Analysis of the intestinal microbiome from fecal samples in health and IBD using NGS

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

M.Sc. Isabell Flade

aus Sindelfingen

Tübingen

2015

Table of Contents

Abbreviation

°C	Degrees Celsius
µg	Microgram
µl	Microliter
µm	Micrometer
µM	Micromolar
2D	Two dimensional
3D	Three dimensional
A	Ampere
B. vulgatus mpk	*Bacteroides vulgatus* (Max von Pettenkofer)
B.d.	*Bacteroides dorei*
B.v.	*Bacteroides vulgatus*
Bac	*Bacteroides*
BHI	Brain Heart Infusion
BMI	Body Mass Index
bp	Base pairs
BP	Bandpass
CD	Crohn's Disease
cDNA	Complementary DNA
CFU	Colony Forming Units
c_t	Threshold cycle
Cy3	Cyanine 3
Cy5	Cyanine 5
DAPI	4',6-diamidino-2-phenylindole
DC	Dendritic Cell
DNA	Deoxynucleic acid
DSS	Dextran sulfate sodium
E. coli$_{MUT}$	*E. coli* JM83 ΔhtrBhtrBpg
E. coli$_{WT}$	Wild type *Escherichia coli*
EAE	Experimental autoimmune encephalomyelitis
EDTA	Ethylenediaminetetraacetic acid
Ent	Enterobacteria
ETH	Animal Facility "Einrichtung für Tierhygiene"
Eub	Eubacteria
FACS	Ethylenediaminetetraacetic acid
FCS	Fetal Calf Serum
Fc	Fragment crystallizable
FISH	Florescence in situ Hybridization
g	Gram
G	Gauge
g	Gravity
GALT	Gut associated Lymphoid Tissue
gDNA	Genomic DNA
GIT	Gastrointestinal Tract
h	Hours
H_2O	Water
HCl	Hydrochloric acid
HGP	Human Genome Project
HNO	Animal Facility "Hals Nasen Ohren Klinik"
i.p.	Intraperitoneal
IBD	Inflammatory Bowel Disease

IgA	Immunoglobulin A
IVC	Individually ventilated Cage
LB	Luria-Bertani
LP	Longpass
LPS	Lipopolysaccharide
LPS_{MUT}	Mutant LPS
LPS_{WT}	Wildtype LPS
M	Molar
Mb	Mega bases
mg	Milligram
min	Minutes
ml	Milliliter
mM	Millimolar
mm	Millimeter
MO	Microorganisms
mRNA	Messenger RNA
MS	Multiple sclerosis
Na_2HPO_4	Disodium phosphate
NaCl	Sodium chloride
NaH_2PO_4	Monosodium phosphate
NaOH	Sodium hydroxide
ng	Nano gram
NGS	Next-generation Sequencing
nm	Nanometer
nr/nt	Nucleotide collection Database
OTU	Orthologous taxonomic Units
PBS	Phosphate-buffered saline
PCA	Principal Component Analysis
PCoA	Principal Coordinates Analysis
PCR	Polymerase Chain Reaction
PFA	Paraformaldehyde
PVPP	Polyvinylpyrrolidon
Qiime	Quantitative Insights Into Microbial Ecology
qPCR	Quantitative Real-time PCR
$Rag1^{-/-}$	Recombination activating gene 1
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
S	Svedberg
SCFA	Short Chain Fatty Acid
SCID	Severe Combined Immunodeficiency
sec	Seconds
SEM	Standard Error of the Mean
SIgA	Secretory immunoglobulin A
SPF	Specific pathogen free
T1D	Type 1 Diabetes
TENP	TRIS, EDTA, NaCl and PVPP containing Buffer
T_H1	T Helper cell type 1
T_H17	T Helper cell type 17
T_H2	T Helper cell type 2
TLR4	Toll like Receptor 4
T_{reg}	Regulatory T-cell
tRNA	Transfer RNA

U	Unit
UC	Ulcerative colitis
V	Volt
v	Hypervariable Region
x	Fold

Abstract

Microorganisms colonizing the body of mammals are not only responsible for infectious diseases, they are also a very important factor for maintaining health. Commensal bacteria have a direct influence on health, as they compete with pathogens and can prevent them from colonizing. Additionally microorganisms have a more indirect influence: The gut microbiota is involved in development and shaping of the immune system, therefore dysbiosis may result in disease development. Within the last decade, many different diseases have been associated with a shift in the gut microbiota. Most studies however do not answer the question whether changes of the intestinal microbiome are cause or result of disease. As a mouse model, we used Rag1$^{-/-}$ mice, lacking B- and T-cells. An immune reaction, and therefor colitis development, is only possible after T-cell transfer. Depending on the composition of the gut microbiome at the time of T-cell transfer, the mice subsequently remain healthy or develop colitis. Changing the microbiome of the mice prior to reconstitution of the immune system influences disease development.

Analysis of the intestinal microbiome can be done using different methods, such as PCR, FISH or by culture. However, these methods are limited to microorganisms, of which the genome sequence is known or that can be grown in the lab. These limitations can be overcome by using Next-Generation-Sequencing (NGS) to identify the members of the gut microbiome. Therefore, here we developed protocols for analyzing the intestinal microbiome from fecal samples.

Our results show, that environmental factors influence the composition of the intestinal microbiome severely. Differences between single experiments were often greater than the differences between different treatments making it impossible to link a single or combination of bacterial species to induction of colitis or maintenance of health. For identifying microbiota compositions that are causative for health or disease, these environmental influences have to be eliminated or at least reduced. Otherwise the strong variations between animals will cover up causative alterations.

Zusammenfassung

Mikroorganismen, die den Körper von Säugetieren besiedeln, sind nicht nur für Infektionen verantwortlich, sondern auch unerlässlich für deren Gesundheit. Kommensale Bakterien können unsere Gesundheit auf direktem Weg beeinflussen, z.B. vermitteln so eine Kolonisierungsresistenz indem sie mit Pathogenen kompetieren und somit diese verdrängen können. Zusätzlich beeinflussen Mikroorgansimen die Gesundheit des Wirts auf indirektem Wege: Darmbakterien sind an der Schulung und Entwicklung des Immunsystems beteiligt. Kommt es zur Dysbiose, kann das somit zur Entwicklung von Krankheiten führen. Im letzten Jahrzehnt wurden zahlreiche Erkrankungen mit einer veränderten Zusammensetzung des Darmmikrobioms assoziiert. Die meisten Forschungsarbeiten geben allerdings keinen Hinweis darauf, ob das veränderte Darmmikrobiom Ursache oder Wirkung einer Erkrankung ist. Um dieser Frage nachzugehen, haben wir als Modelorganismus Rag1$^{-/-}$ Mäuse verwendet. Diesen Tieren fehlen sowohl reife B- wie auch T-Zellen. Eine Kolitis kann sich nur dann entwickeln, wenn das Immunsystem durch einen T-Zell Transfer wiederhergestellt wurde. Somit kann die Zusammensetzung des Darmmikrobioms vor Krankheitsbeginn analysiert und verändert werden. Je nach Zusammensetzung des Darmmikrobioms zum Zeitpunkt des T-Zell Transfers, entwickelte sich eine Kolitis oder die Mäuse blieben gesund. Durch verändern der Ausgangszusammensetzung, konnten wir beeinflussen, ob eine Maus gesund bleibt oder erkrankt.

Die Zusammensetzung des Darmmikrobioms kann mit Hilfe von Methoden wie PCR, FISH oder durch Kultivierung bestimmt werden. Der Nachweis von Bakterien über diese Methoden ist allerdings auf die Stämme beschränkt, deren Genomsequenz bekannt ist, bzw. die im Labor angezüchtet werden können. Die Protokolle, die in dieser Arbeit etabliert wurden, umgehen diese Beschränkungen, indem das Darmmikrobiom mit Hilfe von Next-Generation-Sequencing von DNA aus Stuhlproben charakterisiert wird.

Unsere Ergebnisse legen nahe, dass Umwelteinflüsse das Darmmikrobiom stark beeinflussen. Somit sind Unterschiede zwischen den Tieren verschiedener Experimente meist größer als Verschiebungen, die sich aufgrund unterschiedlicher Behandlungen oder Krankheitsstadien ergeben. Es ist daher unmöglich einzelne oder eine Gruppe von Bakterien zu identifizieren, die dafür verantwortlich sind, dass eine Kolitis entsteht bzw. ein Tier gesund bleibt. Untersuchungen dieser Art, können nur dann zu einer Assoziation zwischen Darmmikrobiom und Gesundheitszustand führen, wenn möglichst alle Umwelteinflüsse identisch sind. Ist dies nicht der Fall, verschleiern die individuellen Unterschiede jene, die mit einer Erkrankung im Zusammenhang stehen.

1. Introduction

1.1 Microbiota and microbiome

Microorganisms are present in various environments and habitats to which they have adapted over time. Some are able to survive under extreme conditions which are lethal for other organisms. The microbial compositions are just as diverse as their natural habitats, ranging from terrestrial hot springs to various human body sites [1–7].

The community of microorganisms within a particular environment is referred to as microbiota. The organisms contributing to such a collective include not only bacteria, viruses and fungi but also archaea, phage and other microbial eukarya, however, bacteria are the most studied group [8]. In 2001, Lederberg defined the term "microbiome" to describe the entirety of microorganisms colonizing the human body including symbiotic as well as pathogenic microorganisms [9,10].

The human body is composed of approximately 10^{13} cells. It is estimated that about 10 times more bacteria, belonging to at least 500 different species, colonize the surfaces of a human and the majority of these are located in the gastrointestinal tract (GIT) [11–14]. Taken together, these bacteria contain approximately 100-150 times more genes than a human [15–18].

The amount and diversity of bacteria differ greatly between body sites. Whereas communities of the oral cavity and stool have been found to be very complex, vaginal communities are particularly simple [4,19].

1.2 Intestinal microbiota

The majority of the bacteria colonizing a human are located in the GIT. The microbiota of the different intestinal sections differs drastically in composition and amount of viable bacteria. Within the duodenum/jejunum the lowest counts (10^2 CFU/g content) have been detected (Figure 1). The number increases and reaches its maximum in the colon where 10^{12} CFU/g content have been recorded [19–21].

Duodenum:
Up to 10^3 MO / ml content

Jejunum:
~10^5 MO / ml content
Lactobacilli, Streptococci

Ileum:
10^8 - 10^9 MO / ml content
Enterobacteriaceae, *Bacteroides*
spp. , *Prevotella*,
Porphyromonas, Fusobacteria,
Bifidobacteria, Enterococci

Colon:
Up to 10^{12} MO / g content
Bacteroides, Prevotella, Porphyromonas spp.
Fusobacteria, Eubacteria, Bifidobacteria,
Streptococci, Peptostreptococcus,
E. coli, Enterobacteriazeae, Enterococci, *Veillonella*,
Lactobacilli, *Clostridium* spp., Staphylococci,
Pseudomonades

Figure 1: Microorganisms (MO) mainly colonizing the different sections of the human intestinal tract [22].

1.2.1 Development of an individual's intestinal microbiota

Depending on method of delivery, a newborn comes into contact with the vaginal as well as the intestinal microbiota of the mother and/or the microorganisms of the surrounding environment during childbirth. For many years this was considered to be the initial colonization and the beginning of a new microbial community [23–26]. It has been recently hypothesized, however, that commensal intestinal bacteria of the mother translocate into the placenta and that the first contact takes place prior to birth [16,27]. Post-partum newborns are also very quickly exposed to the microbiota of the maternal skin, milk and mouth [16].

Breast milk, as a continuous source of bacteria, is one of the most important factors for the colonization of the neonatal intestinal tract [28,29]. The relationship between this bacterial uptake and a child's diet is strongly supported by the fact that the composition of the bacterial community colonizing the intestine differs between breast- and formula-fed children [26].

The bacterial population changes drastically during the first years of life and constantly increases in diversity. Inter-individual variations are pronounced and significantly greater in young children than in adults. It is estimated that at approximately three years of age the intestinal microbial composition has reached a stable adult stage [16,30,31].

The first bacteria to colonize the gastrointestinal tract of infants are facultative aerobic bacteria. After they have consumed the oxygen present in the GIT at birth, the changed environment then allows strictly anaerobic bacteria to reside and grow in the gut [16,26]. The majority of, but not all, studies of the intestinal bacterial community of infants have detected dominating amounts of *Bifidobacteria*, which are considered "beneficial" for the health of the host [16,26,31,32]. The adult intestinal microbiota, conversely, is dominated by bacteria belonging to the phyla of *Bacteroidetes* and *Firmicutes* [13,14].

1.2.2   Intestinal microbiota compositions and variations

Microbiome studies indicate that the presence and abundance of bacterial species varies greatly between individuals; however, on a higher phylogenetic level it seems to be quite conserved [14]. Eckberg et al. discovered approximately 400 different phylotypes belonging to 7 phyla [13] within the human colon. The most abundant phyla are *Bacteroidetes* and *Firmicutes.* Together these two groups account for more than 90 % of the total bacteria, whereas others such as *Proteobacteria*, A*ctinobacteria*, *Fusobacteria* and *Verrucomicrobia* are relatively rare [13,14]. The strictly anaerobic intestinal environment may make it difficult for facultative anaerobic bacteria to successfully compete with *Bacteroidetes* and *Firmicutes,* resulting in them usually being present only in very low quantities [33]. Inter-individual variations can partially be attributed to genetic factors of the host. It is also believed that the order in which bacteria colonize the intestine is important for the establishment of the gut microbiome. Bacteria shape the physiochemical properties of the colonized niche, which may or may not be advantageous to microorganisms that follow, thereby affecting the community assembly [34].

The intestinal microbiota not only differs greatly between individuals, but also underlies smaller fluctuations within a single host [35]. Due to these robust normal variations, pathological modulations may not be detected. In 2011, Arumugam et al. determined that although humans harbor varying microbial compositions in their intestines, people can be divided into one of three groups with different main contributing bacterial species. All studied subjects could be assigned to one of these main compositions. This indicates that non-pathological variations are not continuous but that the microbiome is restrained to certain well-balanced communities for symbiosis with the host [36].

Correlating host properties with the microbial intestinal community by taxa is very difficult and in some cases, even impossible due to large variations. Although bacteria belong to different taxa, their genomes can code for the same or similar functional groups. Studying the genetic information of the microbiota on a functional level instead of only the taxonomic level, may decrease the perceived variations between and within individuals, allowing correlations to host properties [35,36].

1.3 Interactions of intestinal microbiota and host

The microbiota supports its host in many ways. Enhancing digestive efficiency by nutrient processing (germ-free animals need 30 % more calories [33]), promoting proper immune system development and limiting pathogen colonization (colonization resistance) are only a few examples for the benefits the microbiota provides to the host [14,37–39]. Mammalian genomes evolve rather slowly, and the faster evolving colonizing bacteria, can quickly take over functions that the host is unable to fulfill itself, for example, by supplying enzymes. Thus, by changing the microbial composition, otherwise impossible rapid adaptations to new circumstances are possible in mammals [33,40]. In return, the microorganisms are given the advantage of living in a protected and nutrient-rich environment [15,19,33,41]. The prerequisites for this symbiosis, however, are that the microbiota is accepted by the host, is limited to its

habitat and doesn't penetrate host tissue. Commensal bacteria that fulfill these conditions are recognized by the host but do not induce a strong inflammatory reaction, in fact, some are even able to induce tolerance [33,42,43].

Examples of factors suspected to influence the human microbiota include diet, antibiotic treatment, and exposure to environmental bacteria, [30,33,44–48]. Similar bacteria support each other and increase the possibility of colonization of the host by related phylotypes [14]. Influencing the microbiome by oral uptake of bacteria, e.g. by consumption of probiotics, may have beneficial effects to the host, if the correct bacterial subset is varied [49–51]. The opposite, is also possible: certain bacterial community members can be reduced, for example by antibiotic treatment, minimizing the competition for space and nutrients for unaffected bacterial strains enabling these fractions to expand [52]. In some cases, the weakened populations include beneficial bacteria that limit the ability of opportunistic pathogens to induce an immune reaction. The increase of pathogens and reduction of beneficial bacteria can result in manifestation of diseases, as in *Clostridium difficile* infection [52].

1.3.1 Interaction of the intestinal microbiota and the host immune system

The gut-associated lymphoid tissue (GALT) is important for the induction of an immune reaction after the identification of antigens. In addition to the organized lymphoid tissue of the GALT, the lamina propria of the gut wall harbors lymphocytes, which in addition to the architecture of the intestine also limit infection [53].

The enormous amount of bacteria colonizing the gastrointestinal tract represents a high risk of infection for the host. The first line of defense against microbial penetration is the mucosal barrier including epithelial cells, and a mucus layer separating these cells from the gut lumen. Two strata are part of the mucus layer [33,54]. The luminal layer contains bacteria whereas the inner layer is resistant to bacterial penetration in healthy individuals [33,55,56]. This physical barrier, however, is not the only protection of the epithelial layer. Enterocytes and Paneth cells, which are part of the mucosa, protect the host from intestinal microbiota by separation of the deeper host tissue from the outside and by secreting antimicrobial peptides [33,57]. These antimicrobial peptides are hindered from reaching the gut lumen by the mucosal layer [33,58], therefore selectively targeting bacteria threatening to penetrate host tissue. Further protection of the host is provided by the adaptive and innate immune systems, which constantly sample the gut microbiome and activate appropriate immune responses [33]. Secretory immunoglobin A (SIgA) is essential for the host to receive information about the bacterial contents of the gut lumen [56]. SIgA binds commensal bacteria and presents them to tolerogenic CD11c[+] CD11b[+] CD8[-] dendritic cells (DC) [56,59]. This feature of SIgA is important for limiting inflammatory immune response towards commensal bacteria. The second important task of SIgA is to defend the host against invading pathogenic bacteria. SIgA recognizes and binds to pathogenic epitopes which eventually leads to the cross-linkage of these antigens, effectively preventing colonization by these species and reducing the inflammatory potential [56].

While the host must protect itself from some colonizing bacteria, others are essential for the development and shaping of the immune system [37,60–63]. The importance of the intestinal microbiota can be seen in studies with germ-free animals, as these animals do not develop a fully competent immune system [37,63].

The intestinal microbiota is involved in the development of different lymphocyte subsets. These induce the differentiation of pro-inflammatory T_H17 effector T-cells [33,64–67] and anti-inflammatory regulatory FoxP3$^+$ T-cells (T_{reg}) [64,68]. By influencing the T_H1 to T_H2 ratio, the intestinal microbiota determines the outcome of systemic immune response [33,37]. The structure of the LPS (Lipopolysaccharide) of intestinal bacteria is responsible for different TLR4 (Toll-like receptor 4) dependent signaling mechanisms which result in varying responses from the innate immune system [69].
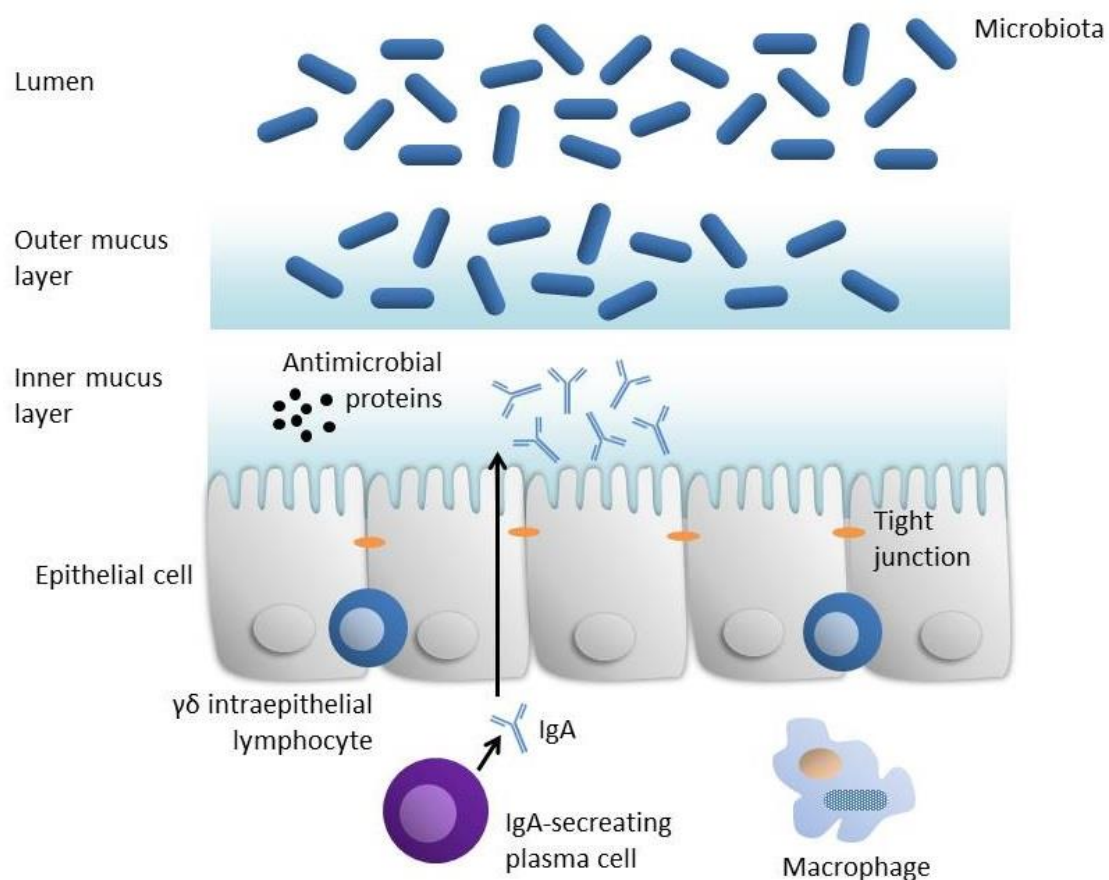
Figure 2: Architecture of the gut helps to defend the intestinal mucosal surface from microbiota [43].

Despite the architecture of the gut and constant interaction with the host immune system, bacteria of the microbiota can overcome the host protection mechanisms and induce pathological states. This can be either due to certain virulence factors of pathogens, for example, facilitating enhanced attachment of pathogens to the epithelium, or to the absence of bacterial products necessary for maintaining homeostasis. One such example is butyrate, an anti-inflammatory substance produced by bacteria of the gut microbiota. The short chain fatty acid (SCFA) butyrate is one of the main energy sources for the epithelial cells and is needed to maintain the tight junctions [70–73]. If butyrate-producing bacteria are

missing, gaps occur between the epithelial cells, allowing increased translocation of antigens from the gut lumen into the host tissue. This phenomenon is referred to as "leaky gut" and leads to an immune reaction resulting in inflammation. SCFAs have been implicated in induction of T_{reg} cell expansion, which is necessary for regulation of the inflammatory response of the immune system [74].

The role of the intestinal microbiota has been discussed in the context of many different diseases. Stimulation of the host immune system and maintenance of the mucosal barrier seem to be primary issues in many of these diseases, as medical conditions, such as type 2 diabetes or obesity, have been associated with general low grade inflammation and an altered gut microbiome[75,76].

1.4 Intestinal microbiota and disease

During the past decade, studies have increasingly been able to associate the intestinal microbiota with a wide variety of diseases ranging from inflammatory bowel disease (IBD) and autoimmune diseases to cancer and mental disorders, as well as conditions like obesity [20,35,44,49,73,77–81]. Although these links have been made, no specific bacterial species have consistently been identified as causative for the medical conditions. Some fractions of the intestinal microbiota are considered to have high potential to trigger inflammation (e.g. *E. coli* [20,82–84]), whereas others are less prone to induce inflammation, or even show the ability to counteract it (e.g. *Lactobacillus, Bifidobacterium, Bacteroides vulgatus*) [20,82]. In addition to working against inflammation-inducing bacteria by decreasing the overall systemic inflammatory level, these beneficial fractions may also improve barrier effects of the intestinal mucosa, allowing less pathogen penetration into host tissue by pathogens, thereby reducing contact with the host's immune system [85]. Yet another possibility is that the anti-inflammatory bacteria directly interact with inflammation-driving components of the immune system to limit the immune response. These hypothesized actions may also intensify their effect by working simultaneously [51]. Some diets and the consumption of probiotics aim to support the beneficial fractions of the intestinal microbiota [85].

Diabetes is an example of a disease that has been associated with an altered intestinal microbiota [65,86,87]. Individuals with type 2 diabetes have lower counts of *Bifidobacterium* and *Faecalibacterium prausnitzii,* whereas children suffering from type 1 diabetes show higher amounts of *Clostridium*, *Bacteroidetes* and *Veillonella* and lower amounts of *Bifidobacteria* and *Lactobacilli*. There is also greater bacterial diversity detected in the feces of healthy children [77].

Autoimmune diseases, such as Type 1 Diabetes (T1D), Rheumatoid arthritis (RA) or multiple sclerosis (MS), have been associated with an altered intestinal microbiota [53,65,86,88,89]. Although genetic predisposition plays an important role in T1D and RA, the low concordance rates in twins suggest that environmental factors strongly impact disease development [88,89]. These autoimmune diseases are characterized by impaired T_{reg} subsets and/or increased

number of autoreactive effector T-cells [89–91]. This favors a proinflammatory immune response, which ultimately leads to tissue damage.

Studies comparing the intestinal microbiota of healthy individuals and T1D patients discovered differences regarding Actinobacteria. Bacteroidetes, Proteobacteria, Firmicutes, Fusobacteria, Tenericutes, and Verrucomicrobia. Despite some disagreements, the conclusions drawn from the experiments are the same: carbohydrate metabolism and therefore the glycemic levels, are affected by the intestinal microbiome, and the decrease of butyrate-producing bacteria reduces mucin and the synthesis of tight junctions, leading to increased permeability of the gut and thus the induction of autoimmunity [65,86,91].

In RA patients, Bifidobacteria and Bacteroides fragilis have been found to be reduced [92].

Studies in mice have associated the intestinal microbiome with MS. Oral application of antibiotics reduced the onset and severity of EAE (experimental autoimmune encephalomyelitis) [93] and mice kept under germ-free conditions did not develop EAE, which is used as a model for MS, or only a mild form. Monocolonization of mice with segmented filamentous bacteria (SFB), on the other hand, is sufficient for inducing EAE [94,95].

Other studies have shown that the intestinal microbiota of obese individuals contains fewer *Bacteroidetes* and more *Actinobacteria* and an overall lower diversity compared to the microbiota of lean subjects, findings on shifts of the amount of *Firmicutes* are controversial [35,77,78] The same studies indicate that the microbiota affects the host; however, the lifestyle of the host also affects the microbiota. The bacteria of obese individuals are more effective in recovering energy from nutrients, therefore transplanting this microbiota into a lean host results in weight gain. On the other hand, if the host loses weight by dieting, the microbiota shifts to a composition resembling that of lean individuals [35].

In general, it is hard to determine whether differences that are observed between diseased and healthy individuals should be considered as cause or effect, as the composition of the microbiota before and during disease development is unknown. In order to answer this question and to better understand normal variation in healthy individuals, an increasing number of prospective studies are monitoring the microbiota of individuals, some beginning immediately after birth [45,96–98]. It is expected that different community compositions or community members that lead to disease development or can be used for diagnostics or therapy in the future will be identified.

1.4.1 Inflammatory bowel disease (IBD)

Inflammatory bowel disease (IBD) comprises two principal types of disease, Crohn's disease (CD) and ulcerative colitis (UC). Symptoms experienced by patients of both include diarrhea, abdominal pain, rectal bleeding, fever, fatigue and weight loss [99]. CD and UC are chronic relapsing inflammatory disorders affecting different parts of the gut. Typical disease onset is during young adulthood; however, 4 % of patients are diagnosed under the age of 5 (early-

onset) or closer to the age of 60 (late onset) [44,100]. These ages correlate with stages of life in which the diversity and stability of intestinal microbiota changes, reinforcing the hypothesis that the gut microbiota composition is involved in disease development [44,101].

Although it is accepted that genetic predisposition plays an important role for the development of IBD, recent developments suggest that the intestinal microbiota and lifestyle of a person are also relevant factors [100,102–104]. The immune system of IBD patients reacts inappropriately toward the intestinal microbiota [105]. Different hypotheses exist on why the immune system overreacts toward the intestinal microbiota:

1. The microbiota is quantitatively and qualitatively normal, but defective mucosal effector T-cells overreact or regulatory T-cells (T_{reg}) do not react strongly enough to suppress the effector T-cell response [50,106].
2. Defects of the immune system lead to inefficient bacterial clearance and therefore to the excessive stimulation of the immune system [44,107,108].
3. The intestinal microbiome is disturbed, leading to either abnormal quantities or types of organisms colonizing the intestinal tract; the immune system is stimulated by the microbial shift [50,99].
4. The mechanisms for the separation of intestinal bacteria and the host immune system are defective. This might be a result of destruction of epithelial barrier integrity due to microbial shifts and changes of the substances available to the host, or to decreased amounts of IgA and antimicrobial peptides [59,74,107,109,110].

The role of the microbiota in IBD has been demonstrated by many human and animal studies. Animal models used for IBD research only develop disease when kept under non-germ-free conditions. As soon as these predisposed animals are colonized with non-pathogenic commensal microbiota, manifestation of IBD can be observed [50,111,112]. The successful use of antibiotics and the treatment with probiotics to reduce intestinal microbiota and effectively treat IBD is another demonstration of this connection [112–117].

IBD has often been associated with either an increase of *Bacteroides* species or Enterobacteria, and with a decrease of Firmicutes. Some studies also link the *Bacteroides*/Firmicutes ratio to development of inflammatory bowel disease [107,118–121]. The reduction of one member of the phylum Firmicutes, *Faecalibacterium prausnitzii (F. prausnitzii)*, has been associated with Crohn's disease [121–124], although a study on pediatric Crohn's disease observed the opposite result [125]. *F. prausnitzii* and some other commensal bacteria that have been isolated from fecal samples are recognized as producers of short chain fatty acids (SCFAs) such as butyrate, which have been found to be reduced in IBD [44,126,127]. Butyrate provides most of the energy for the colonic enterocytes and is reported to be important for maintenance of the integrity of the mucosal barrier [34,107,128].

In IBD patients, certain bacterial strains have not only been reported to vary in presence or quantity, but also the bacterial diversity of the intestinal microbiota have been found to be reduced in patients compared to healthy individuals [118,125,129]. As has been found for other

variations of the microbiota in IBD patients contradicting publications also exist with respect to bacterial diversity [119].The changes of the intestinal microbiome that have been linked to IBD are summarized in Table 1.

Table 1: Changes and their results that have been linked to IBD [44].

Change in:	Result:
Microbial composition	Decrease in α-diversity
	Decrease in *Bacteroides* and *Firmicutes*
	Increase in Gammaproteobacteria
	Presence of *E. coli*, specifically adherent-invasive *E. coli*
	Presence of *Fusobacterium*
	Decrease in Clostridia, Ruminococcaceae, *Bifidobacterium*, *Lactobacillus*
	Decrease in *F. prausnitzii*
Microbial function	Decrease in SCFAs, butyrate
	Decrease in butanoate and propanoate metabolism
	Decrease in amino acid biosynthesis
	Increase in auxotrophy
	Increase in amino acid transport
	Increase in sulfate transport
	Increased oxidative stress
	Increase in type II secretion system, secretion of toxins

Despite the varying results regarding microbiota composition, the common theme is that the microbiota of IBD patients differs from that of healthy individuals [108]. The question that remains is whether the altered microbiota is a cause or a result of dysbiosis. This can be studied best using animal models for IBD as the microbiota can be monitored before and after disease onset.

Various animal models exist, ranging from genetically modified animals with spontaneous disease development to erosive models in which inflammation is chemically induced, for example by the application of Dextran Sodium Sulfate (DSS), which mimics an acute infection [130]. Different animal models have been used to identify the different players, genetic predispositions, host immune systems and microbiota, driving IBD [103,104,131].

Chronic inflammation is a result of a deregulated immune response; therefore, T-cell driven models seem to simulate human disease more accurately regarding to the immunological mechanisms responsible for the induction, perpetuation, and/or regulation of chronic disease [131]. Reconstituting the T-cell population in recombinase activating gene 1 (Rag1) deficient mice which lack B- and T-cells, or in severe combined immunodeficient (SCID) mice, results in chronic inflammation of the colon within a few weeks after the transfer of naïve CD4$^+$CD45^{high} T-cell [132–135].

1.5 Next-Generation-Sequencing (NGS)

The genomic DNA of all living organisms encodes their genetic information. It is the blueprint for all functions needed for development, survival and reproduction of a cell. An abundance of information about an organism can thus be obtained by studying the sequence of the genome. During the past decades many methodologies have been developed to achieve this.

One of the first methods, therefore belonging to the "first generation" of sequencing, was Sanger Sequencing [136]. It relies on a polymerase chain reaction (PCR) that is discontinued by the incorporation of a terminator nucleotide. This leads to DNA strands of various different lengths that are separated by electrophoresis. Each labeled (initially radioactive [136], fluorescent today [137]) DNA strand results in a band according to its size. The terminating base of each band is known, so the sequence of the template DNA can be obtained by analyzing the pattern of the bands (Figure 3).
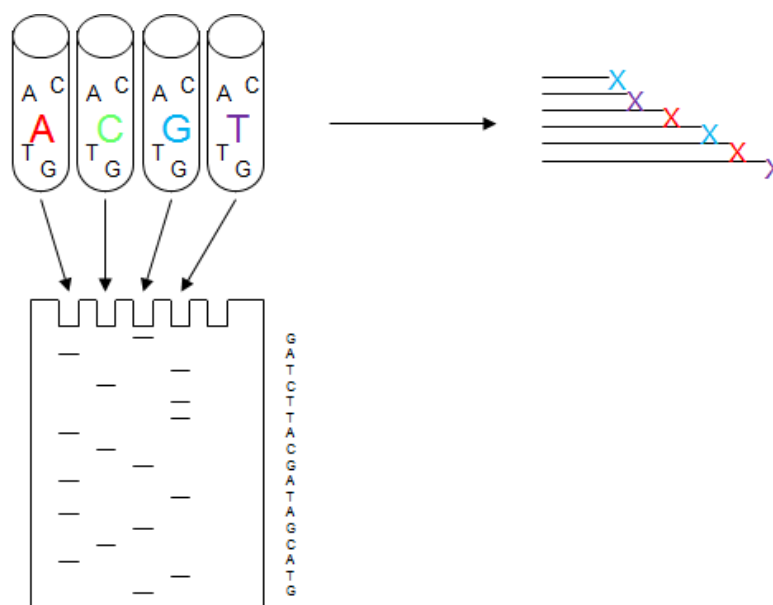
Figure 3: DNA fragments generated during terminated PCR are separated according to size by polyacrylamide gel electrophoresis revealing the DNA sequence.

Due to the Human Genome Project (HGP), which was initiated in the mid 1980's, sequencing strategies were refined and new methods were tested. Whereas strategies based mainly upon Sanger Sequencing were used to complete the HGP [138], it inspired the development of completely new approaches for DNA sequencing. The achievements of this project were therefore not only the generation of the human genome sequence, but also new possibilities for large scale and high-throughput sequencing at constantly decreasing costs and manpower requirements, and increasing speed and accuracy [139]. These new systems were summarized by the term "next-generation sequencing" (NGS). The first NGS system was launched in 2005 by the company 454 [140] and quickly followed by others such as Solexa [141] and Agencourt [142,143].

These new sequencing methods are constantly being improved and have made it possible to sequence large amounts of DNA at low costs. This not only revolutionized the approaches of

studying basic biological mechanisms, but also provided the opportunity to use DNA sequencing in clinical diagnostics.

In contrast to Sanger Sequencing, NGS platforms do not rely on electrophoretic separation of elongation terminated DNA strands produced in individual reactions; instead the DNA sequence is recorded in real-time during the extension of DNA molecules within one reaction [Review 11]. The template DNA molecules are spatially separated in a flow-cell and the systems can detect the sequences of extremely large numbers of different DNA molecules simultaneously, thus dramatically reducing processing time and enormously increasing the data output of each run.

NGS platforms can be separated into systems that detect either the sequence of clonally amplified or single DNA molecules. Clonal amplification is achieved via different methods by each sequencing systems (Figure 4). For 454 pyrosequencing (454/Roche) and SOLiD (Agencourt/Applied Biosystems) for example, single molecules are captured on beads and amplified during emulsion PCR, resulting in beads covered with identical DNA molecules [140]. Illumina/Solexa sequencing, on the other hand, relies on bridge PCR of a template DNA molecule immobilized on the surface of the flow-cell, generating clusters of DNA molecules with the same nucleotide sequence [144,145].
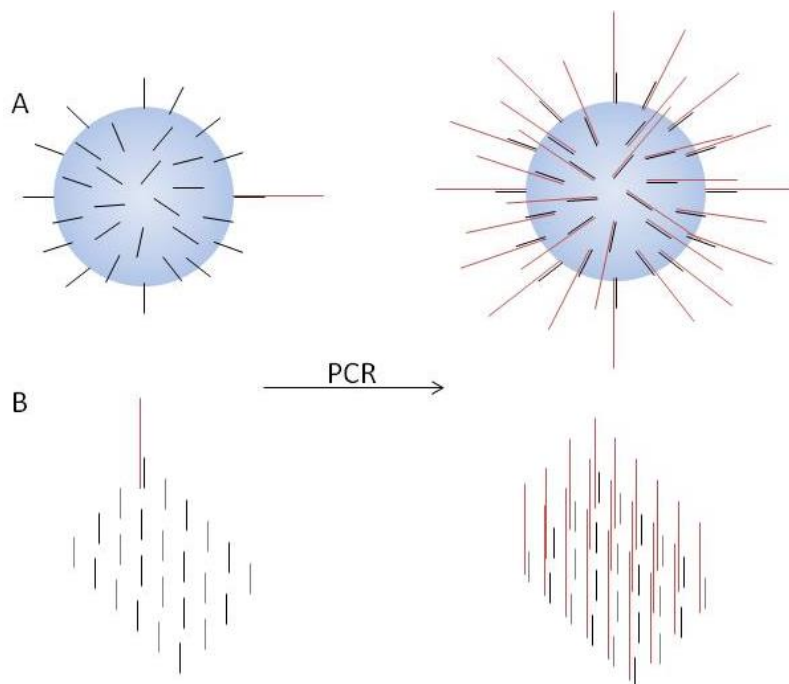
Figure 4: DNA sequences are amplified during an emulsion PCR (A) for 454-Pyrosequencing and SOLiD technologies, or via bridge PCR (B) for Illumina sequencing

Examples of systems based on the sequencing of single DNA molecules are the HeliScope (Helicos Biosciences) [146] and SMRT sequencing (Pacific Biosciences) [147]. Unlike systems that require amplification of the DNA molecules, these systems avoid the introduction of potential errors during PCR, thereby reducing the sources of mistakes within the generated DNA sequence [148].

The specifications of the currently available NGS systems are outlined in Table 2; however, due to the rapid evolvement of the methods they are constantly changing, and accuracy and read lengths are always increasing.

Table 2: Specification of available sequencing platforms

Platform	Reagent Kit or Application	Accuracy	Read length	Data output / run	Sequencing mechanism	Reads	Time / run	Quality	Amplification method	Advantages	Disadvantages
454 GS FLX	GS FLX Titanium XL+	99,997 %	Up to 1000 bp	700 Mb	Pyrosequencing [149]	~1 Million reads	23 h		Emulsion PCR	Long sequencing reads	Error rate for polybases [149]
HiSeq 2500	v4	>80/85 % of bases 99,9 %	1 x 36 bp 2 x 50 bp 2 x 100 bp Or 2x 125 bp	Up to 500 Gb	Sequencing by synthesis [149]	Up to 4 billion paired-end reads	6 days	>85% of bases >Q30 at 2×50 bp >80% of bases >Q30 at 2×100 bp >80% of bases >Q30 at 2×125 bp [150]	Bridge PCR	High throughput [149]	Short reads[149]
MiSeq	v3	>70/85 % of bases 99,9 %	2 x 75 bp 2 x 300 bp	Up to 15 Gb	Sequencing by synthesis	Up to 50 Million paired-end reads	55 h	>85% bases >Q30 at 2 x 75bp >70% reads >Q30[151]	Bridge PCR	High throughput [149]	Short reads [149]
SOLiD 5500 System		99,99 %	Mate-paired: 2 x 60 bp Paired-end 75 bp x 35 bp Fragment 75 bp	Up to 160 Gb	Ligation and two-base coding [149]	Up to 1,4 billion paired reads	Up to 7 days	Greater than 80 % of bases ≥ QV 30	Emulsion PCR	Scalable throughput on single platform; multiple applications in one run	Base calling is difficult
Ion Torrent PGM	Ion 318 Chip v2	98,29 %[152]	200 bp or 400 bp	Up to 2 Gb	Detection of pH differences	Up to 5 million reads	4.4 h or 7.3 h	Mostly Q20[152]	Emulsion PCR		Biased coverage in AT-rich sequences[152]
PacBio RS II		Depending on coverage; up to 99,999 %[153]	50 % of reads >20 kb; maximum >60 kb[153]	500 Mb – 1 Gb[153]	Sequencing-by-synthesis	About 55,000 reads[153]	0,5 – 4 h[154]	QV 20 (single sequence, not consensus sequence)	None	Long reads, no prior amplification, Simultaneous Epigenetic Characterization	Small amounts of reads

1.5.1 Applications of NGS:

There are many applications for DNA sequencing, ranging from determination of the genomic DNA sequence of an organism to medical diagnostics. Depending on the question to be answered, some next-generation sequencing methods are more suitable than others. For example, if a bacterial genome has to be sequenced de novo, it is helpful to obtain long reads due to lack of reference genomes, whereas if the detection of rare bacterial species within a complex community is desired, a high coverage is of greater importance.

Common applications for next-generation DNA sequencing are the analysis of entire genomes of organisms, targeted genomic resequencing to study polymorphisms or gene mutations for diagnostics, microbiome analysis, metagenomics, transcriptomics and identification of DNA-binding proteins or genomic methylation patterns [155].

1.5.2 Studying microbial communities

1.5.2.1 16S rRNA gene analysis

In the past, the identification of bacteria was limited to species that could be cultured. Due to this restriction, only small portions of bacteria present in certain habitats were known and could be associated with specific environments. Of the human-associated microbiome, for example, it is estimated that, depending on the site of interest and taxonomic level, approximately 10 - 60% of the colonizing bacteria cannot be grown in culture [1–3,156–158]. Next-Generation sequencing removes the need to culture bacteria for identification, as this can now be done on the basis of the DNA sequence, and unculturable bacteria are not unintentionally excluded.

The introduction of non-culture dependent identification methods rapidly increased the number of known bacterial species. During evolution some functionally important DNA fragments are conserved, whereas others evolve and rapidly acquire mutations. This enables the determination of relationships between bacterial species by analyzing DNA sequences, even if the actual taxonomy is unknown.

One widely utilized gene for bacterial phylogenetic identification encodes the 16S subunit of ribosomes (16S rDNA) [159–161]. The nucleotide sequence of the 16S rDNA encodes 9 variable regions, as well as regions that are highly conserved within all bacteria [162,163]. By comparing these sequences, especially the variable regions, it is possible, with certain limitations, to determine the taxonomic classification and relationships between bacteria. Bacteria with sequence similarities greater than 99 % are considered to belong to the same strain; 97 % similarity is regarded as a species, 95 % as a genus, 90 % as a family and at least 80 % similarity is considered as a phylum [164,165].

The approach in 16S metagenomics includes DNA isolation, amplification of the selected 16S rRNA gene regions, sequencing and data analysis (Figure 5). Each of these steps holds potential for biasing the result. Depending upon the DNA extraction method, isolation is more efficient from some bacterial species than others [166]. Amplification of the 16S regions can produce chimeric sequences [167], sequencing errors can occur, and bioinformatic analysis

is only as good as the reference databases used and may differ depending on threshold settings. For these experiments it is therefore essential to use the same methods for all samples that will be compared.
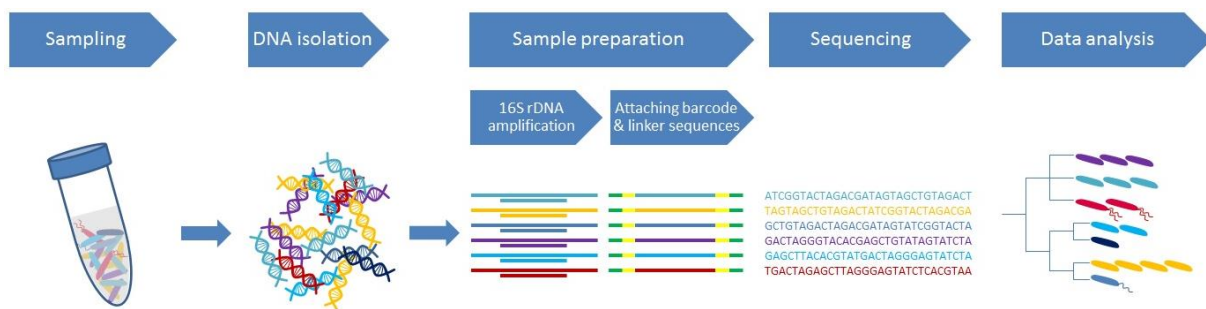

Figure 5: The 16S metagenomics approach includes the steps: DNA isolation, sample preparation, sequencing and data analysis. Information is gained on taxonomic level.

Analyzing the 16S rRNA gene gives information on the composition of the bacterial community within a sample. No conclusions can be drawn as to which biological tasks are taken over by these bacteria, however, or if and which other microorganisms are present in the studied environment.

1.5.2.2   Shotgun metagenomics

Using a shotgun metagenomic approach it is possible to study the entirety of genetic information present within a microbial community. In complex ecosystems it is not feasible to determine what genetic material originated from which organism; however, it is possible to study potential functional differences between samples. For this approach DNA is isolated from a sample and all the genetic material is sequenced. In contrast to 16S metagenomics, no DNA sequences are selected and amplified, rather the complete DNA is sheared into fragments of a defined length and attached to barcodes and sequencing adapters. The total DNA is then sequenced and the data compared to databases for functional and taxonomic assignment (Figure 6) [168,169].
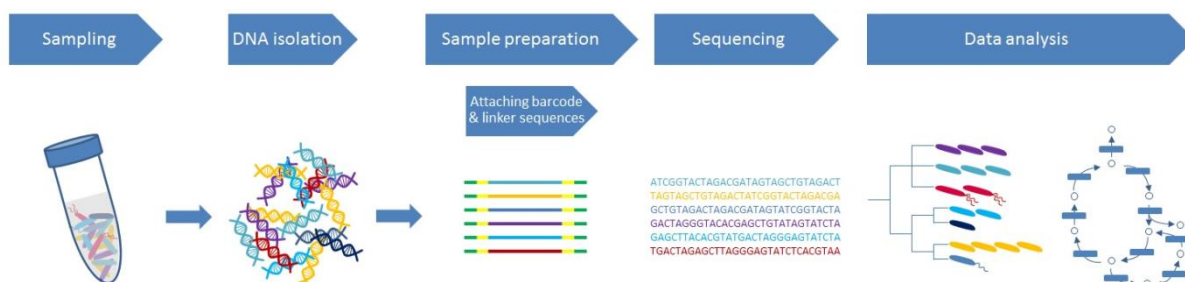
Figure 6: Shotgun metagenomic experiments include DNA isolation, sample preparation, sequencing and data analysis. Information is gained on a taxonomic and a functional level.

Genome sizes vary between bacteria; *Helicobacter pylori* and *Bacteroides thetaiotaomicron*, for example, have genome sizes of 1.67 Mb and 6.3 Mb, respectively. Although the 16S rRNA gene (approximately 1500 bp) copy number ranges from 1 to 15 depending on bacterial species, the ribosomal RNA gene is outnumbered by other genes [170]. The shotgun approach therefore gives information on the genes encoded by bacterial DNA and community

composition. Genes can be classified according to the function of the product encoded by the gene. Commonly used reference databases are KEGG, which maps enzymes onto metabolic pathways, and COG, which uses evolutionary relations to group genes according to related function [18]. This approach enables the detection of all genes present in a metagenome, therefore giving information on all potential tasks carried out by the microbiome, but this does not necessarily mean that all the genes are transcribed and in use.

1.5.2.3 *Metatranscriptomics*

Bacteria transcribe different genes depending on the environmental conditions in which they live. Metagenomic analyses merely give information on which genes can potentially be used for certain biological tasks; metatranscriptomics, on the other hand, can give exact information on which genes are actually transcribed and "in use" [171]. The expression profiles of different samples can be compared to identify specifically regulated genes and composition-independent functional alterations [172].

In contrast to metagenomic approaches, for metatranscriptomic approaches, RNA is isolated instead of DNA. By reverse transcription using random primers, cDNA is created, which is sequenced and analyzed (Figure 7). For the comparison of expression profiles, the mRNA is of interest; however, it only contributes $1 - 5\%$ to total RNA [171]. Various methods can be used for the enrichment of mRNA and reduction of rRNA although some researchers prefer to increase the number of sequences per read to minimize chances of biasing the results instead of enriching mRNA.
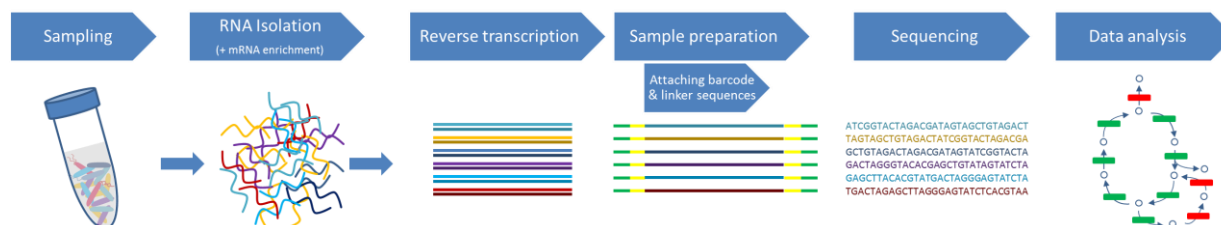

Figure 7: Workflow metatranscriptomics. Total RNA is isolated from a sample and mRNA may be enriched. The RNA is reverse transcribed into cDNA, attached to barcode and linker sequences, sequenced and analyzed.

1.6 Scientific aim of this thesis

Changes of the intestinal microbiota have been linked to many diseases, including IBD, in publications describing alterations of the composition of the microbiota colonizing the intestine of humans. Although some changes have been associated with disease, it is unclear whether alterations induce disease development or if the modified conditions within the inflamed gut lead to the shifts. To answer this question, the microbiota has to be defined before disease onset and during disease development.

The murine T-cell transfer model is helpful for these experiments; disease onset is clearly defined because it is triggered by T-cell reconstitution, allowing determination of the microbial composition colonizing the gut with and without the influence of inflammation.

The first goal of this thesis was to establish a protocol using next-generation sequencing for the analysis of the intestinal microbiota from stool samples. The subsequent aim was to compare the microbiota in the T-cell transfer model to analyze the bacterial composition preventing or leading to development of colitis, as well as the changes induced by disease onset, and related to alterations of the environmental conditions.

The development of IBD, however, may also be influenced by factors other than the bacterial composition of the intestinal microbiome. The colonization of the gut by other microorganisms including viruses and archaea could also impact IBD development; however these organisms are not detected by 16S rRNA gene-based microbiome analysis. In addition, tasks that can be performed by the microorganisms may also be relevant for the development of IBD. The functions performed by these microorganisms may be essential for the availability or removal of substances that are involved in maintaining homoeostasis or in the development of IBD. Establishing suitable methods of shotgun metagenomics to gather information about gut microorganism presence and function based upon their DNA was a further goal of this thesis.

Shotgun metagenomics gives information on which tasks can be performed by the intestinal microbiota, but cannot reveal whether the tasks are actually carried out. Transcriptome analysis of a sample only includes transcribed genes, allowing conclusions to be drawn regarding which functions are carried out by the microorganisms colonizing the gut. To complete the analysis of the intestinal microbiome, the third goal of this thesis was to establish methods needed for metatranscriptome analysis.

2. Materials and Methods

2.1 Materials

2.1.1 Mice

Table 3: List of mouse strains, vendors and housing

Strain	Supplier	Animal housing
C57BL/6Ola Hsd	Harlan Winckelmann	HNO (SPF)
C57BL/6J	Oberberghof Ulm	Oberberghof Ulm (germfree)
C57BL/6J-Rag^{1tm1Mom} (Rag1$^{-/-}$, E^{lo}B^{hi})	Own breeding in FORS, red area	HNO (SPF)
C57BL/6J-Rag^{1tm1Mom} (Rag1$^{-/-}$, E^{hi}B^{lo})	Own breeding	ETH (SPF)

The C57BL/6J animals were bred and kept in isolators under germfree conditions at the animal facility Oberberghof in Ulm. In separate isolators at this animal facility animals were also mono- or co-colonized with *E. coli* mpk and/or *Bacteroides vulgatus*.

C57BL/6Ola Hsd animals were obtained from Harlan Winckelmann at the age of 6-8 weeks and kept under conventional conditions at the animal facility in the ENT-clinic at the University Tübingen (animal facility "HNO"). The spleen of these animals was used to isolate naïve T-cells for transfer into Rag1$^{-/-}$ mice.

The Rag1$^{-/-}$ mice were kept under specific pathogen free (SPF) conditions in isolated ventilated cages (IVC) in two different housing facilities of the University Tübingen ("HNO" and "ETH" (Einrichtung für Tierhygiene).

The animal experiments took place according to the admissions H6/10 as well as the Anzeige §4 01.12.11, §4 13.07.12 and §4 19.12.08.

2.1.2 Antibodies for flow cytometry

Table 4: List of antibodies, Fluorophores, and vendors

Antibody (Ordering Number)	Fluorophore	Clone	Concentration	Supplier
FITC Hamster Anti-Mouse CD3e (553062)	FITC	145 2C11	0.5 µg/µl	BD Biosystems
PerCp Rat Anti-Mouse CD4 (553052)	PerCp	RM4-5	0.2 µg/µl	BD Biosystems
PE Rat Anti-Mouse CD4 (553049)	PE	RM4-5	0.2 µg/µl	BD Biosystems
PE Anti-Mouse CD45RB (553101)	PE	16A	0.2 µg/µl	BD Biosystems
Alexa Fluor 700 Rat Anti-Mouse CD62L (560517)	Alexa Fluor 700	MEL-14	0.2 µg/µl	BD Biosystems

2.1.3 Primers

Table 5: Specifications of primers used for PCR amplification

454-sequencing library preparation		
Name	**Sequence 5'→3'**	**Supplier**
M13 338F	GTAAACGACGGCCAGTGCTCCTACGGGWGGCAGCAGT	biomers.net
revM13 1044R	GGAAACAGCTATGACCATGACTACGAGCTGACGACARCCATG	biomers.net
B_revM13	CTATGCGCCTTGCCAGCCCGCTCAG-GGAAACAGCTATGACCATGA	Eurofins MWG Operon
Forward primer (Linker+Barcode+ M13 338F)	CGTATCGCCTCCCTCGCGCCATCAG-Barcode (→Table 20)-GTAAACGACGGCCAGTG	Eurofins MWG Operon
qPCR		
Name	**Sequence 5'→3'**	**Supplier**
uniF340	ACTCCTACGGGAGGCAGCAGT	biomers.net
uniR514	ATTACCGCGGCTGCTGGC	biomers.net
Bact-F285	GGTTCTGAGAGGAGGTCCC	biomers.net
Univ-R338	GCTGCCTCCCGTAGGAGT	biomers.net
EcoliF395	CATGCCGCGTGTATGAAGAA	biomers.net
EcoliR470	CGGGTAACGTCAATGAGCAAA	biomers.net
qPCR B.v.+B.d. F	TCCATACCCGACTTTATTCCTT	Eurofins MWG Operon
qPCR B.v.+B.d. R	ACTCCTACGGGAGGCAGC	Eurofins MWG Operon

The lyophilized primers were resuspended in ultrapure water (Braun) to a stock concentration of 100 µM. Aliquots of the stocks were stored at -20 °C, and working stocks were diluted to a concentration of 10 µM.

2.1.4 Hardware and software

BiodocAnalyze Ti5 (Biometra)

BD LSRFortesse (BD Bioscience)

BD FACSDiva software (BD Bioscience)

FlowJo (Tree Star Inc., USA)

Bioanalyzer 2100 (Agilent)

2100 Expert software (Agilent)

MEGAN5 software Version 5.2.3 [173]

Qiime[174]

Lightcycler 480 (Roche)

Lightcycler 480 SW1.5 software (Roche)

Leica DMRE microscope (Leica)

Zeiss CLSM 710 NLO (Zeiss)

Zen software (Zeiss)

2.2 Methods

2.2.1 Cultivation of bacteria

Aerobic or facultative anaerobic bacteria such as *E. coli* JM83 were cultivated in 5 ml Luria-Bertani (LB) medium, and *E. coli* JM83 ΔhtrBhtrB$_{pg}$ in 5 ml LB medium supplemented with 0.1 % Ampicillin. Cultures were incubated over night at 37 °C on a shaker at 200 rpm, and were subsequently used to inoculate 50 ml of LB media which was then incubated for another 3-4 hours.

Anaerobic bacteria such as *B. vulgates* mpk were cultivated in soya broth containing liver pieces at 37 °C under anaerobic conditions for two days. The broth was then transferred to 100 ml brain heart infusion (BHI) medium and incubated for another 2 days.

Bacteria grown in liquid culture were harvested by centrifugation at 4000 g for 5 min and the supernatant was discarded.

2.2.2 DNA Isolation from bacterial culture

DNA was isolated from bacterial cultures using the Genomic-tip 100/G Kit (Qiagen). Buffers were prepared according to the specifications given in the Genomic-DNA Handbook.

