2008). Together, they are thought to increase extracellular salt concentrations and subsequently promote apoptosis. *Streptococcus* spp. was among the dominant taxa in active patients, in line with previous reports where a predominance of *Streptococcus* was shown in inflamed mucosa of CD adult patients. Some species of *Streptococcus* are able to produce considerable amounts of hydrogen peroxide or reactive oxygen species that can penetrate through the mucosa and possibly promote necrosis or apoptosis (Fyderek et al., 2009). Several other genera were significantly reduced such as the butyrate producing *Faecalibacterium*, *Blautia*, *Akkermansia* and members belonging to the family *Lachnospiraceae*. The alpha diversity analysis confirmed a decrease of biodiversity in patients with active disease. Interestingly, CD patients with ileal involvement showed a predominance of *Bifidobacterium* and *Lactobacillus* groups. We found it intriguing that while both genera are common probiotics, they rather correlated with inflammation. However, prior literature showed similar trends of increased *Bifidobacterium* in mucosal tissue of UC patients and increased *Lactobacillus* in the mucosal tissue of active CD patients (Wang et al., 2014). Experiments in animal models showed that administration of *Bifidobacterium animalis* in germ-free interleukin-10-deficient mice induced significant duodenal and mild colonic inflammation (James P. et al., 2009). However, other studies showed contradicting results. For instance, oral administration of *Bifidobacterium breve* prevented intestinal inflammation in immunocompromised mice (Jeon et al., 2012). Moreover, one study reported reduced levels of *Bifidobacterium* in UC patients accompanied with increased levels with *Escherichia Shigella* (Mylonaki et al., 2005).

On the other hand, an increase of *Veillonella*, *Escherichia-Shigella* and *Acidominococcus*, *Anaerostipes* and *Klebsiella* was observed in patients with active disease. Compared to patients in remission, members belonging to *Faecalibacterium* genus were sharply decreased in the ileal inflammation. The same trend was observed in *Fusobacterium* and *Blautia* genera. Alpha diversity was significantly reduced in patients with active disease compared to patients in remission. Beta-diversity analysis emphasized the clear difference between ileal-associated microbiota in both

groups and showed better separation than that observed in fecal or colon mucosa-associated microbiota. Collectively these results confirm that the intestinal luminal microbiota and mucosal microbiota are distinct ecosystems that differ in diversity and microbial composition. The better separation of microbial profiles associated to mucosal surfaces confirm that they give better reflection of disease status, Unfortunately, the number of samples at baseline were very low to make predictions about response to therapy. However, this is something remains to be validated in bigger cohorts.

5.3 A Gnotobiotic mouse model to study the functional impact of microbial dysbiosis in IBD

While a great deal of research confirmed that gut dysbiosis is characteristic for patients with IBD, it is still not clear whether this dysbiosis is the disease driver or a mere consequence of inflammation. In this regard, gnotobiotic humanized mice stand as an alternative mechanistic approach to address microbe-host interactions. Recent studies confirmed the success of gnotobiotic humanized mice in recapitulating host pathological phenotype. This implies that dysbiosis is at least not only a consequence, but indeed contributes to development of IBD (Arrieta et al., 2016a; Nagao-Kitamoto et al., 2016; Ridaura et al., 2013; Yatsunenkov, 2011). Clearly, the advantage of using gnotobiotic humanized mice is the possibility to rule out extrinsic confounders, such as diet, genetic background and external environmental factors Besides, it enables the manipulation of microbial communities inhabiting the gut and allows subsequent analysis of host response.

Based on the microbiome analysis of the HSCT CD cohort, identification of microbial signatures predictive of disease was rather challenging. At least at the taxonomic and phylogenetic level, it was not feasible to identify unique microbial signatures specific for disease activity. At this end, we developed a humanized IBD mouse model by colonizing germfree *Il-10*^{-/-} mice with fecal samples from CD patients. We selected 3-paired samples from CD patients representing different disease activities and community clusters Germfree wild type (WT) and *Il-10*^{-/-} mice at the age of 8 weeks were colonized with selected CD patient-derived microbiota for 4 weeks. Interestingly, the transplantation of microbiota from patients with active or inactive disease was reproducibly sufficient to recreate the disease phenotype in recipient *Il-10*^{-/-} mice. We assessed the colitogenic potential of human donor microbiota by means of histopathological evaluation, pro-inflammatory gene expression and assessment of mesenteric lymph nodes immunogenicity in humanized mice. *Il-10*^{-/-} mice colonized with microbiota from patients at baseline or during relapse showed significantly increased inflammation in cecal tissue compared to mice colonized with microbiota

from CD patients in remission. The pro-inflammatory genes response to the different donor microbiota showed a significantly increased production of *IFN γ* and *IL-6* in ceal tissue of humanized mice colonized with disease-associated microbiota. Although the mice are genetically identical and the microbial content inoculated is similar amongst all mice per group, mice colonized with disease-associated microbiota showed a gradient of inflammatory response, in terms of histopathology and pro-inflammatory gene expression. Intriguingly, mice colonized with microbiota from patients during relapse showed higher degrees of inflammation compared to those colonized with microbiota from patients at baseline. A possible explanation for this observation is that while patients during relapse develop an acute severe form of inflammation, patients at baseline suffer chronic inflammation, possibly of less severity. In addition, the heavily administered drugs before initiation of therapy could possibly attenuate the microbiota inflammatory potential. In contrast, all mice colonized with a microbiota from CD patients in remission remained disease-free and showed minimal activation of immune response.

To examine changes in microbial composition, fecal samples from human donors as well as humanized mice were analyzed by 16S rRNA gene sequencing. As reported previously (Nagao-Kitamoto et al., 2016), gnotobiotic humanized mice reflected the composition and the dysbiotic features of their respective human donors in terms of richness and diversity measures, emphasizing the utility of humanized mice as a powerful translational tool to characterize disease-associated dysbiosis in human studies. Interestingly, microbiome analysis of humanized mice revealed that consistent with individually diverse microbiota profiles in donor CD patients, the presence or absence of inflammation in gnotobiotic mice is driven by various individual community profiles, suggesting that different microbiota composition configurations infer similar inflammatory functions in the host, leading eventually to disease pathology. However, when discussing findings in this study, it is important to acknowledge that we are working with a mouse model for a human disease. Even though the humanized mouse model proved to be a great tool for studying the contribution of microbiota to disease outcome, this model has its pitfalls. In this context, we performed a few validation experiments to assess the transfer efficiency of human microbiota to GF mice. Our results showed that the transferred microbiota resemble the human donor microbial community, in terms of established bacterial taxa, however, these taxa changed in their proportions to resemble the microbial communities known to inhabit the murine gut. Furthermore, the transfer of bacteria from human donor to recipient mice was particularly efficient for Bacteroidetes families, including *Bacteroidaceae*, *Rikenellaceae* and *Porphyromonadaceae*, and family *Lachnospiraceae* from the phylum Firmicutes. On the other hand, transfer of bacteria from

the *Ruminococcaceae*, *Streptococcaceae* families of the Firmicutes phylum and *Enterobacteriaceae* from the phylum Proteobacteria was less efficient, suggesting bacterial adaptation to the host environment. Prior literature has addressed this issue of host specificity and the success of microbiota transfer experiments in different animal models. For example, Rawls and colleagues performed reciprocal transplantations of complex microbiota into GF zebrafish and mouse recipients (Rawls et al., 2006). They showed that the transferred microbiota resembles the community of origin in terms of established bacterial taxa, however, these taxa changed in their proportions to resemble the gut microbiota of the recipient host. Similar observations were made by Chung and colleagues, who inoculated GF mice with human microbiota (Chung et al., 2012) Given the multifactorial nature of IBD being driven by a number of other factors; including genetics, diet and environment, it is plausible that these factors drive dysbiosis that is not achieved when human microbiota is transferred to the murine model. That was emphasized by the fact that the transfer of disease-associated microbiota was reproducibly capable of inducing inflammation in the *Il10^{-/-}* colitis model, but not the GF WT mice, indicating that genetic susceptibility is required to enhance the colitogenic potential of disease-associated dysbiotic microbiota.

5.4 Multi-omics approach unravels functional alterations of the gut microbiome

In this work, we adopted an integrative multi-omics approach together with humanized mice to discover and validate biomarker signatures associated with IBD. To identify microbial species and microbial genes that were differentially abundant in IBD patients with active disease compared to those with inactive disease, we performed 16S rRNA gene sequencing and shotgun metagenomics. Our data showed that IBD patients show heterogenous microbial community composition and it was rather challenging to define a biomarker signature based solely on taxonomic information. Going beyond microbial community structure, the quantification of vast numbers of metabolites in an untargeted fashion revealed shared functional metabolic pathways under inflammation. These included pathways involved in sulfur metabolism, bacterial toxins secretion systems and purine metabolism. In this context, prior literature showed that enrichment of pathways involved in purine metabolism drive damage of gut epithelia, and treatment with allopurinol blocked the purine pathways and reversed the damage (Chiaro et al., 2017). Similarly, microbial sulfur metabolism has been implicated in colonic inflammation. For example, hydrogen sulfide can be generated from cysteine degradation and as a metabolic by-product of sulfate-reducing bacteria (e.g *Desulfovibrio*). In turn, hydrogen sulfide inhibits Acyl-COA dehydrogenase; an enzyme required for butyrate oxidation which ultimately leads to impairment of butyrate

oxidation, and disruption of the gut homeostasis. However, enrichment of these compounds was not significant, possibly due to the vast inter-personal variations between IBD patients. Performing the same analysis on humanized mice showed a much better separation of metabolite profiles based on inflammation. We observed a significant enrichment of sulfated compounds in inflamed mice; however, an overall regulation of sulfur metabolism seems to be important. Finally, integrating omics data significantly improved the predictive modelling of disease outcome, emphasizing the utility of multi-omics integration in disentangling the complexity of multifactorial diseases like IBD and to understand the contribution of the various causative factors in disease etiology.

Concluding remarks

This work represents one of the first efforts to investigate and characterize microbial functional signatures associated with IBD using an integrative multi-omics approach together with functional validation in gnotobiotic humanized mice. The use of taxonomic profiling highlighted the relevance of bacterial species differentially abundant in inflamed relative to non-inflamed patients. These patterns go in line with prior literature on IBD. Predicting disease activity based on taxonomic composition was rather challenging and yielded an extended list of discriminative taxa. This could be explained by the heterogeneity of patients' gut microbiome, which is an inherent feature underlining disease progression. We approached this challenge by characterizing the gut microbiome at various omics levels, and in an integrated framework. Further, we used the humanized mouse model to study the functional impact of disease-relevant gut microbiota *in vivo* in a susceptible host. The predictability improved significantly when we integrated differential features from microbiome and metabolome datasets. By transferring dysbiotic microbial communities in ex-germ-free mice, we confirmed that despite the heterogeneity at taxonomic level, patients' gut microbiome is reproducibly capable of driving inflammation in the mouse. This confirmed the causal role of the interacting functional networks of microbiota and metabolite we identified.

