Universiteit
Leiden
The Netherlands

Microbiome-mediated colonization resistance: defense against enteropathogens and multi-drug resistant organisms

Ducarmon, Q.R.

Citation
Ducarmon, Q. R. (2022, March 23). *Microbiome-mediated colonization resistance: defense against enteropathogens and multi-drug resistant organisms*. Retrieved from https://hdl.handle.net/1887/3280022

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/3280022

Note: To cite this publication please use the final published version (if applicable).

Chapter 1

General introduction and
thesis outline

1

The human microbiome

Humans are colonized by microorganisms at different body sites, such as in the oral cavity, on the skin and in the gut. It is estimated that bacterial cells outnumber somatic cells (approximately by a factor of 1.3) and they certainly contain a much wider repertoire of genes than encoded by the human genome[1]. These numbers do not yet take into account other crucial, but understudied, components of the microbiome such as viruses, archaea, fungi and other eukaryotic microorganisms. Defining and distinguishing 'microbiota' and 'microbiome' remains a somewhat controversial topic and one for which extensive debate will likely remain for the coming years. A consensus statement from 2020 defined microbiota as "the assemblage of living microorganisms present in a defined environment" while microbiome was defined as not only including the community of microorganisms, but also their "theatre of activity"[2]. One of the first large-scale projects to characterize microbial communities at different human body sites was the Human Microbiome Project (HMP), officially launched in 2007[3, 4]. The HMP contributed important biological discoveries such as the notion that functional capacity of the microbiome is very stable within a healthy adult over time, but also highly similar between adults, in contrast to taxonomic composition (Figure 1)[4]. While taxonomic composition varied between individuals, general patterns could still be noticed. Most individuals' gut microbiome was dominated by either Firmicutes or Bacteroidetes, with three other phyla (Actinobacteria, Proteobacteria, Verrucomicrobia) being less abundant, but still prevalent. The notion of functional stability between individuals despite taxonomic differences was not only true for the gut microbiome, but also for all other investigated body sites, which included the buccal mucosa, tongue dorsum and anterior nares, amongst others (Figure 1)[4]. In addition, they also released several freely available computational tools (most notably MetaPhlAn for taxonomic profiling and HUManN for functional profiling) which are used by many researchers to this day[3, 5, 6]. Since the launch of the HMP, the HMP and many other research groups and consortia have uncovered that the microbial communities colonizing humans are crucial for maintaining health, and many diseases have been associated with changes in these communities. For example, the gut microbiome can contribute to human health by producing short-chain fatty acids, synthesis of several vitamins and providing resistance against colonization of incoming pathogens[7, 8]. While microbiome research is performed on different body sites, the gut remains the most intensely studied body site.