Following the manufacturer's protocol, pelleted bacteria or bacteria grown on plate were resuspended in 3.5 ml Buffer B1 containing RNase A (0.2 mg/ml). Eighty µl of lysozyme (100 mg/ml; Sigma) and 100 µl of Proteinase K stock solution (Qiagen) were added to the sample prior to incubation for 30 min at 37 °C. After bacterial lysis, 1.2 ml of Buffer B2 were added for deproteinization. The tube was inverted several times to ensure complete mixture of the compounds and incubated at 50 °C for 30 minutes. If the lysate was not clear the incubation time was extended or the particles were pelleted by centrifugation and the supernatant was transferred into a new tube.

A Genomic-tip 100/G column was equilibrated with 4 ml of Buffer QBT, after which the flow-through was removed. The sample was vortexed for 10 seconds and then applied to the column, through which it passed by gravity flow. If liquid remained in the column, slight positive pressure was applied. The column was washed twice using 7.5 ml of Buffer QC each time. The genomic DNA was eluted by the addition of 5 ml Buffer QF and subsequently precipitated by mixing with 3.5 ml of isopropanol (Merck). If the precipitated DNA was visible, it was spooled with a glass rod, transferred into 100 µl Buffer TE, and dissolved on a shaker at 55 °C for 1-2 hours. Alternatively, the DNA was pelleted by centrifugation at 5000 g for 15 min at 4 °C, washed in 3 ml of cold 70 % ethanol and again centrifuged at 5000 g for 10 min at 4 °C. The supernatant was discarded and the DNA resuspended in 100 µl Buffer TE as described earlier.

The genomic DNA was stored at -20 °C.

Gel electrophoresis of the isolated DNA showed that the Genomic Tip 100 kit yielded high quality, non-degraded genomic DNA (Figure 8).
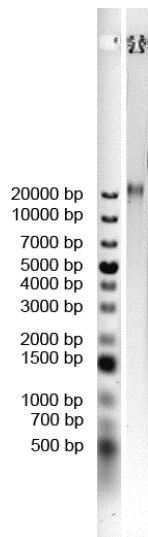
Figure 8: Non-degraded, genomic bacterial DNA isolated using the Genomic Tip 100 kit (Qiagen).

2.2.3 Fluorescence *in Situ* Hybridization (FISH)

2.2.3.1 Sample fixation

2.2.3.1.1 Fecal samples

Fecal samples were resuspended in 1 ml 4 % PFA and stored at 4 °C for 2-4 hours. The PFA was removed by centrifugation for 5 min at 4000 g and the pellet was washed with 1 ml of PBS. After a second centrifugation step, the supernatant was discarded and the pellet was resuspended in 1 ml of 60 % ethanol and stored at –20 °C.

2.2.3.1.2 Bacterial cultures

Bacteria grown on an agar plate were suspended in PBS, resulting in visible clouding. 300 µl of the bacterial suspension was mixed with 900 µl 4 % PFA and incubated at 4 °C for 4 hours. After incubation, the bacteria were pelleted by centrifugation for 5 min at 11000 g. The supernatant was removed and the pellet was washed with 1 ml PBS followed by centrifugation for 5 min at 11000 g. Again the supernatant was removed and the pellet was resuspended in 150 µl 60 % ethanol.

Samples were stored at -20 °C.

2.2.3.2 Sample staining

5 µl of samples, fixed as described in 2.2.3.1 Sample fixation, were pipetted onto a slide. After samples dried they were dehydrated by placing the slide into a reaction tube filled with 50 % ethanol for 3 min and then transferred into 80 % and 100 % ethanol each for another 3 minutes. The slide was dried before the FISH probes (Table 6) were added to the sample.

Table 6: Sequences and flourophores of FISH probes

Name	Sequence 5'→3'	Fluorophore	Supplier
Bac sp. AF488[175]	AF488-CCAATGTGGGGGACCTT	AlexaFluor 488	Eurofins MWG Operon
Ent-CY3[82]	Cy3-CCCCCWCTTTGGTCTTGC	Cy3	Metabion
Bvul-Fluo[82]	Fluo-TCCATACCCGACTTTATTCCTT	Fluorescein	Metabion
Eub338-Cy5[176]	Cy5-GCTGCCTCCCGTAGGAGT	Cy5	Metabion
Bac_spp-Fluo[175]	Fluo-CCAATGTGGGGGACCTT	Fluorescein	Metabion
B. vulgatus-Cy3[82]	TCCATACCCGACTTTATTCCTT-Cy3	Cy3	Metabion

The probes were mixed with freshly prepared ready-to-use hybridization buffer (Table 7).

Table 7: Buffers used for fluorescence in situ hybridization

Prepared hybridization buffer (Stored at 4 °C)	56 ml 5 M NaCl 6.25 ml 1 M Tris-HCl pH8 187.5 ml H_2O
Ready-to-use hybridization buffer	1600 µl prepared hybridization buffer 2 µl 10% SDS 400 µl Formamide
Prepared washing buffer (Stored at 4 °C)	45 ml 5 M NaCl 20 ml 1 M Tris-HCl pH8 Adjust to 1L with H_2O
Ready-to-use washing buffer	100 ml prepared buffer 100 µl 10% SDS

For each sample, 10 µl of the hybridization buffer/probe mix was used. The mix contained 1 µl of each probe. Ready-to-use hybridization buffer was added to the probes to a total volume of 10 µl per sample.

Approximately 1 ml of hybridization buffer was used to moisten a filter paper in the hybridization chamber. The slide was stored in the hybridization chamber in the dark at 46 °C for at least 1 hour. After hybridization, the slide was washed with preheated washing buffer (46 °C), and subsequently stored in the washing buffer for 15 min at 46 °C. To avoid unspecific binding the slide was rinsed with ice cold purified water immediately after removing it from the washing buffer. The samples were dried, covered with 10 µl DAPI (1µg/ml; Merck) and stored at 4 °C for 5 min. The DAPI was rinsed off with ice cold water and the slide was dried at room temperature. Mowiol 4-88 (CALBIOCHEM, prepared according to manufacturer's instructions) was used to conserve the samples and for mounting the cover slip.

Slides were stored in the dark at 4 °C.

A Leica DMRE microscope was used for visualizing the staining of bacteria with the filters listed in Table 8.

Table 8: Filters for Leica DMRE microscope; BP= bandpass filter, LP = longpass filter

Filter	Excitation Range	Excitation Filter	Dichromatic Mirror	Suppression Filter
I3	blue	BP 450-490	510	LP 515
N2.1	green	BP 515-560	580	LP 590
G/R	blue	BP 490/20; 575/30	505;600	BP 525/20; 635/40
Y5	red	BP 620/60	660	BP 700/75
A	UV	BP 340-380	400	LP 425
K3	blue	BP 470-490	510	LP 515

2.2.4 Real-time PCR

Isolated DNA was mixed with reagents for quantitative real-time PCR (qPCR) as listed in Table 9 and amplified according to the thermocycler protocol shown in Table 10. The amount of total bacteria within a sample was measured to provide a reference and allow for the normalization of different samples for comparison. The bacteria detected using specific primers were calculated as proportion of the total amount of bacteria using the following formula:

$$\frac{Amount_{specific}}{Amount_{Total}} = \frac{E^{ct_{Total}}}{E^{ct_{specific}}}$$

Table 9: Components for qPCR

Reagent	Amount	Supplier
2 x QuantiFast SYBR Green PCR Master Mix	12.5 µl	Qiagen
Forward primer (10 mM)	2 µl	Primer dependent
Reverse primer (10 mM)	2 µl	Primer dependent
Ultra pure water	8.5- x	Invitrogen
DNA	50 ng (x µl)	

Table 10: Thermocycler profile for quantitative real-time PCR

Phase	Target temperature [°C]	Duration	Cycles	Ramp rate [°C/sec]	Acquisitions
Denaturation	95	5 min		2	
Program	95	10 sec	40	4.4	
	60	30 sec		2.2	Single
Melting	95	5 sec		4.4	
	46	10 sec		2.2	
	95			0.06	Continuous; 10 acquisitions per °C increase
Cooling	40			2.2	

The specificity of PCR primers was tested *in silico* by comparing the primer sequences with the NCBI nucleotide collection (nr/nt) database using blastn and Primer-BLAST. The ability of the primers to amplify DNA from bacterial strains other than the one of interest was also assessed experimentally. The set of bacteria listed in Table 11 were selected for covering as many different phyla as possible known to colonize the mouse intestine.

Table 11: Experimental assessment of primer specificities with various bacteria from different phyla to minimize the probability of false positive results.

Bacteria Tested	E. coli Primer (EcoliF395 + EcoliR470)	Bacteroidales (BactF285 + UnivR338)	B. vulgatus & B. dorei (qPCR Bv & Bd F + qPCR Bv &Bd R)	Bacteria (uniF340 + uniR514)	Bacteria (M13 388F + revM13 1044R)
B. dorei		✓	✓	✓	
B. fragilis	✗	✓	✗	✓	✓
B. subtilis	✗	✗	✗	✓	✓
B. thetaiotaomicron	✗	✓	✗	✓	✓
B. vulgatus	✗	✓	✓	✓	✓
C. difficile	✗	✗	✗	✓	✓
E. cloacae	✗	✗	✗	✓	✓
E. coli JM83	✓	✗	✗	✓	✓
E. durans	✗	✗	✗	✓	✓
E. faecalis	✗	✗	✗	✓	✓
E. gallinarum	✗	✗	✗	✓	✓
K. pneumonia	✗	✗	✗	✓	✓
L. casei	✗	✗	✗	✓	✓
P. aeruginosa	✗	✗	✗	✓	✓
P. bivia	✗	✓	✗	✓	✓
S. aureus	✗	✗	✗	✓	✓
S. epiderimdis	✗	✗	✗	✓	✓
S. pneumonia	✗	✗	✗	✓	✓
S. thermophilus	✗	✗	✗	✓	✓

The primers were subjected to further evaluation for usability in quantitative real-time PCR (qPCR). Three defined dilutions of the template DNA were generated and amplified. The different c_t (threshold cycle) values resulting from the varying amount of DNA added to the reactions were used for simple linear regression. The slopes of the fitted lines resembled the amplification efficiency of the primer sets and are listed in Table 12.

The melting curve of the primers was assessed to exclude unspecific DNA amplification. None of the three primer pairs evaluated for qPCR showed evidence of random primer binding. Figure 9 displays the melting curve of the universal bacterial primer pair after amplification of DNA extracted from the bacteria listed in Table 11. Figure 9 also presents the melting curves of the *E. coli* and *B. vulgates*/*B. dorei* specific primers. The only differences were the melting temperatures, which varied between 80 °C and 85 °C.
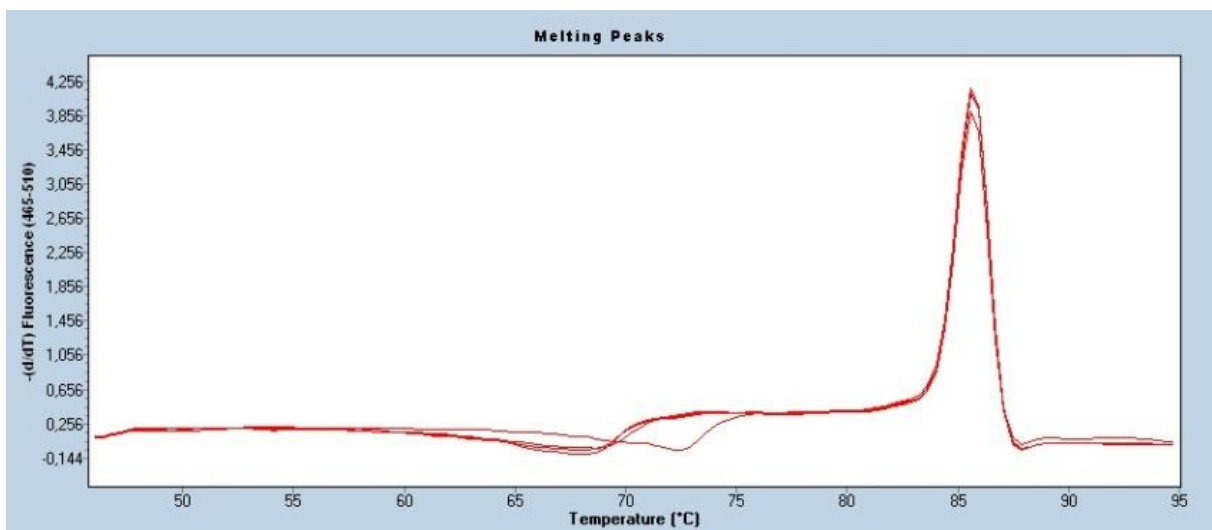

Figure 9: Melting curve recorded during qPCR of bacterial DNA amplified by universal bacterial primer uniF340 and uniR514.

The amount of *E. coli* and *B. vulgatus* / *B. dorei* present in the intestine could be monitored by analyzing fecal samples. As expected, mice treated with antibiotics accommodated fewer bacteria than untreated mice and *E. coli* fed mice (Figure 10). The difference in *B. vulgatus* quantities before and after feeding of bacteria was also evident.
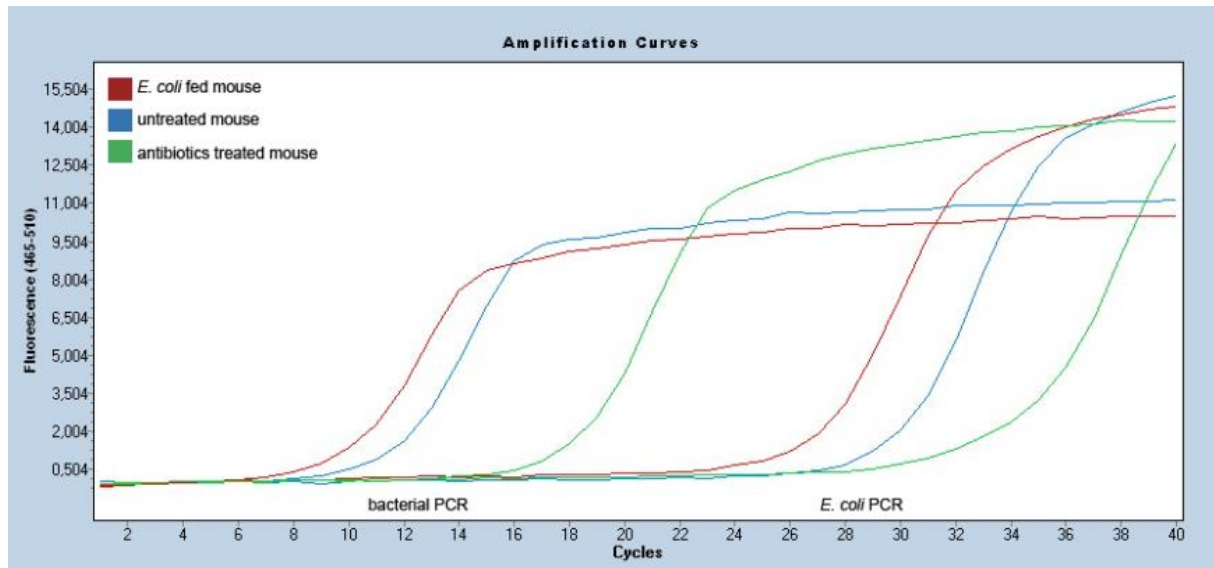
Figure 10: Amplification of bacterial and *E. coli* DNA extracted from the feces of an untreated mouse (blue), after antibiotics treatment (green) and feeding of *E. coli* JM83 (red).

The proportion of each specific species was calculated by combining the species-specific information with that of total bacteria (2.2.4 Real-time PCR). The percentage of *E. coli* colonizing untreated mice was 0.0018 %. After the feeding of *E. coli* JM83, the total bacterial load increased, as did the proportion of *E. coli* (0.0033 %). Streptomycin treatment reduced the total amount of bacteria, however the relative *E. coli* amount increased to 0.0043 %. The same tendencies were observed in animals fed with *E. coli* JM83 ΔhtrBhtrB$_{PG}$. The increase of *B. vulgates* in the stool after the oral application of antibiotics was also detectable by qPCR.

Table 12: Amplification efficiencies of primers used for qPCR

Primer set	Efficiency
E. coli (EcoliF395 +EcoliR470)	1.828
B. vulgatus & *B. dorei* (qPCR Bv & Bd F +qPCR Bv &Bd R)	1.956
Bacteria (uniF340 + uniR514)	1.866

2.2.5 Identification of bacterial strains

For identification of single bacterial strains, they were grown on agar plates to ensure purity.

2.2.5.1 Analyzing 16S gene fragments

DNA was isolated from each bacterial culture by picking a colony from the agar plate and disrupting the cells by heating them for 5 min at 99 °C. The cell particles were separated from the DNA by short centrifugation. 1 µl of the DNA-containing supernatant was mixed with 10 µl 2x Dream Taq Master Mix, 1 µl M13 338F (10 µM), 1µl revM13 1044R (10 µM) and 6.5 µl ultrapure water. The thermocylcler program used for amplification is described in Table 19. The amplicons created during the PCR were purified from the reaction mix using the Agencourt AMPure XP system (Beckman Coulter). A volume equaling 1.8 fold of the PCR volume of Agencourt AMPure XP was added, homogenized with the sample, and incubated

at room temperature for 5 min to ensure binding of the DNA to the magnetic particles. The sample was next placed onto a magnet; the magnetic particles accumulated close to the magnet and the cleared solution was discarded. The DNA was washed twice by the addition of, and after 30 sec, the removal of 200 µl 70 % ethanol (Merck). The reaction tube and the magnet were disassembled and the DNA was separated from the magnetic particles by the addition of 25 µl of AE Buffer (Qiagen). After thorough mixing, the sample was returned to the magnet and the cleared buffer containing the DNA was transferred into a new reaction tube. The DNA was prepared for sequencing according to GATC's specifications and sent to GATC for Sanger Sequencing. The DNA sequence was compared by a standard nucleotide search of the NCBI nucleotide collection database for identification. Uncultured/environmental sample sequences were excluded from the search.

2.2.5.2 *Mass spectrometry*

Bacteria were identified through mass spectrometry using a MALDI-TOF-MS (Shimadzu) and were compared to the SARAMIS Database (Anagnostec) which contains commercially available spectra. They were also compared to spectra from other species which were added individually after verification by 16S sequencing.

2.2.6 Purification of DNA amplicons

PCR amplicons were purified using the Agentcourt AMPure XP system (Beckman Coulter). The volume of the PCR reaction was increased by 1.8 fold with AMPure beads and pipetted up and down until the solution was homogenous. After incubation at room temperature for 5 min the sample was set on a magnet. The magnetic beads to which the DNA was bound accumulated next to the magnet. The cleared supernatant was removed approximately 2 min after the samples were set onto the magnet. Residues of the PCR reaction mix were washed away during two washes of 70 % ethanol (Merck) using 200 µl each. The ethanol remained on the beads for 30 seconds. The sample was removed from the magnet and the DNA was dissociated from the beads by adding 25 µl AE buffer (Qiagen). The sample was placed back onto the magnet for 1 minute, enabling the separation of the dissolved DNA from the beads.

2.2.7 Gel electrophoresis

Gel electrophoresis was used for DNA size separation and visualization. Samples were mixed with the loading buffer Orange G at a ratio of 6:1 and loaded onto a 1 % (for long amplicons) or 2 % (for short amplicons) agarose gel. The electrophoresis was carried out at 100 V for 30-60 minutes. The DNA was visualized by staining with ethidium bromide (1.5 µg/ml; Roth) for 20 minutes, removal of unbound ethidium bromide with distilled water and documentation using the BioDocAnalyze Ti5 (Biometra).

2.2.8 DNA quantification

2.2.8.1 *NanoPhotometer (Implen)*

The absorption of the elution buffer was used as a nucleic acid-free reference. The absorption of a sample was measured between 200 nm and 400 nm. The purity of a sample

was assessed by the ratio of the absorption at 260nm/230 nm; the DNA concentration was calculated from the 260 nm/280 nm ratio.

2.2.8.2 Fluorometric measurement

The Quant-iT™ High-Sensitivity DNA Assay Kit (Life Technologies) was used for fluorometric measurement of the DNA concentration.

2.2.8.2.1 Measurement with the Infinit 200pro

For each of the 8 standards and the samples to be measured, a volume of 100 µl Quant-iT dsDNA HS Buffer was mixed with 0.5 µl Quant-iT dsDNA HS Reagent and transferred into a well of a flat bottom 96 well microplate (Greiner). 10 µl of the standard or 1 µl of the samples were added and mixed thoroughly with the buffer. The absorption was measured using an Infinit 200pro (Tecan) with the settings listed in Table 13. The concentration of the DNA was calculated by using the DNA standard concentrations as a reference.

Table 13: Fluorometer settings for measuring DNA concentrations using the Quant-iT High-Sensitivity DNA Assay Kit

Option	Setting
Shaking (Linear) Duration:	3 sec
Shaking (Linear) Amplitude:	1 mm
Wait (Time)	5 sec
Mode	Fluorescence Top Reading
Multiple Reads per Well (Square (filled))	2 x 2
Multiple Reads per Well (Border)	250 µm
Excitation Wavelength	504 nm
Emission Wavelength	533 nm
Excitation Bandwidth	9 nm
Emission Bandwidth	20 nm
Gain	Optimal
Number of Flashes	50
Integration Time	150 µs
Lag Time	0 µs
Settle Time	0 µs
Z-Position	Calculated automatically from well containing the 10 µg/ml standard

2.2.8.2.2 Measurement with the Qubit2.0

For determination of the DNA concentration using the Qubit 2.0 (Life Technologies), 1 µl sample was mixed with 199 µl Quant-iT dsDNA HS Buffer containing 0.5 % Quant-iT dsDNA HS Reagent in a clear 0.5 ml reaction tube.

2.2.9 **Bioinformatics**

Analysis of the raw sequencing data received from Eurofins MWG Operon was completed by Hans-Joachim Ruscheweyh from the Department for Algorithms in Bioinformatics. The exact pipeline is described in his dissertation and is only outlined in Figure 11 [177].

In the first step, sequences were demultiplexed, meaning the reads were assigned to the sample from which they originated. The sequences were next trimmed to remove the low quality tail of the sequencing reads. Further quality filtering was applied. Data further analyzed using Qiime [178] were clustered into operational taxonomic units (OUTs). Sequences belonging to the same OTU have sequence similarities of at least 97 %. Taxonomic classification was done by RDP using a representative sequence of each OTU.

For analyzing the data using MEGAN, the quality filtered sequences were aligned against the reference Silva database using MALT and imported into MEGAN for further analysis.
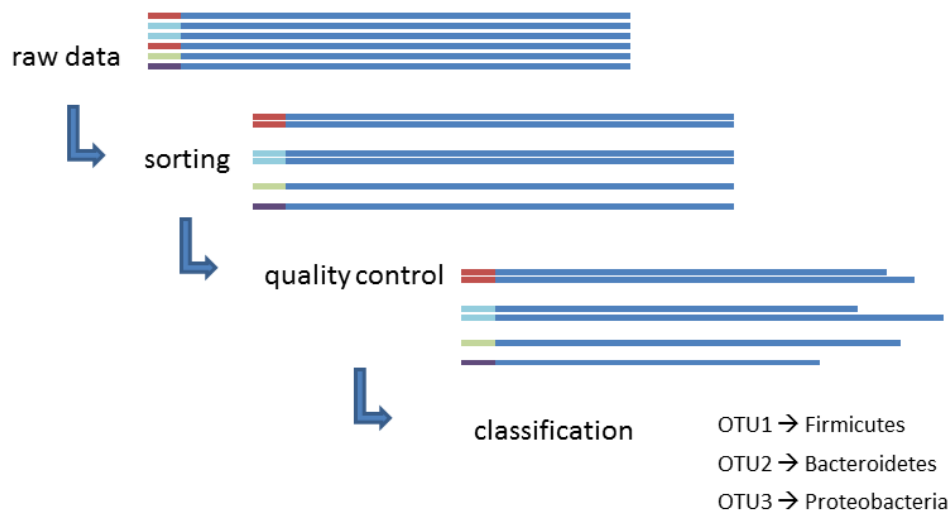

Figure 11: Bioinformatics pipeline for analyzing sequencing data

2.2.10 Adoptive T cell transfer model

2.2.10.1 Isolation of CD4+ T cells from spleen

The spleen of a C57Bl/6 mouse was extracted, transferred to 10 ml of PBS (Invitrogen) supplemented with 1 % FCS (Sigma) (PBS/FCS) and subsequently pressed through a mesh (100 µm). Cells were separated from PBS/FCS by centrifugation at 400 g at room temperature for 5 minutes. Erythrocytes were lysed by the addition of 3-5 ml of Lysis Buffer (155 mM NH_4Cl, 10 mM $KHCO_3$, 10 mM EDTA disodium salt) and incubated for 5 min at room temperature. After adding PBS/FCS to a total volume of 10 ml the cells were centrifuged at 400 g for 5 min. Cells extracted from the spleen were washed with PBS/FCS and counted using a hemocytometer and trypan blue staining to exclude dead cells from count.

Selection of naïve CD4$^+$CD62L$^+$ T-cells was performed using the "CD4$^+$ T Cell Isolation Kit II, mouse" (Miltany Biotec). According to the manufacturer's protocol, for each 1 x 10^7 cells isolated, 40 µl PBS/FCS and 10 µl biotinylated antibody mix were added and the mixture was incubated at 4 °C for 10 min. After the antibodies bound to the cells, for each 1 x 10^7 cells 30 µl PBS/FCS and 20 µl of magnetic beads were added. This mixture was incubated for 15 min at 4 °C. After incubation, PBS/FCS was added to a total volume of 10 ml and the cell/antibody/bead complex was separated by centrifugation for 10 min at 400 g. Next a MACS separation LS column (Miltany Biotec) was equilibrated using 3 ml PBS/FCS. One column was used for a maximum of 1 x 10^8 cells.

After centrifugation, the supernatant was discarded and the cells were resuspended in 1 ml PBS/FCS which was transferred to the prepared MACS separation LS column. Once the cell suspension passed through the column it was washed with 4 x 3 ml of PBS/FCS. The flow-through contained the CD4$^+$CD62L$^+$ naïve T-cells, which were then pelleted by centrifugation for 5 min at 400 g. The supernatant was discarded and the number of living cells was determined as described earlier.

The concentration of the cell suspension was adjusted to 5 x 10^6 T-cell/ml by adding PBS/FCS

5 x 10^5 isolated T cells were transferred intraperitoneally (i.p.) into Rag1$^{-/-}$ mice.

2.2.10.2 Purity and quality control

At least 5 x 10^5 isolated T cells were used to verify the purity and quality. They were suspended in 1 ml CD16/CD32 (Fcγ III/II receptor; clone 2.4G2) antibody solution to avoid unspecific binding of the antibodies to the Fc receptors. After incubation for 15 min at 4 °C, cells were centrifuged at 400 g at room temperature for 5 min, the antibody solution was removed, and the cell pellet was dissolved in 400 µl PBS containing 1 % FCS (PBS/FCS). The cell suspension was divided for 8 different staining combinations as described in (Table 14). For FACS staining at least 0.2 µg of each antibody and 50 µl PBS/FCS were added to the cells and incubated for 30 min at 4 °C. Next the cells were centrifuged at 400 g at room temperature for 5 min, the supernatant was discarded, and the cells were resuspended in 100 µl of PBS/FCS and mixed with 50 µl 4 % PFA. Acquisition of the FACS data was done using a LSR Fortessa (BD) and DIVA software. The data was analyzed using the FlowJo software Version 7.6.1.

Table 14: Antibodies used for verification of quality and purity of the isolated T cells.

Sample	FITC	PE	PerCp	Alexa Fluor 700
1	-	-	-	-
2	CD3e		CD4	
3	CD3e	CD45RB	CD4	
4	CD3e	-	CD4	CD62L
5	CD3e	-	-	-
6	-	CD4	-	-
7	-	-	CD4	-
8	-	-	-	CD62L

2.2.11 Analysis of the murine gut microbiome using 16S metagenomics

2.2.11.1 Sample and metadata collection

Mice sharing a cage were separated from each other and fresh feces was collected in a 1.5 ml tube. The samples were stored at room temperature or on ice until frozen, or DNA was isolated directly from the fresh samples.

For each sample, the metadata listed in Table 15 were collected. Sample metadata that were assessed are listed. As some metadata could only be acquired post mortem, this information is missing for samples collected prior to sacrificing the animals.

Table 15: Metadata associated with assessed samples.

Dataset	Description			
Birthday	Birthday of the mouse from which a sample was collected			
T-cell transfer	Date of T-cell transfer.			
Histology score	Sick (>1, max 3) or healthy (≤1)			
Health state end	Health state of the mouse at the end of the experiment independent of the health state at the time of sampling			
Housing facility	Animal facility in which the mouse lived (ETH or HNO)			
Sequencing run	Groups of samples that were sequenced simultaneously			
Experiment number	Animals of the same experiment were treated simultaneously			
Treatment	Describes the treatment received by a mouse			
Organ morphology (MLN, spleen and colon)	Information about the gross appearance of the spleen, MLN and colon after sacrificing the mice			
Feces texture	Consistence of feces (normal, soft or diarrhea)			
Parents	The breeding pair from which a mouse descended. For purchased animals the company is listed			
Sampling time	The experimental phase in which a sample was collected			
Pool prep	Information about where the sequencing library was prepared			
Gender				
Cell count MLN	Classification of the number of viable cells extracted from the MLN		**Cell count**	**Group**
			< 2 x 10^6	1
Cell count cLP	Classification of the number of viable cells extracted from the colonic tissue		< 3 x 10^6	2
			<4 x 10^6	3
			>4 x 10^6	4

At the end of an experiment the colon was removed and a short piece was used to assess the histology. According to the morphology a histological score was assigned. The scheme used for converting the morphology into a histology score was adopted from Krajina et al. [179]

and is described in Table 16. If attributes of different scores were mixed, the score was accordingly adapted to be as representative as possible. Each section was scored blinded in three independent runs. The final score associated with a section was the mean of the three scorings.

Table 17: Histological scores assigned according to observed morphological changes of colonic sections.

Histological score	Morphological features	Classification
0	normal architecture no inflammatory cells no loss of goblet cells	Healthy
0.5	few inflammatory cells **OR** slight reduction of goblet cells	
1	few inflammatory cells **AND** slight loss of goblet cells Stromal edema	
1.5	features of the histological score 1 and additionally severe loss of goblet cells **OR** intense infiltration of inflammatory cells	Sick
2	Intense inflammatory infiltration **AND** marked reduction of goblet cells hyperplasia of crypts	
2.5	features of the histological score 2 and additionally slight spillover of leukocytes into deeper mucosa **OR** complete loss of goblet cells **OR** distortion of mucosal architecture	
3	spillover of leukocytes beyond mucosa into deeper layers of colonic wall **AND** complete loss of goblet cells **AND** Distortion of mucosal architecture, erosions or ulcerations and crypt abscesses	

2.2.11.2 DNA isolation

DNA was isolated from fecal samples using the QIAamp DNA Stool Kit (Qiagen) with a modified protocol.

One dropping of mouse feces was mixed with 0.5 µl of unwashed glass beads (425-600 µm; Sigma) and 700 µl of ASL buffer and was mixed for 3 min at 30 Hz using a Mixer Mill MM300 (Retsch). After centrifugation at full speed (20 814 g) for 1 min the supernatant was transferred into a new reagent tube. The pellet was treated with 200 µl of lysis buffer (20 mM Tris·HCL, pH 8; 2 mM EDTA; 1.2 % Triton) containing 20 mg/ml lysozyme (Sigma) and incubated at 37 °C for 30 min to enhance the lysis of gram positive bacteria.

Following incubation, 500 µl of ASL buffer was added. The sample was mixed using the Mixer Mill and incubated at 95 °C for 5 minutes. A subsequent physical lysis step, mixing the sample with glass beads in the Mixer Mill, increased the lysis rate of hard-to-lyse bacteria. The samples were then centrifuged for 1 min at full speed and the supernatant was pooled with that collected previously.

To eliminate possible inhibitors of downstream applications, one halfInhibEx (Qiagen) tablet was added and immediately mixed with the supernatant by vortexing until the tablet dissolved. Inhibitors bound to the InhibEx tablet were removed by centrifugation at full speed for 3 min and the supernatant was transferred into a new reaction tube. By repeating the centrifugation step, small residues of the InhibEx tablet that were carried over could be removed.

400 µl of the supernatant was mixed with 15 µl Proteinase K as well as 200 µl of AL buffer and incubated at 70 °C for 10 min. The volume was doubled by adding ethanol. The sample was briefly vortexed to ensure thorough mixing and was then transferred to a QIAamp spin column. After centrifugation at full speed for 1 min the filtrate was discarded and the DNA bound to the spin column was washed with 500 µl of washing buffer AW1. The buffer was removed by centrifugation at full speed for 1 min. A second wash step was performed using 500 µl AW2 buffer and centrifugation at full speed for 3 min. To avoid the carryover of buffer, an additional centrifugation step (1 min at full speed) was done.

The DNA was eluted in 100 µl heated buffer AE (70 °C) and stored at -20°C until used for preparing sequencing libraries.

The purity and quality of the isolated DNA was verified by photometric analysis with a Nanophotometer and gel electrophoresis (Figure 12). The absorption ratios 260/280 nm and 260/230 nm, ideally at 1.8 and between 2.0 and 2.2, respectively, were measured for the majority of the samples.
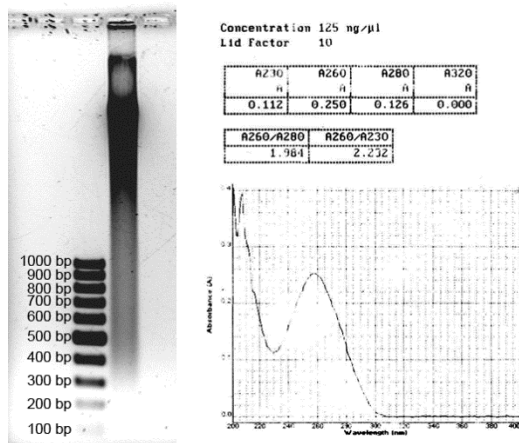
Figure 12: DNA was isolated from fecal samples using the QIAamp DNA Stool Kit (Qiagen) and a modified protocol. Gel electrophoresis was used for determination of degradation; purity and quantity were measured using a Nanophotometer. The DNA was pure and not or only slightly degraded.

2.2.11.3 Amplicon pool preparation

DNA isolated from fecal samples as described above was quantified using a Nanophotometer (Implen) and used as template in a PCR to partially amplify the bacterial 16S gene. The primers (M13 338F and revM13 1044R; Table 5) added to the PCR reaction bind to conserved segments of the 16S gene enclosing the variable regions v3-v6. The composition of the reaction and the amplification conditions are specified in Table 18 and Table 19.

Table 18: PCR reagents for the amplification of 16S gene fragments

Reagent	Amount	Supplier
2x Dream Taq Master Mix	10 µl	Fermentas
M13 338F (10 µM)	1 µl	biomers.net
revM13 1044R (10 µM)	1 µl	biomers.net
DNA Template	50 ng	
Ultrapure water	add to 20 µl	Invitrogen

Table 19: Thermocycler program for amplification of 16S gene fragments

Temperature	Duration	
95 °C	10 min	
95 °C	30 sec	
55 °C	30 sec	20 x
72 °C	45 sec	
72 °C	10 min	
4 °C	∞	

For each 5 samples one negative control was included. The result of the PCR was verified by visualizing the DNA product on a 1% agarose gel as described in 2.2.7 Gel electrophoresis (Figure 13).
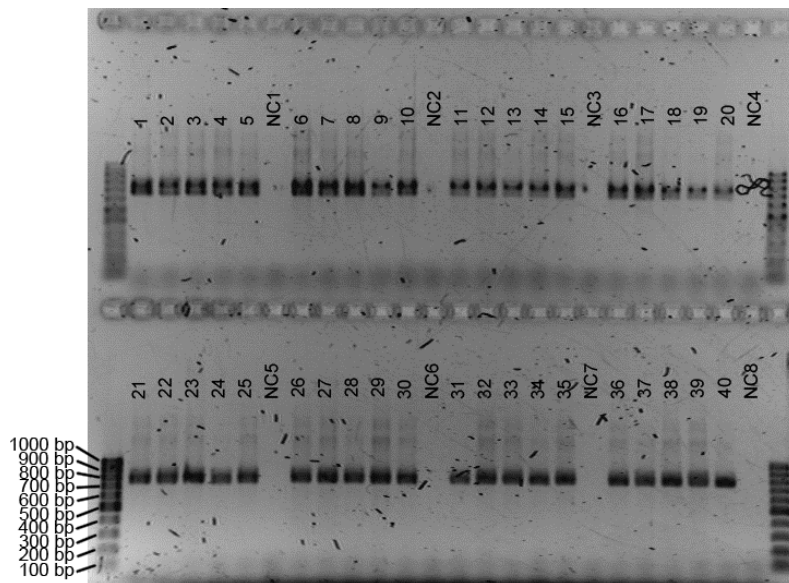
Figure 13: Samples were loaded onto a 1 % agarose gel and DNA was separated according to size during electrophoresis at 100 V for 30 min. Successfully amplified samples were subjected to a further PCR reaction, provided that all negative controls showed no DNA amplification. NC= Negative control, 1-40 = Samples

The amplicons were cleaned using the NucleoSpinExtract II Kit (Macherey-Nagel). The volume of each sample was adjusted to 100 µl with ultrapure water. After the addition and mixing of 200 µl NT Buffer the sample was loaded onto a NucleoSpin Extract II Column and centrifuged at 11 000 g at room temperature for 1 min. The flow-through was discarded and 700 µl of NT3 Buffer was pipetted into the column before repeating the centrifugation. To ensure complete removal of any remaining buffer, the NT3 Buffer in the collection tube was discarded and an additional centrifugation at maximum speed for 2 min was performed, followed by incubation at 70 °C for 4 min with an open lid. The DNA was eluted using 15 µl of ultrapure water which was preheated to 70 °C, and incubated on the column for 1 min. The purified DNA was quantified with a NanoPhotometer. In a second PCR, a barcode sequence (Table 20) and an adapter sequence (Table 5) for 454-sequencing were attached to the 16S fragment previously amplified. The reagents and thermocycler program used for this reaction are described in Table 21 and Table 22.

Table 20: Barcode sequences attached to 16S amplicons for multiplexing

Primer Barcode	MID sequence 5'→3'	Primer Barcode	MID sequence 5'→3'
MID1	ACGAGTGCGT	MID21	TGCGTGAGCA
MID2	ACGCTCGACA	MID22	ACAGCTCGCA
MID3	AGACGCACTC	MID23	CTCACGCAGA
MID4	AGCACTGTAG	MID24	GATGTCACGA
MID5	ATCAGACACG	MID25	GCACAGACTA
MID6	ATATCGCGAG	MID26	GAGCGCTATA
MID7	CGTGTCTCTA	MID27	ATCTCTGTGC
MID8	CTCGCGTGTC	MID28	CTGTGCGCTC
MID9	TAGTATCAGC	MID29	CGACTATGAT
MID10	TCTCTATGCG	MID30	GCGTATCTCT
MID11	TGATACGTCT	MID31	TCTGCATAGT
MID12	TACTGAGCTA	MID32	ATCGAGTCAT
MID13	CATAGTAGTG	MID33	GTGATGATAC
MID14	CGAGAGATAC	MID34	CATAGAGAGC
MID15	ATACGACGTA	MID35	ATGCAGCATA
MID16	TCACGTACTA	MID36	ATCATGCACT
MID17	TCGATCGAGT	MID37	TGAGCTAGCT
MID18	CAGTCAGTAG	MID38	GATGACTGAC
MID19	ACACTGACAC	MID39	CACAGTCACA
MID20	GTACGATCGT	MID40	TGCTAGCATG

Table 21: Reagents used in the PCR for attaching linker and barcode sequences.

Reagent	Amount	Supplier
Primer B_revM13 (10 µM)	1 µl	Eurofins MWG Operon
Forward primer (Linker+Barcode+M13 338F)	1 µl	Eurofins MWG Operon
FastStart Buffer (10x)	2.5 µl	Roche
dNTPs (2.5 mM)	2 µl	Roche
MgCl$_2$ (25mM)	1 µl	Roche
G-C rich solution (5x)	5 µl	Roche
FastStart Taq	0.2 µl	Roche
DNA template	2-5 µl	
Ultrapure water	Add up to 25 µl	Invitrogen

Table 22: Thermocycler program for attaching barcode and linker sequences to the 16S gene amplicons

Temperature	Duration	
95 °C	10 min	
95 °C	15 sec	
60 °C	30 sec	10 x
72 °C	1 min	
72 °C	3 min	
4 °C	∞	

The negative controls of the first reaction were continued as negative controls. Successful amplification of the DNA was observed by visualizing the DNA fragments using gel electrophoresis (1 % agarose gel). By adding several adequately diluted and randomly selected amplicons created during the first PCR, the increase of size and amount generated in the second reaction was verified (Figure 14).
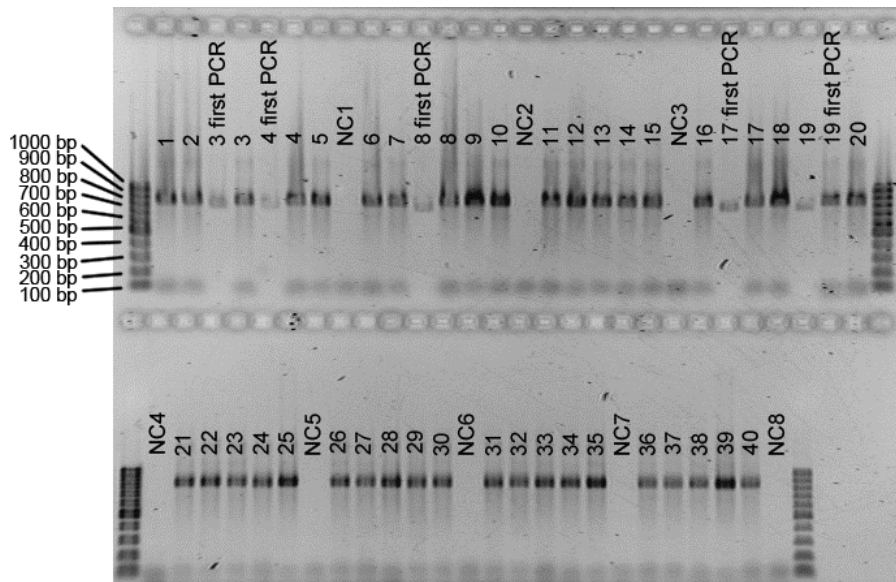

Figure 14: Comparing the bands of a sample after the first and second PCR shows an increase in the amount and length of the amplicons, indicating successful attachment of linker and barcode sequences. NC= Negative control, 1-40 = Samples

In order to verify that the sequences attached to the 16S gene fragment were correct, an amplicon from the second PCR was randomly selected and cloned into chemically-competent *E. coli* DH5α with the CloneJet PCR Cloning Kit (Fermentas; described in 2.2.11.4) and sent to GATC (Konstanz) for Sanger sequencing.

The labeled 16S amplicons originating from different samples were pooled. Different amplification efficiencies were compensated for by increasing or decreasing the amount of DNA added to the pool before purification of the amplicon mixture. The volume of the sample pool was adjusted to at least 400 µl with ultrapure water (Invitrogen) and increased by 1/10 with 3 M sodium acetate and subsequently by 2.5 fold of the pool volume with ethanol (Merck). After the ingredients were thoroughly mixed, the sample was centrifuged at 4 °C and 20,000 g for 30 min. The supernatant was discarded, 500 µl of 70 % ethanol was added to the pellet and the sample was centrifuged for 15 min at 4 °C at 20,000 g. After removing the supernatant the DNA pellet was air dried and resuspended in 50 µl of ultrapure water. Amplicons were size-separated by electrophoresis on a 1 % agarose gel (Figure 15). The band at 750 bp was excised from the gel and purified using the NucleoSpinExtract II Kit (Macherey-Nagel) according to manufacturer's instructions. The DNA was eluted in 20 µl of ultrapure water which was preheated to 70 °C. The DNA concentration of the amplicon pool was determined using a nano photometer (2.2.8.1 NanoPhotometer (Implen)) and sent to Eurofins MWG Operon for sequencing on a GS FLX (Roche). A maximum of 20 or 40 samples were sequenced on quarter and half of a GS FLX sequencing

plate, respectively. The amplicon pool was kept at 4 °C for short term storage and shipping to Eurofins MWG operon.
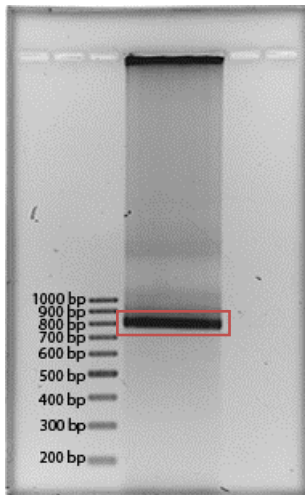

Figure 15: Amplicon pools were purified by electrophoresis and subsequently recovered from the agarose gel

2.2.11.4 Verifying the attachment of the linker and barcode sequences

Prior to 16S amplicon sequencing, it was verified that the appropriate linker and barcode sequence attachment to the 16S gene amplicon, using the CloneJet PCR Cloning Kit (Fermentas).

In the first reaction, the sticky ends of a PCR product were removed in a blunting reaction containing 10 µl of 2x Reaction buffer, 1-2 µl unpurified PCR product, 1 µl DNA blunting enzyme and water to a total volume of 18 µl. The reaction mix was briefly vortexed and centrifuged before incubation at 70 °C for 5 min followed by a quick chill on ice. Next, 1 µl of pJet1.2/blunt Cloning Vector (50 ng/µl) and 1 µl T4 DNA ligase were added, briefly mixed, and centrifuged. The ligation of the PCR product to the vector took place during incubation at room temperature for 5 min. This mixture was used for bacterial transformation. 200 µl of chemically-competent *E. coli* DH5α were thawed on ice before mixing with 5 µl of the ligation product and incubation on ice for 30 min. The bacterial cells were heat shocked at 37 °C for 45 sec and immediately returned to ice for 2 min. The bacteria were grown in 800 µl LB-media at 37 °C and 650 rpm for 1 hour. For selection of the transformed bacteria, bacteria were plated onto a LB-plate containing 100 µg/ml ampicillin and grown over night at 37 °C. One colony was selected and used to inoculate 3 ml of liquid LB-media containing 100 µg/ml ampicillin. After incubation for 4 h at 37 °C and 200 rpm the plasmid was isolated using the NucleoSpin Plasmid Kit (Macherey Nagel). Bacteria were separated from the media by centrifugation for 5 min at 4,000 g and media was discarded. For cell lysis, 250 µl each of Buffers A1 and A2 were added, followed by incubation at room temperature for 5 min, and the addition of 300 µl Buffer A3. The lysate was separated from particles by centrifugation at 11,000 g for 5 min and subsequently transferred to a NucleoSpin Plasmid Column. After centrifugation at 11 000 g for 1 min, the flow-through was discarded and the DNA bound to the silica membrane was washed, first with 500 µl AW Buffer preheated to 50 °C, then with 600 µl Buffer A4. The buffers were removed by centrifugation at 11,000 g for 1 min. Flow-

through was discarded and an additional centrifugation for 2 min at 11,000 g ensured complete removal of buffer residues. DNA was eluted by applying 50 µl of ultrapure water (preheated to 70 °C) to the column for 1 min at room temperature and recaptured by centrifugation for 1 min at 11,000 g.

Amplification of the PCR fragment and the flanking regions of the cloning vector after successful ligation results in an amplicon with a length of approximately 1,000 bp (Figure 16). The reagents used in the PCR reaction are listed in Table 23
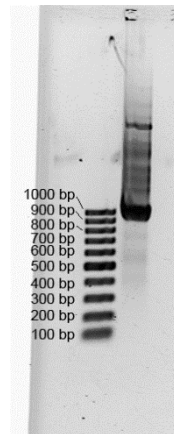
Figure 16: Electrophoresis confirms that the isolated pJet 1.2/blunt cloning vector and the incorporated PCR fragment is 1000 bp in length.

Table 23: Reagents for amplification of the PCR product integrated into the pJet1.2/blunt Cloning Vector

Reagents	Volume	Supplier
10 x Taq buffer with MgCL$_2$	2 µl	Fermentas
dNTP mix (2.5 mM)	2 µl	Fermentas
pJET 1.2 forward sequencing primer (10 µM)	0.4 µl	Fermentas
pJET 1.2 reverse sequencing primer (10 µM)	0.4 µl	Fermentas
Taq DNA polymerase (5u/µl)	0.1 µl	Fermentas
Ultrapure water	15.1 µl	Invitrogen

Sequencing data was analyzed to determine whether the barcode and linker sequences were complete and connected to the 16S rDNA amplicon in the correct manner. Figure 17 shows the sequence received after Sanger sequencing. The linker, barcode and universal primer binding site of the 16S rRNA gene are highlighted.

```
CATTGAATGTATTCTTAGAGCAATTCTAAACAATTGGGGTCCGGCCACGTCCGAAAGTGCACCTACTCTAGAACCAATTA
TTCCATGACTTAACTTAAATAGCGTACACGAGCCGCCCTGCAGCGATAATTATTTGCAATAATTTTAACAAAAGCTCTAA
GTCCTCTCATTAATTCTAGATGATACTTCATCTGGAAATGTCCCATAGTAGCATCACGCTGTGAGTAAGTCTAAACCATT
TTTTATGTGTATATCTCTAATCTACTACTCGATGAGTTTTCGGTATATCTCTATTTTTAACTTGGAGCAGGTTCCATTCA
TTGTTTTTTCATCATAGTGAATAAAATCAACTGCTTTAACACTTGTGCCTGAACACCATATCCATCCGGCGTAATACGAC
TCACTATAGGGAGAGCGGCCGCCAGATCTTCCGGATGGCTCGAGTTTTTCAGCAAGATCGTATCGCCTCCCTCGCGCCAT
CAGCACAGTCACAGTAAACGACGGCCAGTGCTCCTACGGGTGGCAGCAGTGAGGAATATTGGTCAATGGCCGAGAGGCTG
AACCAGCCAAGTCGCGTGAGGGAAGACGGCCCTACGGGTTGTAAACCTCTTTTGTCAGGGAGCAAAGAACGTCACGTGTG
GCGTCTCGAGAGTACCTGAAGAAAAGCATCGGCTAACTCCGTGCCAGCAGCCGCGGTAATACGGAGGATGCGAGCGTTA
TCCGGATTTATTGGGTTTAAAGGGTGCGCAGGCGGGATGCCAAGTCAGCGGTCAAATTTCGGGGCTCAACCCCGACCTGC
CGTTGAAACTGGTGTCCTAGAGTGGGCGAGAAGTATGCGGAATGCGTGGTGTAGCGGTGAAATGCATAGATATCACGCAG
AACTCCGATTGCGAAGGCAGCATACCGGCGCCCAACTGACGCTCATGCACGAAAGCGTGGGTATCGAACAGGATTAGATA
CCCTGGTAGTCCACGCAGTAAACGATGAATACTAACTGTCCGGTCCGAATGAGGACTGGGTGGTACAGCGAAAGCGTTAA
GTATTCCACCTGGGGAGTACGCCGGCAACGGTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGAGGAACATGTG
GTTTAATTCGATGATACGCGAGGAACCTTACCCGGGCTCAAACGACACAGGAATACTTTTGAAAGGAGGTAGCTCTACGG
AGCCTGTGTCGAGGTGCTGCATGGCTGTCGTCAGCTCGTAGTCATGGTCATAGCTGTTTCCCTGAGCGGGCTGGCAAGGC
GCATAGATCTTCTAGAAGATCTCGACAAATGACGG
```

Figure 17: Attachment of linker and barcode sequences was confirmed by Sanger sequencing. Linker sequence Lib-A and barcode sequence are colored in blue and green, the universal binding site for the forward primer in gray. The universal binding site for the reverse primer and the linker sequence Lib-B are colored in light and dark red, respectively.

2.2.11.5 Sequencing

The purified amplicon pool was quantified using a NanoPhotometer (Implen) and sent to Eurofins MWG Operon (Ebersberg) for 454-sequencing on a GS FLX (Roche) with the Titanium series chemistry. Eurofins MWG Operon prepared the sequencing library from the sample pool. A pool of up to 20 Samples was sequenced on quarter segment of a full GS FLX run, and a pool of up to 40 samples was run on half segment.

2.2.12 Analysis of the murine gut metagenome using shotgun metagenomics

2.2.12.1 Sampling

Fecal samples were collected from the individual mice and immediately stored on ice until frozen at -80 °C. Samples were then randomized to minimize technical differences between groups. Samples were collected on the day of T-cell transfer as well as 4, 6 and for some groups, 8 weeks after T-cell transfer.

2.2.12.2 DNA isolation

DNA was isolated from one mouse dropping. 250 µl of Guanidine Thiocyanate and 40 µl of 10 % N-Lauryl sarcosine were added to the frozen fecal sample. the fecal sample was left at room temperature for 10 min to ensure complete thaw before 500 µl of 5 % N-Lauryl sarcosine was added. The feces was subsequently scattered using an inoculation loop, the solution was homogenized by vortexing and then shortly spun down in the centrifuge and incubated at 70 °C for 1 hour.

After incubation, 750 µl of unwashed glass beads (425-600 µm; Sigma) was added. The sample was homogenized through shaking on a horizontal vortexer for 10 min (Vortex Genie 2 (Scientific Industries) speed: 6). 15 mg of polyvinylpyrrolidone (PVPP) was added and the sample was vortexed until the PVPP was no longer at the surface, after which it was centrifuged at 20,814 g for 3 min.

The supernatant was transferred into a new tube and stored on ice, while the pellet was washed with 500 µl TENP (Tris, EDTA, NaCl and PVPP). The resuspended pellet was again centrifuged at 20,814 g for 3 min and the supernatants were pooled. The washing was repeated to a total amount of 3 washes. The pooled supernatants were centrifuged at 20,814 g for 3 min to minimize carryover of contaminants. After centrifugation the supernatant was distributed into two new tubes. 1 ml isopropanol (Merck) was added to each portion of the supernatant and mixed by inverting the tubes. This mixture was incubated at room temperature for 10 min and centrifuged at 20,814 g for 10 min. Supernatant was discarded and the pellet was air-dried under the fume hood for 10 min.

The dried pellet was dissolved in 450 µl Phosphate buffer and 50 µl Potassium Acetate by pipetting up and down. The two samples were pooled, incubated on ice for 90 min and centrifuged at 4 °C for 35 min. The supernatant was transferred into a new tube and 2 µl of RNase (10 mg/ml) were added and mixed into the sample by vortexing. Following a brief centrifugation the sample was incubated at 37 °C for 30 min. 50 µl of Sodium-acetate and 1 ml of ice cold ethanol (Merck) were added and the sample was mixed by inverting the tube several times and incubated at room temperature for 5 min. The ethanol was removed by centrifuging the sample for 5 min at 20,814 g and aspiration of the supernatant. The pellet was washed with ice cold 70 % ethanol and centrifuged at 20,814 g for 5 min. The ethanol was removed and the wash was repeated with 500 µl 70 % ethanol. The DNA pellet was dried at 37 °C for 15 min and resuspended in 100 µl TE Buffer. The DNA was incubated at 4 °C overnight and homogenized by pipetting up and down a several times [180,181].

DNA concentration was assessed as described in 2.2.8.2.2 Measurement with the Qubit2.0. The average DNA yield per sample was approximately 100 ng/µl. Samples were stored at -20 °C until Illumina sequencing.

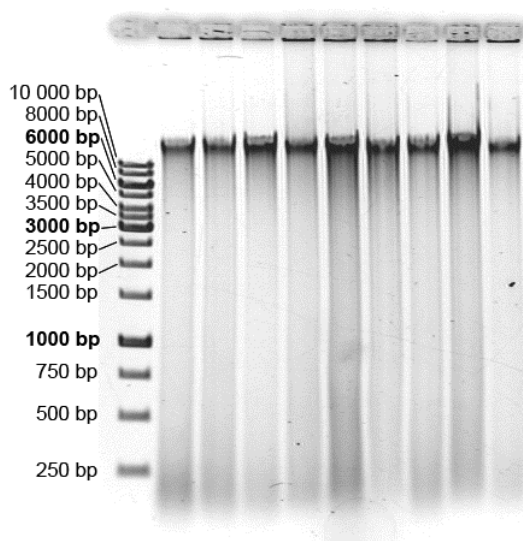Agarose gel electrophoresis was used to verify that extracted DNA was not degraded (Figure 18).

Figure 18: DNA isolated for metagenomic analysis showed no signs of degradation.