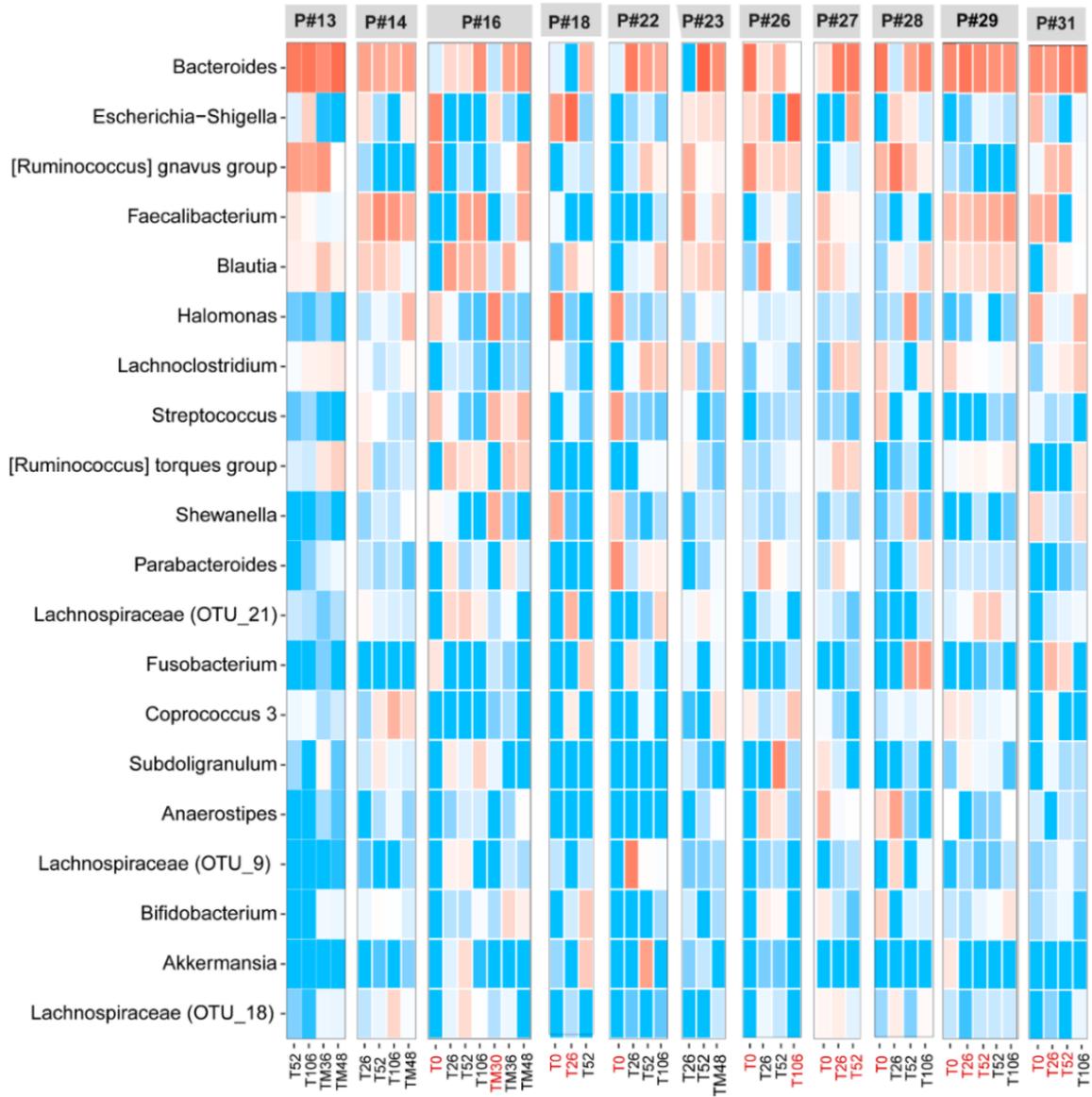
Several future directions are possible to expand this work and to improve the identification and the validation of disease-associated functional signatures in IBD. To identify a functionally relevant interactome of microbiota and metabolite, it is possible to focus on compounds that are interacting with certain bacterial species under inflammatory conditions. The identification, and further quantification of these compounds via quantitative targeting methods could be an approach for mining potential therapeutic targets. Furthermore, integrating the role of environmental factors, such as diet and medications, in addition to omics data integration could

enhance our understanding of disease mechanisms and of medications-driven changes in microbiome. Integrating multi-omics datasets with clinical data of patients longitudinally sampled overtime would further aid in understanding causality and in characterizing the changes preceding disease onset, or foreseeing relapse.

SUPPLEMENTARY FIGURES

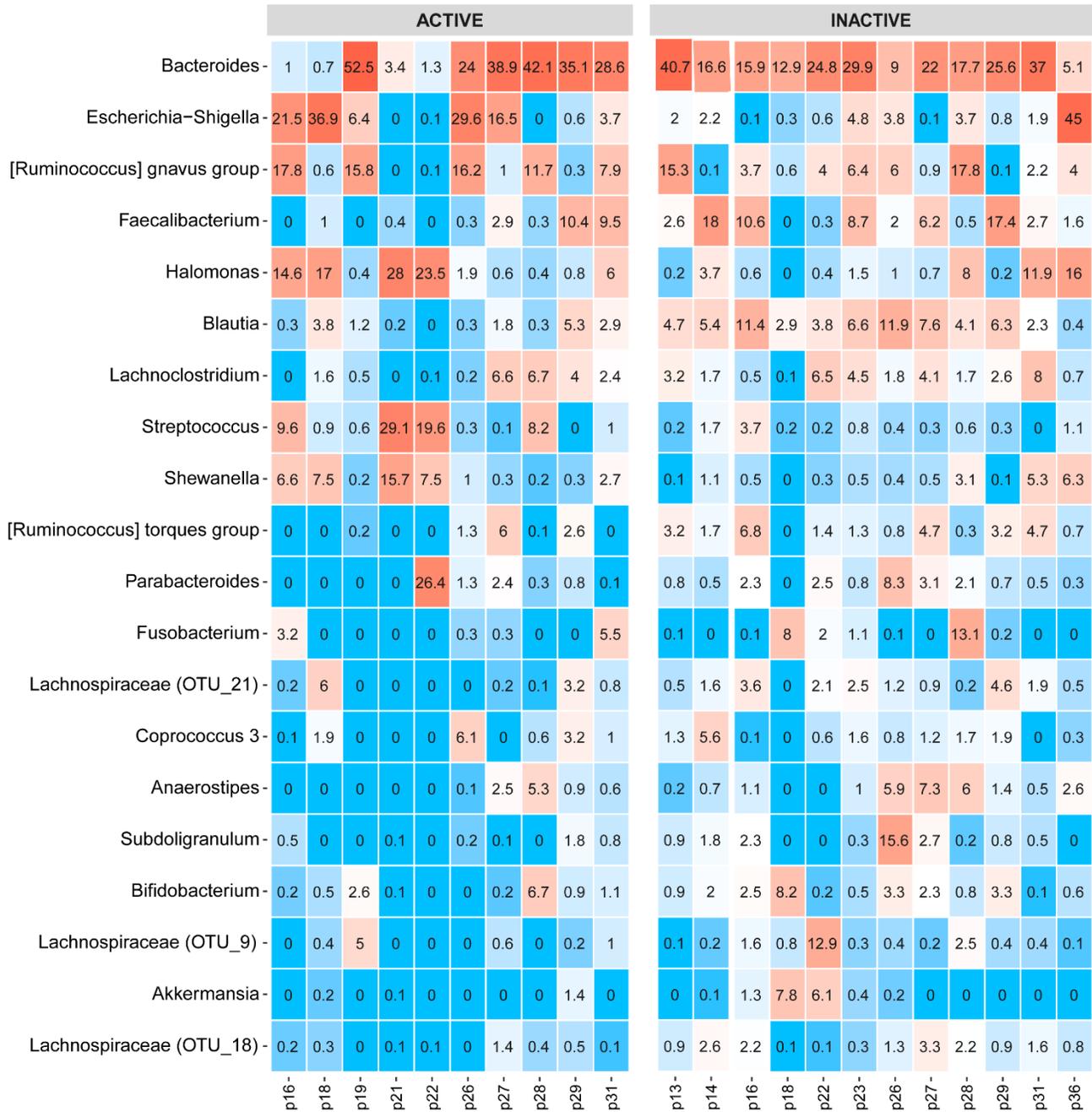
Bacteroides -	22.7	21.7
Escherichia-Shigella -	12.3	3.4
[Ruminococcus] gnavus group -	7.4	6.1
Faecalibacterium -	3.5	6.9
Blautia -	2.3	6.2
Halomonas -	8.2	2.7
Lachnoclostridium -	2.2	3
Streptococcus -	4.8	1.2
[Ruminococcus] torques group -	1.4	2.7
Shewanella -	3.7	1.1
Parabacteroides -	1.8	1.8
Lachnospiraceae (OTU_21) -	1.6	1.8
Fusobacterium -	1.4	1.9
Coprococcus 3 -	1.6	1.5
Subdoligranulum -	0.8	2
Anaerostipes -	0.7	1.9
Lachnospiraceae (OTU_9) -	0.5	1.9
Bifidobacterium -	0.9	1.7
Akkermansia -	0.2	1.9
Lachnospiraceae (OTU_18) -	0.3	1.4
	Active	Inactive

S 1 Relative Abundance of most abundant taxa in colon mucosal biopsies of active and inactive CD patients undergoing HSCT



S 2 Heatmap of most abundant taxa in individual patients, overtime

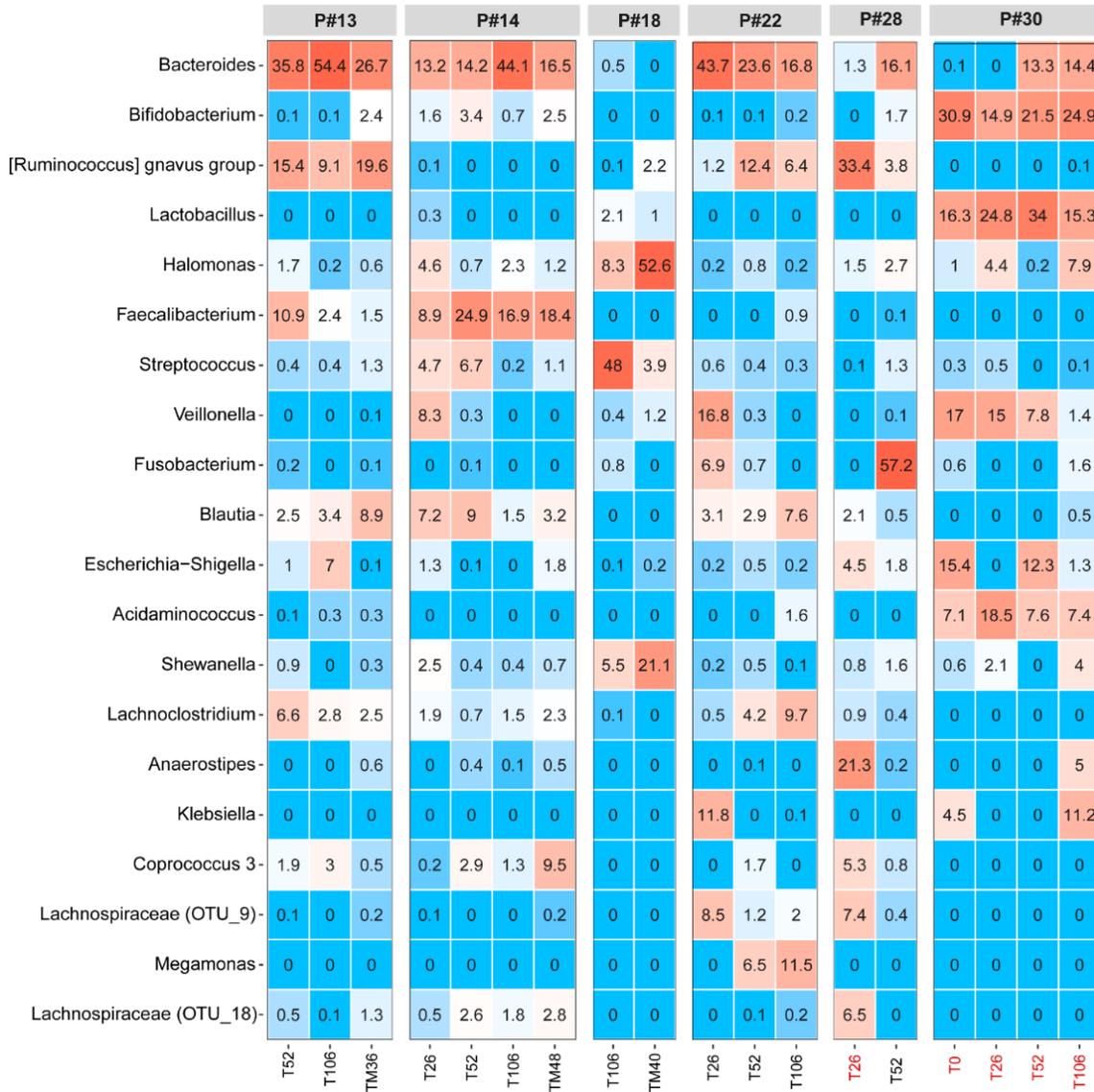
Top: Patient ID number. Bottom: weeks post-HSCT; (Red) active/inflamed and (black) inactive/non-inflamed.