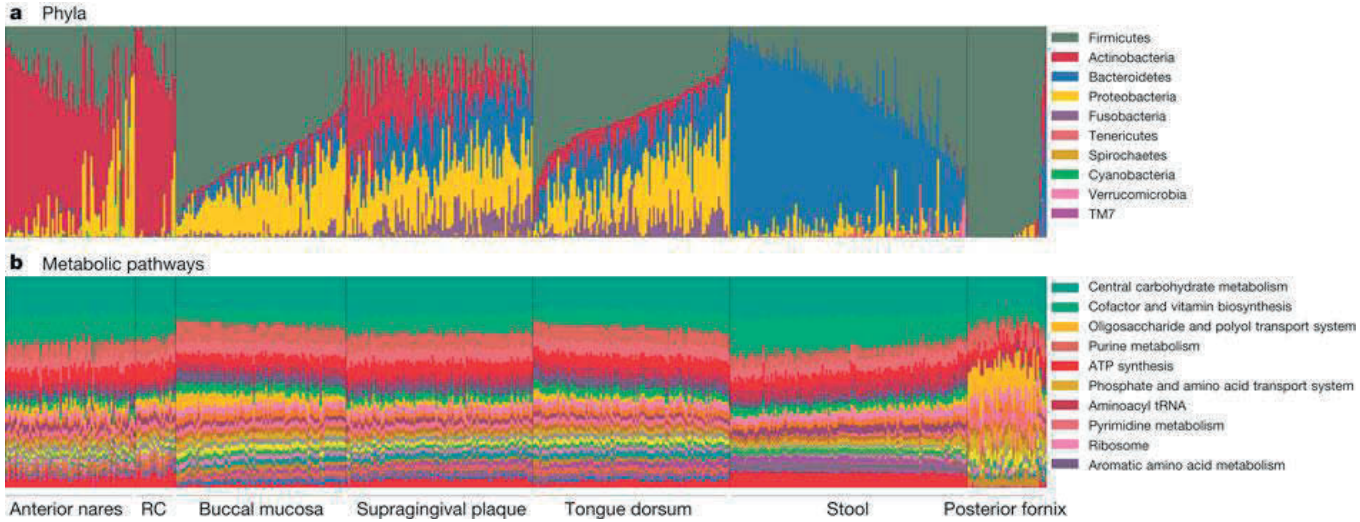

Figure 1: Carriage of microbial taxa varies while metabolic pathways remain stable within a healthy population. Vertical stacked bar charts represent relative abundance from operational taxonomic units (OTUs) obtained from 16S data (a) and metabolic modules obtained from metagenomic shotgun sequencing data (b). Legends indicate the, on average, most abundant phyla and pathways over the different body habitats. RC: retroauricular crease. Retrieved from the HMP consortium publication[4].

Development of the gut microbiome research field

The gut microbiome research field has emerged as an independent research field over the last fifteen to twenty years, even though researchers have already hypothesized about the role of the gut microbiome for far longer. For example, Theodor Escherich (after whom *Escherichia coli* was named) stated that it was crucial to study the microorganisms in the gut to understand (patho)physiological processes in the intestine in a publication from 1885[9]. The major breakthroughs in this field have been made possible by the advent of next-generation sequencing (NGS) technologies, which became more widely adopted in the first decade of this century due to the decreasing costs. NGS allows for characterization of entire microbial communities, which was not possible with traditional microbiological methods (e.g. culturing) or older sequencing methods such as Sanger sequencing. It should be noted that shotgun metagenomic sequencing (one variant of NGS) was already applied on environmental samples (by Craig Venter, amongst others) before it was widely adopted by the human microbiome field[10].

One of the early milestone papers in the gut microbiome field is a study by Turnbaugh et al. Here, the authors showed that obese mice had a gut microbiome with increased capability for energy harvest from the diet and causally linked the gut microbiome to the pathophysiology of obesity through a series of elegant experiments[11]. This included transplanting feces from obese mice into gnotobiotic mice, which led to a greater increase in body fat than when gnotobiotic mice received a fecal microbiota transplantation from lean mice. This study was one of the first to not only find a correlation between the gut microbiome and disease, but to causally link the two, and subsequently triggered a global interest in the role of the gut microbiome in human health and disease.

Most studies in the early days of microbiome research were observational studies where 16S rRNA gene amplicon sequencing of hypervariable regions was performed to compare patient groups, and differences in gut microbiota composition would be associated with disease or health parameters. Using 16S rRNA as an evolutionary marker for classifying bacteria was proposed by Carl Woese and George Fox for the first time in 1977 and preceded the first efficient sequencing technique for the 16S rRNA gene by almost ten years[12, 13]. While 16S rRNA gene amplicon sequencing can be highly valuable for understanding differences in microbiota composition, sequencing of the 16S rRNA gene only provides accurate taxonomic classification up to the genus level and does not provide functional information. Deeper resolution, or application of different methods, is necessary to obtain a more systemic image of the composition and function of a microbial community[14, 15]. In recent years there has been an increase in the number of studies employing metagenomics (sequencing of all DNA in a sample), metabolomics (measuring the metabolites in a sample) and to a smaller extent metatranscriptomics