Table 24: Buffers for DNA isolation for metagenomics

Buffer	Contents
Na$_2$HPO$_4$ (1M)	8.9 g Na$_2$HPO$_4$ (Merck) 50 ml H$_2$O
NaH$_2$PO$_4$ (1M)	6.9 g NaH$_2$PO$_4$ (Applichem) 50 ml H$_2$O
EDTA, 2 H$_2$O (0.5 M) pH 8	9.305 g EDTA (AMRESCO) 50 ml of H$_2$O Dissolve by heating Adjust to pH 8 using NaOH
Tris-HCl (1 M) pH 7.5 and pH 8	6.05 g Tris (Applichem) 50 ml H$_2$O Adjust to pH 7.5 and pH 8 with HCl
NaCl (5 M)	14.6 g NaCl (VWR Chemicals) 50 ml H$_2$O
N-lauroyl sarcosine (10%)	2 g N-lauroyl sarcosine (Sigma) 20 ml H$_2$O
N-lauroyl sarcosine (5%)	1 g N-lauroyl sarcosine (Sigma) 20 ml phosphate buffer (0.1 M, pH 8)
Potassium acetate (5 M acetate, 3 M potassium)	29.44 g potassium acetate (Merck) 11.5 ml glacial acetic acid (Merck) 28.5 ml H$_2$O Adjust to 100 ml with H$_2$O
Sodium acetate (3 M)	12.304 g Sodium acetate (Merck) 40 ml H$_2$O Adjust to pH 5.2 with glacial acetic acid Adjust to 50 ml with H$_2$O
Ethanol (70%)	70 ml ethanol (Merck) 30 ml H$_2$O
Phosphate buffer (0.1 M) pH 8	9,32 mL Na$_2$HPO$_4$ (1M) 0,68 ml NaH$_2$PO$_4$ (1M) 90 ml H$_2$O Adjust to pH 8
TENP (50 M Tris, 20 mM EDTA, 100 mM NaCl, 1 % PVPP)	1.5 ml Tris (1 M, pH 8) 1.2 ml EDTA (0.5 M, pH 8) 0.6 ml NaCl (5 M) 0.3 g PVPP (polyvinylpyrrolidone K90,Flucka 81438) NB: will not be dissolved Adjust to a total of 30 ml with H$_2$O
Guanidine Thiocyanate (4 M)	12.37 g guanidine thiocyanate (work in the fumehood) (Applichem) 13.5 ml H$_2$O 2.6 ml Tris (1 M, pH 7.5) Agitate overnight in the dark Add 26.1 ml H$_2$O
TE buffer (10 mM Tris, 1 M, 1 mM EDTA)pH 8	200 µl Tris, 1 M, pH 8 40 µl EDTA 0.5 M, pH 8 20 ml H$_2$O
RNase (Ribonuclease, Sigma R6513), 10 mg/ml:	200 µl RNase (100 mg/ml Qiagen) 20 µl Tris, 1 M, pH 7.5 6 µl NaCl, 5 M Adjust to 2 ml with H$_2$O

2.2.13 Analysis of the gut transcriptome of gnotobiotic mice

2.2.13.1 Sampling

Germfree C57Bl/6 mice were colonized with either *B. vulgatus* mpk, *E. coli* mpk, or both bacterial strains, administered for 10 days via the drinking water. Fecal samples were collected after colonization, suspended in 1 ml RNAprotect Bacteria Reagent (Qiagen) and frozen at -80 °C.

2.2.13.2 RNA isolation

The ZR Soil/Fecal RNA MicroPrep (Zymo Research) kit was used to isolate RNA from fecal samples. Before starting with the RNA isolation, the Zymo-Spin IV-HRC Spin Filter was prepared as described by the manufacturer by centrifugation at 8,000 g for 3 min.

Fecal samples stored at -80 °C in RNAprotect Bacteria Reagent (Qiagen) were thawed and centrifuged at 4 000 g for 5 min. The RNAprotect Bacteria Reagent was removed, the fecal sample was transferred into the ZR BashingBead Lysis Tube, and 900 µl of S/F RNA Lysis Buffer was added. The sample was mixed on a horizontal vortexer for 3 min at full speed and subsequently centrifuged for 1 min at 12,000 g. 400 µl of the supernatant was transferred to a new tube, mixed with the same volume of RNA Binding Buffer, pipetted onto a Zymo-Spin IIIC Column and centrifuged at 12,000g for 30 sec. The volume of the flow-through was doubled using ethanol (Merck) and mixed thoroughly. Half of the sample was transferred onto a new Zymo-Spin IIIC Column and centrifuged at 12,000g for 30 sec. This step was repeated with the remaining sample. The RNA was washed by loading the column with 400 µl RNA Prep Buffer and centrifugation at 12,000g for 1 min and then eluted by incubating 100 µl of DNase/RNase-Free Water for 1 min at room temperature on the matrix of the column. RNA was collected by centrifugation at 12,000 g for 30 sec, then further cleaned by passing it through a prepared Zymo-Spin IV-HRC Spin Filter via centrifugation for 1 min at 8,000 g.

DNA contamination was removed by adding 10 µl of 10x DNase I Buffer (Ambion, DNA-free Kit), 2 µl rRNasin (Promega) and 4 µl rDNase I (Ambion, DNA-free Kit) to the sample followedby incubation at room temperature for 30 min. The DNA digestion was terminated by adding 12 µl of DNase Inactivation Reagent (Ambion, DNA-free Kit) and incubated for 2 min at room temperature. During this time the sample was regularly mixed on a shaker, then centrifuged for 1 min at room temperature at 10,000 g and the supernatant then transferred into a new reaction tube.

The volume of the RNA sample after DNase treatment was increased by adding twice as much RNA Binding Buffer and mixing the ingredients well by pipetting up and down. The new volume was doubled by adding ethanol, then homogenized and transferred onto a Zymo-Spin IC Column. After centrifugation at 12,000 g for 1 min the flow-through was removed and the column washed by the addition of 400 µl RNA Prep Buffer and centrifugation at 12,000 g for 1 min. Two more washes followed, first with 700 µl then 400 µl RNA Wash Buffer and centrifugation at 12,000 g for 30 sec. Transfer to a new collection tube

followed by centrifugation at 12,000 g for 2 min ensured the complete removal of any remaining buffer. The RNA was eluted from the column by incubating 10 μl of DNase/RNase Free Water on the matrix of the spin column for 1 min at room temperature and centrifugation at full speed for 30 sec.

For quality control and quantification, 1 μl of the sample was run on an Agilent Bioanalyzer 2100 using the RNA 6000 Nano Kit.

2.2.14 Statistics

Statistical analyses were performed using the unpaired Student's t-Test to compare two groups. Results were considered statistically significant if $p < 0.05$, indicated by *. Error bars represent ± SEM.

3. Results

3.1 Fluorescence in situ hybridization (FISH)

3.1.1 Establishing a hybridization protocol

The protocol used for fluorescence *in situ* hybridization (FISH) was established using fresh or PFA fixed (2.2.3.1 Sample fixation) bacteria, before applying it to fecal samples. The signals detected were generally stronger if the bacteria were PFA fixed prior to FISH. Multiple variations of the approach described in 2.2.3 Fluorescence *in Situ* Hybridization (FISH) were tested.

3.1.1.1 Formamide concentration

Formamide reduces the melting point of double stranded nucleic acids, thus reducing the annealing temperature and consequently the temperature needed for hybridization [182]. Lower incubation temperatures better preserve cell morphology; however, formamide is toxic, so the minimal effective amount should be used. Two concentrations (20 % and 30 %) were tested.

No difference between 30% vs. 20% formamide was found in the *E. coli* sample. Probe hybridization to *B. vulgatus* was barely detectable at the higher concentration, and only became visible after strong enhancement of the signal. For these reasons, the lower concentration was chosen for further experiments.

3.1.1.2 Duration of hybridization

Variations of the hybridization time (between 1 and 18 h) showed that longer hybridization didn't generally improve the labeling of bacteria; in contrast, the Fluorescin and Cy5 dyes seemed to fade faster and signals of the Fluorescin were only visible after strong artificial enhancement. After incubation for 2 h 45 min the hybridization results were comparable to those after 18 h incubation. Figure 19 and Figure 20 show *E. coli* and *B. vulgatus* hybridized with the Eub, Ent or Bac probes after 1 h (F-H), 2.75 h (I-L) and 18 h (M-P) incubation.

3.1.1.3 Pre-hybridization

The necessity of a pre-hybridization step was investigated by covering samples with hybridization buffer and incubating them at 46 °C for about 40 min prior to hybridization. The efficiencies with and without pre-hybridization were comparable, demonstrating that the additional step was unnecessary.
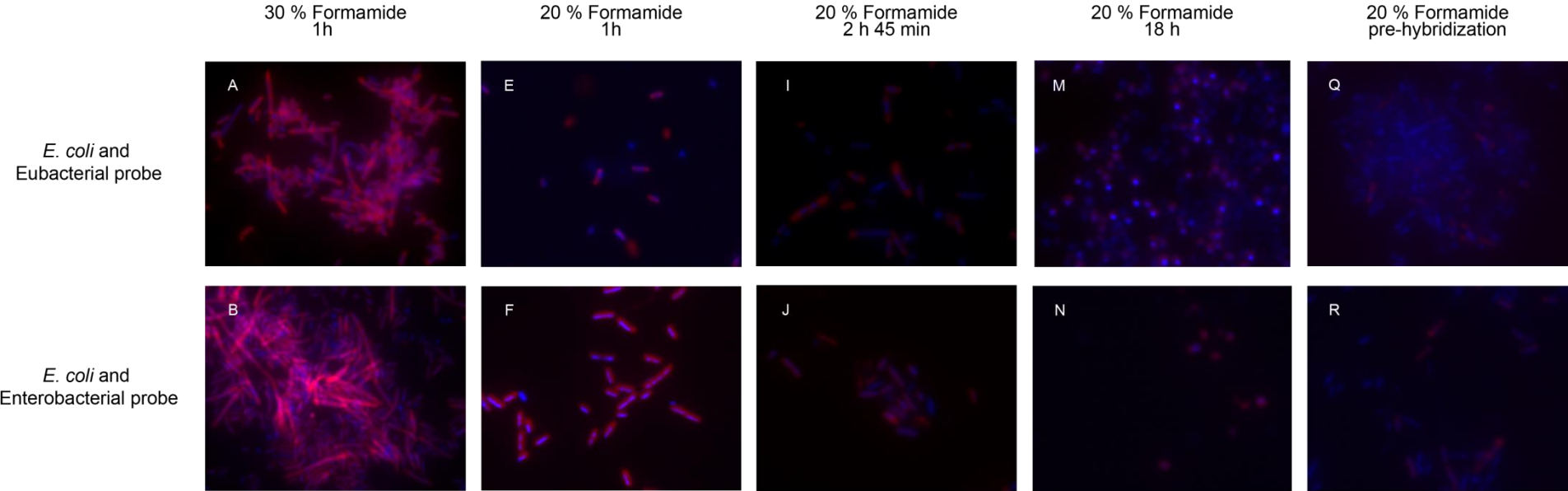
Figure 19: Fluorescent in situ hybridization (FISH) of *E. coli* with an eubacterial, an enterobacterial probe and one specific for *B. vulgatus* with different hybridization conditions. The best results were achieved using 20 % Formamide and hybridizing the probe for 1 h; therefore, these conditions were used in all following experiments.
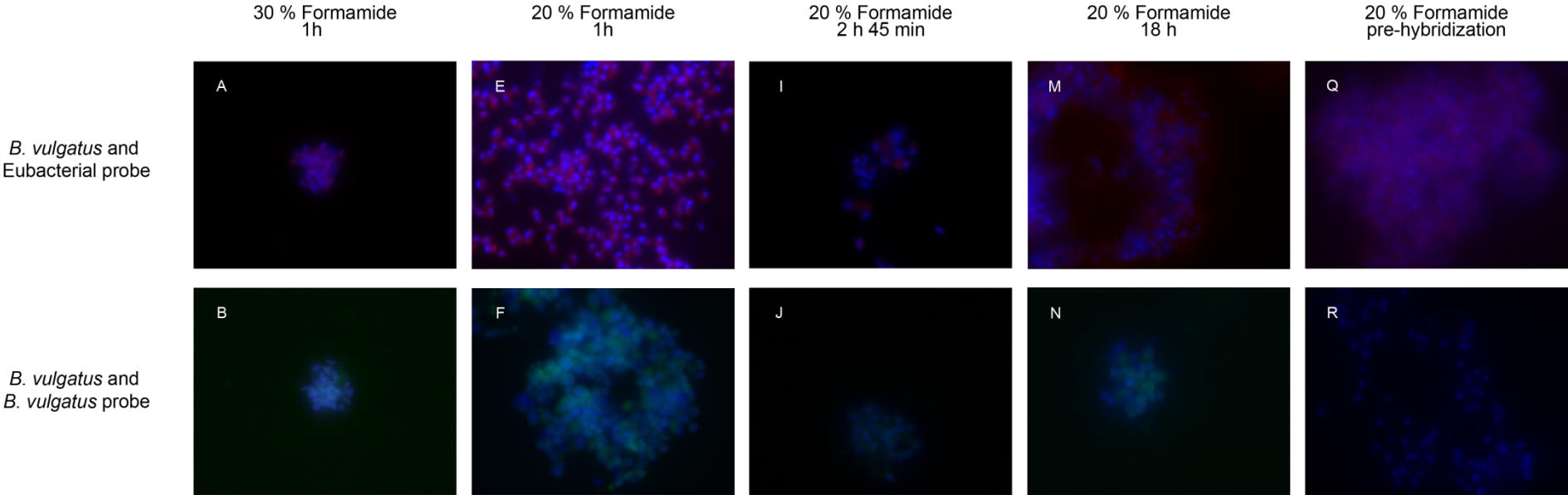
Figure 20: Fluorescent in situ hybridization (FISH) of *B. vulgatus* with an eubacterial, an enterobacterial probe and one specific for *B. vulgatus* with different hybridization conditions. The best results were achieved using 20 % Formamide and hybridizing the probe for 1 h; therefore, these conditions were used in all following experiments.

3.1.2   Selection of appropriate probes

The specificity of the FISH probes used for labeling was appraised by *in silico* comparison of the probe sequences with known bacterial genome sequences, and by experimentally testing if bacteria other than those predicted were recognized by the probes.

Different bacterial species were selected to confirm the probe specificity. Probes with sequences specific for eubacteria, Enterobacteria, Bacteroidetes or *Bacteroides vulgatus* were selected and labeled with Cy5, Cy3, Fluorescin or Alexa Fluor 488. DNA was also stained using DAPI. PFA fixed microbes were hybridized with the probes and visualized using a Leica DMRE microscope with a 100 fold magnification.

Table 25 summarizes which bacterial species were targeted by the selected probes. With the exception of Staphylococci, all bacteria could be stained using the Eubacterial (Eub) probe. Only Enterobacteria (Ent) were marked by that specific probe and the Bacteriodetes (Bac) probe hybridized with all bacterial strains belonging to this genus whereas the *B. vulgatus* specific probe did not bind to other *Bacteroides*.

Table 25: The probes to detect eubacteria, Enterobacteria, Bacteroidetes and *B. vulgatus* were tested with PFA fixated bacteria to confirm specificity. ✓ indicates successful hybridization whereas ✗ indicates that no fluorescence could be detected.

Probe / Bacteria	Eub	Ent	Bac	B.v.
B. vulgatus	✓	✗	✓	✓
B. dorei	✓	✗	✓	✗
B. fragilis	✓	✗	✓	✗
B. thetaiotaomicron	✓	✗	✓	✗
E. cloacae	✓	✓	✗	✗
E. coli	✓	✓	✗	✗
E. durans	✓	✗	✗	✗
E. faecalis	✓	✗	✗	✗
E. gallinarum	✓	✗	✗	✗
K. pneumoniae	✓	✓	✗	✗
P. aeruginosa	✓	✗	✗	✗
S. aureus	✗	✗	✗	✗
S. epidermidis	✗	✗	✗	✗
S. pneumonia	✓	✗	✗	✗

3.1.3   From single to multiple staining

Simultaneous staining of different bacteria within one sample can drastically reduce the time and material needed. The possible combinations of different probes depend on the fluorescent labels and the ability to detect the probes separately; experimental fluorescent dyes were chosen accordingly. The extinction and emission wavelengths of the chosen dyes listed in Table 15 were theoretically far enough apart for differential detection when using appropriate filters.

Table 26: Fluorescent dyes attached to probes for specific labeling of bacteria. Information was retrieved from the Eurofins MWG Operon (www.eurofinsgenomics.eu) and Olympus (http://www.olympusmicro.com) websites

Flourescent dye	Extinction wavelength	Emission wavelength	Color of fluorescence
Fluorescin	495 nm	520 nm	Green
Cy 5	647 nm	673 nm	Red
Cy 3	550 nm	568 nm	Orange
Alexa Fluor 488	495 nm	519 nm	Green
DAPI	350 nm	470 nm	Blue

The filters used with the Leica DMRE for selectively visualizing different dyes are listed in Table 8. Individual Cy3 stained bacteria, however, were also visible through the filter used for the Fluorescin and Alexa Fluor 488 staining. Even though Cy3 fluorescence was not eliminated by this filter, the different bacteria could be distinguished because Cy3 appeared yellow instead of the green specific for Fluorescin and Alexa Fluor 488. The camera used for documentation, however, only detects light intensities, not colors. The signal of the Cy3 dye detected through the filter for green fluorescence was too high to correctly identify the fluorescent dye responsible for the signal, and compensation for intensities was not successful. Cy3 stained bacteria therefore appeared to be stained with Fluorescin or Alexa Fluor 488 (Figure 21).
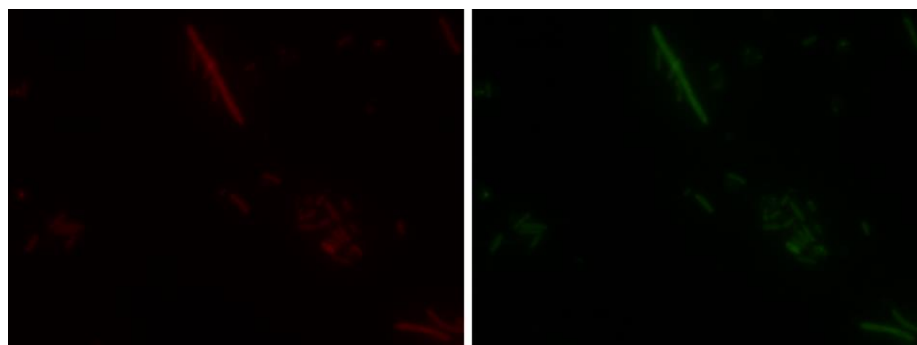
Figure 21: Cy3 labeled bacteria are detected through the filter used for Fluorescin and Alexa Fluor 488.

A mixture of the bacteria listed in Table 25 was stained with the Eubacterial probe (Cy5), the Enterobacterial probe (Cy3) and the Bacteroidetes probe (Alexa Fluor 488). Figure 22 shows the single detection channels (A-D) and overlays (E and F). Bacteria stained with Cy3 were also detected by the filter for Alexa Fluor 488, and therefore appear yellow in the overlay (E). Bacteria selectively stained by Alexa Fluor 488 appear green. The overlay (E) shows that the majority of the DAPI stained bacteria were also colored by the Enterobacterial or Bacteriodetes probes. Only DAPI stained bacteria are highlighted by white circles. The coloring with the Eubacterial probe did not vary from DAPI staining (F).
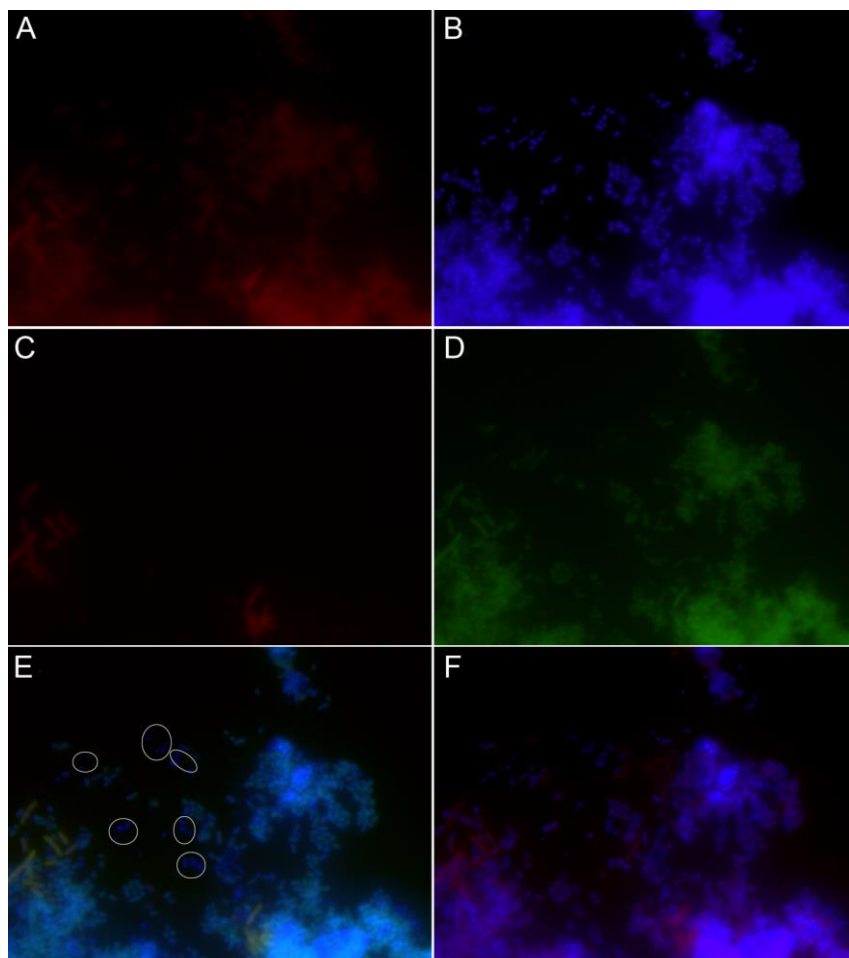
Figure 22: Fluorescent in situ hybridization of a bacterial mixture. The Eubacterial probe (Cy5) colors all bacteria (A) as does the DAPI staining (B). The Enterobacterial probe (Cy5) (C) and the Bacteroidetes probe (Alexa Fluor 488) (D) selectively hybridized with the respective bacteria. An overlay of the Alexa Fluor 488, Cy3 and DAPI staining (E) enables differentiation of bacteria according to coloring. F shows that all DAPI stained bacteria were recognized by the Eubacterial probe.

Changing the microscope used for visualizing the fluorescently dyed bacteria solved this problem. With the confocal scanning microscope LSM 710 NLO (Zeiss) the colors could be differentiated, permitting simultaneous staining and visualization with Cy3, Cy5 and Alexa Fluor 488.

3.2 T-cell isolation

Use of the Rag1$^{-/-}$ T-cell transfer model required the isolation of naïve CD4$^+$ T-cells from the spleen of C57Bl/6 mice. Before transferring them to Rag1$^{-/-}$ mice, the purity of the T-cells was confirmed by FACS analysis. T-cells were stained for the surface markers CD3, CD4, CD45RB and CD62L (L-selectin). T-cells used for the induction of colitis in Rag1$^{-/-}$ animals showed CD62L expression of 78.96% ± 1.69% and CD45RB^{hi} expression of 66,68 % ± 4,59 % (Figure 23). The majority of T-cells were therefore considered naïve.
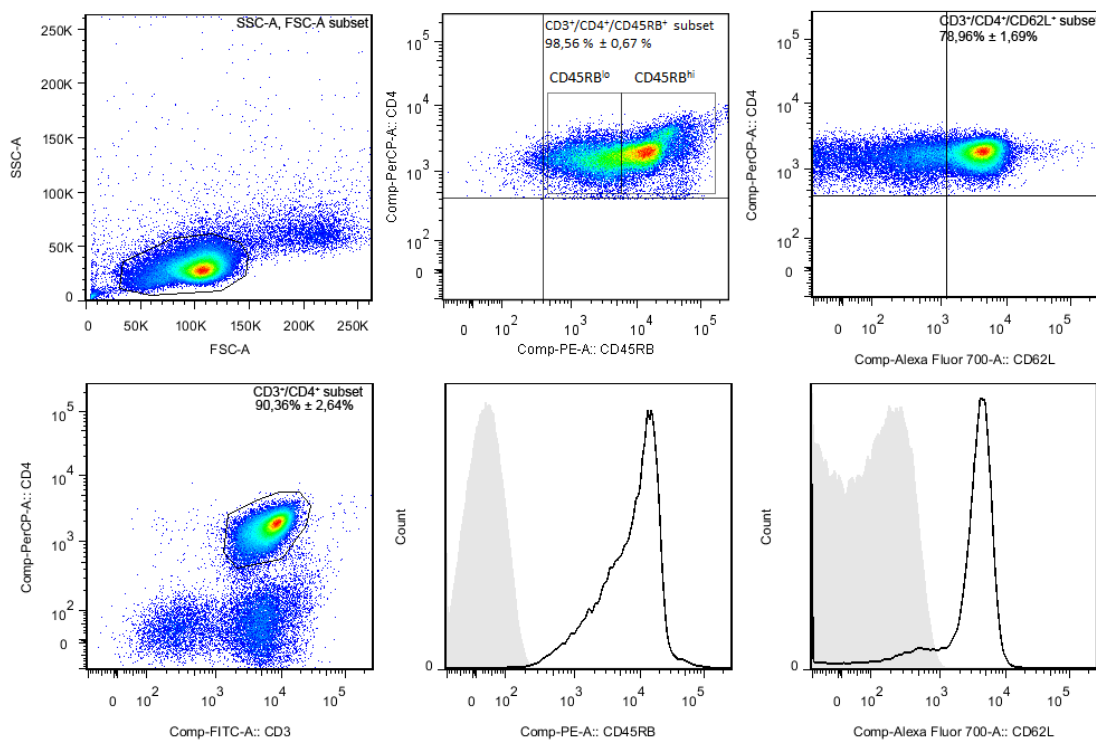
Figure 23: Purity of the isolated naïve T-cells was confirmed by staining for surface markers CD3, CD4, CD45RB and CD62L

The T-cells transferred into Rag1$^{-/-}$ mice could be recovered from the colonic Lamina Propria (cLP) and the Mesenteric Lymph Node (MLN) 6-8 weeks after T-cell transfer and analyzed by FACS (Figure 24 A). In cases where recovery failed, as shown in Figure 24 B, the mice were excluded from the evaluation.
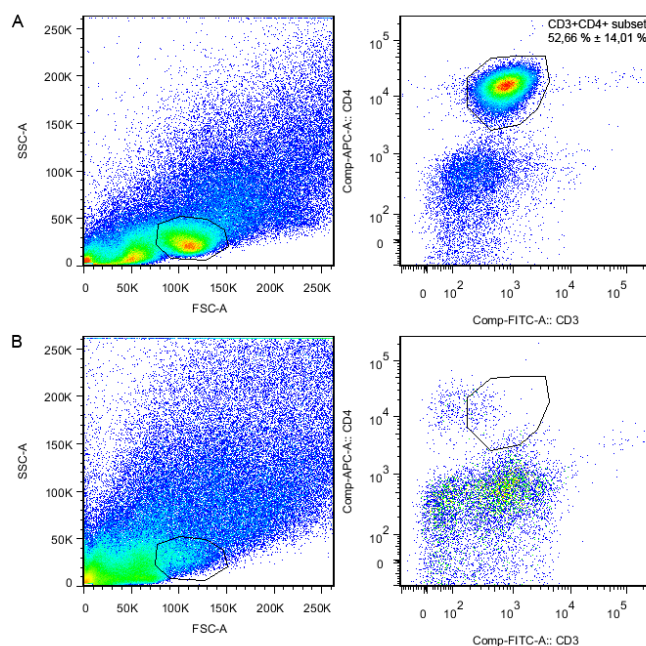
Figure 24: After successful T-cell transfer, T-cells can be isolated from the colonic Lampina Propria. FACS analysis after (A) successful and (B) ineffective T-cell transfer.

3.3 Comparability of individual 454-Sequencing runs

One sequencing run delivers a limited number of reads. If samples are pooled, the reads are divided among these samples resulting in a reduction of the coverage of each individual sample. To ensure sufficient coverage, samples must be split into multiple sequencing runs. The comparability of these runs is assessed by including a control sample. The DNA for this sample was isolated from cultivated bacteria and pooled in an equimolar manner (900 ng per sample). The components of the control are listed in Table 27. Figure 25 shows the 16S amplicon sequencing data of 3 sequencing runs of the control sample on different taxonomic levels, all showing slight differences between the runs. Associations of reads to taxa that were not included in the control sample were seen at the family level and below. These false families were, however, only represented by a maximum of 0.06 % of the reads. The family Peptostreptococcaceae, represented by *C. difficile*, was not detected in two of the three controls.

Table 27: Components of the control sample for 454-sequencing of 16S amplicons. DNA concentrations were measured using the QuantIT HS kit (Invitrogen)

Bacteria	DNA concentration
B. fragilis	276.04 ng/µl
B. subtilis	273.75 ng/µl
B. thetaiotaomicron	330.74 ng/µl
B. vulgatus	279.16 ng/µl
C. difficile	22.85 ng/µl
E. cloacae	293.85 ng/µl
E. coli JM83	286.02 ng/µl
E. durans	40.14 ng/µl
E. faecalis	33.95 ng/µl
E. gallinarum	38.80 ng/µl
K. pneumoniae	275.90 ng/µl
L. casei	10.71 ng/µl
P. aeruginosa	288.35 ng/µl
P. bivia	243.25 ng/µl
S. aureus	67.44 ng/µl
S. epidermidis	29.21 ng/µl
S. pneumonia	202.37 ng/µl
S. thermophilus	163.69 ng/µl
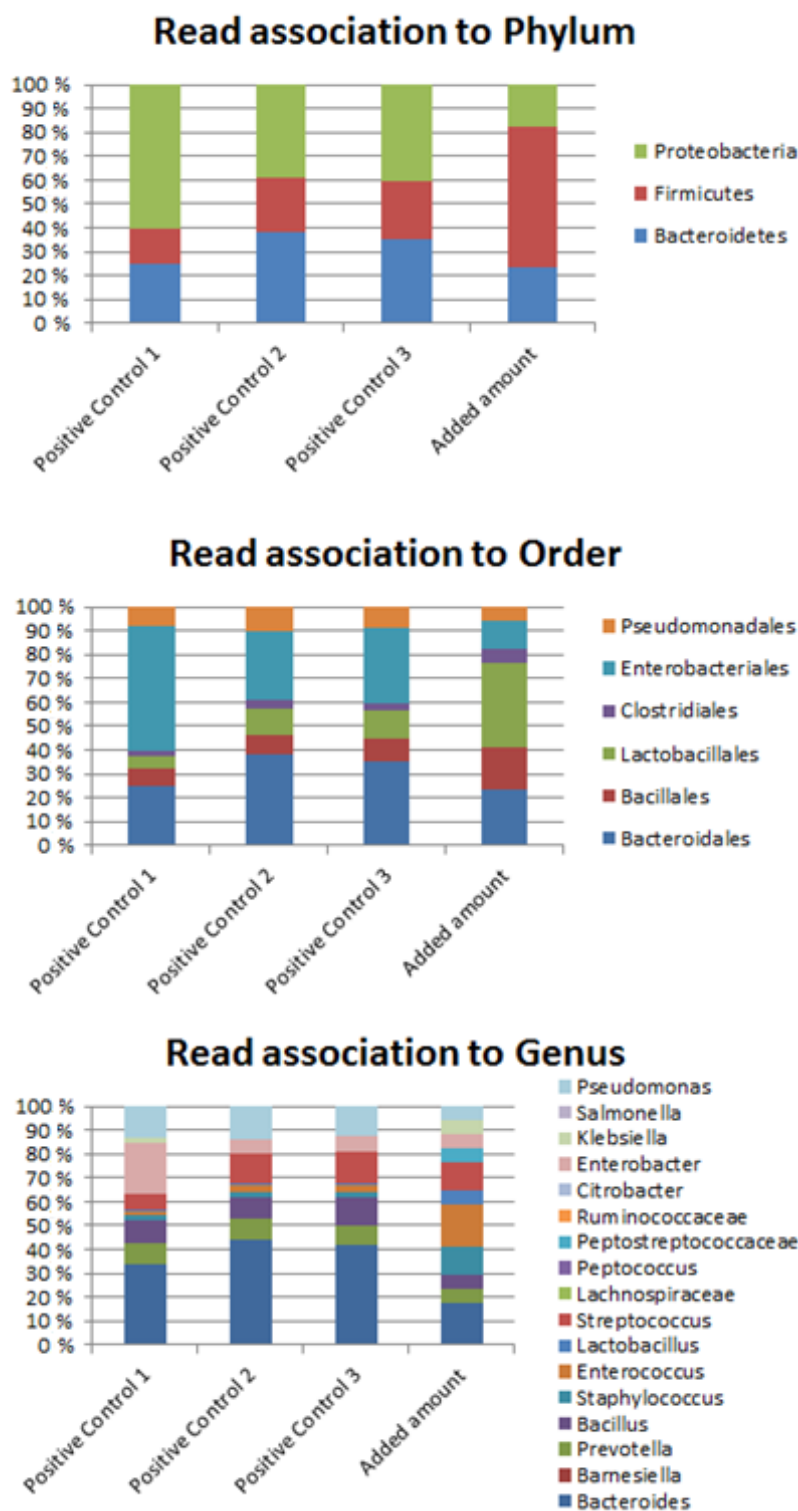
Figure 25: The same positive control was analyzed in three independent sequencing runs and compared to the actual added amounts of bacteria. Only slight variations concerning the amounts of the reads belonging to the different bacterial groups were seen.

3.4 The intestinal microbiome in homoeostasis and inflammation

For analyzing the intestinal microbiome and its influence on maintaining homoeostasis or development of IBD the Rag1$^{-/-}$ T-cell transfer mouse model was used. Genetically identical mice were colonized with two different microbiotas. The Endo^{lo} mice were previously shown to harbor a low endotoxic microbiota and maintain homoeostasis after T-cell reconstitution,

whereas the Endo[hi] mice are colonized with a high endotoxic microbiota and develop colitis 6-8 weeks after transfer of CD4[+] T-cells [183]. The goal was to analyze which components of the microbiota favor IBD development and which components are responsible for maintenance of homeostasis. Rag1[-/-] mice cannot develop an adaptive immune response because the development of mature B and T-cells is impaired. Prior to T-cell transfer, the intestinal microbiota of these mice is therefore not influenced by gut inflammation. After the transfer of naïve T-cells, changes of the intestinal microbiota can be associated to IBD development, and comparing the microbiota of mice with and without manifested IBD can give insight into the influence of inflammation on the composition of the gut microbiota.

The following experimental setup was used to distinguish between the influence of the microbiota on the development of IBD and vice versa, the impact of inflammation on the microbiota:

Endo[lo] and Endo[hi] mice were transferred with naïve CD4[+] T-cells on day 0 after collection of a fecal sample. Further fecal samples were collected over a timeframe of 6-8 weeks (Figure 26). During this time the Endo[lo] mice stayed healthy whereas the Endo[hi] mice developed colitis. The microbiota of Endo[lo] mice and Endo[hi] mice prior to T-cell transfer were compared to identify potential components responsible for predisposition to colitis. In week three the animals did not yet show signs of illness such as diarrhea or a relieving posture, however, the immune reaction has started. Samples were collected at this time to give indication as to which components of the intestinal microbiota were affected by the immune reaction and might even act as drivers of inflammation. Samples collected from Endo[lo] mice in week 3 showed the effect the transfer of T-cells had on the intestinal microbiota without the additional impact of inflammation.
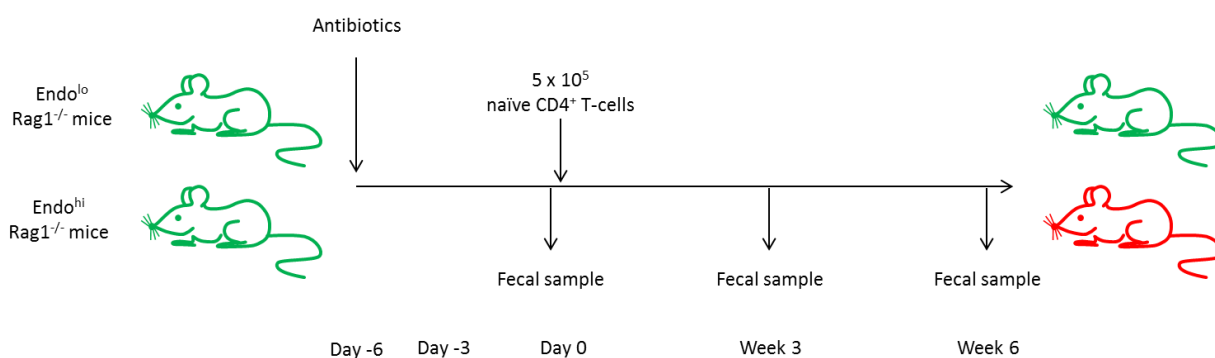

Figure 26: Experimental setup for identification of microbiota components predisposing mice to colitis, and changes during disease development and microbiota during inflammation. Endo[lo] and Endo[hi] mice were treated with antibiotics for 3 days and subsequently transferred with 5 x 10[5] naïve CD4[+] T-cells after collection of a fecal sample. Further fecal samples were collected 3 and 6 weeks after T-cell transfer. Endo[lo] mice stayed healthy (green) after T-cell transfer, whereas Endo[hi] mice developed colitis (red).

All fecal samples were collected and frozen at -80 °C as soon as possible until DNA was isolated and sequenced.

The raw sequencing data was filtered for quality, sorted according to the barcode sequences and analyzed using Qiime [178] by Hans-Joachim Ruscheweyh, from the Department for

Algorithms in Bioinformatics, Tübingen. The data was visualized using Emperor [184] and Megan5 [173].

The similarity of the composition of the gut microbiome of different mice can be analyzed using principal coordinates analysis (PCoA). The distances between the samples within the resulting dot plots represent their similarity. The greater the difference of the intestinal microbiome, the further apart the samples are located on the plot.

Distances among the samples, estimated by PCoA, were visualized using the Emperor software or Megan. For calculating distances between the samples, Emperor uses a phylogenetic tree created by dividing the sequencing reads into operational taxonomic units (OTUs). The reads within one OTU are at least 97 % identical. In contrast, the basis of the analysis with Megan is a phylogenetic tree created after association of the sequencing reads to different taxonomic groups. Association of a read to the lowest (genus) possible, but highest (kingdom) necessary, taxa minimizes false correlations. For example if a read maps to two genera with similar identity it will be linked to the lowest common taxa e.g. family instead.

Emperor is an interactive viewer depicting the PCoA results as 3D plots whereas Megan uses a 2D image. In the first case the three components explaining the highest variances are considered whereas in the latter only the two most prominent components influence the graph. The unweighted unifrac method was used in both cases to calculate the distances between samples creating the distance matrices needed for PCoA.

3.4.1 Microbiota differs between mice resistant or prone to disease

The initial microbiota (= microbiome present when T-cells are transferred) of the Rag1$^{-/-}$ mice is represented by the animal facility in which they were housed (Endo^{lo} = HNO, Endo^{hi} = ETH). The differences between the starting microbiomes was analyzed first. Irrespective of treatment, time elapsed between T-cell transfer and sampling, health state or other metadata, the principal component analysis shows that samples that accumulate within one region mainly originate from Endo^{hi} or Endo^{lo} mice (Figure 27). Notably, not all samples from the same animal housing facility cluster, although a general separation between the samples collected in the other animal facility is apparent, indicating that environmental factors of the housing facilities impacts the microbiota more strongly than other metadata, and even treatment prior to day 0 (Figure 27). The PCoA computed using Megan, however, does not show the same subdivisions (Figure 27) or divisions correlating with other metadata.
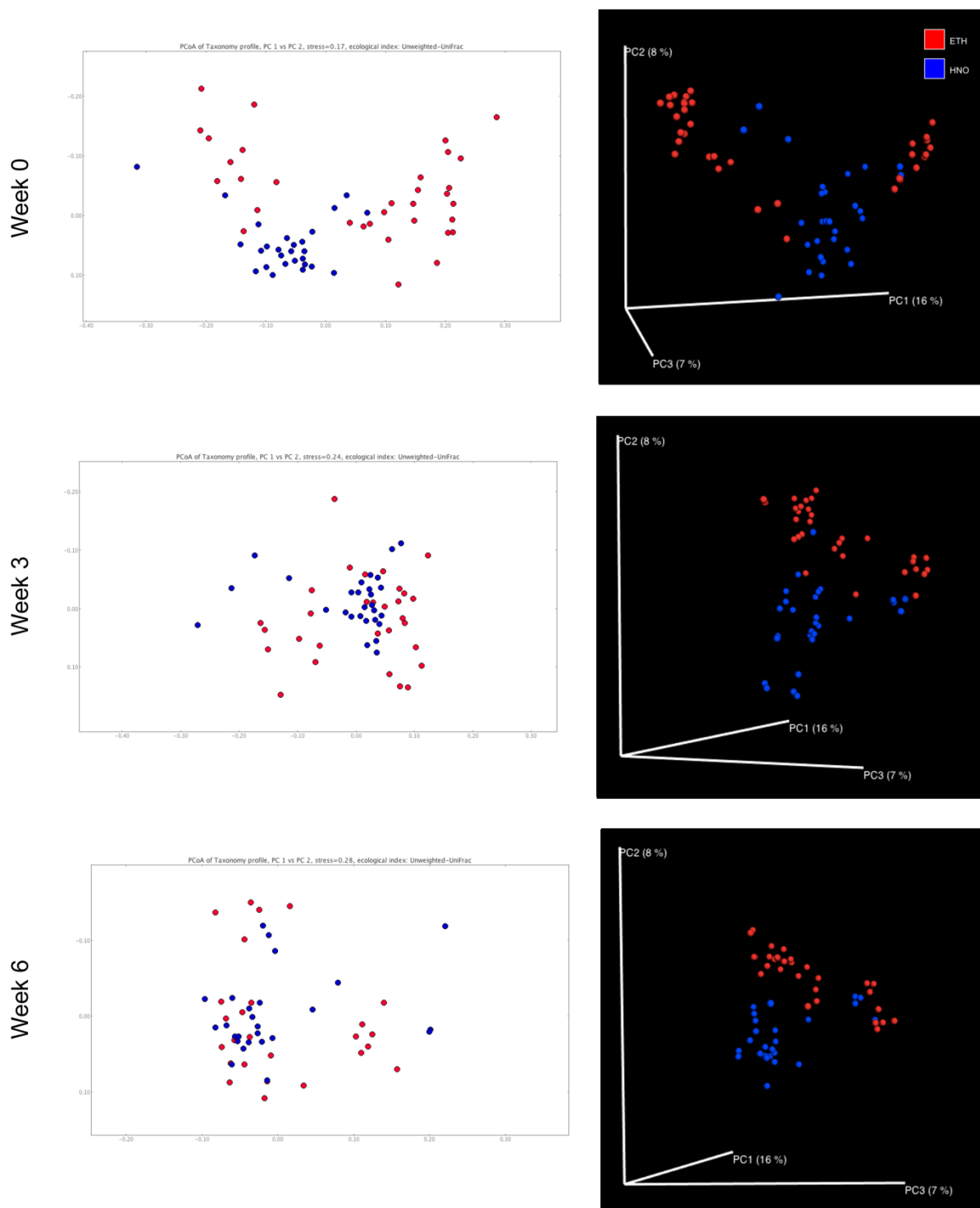
Figure 27: Principal Coordinates Analysis of intestinal microbiota of Endo^hi and Endo^lo mice in week 0, 3 or 6. Each dot represents one sample. Coloring depends upon the housing facility of the specific mouse. Red dots represent Endo^hi mice and blue data points Endo^lo mice. The same PCoA is depicted twice from two different angles. All samples analyzed are included, irrespective of metadata other than sampling time. Percentages associated with the axes quantify the variances explained by the corresponding principal component. The left graphs were created with Megan, the right graphs were made using Emperor.

As the immune system and the intestinal microbiome are known to influence each other, it was expected that after T-cell transfer, changes of the microbiome would become apparent, especially in mice that develop colitis. To analyze differences of the initial microbiota of Endo^hi and Endo^lo animals, samples from week 3 and 6 were eliminated from the plot,

59

resulting in a more evident separation of the Endo[hi] and Endo[lo] animals. The clustering of Endo[hi] and Endo[lo] mice prior to T-cell transfer was evident in the PCoA visualized using Emperor (Figure 28) as well as that created using Megan (Figure 29). In addition, both plots show that the Endo[hi] (red, ETH) animals were divided into two clusters, both of which were separated from the Endo[lo] (blue, HNO) cluster.
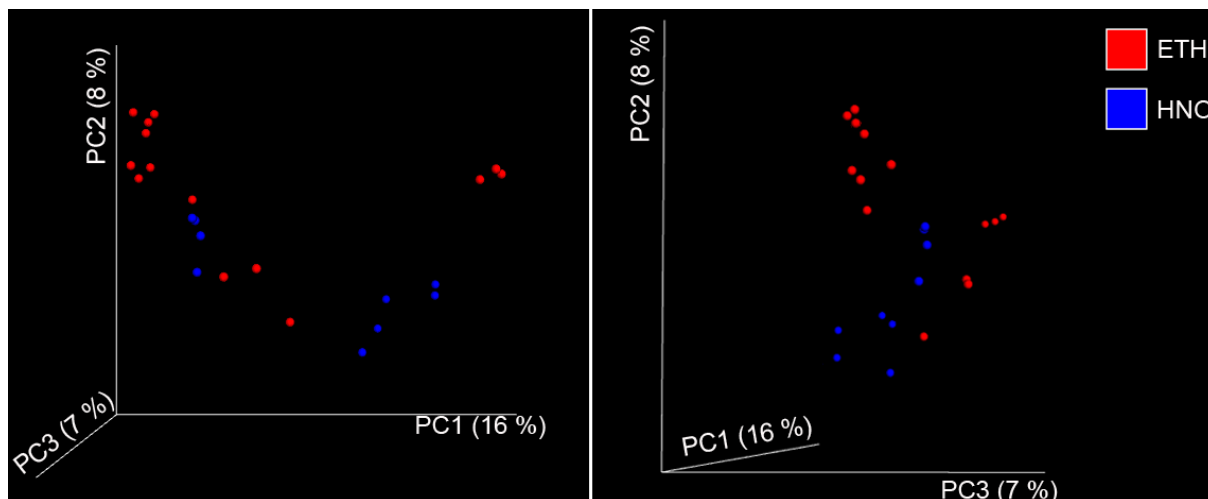

Figure 28: PCoA, visualized using Emperor, of all samples collected in week 0 without selection due to metadata. The samples are colored according to the animal housing facility of the mice. The same plot is shown from two different perspectives.
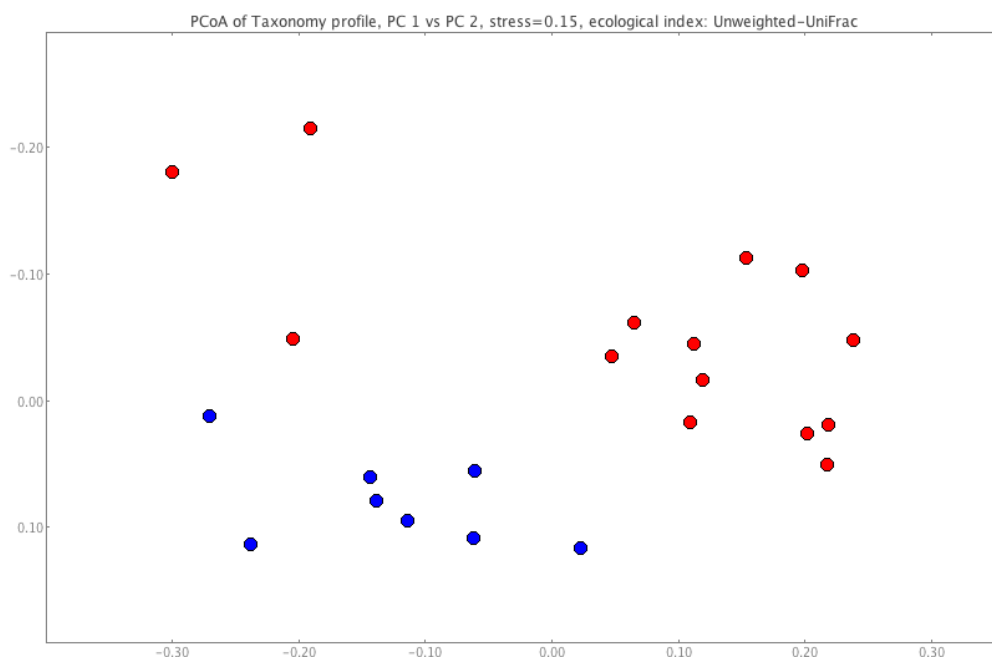
Figure 29: PCoA analysis of all samples collected on the day of T-cell transfer, created using Megan. No pre-selection occurred. Coloring represents the animal facility in which the mouse was housed. Red indicates samples collected in "ETH", blue in "HNO".

3.4.2 Contribution of different metadata to clustering of datasets

Metadata such as the treatment of a mouse, its age, and health state at the end of an experiment or at the time of sampling were collected for every sample analyzed. If one of

these properties severely influenced the composition of the intestinal microbiome, it would be expected that all samples of the same group sharing that property cluster. It could, for example, be assumed that all mice with manifested inflammation harbor similar bacteria that differ from those present in healthy mice. The major factor influencing dataset clustering was correlated with the facility in which mice in an experiment were housed (Figure 28). This correlation was even obvious when including all the sequenced samples, irrespective of any other metadata or treatment and when comparing samples collected prior to T-cell Transfer. The only other Metadata that seemed to be responsible for clustering of datasets, was the experiment an animal was included into (Figure 30). Principal Coordinate Analysis visualized by Emperor and Megan showed the same results.
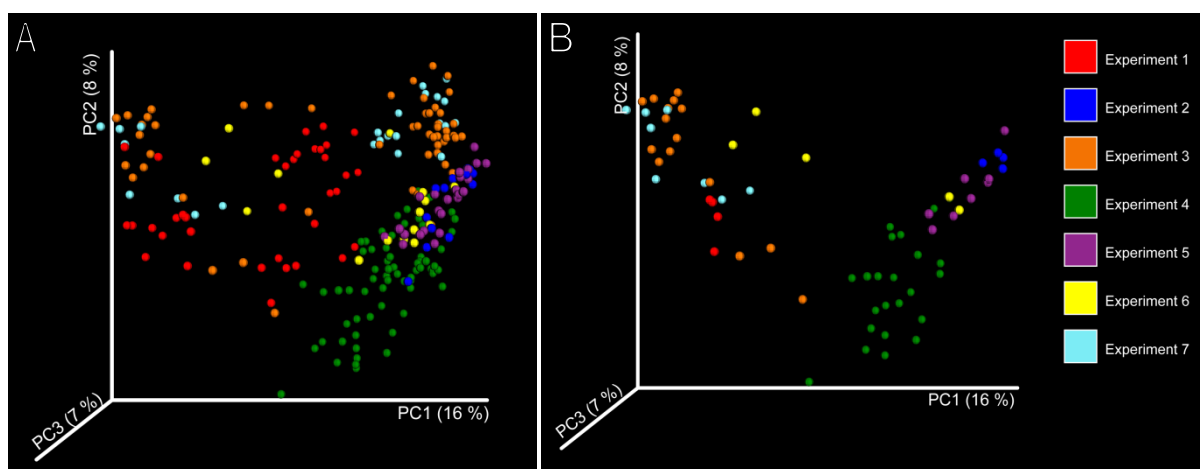

Figure 30: PCoA of all samples (Endo^{hi} and Endo^{lo}) independent of any metadata (A) and all samples collected before T-cell reconstitution (B). The data points are colored according to the experiments in which a mouse was included. The figure was created using Emperor.

To avoid that environmental influences impact the results to an extent that shifts due to treatment and/or disease state are concealed, for further analysis only comparable samples were included into the various analysis. As the PCoA calculated using Qiime and visualized by Emperor include 3 principal components, whereas the visualization using Megan only includes 2 principal components, in the following analyses, only the PCoA created using Emperor are used for analysis of the microbiome composition.

3.4.3 Influence of T-cell transfer on the composition of the intestinal microbiome

Injection of naïve CD4^{+} T-cells induced intestinal inflammation in Endo^{hi} mice whereas Endo^{lo} mice maintained homoeostasis (Figure 31). During development of colitis the composition of the gut microbiome of Endo^{hi} animals appeared to change drastically. The shift in the bacterial community is evident in Figure 32 A. The data sets representing day 0 are highlighted in green. The composition of the microbiome changed most during the first 3 weeks after T-cell transfer, after which no further severe shifts could be detected so later data points are located in close proximity of each other in the PCoA (highlighted by the white mark). The microbes colonizing the gut of mice that stayed healthy, however, did not appear to be greatly affected after T-cell transfer (Figure 32 B). Even though a shift took place within the first 3 weeks, the bacterial community seemed to return to the original composition within the monitored 6 week period. The datasets of weeks 0 and 6 are therefore located

close to each other and those representing the samples collected in week 3 are found at a slight distance.
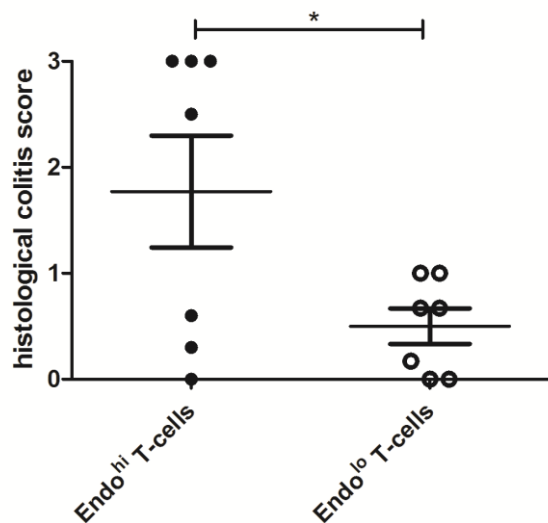

Figure 31: The histological colitis score quantifies the inflammation of the colon of mice. Whereas Endo[hi] mice developed colitis approximately 6 weeks after T-cell transfer, Endo[lo] mice remained healthy. Statistical analysis was done using Student t Test. * p<0.05
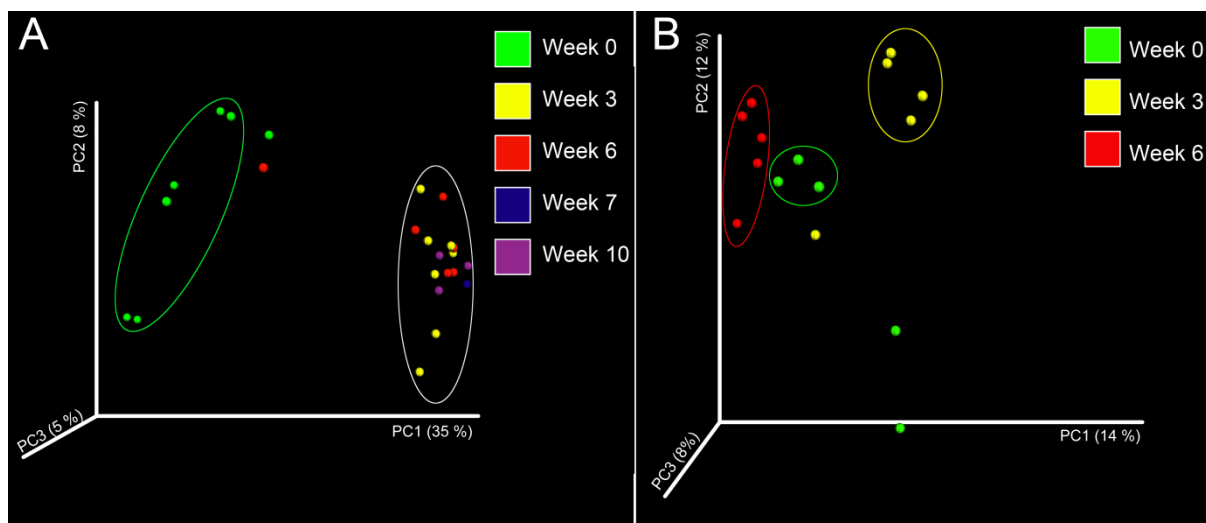

Figure 32: PCoA of Endo[hi] (A) and Endo[lo] (B) mice treated with T-cells. During colitis development a shift of the microbiota was seen (A), whereas the microbiota of mice that maintained homeostasis showed slight variations of their microbiome after 3 weeks but returned towards the initial composition. The figure was created using Emperor.