S 3 Heatmap of most abundant taxa in individual patients grouped by activity

Bacteroides-	5.8	23.5
Bifidobacterium-	18.4	1
[Ruminococcus] gnavus group-	6.7	5.4
Lactobacillus-	18.1	0.3
Halomonas-	3	5.8
Faecalibacterium-	0	6.5
Streptococcus-	0.2	5.3
Veillonella-	8.3	2.1
Fusobacterium-	0.4	5.1
Blautia-	0.5	3.8
Escherichia-Shigella-	6.7	1.1
Acidaminococcus-	8.1	0.2
Shewanella-	1.5	2.6
Lachnospiraceae (OTU_9) -	0.2	2.6
Anaerostipes-	5.3	0.2
Klebsiella-	3.1	0.9
Coprococcus 3-	1.1	1.7
Lachnospiraceae (OTU_9) -	1.5	1
Megamonas-	0	1.4
Lachnospiraceae (OTU_18) -	1.3	0.8
	Active	Inactive

S 4 Relative Abundance of most abundant taxa in ileal mucosal biopsies of active and inactive CD patients undergoing HSCT



S 5 Heatmap of most abundant taxa in individual patients, overtime

Top: Patient ID number. Bottom: weeks post-HSCT; (Red) active/inflamed and (black) inactive/non-inflamed.

	ACTIVE		INACTIVE				
	p28-	p30-	p13-	p14-	p18-	p22-	p28-
Bacteroides-	1.3	7	39	22	0.2	28.1	16.1
Bifidobacterium-	0	23.1	0.9	2	0	0.1	1.7
[Ruminococcus] grnavus group-	33.4	0	14.7	0	1.2	6.7	3.8
Lactobacillus-	0	22.6	0	0.1	1.5	0	0
Halomonas-	1.5	3.4	0.8	2.2	30.4	0.4	2.7
Faecalibacterium-	0	0	4.9	17.3	0	0.3	0.1
Streptococcus-	0.1	0.3	0.7	3.2	25.9	0.4	1.3
Veillonella-	0	10.3	0	2.2	0.8	5.7	0.1
Fusobacterium-	0	0.6	0.1	0	0.4	2.5	57.2
Blautia-	2.1	0.1	5	5.2	0	4.5	0.5
Escherichia-Shigella-	4.5	7.3	2.7	0.8	0.1	0.3	1.8
Acidaminococcus-	0	10.2	0.2	0	0	0.6	0
Shewanella-	0.8	1.7	0.4	1	13.3	0.2	1.6
Lachnoclostridium-	0.9	0	4	1.6	0.1	4.8	0.4
Anaerostipes-	21.3	1.3	0.2	0.3	0	0	0.2
Klebsiella-	0	3.9	0	0	0	3.9	0
Coprococcus 3-	5.3	0	1.8	3.5	0	0.6	0.8
Lachnospiraceae (OTU_9) -	7.4	0	0.1	0.1	0	3.9	0.4
Megamonas-	0	0	0	0	0	6	0
Lachnospiraceae (OTU_18) -	6.5	0	0.6	1.9	0	0.1	0

S 6 Heatmap of most abundant taxa in individual patients grouped by activity

LIST OF FIGURES

Figure 1 The Human Gastrointestinal Tract	1
Figure 2 Bile acid metabolism by gut microbiota.....	5
Figure 3 The process of autologous Hematopoietic stem cell transplantation.....	10
Figure 4 Hematopoietic Stem Cell Transplantation Study Design.....	24
Figure 5 REMIND Post-surgical cohort Study Design	26
Figure 6 Biotherapy Study Design	28
Figure 7 HSCT cohort and patients' groups.....	38
Figure 8 IBD patients show fluctuations through the course of disease	40
Figure 9 Time course of intra- and interpersonal variations in fecal bacterial communities	41
Figure 10 Stratification of IBD patients based on disease activity	42
Figure 11 Differential abundant taxa in IBD patients with active disease	43
Figure 12 Stratification of IBD patients Microbiota Composition by timepoint.....	44
Figure 13 Pairwise comparisons and differentially abundant taxa between the subgroups	45
Figure 14 Effect of Co-administration of biological and Immunosuppressive therapy on Shifts of Gut Microbiota in IBD patients post- HSCT.....	46
Figure 15 Pairwise comparisons and differentially abundant taxa in subgroups of IBD patients stratified based on administration of additional drugs	47
Figure 16 Microbial community diversity and richness at baseline could not predict response to HSCT therapy	49
Figure 17 Dynamic fluctuations of microbial community composition overtime in HSCT-treated CD patients	50
Figure 18 Unsupervised clustering emphasized the heterogeneity of gut microbiota structure in CD patients	52
Figure 19 Mucosal-associated microbiota samples included in the analysis	54
Figure 20 Characterization of colonic mucosa-associated dysbiosis.....	55
Figure 21 Temporal and Inter-Individual Variations in colonic-mucosa associated microbiota ...	56
Figure 22 Characterization of ileal mucosa-associated dysbiosis	58
Figure 23 Temporal and Inter-Individual Variations in ileal mucosa-associated microbiota.....	59
Figure 24 Selection of patients for transfer experiments in germfree mice.....	60
Figure 25 Patient's clinical course and microbial composition of fecal samples collected before and over-time after HSCT.....	61
Figure 26 Functional characterization of microbial signatures in IBD using gnotobiotic humanized mice	62
Figure 27 Morphometric and immunologic changes in humanized mice	63
Figure 28 Humanized mice reflect the dysbiotic features of their respective donors	64
Figure 29 Alteration of bile acid metabolism in mice associated with dysbiotic human microbiota	65
Figure 30 Identification of discriminative bile acid repertoire in humanized mice	66
Figure 31 Functional pathways enriched in the fecal microbiota of IBD patients under inflammatory conditions.....	68
Figure 32 Functional pathways enriched in the fecal microbiota of humanized mice under inflammatory conditions.....	69
Figure 33 Comparison of metagenomic profiling of human donors and humanized mice under inflammation.....	71
Figure 34 Alterations of fecal microbiota metabolite profiles in IBD patients	72
Figure 35 Alterations of fecal microbiota metabolite profiles in humanized mice.....	73

Figure 36 Integrative multi-omics signature improves the predictive modelling of disease outcome	74
Figure 37 Transfer efficiency of human microbiota in comparison with mouse microbiota	75
Figure 38 Dissimilarity of humanized mice and respective human donors	76
Figure 39 Intergenerational Microbiota Transfer Efficiency and stability	78
Figure 40 Higher Inter-individual variation in CD compared to UC patients.....	79
Figure 41 Effect of biological therapy on shifts in microbial community composition in IBD patients	80
Figure 42 Baseline Composition and Indication for Response to biotherapy.....	81
Figure 43 Changes in community structure at week 14 post-biotherapy	82
Figure 44 Characterization of post-surgical recurrence associated microbial dysbiosis	83
Figure 45 Microbial shifts pre- and post-surgery in remitters and non-remitters	84

LIST OF TABLES

Table 1 Use of gnotobiotic humanized mice in modelling human disease.....	17
Table 2 Validation studies on transplanted microbiota	18
Table 3 Characteristics of patients initiating Hematopoietic Stem Cell Transplantation (HSCT) therapy	25
Table 4 Rutgeerts Scoring System for assessment of post-surgical recurrence in CD patients .	27
Table 5 Characteristics of the included patients initiating biotherapy treatment	29
Table 6 Protocol for qPCR gene expression analysis	37
Table 7 Changes of bacterial taxa in humanization	76
Table 8 Core and rare bacterial phylotypes in human/mouse donors and respective recipient mice	77

SUPPLEMENTARY FIGURES

S 1 Relative Abundance of most abundant taxa in colon mucosal biopsies of active and inactive CD patients undergoing HSCT	95
S 2 Heatmap of most abundant taxa in individual patients, overtime	96
S 3 Heatmap of most abundant taxa in individual patients grouped by activity	97
S 4 Relative Abundance of most abundant taxa in ileal mucosal biopsies of active and inactive CD patients undergoing HSCT	98
S 5 Heatmap of most abundant taxa in individual patients, overtime	99
S 6 Heatmap of most abundant taxa in individual patients grouped by activity	100

Abbreviations

AHSCT	Autologous Hematopoetic Stem Cell Transplantation
ASTIC	The Autologous Stem Cell Transplantation International Crohn's Disease trial
AUC	Area Under the Curve
BA	Baseline Active
BSH	bile salt hydrolase
C3	Complement C3
CA	Cholic Acid
CD	Crohn's Disease
CDAI	Crohn's Disease Activity Index (Clinical)
CDCA	Chenodeoxycholic Acid
CDEIS	Crohn's disease index of severity (Endoscopic)
CFU	Colony Forming Unit
COG	Clusters of Orthologous Groups
DCA	Deoxycholic Acid
E.coli	Escherichia coli
E.faecalis	Enterococcus faecalis
F. prausnitzii	Faecalibacterium prausnitzii
FMT	Faecal Microbiota Transplantation
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GF	Germ-free
GI	Gastrointestinal
GWAS	Genome-wide Association Studies
H&E	Hematoxylin & Eosin
H.pylori	Helicobacter pylori
HBI	Harvey-bradshaw index (Clinical)
HMb	Human Microbiota
HMP	Human Microbiome Project
HPLC	High Performance Liquid Chromatography
HSCT	Hematopoetic Stem Cell Transplantation
IBD	Inflammatory Bowel Disease
IFNγ	Interferon gamma
IgA	Immunoglobulin A
IL	interleukin
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Lysogeny Broth
LCA	lithocholic acid
LC-MS/MS	Liquid chromatography–Mass spectrometry
LC–TOF-MS	Liquid Chromatography–Time of Flight–Mass Spectrometry
LEfSe	Linear discriminant analysis Effect Size
Ln	lean

LPS	Lipopolysaccharides
MCA	Muricholic acid
MDS	Multi-dimensional Scaling plot
MetaHIT Project	Metagenomics of the Human Intestinal Tract Project
MLN	Mesenteric Lymph Node
MMb	Mouse Microbiota
MS	Mass Spectrometry
NGS	Next-Generation Sequencing
NMDS	Non-Metric Multidimensional Scaling
NMR	Nuclear Magnetic Resonance
NR	Non-responder
NR W TMT	Non-responder with treatment
NR WO TMT	Non-responder without treatment
Ob	Obese
OTU	Operational Taxonomic Unit
PBS	Phosphate Buffered Saline
PCA	Principle Component Analysis
PCR	Polymerase Chain Reaction
pH	potential of hydrogen
PICRUSt tool	Phylogenetic Investigation of Communities by Reconstruction of Unobserved States
PLS	Partial Least Squares
PLSDA	Partial least squares regression
PR	Primary response
qPCR	Quantitative Polymerase Chain Reaction
R W TMT	Responder with treatment
R WO TMT	Responder without treatment
RDP	The Ribosomal Database Project
Re	Relapse
Rm	Remission
ROC	Receiver Operating Curve
rRNA	Ribosomal ribonucleic acid
RS	Rutgeerts score
RT	Room Temperature
SCFA	Short Chain Fatty Acids
SES-CD	Simple endoscopic score for Crohn's disease
SPF	Specific Pathogen Free
TCA	Taurocholic Acid
TNFα	Tumor necrosis factor alpha
UC	Ulcerative Colitis
UCEIS	Ulcerative Colitis Endoscopic Index of Severity
UHPLC	Ultra High Performance Liquid Chromatographs

UPL	Universal Probe Library
V3/V4	Variable regions 3 and 4
VIP	Variable importance in projection
W/V	Weigh per volume
WCA	Wilkins-Chalgren-Agar
WGS	Whole-genome Sequencing
WT	Wildtype

References

- Acheson, E.D. (1960). The distribution of ulcerative colitis and regional enteritis in United States veterans with particular reference to the Jewish religion. *Gut* 1, 291–293.
- Amre, D.K., D’Souza, S., Morgan, K., Seidman, G., Lambrette, P., Grimard, G., Israel, D., MacK, D., Ghadirian, P., Deslandres, C., et al. (2007). Imbalances in dietary consumption of fatty acids, vegetables, and fruits are associated with risk for crohn’s disease in children. *Am. J. Gastroenterol.* 102, 2016–2025.
- Ananthakrishnan, A.N. et al. (2017). Gut Microbiome Function Predicts Response to Anti-integrin Biologic Therapy in Inflammatory Bowel Diseases. *Cell Host Microbe* 21, 603-610.e3.
- Anderson, C. a, Boucher, G., Lees, C.W., and Franke, A. (2011). NIH Public Access. *Nat. Genet.* 43, 246–252.
- Andres Ekblom, Charles Helmck, M.Z., and Hans-Olov, A. (1991). The Epidemiology of Inflammatory Bowel Disease: A Large, Population-Based Study in Sweden. *Gastroenterology* 100, 350–358.
- Arrieta, M., Sadarangani, M., Brown, E.M., Russell, S.L., and Nimmo, M. (2016a). A humanized microbiota mouse model of ovalbumin-induced lung inflammation. *Gut Microbes* 7, 342–352.
- Arrieta, M.C., Walter, J., and Finlay, B.B. (2016b). Human Microbiota-Associated Mice: A Model with Challenges. *Cell Host Microbe*.
- Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.M., et al. (2011). Enterotypes of the human gut microbiome. *Nature*.
- Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A., and Gordon, J.I. (2005). Host[ndash]bacterial mutualism in the human intestine. *Science* (80-.). 307, 1915–1920.
- Bäckhed, 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*.
- Baker, G.C., Smith, J.J., and Cowan, D.A. (2003). Review and re-analysis of domain-specific 16S primers. *J. Microbiol. Methods* 55, 541–555.
- Balish, E., and Warner, T. (2002). Enterococcus faecalis induces inflammatory bowel disease in interleukin-10 knockout mice. *Am. J. Pathol.* 160, 2253–2257.
- Becker, C., Neurath, M.F., and Wirtz, S. (2015). The intestinal microbiota in inflammatory bowel disease. *ILAR J*.
- Berg, D.J., Davidson, N., Kühn, R., Müller, 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. *J. Clin. Invest.* 98, 1010–1020.
- Bernstein, C.N., and Shanahan, F. (2008a). Disorders of a modern lifestyle-reconciling the epidemiology of inflammatory bowel diseases. *Gut* 1185–1192.
- Bernstein, C.N., and Shanahan, F. (2008b). Disorders of a modern lifestyle-reconciling the epidemiology of inflammatory bowel diseases. *Gut* 1185–1192.
- Bernstein, C.N., Rawsthorne, P., Cheang, M., and Blanchard, J.F. (2006). A population-based case control study of potential risk factors for IBD. *Am. J. Gastroenterol.* 101, 993–1002.
- den Besten, G., van Eunen, K., Groen, A.K., Venema, K., Reijngoud, D.-J., and Bakker, B.M. (2013). The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J. Lipid Res.*
- Blanton, L. V. (2016). Gut bacteria that rescue growth impairments transmitted by immature microbiota from undernourished children. *351*, 1–18.
- Bleich, A., Mähler, M., Most, C., Leiter, E., Liebler-Tenorio, E., Elson, C., Hedrich, H., Schlegelberger, B., and Sundberg, J. (2004). Refined histopathologic scoring system improves power to detect colitis QTL in mice. *Mamm. Genome* 15, 865–871.
- BM, C. (1989). A meta-analysis of the role of smoking in inflammatory bowel disease. *Dig Dis Sci.* 34, 1841–18.
- Bringer, S., Swidsinski, A., Glasser, A.L., Barnich, N., Agne, M., Beaugerie, L., and Fre, J. (2004). High Prevalence of Adherent-Invasive *Escherichia coli* Associated With Ileal Mucosa in Crohn's Disease. *Gastroenterology* 412–421.
- Burt, R.K., Craig, R.M., Milanetti, F., Quigley, K., Gozdzia, P., Bucha, J., Testori, A., Halverson, A., Verda, L., Villiers, W.J.S. De, et al. (2011). Autologous nonmyeloablative hematopoietic stem cell transplantation in patients with severe anti-TNF refractory Crohn disease : long-term follow-up. *Blood* 116, 6123–6132.
- Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al. (2010). correspondence QIIME allows analysis of high-throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nat. Publ. Gr.* 7, 335–336.
- Caporaso, J.G., Lauber, C.L., Costello, E.K., Berg-Lyons, D., Gonzalez, A., Stombaugh, J., Knights, D., Gajer, P., Ravel, J., Fierer, N., et al. (2011). Moving pictures of the human microbiome. *Genome Biol.* 12, R50.
- Chamaillard, M., and Radulovic, K. (2016). Defining dysbiosis threatens Koch's postulates and current dogma on the role of Paneth cells in Crohn's disease. *Gut*.
- Chiaro, T.R., Soto, R., Stephens, W.Z., Kubinak, J.L., Petersen, C., Gogokhia, L., Bell, R., Delgado, J.C., Cox, J., Voth, W., et al. (2017). A member of the gut mycobiota modulates host purine metabolism exacerbating colitis in mice. *Sci. Transl. Med.* 9044, 1–12.

- Chung, H., Pamp, S.J., Hill, J.A., Surana, N.K., Edelman, S.M., Troy, E.B., Reading, N.C., Villablanca, E.J., Wang, S., Mora, J.R., et al. (2012). Gut immune maturation depends on colonization with a host-specific microbiota. *Cell*.
- Consortium, H. (2013). Structure, Function and Diversity of the Healthy Human Microbiome. *Nature* 486, 207–214.
- Costea, P.I., Hildebrand, F., Manimozhayan, A., Bäckhed, F., Blaser, M.J., Bushman, F.D., de Vos, W.M., Ehrlich, S.D., Fraser, C.M., Hattori, M., et al. (2018). Enterotypes in the landscape of gut microbial community composition. *Nat. Microbiol.* 3, 8–16.
- DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., and Andersen, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* 72, 5069–5072.
- Dhiman, R. (2012). Gut Microbiota , Inflammation and Hepatic Encephalopathy : A Puzzle with a Solution in Sight. *J. Clin. Exp. Hepatol.* 2, 1–4.
- Ding, H., Wang, T., Hooper, L. V, Koh, G.Y., Nagy, A., Semenkovich, C.F., and Gordon, J.I. (2004). The gut microbiota as an environmental factor that regulates fat storage. 101.
- Dorrestein, P.C., Mazmanian, S.K., Knight, R., Jolla, L., States, U., Diego, C.S., Jolla, L., and Engineering, B. (2015). From microbiomess to metabolomes to function during host-microbial interactions. *Immunity* 40, 824–832.
- Duncan, S.H., Hold, G.L., Barcenilla, A., Stewart, C.S., and Flint, H.J. (2002). *Roseburia intestinalis* sp . nov ., a novel saccharolytic , butyrate-producing bacterium from human faeces. *Int. J. Syst. Evolutionary Microbiol.* 1615–1620.
- Edgar, R.C. (2013). UPARSE : highly accurate OTU sequences from microbial amplicon reads. *Nat. Commun.* 10.
- 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.
- Ellinghaus, D., Jostins, L., Cortes, A., Park, Y.R., Raychaudhuri, S., Pouget, J.G., Folseraas, T., Wang, Y., Esko, T., Metspalu, A., et al. (2016). Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. *Nat. Genet.* 48, 510–518.
- Engels, C., Ruscheweyh, H., Beerenwinkel, N., and Broderick, N.A. (2016). The Common Gut Microbe *Eubacterium hallii* also Contributes to Intestinal Propionate Formation. 7, 1–12.
- Erben, U., Loddenkemper, C., Doerfel, K., Spieckermann, S., Haller, D., Heimesaat, M.M., Zeitz, M., Siegmund, B., Kuhl, A.A., and Kuhl, A.A. (2014). Original Article A guide to histomorphological evaluation of intestinal inflammation in mouse models. *Int J Clin Exp Pathol* 7, 4557–4576.

- Eun, C.S., Mishima, Y., Wohlgemuth, S., Liu, B., Bower, M., Carroll, I.M., and Sartor, R.B. (2014). Induction of bacterial antigen-specific colitis by a simplified human microbiota consortium in gnotobiotic interleukin-10^{-/-} mice. *Infect. Immun.* *82*, 2239–2246.
- Flemer, B., Lynch, D.B., Brown, J.M.R., Jeffery, I.B., Ryan, F.J., Claesson, M.J., Riordain, M.O., Shanahan, F., and Toole, P.W.O. (2017). Tumour-associated and non-tumour-associated microbiota in colorectal cancer. *Gut* 633–643.
- Franke, A., Balschun, T., Karlsen, T.H., Sventoraityte, J., Nikolaus, S., Mayr, G., Domingues, F.S., Albrecht, M., Nothnagel, M., Ellinghaus, D., et al. (2008). Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nat. Genet.* *40*, 1319–1323.
- Franke, A., McGovern, D.P.B., Barrett, J.C., Wang, K., Radford-Smith, G.L., Ahmad, T., Lees, C.W., Balschun, T., Lee, J., Roberts, R., et al. (2010). Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat. Genet.* *42*, 1118–1125.
- Fuentes, S., Rossen, N.G., Spek, M.J. Van Der, Hartman, J.H.A., Huuskonen, L., Korpela, K., Salojärvi, J., Aalvink, S., Vos, W.M. De, Haens, G.R.D., et al. (2002). Microbial shifts and signatures of long-term remission in ulcerative colitis after faecal microbiota transplantation. *FEMS Microbiol. Ecol.* *11*, 1877–1889.
- Fujimoto, T., Imaeda, H., Takahashi, K., Kasumi, E., and Bamba, S. (2013). Decreased abundance of *Faecalibacterium prausnitzii* in the gut microbiota of Crohn's disease. *28*, 613–619.
- Fyderek, K., Strus, M., Kowalska-duplaga, K., Gosiewski, T., Wędrychowicz, A., Jedynak-wąsowicz, U., Śladek, M., Pieczarkowski, S., Adamski, P., and Kochan, P. (2009). Mucosal bacterial microflora and mucus layer thickness in adolescents with inflammatory bowel disease. *World J Gastroenterol.* *15*, 5287–5294.
- Gensollen, T., Iyer, S.S., Kasper, D.L., and Blumberg, R.S. (2016). How colonization by microbiota in early life shapes the immune system. *Science* (80-).
- Gent, A.E., Hellier, M.D., Grace, R.H., Swarbrick, E.T., and Coggon, D. (1994). Inflammatory bowel disease and domestic hygiene in infancy. *Lancet* *343*, 766–767.
- Gerritsen, J., Hornung, B., Renckens, B., Hijum, S.A.F.T. Van, Martins, V.A.P., Rijkers, G.T., Schaap, P.J., Vos, W.M. De, and Smidt, H. (2017). Genomic and functional analysis of *Romboutsia ilealis* CRIB T reveals adaptation to the small intestine. 1–28.
- Gevers, D., Kugathasan, S., Denson, L.A., Vázquez-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.
- Glocker, E., and Kotlarz, D. (2009). Inflammatory bowel disease and mutations affecting the interleukin-10 receptor. ... *Engl. J. ...* *361*, 2033–2045.
- Glocker, E.-O., Kotlarz, D., Boztug, K., Gertz, E.M., Schäffer, A.A., Noyan, F., Perro, M.,