Independent of the software or algorithm used for computing the PCoAs, the reads of the different bacteria present could be associated with the corresponding taxa, thus enabling studies of the gut microbiome on different taxonomic levels. The composition on the phylum and genus level of the Endo[hi] mice before as well as 3 and 6 weeks after T-cell transfer is shown in Figure 33, Figure 34 and Figure 35, which were created by Jörg Bernhardt using the Paver software (Decodon)[185]. The read associations for the taxonomic levels class, order and family are included in the Appendix. The darker the coloring, the more reads were mapped to a certain taxa. According to the sequencing reads, the amount of Firmicutes slightly decreased during development of colitis whereas Bacteroidetes reads increased. Although

the total amount of reads mapping to Firmicutes decreased after disease onset, the variety of different Firmicutes species detected in sick mice increased as disease progressed.
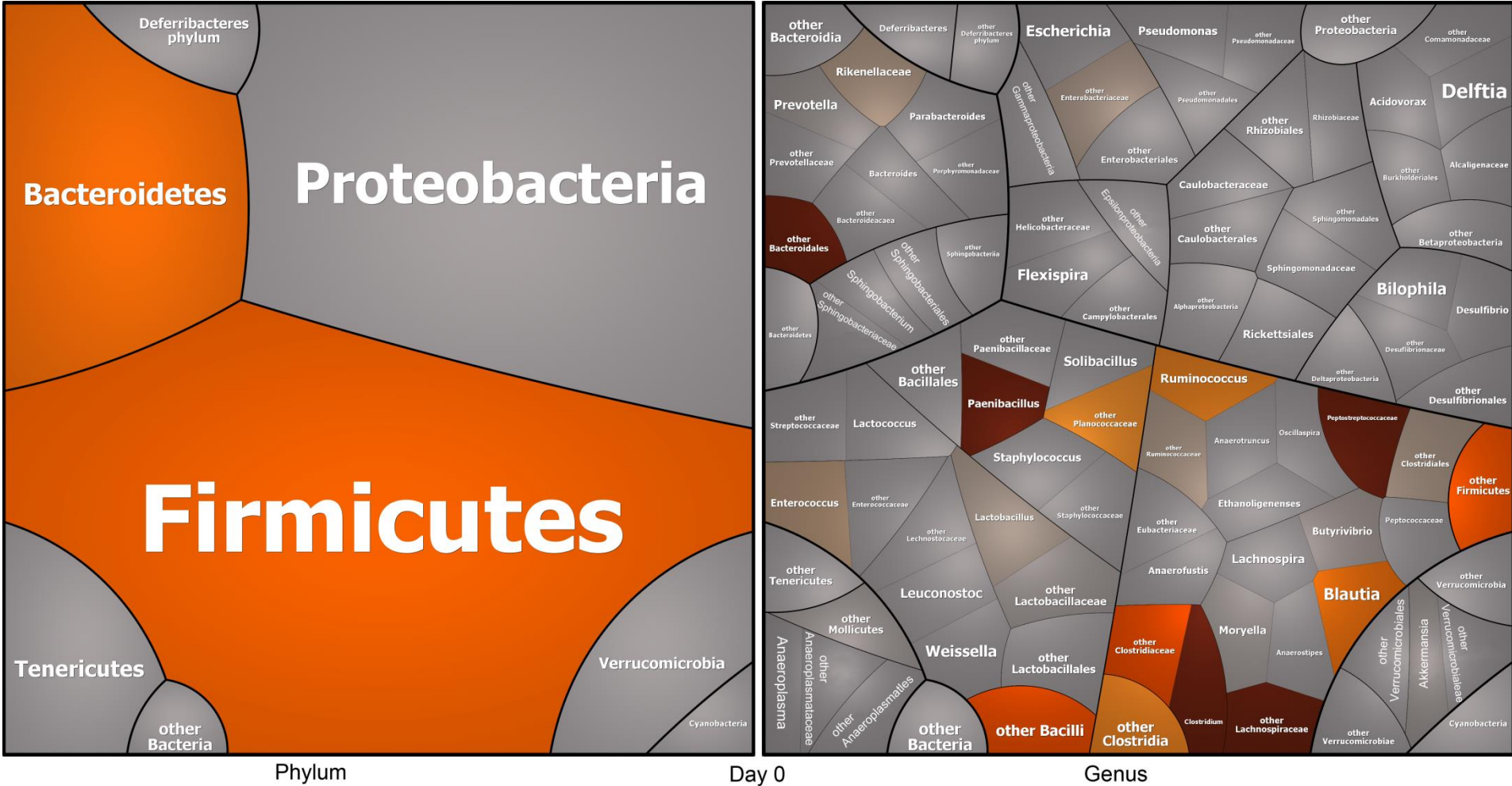
Figure 33: The composition of Endo^hi mice prior to T-cell transfer (n=7; all mice were part of experiment 3). The left picture shows the correlation of the sequencing reads with taxonomy on phylum level. In the right picture, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
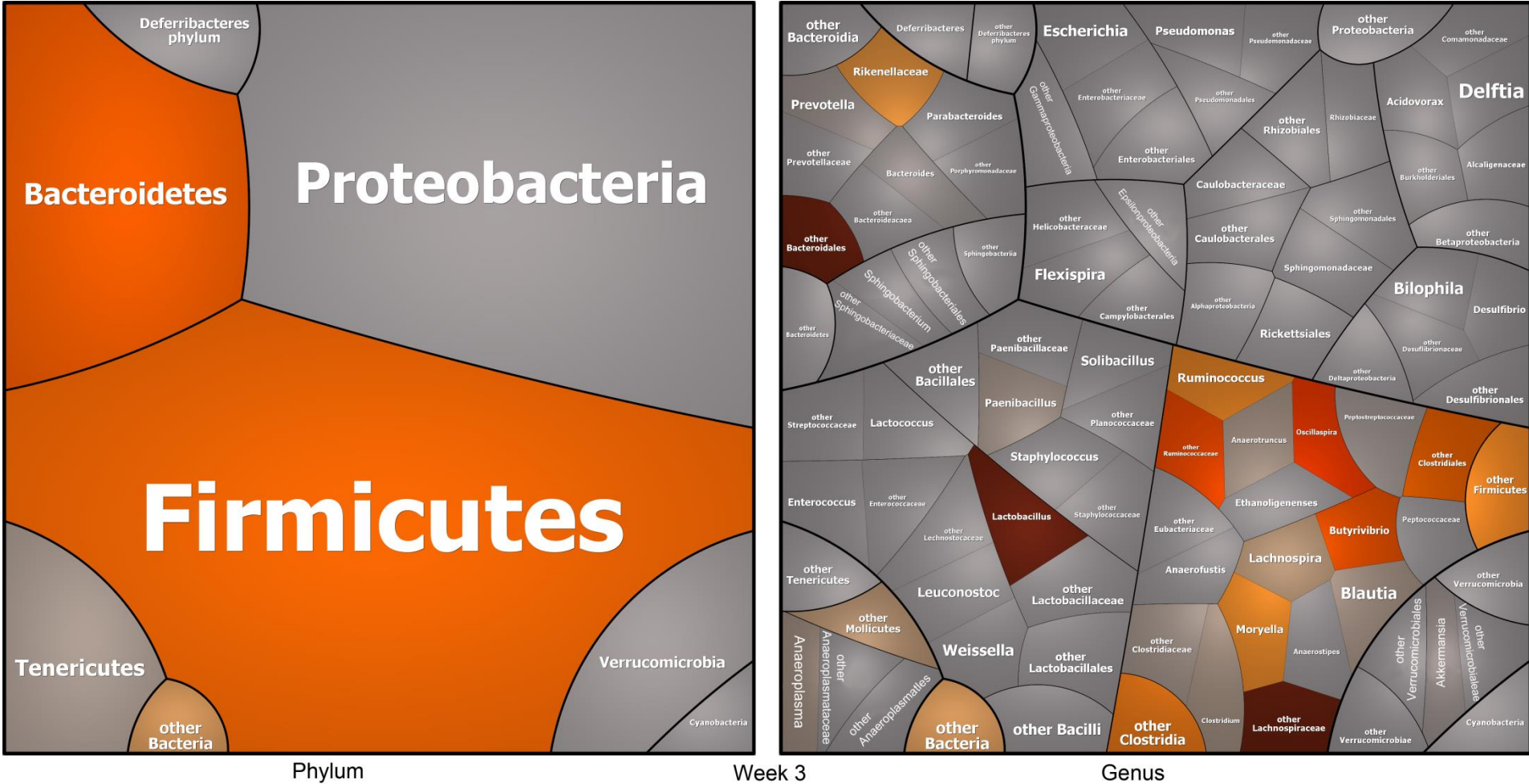
Figure 34: The composition of Endo^{hi} mice 3 weeks after T-cell transfer (n=7; all mice were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right diagram, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
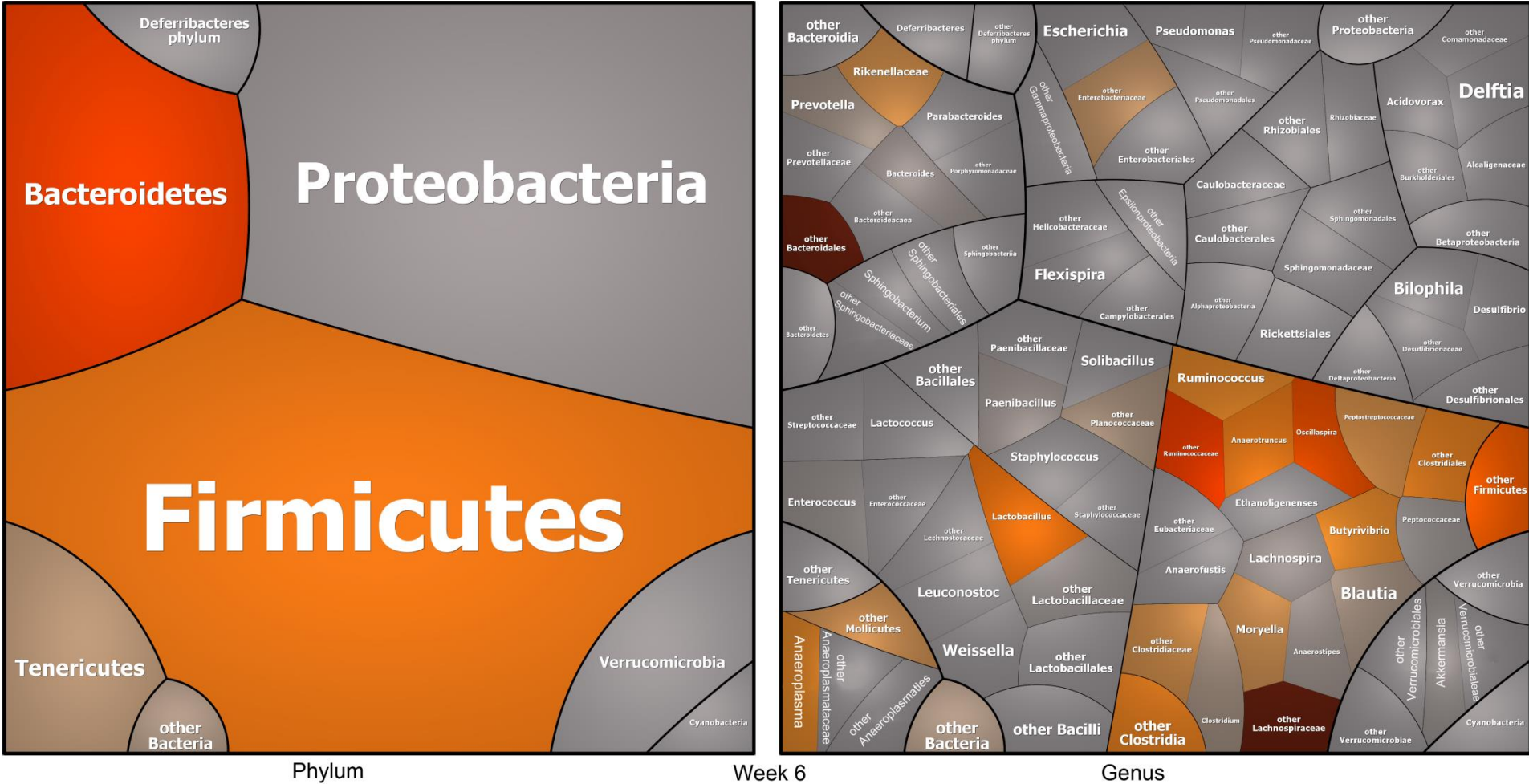
Figure 35: The composition of Endo^{hi} mice 6 weeks after T-cell transfer (n=7; all mice were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

The exact opposite, an increase in Firmicutes and decrease in Bacteroidetes, was observed for mice that stayed healthy after T-cell transfer (Figure 36, Figure 37 and Figure 38). As observed for the Firmicutes in Endo^{hi} mice, the Endo^{lo} mice showed an overall decrease in reads mapping to Bacteroidetes, but the genera of Rikenellacaea increased during the experiment. The phylum of Verrucomicrobia decreased after T-cell transfer and could not be detected 3 weeks later; however, they did recover before the end of the experiment in week 6. A steady increase of Tenericutes and bacteria not associated to a phylum was seen post T-cell transfer. The variety of genera detected in the fecal samples of the Endo^{lo} mice did not differ between starting and end point, but did slightly decrease in between.

Figures 33 - 38 also show that Proteobacteria (visible only on genus level) were only detectable within the microbiota of the Endo^{hi} mice that developed colitis. These results also indicated that the overall variety of detectable genera was greater in Endo^{hi} mice than in Endo^{lo} mice. The proportion of Bacteroidetes was also greater in Endo^{lo} mice than in Endo^{hi} mice, and the fraction of Firmicutes was smaller in Endo^{lo} mice. Tenericutes seemed to increase in both groups after T-cell transfer.
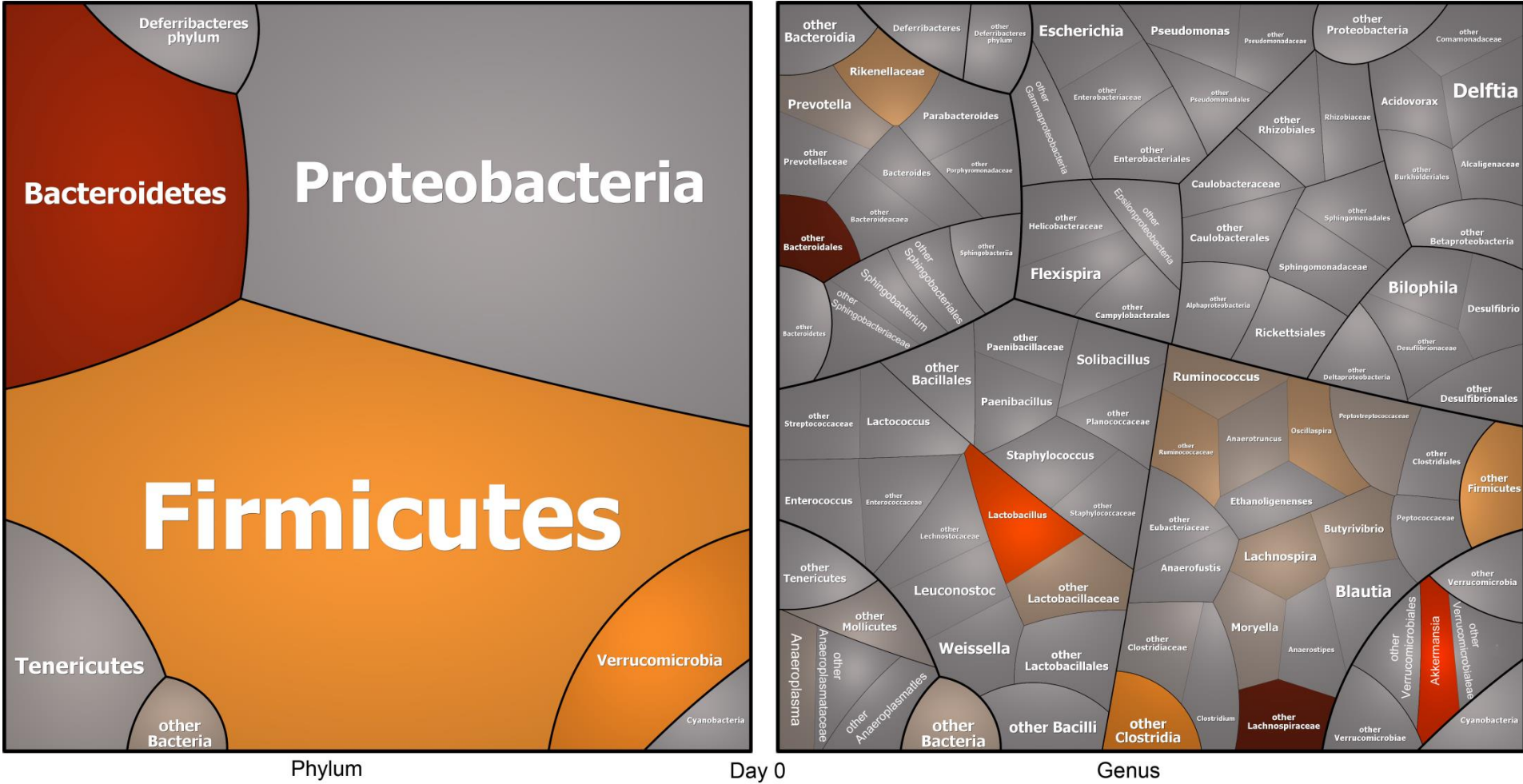
Figure 36: The composition of Endo^{lo} mice prior to T-cell transfer (n=5; all mice were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
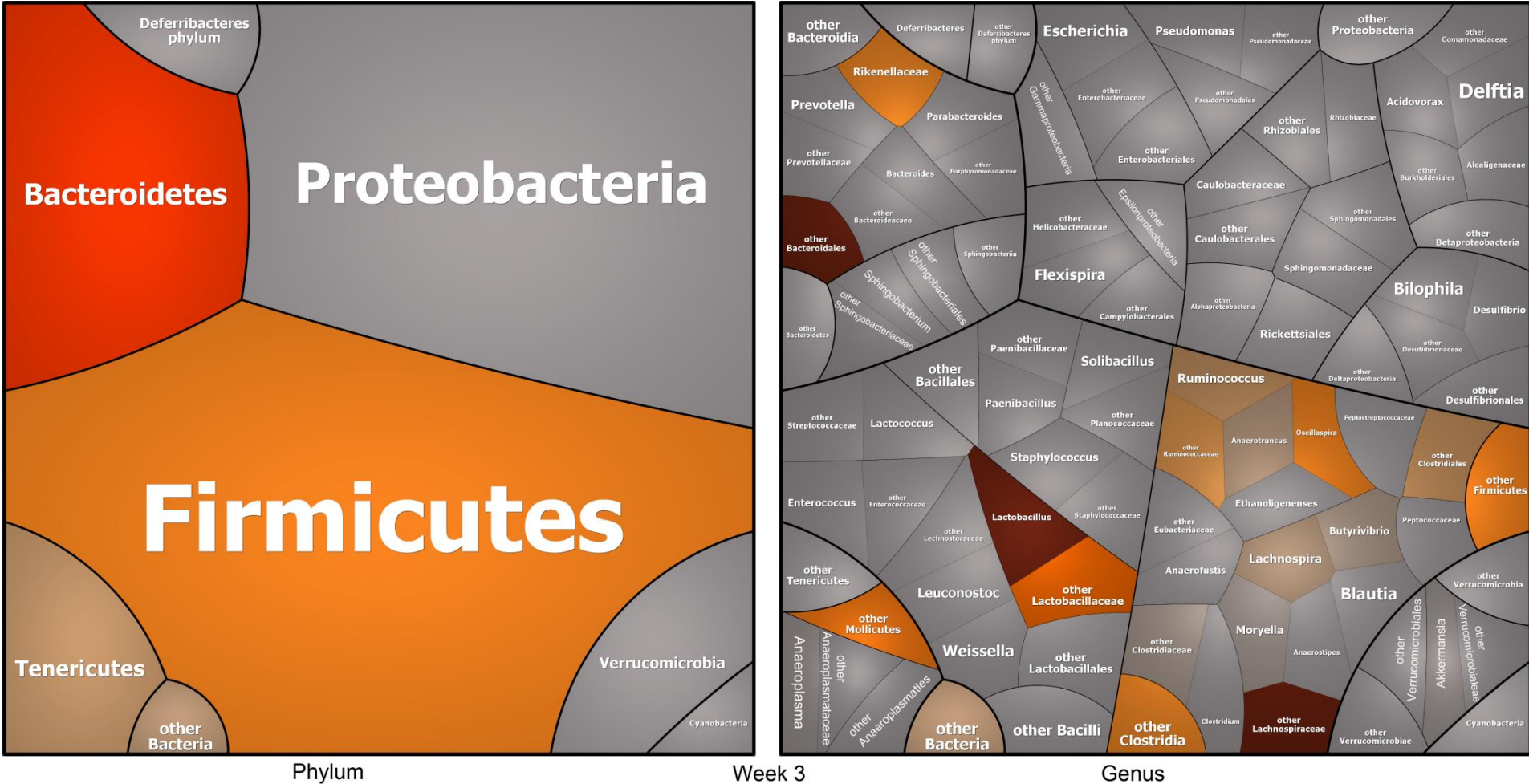
Figure 37: The composition of Endo^lo mice 3 weeks after T-cell transfer (n=5; all mice were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

Figure 38: The composition of Endo^lo mice 3 weeks after T-cell transfer (n=5; all mice were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
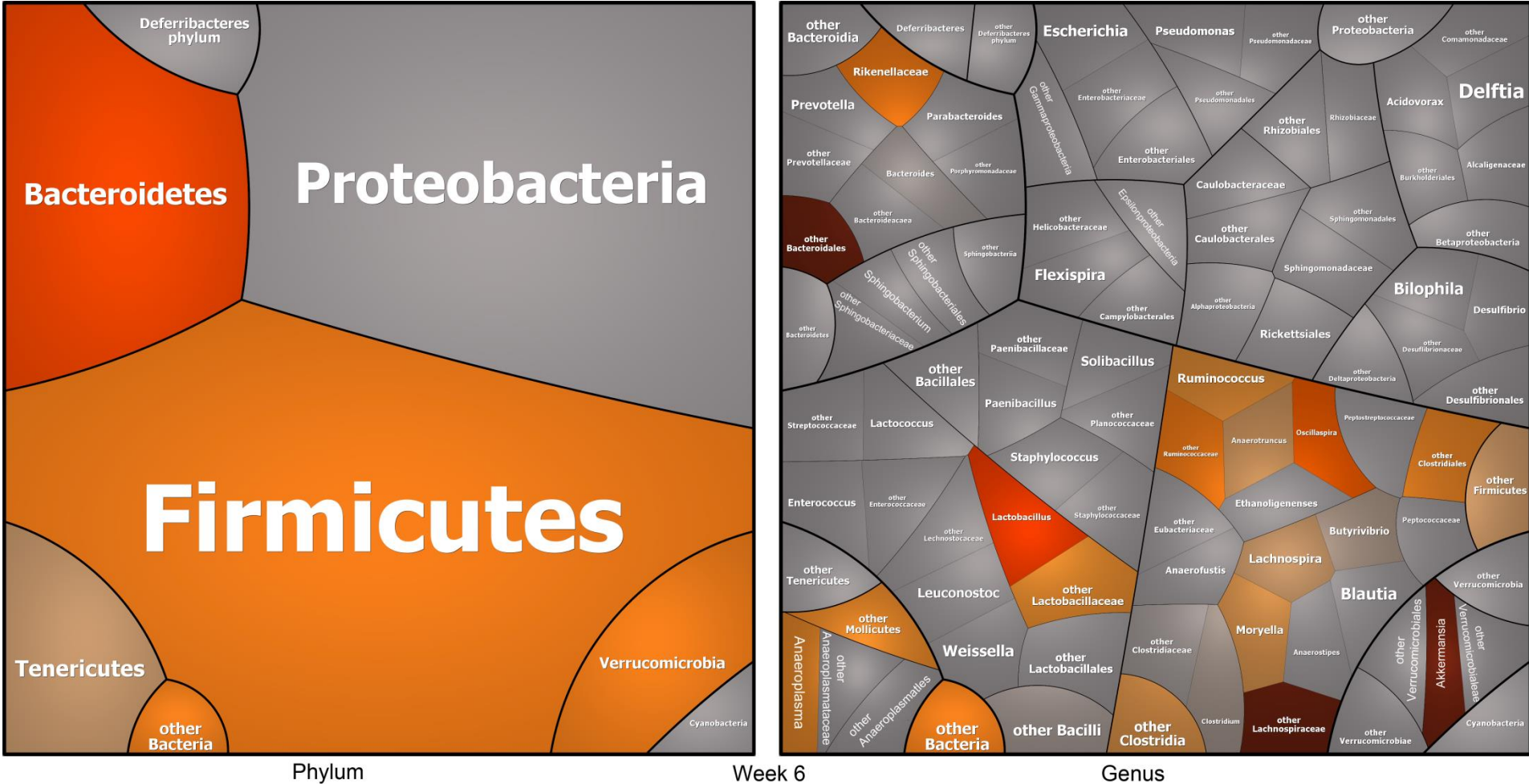
To verify the results on Enterobacteriaceae and *Bacteroides* of the NGS microbiome analysis, FISH was performed on samples collected 6 weeks post T-cell transfer. Representative photomicrographs of the outcome are shown in Figure 39. In contrast to NGS, no Enterobacteriaceae were detected in the sample of the Endo[hi] mouse. *Bacteroides* were only detected by FISH in samples collected from Endo[lo] mice. NGS, however, did not detect *Bacteroides*, only Bacteroidales in samples collected from Endo[hi] and Endo[lo] animals.
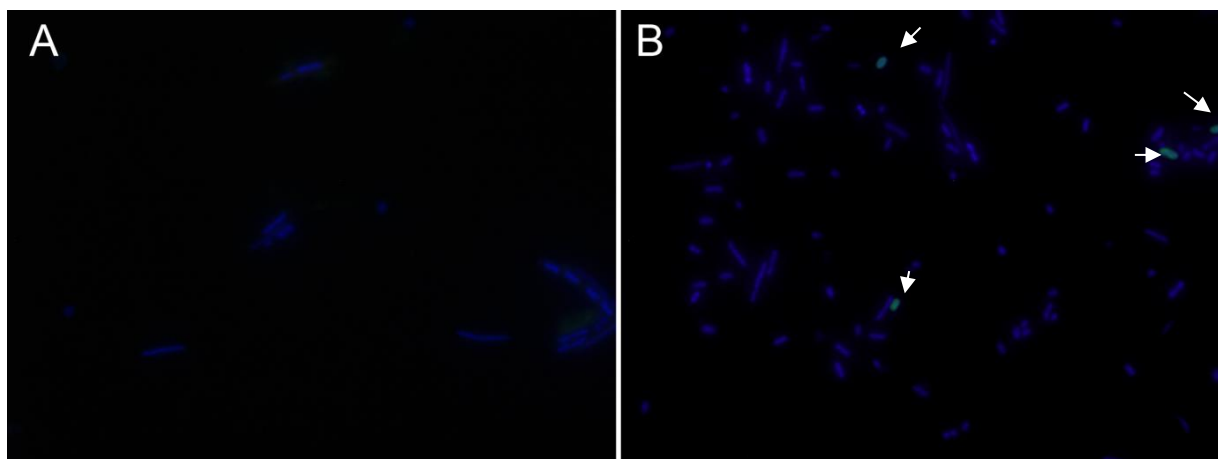

Figure 39: FISH analysis showing the representative composition of bacteria present in the feces of Endo[hi] (A) and Endo[lo] (B) mice 6 weeks after T-cell transfer. All bacteria present were stained using DAPI while Enterobacteriaceae were stained red with a Cy3-labeled probe and *Bacteroides* were colored green by an Alexa Fluor 488 linked probe. No Enterobacteriaceae could be detected in either sample. *Bacteroides* were present in the Endo[lo] sample, indicated by white arrows.

3.4.4 Feeding bacteria influences the intestinal microbiome

The LPS of gram-negative bacteria is recognized by the mammalian immune system. Depending on the structure of the Lipid A of LPS, different biological activity has been observed [69]. Some bacteria therefore seem to be immune stimulating whereas others appear to inhibit immune reaction[63,82]. This, combined with our results showing that mice with a high endotoxic microbiome developed colitis and those harboring a low endotoxic microbiome didn't, led us to the hypothesis that, changing the endotoxicity of the intestinal microbiome prior to T-cell transfer by feeding of high or low endotoxic bacteria could influence colitis development. Prof. R. Darveau from the University of Washington in Seattle kindly provided us with a high endotoxic *E. coli* strain (*E. coli* JM83; *E. coli*$_{WT}$) and a low endotoxic mutant of this strain (*E. coli* JM83 ΔhtrBhtrB$_{pg}$; *E. coli* $_{MUT}$). The difference between the two bacterial strains lies in the structure of their LPS; the wild-type strain Lipid A contains a Laurate molecule. In the Lipid A molecule of the mutant strain, this Laurate is substituted by a palmitic acid, changing the endotoxicity of the LPS molecule from high to low.

To test our hypothesis, Endo[lo] and Endo[hi] animals were fed *E. coli*$_{WT}$ or *E. coli*$_{MUT}$ prior to T-cell transfer. Antibiotics were given to the animals before treatment with bacteria to weaken colonization resistance of the initial microbiome, enabling colonization by the bacterial strains fed to the animals.

The experimental setup to study the influence of the two bacterial strains on the composition of the microbiome and development of colitis is shown in (Figure 40). Six days prior to T-cell transfer, animals were treated with antibiotics for 3 days through their drinking water. Endo^{hi} mice were treated with 20 g/l Streptomycin due to the high amount of gram- bacteria of their initial microbiome. Enod^{lo} mice were treated with 1 g/l Metronidazole targeting anaerobic bacteria such as *Bacteroides*. Three days before T-cell transfer, the mice were treated with 1×10^8 bacteria i.g. Feeding of bacteria continued over the course of the experiment via drinking water. The first fecal samples for analysis of the intestinal microbiome were collected immediately before T-cell transfer on day 0. Further fecal samples were collected and frozen at -80 °C in week 3 and at the end of the experiment in week 6.
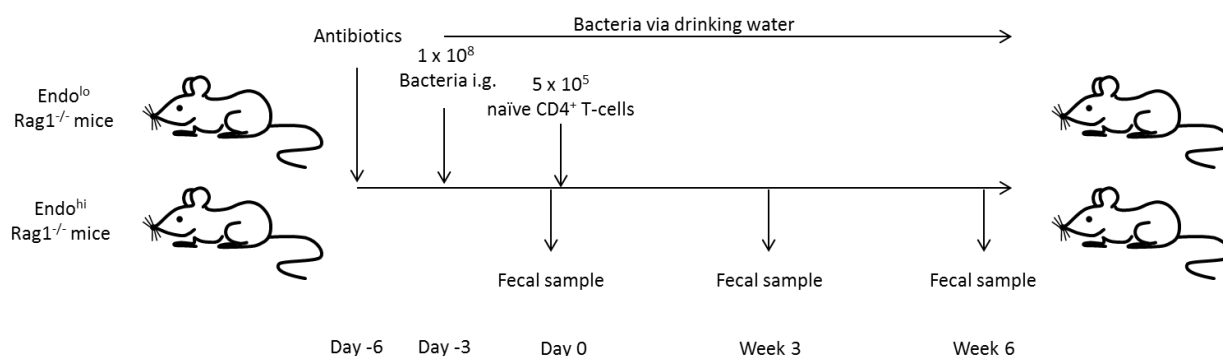

Figure 40: Endo^{lo} and Endo^{hi} mice were treated with metronidazole and streptomycin, respectively, 6 days prior to T-cell transfer. Three days before T-cell transfer, the animals were treated with 1×10^8 bacteria i.g. and via drinking water for the rest of the experiment. On day 0, 5×10^5 naïve CD4+ T-cells were transferred. Fecal samples were collected on day 0 prior to T-cell transfer as well as 3 and 6 weeks after T-cell transfer.

Mice treated with *E. coli*$_{WT}$ developed colitis irrespective of the original microbiota (Endo^{hi} or Endo^{lo}) (Figure 41). Approximately 4 weeks after T-cell transfer the first signs of colitis were visible; the mice first developed diarrhea followed by a display of a relieving posture and weight loss.

In contrast, mice treated with *E. coli*$_{MUT}$ maintained homoeostasis and did not develop colitis within the 6 weeks of the experiment, irrespective of the initial microbiota (Figure 41). The inflammation of the colon was quantified by the histological colitis score and was used to distinguish between healthy and diseased mice.
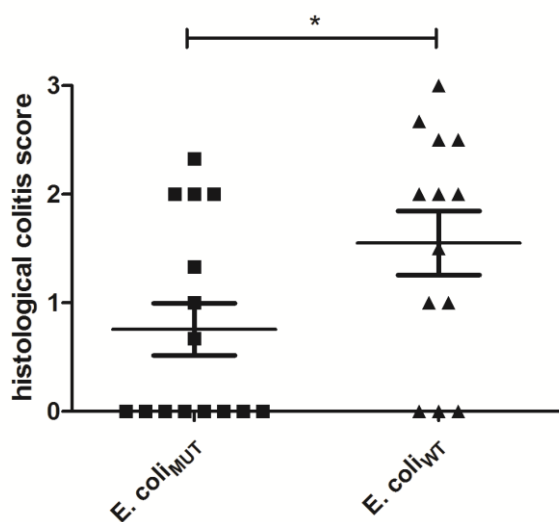
Figure 41: The histological colitis score shows that animals fed with *E. coli*WT develop colitis and animals fed with *E. coli*MUT maintain homoeostasis irrespective of the initial Endo^lo or Endo^hi microbiome. Statistical analysis was done using Student t Test. * p<0.05.

3.4.4.1 Treatment with E. coli_WT induces colitis

Colonizing Endo^hi animals with *E. coli*_WT_ prior to T-cell transfer did not change the health state of the mice at the end of the experiment. These mice developed the same indicators for disease as the non-pre-treated Endo^hi mice.

Visualization of sequencing reads on the phylum level showed no Proteobacteria for non-pre-treated Endo^hi mice, but they were apparent in *E. coli*_WT_ fed animals. The proportion of Proteobacteria was primarily influenced by the reads mapping to the genus *Escherichia*. No major changes of the Firmicutes were evident during the experiment; an increase of Tenericutes was seen between week 3 and 6 and no Bacteroidetes could be detected. The variety of genera detected in samples from Endo^hi mice fed with *E. coli*_WT_ was greater at the end of the experiment than in week 0.
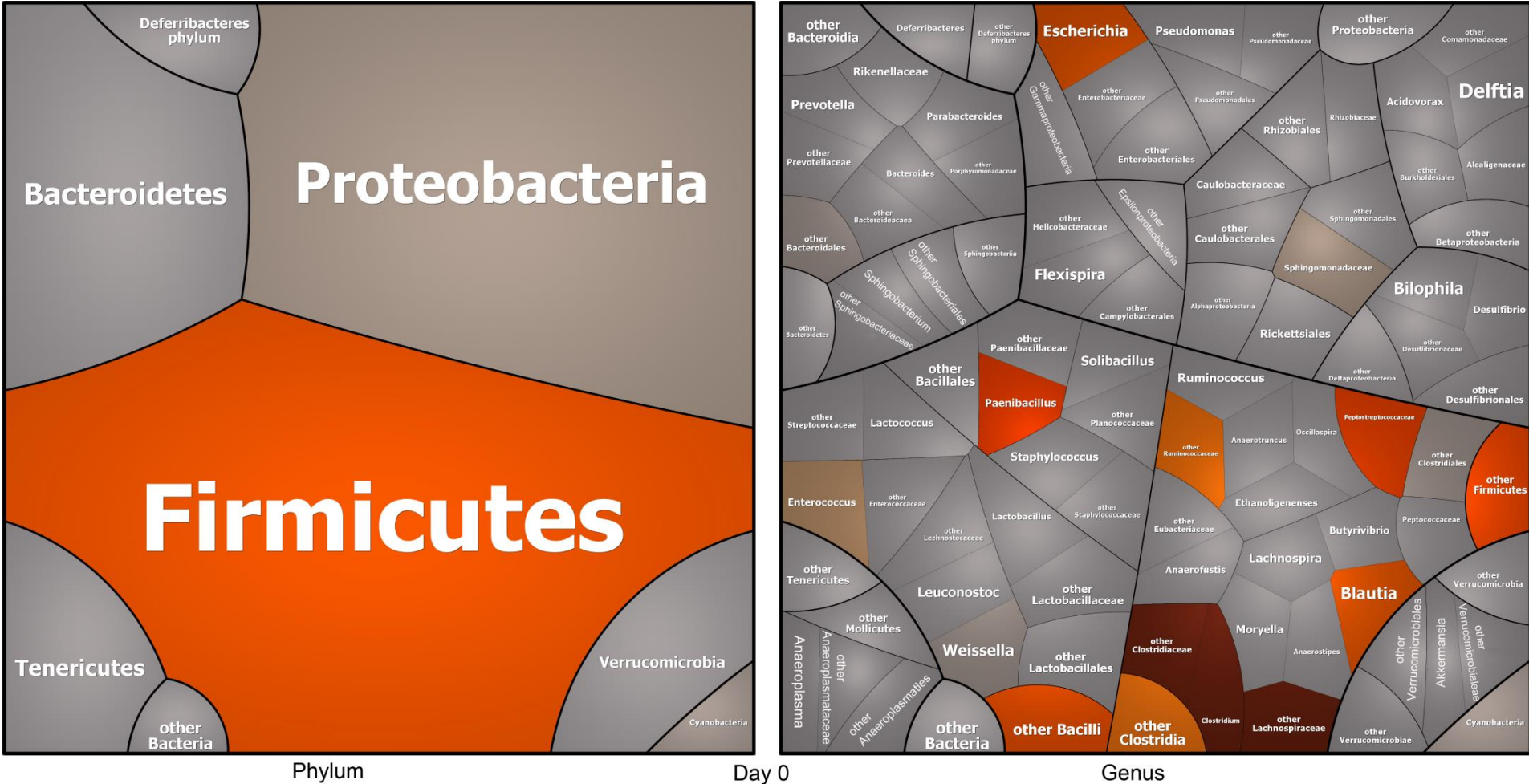
Figure 42: Composition of the intestinal microbiota of *E. coli*WT treated Endohi mice before T-cell transfer (n = 2; animals were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

Figure 43: Composition of the intestinal microbiota of *E. coli*WT treated Endohi mice, 3 weeks after T-cell transfer (n = 2; animals were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
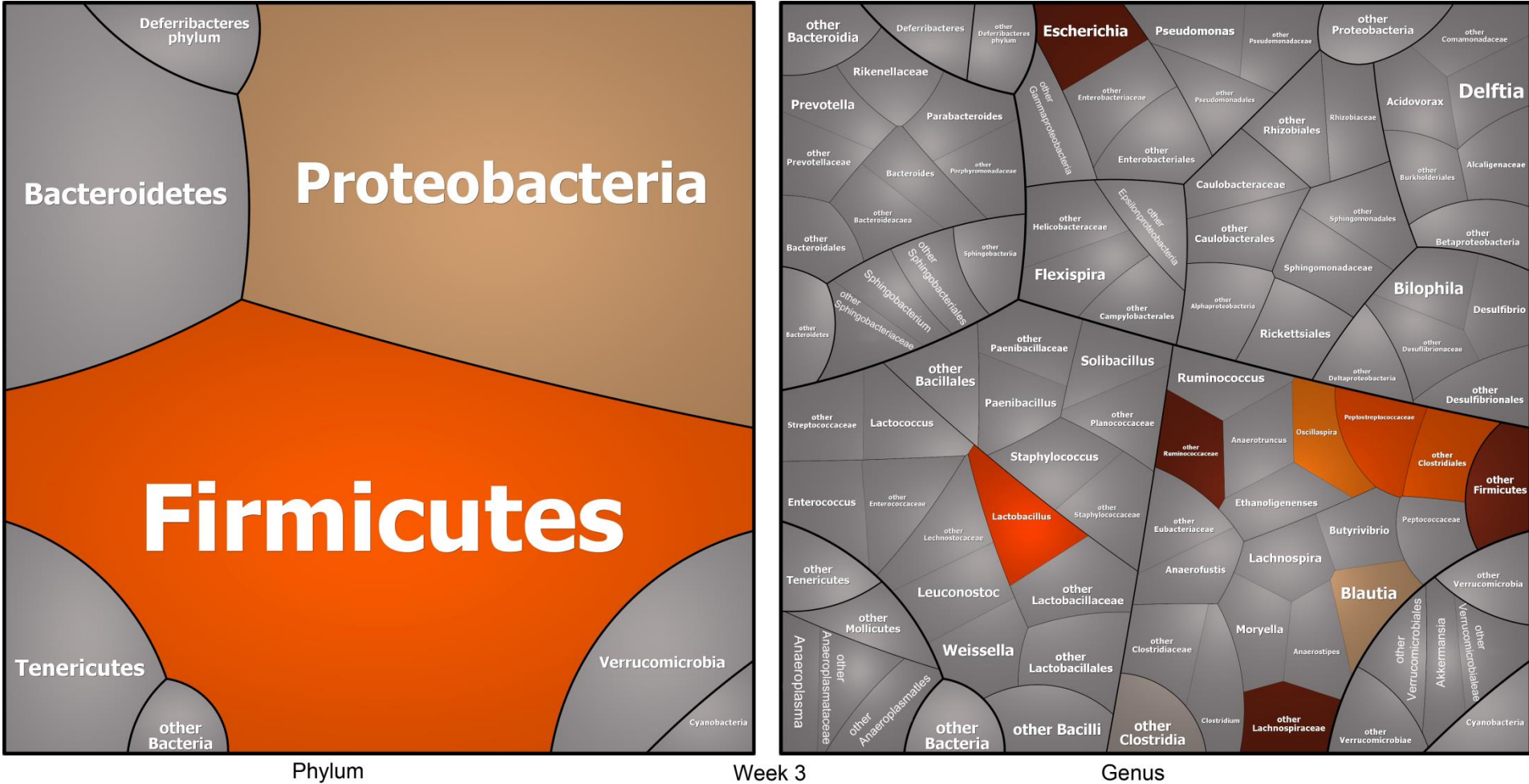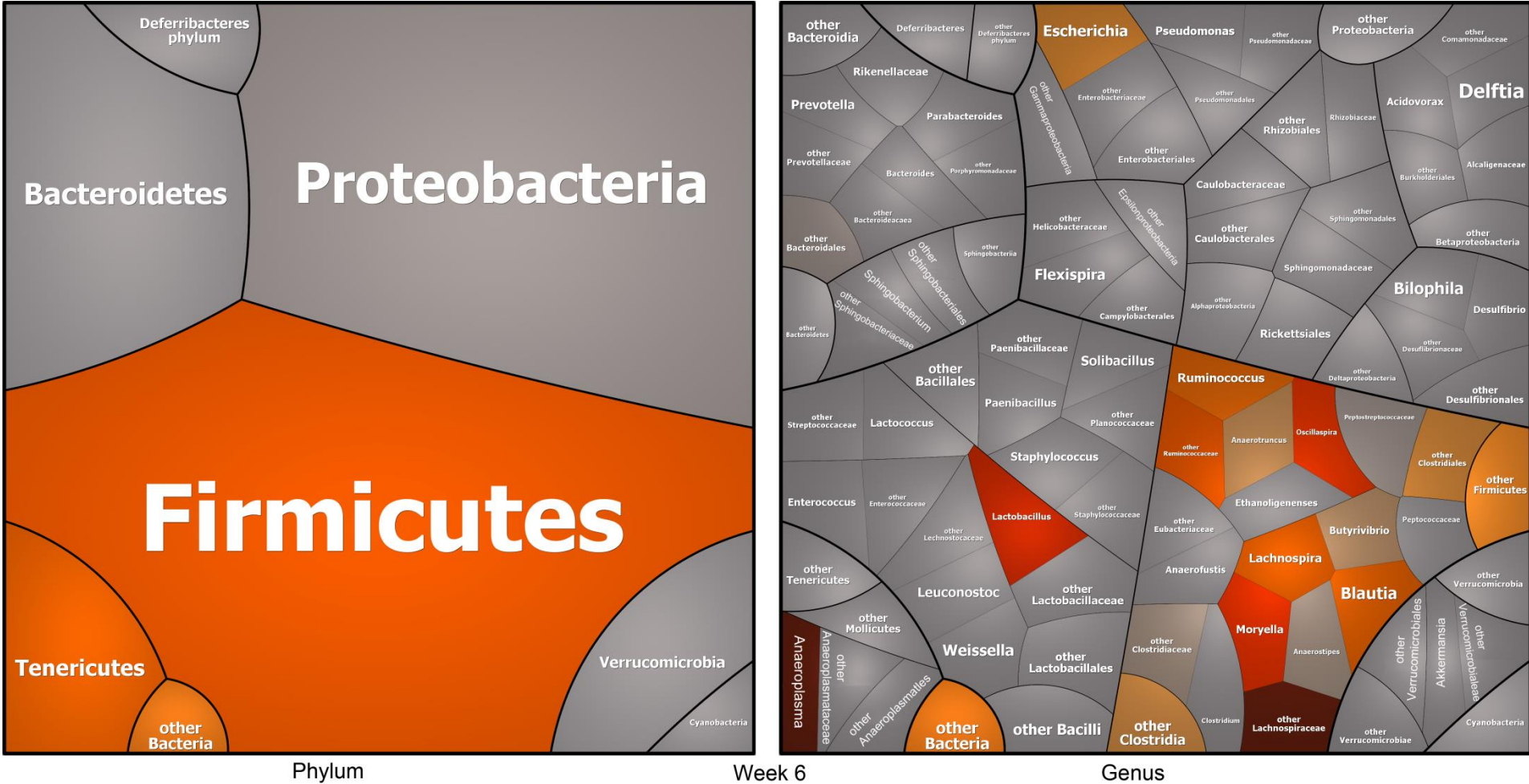
Figure 44: Composition of the intestinal microbiota of *E. coli*WT treated Endo^{hi} mice 6 weeks after T-cell transfer (n = 2; animals were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

Colonizing Endo^{lo} mice with $E.\ coli_{WT}$ prior to T-cell reconstitution resulted in the development of colitis. Figures 45 - 47 show that although $E.\ coli_{WT}$ was fed prior to T-cell transfer in week 0, and during the complete timespan of the experiment, Proteobacteria remained below the limit of detection and was only visible on the family level (0 Appendix) or lower in week 6. Changes of the microbiota between week 0 and week 6 appeared to be similar in non- and $E.\ coli_{WT}$ pre-treated Endo^{lo} animals on the phylum level with one exception: the increase of bacteria that could not be associated to a phylum seen in non-pre-treated mice did not appear in mice fed with $E.\ coli_{MUT}$.

Resolving the sequencing data on the genus level indicated that generally the same genera could be detected in the microbiome of $E.\ coli_{WT}$ Endo^{lo} animals and Endo^{lo} animals that only had T-cells transferred. Although some fluctuation is seen in the relative amounts of the different genera, the genera themselves did not change substantially.

Figure 45: Composition of the intestinal microbiota of *E. coli*_{WT} treated Endo^{lo} mice before T-cell transfer (n = 5; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
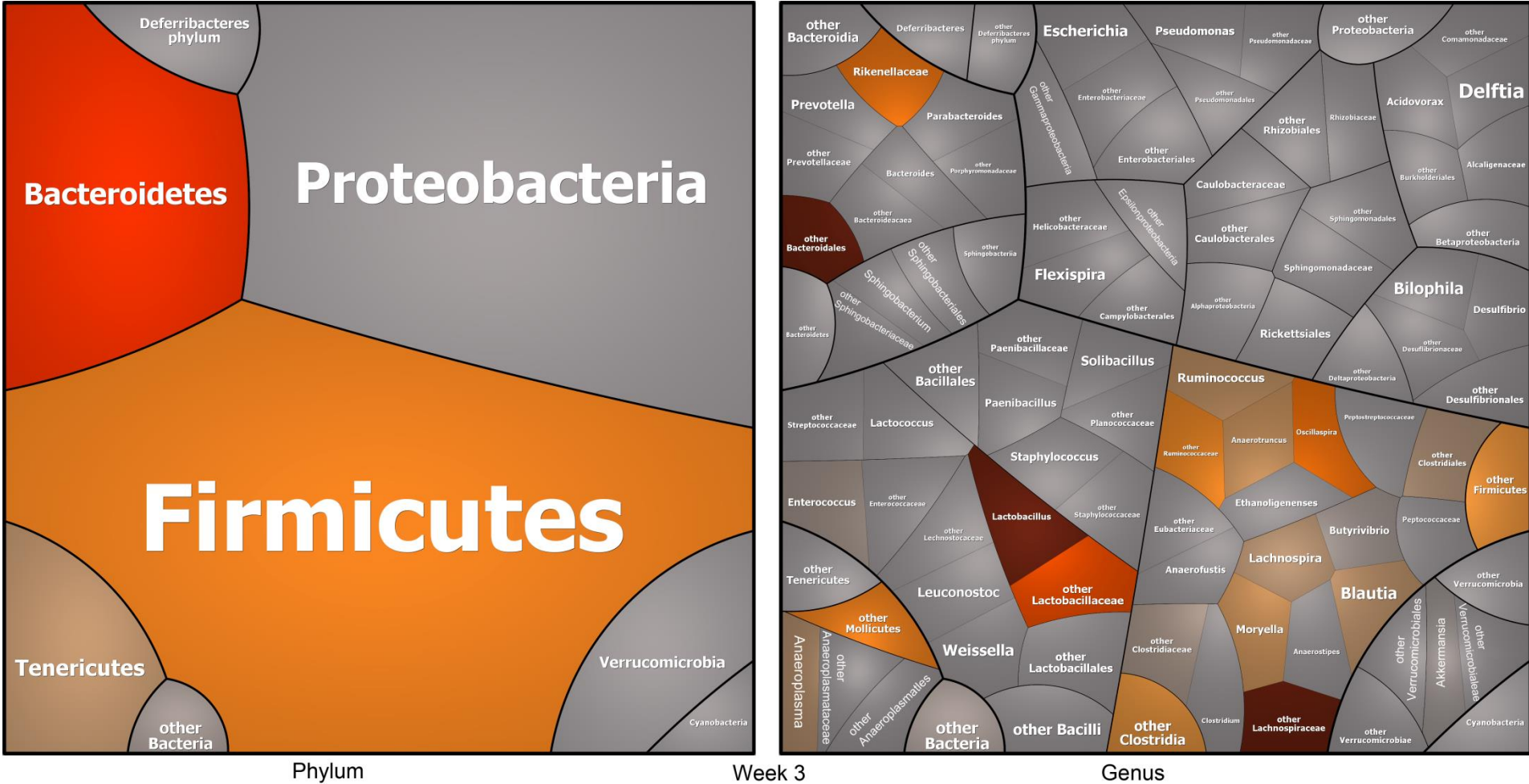
Figure 46: Composition of the intestinal microbiota of *E. coli*WT treated Endo^lo mice, 3 weeks after T-cell transfer (n = 5; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
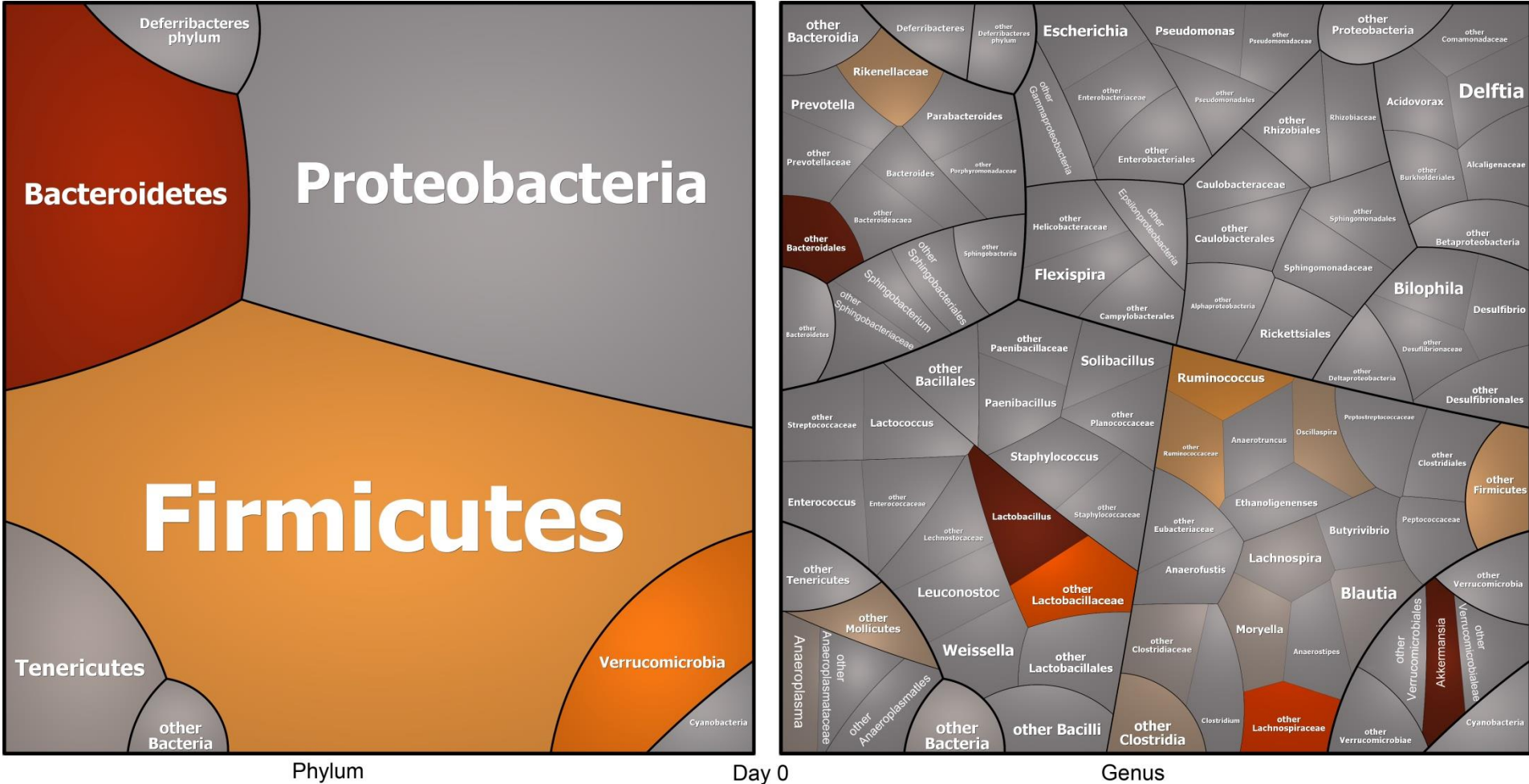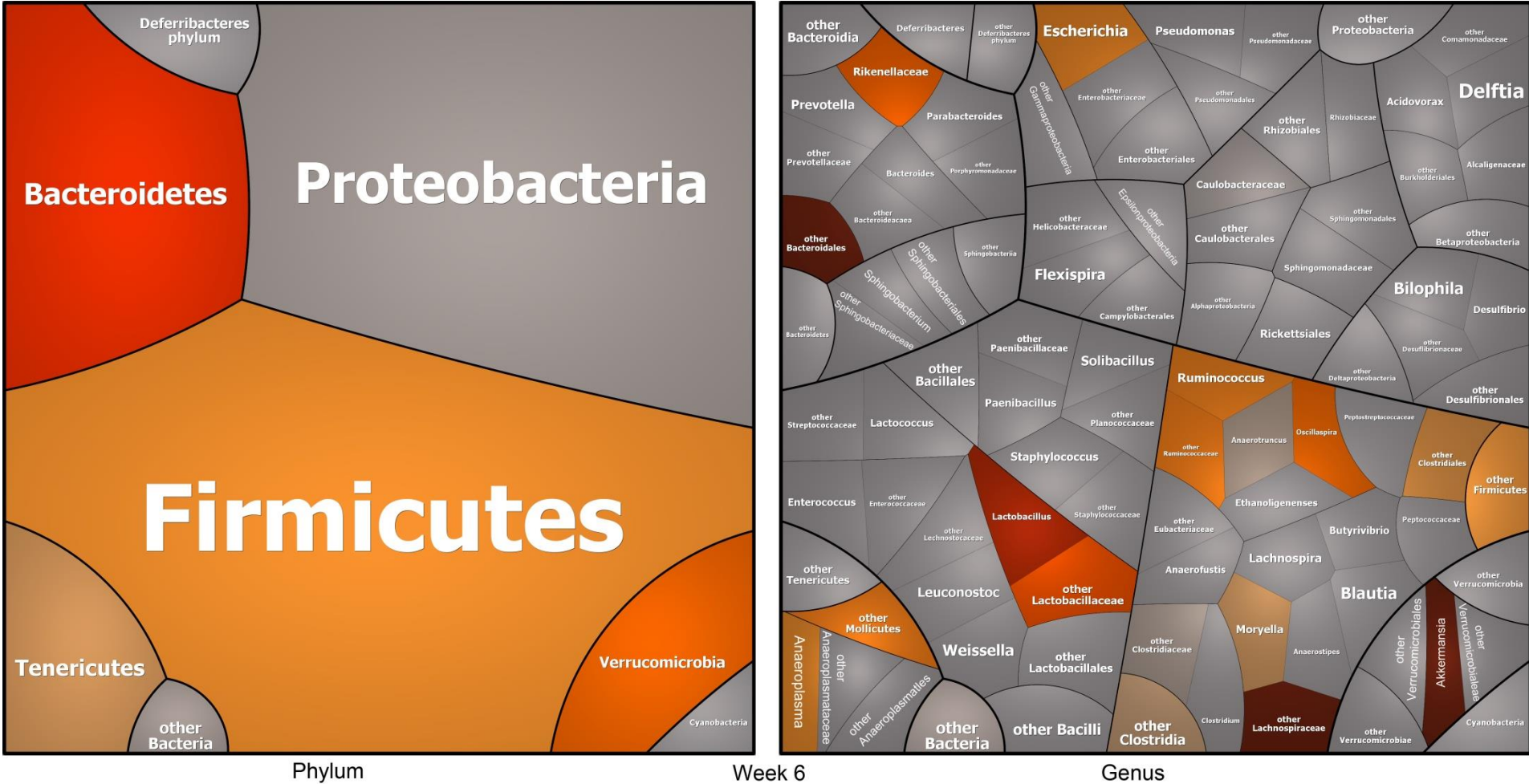
Figure 47: Composition of the intestinal microbiota of *E. coli*WT treated Endo^lo mice, 6 weeks after T-cell transfer (n = 5; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

FISH analysis of feces collected in week 6 from *E. coli*_{WT} treated Endo^{hi} and Endo^{lo} mice confirmed the findings of the sequencing experiments. *Bacteroides* were only found within the samples of *E. coli*_{WT} treated Endo^{lo} mice and *Enterobacteriaceae* were sporadically found in the Endo^{hi} animals treated with *E. coli*_{WT} but not in the *E. coli*_{WT} treated Endo^{lo} mice (Figure 48).
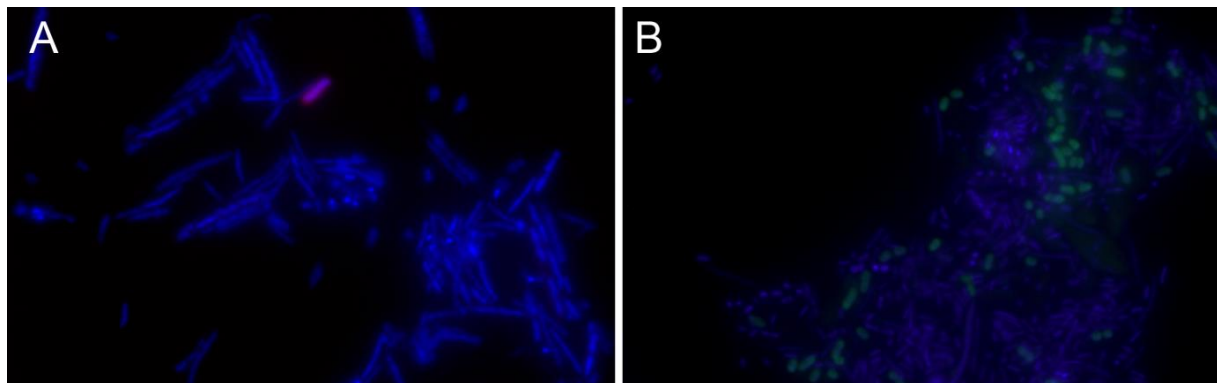

Figure 48: FISH of representative samples collected from *E. coli*_{WT} treated Endo^{hi} (A) and Endo^{lo} (B) mice in week 6. DAPI staining visualizes all bacteria, Enterobacteriaceae are labeled by Cy3 (red) and *Bacteroides* by Alexa Fluor 488 (green). Enterobacteriaceae were only detected in samples from Endo^{hi} animals and Bacteroides only in samples collected from the Endo^{lo} mice.