- Diestelhorst, J., Allroth, A., Murugan, D., et al. (2009). Inflammatory Bowel Disease and Mutations Affecting the Interleukin-10 Receptor. *N. Engl. J. Med.* 361, 2033–2045.
- Godon, J., Zumstein, E., Dabert, P., and Habouzit, R.I.C. (1997). Molecular microbial diversity of an anaerobic digester as determined by small-subunit rDNA sequence analysis . *Molecular Microbial Diversity of an Anaerobic Digester as Determined by Small-Subunit rDNA Sequence Analysis. Applied Environ. Microbiol.* 63, 2802–2813.
- Goodacre, R., Vaidyanathan, S., Dunn, W.B., Harrigan, G.G., and Kell, D.B. (2004). Metabolomics by numbers: Acquiring and understanding global metabolite data. *Trends Biotechnol.* 22, 245–252.
- Gootenberg, D.B., and Turnbaugh, P.J. (2011). Companion animals symposium: Humanized animal models of the microbiome. *J. Anim. Sci.*
- Hai-Qin, M., Ting-Ting, Y., Xiao-Jing, Z., and Yi Zhang, H.-J.Z. (2018). Fecal microbial dysbiosis in Chinese patients with inflammatory bowel disease. *World J Gastroenterol.* 24, 1464–1477.
- Halfvarson, J., Brislawn, C.J., Lamendella, R., Walters, W.A., Bramer, L.M., Bonfiglio, F., Gonzalez, A., McClure, E.E., Dunklebarger, M.F., Knight, R., et al. (2017). Dynamics of the human gut microbiome in Inflammatory Bowel Disease. *Nat. Microbiol.* 2.
- Hansen, R., Berry, S.H., Mukhopadhyay, I., Thomson, J.M., Saunders, K.A., Nicholl, C.E., Bisset, W.M., Loganathan, S., Mahdi, G., Kastner, D., et al. (2013). The Microaerophilic Microbiota of De-Novo Paediatric Inflammatory Bowel Disease : The BISCUIT Study. 8.
- Hawkey, C.J., Allez, M., Clark, M.M., Labopin, M., Lindsay, J.O., Ricart, E., Rogler, G., Rovira, M., Satsangi, J., Danese, S., et al. (2015). Autologous hematopoietic stem cell transplantation for refractory Crohn disease a randomized clinical trial. *JAMA - J. Am. Med. Assoc.*
- Herington, J.L., Crispens, M.A., Carvalho-macedo, A.C., Camargos, A.F., Lebovic, D.I., Bruner, K.L., and Osteen, K.G. (2012). *NIH Public Access.* 95, 1295–1301.
- Heslop, J.A., Hammond, T.G., Santeramo, I., Tort Piella, A., Hopp, I., Zhou, J., Baty, R., Graziano, E.I., Proto Marco, B., Caron, A., et al. (2015). Concise Review: Workshop Review: Understanding and Assessing the Risks of Stem Cell-Based Therapies. *Stem Cells Transl. Med.* 4, 389–400.
- Hommes, D., Colombel, J.F., Emery, P., Greco, M., and Sandborn, W.J. (2012). Changing Crohn’s disease management: Need for new goals and indices to prevent disability and improve quality of life. *J. Crohn’s Colitis* 6, 224–234.
- Hwang, J.-M. (2008). Surgery for inflammatory bowel disease. *World J. Gastroenterol.* 14, 2678.
- Ianiro, J.M.G., and Hansen, I.M.R. (2018). Review article : the gut microbiome in inflammatory bowel disease — avenues for microbial management. 26–42.
- Im, J.P., Ye, B.D., Kim, Y.S., and Kim, J.S. (2018). Changing treatment paradigms for

the management of inflammatory bowel disease. *Korean J. Intern. Med.* 33, 28–35.

INGER CAMILLA SOLBERG, MORTEN H. VATN, OLE HØIE, NJAAL STRAY, JOSTEIN SAUAR, JØRGEN JAHNSEN, BJØRN MOUM, IDAR LYGREN, and T.I.S.G. (2007). Clinical Course in Crohn's Disease: Results of a Norwegian Population-Based Ten-Year Follow-Up Study. *Clin. Gastroenterol. Hepatol.* 5, 1430–1438.

Jakobsson, H.E., Jernberg, C., Andersson, A.F., Sjölund-Karlsson, M., Jansson, J.K., and Engstrand, L. (2010). Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *PLoS One* 5.

James P., M., Jens, W., Gerald W., T., Susan L., T., and Balfour, S. (2009). *Bifidobacterium animalis* Causes Extensive Duodenitis and Mild Colonic Inflammation in Monoassociated Interleukin-10 Deficient Mice. *Inflamm. Bowel Dis.* 15, 1022–1031.

Jauregui-Amezaga, A., Rovira, M., Marín, P., Salas, A., Pinó-Donnay, S., Feu, F., Elizalde, J.I., Fernández-Avilés, F., Martínez, C., Gutiérrez, G., et al. (2016). Improving safety of autologous haematopoietic stem cell transplantation in patients with Crohn's disease. *Gut*.

Jeon, S.G., Kayama, H., Ueda, Y., Takahashi, T., Asahara, T., Tsuji, H., Tsuji, N.M., Kiyono, H., Ma, J.S., Kusu, T., et al. (2012). Probiotic *Bifidobacterium breve* Induces IL-10-Producing Tr1 Cells in the Colon. *Plos Pathog.* 8, 1–15.

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.

Julià, A., Domènech, E., Chaparro, M., García-Sánchez, V., Gomollón, F., Panés, J., Mañosa, M., Barreiro-De Acosta, M., Gutiérrez, A., Garcia-Planella, E., et al. (2014). A genome-wide association study identifies a novel locus at 6q22.1 associated with ulcerative colitis. *Hum. Mol. Genet.* 23, 6927–6934.

Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M. (2004). The KEGG resource for deciphering the genome. *Nucleic Acids Res.* 32.

Katrina M de Lange, L.M.C.B. (2017). Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat. Genet.* 49, 256–261.

Keating, A., DaSilva, G., Pérez, W.S., Gupta, V., Cutler, C.S., Ballen, K.K., Cairo, M.S., Camitta, B.M., Champlin, R.E., Gajewski, J.L., et al. (2013). Autologous blood cell transplantation versus HLA-identical sibling transplantation for acute myeloid leukemia in first complete remission: A registry study from the center for international blood and marrow transplantation research. *Haematologica* 98, 185–192.

Kenny, E.E., Pe'er, I., Karban, A., Ozelius, L., Mitchell, A.A., Ng, S.M., Erazo, M., Ostrer, H., Abraham, C., Abreu, M.T., et al. (2012). A genome-wide scan of ashkenazi jewish crohn's disease suggests novel susceptibility loci. *PLoS Genet.* 8.

Kibe, R., Sakamoto, M., Yokota, H., Aiba, Y., Koga, Y., Benno, Y., and Ishikawa, H.

- (2005). Movement and Fixation of Intestinal Microbiota after Administration of Human Feces to Germfree Mice. *Appl Env. Microbiol* *71*, 3171–3178.
- Kim, S.C., Tonkonogy, S.L., Albright, C.A., Tsang, J., Balish, E.J., Braun, J., Huycke, M.M., and Sartor, R.B. (2005). Variable phenotypes of enterocolitis in interleukin 10-deficient mice monoassociated with two different commensal bacteria. *Gastroenterology*.
- Kim, S.C., Tonkonogy, S.L., Karrasch, T., Jobin, C., and Balfour Sartor, R. (2007). Dual-association of gnotobiotic IL-10-/- mice with 2 nonpathogenic commensal bacteria induces aggressive pancolitis. *Inflamm. Bowel Dis.* *13*, 1457–1466.
- Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F.O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* *41*, 1–11.
- Knights, D., Costello, E.K., and Knight, R. (2011a). Supervised classification of human microbiota. *FEMS Microbiol. Rev.* *35*, 343–359.
- Knights, D., Parfrey, L.W., Zaneveld, J., Lozupone, C., and Knight, R. (2011b). Human-associated microbial signatures: Examining their predictive value. *Cell Host Microbe*.
- Knights, D., Ward, T.L., McKinlay, C.E., Miller, H., Gonzalez, A., McDonald, D., and Knight, R. (2014). Rethinking enterotypes. *Cell Host Microbe*.
- Koloski, N.A., Bret, L., and Radford-Smith, G. (2008). Hygiene hypothesis in inflammatory bowel disease: A critical review of the literature. *World J. Gastroenterol.* *14*, 165–173.
- Kostic, A.D., Gevers, D., Pedamallu, C.S., Michaud, M., Duke, F., Earl, A.M., Ojesina, A.I., Jung, J., Bass, A.J., Liu, C., et al. (2012). Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. 292–298.
- Kostic, A.D., Gevers, D., Siljander, H., Vatanen, T., Peet, A., Tillmann, V., Mattila, I., Franzosa, E.A., Vaarala, O., Goffau, M. De, et al. (2016). The Dynamics of the Human Infant Gut Microbiome in Development and in Progression towards Type 1 Diabetes. *Cell Host Microbe* *17*, 260–273.
- Kühn, R., Löhler, J., Rennick, D., Rajewsky, K., and Müller, W. (1993). Interleukin-10-deficient mice develop chronic enterocolitis. *Cell* *75*, 263–274.
- Lagkouvardos, I., and Kl, K. (2015). Gut metabolites and bacterial community networks during a pilot intervention study with flaxseeds. 1614–1628.
- Lagkouvardos, I., Joseph, D., Kapfhammer, M., Giritli, S., Horn, M., Haller, D., and Clavel, T. (2016). IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. *Sci. Rep.* *6*, 1–9.
- Lagkouvardos, I., Fischer, S., Kumar, N., and Clavel, T. (2017). Rhea : a transparent and modular R pipeline for microbial profiling based on 16S rRNA gene amplicons. *PeerJ*.

- Lashner, B.A., and Loftus, E. V. (2006). True or false? The hygiene hypothesis for Crohn's disease. *Am. J. Gastroenterol.* *101*, 1003–1004.
- Lee, W., and Hase, K. (2014). Gut microbiota-generated metabolites in animal health and disease. *Nat. Chem. Biol.* *10*, 416–424.
- Lepage P, Häslér R, Spehlmann ME, Rehman A, Zvirbliene A, Begun A, Ott S, Kupcinkas L, Doré J, Raedler A, S.S. (2011). Twin study indicates loss of interaction between microbiota and mucosa patients with ulcerative colitis. *Gastroenterology* *141*, 227–236.
- Li, H., and Sykes, M. (2012). Emerging concepts in haematopoietic cell transplantation. *Nat. Rev. Immunol.* *12*, 403–416.
- Liu, J.Z., Sommeren, S. Van, Huang, H., Ng, S.C., and Alberts, R. (2015). Europe PMC Funders Group Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* *47*, 979–986.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* *25*, 402–408.
- Lloyd-Price, J., Andrews, E., Ajami, N.J., Bonham, K.S., Brislawn, C.J., Casero, D., Courtney, H., Gonzalez, A., Graeber, T.G., Hall, A.B., et al. (2019). Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* *569*, 655–662.
- López-García, A., Rovira, M., Jauregui-Amezaga, A., Marín, P., Barastegui, R., Salas, A., Ribas, V., Feu, F., Elizalde, J.I., Fernández-Avilés, F., et al. (2017). Autologous haematopoietic stem cell transplantation for refractory crohn's disease: Efficacy in a single-centre cohort. *J. Crohn's Colitis*.
- Lopez-siles, M., Khan, T.M., Duncan, S.H., Harmsen, H.J.M., Garcia-gil, L.J., and Flint, H.J. (2012). Cultured Representatives of Two Major Phylogroups of Human Colonic *Faecalibacterium prausnitzii* Can Utilize Pectin , Uronic Acids , and Host-Derived Substrates for Growth. *420–428*.
- Lopez-Cubero, S.O., Sullivan, K.M., and McDonald, G.B. (1998). Course of Crohn's disease after allogeneic marrow transplantation. *Gastroenterology* *114*, 433–440.
- Loubinoux, J., Bronowicki, J., Pereira, I.A.C., Mougénel, J., and Le, A.E. (2002). Sulfate-reducing bacteria in human feces and their association with inflammatory bowel diseases. *40*.
- Lozupone, C.A., Stombaugh, J.I., Gordon, J.I., Jansson, J.K., and Knight, R. (2012). Diversity, stability and resilience of the human gut microbiota. *Nature*.
- Lozupone, C.A., Stombaugh, J., Gonzalez, A., Ackermann, G., Wendel, D., Vázquez-Baeza, Y., Jansson, J.K., Gordon, J.I., and Knight, R. (2013). Meta-analyses of studies of the human microbiota. *Genome Res.* *23*, 1704–1714.
- Lu, J.J., Perng, C.L., Lee, S.Y., and Wan, C.C. (2000). Use of PCR with universal