3.4.4.2   Treatment of mice with E. coli JM83 mutant strain prevents colitis

Feeding of *E. coli* JM83 ΔhtrBhtrB_{pg} (*E. coli*_{MUT}) to Endo^{hi} animals prior to T-cell transfer resulted in maintenance of homoeostasis. On the day of T-cell transfer, no Bacteroidetes were detected, but the Proteobacteria are evident, with most reads mapping to Escherichia. During the course of the experiment, the Proteobacteria fell below the detection limit and Bacteroidetes and Tenericutes were detected for the first time.

The bacterial composition and the changes of the intestinal microbiome of the *E. coli*_{MUT} treated Endo^{hi} mice (Figures 49 - 51) were similar to the Endo^{hi} mice without bacterial pre-treatment (Figures 33 - 35). The main difference between these two groups was that Proteobacteria were only detectable on the day of T-cell transfer in the *E. coli*_{MUT} pre-treated mice. The bacterial diversity on genus level, however, was lower than the only T-cell treated Endo^{hi} mice. The variety of identified genera was greater in the microbiome after disease development.

Figure 49: Composition of the intestinal microbiota of *E. coli*MUT treated Endo[hi] mice, before T-cell transfer (n = 2; animals were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
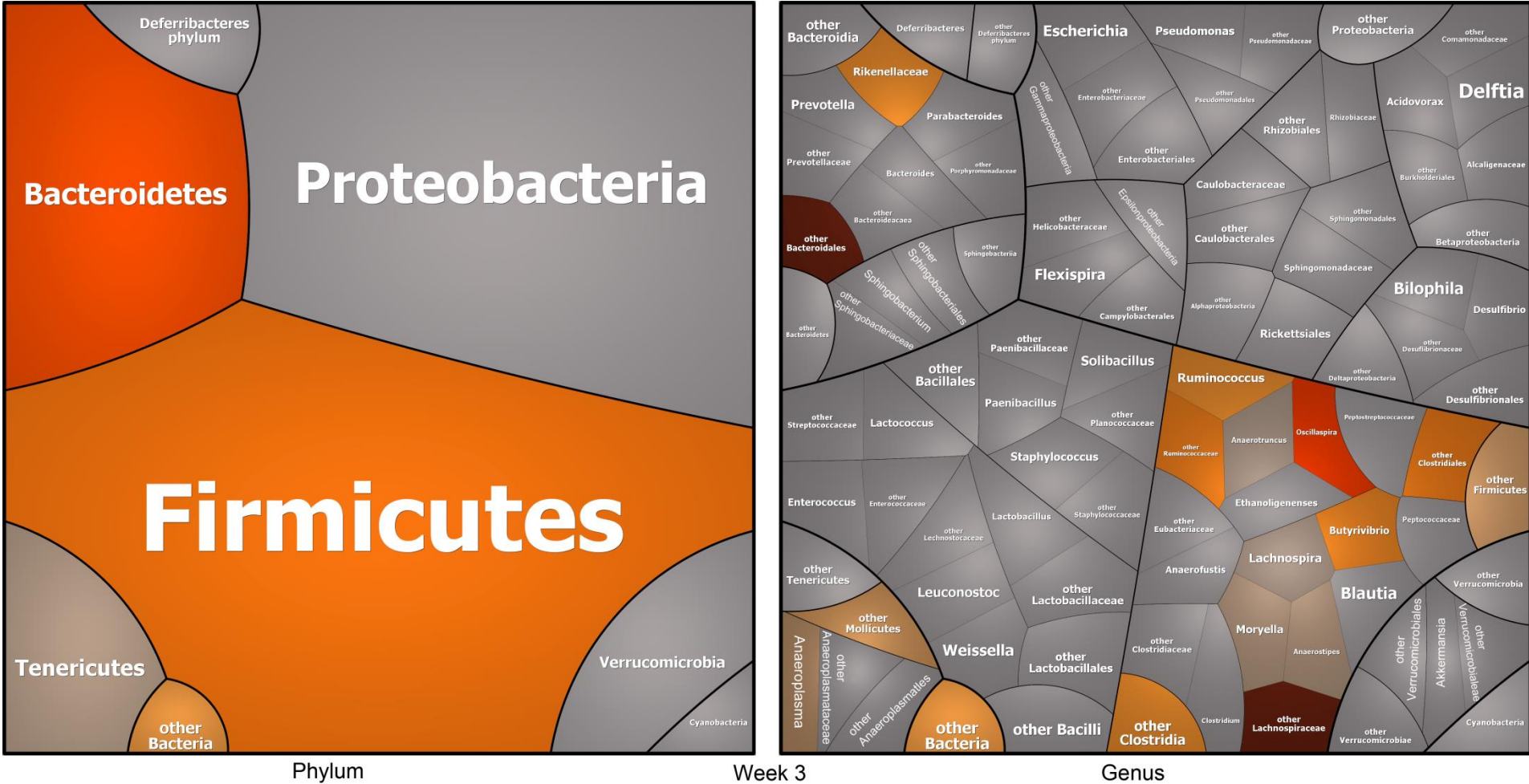
Figure 50: Composition of the intestinal microbiota of *E. coli*MUT treated Endohi mice, 3 weeks after T-cell transfer (n = 2; animals were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
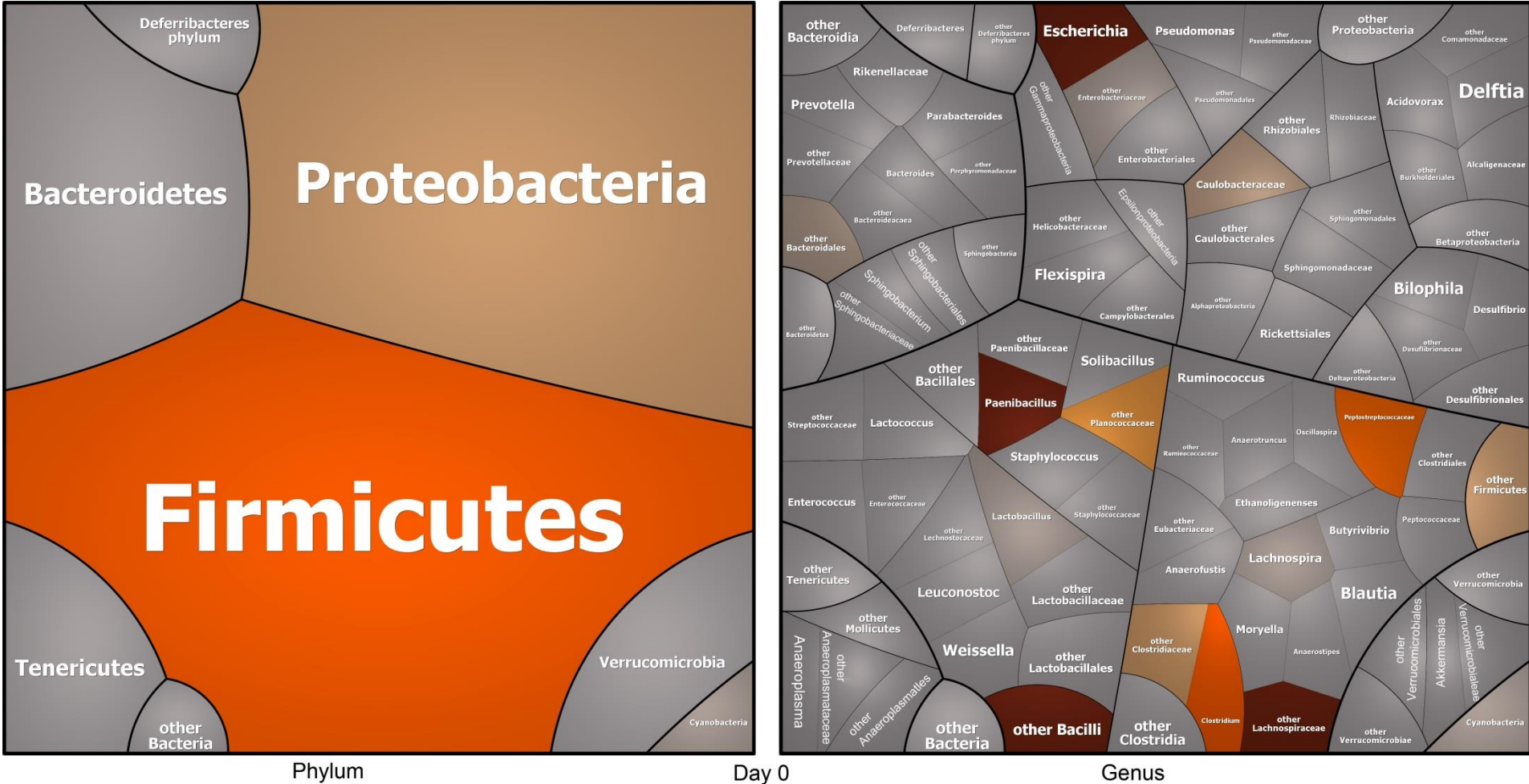
Figure 51: Composition of the intestinal microbiota of *E. coli*MUT treated Endo^hi mice, 6 weeks after T-cell transfer (n = 2; animals were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

Figures 52 - 54 show that despite colonizing Endo^{lo} mice with *E. coli*$_{MUT}$ prior to T-cell transfer and via the drinking water during the entire experiment, the amount of Proteobacteria remained below the limit of detection. On the phylum level the only obvious changes of microbiota composition was the decrease of Bacteroidetes between weeks 0 and 3, followed by their increase and the steady increase of Tenericutes. The abundances of the reads mapping to the different genera differed slightly over the course of the experiment, however, the genera with which the reads were associated to did not change substantially.

Figure 52: Composition of the intestinal microbiota of *E. coli*MUT treated Endo^lo mice, before T-cell transfer (n = 5; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
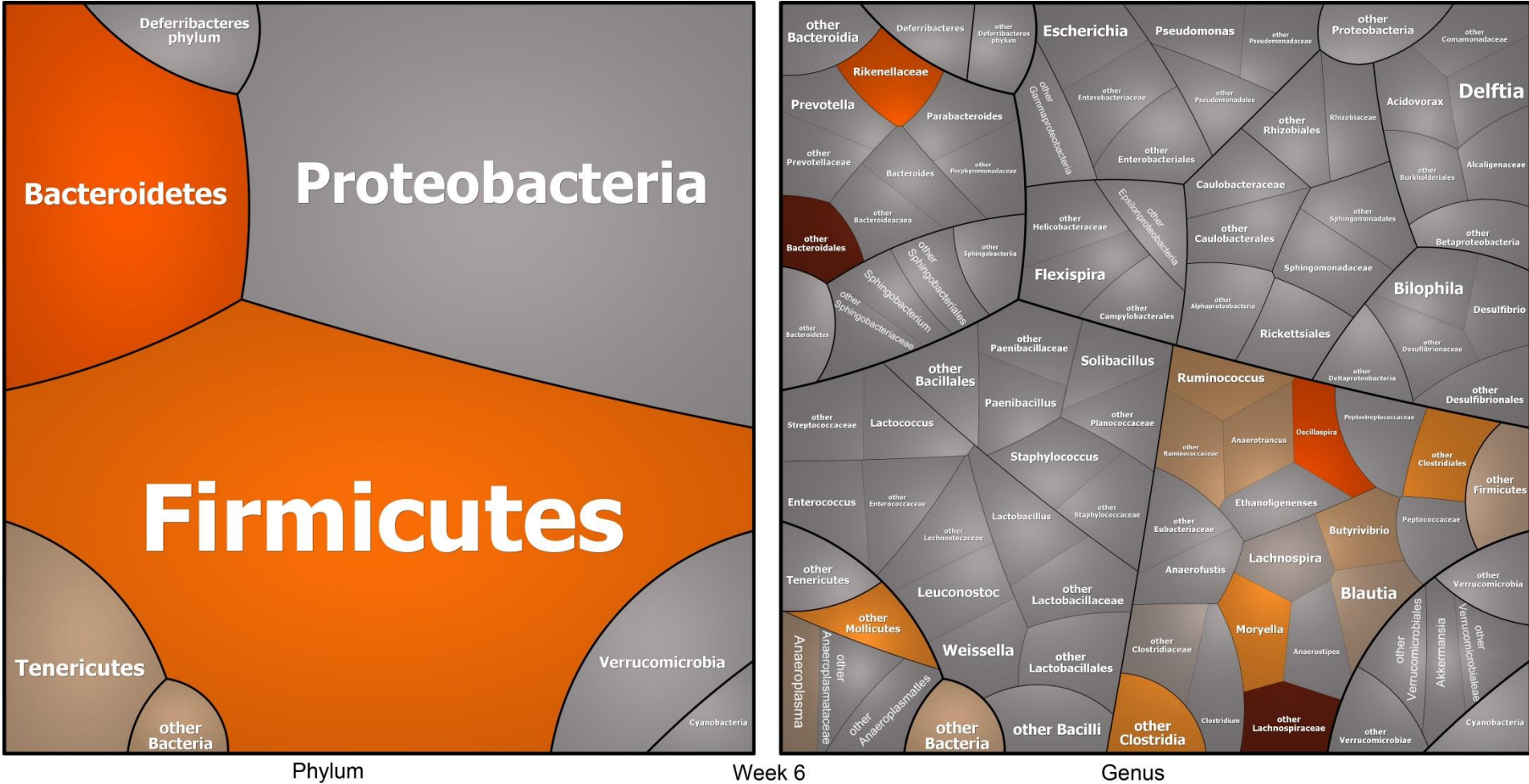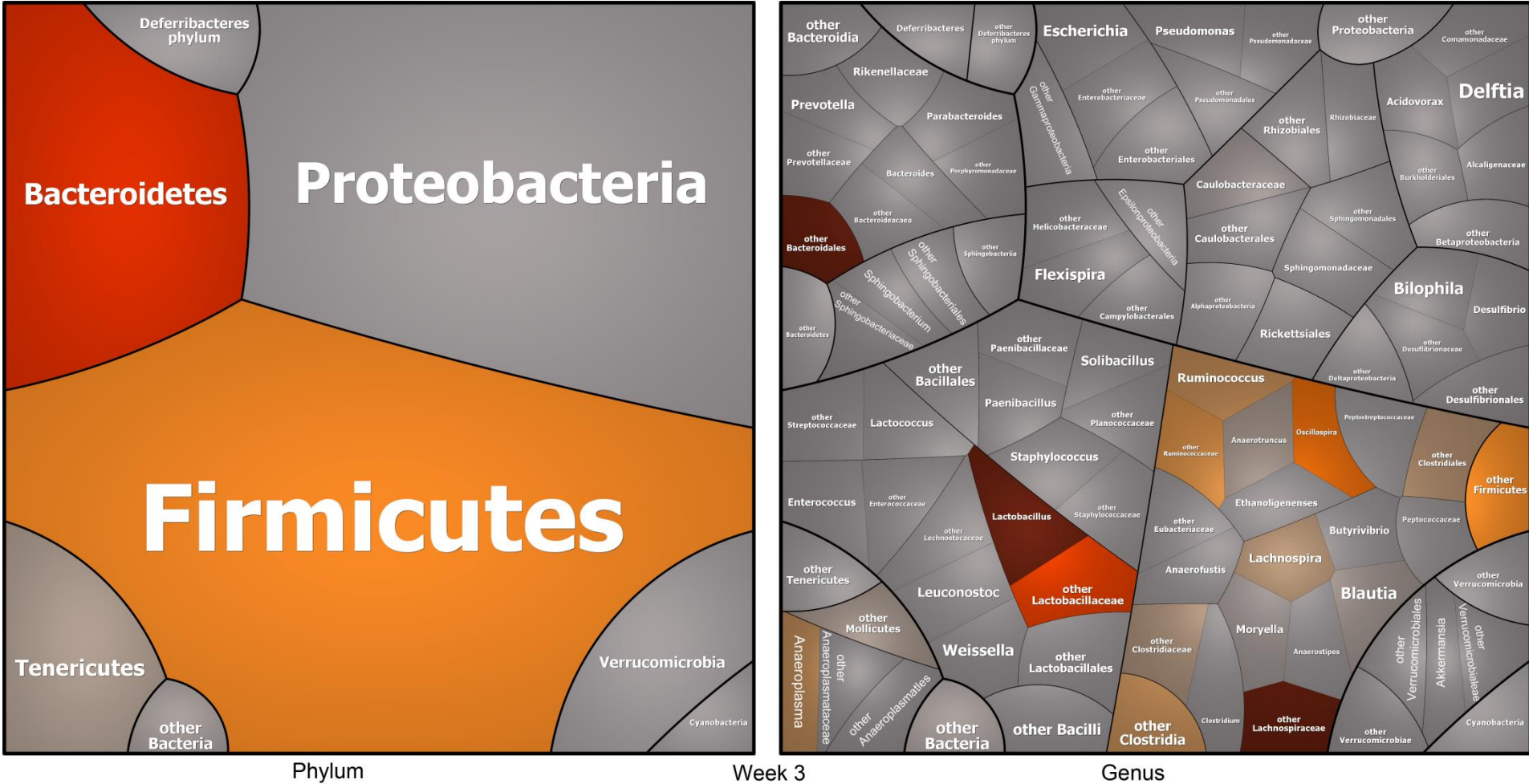
Figure 53: Composition of the intestinal microbiota of *E. coli*MUT treated Endo^lo mice, 3 weeks after T-cell transfer (n = 5; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
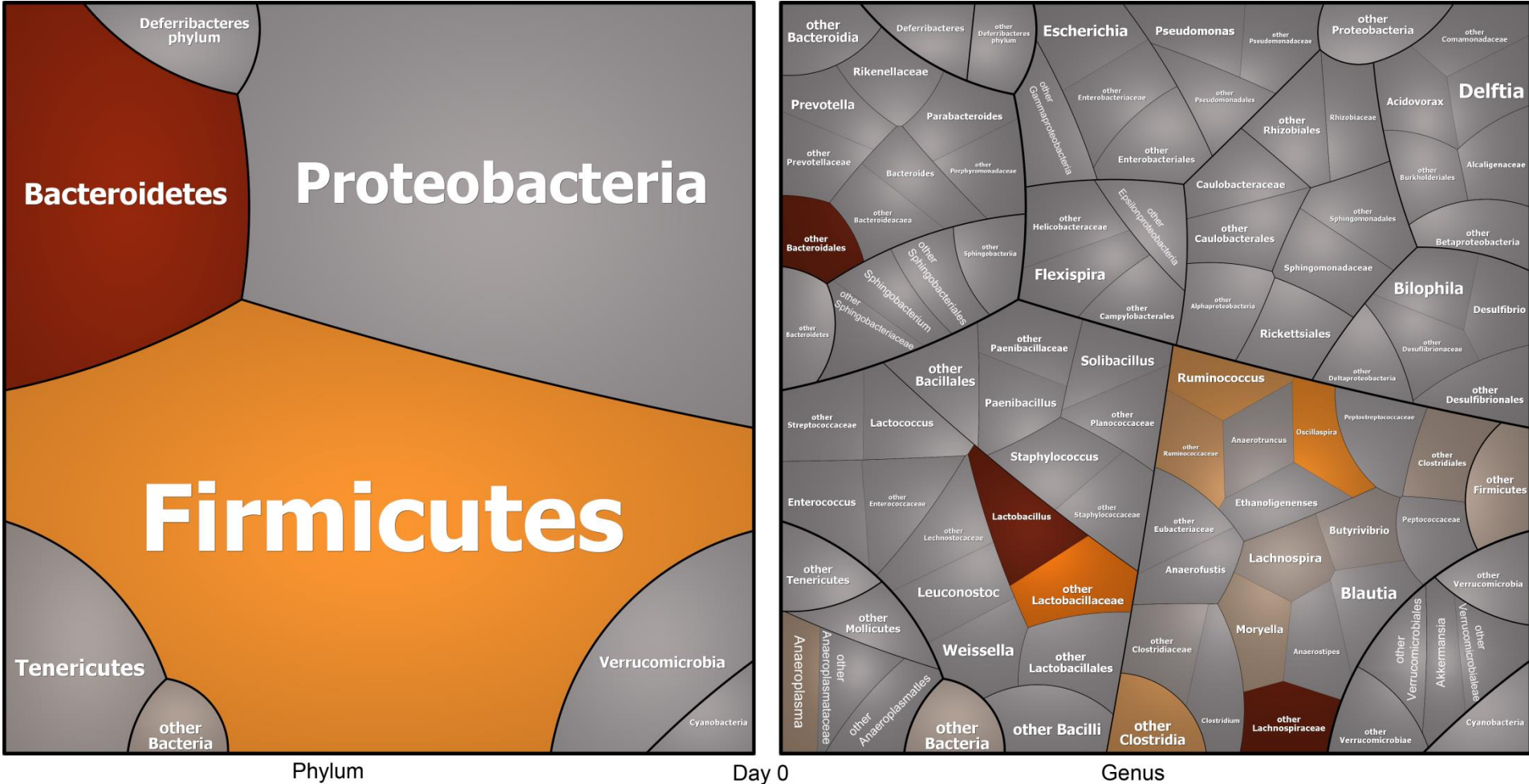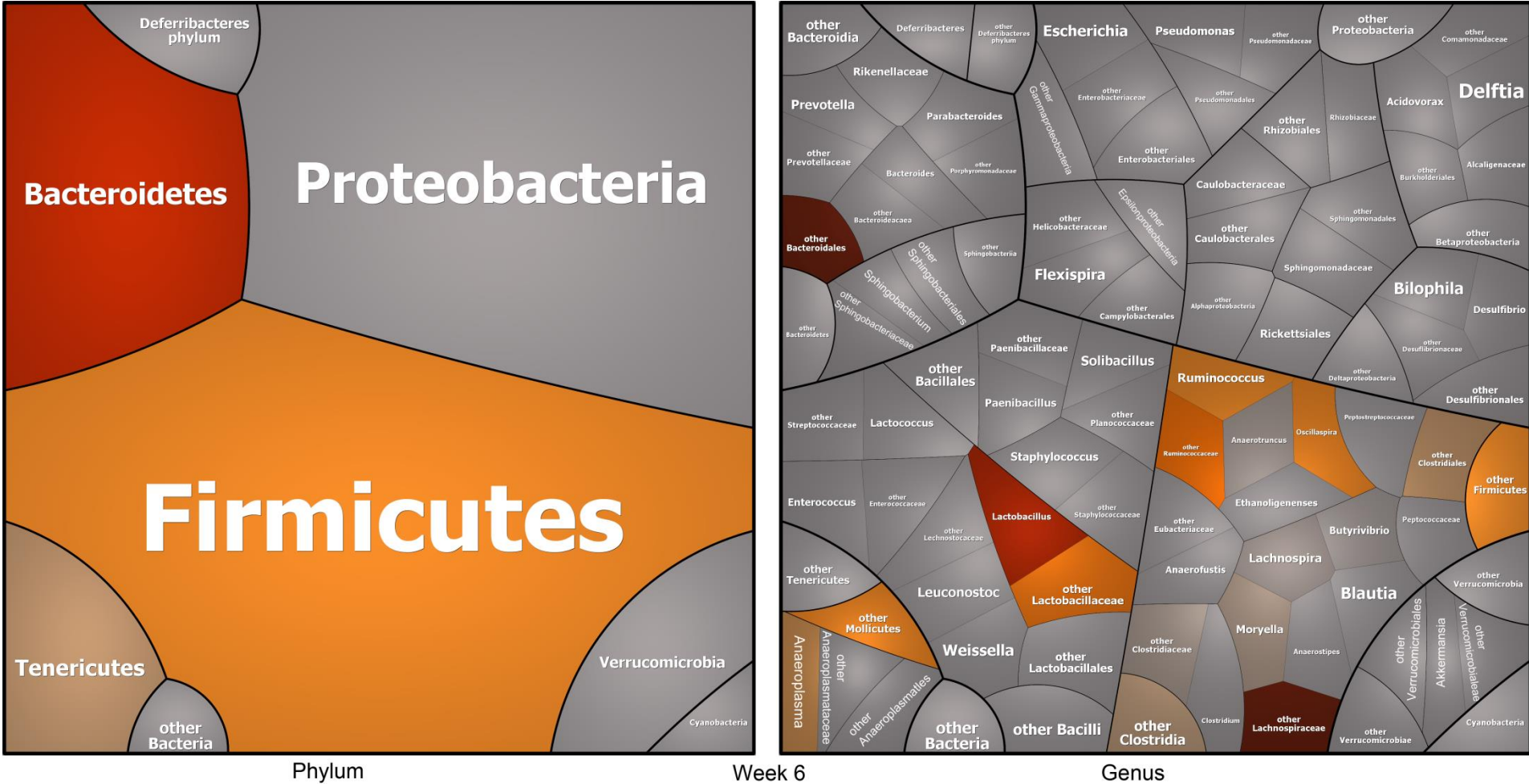
Figure 54: Composition of the intestinal microbiota of *E. coli*MUT treated Endolo mice, 6 weeks after T-cell transfer (n = 5; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

FISH of samples collected in week 6 from *E. coli*$_{MUT}$ treated Endo^{hi} and Endo^{lo} animals is shown in (Figure 55). *Bacteroides* were detected in the sample of the Endo^{lo} animals, but not in the Endo^{hi} animals. As the proportion of *Bacteroides* was higher in the microbiome of Endo^{lo} animals than that of Endo^{hi} animals, the failure to detect *Bacteroides* in Endo^{hi} animals might be due to the lower detection limit of FISH than NGS. The 16S analysis detected no Proteobacteria in week 6 in Endo^{lo} or Endo^{hi} animals. In contrast, FISH analysis showed hybridization of the Enterobacterial probe in the samples collected from Endo^{hi} mice.
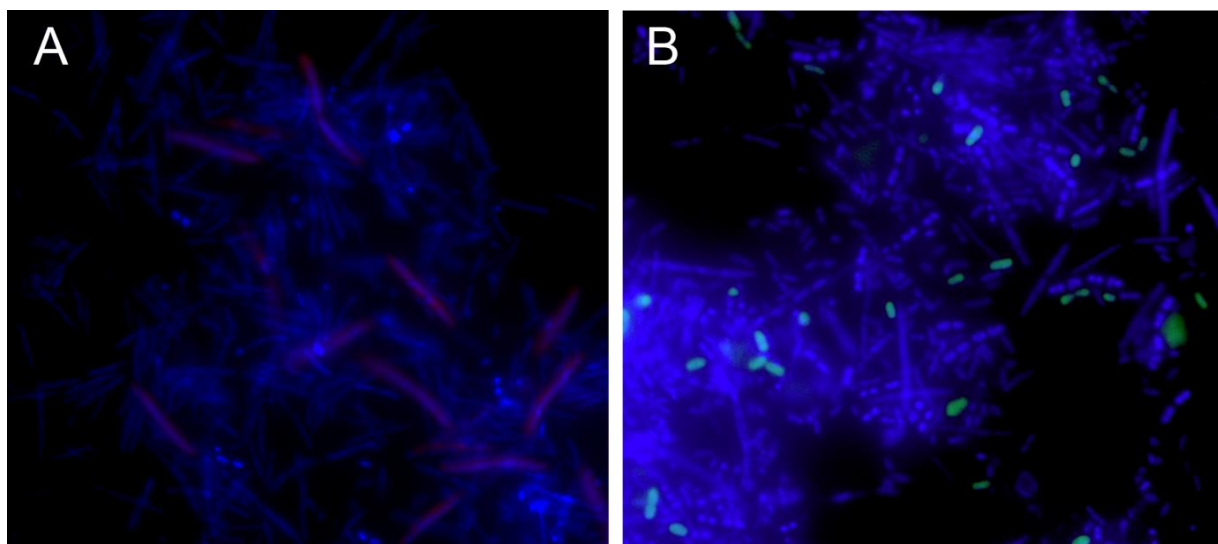
Figure 55: Representative composition of the bacteria present in the feces of *E. coli*$_{MUT}$ treated Endo^{hi} (A) and Endo^{lo} (B) mice 6 weeks after T-cell transfer analyzed by FISH. All bacteria present were stained by DAPI while Enterobacteriaceae were stained red with a Cy3 labeled probe and *Bacteroides* were colored green by an Alexa Fluor 488 linked probe.

3.4.5 The LPS expressed by bacteria fed decides on the disease state

In theory, the only difference between the two *E. coli* strains used for the experiments described above is the structure of the LPS. The Lipid A of the isolated LPS$_{MUT}$ contained palmitic acid in addition to Laurate. The Lipid A of LPS$_{WT}$ contained only Laurate [69]. This indicates that the structure of the LPS decides the colitogenic potential of the bacteria. Creating a mutant strain, however, could also cause other non-evident changes to the bacterium that could affect the pro- or anti-inflammatory potential. To confirm LPS as the molecule responsible for the outcome of the experiments, wild-type and mutant LPS (LPS$_{WT}$ and LPS$_{MUT}$) were isolated and fed to the mice.

To study the impact of LPS$_{WT}$ and LPS$_{MUT}$ on the composition of the microbiota, an experimental setup analogous to that used to study the impact of feeding complete bacteria was chosen (Figure 56). 160 µg LPS, equivalent to 1×10^8 bacteria, was given i.g. Mice were also fed with LPS through their drinking water during the rest of the experiment. 2 mg of LPS were added to their water per week, which was refilled to 250 ml.
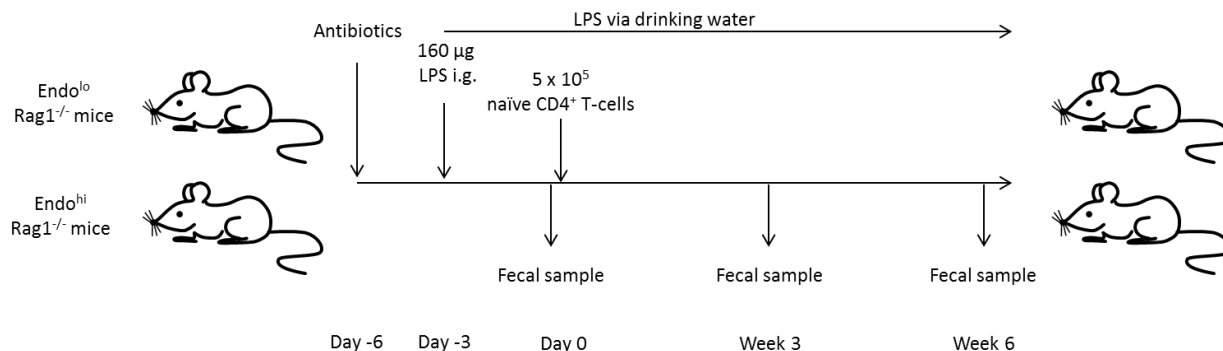
Figure 56: Endo^{lo} and Endo^{hi} mice were treated with metronidazole and streptomycin, respectively, for 3 days beginning 6 days prior to T-cell transfer. Three days before T-cell transfer, the animals were treated with 160 µg LPS i.g. and 2 mg /week via drinking water for the following course of the experiment. 5 x 10^5 naïve CD4$^+$ T-cells were transferred on day 0. Fecal samples were collected on this day prior to T-cell transfer as well as 3 and 6 weeks after T-cell transfer and frozen at -80°C.

As when mice were fed viable bacteria, irrespective of the initial microbiota, mice fed with LPS$_{WT}$ developed colitis whereas mice fed with LPS$_{MUT}$ maintained homoeostasis (Figure 57).
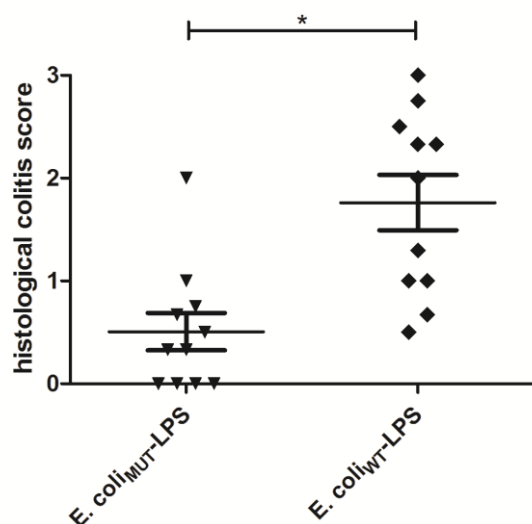
Figure 57: Irrespective of the initial microbiota (Endo^{lo} or Endo^{hi}), mice maintained homoeostasis if they were fed with LPS$_{MUT}$ but developed colitis after feeding of LPS$_{WT}$. Histological colitis score of animals fed with LPS$_{MUT}$ vs. LPS$_{WT}$ irrespective of the initial microbiota. Statistical analysis was performed using the Student t-test, * indicates $p < 0.05$.

The composition of the intestinal microbiome of Endo^{hi} animals treated with LPS$_{WT}$ on day 0, week 3 and week 6 are shown in Figures 58 - 60. Feeding of LPS$_{WT}$ to Endo^{hi} animals prior to T-cell reconstitution induced colitis. In week 0 no Bacteroidetes were detected; (Figure 58), however, this phylum increased throughout the first 3 weeks post T-cell transfer (Figure 59). The microbiota on day 0 was only made up of Firmicutes, or more precisely, the microbiome only contained Bacilli. The Endo^{hi} animals that were not pre-treated with bacteria or LPS also contained Bacteroidetes and Clostridia in their microbiome on day 0. During the first 3 weeks of the experiment these bacterial groups also appeared in the microbiome of the LPS$_{WT}$ treated Endo^{hi} mice. The Tenericutes that appeared in the microbiome of the T-cell only treated Endo^{hi} mice in week 3 were detected in the sample collected in week 6 of the

LPS$_{WT}$ treated Endo^{hi} mice. These mice show an increasing diversity of their intestinal microbiome over the course of the experiment.

Figure 58: Composition of the intestinal microbiota of LPS$_{WT}$ treated Endo^{hi} mice, before T-cell transfer (n = 3; animals were part of experiment 7). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

Figure 59: Composition of the intestinal microbiota of LPS_{WT} treated Endo^{hi} mice, 3 weeks after T-cell transfer (n = 3; animals were part of experiment 7). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

Figure 60: Composition of the intestinal microbiota of LPS$_{WT}$ treated Endo^{hi} mice, 6 weeks after T-cell transfer (n = 3; animals were part of experiment 7). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

Figures 61 - 63 show the composition of the intestinal microbiome of Endo^{lo} mice pre-treated with LPS$_{WT}$. These mice developed colitis during the course of the experiment in the same way as the *E. coli*$_{WT}$ treated Endo^{lo} mice. Both groups of mice showed very similar intestinal microbiome compositions over the whole course of the experiment. High amounts of Bacteroidetes decreased during the 6 weeks whereas Firmicutes showed no apparent change. Verrucomicrobia were present in the samples collected on day 0 and week 6, but not in week 3, and Tenericutes increased during the experiment. All of this however was also found in the T-cell only treated Endo^{lo} animals. The only difference on phylum level seen between the two Endo^{lo} groups that develop colitis, and the T-cell only treated group that stays healthy, is that the microbiome of the latter group also contains an increasing proportion of sequencing reads that could not be associated to a phylum.
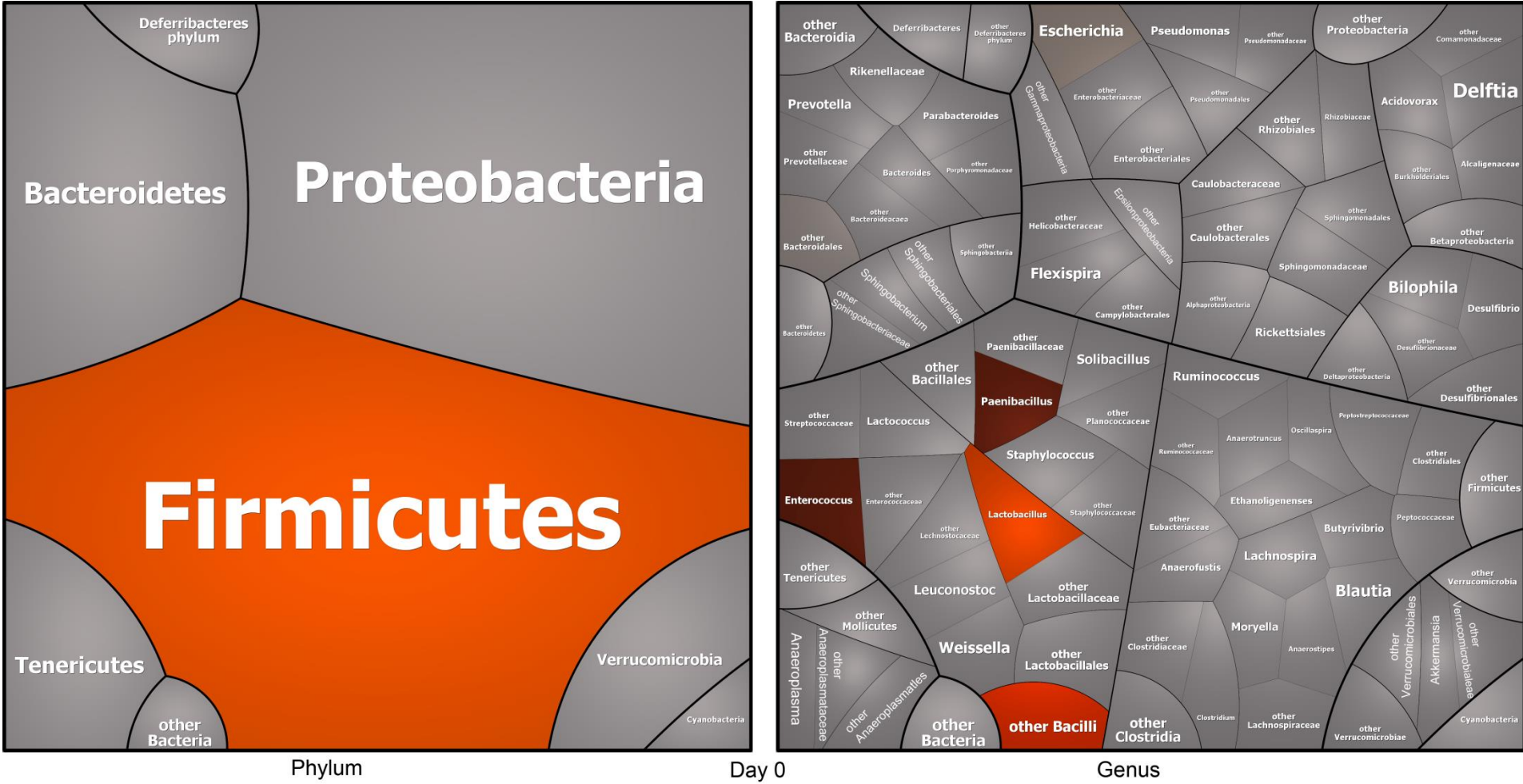
Figure 61: Composition of the intestinal microbiota of LPS$_{WT}$ treated Endo^{lo} mice, before T-cell transfer (n = 4; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
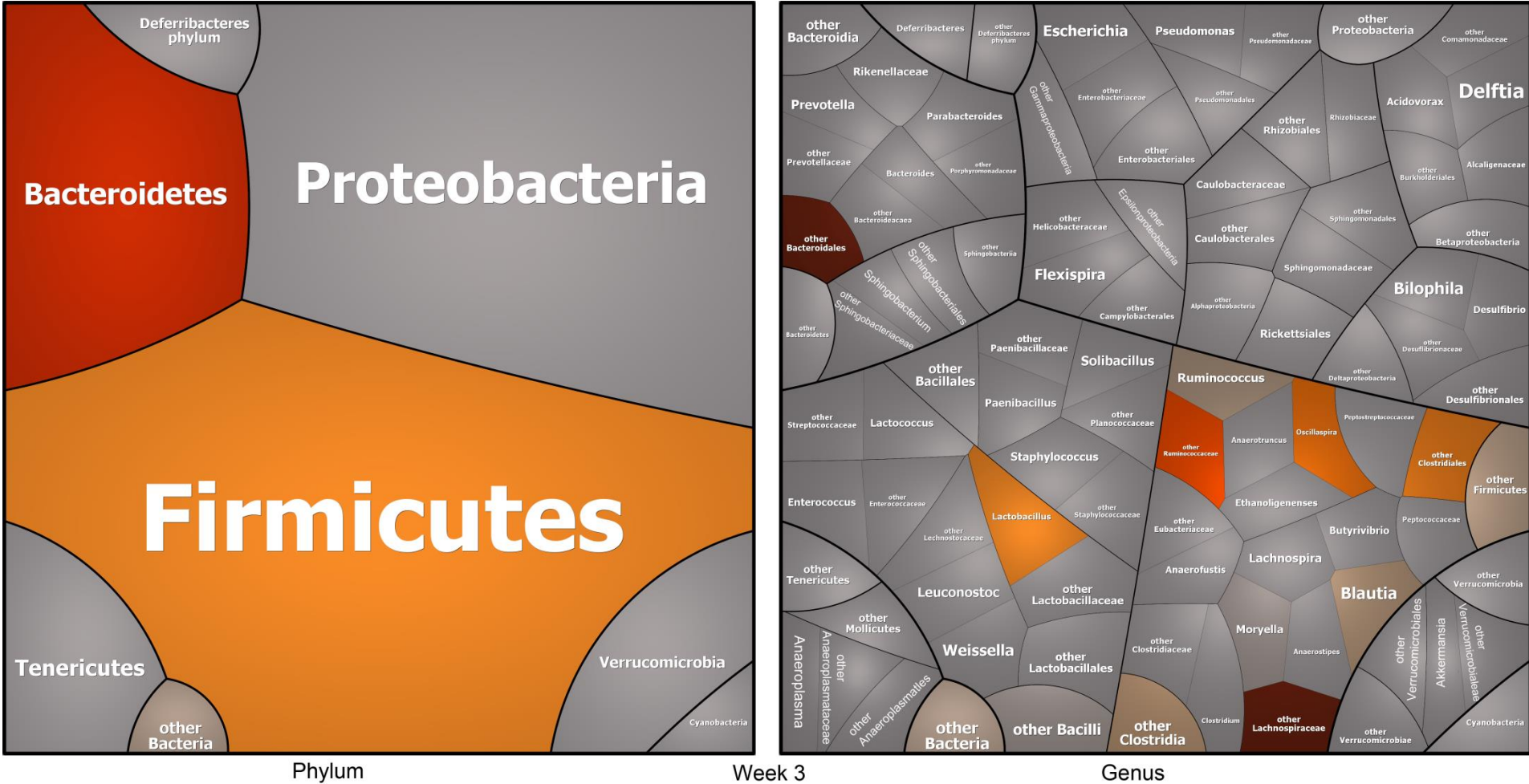
Figure 62: Composition of the intestinal microbiota of LPS$_{WT}$ treated Endo^{lo} mice, 3 weeks after T-cell transfer (n = 4; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
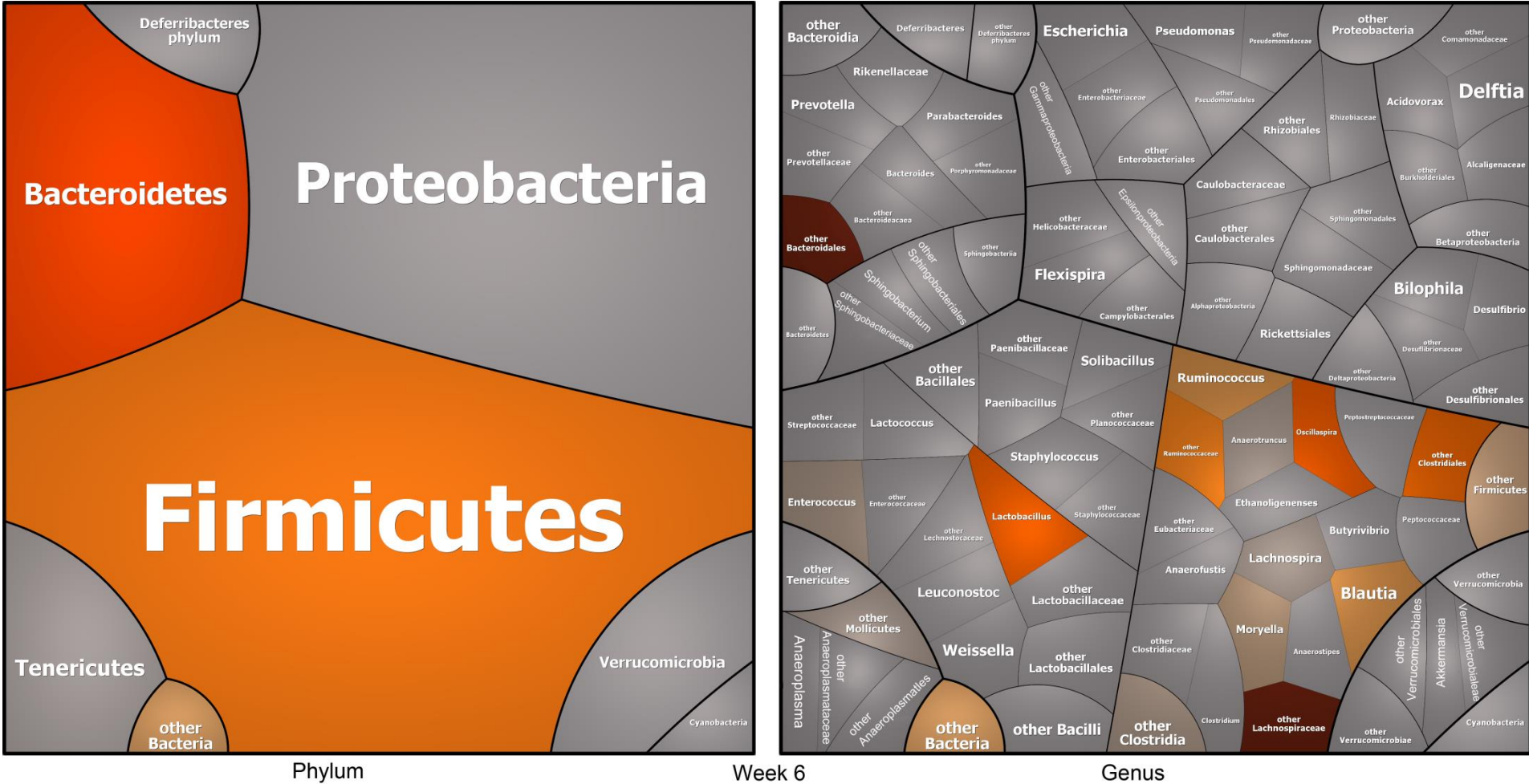
Figure 63: Composition of the intestinal microbiota of LPS$_{WT}$ treated Endo^{lo} mice, 6 weeks after T-cell transfer (n = 4; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

The portions of *Bacteroides* and Enterobactereaceae in the samples collected in week 6 from Endo[hi] and Endo[lo] mice treated with LPS[WT] were also analyzed by FISH (representative samples are shown in Figure 64). These results confirmed the NGS results; more *Bacteroides* were present in the Endo[lo] group and no *Enterobacteriaceae* were detected in either sample. In contrast to the sequencing results, no *Bacteroides* could be found in the Endo[hi] LPS[WT] group.
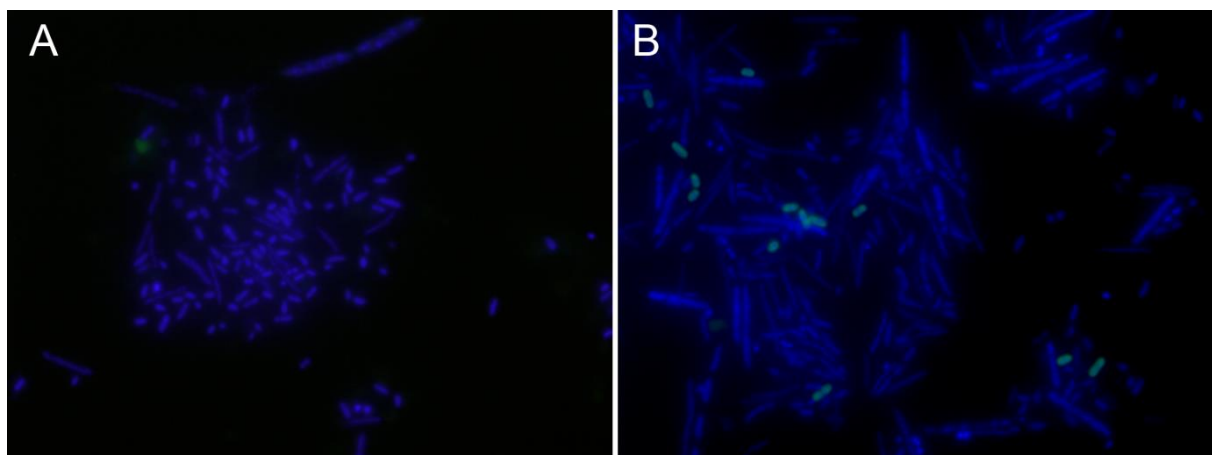
Figure 64: FISH of representative samples collected from LPS[WT] treated Endo[hi] (A) and Endo[lo] (B) mice in week 6. DAPI staining visualizes all bacteria, Enterobacteriaceae are labeled by Cy3 (red) and *Bacteroides* by Alexa Fluor 488 (green).

Feeding of LPS[MUT] to Endo[hi] or Endo[lo] animals prevented colitis development, as did treatment with *E. coli*[MUT] (Figure 57).

The intestinal microbiome of LPS[MUT] treated Endo[hi] mice on day 0 contained mainly Bacteroidetes and Firmicutes (Figure 65). Small amounts of bacteria that could not be assigned to a phylum, and Tenericutes were also detected. The proportions of Bacteroidetes were higher than in T-cell only treated Endo[hi] mice and decreased during the first 3 weeks of the experiment but increased again during the last 3 weeks. The amount of different genera belonging to the phylum Bacteroidetes present in the microbiome of Endo[hi] mice treated with LPS[MUT] was higher than in T-cell only treated Endo[hi] animals.
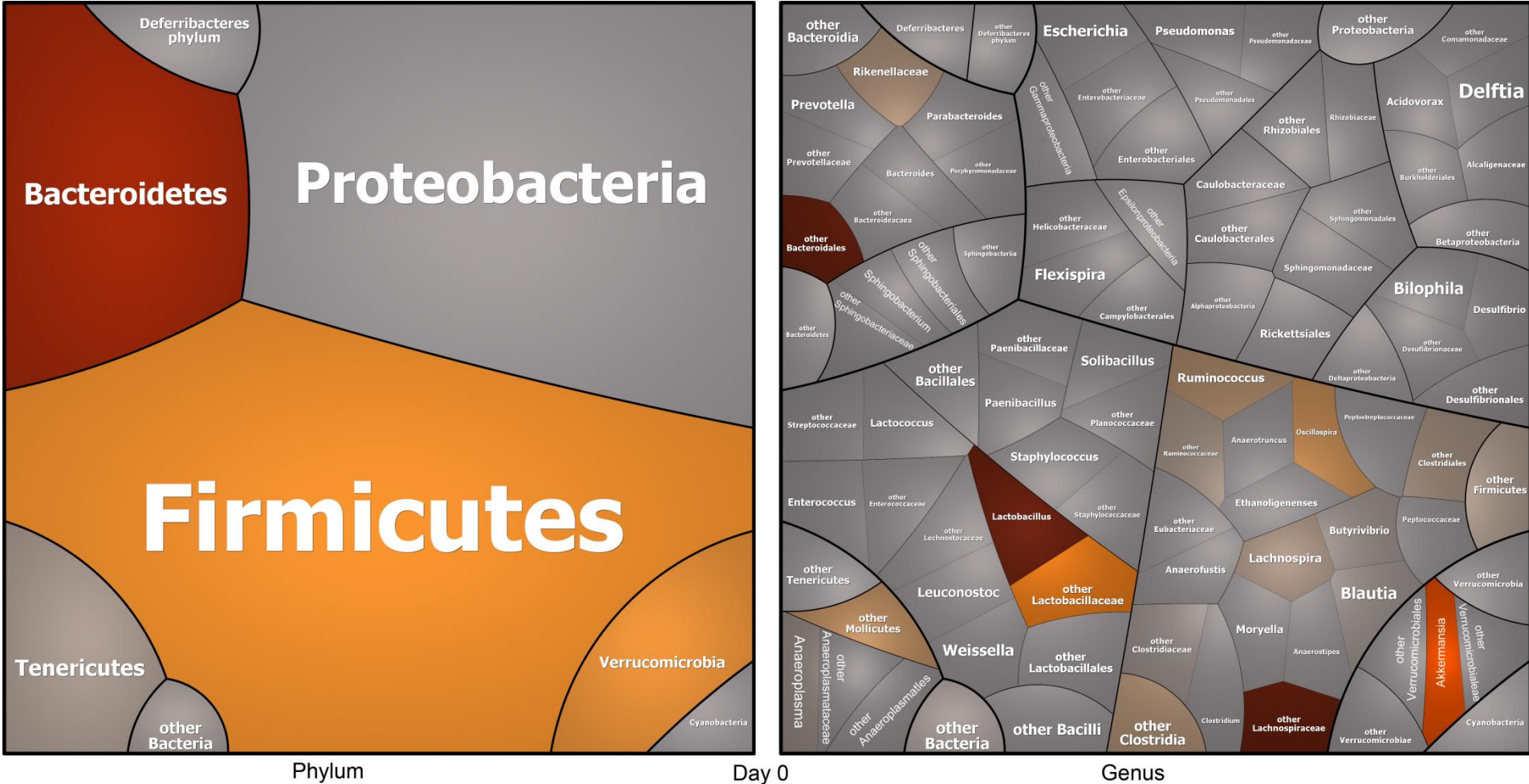
Figure 65: Composition of the intestinal microbiota of LPS$_{MUT}$ treated Endo^{hi} mice, before T-cell transfer (n = 5; animals were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
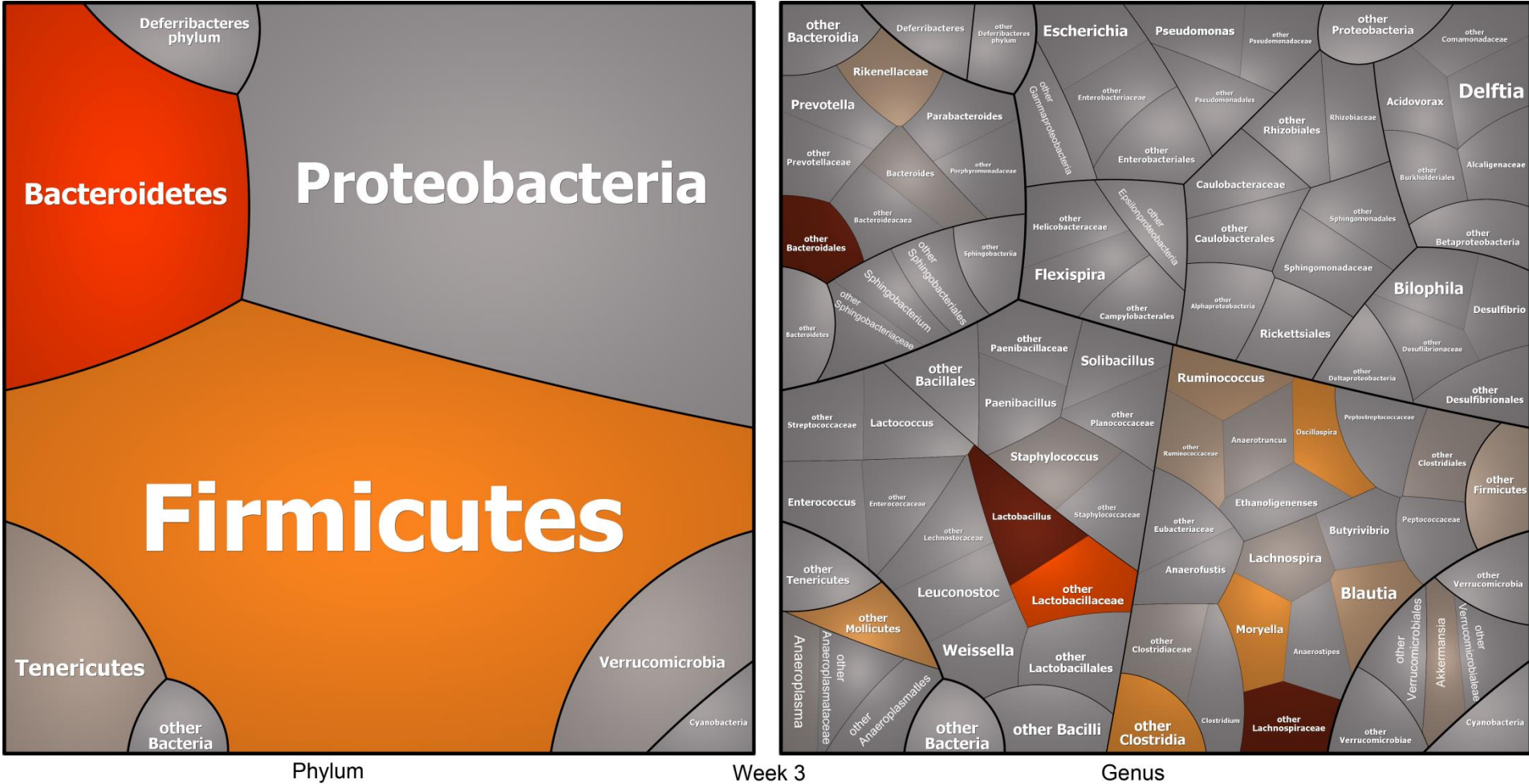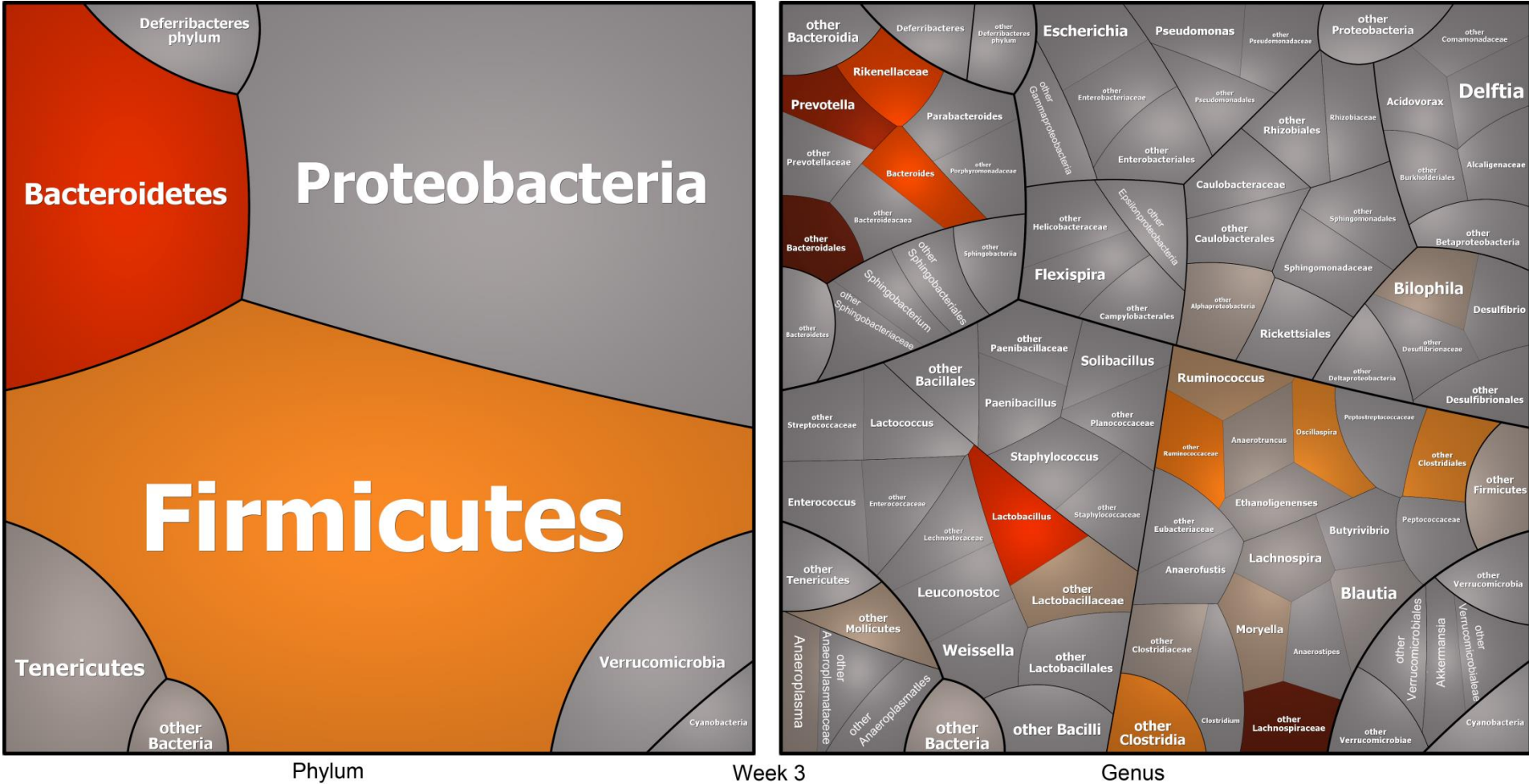
Figure 66: Composition of the intestinal microbiota of LPS_MUT treated Endo^hi mice, 3 weeks after T-cell transfer (n = 5; animals were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
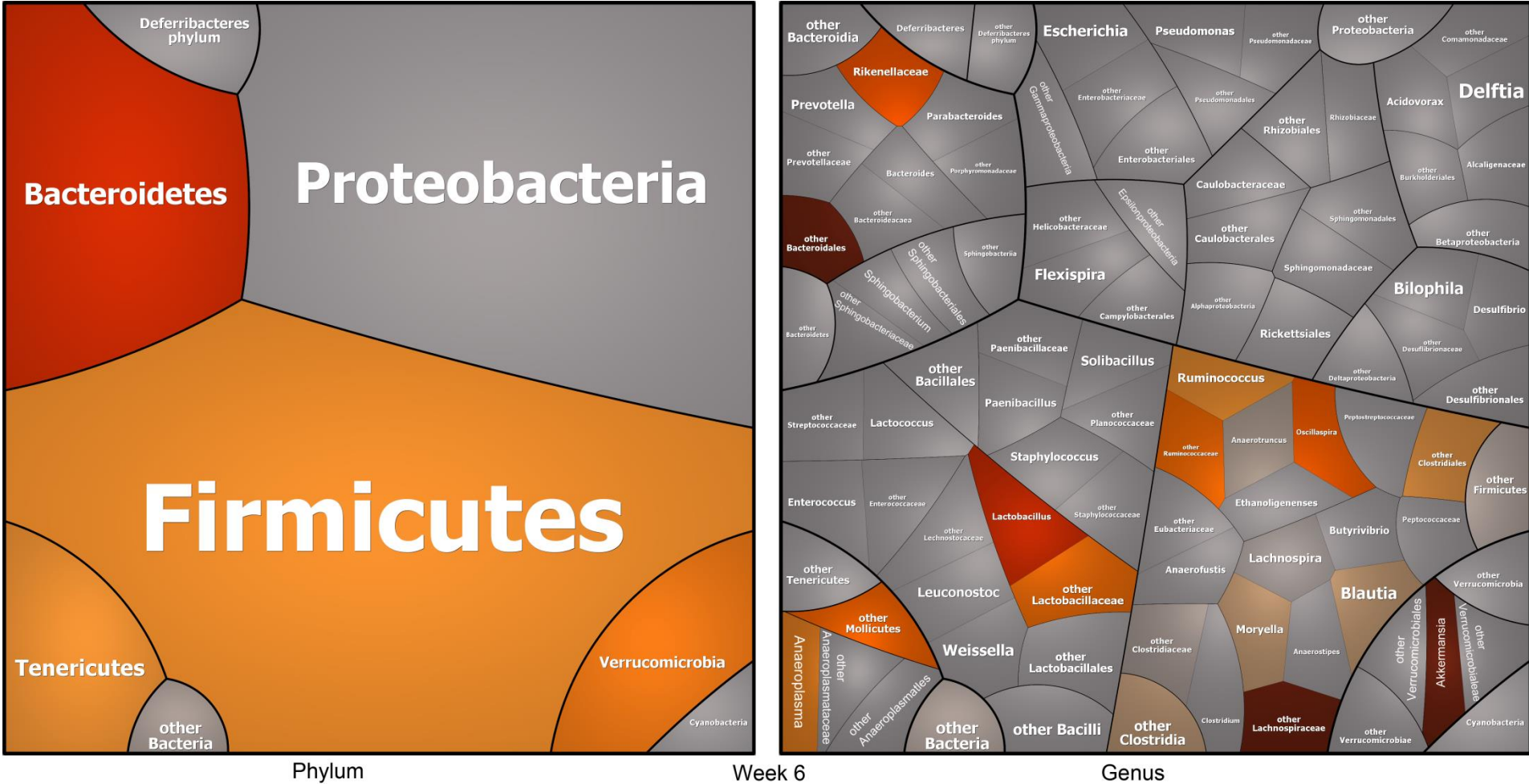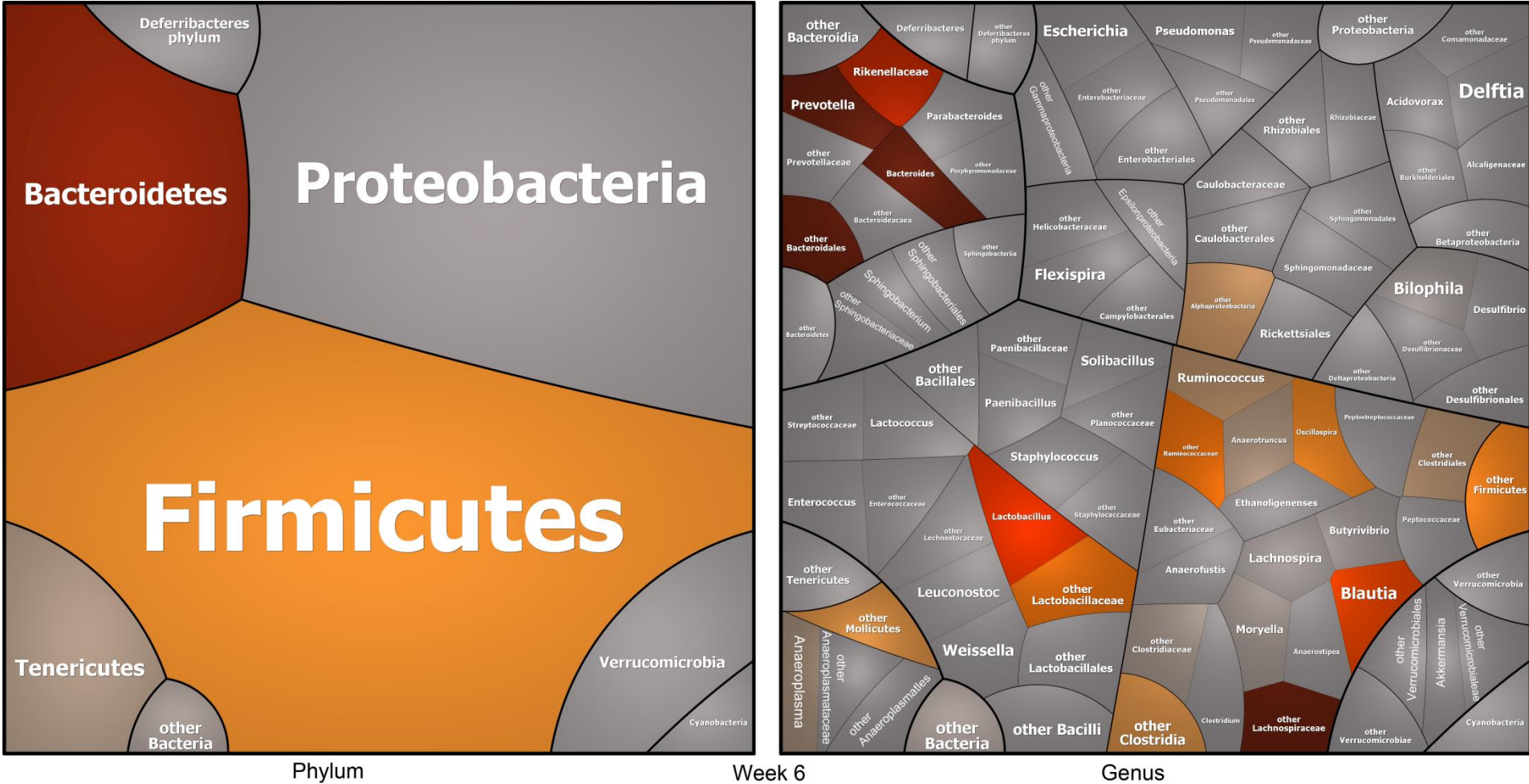
Figure 67: Composition of the intestinal microbiota of LPS$_{MUT}$ treated Endo^{hi} mice, 6 weeks after T-cell transfer (n = 5; animals were part of experiment 3). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

The intestinal microbiota of Endo^{lo} animals fed with LPS$_{MUT}$ contained a high proportion of Bacteroidetes, Firmicutes, and Tenericutes on day 0 (Figure 68). During the experiment, the reads associated to Bacteroidetes decreased. Tenericutes were also detected in the sample collected in week 6, but not in week 3. Reads that could not be associated to a phylum were only found in the samples collected in week 6. Verrucomicrobia (genus *Akkermansia*) were detected in week 3 and week 6 after T-cell transfer (Figures 68 - 70).
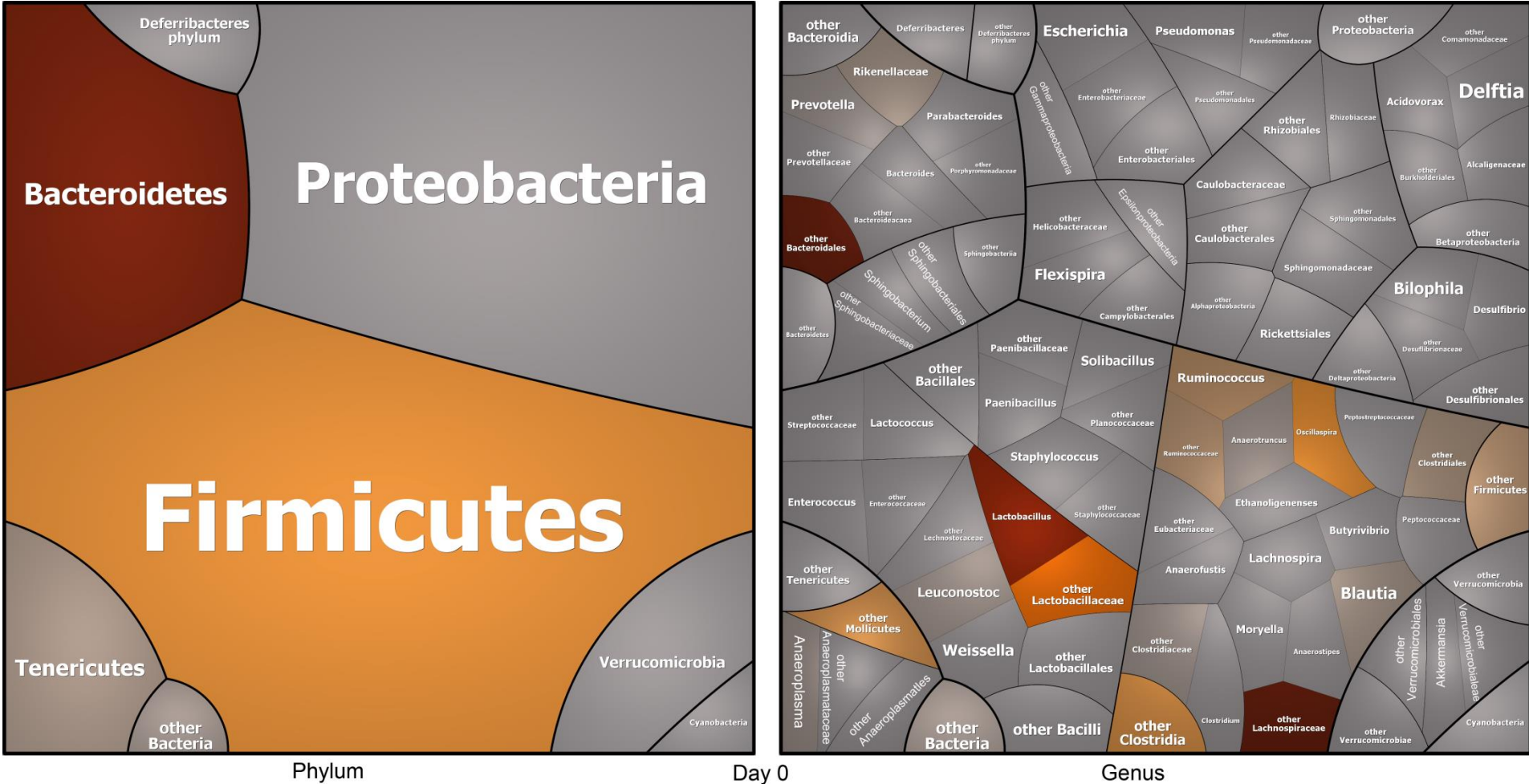
Figure 68: Composition of the intestinal microbiota of LPS_{MUT} treated Endo^{lo} mice, before T-cell transfer (n = 4; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
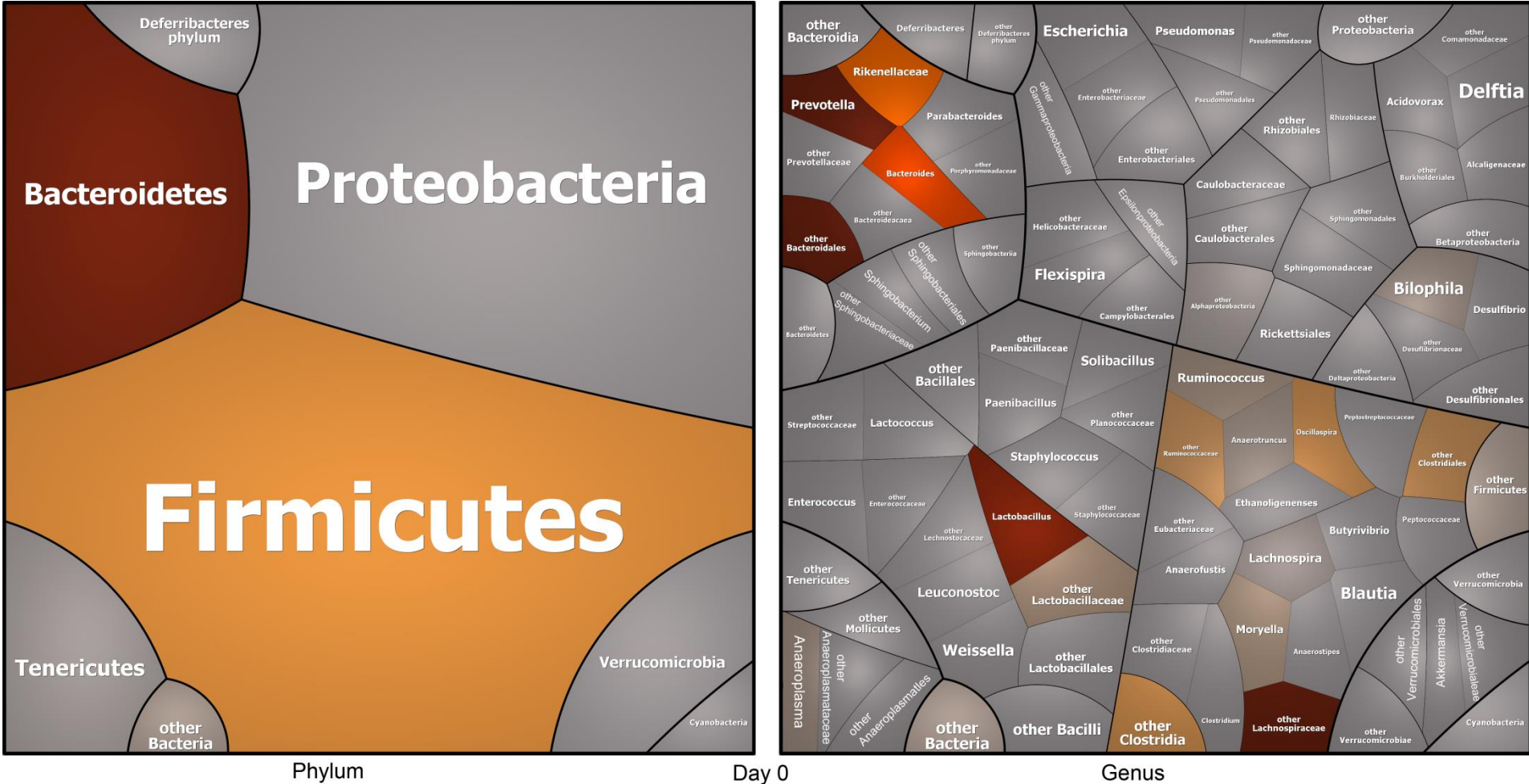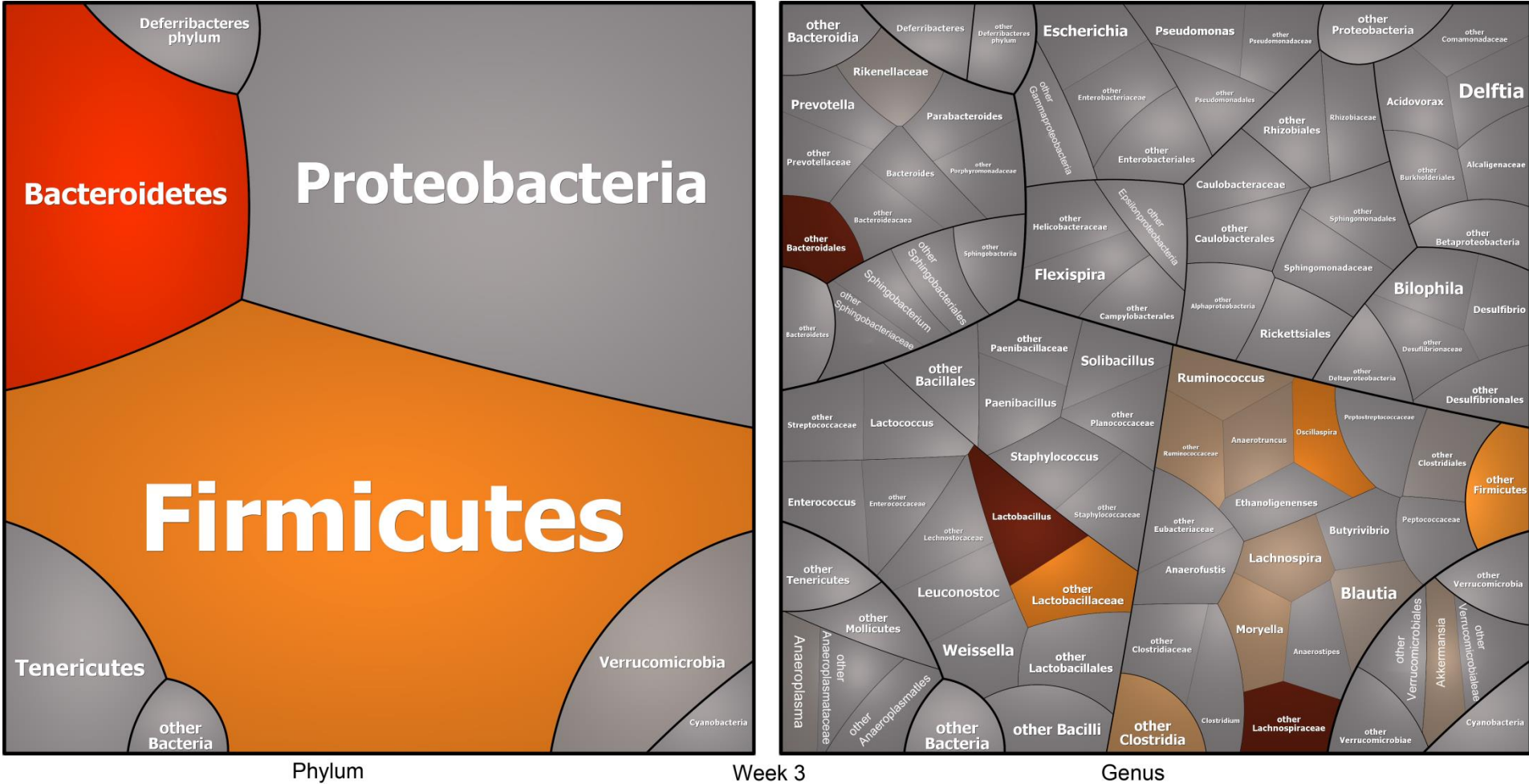
Figure 69: Composition of the intestinal microbiota of LPS$_{MUT}$ treated Endo^{lo} mice, 3 weeks after T-cell transfer (n = 4; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
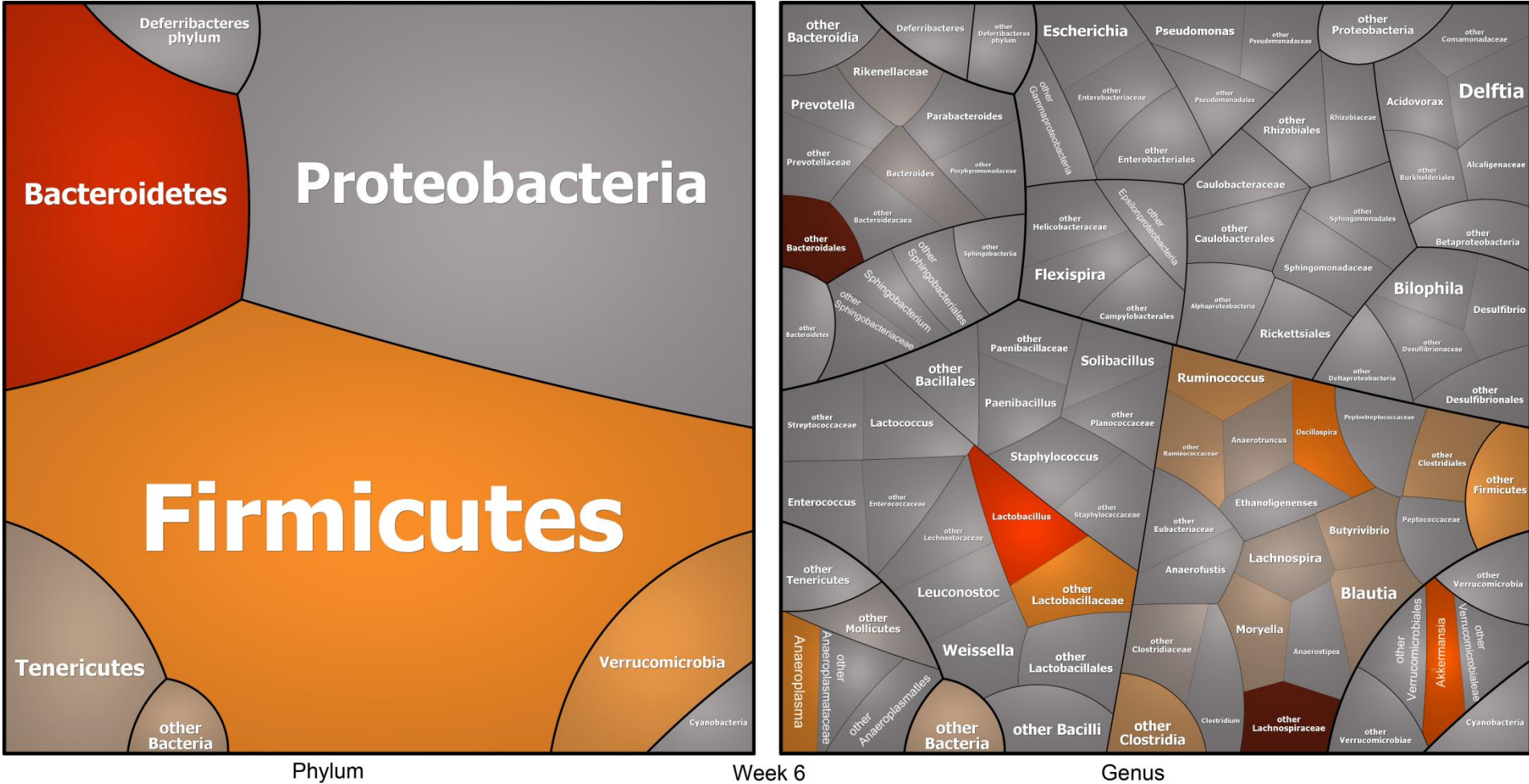
Figure 70: Composition of the intestinal microbiota of LPS_MUT treated Endo^lo mice, 3 weeks after T-cell transfer (n = 4; animals were part of experiment 4). The left image shows the correlation of the sequencing reads with taxonomy on phylum level. In the right image, the reads were associated with the corresponding taxonomy on the genus level where possible. Reads that could not be assigned to a genus were declared as „others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

FISH analyses of samples collected in week 6 of the Endo^{hi} and Endo^{lo} animals treated with LPS_{MUT} confirmed the findings of the sequencing analysis. More *Bacteroides* were detected in the sample from Endo^{hi} animals (A) than Endo^{lo} (B) mice (Figure 71). *Enterobacteriaceae* were not detectable in either sample.
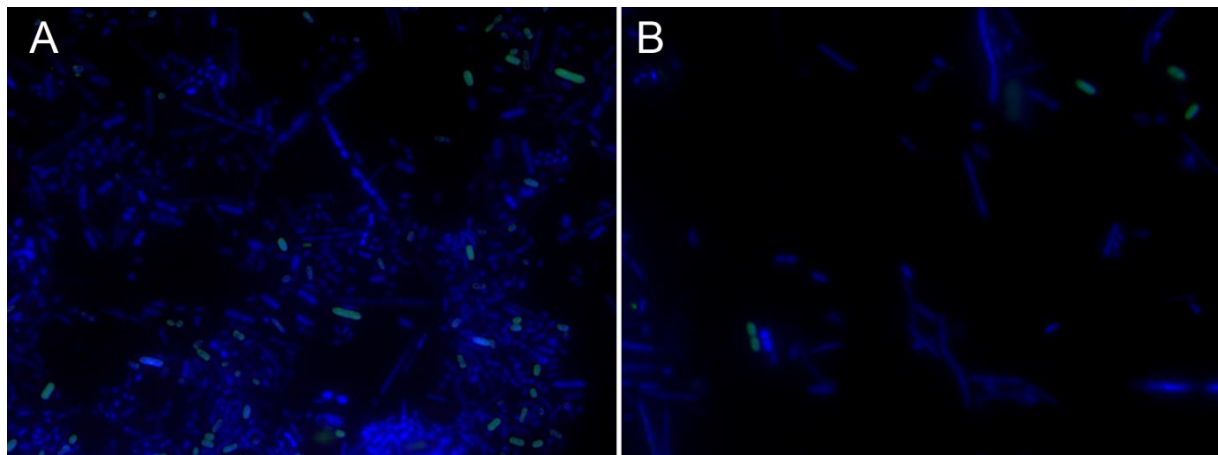
Figure 71: representative FISH analysis of samples collected in week 6 from Endo^{hi} (A) and Endo^{lo} (B) animals treated with LPS_{MUT}. DAPI staining visualizes all bacteria, Enterobacteriaceae are labeled by Cy3 (red) and *Bacteroides* by Alexa Fluor 488 (green). *Bacteroides* were detected in Endo^{hi} and Endo^{lo} mice.

3.4.6 Metagenomics

Mammalian intestines are not only colonized by bacteria, but also by microorganisms such as viruses and archaea. To analyze the complete composition of the microbiota, shotgun metagenomic analysis can be used. These analyses also give information about the functions that can be carried out by the microorganisms. The bacterial composition of the intestinal microbiota can vary extremely between individuals, or even over time within the same individual. As different bacteria can perform the same tasks, however, variability between samples can be lower on the functional level. As no bacterial species or bacterial composition could conclusively be linked to colitis development, additional experiments were performed to analyze the complete microbial composition of the intestinal microbiota as well as the functions that could be performed by the microorganisms.

Other experiments have shown that application of antibiotics, as done for the experiments to analyze the bacterial composition of the gut microbiome, is not necessary to influence the development of colitis by treatment with high or low endotoxic bacteria. No antibiotic treatment was therefore performed for the shotgun metagenome experiments. In addition to prevention of colitis by feeding low endotoxic bacteria before T-cell transfer, we also wanted to study the potential of low endotoxic bacteria to treat colitis. The experimental setup used for these studies is described in figures 72-75.

Endo^{hi} mice were transferred with 5 x 10^5 CD4⁺ naïve T-cells on day 0 and developed colitis over the next 4 weeks. Fecal samples were collected just before T-cell transfer and kept on ice until frozen at -80 °C. Additional fecal samples were collected in week 4 and 6 after T-cell transfer.
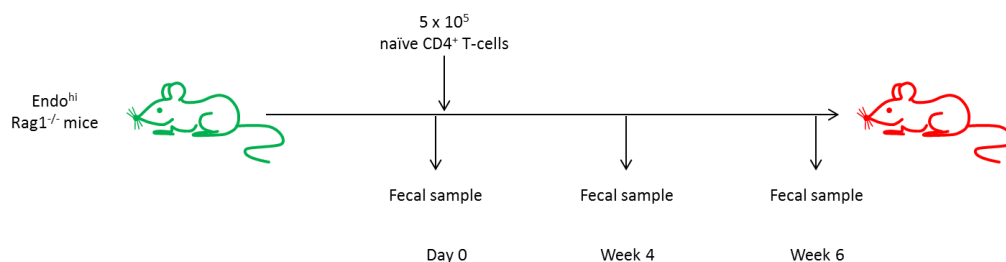
Figure 72: Experimental setup for shotgun metagenomic analysis of Endo[hi]** mice transferred with naïve CD4**[+]** T-cells. Approximately 4 weeks after T-cell reconstitution, the animals show first signs of disease.**

A second group of Endo[hi] mice were treated in the same way but additionally treated with low endotoxic *Bacteroides vulgatus* mpk (1 x 10^8 bacteria) 3 days prior to T-cell reconstitution. The bacterial treatment was continued via the drinking water during the complete experiment. Fecal samples were collected as described for the first group on day 0 as well as in weeks 4 and 6 after T-cell transfer. These animals did not show signs of colitis.
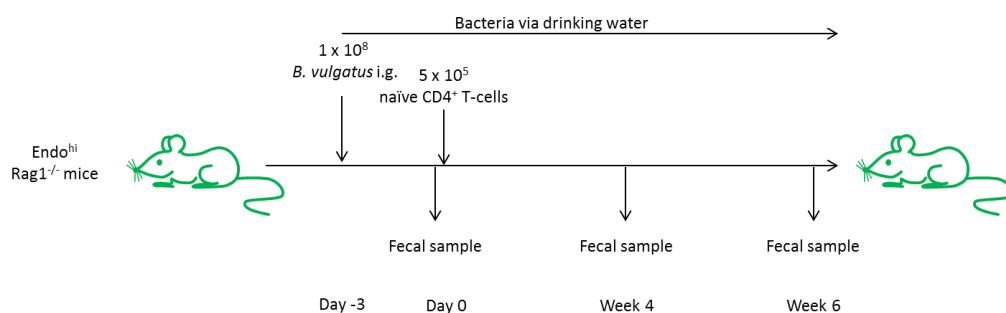
Figure 73: Experimental setup for analyzing the impact of preventive *B. vulgatus* treatment in Endo[hi]** mice transferred with naïve CD4**[+]** T-cells. Mice pretreated with *B. vulgatus* maintain homeostasis.**

A third group of Endo[hi] mice were also reconstituted with T-cells on day 0. In week 4 after T-cell transfer, these mice were treated with *B. vulgatus* mpk i.g (1 x 10^8 bacteria). Bacterial treatment continued for the rest of the experiment through the drinking water of the mice. The experiment with these mice continued for 8 weeks. Fecal samples were collected as in the other groups on day 0, in week 4 before *B. vulgatus* application, in week 6 and in week 8. Diarrhea, indicating disease onset that was present in week 4 improved or even disappeared after treatment with *B. vulgatus*.
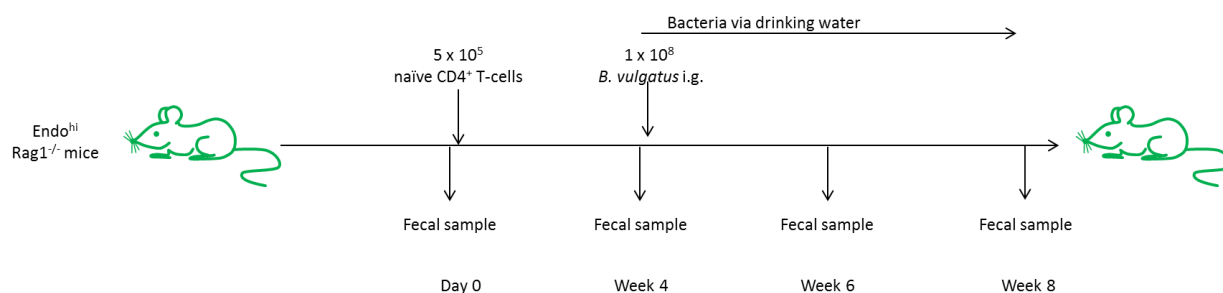
Figure 74: Experimental setup for analysing the impact of colitis and intervention by treantment with *B vulgatus,* on the microbiome using shotgun metagenomics.

To monitor the changes of the intestinal microbiota that were not related to the treatment or colitis development but due to environmental factors, an untreated control group was included. Fecal samples of these mice were collected at the same time as those of the other groups, on day 0, and in weeks 3, 6 and 8.
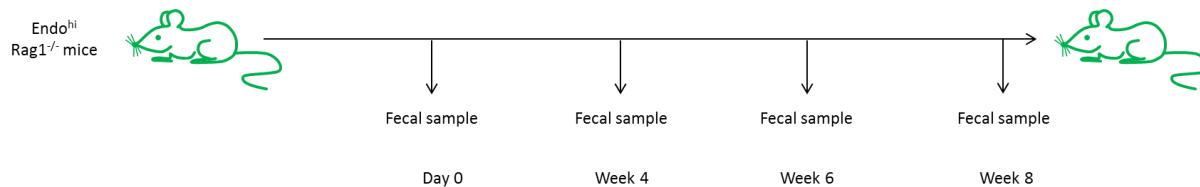
Figure 75: Experimental setup of the reference group for monitoring the impact of environmental factor on the metagenome.

The DNA was isolated from all the samples as described 2.2.12.2 DNA isolation. Sequencing and analysis of these samples is the subject of ongoing studies.

3.4.7 Transcriptomics

Analyzing the 16S genes from a fecal sample is a way to identify bacterial species colonizing the gut of an individual. By studying total DNA (shotgun metagenomics) instead of amplicons, all the genes present within the intestinal microbiome can be evaluated and conclusions can be drawn about the hypothetical functional complexity provided by the intestine's bacterial population. The simple presence of a gene, however, does not imply that it is active. By studying the RNA found within a sample instead of the DNA, the discrepancy can be reduced between the genes present and those actually influencing functions of the microbiome. Only the active genes are translated into mRNA and have the ability to impact the features of the microbiota. The potential for false associations between a phenotype and the role of the bacterial community within the gut can be reduced by studying transcriptomics instead of metagenomics.

Previous experiments have shown that low endotoxic bacteria can prevent colitis development, whereas high endotoxic bacteria trigger colitis. In the course of other projects, germ-free mice were colonized with *E. coli* mpk and/or *B. vulgatus* mpk. IL2[-/-] mice colonized with *E. coli* mpk developed colitis, while those colonized with *B. vulgatus* mpk stayed healthy, as did mice colonized with both bacterial strains [82,186]. It is therefore interesting to analyze how the bacteria influence each other and the host. Studying which genes are expressed in the different colonization scenarios can help to understand what triggers colitis, and how *B. vulgatus* mpk prevents *E. coli* mpk induced colitis development.

For this study, germ-free C57Bl/6 mice were colonized with *E. coli* mpk through the drinking water. After 4 weeks, fecal samples for transcriptomics were collected and frozen in RNAprotect bacteria at -80 °C.

A second group of mice was colonized with *B. vulgatus* mpk through the drinking water. Fecal samples for transcriptomics from these mice were also collected 4 weeks later.

The mice colonized with both bacterial strains were first colonized with *E. coli* mpk und 5 weeks later colonized with *B. vulgatus* mpk through their drinking water. Fecal samples for transcriptomics were collected 8 weeks after the first colonization and frozen as described above.

The samples were used to establish RNA isolation for transcriptomic analysis. Further analysis is a part of ongoing studies.

RNA was isolated from fecal samples as described in 2.2.13.2 RNA isolation. The quality of the isolated RNA was checked using a Bioanalyzer 2100 and the RNA nano kit. The electropherogram of RNA isolated from freshly collected mouse feces is depicted in Figure 76 A. The typical peaks for bacterial 16S and 23S ribosomal RNA (rRNA) were clearly visible. A third peak within the 5S rRNA region indicated accumulation of small RNA fragments. This analysis showed no signs of RNA degradation. A bump in the electropherogram under the 16S and 23S rRNA peaks indicated DNA contamination. Applying DNase treatment removed DNA residues (Figure 76 B). The amount of 1 000 – 4 000 bases long nucleic acids was reduced and a second peak within the 5S rRNA region marked in Figure 76 A was also eliminated.
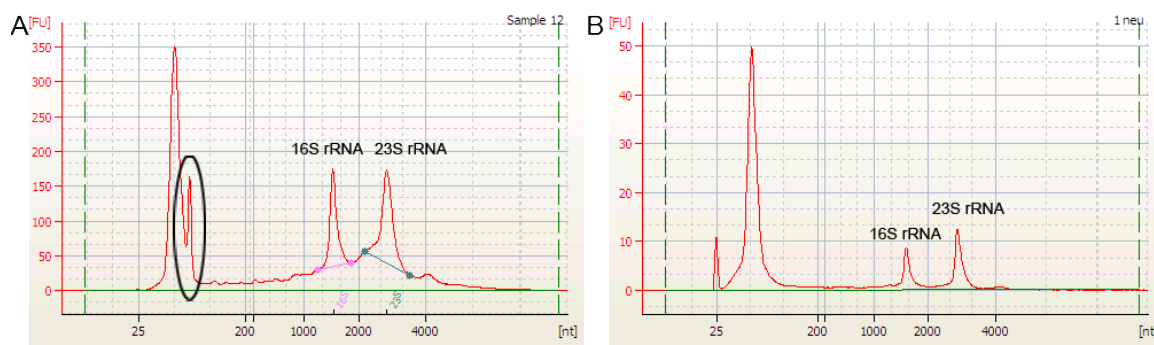
Figure 76: RNA isolated from fecal samples of conventionally colonized mice were contaminated by DNA residues (A) that were eliminated by DNase treatment. (B). The RNA quality was analyzed using a Bioanalyzer 2100 (Agilent). The RNA was loaded onto a RNA chip using the RNA nano reagents.

Feces from germfree animals should only contain host RNA or ribonucleic acids that were part of the diet. These samples were used to determine the amount of contaminating RNA not originating from intestinal bacteria. As shown in Figure 77 no 16S or 23S rRNA could be detected within these samples. A very small amount of RNA within the 5S rRNA region was visible, however in much lower concentrations compared to samples from conventional mice.
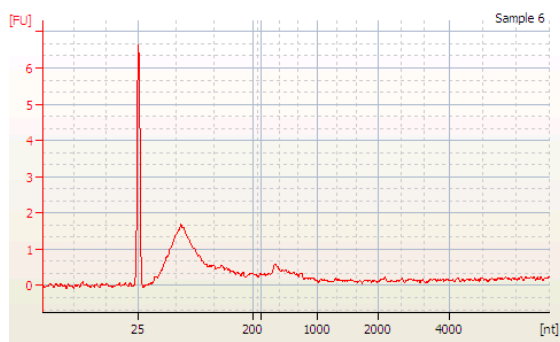
Figure 77: Electropherogram of RNA isolated from feces of germfree mice.

Isolation of RNA using the Zymo Research kit "Soil/Fecal RNA MicroPrep" from cultured bacteria shows the same peaks in the electropherogram (Figure 78) as already seen for the conventional mice (Figure 76).
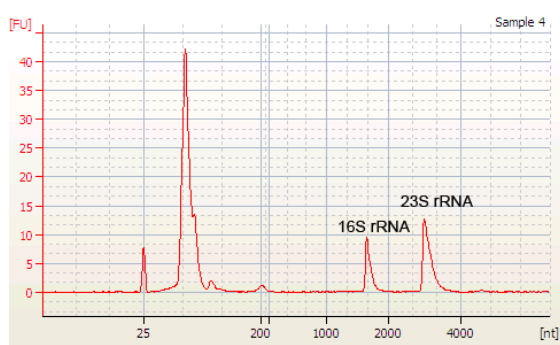
Figure 78: Next to the expected 16S and 23S rRNA, RNA isolated from cultured bacteria with the Soil/Fecal RNA MicroPrep kit contains large amounts of small RNA. Electropherogram created using a Bioanalyzer 2100 and the RNA nano reagents of RNA isolated from cultivated *B. vulgatus*.

In order to determine whether RNA has to be isolated from fresh fecal samples or if the samples can be frozen and collected before retrieving the RNA, samples were collected and stored at different conditions. For each group three samples were analyzed. The conditions tested were fresh and frozen native samples as well as fresh and frozen samples homogenized in RNAprotect Bacteria. Figure 79 shows representative electropherograms for each group of samples. No obvious differences were apparent between the fresh samples and the samples frozen in RNAprotect Bacteria. The amount of shorter RNA fragments increased in native frozen samples. The ratio of 16S and 23S rRNA was also different in these samples compared to the other RNA isolates. For metatranscriptomic analysis the RNA from fecal samples should therefore be isolated immediately after sample collection or frozen in RNAprotect Bacteria at -80 °C until the RNA can be isolated.
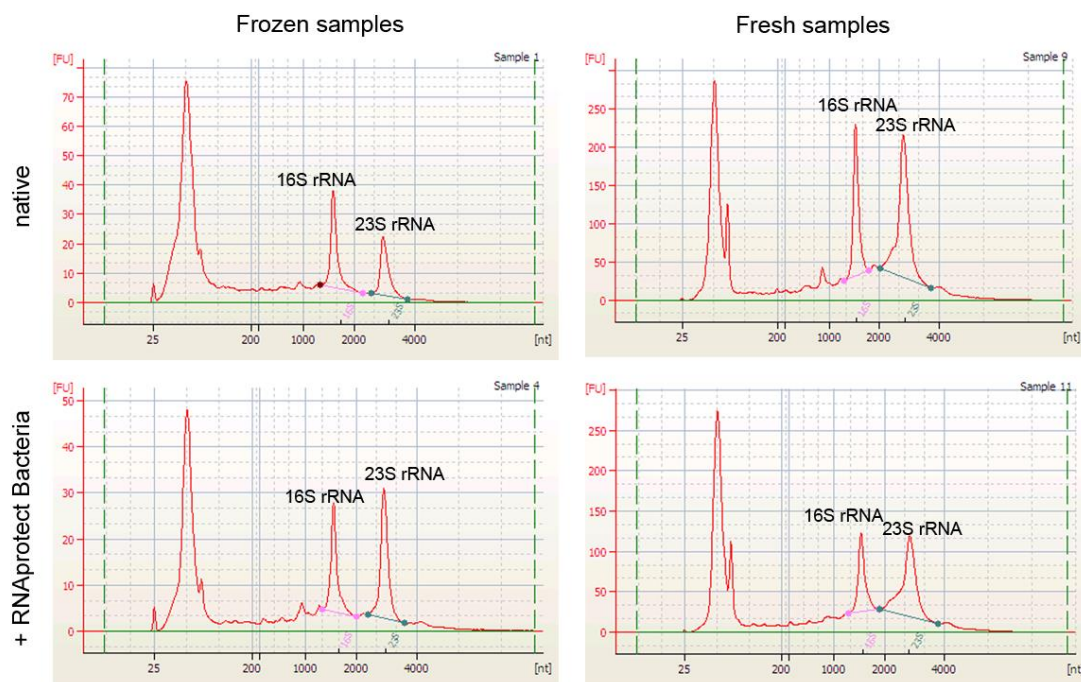
Figure 79: Elektropherograms of RNA isolated from fecal samples of conventional mice. Fecal samples were either homogenizedin RNAprotect Bacteria or native and were used freshly or frozen previous to RNA isolation. Freezing of native fecal samples leads to RNA degradation, whereas the three other conditions do not show apparent differences.

RNA stability was further assessed by analyzing its quality after a cycle of freezing and thawing. The electropherograms (Figure 80 A and B) show an increase of short and decrease of long RNA sequences indicating RNA degradation. RNA should therefore be transcribed into DNA immediately after isolation and not frozen previous to this step.
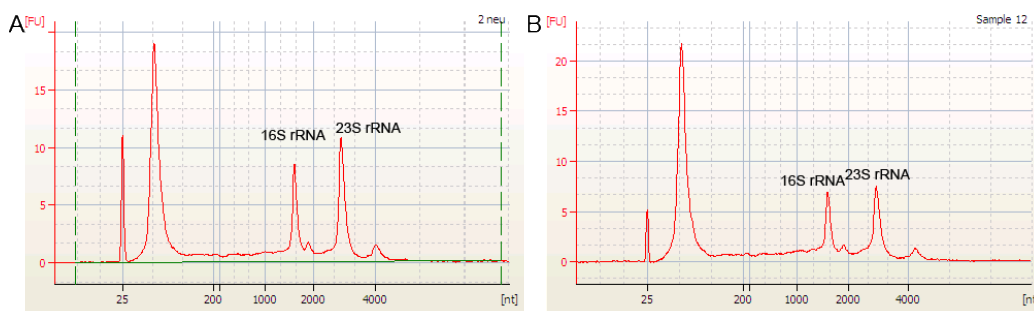
Figure 80: Electropherogram of isolated RNA immediately after isolation (A) and after freezing and thawing (B). Freezing the RNA after isolation leads to RNA degradation, therefore only freshly isolated RNA should be used for reverse transcription and subsequent sequencing.

4. Discussion

4.1 Establishment of Methods

4.1.1 qPCR: verification of primer specificities

Quantitative real-time PCR (qPCR) is a method for determining the amount of defined DNA copies in a sample. Samples with defined concentrations are used to determine absolute amounts of DNA. Without standards, relative quantities of DNA can be determined by comparing multiple samples.

Primer specificity determines which defined DNA sequences can be analyzed. Primer sequences were selected from publications or designed *in silico* based on publicly available bacterial genome sequences (www.ncbi.nlm.nih.gov). The specificity of the chosen primers was validated in two steps. First, a Primer-BLAST search was conducted for each primer pair to determine its specificity. The results of these searches, however, are only as good as the reference database and the results vary depending on parameter selection. In addition to computational analysis, species specificity was ensured in an experimental approach. DNA was isolated from a variety of different bacteria representing many bacterial phyla, classes, orders, families and genera. It is widely accepted that bacteria can be grouped according to the sequence identity of their 16S rRNA gene [164,165]. Selection of primers complementary to extremely conserved regions will amplify a wide variety of bacteria and can therefore be used as universal primers. Designing primers to variable regions is useful to target subsets of bacteria. Depending upon the degree of variance of the primer binding site, bacteria grouped at different taxonomic levels can be targeted. Regions of high variability will be chosen to differentiate bacteria on a species level. The closer the phylogenic relationship between bacteria, the more likely a primer set will amplify both strains and lead to a false positive result. The bacterial strains chosen for experimental validation of primer specificity were selected to include closer and more distantly related species to cover different specificity levels. The primers used for quantification of *B. vulgatus* and *B. dorei* did not amplify other members of the genus *Bacteroides* and were therefore considered to be specific for these two closely related species. Another primer set could be used to detect all the members of the order Bacteroidales included in the experiments. The experimental approach verified the results of the *in silico* sequence comparison, indicating that the selected primers were specific for the targeted bacteria.

Amplification efficacy is important for calculation of the template concentration. Theoretically, during one cycle of amplification every template strand is doubled. After n cycles the initial concentration will accordingly be increased by a factor of 2^n. However, in reality the amplification efficiency is slightly lower, including factors such as less-than-optimal primer properties, PCR inhibitors in the sample, or primer dimerization [187,188]. By amplification of defined template dilutions, it is possible to calculate the efficiency of a primer set. This is done by plotting the concentration vs. the threshold cycle and performing simple linear regression.

4.1.2 Fluorescence in situ hybridization

Fluorescence in situ hybridization (FISH) is a non-PCR dependent method and can therefore be considered independent from the sequencing and qPCR approaches. For this reason, the method was selected to verify the results of 454 sequencing. One great advantage of this method is that it is culture independent: if the DNA sequence of a bacterial strain is known, probes can be designed and bacteria can be detected directly in a sample.

16S rRNA sequencing and qPCR are both based on PCR and are therefore likely to be biased in the same way. FISH on the other hand does not rely on amplification of DNA sequences, although similar to primers, specific probes are designed to detect certain bacteria. This method is therefore targeted and cannot be used for the identification of unknown bacteria or bacteria that have not been sequenced.

An additional limitation of this method is the high threshold for the detection of bacteria. This is because only a very small amount of sample is analyzed, and it must be diluted enough to ensure single bacteria can be identified. Bacteria which are present in low concentrations are therefore unlikely to be detected.

Our results show that bacteria can be detected in stool samples by FISH. We were able to establish a protocol without the need to remove stool particles prior to probe hybridization. The background fluorescence could clearly be distinguished from specific staining of the probes.

Two different methods can be used for absolute quantification; 1: counting bacteria stained by a certain probe; 2. Determining the area stained by the probe in question [189,190]. For counting bacteria, it has to be ensured that the sample is dilute enough to clearly identify single bacteria on the slide. Differences in morphology do not influence the count. Quantification of bacteria according to area stained will be biased by the size of the bacterial species. The larger a bacterial cell, the bigger the area will be that is colored by the probe. Comparison of a small and a large bacterial cell will therefore result in supposedly larger amounts of the bigger species. For exact quantification, normalization is necessary however, if groups of bacteria with different morphology are targeted, normalization becomes impossible. Relative quantification may also be useful, including cases where not all bacteria can clearly be separated from each other. The result will be somewhat biased but, an estimation of the relative amount is possible and useful.

4.1.2.1 Probe selection

Complete verification of sequencing results would require designing probes for all different taxonomic levels of the detected bacteria. We hypothesized, based on earlier experiments and published data that *Bacteroides* and Enterobacteriacae would be of interest in IBD development. Bacterial identification was limited to these groups and probes targeting all bacteria or specific for Enterobacteriaceae, *Bacteroides* or *Bacteroides vulgatus* were selected.

These targets were selected on the basis of published data indicating that certain *B. vulgatus* strains are able to inhibit *E. coli* induced colitis and the observation that Rag1$^{-/-}$ mice with high or low CFU counts of *Bacteroidetes* or *Enterobacteriaceae*, respectively, do not develop colitis after T-cell transfer, although genetically identical mice with opposite relative amounts of *Bacteroidetes* and *Enterobacteriaceae* did [41,183]. In addition to these findings, metagenome analyses have associated IBD with altered *Bacteroides* and *Enterobactericeae* colonization [42,44,191].

4.1.2.2 Multiple staining

Staining one sample simultaneously with multiple probes reduces the workload and time necessary for determining the relative abundance of targeted bacteria. It also ensures that the conditions are consistent for all bacteria, ensuring comparability.

A pre-requisite for simultaneous staining with various probes specific for different bacteria is that staining is successful using identical hybridization conditions. After a protocol was established that could be followed for staining with any of the selected probes, simultaneous staining with 3 probes was successfully attempted. Effective analysis of FISH using multiple probes strongly depends on the microscope, camera and filters used for visualization of the samples. Although the coloring of the bacteria could clearly be differentiated by eye, when using the Leica DMRE microscope, the signals originating from the Cy3, Flourescin or Alexa Fluor 488 dye were not clearly differentiated.

The cameras used with our microscopes do not detect different colors, but only capture light signals. The color is added to the picture artificially by the software. The staining of bacteria in a multi-staining setting is therefore detected successively. The fluorescent dyes can be separated according to their varying wavelengths. In combination with the correct filter, the camera will only detect signals of a single dye. The light produced by the others, will not be able to pass the filter. Using the Zeiss CLSM 710 NLO instead of the Leica DMRE microscope solved the problem of detecting multiple dyes at once. These results show that the Zeiss CLSM 710 NLO should be preferred when analyzing FISH samples stained using Cy3 in combination with Alexa Fluor 488 or Flourescin.

4.1.3 **DNA isolation from fecal samples**

For analysis of the microbiome by 16S or shotgun metagenomics, DNA of the bacteria within the feces samples has to be isolated. The data resulting from DNA based identification of community members varies depending on the isolation method [180,192–195]. The greatest difference between envelopes of different bacterial species is the structure of the cell-wall. Gram-positive (gram$^+$) bacteria are much more stable than gram-negative (gram$^-$). To ensure that gram$^+$ bacteria are also lysed properly, samples were frozen and the protocol of the QIAamp DNA Stool Kit was modified to include an additional bead beating step and enzymatic lysis using Lysozyme [180,192,196]. Bonot et al [195] tested 10 different DNA extraction protocols and only two methods, of which one was the QIAamp DNA Stool Kit, yielded DNA that could successfully be used for subsequent amplification. Different parameters were assessed for determining the quality of the extracted DNA. Purity was assessed

spectrophotometrically. The ratios of absorption at 260/280 nm and 260/230 nm provide an indication of possible contaminants, protein or carbohydrates, aromatic compounds, humic acids and phenolics, respectively. Ideally the 260/280 nm ratio is 1.8 and the 260/280 nm ratio is between 2.0 and 2.2 [194]. Using the QIAamp DNA Stool Kit usually yielded DNA that met these criteria. The extracted DNA was also tested for degradation by gel electrophoresis. Although a smear indicating slight degradation was seen for most of the samples, the majority was high molecular DNA of satisfactory integrity. The 16S rRNA gene region which was amplified and later sequenced is only approximately 700 bp long. Even the small amount of fragmented DNA present in some samples was longer than that, indicating that the amplification of the 16S rRNA gene fragment was not affected.

Compared to the phenol-based DNA isolation used for DNA extraction for shotgun metagenomics, isolation with the QIAamp DNA Stool Kit was much faster and more samples could be processed simultaneously. The slightly lower quality, due to degradation, was accepted as a tradeoff, as only short DNA fragments were of interest for subsequent applications and amplicon generation did not seem to be affected.

For shotgun metagenomics, the total DNA is sheared into fragments of defined length. Using non-degraded DNA as starting material minimizes the amount of sequences of incorrect length. Depending on the method used to fragment the DNA, sequence specific biases are introduced, which are exacerbated if degraded DNA is used as starting material and may even result in complete loss of genomic regions [197]. For shotgun metagenomics of the intestinal microbiome, the DNA extracted using the QIAamp Stool Kit was not of high enough quality. DNA for this application was therefore isolated using a phenol-based DNA extraction method, yielding non-degraded DNA that was longer than 10 000 bp. Freezing of the samples at -80 °C and a bead-beating step were included to enhance lysis of gram$^+$ bacteria and to minimize biases due to DNA isolation.

4.1.4 RNA isolation from fecal samples

RNA was isolated from fecal samples in preparation of transcriptomic analysis of the intestinal microbiome. Expression profiles can be used to identify specifically regulated genes under different environmental conditions.