- primers and restriction endonuclease digestions for detection and identification of common bacterial pathogens in cerebrospinal fluid. *J. Clin. Microbiol.* **38**, 2076–2080.
- Maidak, B.L., Olsen, G.J., Larsen, N., Overbeek, R., McCaughey, M.J., and Woese, C.R. (1997). The RDP (Ribosomal Database Project). *Nucleic Acids Res.* **25**, 109–110.
- Majhail, N.S., Farnia, S.H., Carpenter, P.A., Champlin, R.E., Crawford, S., Marks, D.I., Omel, J.L., Orchard, P.J., and Palmer, J. (2015). Indications for Autologous and Allogeneic Hematopoietic Cell Transplantation: Guidelines from the American Society for Blood and Marrow Transplantation HHS Public Access. *Biol Blood Marrow Transpl.* **21**, 1863–1869.
- Manchester, J.K., Semenkovich, C.F., and Gordon, J.I. (2007). Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *104*.
- Markus, G., Gronow, S., Zeytun, A., Nolan, M., Lucas, S., Hammon, N., Deshpande, S., Cheng, J., Pitluck, S., Liolios, K., et al. (2011). Complete genome sequence of *Odoribacter splanchnicus* type strain (1651/6T). *Stand. Genomic Sci.* **4**, 200–209.
- Martinez C, Antolin M, Santos J, Torrejon A, Casellas F, Borrueal N, Guarner F, M.J. (2008). Unstable composition of the fecal microbiota in ulcerative colitis during clinical remission. *Am. J. Gastroenterol.* **103**, 643–648.
- Martiny, J.B.H., Jones, S.E., Lennon, J.T., and Martiny, A.C. (2015). Microbiomes in light of traits: A phylogenetic perspective. *Science (80-.)*. **350**.
- Maubert, M.A., Quevrain, E., Chain, F., Marquant, R., Kharrat, P., Carlier, L., Bermudez-Humaran, L.G., Pigneur, B., Lequin, O., Bridonneau, C., et al. (2014). Identification of an anti-inflammatory protein from faecalibacterium *prausnitzii*, a deficient commensal bacteria implicated in Crohn's disease. *J. Crohn's Colitis* **1**), S-23.
- McCabe, K.M., Zhang, Y.-H., Huang, B.-L., Wagar, E.A., and McCabe, E.R.B. (1999). Bacterial Species Identification after DNA Amplification with a Universal Primer Pair. *Mol. Genet. Metab.* **66**, 205–211.
- McCarthy, J., O'Mahony, L., O'Callaghan, L., Sheil, B., Vaughan, E.E., Fitzsimons, N., Fitzgibbon, J., O'Sullivan, G.C., Kiely, B., Collins, J.K., et al. (2003). Double blind, placebo controlled trial of two probiotic strains in interleukin 10 knockout mice and mechanistic link with cytokine balance. *Gut* **52**, 975–980.
- Mehta, R.S., Abu-ali, G.S., Drew, D.A., Lloyd-price, J., Subramanian, A., Lochhead, P., Joshi, A.D., Ivey, K.L., Khalili, H., Brown, G.T., et al. (2018). Stability of the human faecal microbiome in a cohort of adult men. *Nat. Microbiol.*
- Mikhailov, T.A., and Furner, S.E. (2009). Breastfeeding and genetic factors in the etiology of inflammatory bowel disease in children. *World J. Gastroenterol.* **15**, 270–279.
- 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., et al. (2012). Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* **142**, 46-54.e42.

- 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 Biol.*
- Munkholm P1, Langholz E, Nielsen OH, Kreiner S, B. V. (1992). Incidence and prevalence of Crohn's disease in the county of Copenhagen, 1962-87: a sixfold increase in incidence. Author information. *Scand J Gastroenterol.* 27, 609–614.
- Muraro, P.A., Robins, H., Malhotra, S., Howell, M., Phippard, D., Desmarais, C., Paula, A. De, Sousa, A., Griffith, L.M., Lim, N., et al. (2014). Brief report T cell repertoire following autologous stem cell transplantation for multiple sclerosis. *J. Clin. Invest.* 124, 1168–1172.
- 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. *Inflamm. Bowel Dis.* 11, 481–487.
- N.Barnichab, J.Denizotab, A.D.-M. (2013). E. coli-mediated gut inflammation in genetically predisposed Crohn's disease patients. *Pathol. Biol.* 61, e65–e69.
- Nagao-kitamoto, H., Shreiner, A.B., Iii, M.G.G., Kitamoto, S., Ishii, C., Hirayama, A., Kuffa, P., El-zaatari, M., Grasberger, H., Seekatz, A.M., et al. (2016). Functional Characterization of In fl ammatory Bowel. *Cell. Mol. Gastroenterol. Hepatol.* 2, 468–481.
- Nagao-Kitamoto, H., Shreiner, A.B., Gilliland, M.G., Kitamoto, S., Ishii, C., Hirayama, A., Kuffa, P., El-Zaatari, M., Grasberger, H., Seekatz, A.M., et al. (2016). Functional Characterization of Inflammatory Bowel Disease-Associated Gut Dysbiosis in Gnotobiotic Mice. *CMGH.*
- Naito, Y., Uchiyama, K., and Takagi, T. (2018). next generation beneficial microbe : *Akkermansia muciniphila.* 63, 33–35.
- Naser, S.A., Sagramsingh, S.R., Naser, A.S., and Thanigachalam, S. (2014). *Mycobacterium avium* subspecies paratuberculosis causes Crohn's disease in some inflammatory bowel disease patients. *World J. Gastroenterol.* 20, 7403–7415.
- Natividad, J.M.M., and Verdu, E.F. (2013). Modulation of intestinal barrier by intestinal microbiota: Pathological and therapeutic implications. *Pharmacol. Res.* 69, 42–51.
- Ng, S.C. et al. (2017). Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. *Lancet* 390, 2769–2778.
- Pascal, V., Pozuelo, M., Borruel, N., Casellas, F., Campos, D., Santiago, A., Martinez, X., Varela, E., Sarrabayrouse, G., Machiels, K., et al. (2017). A microbial signature for Crohn's disease. *Gut* 66, 813–822.
- Van de Peer, Y. (1996). A quantitative map of nucleotide substitution rates in bacterial rRNA. *Nucleic Acids Res.* 24, 3381–3391.
- Peeters, M., Nevens, H., Baert, F., Hiele, M., Vlietinck, R., Rutgeerts, P., and Collection,

- D. (1996). Adjusted Risk and Concordance in Clinical Characteristics. *Gastroenterology* 597–603.
- Peyrin-Biroulet, L., and Lémann, M. (2011). Review article: Remission rates achievable by current therapies for inflammatory bowel disease. *Aliment. Pharmacol. Ther.* 33, 870–879.
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, S., Manichanh, C., Nielsen, T., Pons, N., Yamada, T., Mende, D.R., et al. (2010). A human gut microbial gene catalog established by metagenomic sequencing. *Nature* 464, 59–65.
- Qin, J., Li, Y., Cai, Z., Li, S., Zhu, J., Zhang, F., Liang, S., Zhang, W., Guan, Y., Shen, D., et al. (2012). A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* 490, 55–60.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F.O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* 41, 590–596.
- Quesada, E., Llamas, I., and Jose, M. (2018). *Halomonas ventosae* sp. nov., a moderately halophilic, denitrifying, exopolysaccharide-producing bacterium. *Arch. Microbiol.* 190, 733–737.
- 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*.
- Ridaura, V.K., Faith, J.J., Rey, F.E., Cheng, J., Duncan, A.E., Kau, A.L., Griffin, N.W., Lombard, V., Henrissat, B., Bain, J.R., et al. (2013). Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* (80-).
- Ridlon, J.M., Kang, D., and Hylemon, P.B. (2006). Bile salt biotransformations by human intestinal bacteria. *Microbiol. Rev.* 70, 241–259.
- Ringel, Y., Maharshak, N., Ringel-kulka, T., Wolber, E.A., Sartor, R.B., and Carroll, I.M. (2015). High throughput sequencing reveals distinct microbial populations within the mucosal and luminal niches in healthy individuals. *Gut Microbes* 173–181.
- Rogler, G., Vavricka, S., Schoepfer, A., and Lakatos, P.L. (2013). Mucosal healing and deep remission: What does it mean? *World J. Gastroenterol.* 19, 7552–7560.
- Rolhion, N., and Darfeuille-michaud, A. (2018). Adherent-Invasive *Escherichia coli* in Inflammatory Bowel. *Gastroenterology* 13, 1277–1283.
- Roth, M.P., Petersen, G.M., McElree, C., Feldman, E., and Rotter, J.I. (1989). Geographic origins of Jewish patients with inflammatory bowel disease. *Gastroenterology* 97, 900–904.
- Ruiz-ruiz, C., Srivastava, G.K., Carranza, D., Mata, J.A., Llamas, I., Santamaría, M., Quesada, E., and Molina, I.J. (2011). An exopolysaccharide produced by the novel halophilic bacterium *Halomonas stenophila* strain B100 selectively induces apoptosis in human T leukaemia cells. *Microbiol. Res.* 166, 345–355.

- Russell, D.W. (2003). THE ENZYMES, REGULATION, OF BILE ACID SYNTHESIS. 137–174.
- Rutgeerts, P., Geboes, K., and Vantrappen, G. (1990). Predictability of the Postoperative Course of Crohn's Disease. *Gastroenterology* 99, 956–963.
- Rutgeerts, P., Sandborn, W.J., Feagan, B.G., Reinisch, W., Olson, A., Johanns, J., Travers, S., Rachmilewitz, D., Hanauer, S.B., Lichtenstein, G.R., et al. (2005). Infliximab for Induction and Maintenance Therapy for Ulcerative Colitis. *N. Engl. J. Med.* 353.
- Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., et al. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541.
- Schultz, M., Veltkamp, C., Dieleman, L.A., Grenther, W.B., Wyrick, P.B., Tonkonogy, S.L., and Balfour Sartor, R. (2002). *Lactobacillus plantarum* 299V in the treatment and prevention of spontaneous colitis in interleukin-10-deficient mice. *Inflamm. Bowel Dis.* 8, 71–80.
- Segata, N., Haake, S.K., Mannon, P., Lemon, K.P., Waldron, L., Gevers, D., Huttenhower, C., and Izard, J. (2012). Composition of the adult digestive tract bacterial microbiome based on seven mouth surfaces, tonsils, throat and stool samples. *Genome Biol.* 13, R42.
- Sellon, R.K., Tonkonogy, S., Schultz, M., Dieleman, L.A., Grenther, W., Balish, E., Rennick, D.M., and Sartor, R.B. (1998). Resident enteric bacteria are necessary for development of spontaneous colitis and immune system activation in interleukin-10-deficient mice. *Infect. Immun.* 66, 5224–5231.
- Shanahan, F., and Bernstein, C.N. (2009). The evolving epidemiology of inflammatory bowel disease. *Curr. Opin. Gastroenterol.* 25, 301–305.
- Shivashankar, R. et al. (2017). Incidence and Prevalence of Crohn's Disease and Ulcerative Colitis in Olmsted County, Minnesota From 1970 Through 2010. *Clin. Gastroenterol. Hepatol.* 15.
- Silverberg, M.S., Satsangi, J., Ahmad, T., Arnott, I.D., Bernstein, C.N., Brant, S.R., Caprilli, R., Colombel, J.-F., Gasche, C., Geboes, K., et al. (2005). Toward an Integrated Clinical, Molecular and Serological Classification of Inflammatory Bowel Disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can. J. Gastroenterol.* 19, 5A-36A.
- Smith, J.L., Bayles, D.O., Smith, J.L., and Bayles, D.O. (2018). Critical Reviews in Microbiology The Contribution of Cytolethal Distending Toxin to Bacterial Pathogenesis The Contribution of Cytolethal Distending Toxin to Bacterial Pathogenesis. *World J Gastroenterol.* 7828.
- Snowden, J.A., Saccardi, R., Allez, M., Ardizzone, S., Arnold, R., Cervera, R., Denton, C., Hawkey, C., Labopin, M., Mancardi, G., et al. (2012). Haematopoietic SCT in severe

autoimmune diseases: Updated guidelines of the European group for blood and marrow transplantation. *Bone Marrow Transplant.* 47, 770–790.

Söderholm, J.D., Yang, P.C., Ceponis, P., Vohra, A., Riddell, R., Sherman, P.M., and Perdue, M.H. (2002). Chronic stress induces mast cell-dependent bacterial adherence and initiates mucosal inflammation in rat intestine. *Gastroenterology* 123, 1099–1108.

Sokol, H., and Seksik, P. (2010). The intestinal microbiota in inflammatory bowel diseases: Time to connect with the host. *Curr. Opin. Gastroenterol.* 26, 327–331.

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. *Proc. Natl. Acad. Sci.* 105, 16731–16736.

Sonnenberg, a, McCarty, D.J., and Jacobsen, S.J. (1991). Geographic variation of inflammatory bowel disease within the United States. *Gastroenterology* 100, 143–149.