4.1.4.1 RNA quality

RNA was isolated with the ZR Soil/Fecal RNA MicroPrep kit following manufacturer's instructions. Quality control was done using an Agilent Bioanalyzer and the RNA 6000 Nano Kit. The electropherogram produced by the Bioanalyzer gives information on RNA integrity and amount. The analysis of the RNA using the Bioanalyzer is size-dependent separation of nucleic acids by gel electrophoresis [198]. The small tRNA (4S) and the 5S rRNA will be detected within the same region and separated from 16S and 23S rRNA. The mRNA is not of defined length therefore, no defined peak can be associated to mRNA [199]. As the mRNA only contributes 1-5 % to total RNA [171], it is not explicitly visible in the spectrum of total RNA. The ratio of 23S/16S rRNA (approximately 1.5 - 2 for good quality) is one of the most common used indicators for RNA integrity, however, exceptions to this rule do exist [200,201]. If RNA is

degraded, the amount of short RNA strands increases and consequently the long strands decrease, which can be detected in the electropherogram. As a measure of RNA quality, the level of the baseline in the electropherogram was evaluated, which is increased across the entire RNA range if degradation occurs.

The large amount of RNA accumulating within the range of the 5S rRNA was surprising. The result expected according to the documentation provided by Zymo Research for the RNA isolation kit and Agilent Technology on the RNA Nano assay were that only slight accumulation of RNA would occur in the 5S region. Degradation of RNA was ruled out, because if this had occurred, the complete baseline of the electropherogram would have been increased and the peak resembling the 23S rRNA would be as high as or lower than the 16S rRNA peak, depending on the extend of degradation.

The theory that DNA contamination accounts for the large 5S peak, was tested by including a DNA degradation step into the RNA isolation protocol; although an offset peak connected to 5S peak in question could be eliminated, this was not the case for the main 5S peak. This result indicates that without an extra DNA degradation step, the RNA is slightly contaminated with DNA, but that the unexpected 5S peak is not a result of DNA contamination.

The electropherogram of RNA isolated from bacterial cell cultures using the ZR Soil/Fecal RNA MicroPrep kit was very similar to that of RNA isolated from fecal samples, ruling out the hypothesis that the 5S peak is something specific to fecal samples. Interestingly, RNA isolation from feces collected from germ-free mice also displayed an accumulation of molecules similar to 5S rRNA in size, which lead to the hypothesis that the large amount of small RNA molecules present in the total RNA isolated from fecal samples might be a combination of 5S rRNA and tRNA. As the 5S peak is lower, in comparison to the 16S and 23S peak, in samples isolated from pure bacterial culture than when isolated from fecal samples, the RNA within the 5S rRNA region might partially originate from the host and hasn't been degraded due to the tertiary structure.

4.1.4.2   *Storage of samples for transcriptome analysis*

Factors such as the housing of animals in the animal facility and time required to sampling multiple animals, makes it impossible to start isolating RNA immediately after sample collection. It was therefore important to determine storage conditions which would have the least influence on RNA quality. As a reference, RNA was isolated from a sample collected from a mouse in the animal facility immediately stored on ice and transported into the laboratory. Simultaneous sampling of multiple mice minimized the time between sample collection and RNA isolation and storage on ice reduced possible RNA degradation during transport. This is the most optimal way feasible for RNA isolation from fresh fecal samples. The RNA retrieved from this reference sample was compared to RNA isolated from samples stored in RNAprotect Bacteria for the transport from the animal facility to the lab or natively frozen at −80 °C as fast as possible after collection or stored in RNAprotect and frozen at -80 °C as fast as possible. Freezing samples prior to RNA isolation can minimize the variability

between samples isolated in different runs, because samples can be collected and isolated simultaneously. The RNA of the reference was comparable to the RNA isolated from samples that were stored or frozen in RNAprotect Bacteria prior to RNA isolation, indicating that these storage conditions do not substantially impact RNA quality. RNA isolated form fecal samples that were natively frozen however, was degraded. For metatranscriptomic analysis RNA should therefore either be isolated immediately after sample collection or stored in RNAprotect Bacteria until the RNA can be isolated.

4.1.4.3   RNA stability

The stability of the RNA was assessed by analyzing the same RNA sample immediately after isolation and after one cycle of freezing and thawing on ice, which was enough for slight degradation to be apparent. This was indicated by the reduction of the peak height of the 23S rRNA compared to the 16S rRNA peak and the increase in short RNA strands represented by the increased overall baseline level and an even larger peak in the 5S rRNA region. RNA samples used for transcriptomic analysis should therefore ideally be reverse transcribed immediately after isolation or gone through a maximum of one freezing and thawing cycle. In both cases it is important to treat all the samples that will be compared in the same way to ensure comparability.

4.1.5   Library preparation and 454-sequencing of 16S rRNA amplicons

The composition of a bacterial community can be analyzed on the basis of 16S rRNA gene sequences. Depending on the variable region(s) chosen, diversity estimates and differentiation of bacteria varies [202,203]. The more sequencing information is included into an analysis, the more accurate the results are. The 454 sequencing technology was chosen due to its superior read length at the beginning of the study; technologies such as Illumina however, have improved in the meantime and now also deliver long sequencing reads.

The literature is not consistent in regards to the discriminating power of the different hypervariable regions, although v1-4 (single, or combination of two variable regions) are most often mentioned favorably for this application [202–206]. Due to the difference in results obtained depending upon the hypervariable regions that are analyzed, we would suggest, that data comparison should be done cautiously and preferably only with data obtained from the same variable regions. For our study, primers spanning the v3 – v6 regions were used to create amplicons, which then were sequenced one-sided, starting at the v3 hypervariable region, on a GS FLX. The mean length of the sequencing reads after removal of the barcode and linker sequences and quality trimming, was approximately 300 bp [177]. The classification of our reads was therefore mostly based on the v3 and v4 hypervariable regions. Although Roche is shutting down the 454 platform [207], with the increased read length achieved by the Illumina platforms, the combination of v3 and v4 has become one of the standard regions examined for analysis of bacterial community compositions [208–211].

As previously mentioned, the 454 sequencing technology was chosen due to the long sequencing reads that were created (up to 700 bp) compared to the much shorter Illumina reads (2 x 150 bp for paired end reads in 2012) [149]. Taxonomic classification using 16S rRNA

gene analysis increases in accuracy with sequences length and available sequence information in reference databases. As the main goal of the project was to analyze the bacterial community composition, it was most important to be able to rely on the taxonomic assignment, therefore read length was prioritized. One disadvantage of 454-sequencing is the low coverage; at the time of the wetlab experiments, Illumina produced up to 600 Gb per run whereas 454 could only sequence 14 Gb per run [149]. The coverage, however, can easily be adjusted by varying the amount of samples pooled for sequencing or even by dividing one sequencing library and sequencing it in two runs, but this significantly increases cost.

Both of these sequencing technologies are based on the sequencing-by-synthesis principal. Illumina detects the individual fluorescent label of the incorporated nucleotide, whereas Roche detects light signals. Incorporation of nucleotides during synthesis, leads to release of pyrophosphate which is transformed into ATP by sulferylase and subsequently converted into light by luciferase. The 4 nucleotides are added to the reaction one after the other. The number of identical nucleotides added to the synthesized strand is not limited, and each additional incorporation increases the light signal. Properties of the light signal therefore limit the ability of the platform to detect the addition of multiple nucleotides in one cycle. Homopolymer stretches therefore lead to sequencing errors [140,212–214].

4.2 Microbiome/Metagenome analysis

4.2.1 Selection and impact of metadata

After the launch of the first NGS Platform by the company 454 in 2005, an increasing number of studies have been performed analyzing various microbiotas in health and disease [64,65,79,89,215–217]. Although the microbiota of various body sites has been analyzed, the intestinal microbiome composition has been most commonly associated with disease. Most studies have been conducted by comparing samples from healthy people to patients affected by diseases of interest. The detected differences therefore do not give any insight into whether the microbiota variations are cause or result of disease. In humans the microbiota would have to be monitored over long periods of time, and it is very hard to predict how many of the individuals participating in a monitoring study will ever develop a disease of interest. Animal models with a defined induction of disease, e.g. T-cell transfer or the application of chemicals for IBD induction [130,132–135], give the opportunity to define the microbiota in a healthy initial state, during development and after manifestation of disease. By analyzing the microbiome during the different disease stages, it is possible to determine whether changes are the cause or the result of disease. Large inter-individual differences of the intestinal microbiota exist, therefore examination of the same individual during different phases of health and disease helps to minimize artificial variations of the data resulting from strong inter-individual differences [4]. When studying the intestinal microbiome, there are two main sources for acquiring samples: populations of microbiota associated with mucosa can be assessed by taking biopsies, or microorganisms associated with both mucosa and lumen can be assessed through the analysis of fecal samples [34]. Protocols were established for

microbiome analysis from fecal samples to minimize the burden on the study animals; non-invasive sampling is a major advantage when repeated sampling must be done on the same individual.

Many environmental factors, such as diet, pharmaceutical intervention, age, stress or infection, are believed to play an important role in determining the composition of the intestinal microbiome [218,219]. It has been proposed that these environmental conditions are more important in forming the microbiome than genetic background, as the microbiota of mono- and dizygotic twins is equally similar [35].

A wide variety of metadata were collected for the animals included in the study. We included as many factors as possible to test for correlation with changes of the microbiota composition but also to exclude redundant information to keep the analysis as simple as possible. As some factors imply others, for example all animals housed in the same facility receive the same food, these were only recorded as one metadata "housing facility". Continuous metadata, as for example the number of living cells isolated from the caecum and colon, were combined into groups with similar values for simplified distinction between the animals.

The influence of metadata on the microbial composition was assessed by analyzing whether metadata correlated with the clustering of the data. The most obvious correlation was established between clustering and animal housing facility, indicating that this had the greatest influence. This metadata includes the environment of the housing facility, components of the food and water, hygiene state and personal working with the animals, which may induce stress and in turn influence the microbiome [220]. After separation of the data according to the animal facility in which animals were housed, clustering of the data correlated with the experiment number. As all experiments ran sequentially, this indicates that the environmental conditions changed between the experiments. The strong variation between the experiments makes it impossible to compare the results of the different runs, because similarities due to treatment protocols were concealed by the strong inter-experimental variation. Despite the strong variations of the gut microbiome, the clinical outcome was the same in the different experiments. This indicates that not only one certain composition is responsible for triggering colitis. Also no single bacterial species could be associated with disease onset, indicating that something other than the bacterial community composition is involved in disease development.

The human microbiome consortium has determined that large variations seen on a taxonomic level do not necessarily imply that the biological functions carried out by the microbiome differ [4]. As different bacteria can carry out identical (metabolic) functions this could be an important fact in colitis development. Analyzing the molecular functions of the microbiome may result in less variation between the mice of the different experiments which in turn could lead to identification of factors of the microbiome that influence colonic inflammation.

4.2.2 Differences between PCoA (<u>P</u>rincipal <u>Co</u>ordinates <u>A</u>nalysis) results computed by Qiime and Megan:

The first analysis of the sequencing data, using Qiime for PCoA, revealed differences between the Endo^{hi} and Endo^{lo} mice (Figure 27). A separation of the samples collected from the animals according to the animal facility where they were housed, was apparent. The clustering of samples within Endo^{hi} or Endo^{lo} animals indicated that they were more similar within one group than the groups were to each other. The PCoA analysis also displayed three sub-clusters of Endo^{hi} animals. This finding suggested that the microbiota of some animals were more similar to each other than others. The PCoA analysis created using Megan, however, did not show the same clustering patterns (Figure 27). The compositions of the intestinal microbiome of Endo^{lo} animals were not more similar within this group than compared to Endo^{hi} animals. Different factors are responsible for the contrasting results; two factors with large impact on the results are:

1. Qiime PCoA analysis creates plots including the 3 principal components responsible for the largest variation between the samples. Our Megan PCoA only considered two principal components simultaneously for calculating the similarity of multiple samples.

2. The sequencing reads are grouped into OTUs (Orthologous Taxonomic Units) according to sequence similarity and a phylogenetic tree is built from this information. The Qiime PCoA analysis was calculated from the OTUs, which contain sequences that were 97 % identical. Phylogenic trees were created, and the distances between two samples were then calculated. In contrast, the PCoA analysis performed by Megan was based on a taxonomic tree built after taxonomic binning of the sequencing reads [177]. Sequences grouped into different OTUs were not necessarily associated with different taxonomies and were therefore considered to be the same (as described in Figure 81) in the Megan analysis. The basis for calculation of the similarities between samples is therefore different for the Qiime and Megan analysis and can lead to different results.
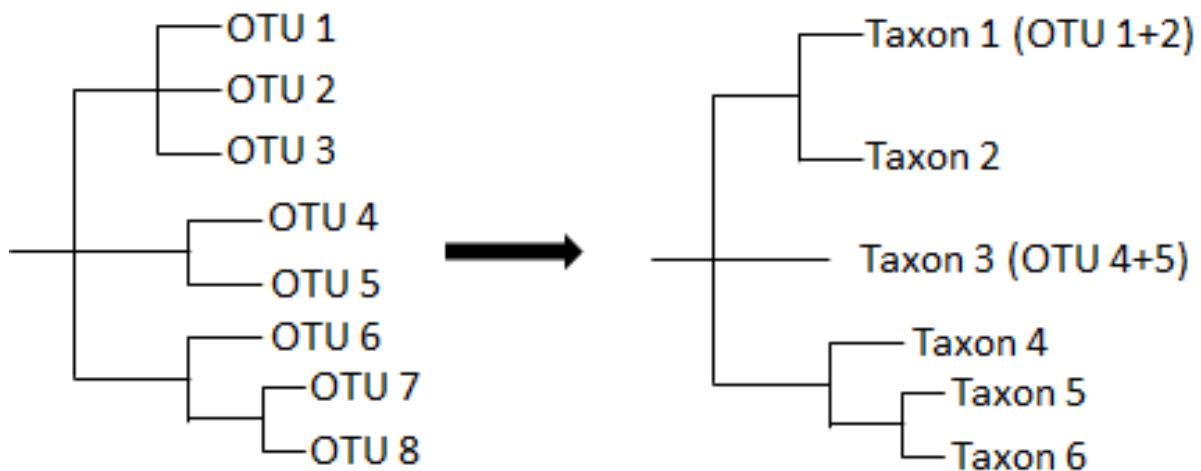
Figure 81: Qiime uses a OTU based phylogenetic tree especially created for the sequencing data for PCoA whereas PCoA calculated by Megan is based on a taxonomic association. As different OTUs may be associated to the same taxon, the phylogenetic trees used for calculation of similarities can differ.

Multivariate analyses are widely used for cluster analysis and comparing microbiota compositions [13,36,90,194,221–223]. It has to be kept in mind however, that the relationships between large amounts of datasets are simplified and projected onto only two or three dimensions, which leads to loss of information.

Metagenomic analysis addresses the question of which differences are responsible for deviating clinical phenotypes by correlating taxonomic and functional information with sequencing data. The goal is to determine which microorganisms are responsible for disease, and/or which microbiome-specific physiologic functions are missing, acquired, or increased in the microorganisms and influence the patients' health.

As the revealed community structure and function analysis strongly depend on the analysis methods used and on how the underlying data has been processed, the analysis pipeline must be adapted to the problem one wants to solve.

For analyzing which metadata factors influence the observed clustering of our data we decided to use the PCoA created by Qiime. Qiime has two properties that guided our decision: first, it is specific to our experiment, because the phylogenetic trees were created using our actual sequencing data, and second, three principal components were considered simultaneously. As more components were considered, we received a more refined clustering.

4.2.3   Impact of metadata on clustering:
Analysis of the microbiome prior to T-cell transfer showed that the separation of Endo^{hi} and Endo^{lo} mice is most prominent. The data did not cluster according to other factors, such as treatment, sampling time or disease outcome. After selection of samples belonging to the same cluster (e.g. Endo^{hi}), however, clustering according to the experiment number is apparent. This demonstrates that, if one wants to detect components of the microbiome

important for health and disease, it is important to select data that is comparable according to the metadata and thereby simplify analysis and reveal otherwise obscured results.

After analyzing all varieties of metadata in our experiment it became apparent that animals first clustered depending upon their housing facility, and then according to the experiment in which they were included. This indicates that the differences due to treatment are less than differences related to environmental factors between the experiments themselves. This was unexpected and undesirable, as it meant that differences between groups in different experiments could not be compared. The large number of experiments in this study could not be conducted simultaneously. Because of the sequential nature of the experimental set up, animals in the different experiments may have been exposed to different environmental conditions leading to changes of the composition of the intestinal microbiota. Our results indicate, that these differences between the experiments impact the microbiome to a larger extend than the different types of treatment.

Accepting that the environmental influence of housing was the most predominant factor, this variable was eliminated from analyses by only comparing data of animals within the same experiments. This would allow prominent influences from other metadata to be unmasked. Although this decreased effective sample sizes, analyses of the remaining variables of interest, such as treatment group, included only the samples collected from animals of the same or a comparable experiment that had clustered in the first analysis.

Our data show that, for microbiome analysis, it is important to run the experiments of groups one wants to compare to each other, simultaneously. This measure will reduce the influence of environmental factors on the results and differences seen between groups are more likely to be related to disease outcome. The large variations of the microbiome composition of individuals indicate that large group sizes are necessary to distinguish between physiologic and causative differences. If experiments can't be performed in parallel (e.g. due to a limited number of available animals), it is important to verify that the results are comparable before analyzing them. This can be done for example by comparing the initial microbiomes of animal belonging to the same group but run in separate experiments.

A second measure that can be taken to reduce differences between animals is using offspring of the same mice. As the microbiome is transferred from the mother to the newborns [224]and they share a cage for the first 3 weeks, siblings should be inoculated and exposed to the same microorganisms.

4.2.4   Bacterial gut community and its effect on the host

Clinical experimental outcome was classified on the basis of the histological score (Table 17). As expected Endo^{hi} Rag1$^{-/-}$ mice developed colitis after reconstitution of the T-cell population, while Endo^{lo} animals remained healthy and displayed no signs of inflammation. The PCoA analysis distinguished two clusters for the Endo^{hi} animals. One cluster consisted of samples collected prior to T-cell transfer, and the other included the samples collected from the same mice later during the experiment. This revealed that the bacterial composition of

the animals developing colitis seemed to change drastically within the first 3 weeks post T-cell transfer. A new stable bacterial community was established in the inflamed gut, which did not continue to change during the remainder of the experiment.

In contrast, the intestinal microbiome of the Endo^{lo} animals showed only slight variations 3 weeks after T-cell transfer, and returned to a composition very similar to the initial one by the end of the study. These results indicate that the microbiome of Endo^{lo} animals is stable, specific for healthy animals, and promotes the maintenance of homoeostasis. The slight variation of the microbiome during the first three weeks of the experiment might be the result of factors such as the following:

1. Stress can influence the bacterial colonization of the intestine [225]. The process of T-cell transfer is very stressful for the animals, which could impact the composition of the microbiota. This potential stress reaction variable could be investigated in future studies by analyzing changes of the intestinal microbiome of mice as a result of exposure to stress.

2. The host immune system and the bacteria colonizing the intestine interact with each other [33,34,64]. As Rag1$^{-/-}$ mice lack B- and T-cells, they do not have a functional adaptive immune response. With the introduction of T-cells, immune reactions that could not occur prior to T-cell reconstitution, could then take place. This interaction could result in the slight shift of the microbiota until a stable state was reached, than the original composition resurfaces.

Published studies associating diseases with certain microbial compositions usually cannot answer whether the variations between healthy and diseased individuals are the cause or result of disease due to the experimental setup (no longitudinal analysis, just a comparison of healthy and affected individuals) [226].

Our work demonstrates that genetically identical mice react differently to transfer of naïve T-cells depending upon the composition of the intestinal microbiome. The outcome of our experiments suggests that the intestinal bacteria play an important role in colitis induction. This hypothesis is strengthened by further experiments showing that by changing the composition of bacteria colonizing the mouse intestine, the development of colitis after T-cell transfer can be induced or prevented. This is reciprocal: the changes of the bacterial composition occurring during disease development also indicate that disease affects the microbiome. This is not surprising, as bacterial fitness and survival depend highly upon environmental conditions, such as an altered pH, which change vastly in an inflamed gut [51,227].

The PCoA analysis of all analyzed samples showed that in some cases the microbiome of animals developing colitis is very similar to that of animals maintaining homoeostasis, and in other cases animals with the same disease state have very different microbiota. This indicates, that the composition determined by our 16S rRNA gene sequencing assay is not sufficient to classify a given microbiota as colitogenic or not.

One possible explanation for the apparent similarity of the microbiomes of animals which will develop colitis and those which will maintain homoeostasis may simply be a result of the sequencing not being deep enough. That is, there may not be sufficient amounts of sequencing reads per sample to detect very low abundance species, which can greatly impact the host [228]. If this is the case in our model, then the bacterial species responsible for disease development might be under the detection limit and was not found in the analysis. Deeper sequencing could give more insight into this issue. This can either be done by pooling fewer samples for a 454 sequencing run, or switching to the Illumina system which produces much more sequences per run and with advances in technology is now able to produce longer read lengths as well.

One difference seen in the microbiome of Endo^{hi} and Endo^{lo} mice is the presence of *Akkermansia*. One species, *Akkermansia muciniphilia*, is known to degrade mucus [229]. One could assume that this would negatively affect the barrier function of the gut, as the mucus is important for preventing intestinal bacteria from infecting the host. However, the presence of *Akkermansia muciniphila* has been linked to healthy individuals. The reason for this might be that this bacterial strain produces high amounts of butyrate, which is used as an energy source by the epithelial cells and therefore promotes barrier integrity of the gut[85,229,230].

Due to the great complexity of the samples, for group comparisons the dimensions had to be reduced by PCoA. Only the three components accounting for the greatest variations among samples were considered for analyzing and grouping the samples. This analysis is therefore based on the assumption that the experimental outcome is a result of the components responsible for the greatest variations between the samples. If only two factors could possibly influence clustering of samples, a graph showing both factors could be created and clustering would be apparent. Due to the high complexity of the microbiome data however, a simple visualization of the data is impossible. Therefor we have to except a tradeoff: For visualization and cluster analysis, dimensions have to be reduced, therefore not all factors are considered, and some important variables clustering the data into the desired groups may be lost if they do not account for high variability.

Feeding *E. coli*$_{WT}$ to Endo^{lo} animals prior to T-cell transfer resulted in the development of disease. We demonstrated that the endotoxicity of the intestinal microbiom is important for colitis induction. The wild-type *E. coli* strain is known to be highly endotoxic [183], suggesting, that the feeding of this bacterial strain increases the endotoxicity of the microbiome to a level which leads to pathology. On the other hand, treating Endo^{hi} animals with the low endotoxic *E. coli*$_{MUT}$ strain prevents colitis, suggesting that the overall endotoxicity is lowered and does not induce inflammation. Importantly, the results were confirmed after feeding mice purified wild-type and mutant LPS. The intestinal microbial community composition of mice developing disease was not the same in all the mice treated for colitis induction, and no bacterial strain could consistently be linked to inflammation. This leads to the hypothesis, that the overall endotoxicity of the different communities may predispose mice to disease,

and be higher in mice prone to disease compared to those maintaining homeostasis. This thesis is supported by the fact that feeding purified LPS with different endotoxicities is sufficient to induce or prevent inflammation. Whether the molecules directly interact with the host or if they influence the gut community, which then in turn influences the host, however, still remains to be elucidated.

Wild-type and mutant *E. coli* strains can both be detected in Endo[hi] animals on the day of T-cell transfer. Whereas sequencing reads assigned to the genus of *Escherichia* are detectable at all three time points in samples collected from the mice fed with the wild-type strain, the samples from Endo[hi] animals fed with the mutant strain contained no *Proteobacteria* reads, but from week 3 on many *Bacteroidetes* reads were found.

Before T-cell transfer, the Endo[hi] animal were treated with antibiotics and subsequently fed with *E. coli*[WT] and *E. coli*[MUT]. Antibiotics reduce the amount of native bacteria in the gastrointestinal tract, leading the introduced bacteria that are fed to the mice to be present in the gut in a high proportion and can therefore be detected in the 16S microbiome analysis. The amount of *Escherichia* detected in Endo[hi] mice fed with *E. coli*[WT] increases during the first 3 weeks. This can be ascribed to the constant feeding of *E. coli*[WT] in combination with growth of introduced bacteria that have colonized the gut. After week 3, the proportion of reads associated with *E. coli* was reduced although the amount of bacteria that were fed to the mice did not decrease. At the same time, reads were correlated to Tenericutes and other bacteria that were not previously detected. The evident growth of these bacteria is either a result of the changing conditions during inflammation of the gut or due to regeneration of native bacteria after antibiotic treatment. Another possibility is that the growing amount of *Escherichia* in the gut directly or indirectly positively influenced the growth of these other bacteria. As the sequencing reads only describe the proportions of bacterial species within the gut and not the absolute amount of bacteria, this decrease in reads associated with *Escherichia* does not imply that the absolute amount of *Escherichia* also decreases after week 3.

As analysis by NGS detects all bacteria present within a sample, it could be argued that the *Escherichia* detected are only from fed bacteria that are passing through the gut, not from bacteria colonizing the gut. If this were the case, however, we would expect to see the same amount of reads belonging to *Escherichia* in the group of animals fed the mutant strain. As this is not the case, it is assumed that in contrast to the mutant strain, *E. coli*[WT] was able to colonize the intestine of the Endo[hi] mice.

Other groups have described differences of the microbiota during IBD and in a healthy state [118–120]. Most described differences were correlated to *Firmicutes* and *Bacteroidetes*. Manichanh et al., noted a loss of diversity in Crohn's disease, which was mainly due to far fewer OTUs associated to the *Firmicutes* phylum [118]. Our results did not indicate loss of diversity resulting from inflammation. An equal number of genera were detected in mice prior to and during inflammation. It is possible that in order to see an overall loss of

diversity, the experiments would have had to be run over a longer time period after onset of inflammation. Environmental conditions within the gut change during disease, and while many bacteria can survive, some bacteria are better equipped than others to cope with an altered environment. Selection of these bacteria and their subsequent expansion in turn can alter the conditions for the remaining bacteria. These factors can lead to a less diverse microbiome over time. Even if a reduction of diversity has taken place, this might not immediately be visible by NGS analysis, because this method does not distinguish between DNA isolated from viable or dead bacteria. Therefore bacterial species can be detected, even if they are no longer an active part of the intestinal microbiome and until all bacterial residues have passed through the gut, the diversity detected by NGS will be overestimated.

Some publications have connected colitis development with the genera of *Bacteroides or Bifidobacterium*, as well as single bacterial species such as *F. prausnitzii* [77,124,125,231,232]. In our study, no bacterial species or groups could be consistently linked to development of colitis. Changing the initial composition of the intestinal microbiota, however, resulted in the prevention or induction of colitis. This indicates that the intestinal microbiota plays a crucial role in colitis development and maintenance of homeostasis. It is possible that not one strain of bacteria is responsible for the development of inflammation, but instead combinations of bacteria that were not recognized as being involved in this process as a group. As the endotoxicity of the gut microbiome seems to be important, it is possible that a collection of high endotoxic bacteria are increased in Endo^{hi} mice compared to Endo^{lo} mice, but this is not statistically significant for the single bacteria and therefore not evident in our analysis.

The great differences seen between data described in different publications and our own must be compared with caution. Sequencing results are dependent upon the sample handling and methods used for DNA extraction, sequencing and analysis [233]. Selection bias may be introduced at various levels: different DNA extraction methods extract DNA better from certain microorganisms than from others [180,194], detected phylogenetic association varies depending upon the variable region of the 16S rRNA gene analyzed [203], and OTU grouping is influenced by the selected algorithms and thresholds and taxonomic classification depends on the information stored in the database used as reference.

Due to the preparation of the DNA for sequencing (e.g. PCR prior to 454 and Illumina-sequencing) and the specifications of the different sequencing techniques and chemistries (Table 2), the sequencing results vary between the various sequencing platforms. This can also lead to different taxonomic classification during subsequent database comparison.

4.2.5 Influence of the host on bacterial gut community

The changes of the microbial composition during colitis development (Figure 32) indicate that inflammation impacts the composition of the microbial community within the intestine. This shows that, when comparing the microbiomes of healthy individuals and those suffering from IBD, the microbiome differences are not likely to be the cause of disease. We would suggest that the differences are due to a combination of an initial microbiota that promotes inflammation, which in turn induces changes of the bacterial gut community. In order to

identify microbial compositions capable of triggering onset of disease it is therefore necessary to perform longer-term studies. Longitudinal studies must include enough subjects that some will potentially develop disease. Studying the microbiome during different stages of disease (prior to disease development as well as during its onset and progression) can allow the identification of microbial communities involved in promoting or preventing disease. If certain community members or functions can be identified, microbial screenings as described for colorectal cancer by Zackular et al. [234] can be utilized for improving non-invasive preventive medical examination.

4.3   Functional microbiome analyses – Metagenomics & -transcriptomics

Analysis of the bacterial composition was used for detection of differences within the intestinal bacterial population in different health conditions. We saw large inter-individual variability of our 16S metagenomics data and no certain bacterial species could be linked to induction of colitis or manifested disease. The Human Microbiome Consortium found that variation of the community members within the microbiota does not necessarily lead to differences within the (metabolic) functions of the bacteria [4]. Some functions are necessary for survival and thus are carried out by all bacteria. Other genes are "nice to have" because they lead to increased fitness of the microbes in certain environmental conditions. As genes can be transferred between different organisms, the phylogenetic composition cannot be correlated to biological tasks that are carried out. Additionally, Arumugam et al. [36] found that high abundance functions can be performed by a variety of low abundance species. Even if these functions were unambiguously connected to defined bacterial species, the limit of detection would have to be very low to identify those species as a driving force.

Methods that are available for functional analysis include shotgun metagenomics and metatranscriptomics. The genomic DNA of the microbiota gives information on which biological tasks can potentially be performed by the members of the microbiota; studying the RNA answers the question which genes are in use. The levels of mRNA and regulatory RNAs, such as miRNA or siRNA indicate up- or down regulation of genes.

The intestinal microbiota clearly plays an important role in colitis induction. However, the 16S metagenomics data do not indicate a specific bacterial species to be responsible for disease development, suggesting that the crucial difference is on the functional level.

Besides the aforementioned differences between shotgun metagenomics and metatranscriptomics, the stability of the DNA and RNA molecules influence the results. Most RNA molecules are very unstable, which both complicates working with the RNA and limits the detection of residual RNA molecules from prior disease stages. The stability of DNA isolated from dead bacteria that had not yet cleared the mouse intestine, but were no longer able to interact with the host could contribute to the similarity of the 16S metagenomics results of treated and untreated mice housed under identical conditions.

Definition of functions carried out by the intestinal microbiome and its interaction with the host in health and disease is an important step on the way to prevent and heal IBD. Functional analysis can therefore help to identify targets for IBD treatment.

4.4 Verification of sequencing data

In this project we used FISH to verify the sequencing results with respect to *Bacteroides* and Enterobacteriaceae. FISH was chosen because it is independent of PCR and can directly detect various bacterial strains without prior cultivation. However, differences were often seen between the sequencing and FISH results.

According to the sequencing results, the intestinal microbiomes of T-cell transferred Endo^{hi} animals and *E. coli*$_{WT}$ treated Endo^{lo} animals contain Enterobacteriaceae, but these were not detected by FISH. Innate properties of the technique mean that the identification of single bacteria by FISH is only possible if the bacteria are separated from each other, but this reduces the number of bacteria that are analyzed by this method. The lower the number of bacteria analyzed, the lower the probability of detecting rare bacterial species. The sequencing data showed very low read counts for Enterobacteriaceae in the T-cell transferred Endo^{hi} animals and *E. coli*$_{WT}$ treated Endo^{lo} mice, indicating low amounts of Enterobacteriaceae specific 16S rRNA genes. As the detection limit of NGS (depending on sequencing depth) is lower than that of FISH, it is not surprising that rare bacteria were only detected by NGS.

A second issue that might account for the differences found between the amplicon sequencing results and FISH could be that by FISH, only viable bacteria are detected. Our probes target 16S rRNA, which is distributed in the cytoplasm of the bacteria, allowing the bacteria to be visualized using a fluorescence microscope. Dead bacteria that no longer contain enough intact rRNA are not detected. DNA, however, is much more stable than RNA, and it is therefore possible that sequencing also detects sequences from dead bacteria and they are treated as part of the current intestinal microbiota. Bacterial DNA that was ingested by the mice and not degraded enough while passing the GIT would also be detected by NGS but not by FISH.

Hybridization of the *Bacteroides* probe was detected for all samples collected from Endo^{lo} mice, irrespective of treatment. The NGS results, however, did not confirm the FISH results. No sequencing reads were assigned to the genus *Bacteroides*, however all the samples contained high amounts of sequencing reads that were mapped to the order Bacteroidales. The mapping of the reads to the corresponding taxon was performed using a LCA (lowest common ancestor) algorithm. If a read mapped to two different reference strains with a similar score, the read was mapped to the taxon that is the lowest common ancestor of the two reference strains. The NGS results therefore do not necessarily mean that the sample did not contain *Bacteroides*, but that *Bacteroides* sequences could not be clearly distinguished from sequences of other bacteria belonging to the family Bacteroideaceae.

The differences between the results of the two methods may also be method unrelated; the bacteria are not uniformly distributed within the gut [19–21], therefore it is likely that the microbial composition varies as the feces passes through the intestine. Instead of being collected as defecated pellets, the fecal samples collected in week 6 were taken directly from the intestine after the mice were sacrificed. Samples collected for the first two time points for NGS had passed through the gut completely, whereas samples for FISH were extracted from more proximal parts of the gut. The more or less incomplete passage of the fecal samples used for FISH could therefore also be responsible for variations between the NGS data and the FISH results.

FISH was selected to verify our NGS results because it is a PCR independent method and no cultivation of bacteria is necessary. The differences of NGS and FISH results, however, indicate that FISH may not be suitable as a reference. Using just these two methods is not sufficient to discriminate if the results differ due to technical reasons or due to false bacterial identification of NGS. In addition, FISH can only verify the presence or absence of defined bacterial strains, whereas amplicon sequencing will also detect unknown bacteria.

Due to technical differences FISH can be used to verify that bacteria detected by amplicon sequencing are viable and present in a sample; however, if bacteria are not detected by FISH this does not necessarily imply that amplicon sequencing produces false positive results. Cultivation of bacteria could verify presence of rare bacteria but this method is limited to bacteria that can be cultivated in the lab. Another possibility to verify the results of amplicon sequencing would be by PCR. However, as our NGS approach is also based on PCR, the methods could be biased in the same way. These methods can only be used to identify known bacteria; verification of the sequencing results with these methods is therefore limited. FISH and other methods can be used to verify certain parts of the bacterial composition, but their limitations have to be considered as none of them can completely verify the sequencing results. In some cases, a combination of different methods might be the best strategy. It also has to be recognized that NGS is not perfect and therefore not only limitations of other methods are responsible for variations of results, but that errors of amplicon sequencing and subsequent bioinformatic analysis can also account for differences.

4.5 Conclusions

The protocols developed here can be used for analyzing the microbiome from fecal samples using next-generation sequencing. However, this work also shows that different measures should be taken to increase the impact of these kinds of studies. First, the experiments for microbiome analyses should be done in one large study rather than multiple, serial, smaller experiments to reduce environmental influences on the results. Cohousing of animals prior to the experiments, irrespective of the later treatment, may also help to minimize differences in the native microbiota between the animals of individual cages and increase the chance of identifying responsible microorganisms.

Using siblings, or offspring of sisters can minimize the maternal effect and reduce differences between the individual animals. Housing the animals in isolators, can also contribute to a maintaining the microbiota of the animals as the influence of environmental factors is limited. It is also important to ensure that all the metadata are represented by a sufficient amount of animals in order to analyze the influence of the different aspects. As some may not be predictable, it is important to include enough extra individuals at the same time however, consider if the prospective benefit justifies larger cohort sizes.

As the sequencing method and chemistry also strongly impact the results, it is important to ensure, that these do not change during the course of the experiments. One should also make sure that the same reference data sets were used for all the data that will be compared.

If microbiome analysis is done for the first time on a new type of sample and the magnitude of number of microorganisms or genes is unknown it is helpful to perform a pilot study with a small amount of individuals. Analyzing the rarefaction curve of the sequencing data will indicate, how many reads per sample should be sequenced or in other words how deep sequencing should be.

Bacteria are not the only microorganisms that may be of interest, when correlating a microbiome to different phenotypes. In addition to this, the functions the microbiota may also be important for disease development and they are not necessarily correlated to the microorganisms present. In many cases it is therefore advisable to perform shotgun sequencing instead of 16S analysis. This will also give a better resolution of the data, because bacteria can usually be identified down to the species or even subspecies level whereas 16S analysis can normally only identify bacteria down to the genus level.

References

1. Pei, Z. *et al.* Bacterial biota in the human distal esophagus. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 4250–4255 (2004).

2. Verhelst, R. *et al.* Cloning of 16S rRNA genes amplified from normal and disturbed vaginal microflora suggests a strong association between Atopobium vaginae, Gardnerella vaginalis and bacterial vaginosis. *BMC Microbiol.* **4,** 16 (2004).

3. Zhou, X. *et al.* Characterization of vaginal microbial communities in adult healthy women using cultivation-independent methods. *Microbiol. Read. Engl.* **150,** 2565–2573 (2004).

4. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature* **486,** 207–214 (2012).

5. Ward, D. M., Ferris, M. J., Nold, S. C. & Bateson, M. M. A natural view of microbial biodiversity within hot spring cyanobacterial mat communities. *Microbiol. Mol. Biol. Rev. MMBR* **62,** 1353–1370 (1998).

6. Ward, D. M. *et al.* Biodiversity within hot spring microbial mat communities: molecular monitoring of enrichment cultures. *Antonie Van Leeuwenhoek* **71,** 143–150 (1997).

7. McCaig, A. E., Glover, L. A. & Prosser, J. I. Molecular analysis of bacterial community structure and diversity in unimproved and improved upland grass pastures. *Appl. Environ. Microbiol.* **65,** 1721–1730 (1999).

8. Cox, M. J., Cookson, W. O. C. M. & Moffatt, M. F. Sequencing the human microbiome in health and disease. *Hum. Mol. Genet.* **22,** R88–94 (2013).

9. Lederberg, J. 'Ome Sweet 'Omics-- A Genealogical Treasury of Words | The Scientist Magazine®. *The Scientist* (2001). at <http://www.the-scientist.com/?articles.view/articleNo/13313/title/-Ome-Sweet--Omics---A-Genealogical-Treasury-of-Words/>

10. NIH HMP Working Group *et al.* The NIH Human Microbiome Project. *Genome Res.* **19,** 2317–2323 (2009).

11. Savage, D. C. Microbial ecology of the gastrointestinal tract. *Annu. Rev. Microbiol.* **31,** 107–133 (1977).

12. Luckey, T. D. Introduction to intestinal microecology. *Am. J. Clin. Nutr.* **25,** 1292–1294 (1972).

13. Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308,** 1635–1638 (2005).

14. Stecher, B. *et al.* Like will to like: abundances of closely related species can predict susceptibility to intestinal colonization by pathogenic and commensal bacteria. *PLoS Pathog.* **6,** e1000711 (2010).

15. Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464,** 59–65 (2010).

16. Wopereis, H., Oozeer, R., Knipping, K., Belzer, C. & Knol, J. The first thousand days - intestinal microbiology of early life: establishing a symbiosis. *Pediatr. Allergy Immunol. Off. Publ. Eur. Soc. Pediatr. Allergy Immunol.* **25,** 428–438 (2014).

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

18. Gill, S. R. *et al.* Metagenomic analysis of the human distal gut microbiome. *Science* **312,** 1355–1359 (2006).

19. Costello, E. K. *et al.* Bacterial community variation in human body habitats across space and time. *Science* **326,** 1694–1697 (2009).

20. Hakansson, A. & Molin, G. Gut microbiota and inflammation. *Nutrients* **3,** 637–682 (2011).

21. Wang, M., Ahrné, S., Jeppsson, B. & Molin, G. Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes. *FEMS Microbiol. Ecol.* **54,** 219–231 (2005).

22. *Medizinische Mikrobiologie*. (Urban & Fischer, 2001).

23. Matsumiya, Y., Kato, N., Watanabe, K. & Kato, H. Molecular epidemiological study of vertical transmission of vaginal Lactobacillus species from mothers to newborn infants in Japanese, by arbitrarily primed polymerase chain reaction. *J. Infect. Chemother. Off. J. Jpn. Soc. Chemother.* **8,** 43–49 (2002).

24. Grönlund, M. M., Lehtonen, O. P., Eerola, E. & Kero, P. Fecal microflora in healthy infants born by different methods of delivery: permanent changes in intestinal flora after cesarean delivery. *J. Pediatr. Gastroenterol. Nutr.* **28,** 19–25 (1999).

25. Long, S. S. & Swenson, R. M. Development of anaerobic fecal flora in healthy newborn infants. *J. Pediatr.* **91,** 298–301 (1977).

26. Penders, J. *et al.* Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* **118,** 511–521 (2006).

27. Satokari, R., Grönroos, T., Laitinen, K., Salminen, S. & Isolauri, E. Bifidobacterium and Lactobacillus DNA in the human placenta. *Lett. Appl. Microbiol.* **48,** 8–12 (2009).

28. Martín, R. *et al.* Human milk is a source of lactic acid bacteria for the infant gut. *J. Pediatr.* **143,** 754–758 (2003).

29. Martín, R. *et al.* Isolation of bifidobacteria from breast milk and assessment of the bifidobacterial population by PCR-denaturing gradient gel electrophoresis and quantitative real-time PCR. *Appl. Environ. Microbiol.* **75,** 965–969 (2009).

30. Koenig, J. E. *et al.* Succession of microbial consortia in the developing infant gut microbiome. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl 1,** 4578–4585 (2011).

31. Yatsunenko, T. *et al.* Human gut microbiome viewed across age and geography. *Nature* **486,** 222–227 (2012).

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

33. Duerkop, B. A., Vaishnava, S. & Hooper, L. V. Immune responses to the microbiota at the intestinal mucosal surface. *Immunity* **31,** 368–376 (2009).

34. Walter, J. & Ley, R. The human gut microbiome: ecology and recent evolutionary changes. *Annu. Rev. Microbiol.* **65,** 411–429 (2011).

35. Turnbaugh, P. J. *et al.* A core gut microbiome in obese and lean twins. *Nature* **457,** 480–484 (2009).

36. Arumugam, M. *et al.* Enterotypes of the human gut microbiome. *Nature* **473,** 174–180 (2011).

37. Mazmanian, S. K., Liu, C. H., Tzianabos, A. O. & Kasper, D. L. An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell* **122,** 107–118 (2005).

38. Blumberg, R. & Powrie, F. Microbiota, disease, and back to health: a metastable journey. *Sci. Transl. Med.* **4,** 137rv7 (2012).

39. Van der Waaij, D., Berghuis-de Vries, J. M. & Lekkerkerk Lekkerkerk-v, null. Colonization resistance of the digestive tract in conventional and antibiotic-treated mice. *J. Hyg. (Lond.)* **69,** 405–411 (1971).

40. Rosenberg, E., Sharon, G., Atad, I. & Zilber-Rosenberg, I. The evolution of animals and plants via symbiosis with microorganisms. *Environ. Microbiol. Rep.* **2,** 500–506 (2010).

41. Ley, R. E. *et al.* Evolution of mammals and their gut microbes. *Science* **320,** 1647–1651 (2008).

42. Fava, F. & Danese, S. Intestinal microbiota in inflammatory bowel disease: friend of foe? *World J. Gastroenterol. WJG* **17,** 557–566 (2011).

43. Lee, J. *et al.* Maintenance of colonic homeostasis by distinctive apical TLR9 signalling in intestinal epithelial cells. *Nat. Cell Biol.* **8,** 1327–1336 (2006).

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

45. David, L. A. *et al.* Host lifestyle affects human microbiota on daily timescales. *Genome Biol.* **15,** R89 (2014).

46. Wu, G. D. *et al.* Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334,** 105–108 (2011).

47. Walker, A. W. *et al.* Dominant and diet-responsive groups of bacteria within the human colonic microbiota. *ISME J.* **5,** 220–230 (2011).

48. Cotillard, A. *et al.* Dietary intervention impact on gut microbial gene richness. *Nature* **500,** 585–588 (2013).

49. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444,** 1027–1031 (2006).

50. Verbeke, K. A., Boesmans, L. & Boets, E. Modulating the microbiota in inflammatory bowel diseases: prebiotics, probiotics or faecal transplantation? *Proc. Nutr. Soc.* **73,** 490–497 (2014).

51. Round, J. L. & Mazmanian, S. K. The gut microbiota shapes intestinal immune responses during health and disease. *Nat. Rev. Immunol.* **9,** 313–323 (2009).

52. Aroniadis, O. C. & Brandt, L. J. Intestinal microbiota and the efficacy of fecal microbiota transplantation in gastrointestinal disease. *Gastroenterol. Hepatol.* **10,** 230–237 (2014).

53. *Immunobiology: the immune system in health and disease ; [animated CD-ROM inside]*. (Garland Publ. [u.a.], 2001).

54. Atuma, C., Strugala, V., Allen, A. & Holm, L. The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. *Am. J. Physiol. Gastrointest. Liver Physiol.* **280,** G922–929 (2001).

55. Johansson, M. E. V. *et al.* The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 15064–15069 (2008).

56. Alexander, K. L., Targan, S. R. & Elson, C. O. Microbiota activation and regulation of innate and adaptive immunity. *Immunol. Rev.* **260,** 206–220 (2014).

57. Mukherjee, S., Vaishnava, S. & Hooper, L. V. Multi-layered regulation of intestinal antimicrobial defense. *Cell. Mol. Life Sci. CMLS* **65,** 3019–3027 (2008).

58. Meyer-Hoffert, U. *et al.* Secreted enteric antimicrobial activity localises to the mucus surface layer. *Gut* **57,** 764–771 (2008).

59. Macpherson, A. J. *et al.* A primitive T cell-independent mechanism of intestinal mucosal IgA responses to commensal bacteria. *Science* **288,** 2222–2226 (2000).

60. Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S. & Medzhitov, R. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* **118,** 229–241 (2004).

61. Macpherson, A. J. & Harris, N. L. Interactions between commensal intestinal bacteria and the immune system. *Nat. Rev. Immunol.* **4,** 478–485 (2004).

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

63. Maier, E., Anderson, R. C. & Roy, N. C. Understanding How Commensal Obligate Anaerobic Bacteria Regulate Immune Functions in the Large Intestine. *Nutrients* **7,** 45–73 (2014).

64. Bhargava, P. & Mowry, E. M. Gut microbiome and multiple sclerosis. *Curr. Neurol. Neurosci. Rep.* **14,** 492 (2014).

65. Murri, M. *et al.* Gut microbiota in children with type 1 diabetes differs from that in healthy children: a case-control study. *BMC Med.* **11,** 46 (2013).

66. Ivanov, I. I. *et al.* Specific microbiota direct the differentiation of Th17 cells in the mucosa of the small intestine. *Cell Host Microbe* **4,** 337–349 (2008).

67. Ivanov, I. I. *et al.* Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell* **139,** 485–498 (2009).

68. Smith, P. M. *et al.* The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science* **341,** 569–573 (2013).

69. Bainbridge, B. W., Coats, S. R., Pham, T.-T. T., Reife, R. A. & Darveau, R. P. Expression of a Porphyromonas gingivalis lipid A palmitylacyltransferase in Escherichia coli yields a chimeric lipid A with altered ability to stimulate interleukin-8 secretion. *Cell. Microbiol.* **8,** 120–129 (2006).

70. Hague, A., Butt, A. J. & Paraskeva, C. The role of butyrate in human colonic epithelial cells: an energy source or inducer of differentiation and apoptosis? *Proc. Nutr. Soc.* **55,** 937–943 (1996).

71. Hague, A., Singh, B. & Paraskeva, C. Butyrate acts as a survival factor for colonic epithelial cells: further fuel for the in vivo versus in vitro debate. *Gastroenterology* **112,** 1036–1040 (1997).

72. Peng, L., Li, Z.-R., Green, R. S., Holzman, I. R. & Lin, J. Butyrate Enhances the Intestinal Barrier by Facilitating Tight Junction Assembly via Activation of AMP-Activated Protein Kinase in Caco-2 Cell Monolayers. *J. Nutr.* **139,** 1619–1625 (2009).

73. Fernandes, J., Su, W., Rahat-Rozenbloom, S., Wolever, T. M. S. & Comelli, E. M. Adiposity, gut microbiota and faecal short chain fatty acids are linked in adult humans. *Nutr. Diabetes* **4,** e121 (2014).

74. Nuding, S., Fellermann, K., Wehkamp, J. & Stange, E. F. Reduced mucosal antimicrobial activity in Crohn's disease of the colon. *Gut* **56,** 1240–1247 (2007).

75. Cani, P. D., Osto, M., Geurts, L. & Everard, A. Involvement of gut microbiota in the development of low-grade inflammation and type 2 diabetes associated with obesity. *Gut Microbes* **3,** 279–288 (2012).

76. Furet, J.-P. *et al.* Differential Adaptation of Human Gut Microbiota to Bariatric Surgery-Induced Weight Loss: Links With Metabolic and Low-Grade Inflammation Markers. *Diabetes* **59,** 3049–3057 (2010).

77. Gomes, A. C., Bueno, A. A., de Souza, R. G. M. & Mota, J. F. Gut microbiota, probiotics and diabetes. *Nutr. J.* **13,** 60 (2014).

78. Ley, R. E. *et al.* Obesity alters gut microbial ecology. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 11070–11075 (2005).

79. Foster, J. A. & McVey Neufeld, K.-A. Gut-brain axis: how the microbiome influences anxiety and depression. *Trends Neurosci.* **36,** 305–312 (2013).

80. Arthur, J. C. *et al.* Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* **338,** 120–123 (2012).

81. Arthur, J. C. & Jobin, C. The struggle within: microbial influences on colorectal cancer. *Inflamm. Bowel Dis.* **17,** 396–409 (2011).

82. Waidmann, M. *et al.* Bacteroides vulgatus protects against Escherichia coli-induced colitis in gnotobiotic interleukin-2-deficient mice. *Gastroenterology* **125,** 162–177 (2003).

83. Kleessen, B., Kroesen, A. J., Buhr, H. J. & Blaut, M. Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. *Scand. J. Gastroenterol.* **37,** 1034–1041 (2002).

84. Darfeuille-Michaud, A. *et al.* Presence of adherent Escherichia coli strains in ileal mucosa of patients with Crohn's disease. *Gastroenterology* **115,** 1405–1413 (1998).

85. Maslowski, K. M. & Mackay, C. R. Diet, gut microbiota and immune responses. *Nat. Immunol.* **12,** 5–9 (2011).

86. Brown, C. T. *et al.* Gut microbiome metagenomics analysis suggests a functional model for the development of autoimmunity for type 1 diabetes. *PloS One* **6,** e25792 (2011).

87. Alam, M. Z., Alam, Q., Kamal, M. A., Abuzenadah, A. M. & Haque, A. A possible link of gut microbiota alteration in type 2 diabetes and Alzheimer's disease pathogenicity: an update. *CNS Neurol. Disord. Drug Targets* **13,** 383–390 (2014).

88. Chervonsky, A. V. Influence of microbial environment on autoimmunity. *Nat. Immunol.* **11,** 28–35 (2010).

89. Scher, J. U. & Abramson, S. B. The microbiome and rheumatoid arthritis. *Nat. Rev. Rheumatol.* **7,** 569–578 (2011).

90. De Goffau, M. C. *et al.* Fecal microbiota composition differs between children with β-cell autoimmunity and those without. *Diabetes* **62,** 1238–1244 (2013).

91. Vaarala, O., Atkinson, M. A. & Neu, J. The 'perfect storm' for type 1 diabetes: the complex interplay between intestinal microbiota, gut permeability, and mucosal immunity. *Diabetes* **57,** 2555–2562 (2008).

92. Yeoh, N., Burton, J. P., Suppiah, P., Reid, G. & Stebbings, S. The role of the microbiome in rheumatic diseases. *Curr. Rheumatol. Rep.* **15,** 314 (2013).

93. Ochoa-Repáraz, J. *et al.* Role of gut commensal microflora in the development of experimental autoimmune encephalomyelitis. *J. Immunol. Baltim. Md 1950* **183,** 6041–6050 (2009).

94. Lee, Y. K., Menezes, J. S., Umesaki, Y. & Mazmanian, S. K. Proinflammatory T-cell responses to gut microbiota promote experimental autoimmune encephalomyelitis. *Proc. Natl. Acad. Sci. U. S. A.* **108 Suppl 1,** 4615–4622 (2011).

95. Wu, H.-J. *et al.* Gut-residing segmented filamentous bacteria drive autoimmune arthritis via T helper 17 cells. *Immunity* **32,** 815–827 (2010).

96. Flores, G. E. *et al.* Temporal variability is a personalized feature of the human microbiome. *Genome Biol.* **15,** 531 (2014).

97. Thompson, A. L., Monteagudo-Mera, A., Cadenas, M. B., Lampl, M. L. & Azcarate-Peril, M. A. Milk- and solid-feeding practices and daycare attendance are associated with differences in bacterial diversity, predominant communities, and metabolic and immune function of the infant gut microbiome. *Front. Cell. Infect. Microbiol.* **5,** 3 (2015).

98. Koren, O. *et al.* Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell* **150,** 470–480 (2012).

99. Strober, W., Fuss, I. & Mannon, P. The fundamental basis of inflammatory bowel disease. *J. Clin. Invest.* **117,** 514–521 (2007).

100. Kelsen, J. & Baldassano, R. N. Inflammatory bowel disease: the difference between children and adults. *Inflamm. Bowel Dis.* **14 Suppl 2,** S9–11 (2008).

101. Spor, A., Koren, O. & Ley, R. Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat. Rev. Microbiol.* **9,** 279–290 (2011).

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

103. Fiocchi, C. Inflammatory bowel disease: etiology and pathogenesis. *Gastroenterology* **115,** 182–205 (1998).

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

105. Tlaskalová-Hogenová, H. *et al.* The role of gut microbiota (commensal bacteria) and the mucosal barrier in the pathogenesis of inflammatory and autoimmune diseases and cancer: contribution of germ-free and gnotobiotic animal models of human diseases. *Cell. Mol. Immunol.* **8,** 110–120 (2011).

106. Podolsky, D. K. Inflammatory bowel disease. *N. Engl. J. Med.* **347,** 417–429 (2002).

107. Comito, D., Cascio, A. & Romano, C. Microbiota biodiversity in inflammatory bowel disease. *Ital. J. Pediatr.* **40,** 32 (2014).

108. Hold, G. L. *et al.* Role of the gut microbiota in inflammatory bowel disease pathogenesis: what have we learnt in the past 10 years? *World J. Gastroenterol. WJG* **20,** 1192–1210 (2014).

109. Schwiertz, A. *et al.* Microbiota in pediatric inflammatory bowel disease. *J. Pediatr.* **157,** 240–244.e1 (2010).

110. Wehkamp, J. *et al.* Reduced Paneth cell alpha-defensins in ileal Crohn's disease. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 18129–18134 (2005).

111. Taurog, J. D. *et al.* The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. *J. Exp. Med.* **180,** 2359–2364 (1994).

112. Schultz, M. *et al.* Lactobacillus plantarum 299V in the treatment and prevention of spontaneous colitis in interleukin-10-deficient mice. *Inflamm. Bowel Dis.* **8,** 71–80 (2002).

113. Arnold, G. L., Beaves, M. R., Pryjdun, V. O. & Mook, W. J. Preliminary study of ciprofloxacin in active Crohn's disease. *Inflamm. Bowel Dis.* **8,** 10–15 (2002).

114. Greenbloom, S. L., Steinhart, A. H. & Greenberg, G. R. Combination ciprofloxacin and metronidazole for active Crohn's disease. *Can. J. Gastroenterol. J. Can. Gastroenterol.* **12,** 53–56 (1998).

115. Guslandi, M., Mezzi, G., Sorghi, M. & Testoni, P. A. Saccharomyces boulardii in maintenance treatment of Crohn's disease. *Dig. Dis. Sci.* **45,** 1462–1464 (2000).

116. Kruis, W. *et al.* Double-blind comparison of an oral Escherichia coli preparation and mesalazine in maintaining remission of ulcerative colitis. *Aliment. Pharmacol. Ther.* **11,** 853–858 (1997).

117. Jonkers, D. & Stockbrügger, R. Probiotics and inflammatory bowel disease. *J. R. Soc. Med.* **96,** 167–171 (2003).

118. Manichanh, C. *et al.* Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* **55,** 205–211 (2006).

119. Seksik, P. *et al.* Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. *Gut* **52,** 237–242 (2003).

120. Lucke, K., Miehlke, S., Jacobs, E. & Schuppler, M. Prevalence of Bacteroides and Prevotella spp. in ulcerative colitis. *J. Med. Microbiol.* **55,** 617–624 (2006).

121. Sokol, H. *et al.* Low counts of Faecalibacterium prausnitzii in colitis microbiota. *Inflamm. Bowel Dis.* **15,** 1183–1189 (2009).

122. Sokol, H. *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 16731–16736 (2008).

123. Swidsinski, A., Loening-Baucke, V., Vaneechoutte, M. & Doerffel, Y. Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. *Inflamm. Bowel Dis.* **14,** 147–161 (2008).

124. Fujimoto, T. *et al.* Decreased abundance of Faecalibacterium prausnitzii in the gut microbiota of Crohn's disease. *J. Gastroenterol. Hepatol.* **28,** 613–619 (2013).

125. Hansen, R. *et al.* Microbiota of de-novo pediatric IBD: increased Faecalibacterium prausnitzii and reduced bacterial diversity in Crohn's but not in ulcerative colitis. *Am. J. Gastroenterol.* **107,** 1913–1922 (2012).

126. Duncan, S. H., Hold, G. L., Harmsen, H. J. M., Stewart, C. S. & Flint, H. J. Growth requirements and fermentation products of Fusobacterium prausnitzii, and a proposal to reclassify it as Faecalibacterium prausnitzii gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* **52,** 2141–2146 (2002).

127. Duncan, S. H., Louis, P. & Flint, H. J. Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. *Appl. Environ. Microbiol.* **70,** 5810–5817 (2004).

128. Mortensen, P. B. & Clausen, M. R. Short-chain fatty acids in the human colon: relation to gastrointestinal health and disease. *Scand. J. Gastroenterol. Suppl.* **216,** 132–148 (1996).

129. Ott, S. J. *et al.* Reduction in diversity of the colonic mucosa associated bacterial microflora in patients with active inflammatory bowel disease. *Gut* **53,** 685–693 (2004).

130. Randhawa, P. K., Singh, K., Singh, N. & Jaggi, A. S. A review on chemical-induced inflammatory bowel disease models in rodents. *Korean J. Physiol. Pharmacol. Off. J. Korean Physiol. Soc. Korean Soc. Pharmacol.* **18,** 279–288 (2014).

131. Elson, C. O., Sartor, R. B., Tennyson, G. S. & Riddell, R. H. Experimental models of inflammatory bowel disease. *Gastroenterology* **109,** 1344–1367 (1995).

132. Ostanin, D. V. *et al.* T cell transfer model of chronic colitis: concepts, considerations, and tricks of the trade. *Am. J. Physiol. Gastrointest. Liver Physiol.* **296,** G135–146 (2009).

133. Morrissey, P. J., Charrier, K., Braddy, S., Liggitt, D. & Watson, J. D. CD4+ T cells that express high levels of CD45RB induce wasting disease when transferred into congenic severe combined immunodeficient mice. Disease development is prevented by cotransfer of purified CD4+ T cells. *J. Exp. Med.* **178,** 237–244 (1993).

134. Powrie, F., Leach, M. W., Mauze, S., Caddle, L. B. & Coffman, R. L. Phenotypically distinct subsets of CD4+ T cells induce or protect from chronic intestinal inflammation in C. B-17 scid mice. *Int. Immunol.* **5,** 1461–1471 (1993).

135. Mombaerts, P. *et al.* RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* **68,** 869–877 (1992).

136. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5463–5467 (1977).

137. Prober, J. M. *et al.* A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* **238,** 336–341 (1987).

138. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* **431,** 931–945 (2004).

139.   Collins, F. S., Morgan, M. & Patrinos, A. The Human Genome Project: lessons from large-scale

biology. *Science* **300,** 286–290 (2003).

140.   Margulies, M. *et al.* Genome sequencing in microfabricated high-density picolitre reactors.

Nature **437,** 376–380 (2005).

141.   Bennett, S. Solexa Ltd. *Pharmacogenomics* **5,** 433–438 (2004).

142.   Breu, H. A Theoretical Understanding of 2 Base Color Codes and Its Application to

Annotation, Error Detection, and Error Correction. at

<https://www3.appliedbiosystems.com/cms/groups/mcb_marketing/documents/generaldocum

ents/cms_058265.pdf>

143.   Shendure, J. *et al.* Accurate multiplex polony sequencing of an evolved bacterial genome.

Science **309,** 1728–1732 (2005).

144.   Adessi, C. *et al.* Solid phase DNA amplification: characterisation of primer attachment and

amplification mechanisms. *Nucleic Acids Res.* **28,** E87 (2000).

145.   Fedurco, M., Romieu, A., Williams, S., Lawrence, I. & Turcatti, G. BTA, a novel reagent for

DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic

Acids Res.* **34,** e22 (2006).

146.   Braslavsky, I., Hebert, B., Kartalov, E. & Quake, S. R. Sequence information can be obtained

from single DNA molecules. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 3960–3964 (2003).

147.   Eid, J. *et al.* Real-time DNA sequencing from single polymerase molecules. *Science* **323,** 133–

138 (2009).

148.   Kobayashi, N., Tamura, K. & Aotsuka, T. PCR error and molecular population genetics.

Biochem. Genet. **37,** 317–321 (1999).

149.   Liu, L. *et al.* Comparison of next-generation sequencing systems. *J. Biomed. Biotechnol.* **2012,**

251364 (2012).

150.   Illumina. HiSeq 2500 Spceifications. *HiSeq 2500 Specifications* at

<http://www.illumina.com/systems/hiseq_2500_1500/performance_specifications.html>

151.   Illumina. MiSeq Specifications. *MiSeq Specifications* at

        <http://www.illumina.com/systems/miseq/performance_specifications.html>

152.   Quail, M. A. *et al.* A tale of three next generation sequencing platforms: comparison of Ion

        Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* **13,** 341 (2012).

153.   Revolutionize Genomics with SMRT Sequencing - PacBio_RS_II_Brochure.pdf. at

        <http://files.pacb.com/pdf/PacBio_RS_II_Brochure.pdf>

154.   Pacific Biosciences: Workflow www.pacificbiosciences.com. at

        <http://www.pacificbiosciences.com/products/pacificbio-rs-workflow-main/>

155.   Voelkerding, K. V., Dames, S. A. & Durtschi, J. D. Next-generation sequencing: from basic

        research to diagnostics. *Clin. Chem.* **55,** 641–658 (2009).

156.   Aas, J. A., Paster, B. J., Stokes, L. N., Olsen, I. & Dewhirst, F. E. Defining the normal bacterial

        flora of the oral cavity. *J. Clin. Microbiol.* **43,** 5721–5732 (2005).

157.   Bik, E. M. *et al.* Molecular analysis of the bacterial microbiota in the human stomach. *Proc.

        Natl. Acad. Sci. U. S. A.* **103,** 732–737 (2006).

158.   Goodman, A. L. *et al.* Extensive personal human gut microbiota culture collections

        characterized and manipulated in gnotobiotic mice. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 6252–6257

        (2011).

159.   Woese, C. R. & Fox, G. E. Phylogenetic structure of the prokaryotic domain: the primary

        kingdoms. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 5088–5090 (1977).

160.   Fox, G. E. *et al.* The phylogeny of prokaryotes. *Science* **209,** 457–463 (1980).

161.   Kolbert, C. P. & Persing, D. H. Ribosomal DNA sequencing as a tool for identification of

        bacterial pathogens. *Curr. Opin. Microbiol.* **2,** 299–305 (1999).

162.   Woese, C. R. *et al.* Conservation of primary structure in 16S ribosomal RNA. *Nature* **254,** 83–

        86 (1975).

163.   Van de Peer, Y., Chapelle, S. & De Wachter, R. A quantitative map of nucleotide substitution

        rates in bacterial rRNA. *Nucleic Acids Res.* **24,** 3381–3391 (1996).

164. Peterson, D. A., Frank, D. N., Pace, N. R. & Gordon, J. I. Metagenomic approaches for defining the pathogenesis of inflammatory bowel diseases. *Cell Host Microbe* **3,** 417–427 (2008).

165. Schloss, P. D. & Handelsman, J. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* **71,** 1501–1506 (2005).

166. Morgan, J. L., Darling, A. E. & Eisen, J. A. Metagenomic sequencing of an in vitro-simulated microbial community. *PloS One* **5,** e10209 (2010).

167. Wang, G. C. & Wang, Y. The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. *Microbiol. Read. Engl.* **142 (Pt 5),** 1107–1114 (1996).

168. Morgan, X. C. & Huttenhower, C. Chapter 12: Human microbiome analysis. *PLoS Comput. Biol.* **8,** e1002808 (2012).

169. Streit, W. R. & Schmitz, R. A. Metagenomics--the key to the uncultured microbes. *Curr. Opin. Microbiol.* **7,** 492–498 (2004).

170. Klappenbach, J. A., Saxman, P. R., Cole, J. R. & Schmidt, T. M. rrndb: the Ribosomal RNA Operon Copy Number Database. *Nucleic Acids Res.* **29,** 181–184 (2001).

171. Warnecke, F. & Hess, M. A perspective: metatranscriptomics as a tool for the discovery of novel biocatalysts. *J. Biotechnol.* **142,** 91–95 (2009).

172. Wu, G. D. & Lewis, J. D. Analysis of the human gut microbiome and association with disease. *Clin. Gastroenterol. Hepatol. Off. Clin. Pract. J. Am. Gastroenterol. Assoc.* **11,** 774–777 (2013).

173. MEGAN5 - MEtaGenome ANalyzer — Algorithms in Bioinformatics. at <http://ab.inf.uni-tuebingen.de/software/megan5/>

174. Qiime. Analysis of shotgun sequencing data — Homepage. (2015). at <http://qiime.org/tutorials/shotgun_analysis.html>

175. Kolida, S., Meyer, D. & Gibson, G. R. A double-blind placebo-controlled study to establish the bifidogenic dose of inulin in healthy humans. *Eur. J. Clin. Nutr.* **61,** 1189–1195 (2007).

176. Sangwan, P., Kovac, S., Davis, K. E. R., Sait, M. & Janssen, P. H. Detection and Cultivation of Soil Verrucomicrobia. *Appl. Environ. Microbiol.* **71,** 8402–8410 (2005).

177. Ruscheweyh, Hans-Joachim. Methods for large-scale Microbiome Analysis using MEGAN. (2015). doi:10.15496/publikation-5239

178. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **7,** 335–336 (2010).

179. Krajina, T., Leithäuser, F. & Reimann, J. MHC class II-independent CD25+ CD4+ CD8alpha beta+ alpha beta T cells attenuate CD4+ T cell-induced transfer colitis. *Eur. J. Immunol.* **34,** 705–714 (2004).

180. Wesolowska-Andersen, A. *et al.* Choice of bacterial DNA extraction method from fecal material influences community structure as evaluated by metagenomic analysis. *Microbiome* **2,** 19 (2014).

181. Godon, J. J., Zumstein, E., Dabert, P., Habouzit, F. & Moletta, R. Molecular microbial diversity of an anaerobic digestor as determined by small-subunit rDNA sequence analysis. *Appl. Environ. Microbiol.* **63,** 2802–2813 (1997).

182. Matthiesen, S. H. & Hansen, C. M. Fast and non-toxic in situ hybridization without blocking of repetitive sequences. *PloS One* **7,** e40675 (2012).

183. Gronbach, K. *et al.* Endotoxicity of lipopolysaccharide as a determinant of T-cell-mediated colitis induction in mice. *Gastroenterology* **146,** 765–775 (2014).

184. Vázquez-Baeza, Y., Pirrung, M., Gonzalez, A. & Knight, R. EMPeror: a tool for visualizing high-throughput microbial community data. *GigaScience* **2,** 16 (2013).

185. Bernhardt, J., Michalik, S., Wollscheid, B., Völker, U. & Schmidt, F. Proteomics approaches for the analysis of enriched microbial subpopulations and visualization of complex functional information. *Curr. Opin. Biotechnol.* **24,** 112–119 (2013).

186. Müller, M. *et al.* Intestinal Colonization of IL-2 Deficient Mice with Non-Colitogenic B. vulgatus Prevents DC Maturation and T-Cell Polarization. *PLoS ONE* **3,** (2008).

187.   Promega. Protocols Applications Guide. at

<https://www.promega.de/~/media/files/resources/paguide/a4/chap1a4.pdf?la=de-de>

188.   Poor PCR Efficiency. at <http://www.lifetechnologies.com/de/de/home/life-science/pcr/real-

time-pcr/qpcr-education/real-time-pcr-troubleshooting-tool/gene-expression-quantitation-

troubleshooting/poor-pcr-efficiency.html>

189.   Hoshino, T., Yilmaz, L. S., Noguera, D. R., Daims, H. & Wagner, M. Quantification of target

molecules needed to detect microorganisms by fluorescence in situ hybridization (FISH) and

catalyzed reporter deposition-FISH. *Appl. Environ. Microbiol.* **74,** 5068–5077 (2008).

190.   daime: Digital Image Analysis in Microbial Ecology | Microbial Ecology, University of Vienna.

at <http://www.microbial-ecology.net/daime#abundance>

191.   Bloom, S. M. *et al.* Commensal Bacteroides species induce colitis in host-genotype-specific

fashion in a mouse model of inflammatory bowel disease. *Cell Host Microbe* **9,** 390–403 (2011).

192.   Guo, F. & Zhang, T. Biases during DNA extraction of activated sludge samples revealed by

high throughput sequencing. *Appl. Microbiol. Biotechnol.* **97,** 4607–4616 (2013).

193.   Vishnivetskaya, T. A. *et al.* Commercial DNA extraction kits impact observed microbial

community composition in permafrost samples. *FEMS Microbiol. Ecol.* **87,** 217–230 (2014).

194.   Henderson, G. *et al.* Effect of DNA extraction methods and sampling techniques on the

apparent structure of cow and sheep rumen microbial communities. *PloS One* **8,** e74787 (2013).

195.   Bonot, S., Courtois, S., Block, J.-C. & Merlin, C. Improving the recovery of qPCR-grade DNA

from sludge and sediment. *Appl. Microbiol. Biotechnol.* **87,** 2303–2311 (2010).

196.   Bahl, M. I., Bergström, A. & Licht, T. R. Freezing fecal samples prior to DNA extraction affects

the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis.

FEMS Microbiol. Lett. **329,** 193–197 (2012).

197.   Lutz, K. A., Wang, W., Zdepski, A. & Michael, T. P. Isolation and analysis of high quality

nuclear DNA with reduced organellar DNA for plant genome sequencing and resequencing. *BMC

Biotechnol.* **11,** 54 (2011).

198.   Agilent Technologies. Agilent 2100 Bioanalyzer 2100 Expert User's Guide. (2003). at
    <http://gcf.pbrc.edu/docs/Agilent/Agilent%20Manual.pdf>

199.   Hirsch-Kauffmann, M. & Schweiger, M. *Biologie für Mediziner und Naturwissenschaftler: 74 Tabellen*. (Thieme, 2006).

200.   Bhagwat, A. A., Ying, Z. I., Karns, J. & Smith, A. Determining RNA quality for NextGen sequencing: some exceptions to the gold standard rule of 23S to 16S rRNA ratio§. *Microbiol. Discov.* **1,** 10 (2013).

201.   Biomedical Genomics. RNA Quality Control. at
    <http://biomedicalgenomics.org/How_does_intact_total_RNA_looks_like.html>

202.   Kim, M., Morrison, M. & Yu, Z. Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J. Microbiol. Methods* **84,** 81–87 (2011).

203.   Shah, N., Tang, H., Doak, T. G. & Ye, Y. Comparing bacterial communities inferred from 16S rRNA gene sequencing and shotgun metagenomics. *Pac. Symp. Biocomput. Pac. Symp. Biocomput.* 165–176 (2011).

204.   Mizrahi-Man, O., Davenport, E. R. & Gilad, Y. Taxonomic classification of bacterial 16S rRNA genes using short sequencing reads: evaluation of effective study designs. *PloS One* **8,** e53608 (2013).

205.   Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73,** 5261–5267 (2007).

206.   Salipante, S. J. *et al.* Rapid 16S rRNA next-generation sequencing of polymicrobial clinical samples for diagnosis of complex bacterial infections. *PloS One* **8,** e65226 (2013).

207.   Roche Shutting Down 454 Sequencing Business. *GenomeWeb* (2013). at
    <https://www.genomeweb.com/sequencing/roche-shutting-down-454-sequencing-business>

208.   Illumina. 16S Metagenomic Sequencing Library Preparation. at

<http://res.illumina.com/documents/products/appnotes/16s-metagenomic-library-prep-

guide.pdf>

209.   Fadrosh, D. W. *et al.* An improved dual-indexing approach for multiplexed 16S rRNA gene

sequencing on the Illumina MiSeq platform. *Microbiome* **2,** 6 (2014).

210.   Takahashi, S., Tomita, J., Nishioka, K., Hisada, T. & Nishijima, M. Development of a

prokaryotic universal primer for simultaneous analysis of Bacteria and Archaea using next-

generation sequencing. *PloS One* **9,** e105592 (2014).

211.   Kianoush, N. *et al.* Bacterial profile of dentine caries and the impact of pH on bacterial

population diversity. *PloS One* **9,** e92940 (2014).

212.   Ronaghi, M., Uhlén, M. & Nyrén, P. A sequencing method based on real-time pyrophosphate.

Science **281,** 363, 365 (1998).

213.   Marsh, S. *Pyrosequencing protocols*. (Humana Press, 2007). at

<http://site.ebrary.com/id/10190527>

214.   Luo, C., Tsementzi, D., Kyrpides, N., Read, T. & Konstantinidis, K. T. Direct comparisons of

Illumina vs. Roche 454 sequencing technologies on the same microbial community DNA sample.

PloS One **7,** e30087 (2012).

215.   Larsen, N. *et al.* Gut microbiota in human adults with type 2 diabetes differs from non-

diabetic adults. *PloS One* **5,** e9085 (2010).

216.   Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Microbial ecology: human gut microbes

associated with obesity. *Nature* **444,** 1022–1023 (2006).

217.   Willing, B. P. *et al.* A pyrosequencing study in twins shows that gastrointestinal microbial

profiles vary with inflammatory bowel disease phenotypes. *Gastroenterology* **139,** 1844–1854.e1

(2010).

218.   Moloney, R. D., Desbonnet, L., Clarke, G., Dinan, T. G. & Cryan, J. F. The microbiome: stress,

health and disease. *Mamm. Genome Off. J. Int. Mamm. Genome Soc.* **25,** 49–74 (2014).

219. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489,** 220–230 (2012).

220. Sorge, R. E. *et al.* Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nat. Methods* **11,** 629–632 (2014).

221. Tannock, G. W. & Savage, D. C. Influences of dietary and environmental stress on microbial populations in the murine gastrointestinal tract. *Infect. Immun.* **9,** 591–598 (1974).

222. Walters, W. A., Xu, Z. & Knight, R. Meta-analyses of human gut microbes associated with obesity and IBD. *FEBS Lett.* **588,** 4223–4233 (2014).

223. Carvalho, F. A. *et al.* Transient inability to manage proteobacteria promotes chronic gut inflammation in TLR5-deficient mice. *Cell Host Microbe* **12,** 139–152 (2012).

224. Collado, M. C., Cernada, M., Baüerl, C., Vento, M. & Pérez-Martínez, G. Microbial ecology and host-microbiota interactions during early life stages. *Gut Microbes* **3,** 352–365 (2012).

225. Rogers, G. B. *et al.* Functional divergence in gastrointestinal microbiota in physically-separated genetically identical mice. *Sci. Rep.* **4,** 5437 (2014).

226. Frank, D. N., Zhu, W., Sartor, R. B. & Li, E. Investigating the biological and clinical significance of human dysbioses. *Trends Microbiol.* **19,** 427–434 (2011).

227. Nugent, S. G., Kumar, D., Rampton, D. S. & Evans, D. F. Intestinal luminal pH in inflammatory bowel disease: possible determinants and implications for therapy with aminosalicylates and other drugs. *Gut* **48,** 571–577 (2001).

228. Hajishengallis, G. *et al.* Low-abundance biofilm species orchestrates inflammatory periodontal disease through the commensal microbiota and complement. *Cell Host Microbe* **10,** 497–506 (2011).

229. Derrien, M. Akkermansia muciniphila gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. *Int. J. Syst. Evol. Microbiol.* **54,** 1469–1476 (2004).

230. Everard, A. *et al.* Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. *Proc. Natl. Acad. Sci.* **110,** 9066–9071 (2013).

231.   Miquel, S. *et al.* Faecalibacterium prausnitzii and human intestinal health. *Curr. Opin.*

Microbiol. **16,** 255–261 (2013).

232.   Rath, H. C. *et al.* Normal luminal bacteria, especially Bacteroides species, mediate chronic

colitis, gastritis, and arthritis in HLA-B27/human beta2 microglobulin transgenic rats. *J. Clin.*

Invest. **98,** 945–953 (1996).

233.   Cardona, S. *et al.* Storage conditions of intestinal microbiota matter in metagenomic analysis.

BMC Microbiol. **12,** 158 (2012).

234.   Zackular, J. P., Rogers, M. A. M., Ruffin, M. T. & Schloss, P. D. The human gut microbiome as a

screening tool for colorectal cancer. *Cancer Prev. Res. Phila. Pa* **7,** 1112–1121 (2014).

Appendix

| Week 0 | Week 3 | Week 6 |

Figure 82: Composition of the intestinal microbiome of Endo[hi] mice (n=7; all mice were part of experiment 3) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

Week 0   Week 3   Week 6

Figure 83: Composition of the intestinal microbiome of Endo^{lo} mice (n=5; all mice were part of experiment 4) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
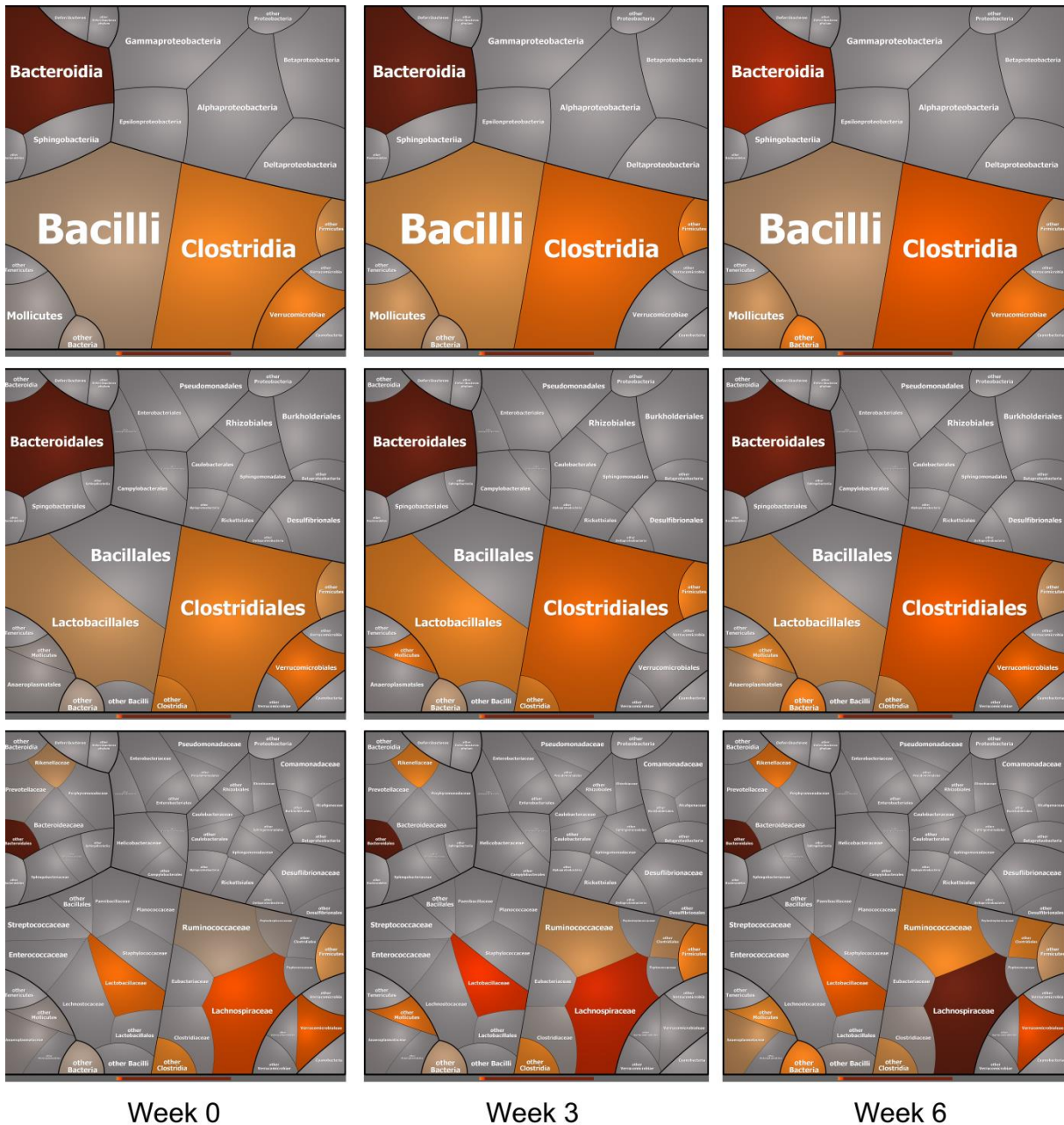
Week 0   Week 3   Week 6

Figure 84: Composition of the intestinal microbiome of *E. coli*_{WT} pretreated Endo^{hi} mice (n=2; all mice were part of experiment 3) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
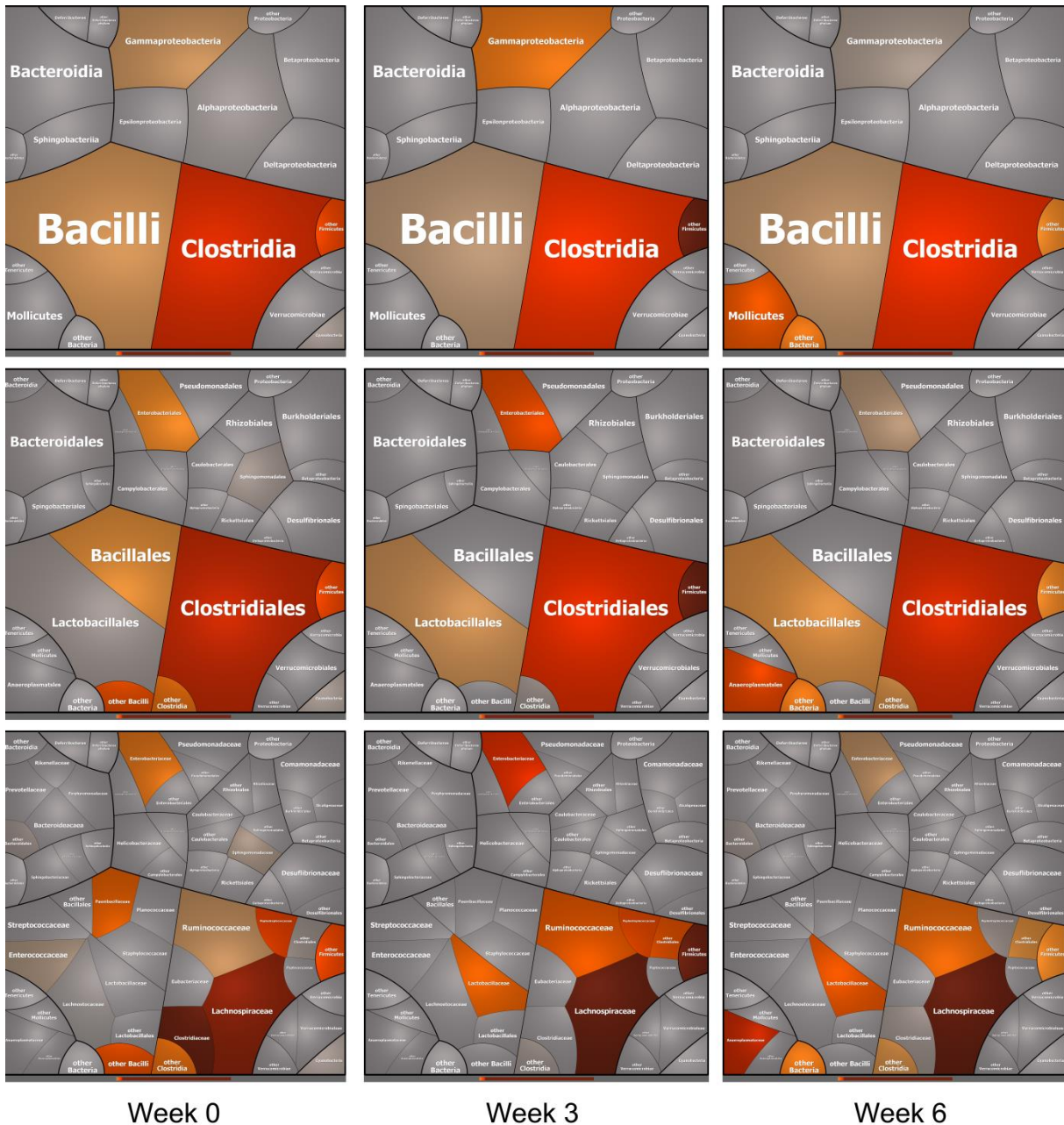
Week 0   Week 3   Week 6

Figure 85: Composition of the intestinal microbiome of _E. coli_WT pretreated Endo^{lo} mice (n=5; all mice were part of experiment 4) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
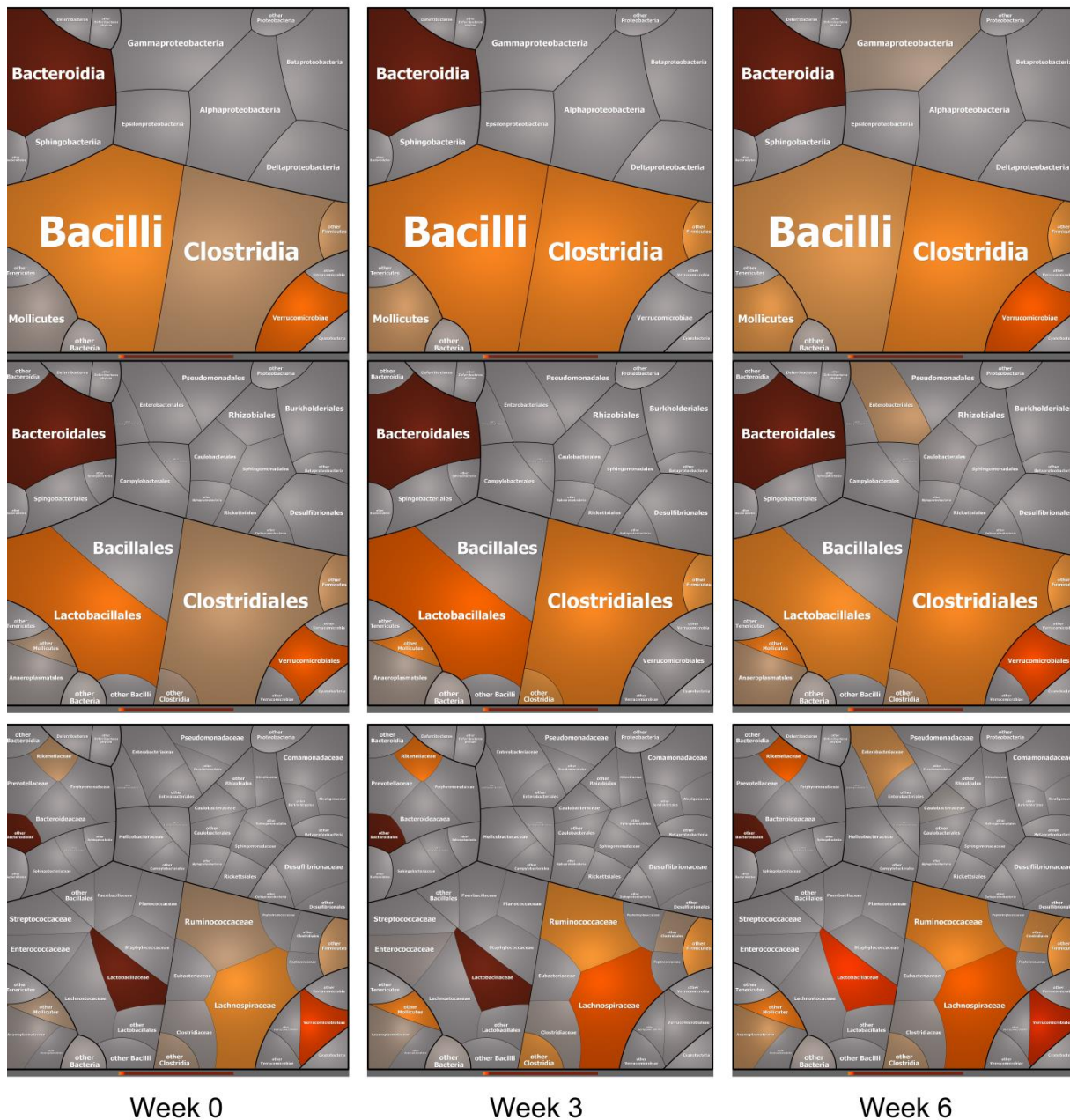
Figure 86: Composition of the intestinal microbiome of *E. coli*MUT pretreated Endohi mice (n=2; all mice were part of experiment 3) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
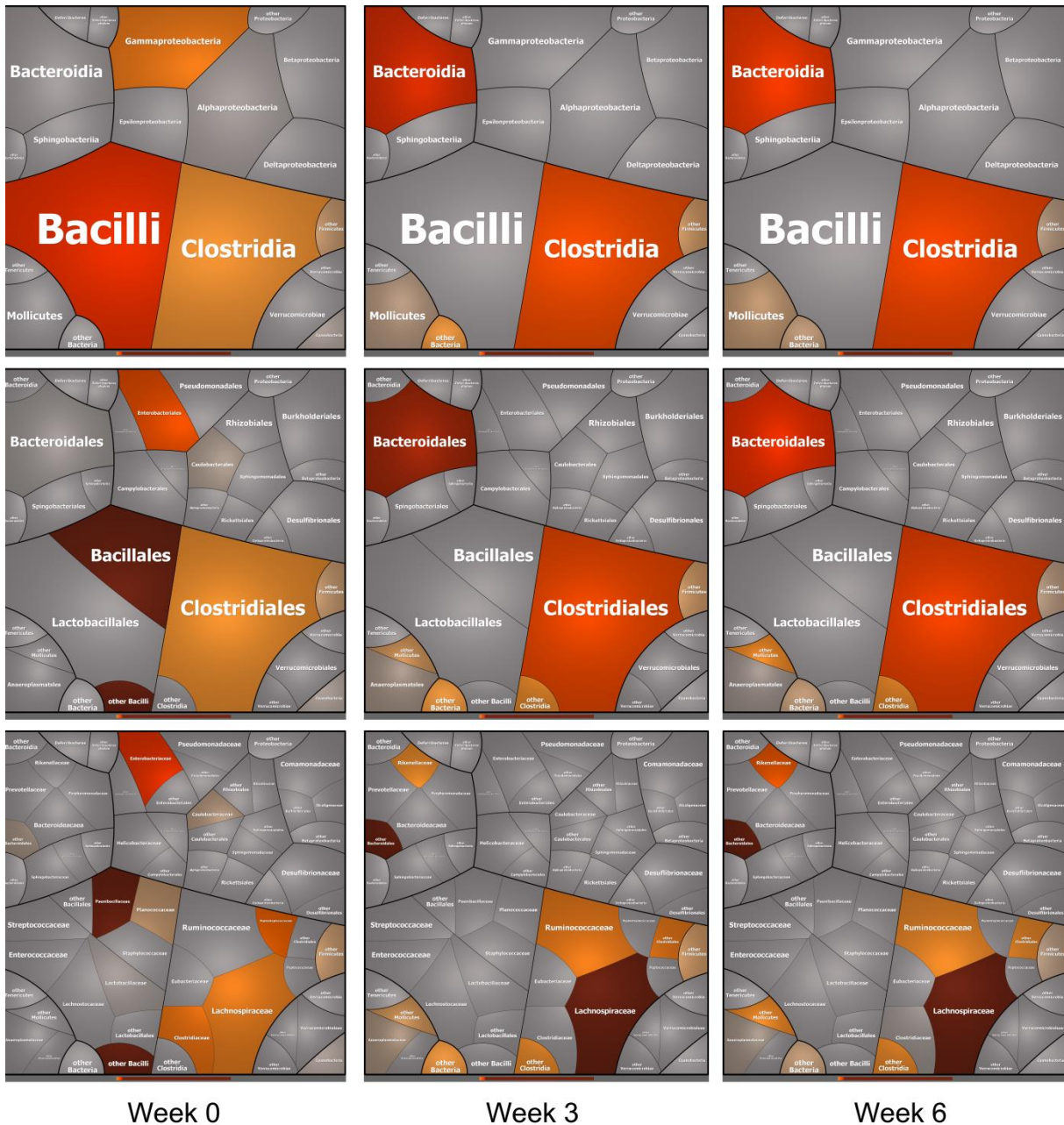
Figure 87: Composition of the intestinal microbiome of *E. coli*_{MUT} pretreated Endo^{lo} mice (n=5; all mice were part of experiment 4) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
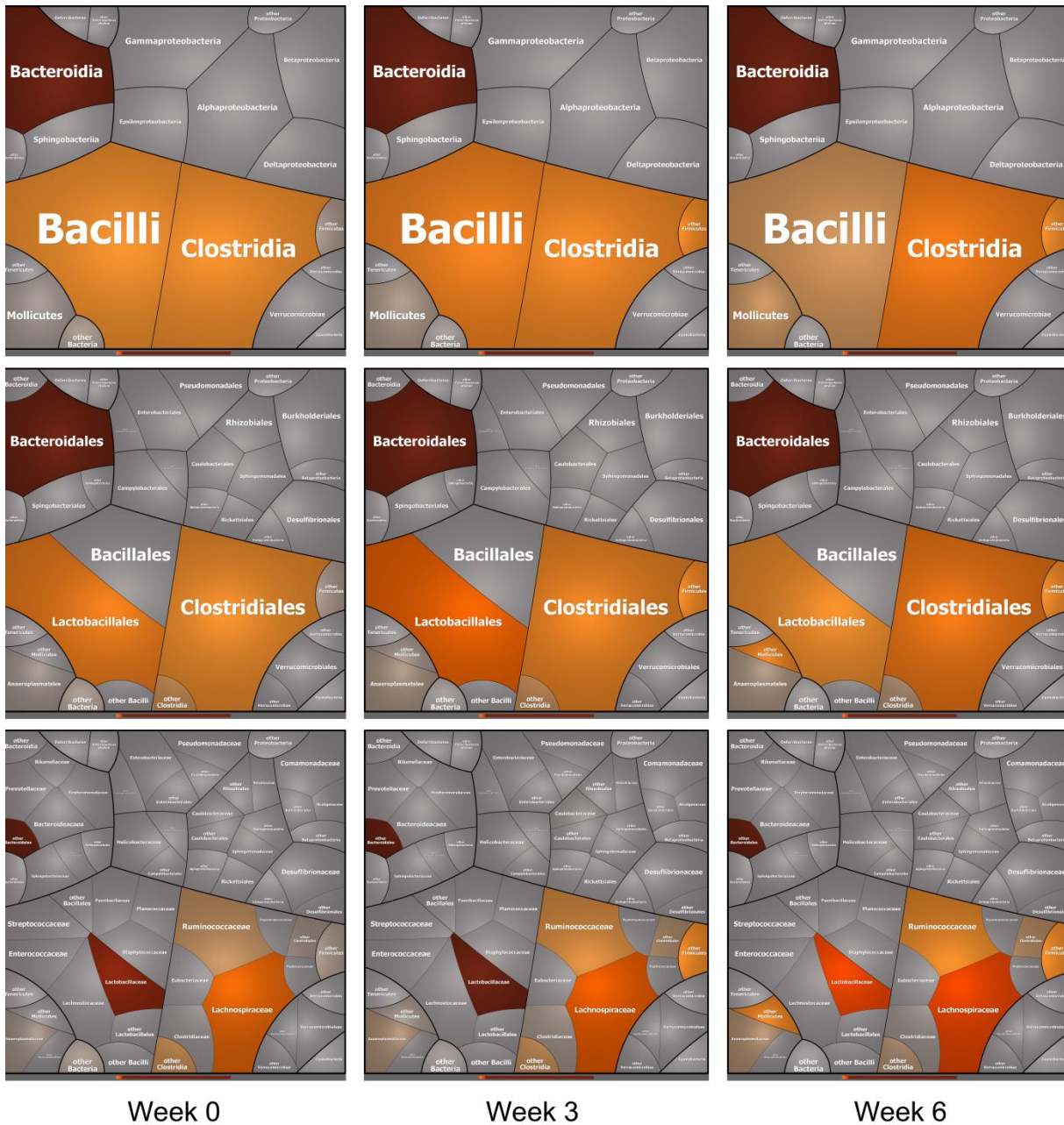
Figure 88: Composition of the intestinal microbiome of LPS$_{WT}$ pretreated Endo^{hi} mice (n=3; all mice were part of experiment 7) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
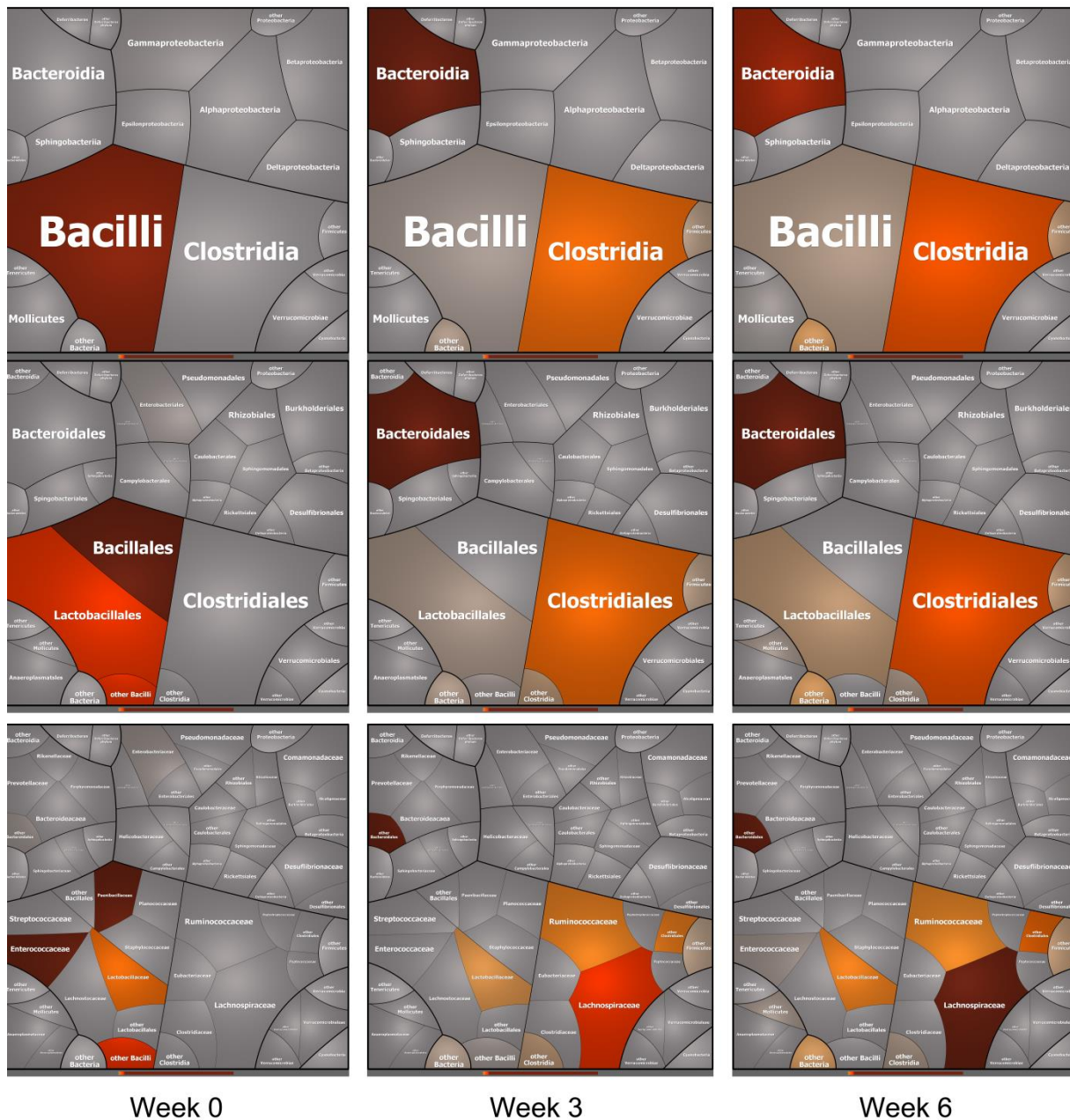
Week 0   Week 3   Week 6

Figure 89: Composition of the intestinal microbiome of LPS$_{WT}$ pretreated Endo^{lo} mice (n=4; all mice were part of experiment 4) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.

Figure 90: Composition of the intestinal microbiome of LPS$_{MUT}$ pretreated Endo^{hi} mice (n=5; all mice were part of experiment 3) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
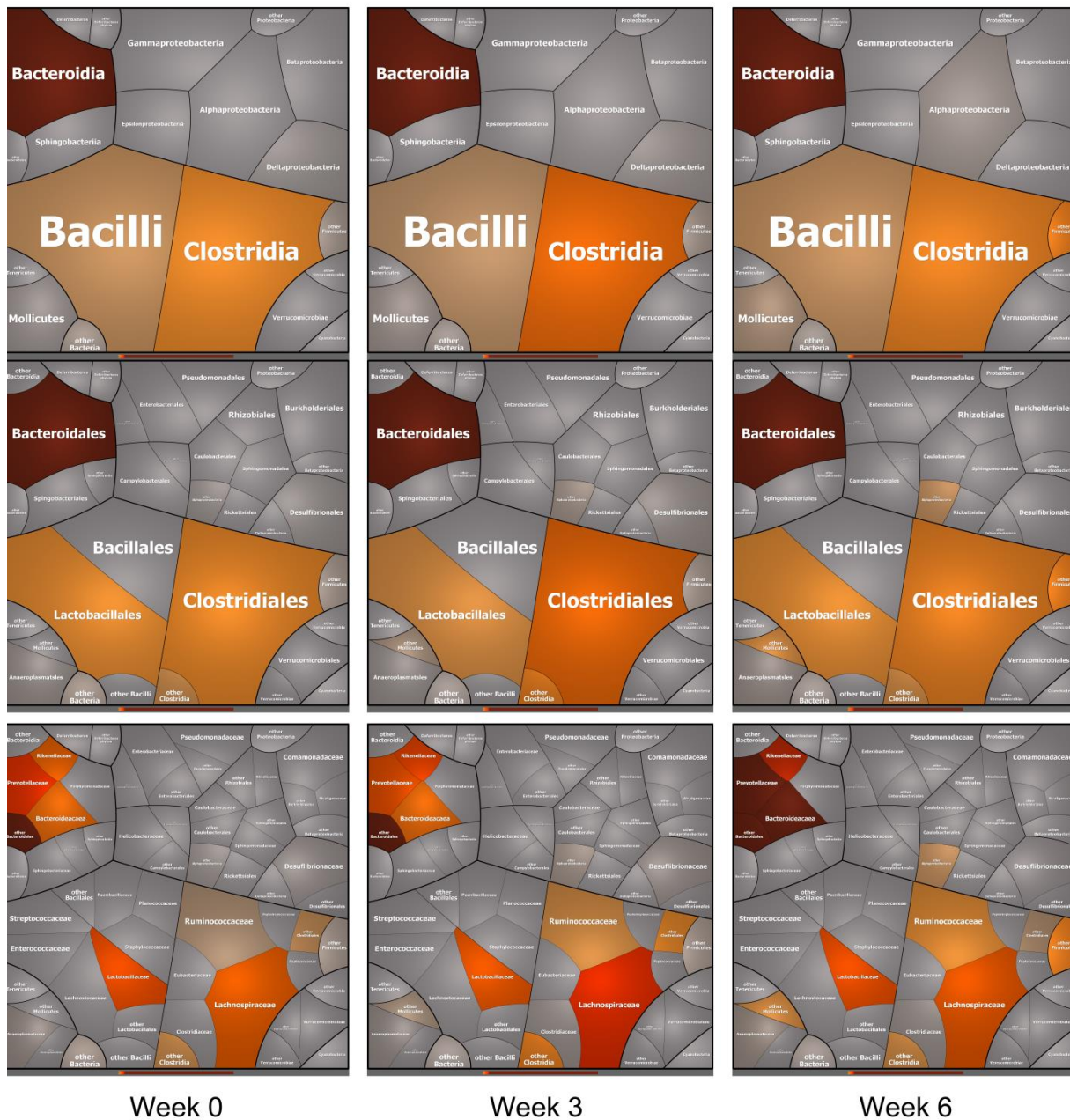
Figure 91: Composition of the intestinal microbiome of LPS_{MUT} pretreated Endo^{lo} mice (n=4; all mice were part of experiment 4) before, 3 or 6 weeks post T-cell transfer. The top row shows the composition on the class level, the row in the middle on the order level and the last row on the family level. Reads that could not be assigned, were declared as "others" of the lowest taxonomic level possible. The figures were created using Paver. The darker the color, the more reads were associated with the corresponding taxon.
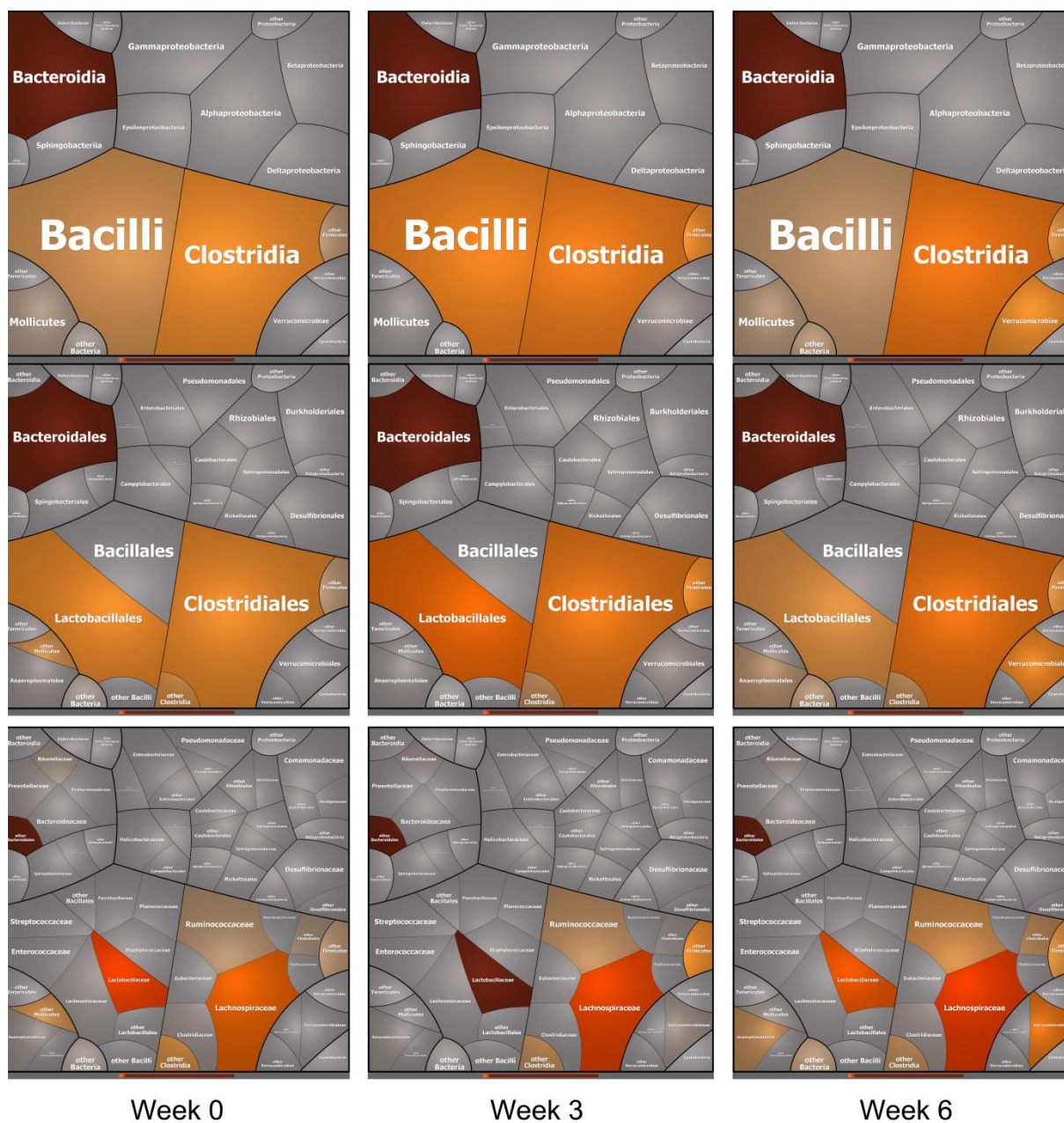
Acknowledgements

First and foremost, I thank Prof. Dr. Julia-Stefanie Frick for accepting me into her group and giving me the chance to work on this project. I am very grateful for the supervision of my project, our discussions and all her support, patience and advice.

My appreciation also extends to Prof. Dr. Ingo Autenrieth for the opportunity to work at the Institute for Medical Microbiology and Hygiene under great working conditions, and for the support of my project.

Many thanks to Prof. Dr. Andreas Peschel, Prof. Dr. Christoph Mayer and Prof. Dr. Friedrich Götz for taking the time to be part of my examination committee.

A special thanks goes to Dr. Kerstin Gronbach for the incredible help from the beginning to the end of my project, for proofreading the dissertation, for everything she has taught me and the great cooperation on the mouse experiments.

I also thank Dr. Hans-Joachim Ruscheweyh for the bioinformatic processing of our data, for the patience answering all my questions concerning the bioinformatic analysis and the inspiring discussions.

Thanks also to Jörg Bernhard for helping out to create the figures using the Paver software.

My gratitude also goes to everyone who is or was part of our working group, especially Sarah, Alex W. Annika, Andi, Anna, Alex S and Raphael, for the fun time we had together in- and outside of the institute, the pleasant working environment and especially the willingness to help out whenever it was needed.

I further thank all the people working at the Institute of Medical Microbiology for the countless breaks we had together with a lot of fun as well as encouraging talks and helpful ideas.

Many thanks to Dr. Melissa Cox for proofreading my dissertation.

Last but not least, I thank my parents for all the support, encouragement and for believing in me all these years.

Curriculum Vitae

Personal Information

Name:	Isabell Flade
Date and Place of Birth:	09.08.1985, Sindelfingen

Education

01.2011- 07.2014	PhD Student at the Institute for Medical Microbiology and Hygiene at the University clinic Tübingen
22.10.2010	Master of Science
10.2008 – 10.2010	Masters program Molecular Life Science at the University of Luebeck
04.08.2008	Bachelor of Science
10.2005 - 08.2008	Bachelors program Molecular Medicine at the University of Ulm
24.06.2005	Abitur
09.2002 - 06.2005	Biotechnologisches Gymnasium Mildred-Scheel-Schule Boeblingen
02.1998 - 08.2002	Andreae-Gymnasium Herrenberg
09.1996 – 02.1998	Jerg-Ratgeb-Realschule Herrenberg

Work experience

08.2014 - current	Managing Director of CeMeT GmbH
09.11.2009 – 08.02.2010	Internship at the Dana Farber Cancer Institute; Department Cancer, Immunology and AIDS (analysis of miRNAs in T-cell development)
14.08. – 31.10.2009	Internship at Karolinska Institute in Stockholm Department of Laboratory Medicine (DNA isolation from Blood serum and detection via PCR)
15.08.2006-14.10.2007	Working Student at Philips Medical systems Department Molecular Diagnostics
01.03. – 31.03.2006	Internship at Philips Medical systems Department Molecular Diagnostics
01.08 – 30.09. 2005	Internship at Philips Medical systems Department Molecular Diagnostics (Main focus on isolation of bacterial DNA)

Publications

Paper

Gronbach K*, **Flade I***, Holst O, Lindner B, Ruscheweyh HJ, Wittmann A, Menz S, Schwiertz A, Adam P, Stecher B, Josenhans C, Suerbaum S, Gruber AD, Kulik A, Huson D, Autenrieth IB, Frick JS.
Endotoxicity of lipopolysaccharide as a determinant of T-cell-mediated colitis induction in mice.
Gastroenterology. 2014 Mar;146(3):765-75.

* equally contributing

Wittmann A, Bron PA, van Swam II, Kleerebezem M, Adam P, Gronbach K, Menz S**, Flade I**, Bender A, Schäfer A, Korkmaz AG, Parusel R, Autenrieth IB, Frick JS.
TLR signaling-induced CD103-expressing cells protect against intestinal inflammation.
Inflamm Bowel Dis. 2015 Mar;21(3):507-19.

Posters

Impact of the intestinal microbiota on mucosal homeostasis
Flade I, Gronbach K, Stecher B, Huson D, Ruscheweyh HJ, Autenrieth IB, Frick JS
634th Annual Meeting of the German Society for Hygiene and Microbiology (DGHM) e. V.(2011)

Impact of the intestinal microbiota on mucosal homeostasis
Flade I, Gronbach K, Stecher B, Huson D, Ruscheweyh HJ, Autenrieth IB, Frick JS
Annual Conference of the Association for General and Applied Microbiology (2012)

Impact of the intestinal microbiota on mucosal homeostasis
Flade I, Gronbach K, Stecher B, Huson D, Ruscheweyh HJ, Autenrieth IB, Frick JS
5th Seeon Conference (2012)

Impact of the intestinal microbiota on homeostasis
I. Flade, K. Gronbach, B. Stecher, D. Huson, H.-J. Ruscheweyh, I.B. Autenrieth, J.-S. Frick
64th Annual Meeting of the German Society for Hygiene and Microbiology (2012)

Influence of the gut microbiota on IBD development Influence of the gut microbiota on IBD development
Flade I, Gronbach K, Stecher B, Huson D, Ruscheweyh HJ, Autenrieth IB, Frick JS
65th Annual Meeting of the German Society for Hygiene and Microbiology (DGHM) e. V.(2013)

The dichotomic role of LPS: induction or prevention of intestinal inflammation
Gronbach K, **Flade I***, Holst O, Lindner B, Ruscheweyh HJ A, Menz S, Schwiertz A, Adam P, Stecher B, Josenhans C, Suerbaum S, Gruber AD, Kulik A, Huson D, Autenrieth IB, Frick JS
1. International RTG 1708 Workshop on "Bacterial survival strategies" (30.09.-02.10.2013)

Impact of the intestinal microbiota on mucosal homeostasis
Flade I, Gronbach K, Stecher B, Huson D, Ruscheweyh HJ, Autenrieth IB, Frick JS
7th Seeon Conference (2014)

Presentations

The dichotomic role of LPS: induction or prevention of intestinal inflammation
6th Seeon Conference (2013)

16S Sequencing Challenges
QBIC Tutorial on Metagenom Analysis (2013)