Steck, N., Hoffmann, M., Sava, I.G., Kim, S.C., Hahne, H., Tonkonogy, S.L., Mair, K., Krueger, D., Pruteanu, M., Shanahan, F., et al. (2011). *Enterococcus faecalis* metalloprotease compromises epithelial barrier and contributes to intestinal inflammation. *Gastroenterology* 141, 959–971.

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. 17, 1971–1978.

Sydora BC, Tavernini MM, Wessler A, Jewell LD, F.R. (2003). Lack of interleukin-10 leads to intestinal inflammation, independent of the time at which luminal microbial colonization occurs. *Inflamm. Bowel Dis.* 9, 87–97.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin, E. V (2000). The COG database : a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 28, 33–36.

Trauner, M., and Boyer, J.L. (2018). Bile Salt Transporters : Molecular Characterization , Function , and Regulation TRANSPORT. 633–671.

Triantafillidis, J.K., Merikas, E., and Georgopoulos, F. (2011). Current and emerging drugs for the treatment of inflammatory bowel disease. *Drug Des. Devel. Ther.* 5, 185–210.

Turnbaugh, P.J., and Gordon, J.I. (2009). The core gut microbiome, energy balance and obesity. *J. Physiol.* 587, 4153–4158.

Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., and Gordon, J.I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. 444, 1027–1031.

- Turnbaugh, P.J., Hamady, M., Yatsunencko, T., Cantarel, B.L., Ley, R.E., Sogin, M.L., Jones, W.J., Roe, B. a, Jason, P., Egholm, M., et al. (2009). A core gut microbiom in obese and lean twins. *Nature* 457, 480–484.
- Ubeda, C. (2013). Intestinal Microbiota Containing *Barnesiella* Species Cures Vancomycin-Resistant *Enterococcus faecium* Colonization. *Infect. Immun.* 81, 965–973.
- Udayappan, S., Manneras-holm, L., Chaplin-scott, A., Belzer, C., Herrema, H., Dallinga-thie, G.M., Duncan, S.H., Stroes, E.S.G., Groen, A.K., Flint, H.J., et al. (2016). Oral treatment with *Eubacterium hallii* improves insulin sensitivity in db / db mice. *Nat. Publ. Gr.*
- Vila, A.V., Imhann, F., Collij, V., Jankipersadsing, S.A., Gurry, T., Mujagic, Z., Kurilshikov, A., Bonder, M.J., Jiang, X., Tigchelaar, E.F., et al. (2019). Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. *Sci. Transl. Med.* 10.
- Wang, B., Yao, M., Lv, L., Ling, Z., and Li, L. (2017). The Human Microbiota in Health and Disease. *Engineering* 3, 71–82.
- Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., and Al, W.E.T. (2007). Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy □ †. *Appl. Environ. Microbiol.* 73, 5261–5267.
- Wang, W., Chen, L., Zhou, R., Wang, X., Song, L., Huang, S., Wang, G., and Xia, B. (2014). Increased Proportions of *Bifidobacterium* and the *Lactobacillus* Group and Loss of Butyrate-Producing Bacteria in Inflammatory Bowel. *Jcm* 52, 398–406.
- Wang, X., Yu, R., Luo, X., Zhou, M., and Lin, X. (2008). Toxicon Toxin-screening and identification of bacteria isolated from highly toxic marine gastropod *Nassarius semiplicatus*. 52, 55–61.
- Willing, B.P., Dicksved, J., Halfverson, J., ANDERSSON, ANDERS F. LUCIO, M., ZHENG, Z., JÄRNEROT, G., TYSK, C., JANSSON, J.K., and ENGSTRAND, L. (2011). A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes: Commentary. *Inflamm. Bowel Dis. Monit.* 11, 166.
- Wills, E.S., Jonkers, D.M.A.E., Savelkoul, P.H., Masclee, A.A., Pierik, M.J., and Penders, J. (2014). Fecal Microbial Composition of Ulcerative Colitis and Crohn ' s Disease Patients in Remission and Subsequent Exacerbation. *PLoS One* 9, 1–10.
- Wong, W.C., Hentges, D.J., and Dougherty, S.H. (1996). Adequacy of the human faecal microbiota associated mouse as a model for studying the ecology of the human intestinal tract. *Microb. Ecol. Health Dis.* 9, 187–198.
- Wos-Oxley, M.L., Bleich, A., Oxley, A.P.A., Kahl, S., Janus, L.M., Smoczek, A., Nahrstedt, H., Pils, M.C., Taudien, S., Platzer, M., et al. (2012). Comparative evaluation of establishing a human gut microbial community within rodent models. *Gut Microbes* 3, 234–249.

- Wu, G.D., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y., Sue, a, Bewtra, M., Knights, D., Walters, W. a, Knight, R., et al. (2012). NIH Public Access. *Science* (80-). 334, 105–108.
- Xu, Z., Malmer, D., Langille, M.G.I., Way, S.F., and Knight, R. (2014). Which is more important for classifying microbial communities: Who's there or what they can do? *ISME J.* 8, 2357–2359.
- Yan, S., Huang, J., Chen, Z., Jiang, Z., Li, X., and Chen, Z. (2016). Metabolomics in gut microbiota: applications and challenges. *Sci. Bull.* 61, 1151–1153.
- Yang, S.K., Hong, M., Zhao, W., Jung, Y., Baek, J., Tayebi, N., Kim, K.M., Ye, B.D., Kim, K.J., Park, S.H., et al. (2014). Genome-wide association study of Crohn's disease in Koreans revealed three new susceptibility loci and common attributes of genetic susceptibility across ethnic populations. *Gut* 63, 80–87.
- Yatsunenکو, T. (2011). The Gut Microbiome In Healthy And Severely Malnourished Humans.
- Yatsunenکو, 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*.
- Yilmaz, B., Juillerat, P., Øyås, O., Ramon, C., Bravo, F.D., Franc, Y., Fournier, N., Michetti, P., Mueller, C., Geuking, M., et al. (2019). Microbial network disturbances in relapsing refractory Crohn's disease. *Nat. Med.* 25.
- Young, V.B., Raffals, L.H., Huse, S.M., Vital, M., Dai, D., Schloss, P.D., Brulc, J.M., Antonopoulos, D.A., Arrieta, R.L., Kwon, J.H., et al. (2013). Multiphasic analysis of the temporal development of the distal gut microbiota in patients following ileal pouch anal anastomosis. *Microbiome* 1.
- Zhang, Y., Yu, X., Yu, E., Wang, N., Cai, Q., Shuai, Q., Yan, F., Jiang, L., Wang, H., Liu, J., et al. (2018). Changes in gut microbiota and plasma inflammatory factors across the stages of colorectal tumorigenesis : a case-control study. 1–10.
- Zhou, Y., Chen, H., He, H., Du, Y., Hu, J., Wang, H., Li, Y., Chen, Y., and Nie, Y. (2016). Increased *Enterococcus faecalis* infection is associated with clinically active Crohn disease. *Medicine (Baltimore)*. 1578–1585.

Acknowledgments

First and foremost, I would like to thank Prof. Dirk Haller, for introducing me to the fascinating world of gut microbiome, for his support and guidance during my PhD, and for giving me the trust and freedom to explore research topics upon my own interest. I sincerely thank you for guiding me to excel on all fronts.

A special thanks to all members of the examination committee, Prof. Scherer and Prof. Stecher for their support and critical reviewing of my thesis.

This work would not be possible without the support and encouragement of many great individuals. In particular, I would like to thank Dr. Nadine Waldschmitt and Dr. Ilias Lagkouvardos for their support and constructive feedback. Nadine, thank you for the productive discussions and open exchange of ideas. I very much enjoyed working with you. Ilias, our discussions were always lively. Your critical feedback was always useful in pushing me thinking different ways.

I would like to thank the awesome people of Haller's lab. In particular, I would like to thank Dr. Thomas Clavel for his mentorship and support in my early steps in the lab. Thanks are due to Dr. Sigrid Kisling for her assistance with the histological assessment and animal approvals, Silvia, Alex, Sandra H. for their help in taking care of gnotobiotic animals, Caro for her constant help in getting even the trickiest sample working, Nico for lending a helping hand whenever it was needed and Brita making our daily life in the lab a lot easier.

I have been so lucky to work with such intelligent and warm-hearted people; I would like to thank all the former and current members of Haller's lab for making the journey memorable and enjoyable. I would like to thank all my PhD colleagues; with whom I have shared moments of deep anxiety but also of big excitement. In particular, I would like to thank, Isabella, Jelena, Valentina, Sevana, Sandra R., Elena, Sarah, Hong, Adam and Mohammed for the discussions and exchanges during our PhD time. I would like to thank the post-doc colleagues, Eva, Silke, Olivia and Katharina for being great colleagues and inspiring women in science.

In addition, I have been very privileged to get to know and to collaborate with many other great people, who became friends over the last several years. I learned a lot from you about life, research, how to tackle research questions. Special thanks to Prof. Julian Panes, Dr. Azucena Salas and the team in Barcelona for collaborating on the HSCT human cohort and for the valuable discussion along the way. Thanks are due to Prof. Matthiue Allez, Dr. Lionel L-Bourhis and the team in Paris for their collaboration on the post-surgical and biotherapy human cohorts.

This work was supported by the Helmsley Charitable Trust (IBD overtime Consortium). My dearest thanks to the people in this unique consortium: Garabet Yertessian, Shefali Sonni and Susan Painter.

I would like to thank my friends, who were always there for me. Douaa, Aya, Azza, Marwa T., Loubna, Nora, Nour, thanks for being the best friends one can ask for. Thanks for being my crying shoulder when the going got tough, and for celebrating me when the hard work paid off.

Finally, I would like to thank my parents for raising me up with a love for science and supporting me in all my pursuits. Thanks to my sister and my brothers for your unconditional love and care and for being always my biggest fans. My in-laws, thanks for being the loving, caring and amazing people you are.

This last word of acknowledgment I saved for my dear husband Ahmed and my beautiful daughter Laila, who shared this entire amazing journey with me. Ahmed, thanks for your unshakable love and support. You were there helped me to keep going, and to keep things in perspective. Laila, having you was the best thing that happened to me. Thanks for the patience you showed during the process. You two made me stronger and better than I could have ever imagined. So, it only feels right that I dedicate this dissertation to you two.

Publications and Presentations

Published Manuscripts

Metwaly A and Haller D. Microbial Signals Link Westernized Diet to Metabolic Inflammation: More Evidence to Resolve Controversies. *Cellular and Molecular Gastroenterology and Hepatology* (2019)

Metwaly A and Haller D. Strain-level diversity in the gut– the *P. copri* case. *Cell Host& Microbe* (2019)

Metwaly A and Haller D. Multi-omics in IBD biomarker discovery: the missing links. *Nature Gastroenterology& Hepatology* (2019).

Waldschmitt N, **Metwaly A**, Fischer S, Haller D. Microbial Signatures as a Predictive Tool in IBD- Pearls and Pitfalls. *Inflamm Bowel Dis.* (2018).

Submitted Manuscripts

Metwaly A, Dunkel A, Waldschmitt N, Buttó LF, Lagkouvardos I, Corraliza AM, Mayorgas A, Le Bourhis L, Schmidt S, Durai Raj AC, Schloter M, Hofmann T, Allez M, Panes J, Salas A and Haller D. Integrated Metabolomic and Microbiome approach revealed functional signatures associated with disease Severity in inflammatory bowel disease

Metwaly A*, Calasan J*, Waldschmitt N*, Khaloian S, Ahmed M, Basic M, Bleich A, Cominelli F, Haller D. Segmented filamentous bacteria are enriched in Crohn's disease and cause ileo-colonic inflammation in TNFdARE mice (* Authors contributed equally to this work).

Khaloian S, Rath E, Gleisinger E, Berger E, **Metwaly A**, Waldschmitt N, Hammoudi N, Allez M, Haller D. Mitochondrial impairment in Crohn's disease-like inflammation drives intestinal stem cell transition towards dysfunctional paneth cells

Brückner A, Frivolt K, Werkstetter K., Krohn K, Hajji M, Otte S, Bufler P, Sibylle Koletzko S, Liptay S, Bechtold-Dalla Pozza S, Ahmed M, **Metwaly A**, Shokry E, Jair GM, Uhl O, Koletzko B, Haller D, Schwerd T. Partial enteral nutrition does not influence bone geometry but improves growth in non-inflamed paediatric Crohn's disease patients

Published Abstracts

DDW 2019: ***Integrated Metabolomic and Microbiome approach revealed functional signatures associated with disease Severity in inflammatory bowel disease***

Metwaly A, Waldschmitt N, Buttó L, Lagkouvardos I, Corraliza A, Mayorgas A, Le Bourhis L, Schmidt S, Dunkel A, Durai Raj AC, Schloter M, Hofmann T, Allez M, Panes J, Salas A and Haller D. *Gastroenterology*, Volume 156, Issue 6, S-49

DDW 2019: ***Segmented filamentous bacteria are enriched in Crohn's disease and cause ileo-colonic inflammation in TNFdARE mice***

Calasan J*, Waldschmitt N*, **Metwaly A***, Khaloian S, Ahmed M, Basic M, Bleich A, Cominelli F, Haller D (* Authors contributed equally to this work). *Gastroenterology*, Volume 156, Issue 6, S-49

DDW 2019: *Mitochondrial impairment in Crohn's disease-like inflammation drives intestinal stem cell transition towards dysfunctional Paneth cells*

Khaloian S, Rath E, Gleisinger E, Berger E, Metwaly A, Waldschmitt N, Hammoudi N, Allez M, Haller D. Gastroenterology, Volume 156, Issue 6, S-49

DDW 2017: *Identification of Disease-Relevant Bacterial Signatures in Gnotobiotic IL-10 Deficient Mice using Fecal Samples from IBD Patients Undergoing Hematopoietic Stem Cell Transplantation.*

Metwaly A., Butto LF, Waldschmitt N, Lagkouravdos I, Corraliza AM, Mayorgas A, Martinez-Medina M, Allez M, Panes J, Salas A, Haller D. Gastroenterology, Volume 152, Issue 5, S989

ECCO 2017: *Identification of Bacterial Signatures in Gnotobiotic IL-10 Deficient Mice Using Fecal Samples from IBD Patients Undergoing Hematopoietic Stem Cell Transplantation*

Metwaly A, Butto LF, Waldschmitt N, Lagkouravdos I, Corraliza AM, Mayorgas A, Martinez-Medina M, Allez M, Panes J, Salas A, Haller D. Journal of Crohn's and Colitis 11(suppl. 1): S18 S18

Oral Presentations

2019

Digestive Disease Week 2019 (San Diego)

Integrated Metabolomic and Microbiome approach revealed functional signatures associated with disease Severity in inflammatory bowel disease.

Metwaly A, Waldschmitt N, Buttó LF, Lagkouravdos I, Corraliza AM, Mayorgas A, Le Bourhis L, Schmidt S, Dunkel A, Durai Raj AC, Schloter M, Hofmann T, Allez M, Panes J, Salas A and Haller D.

Keystone Symposium: Host and the Environment in IBD 2019 (Taos, New Mexico)

Integrative multi-omics Identification of Biomarker Microbial Signatures in Inflammatory Bowel Disease.

Metwaly A, Waldschmitt N, Buttó LF, Lagkouravdos I, Corraliza AM, Mayorgas A, Le Bourhis L, Schmidt S, Dunkel A, Durai Raj AC, Schloter M, Hofmann T, Allez M, Panes J, Salas A and Haller D.

71st Annual Conference of the German Society for Hygiene and Microbiology (DGHM) (Göttingen, Germany)

Integrative multi-omics Identification of Biomarker Microbial Signatures in Inflammatory Bowel Disease.

Metwaly A, Waldschmitt N, Buttó LF, Lagkouravdos I, Corraliza AM, Mayorgas A, Le Bourhis L, Schmidt S, Dunkel A, Durai Raj AC, Schloter M, Hofmann T, Allez M, Panes J, Salas A and Haller D.

12th Seeon Conference, Microbiota, Probiota and Host (Seeon, Germany)

Integrated Metabolomic and Microbiome approach revealed functional signatures associated with disease Severity in inflammatory bowel disease.

Metwaly A, Dunkel A, Waldschmitt N, Buttó LF, Lagkouravdos I, Corraliza AM, Mayorgas A, Le Bourhis L, Schmidt S, Durai Raj AC, Schloter M, Hofmann T, Allez M, Panes J, Salas A and Haller D.

2018

11th Seeon Conference, Microbiota, Probiota and Host (Seeon, Germany)

Functional Characterization of Microbial Signatures in Inflammatory Bowel disease using Gnotobiotic Humanized Mice.

Metwaly A, Waldschmitt N, Buttó L, Lagkouvardos I, Corraliza A, Mayorgas A, Le Bourhis L, Schmidt S, Dunkel A, Durai Raj AC, Schloter M, Hofmann T, Allez M, Panes J, Salas A and Haller D.

68. Annual meeting for Oral Surgery 2018 (Bad Homburg, Germany)

Oral microbiota and systemic diseases.

2017

12th Congress of European Crohn's and Colitis Organization (ECCO) (Barcelona, Spain)

Identification of disease-relevant bacterial signatures in gnotobiotic IL-10 deficient mice using fecal samples from IBD patients undergoing hematopoietic stem cell transplantation

Metwaly A, Buttó LF, Waldschmitt N, Lagkouvardos I, Corraliza AM, Mayorgas A, Martinez-Medina M, Allez M, Panes J, Salas A, Haller

5th Joint Conference of DGHM and VAAM (Würzburg, Germany)

Identification of disease-relevant bacterial signatures in gnotobiotic IL-10 deficient mice using fecal samples from IBD patients undergoing hematopoietic stem cell transplantation

Metwaly A, Buttó LF, Waldschmitt N, Lagkouvardos I, Corraliza AM, Mayorgas A, Martinez-Medina M, Allez M, Panes J, Salas A, Haller D

10th Seeon Conference, Microbiota, Probiota and Host (Seeon, Germany)

Identification of disease-relevant bacterial signatures in gnotobiotic IL-10 deficient mice using fecal samples from IBD patients undergoing hematopoietic stem cell transplantation

Metwaly A, Buttó LF, Waldschmitt N, Lagkouvardos I, Corraliza AM, Mayorgas A, Martinez-Medina M, Allez M, Panes J, Salas A, Haller D

Poster Presentations

2019

12th Seeon Conference, Microbiota, Probiota and Host (Seeon, Germany)

Segmented filamentous bacteria are enriched in Crohn's disease and cause ileo-colonic inflammation in TNFdARE mice

Metwaly A*, Calasan J*, Waldschmitt N, Ahmed M, *, Khaloian S, Basic M, Bleich A, Cominelli F, Haller D. (* Authors contributed equally to this work).

2018

International Human Microbiome Consortium meeting (IHMC) (Killanery, Ireland)

Functional Characterization of Microbial Signatures in Inflammatory Bowel Disease (IBD) using Gnotobiotic Humanized Mice Models.

Metwaly A, Waldschmitt N, Buttó LF, Lagkouvardos I, Corraliza AM, Mayorgas A, Le Bourhis L, Schmidt S, Dunkel A, Allez M, Panes J, Salas A and Haller D.

2017

Digestive Disease Week 2017 (Chicago)

Identification of disease-relevant bacterial signatures in gnotobiotic IL-10 deficient mice using fecal samples from IBD patients undergoing hematopoietic stem cell transplantation

Metwaly A, Buttó LF, Waldschmitt N, Lagkouvardos I, Corraliza AM, Mayorgas A, Martinez-Medina M, Allez M, Panes J, Salas A, Haller D

Keystone Symposia on Microbiome in health and disease (Keystone, Colorado)

Validation of microbiota humanization in mouse models of chronic inflammation and metabolic disorders. Metwaly A, Schüppel VL, Butto LF, Lagkouvardos I, Nieuwdorp M and Haller D.

EMBO course on humanized mice (Heidelberg, Germany)

Functional Characterization of Microbial Signatures in Inflammatory Bowel Disease (IBD) using Gnotobiotic Humanized Mice Models.

Metwaly A, Waldschmitt N, Buttó LF, Lagkouvardos I, Corraliza AM, Mayorgas A, Le Bourhis L, Chardiny V, Allez M, Panes J, Salas A and Haller D

2016

9th Seeon Conference, Microbiota, Probiota and Host 2016 (Seeon, Germany)

Identification of disease-relevant bacterial signatures in gnotobiotic IL-10 deficient mice using fecal samples from IBD patients undergoing hematopoietic stem cell transplantation

Metwaly A, Buttó LF, Waldschmitt N, Lagkouvardos I, Corraliza AM, Mayorgas A, Martinez-Medina M, Allez M, Panes J, Salas A, Haller D

Spring School of Immunology (Ettal, Germany)

Humanized Mouse model of inflammatory bowel disease: functional characterization of dysbiotic gut microbiota.

Metwaly A., Butto LF, Lagkouvardos I, Corraliza AM, Panes J, Salas A, Haller D

2015

7th Seeon Conference, Microbiota, Probiota and Host 2015, Germany

Intestinal microbiota composition influences recurrence of ileitis in a mouse model of Crohn's disease. Schauback M, Metwaly A, Clavel T, Waldschmitt N, Wehkamp J, Martinez I, Walter J, Roulis M, Kollias G and Haller

List of cited publications

1. **Metwaly A**, Dunkel A, Waldschmitt N, Buttó LF, Lagkouvardos I, Corraliza AM, Mayorgas A, Le Bourhis L, Schmidt S, Durai Raj AC, Schlöter M, Hofmann T, Allez M, Panes J, Salas A and Haller D. Integrated Metabolomic and Microbiome approach revealed functional signatures associated with disease Severity in inflammatory bowel disease
2. Waldschmitt N, **Metwaly A**, Fischer S, Haller D. Microbial Signatures as a Predictive Tool in IBD-Pearls and Pitfalls. *Inflamm Bowel Dis.* 2018 May 18;24(6):1123-1132.

Hiermit erkläre ich an Eides statt, dass ich alleiniger, federführender Hauptautor der zwei oben genannten Publikationen und Studien bin, die in dieser Arbeit wörtlich zitiert wurden. Die betreffenden Passagen wurden ausschließlich von mir verfasst.

Ort, den

.....

Unterschrift

Curriculum Vitae

MSc. Amira Metwaly

Akademische Ausbildung

- 07/2013–06/2018 **Doktorarbeit, Lehrstuhl für Ernährung und Immunologie von Prof. Dirk Haller, (TU München)**
 “Titel: “Functional Characterizations of Human Gut Microbiota in Inflammatory Bowel Disease Patients Using Gnotobiotic Humanized Mice”
- 10/2008 –12/2010 **Masterstudium, Molekulare Biowissenschaften (Ruprecht-Karls-Universität Heidelberg)**
 Titel “Role of CD95 in Glioblastoma multiform” (**GPA 1.3**)
 Department of Molecular Neurobiology, Deutsches Krebsforschungszentrum (DKFZ)
- 06/2009 -12/2009 **Bioinformatik, A* STAR Bioinformatics Institut, Singapore und National University of Singapore (NUS)**
- 09/2007 – 01/2008 **Bachelorstudium, Pharmazie und Biotechnologie, Deutsche Universität in Kairo (GUC), (Ausgezeichnet, A+).**
- 10/2000 – 07/2003 **Gymnasium Abschluss (GPA: 1,0)**
 Top of class (1st / ca. 200,000 Studenten)

Berufserfahrung

- 06/2018 – Heute **Wissenschaftliche Angestellte**
 Lehrstuhl für Ernährung und Immunologie von Prof. Dirk Haller (TU München)
- 03/2012-06/2018 **Doktorandin am Lehrstuhl für Ernährung und Immunologie von Prof. Dirk Haller, (TU München)**
- 08/2009 – 12/2009 **Intern and research associate**
 A* STAR Bioinformatics Institute (BII), Singapore
- 01/2011 – 04/2011 **Research assistant at the department of basic research**
 Children's Cancer Hospital 57357, Kairo, Ägypten

Auszeichnungen

- 2003** Top of class (200,000 candidates) – Gymnasium Abschluss
- 2003-2008** Full academic scholarship – German University in Cairo (GUC)
- 2006** GUC Travel Award –International Pharmaceutical Federation Congress
- 2007** DAAD Travel grant – Drug design summer school, University of Tübingen
- 2008 – 2010** Citadel Capital academic scholarship – University of Heidelberg
- 2009-2010** Singapore International Pre-Graduate Award (SIPGA)
- 2019** Poster Presentation Award (71st Annual Conference of the German Society for Hygiene and Microbiology (DGHM))