(sequencing all RNA in a sample) and metaproteomics (measuring all proteins in a sample). Metagenomics allows for accurate taxonomic classification at species level, and sometimes strain level, and for profiling functional potential. While presence of a gene can be detected using metagenomics, this does not necessarily mean that the gene is expressed. This is why metatranscriptomics and metaproteomics are becoming increasingly important, as these techniques directly measure transcripts and proteins[15]. In addition, the metabolome is viewed as a functional readout of microbial metabolism and provides an important link between composition and function (Figure 2)[16].

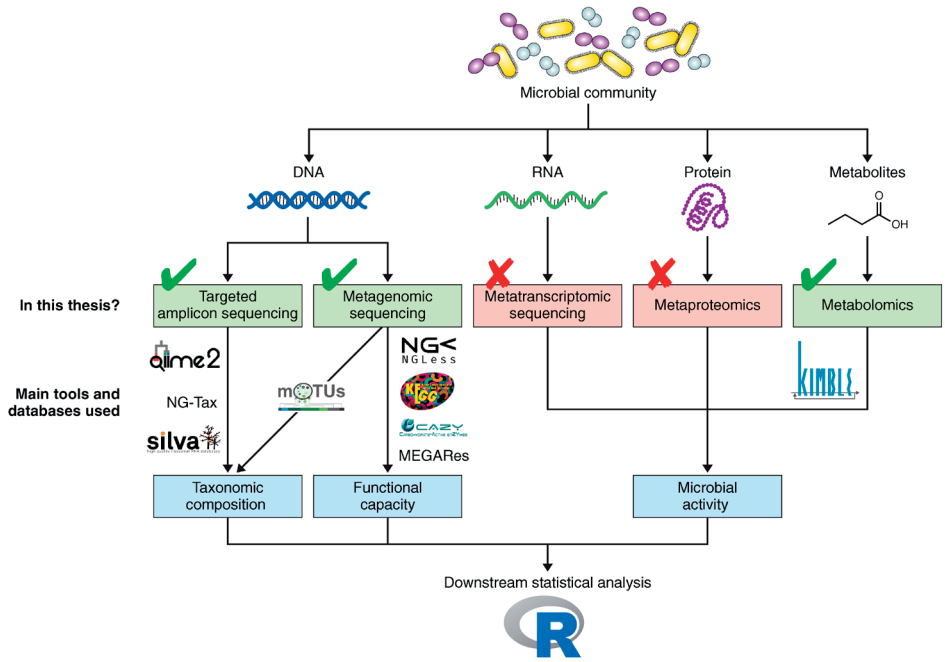

Figure 2: Multi-omics to investigate microbial communities. Each method provides specific information about the community and methods are generally complementary. The computational tools used in this thesis to process raw data are indicated, as well as the main tool used for statistical analysis (R). The reason for choosing these tools is further explained in the section below and in the respective chapters where they are employed. Logos of the tools are obtained from their respective publications and corresponding material[17-26].

Application of -omics techniques (metagenomics, metabolomics, metatranscriptomics, metaproteomics) pose bioinformatic and computational challenges. Expert knowledge is generally necessary to process raw data obtained from these techniques and extensive computational infrastructure can be required. At the Leiden University Medical Center, researchers are fortunate enough to have the luxury of working on a high-performance computing cluster, which allows for processing of large amounts of (sequencing) data.

Shotgun metagenomics and 16S rRNA gene amplicon sequencing data have been processed in this thesis using a variety of techniques (Figure 2). For 16S rRNA gene amplicon data we have used and evaluated two different tools (QIIME 2 and NG-Tax)[17, 18] and concluded that both methods work very well in combination with the SILVA database, which is the most often used database for 16S rRNA gene amplicon data[19, 27]. An enormous variety of tools exist for taxonomic profiling of shotgun metagenomic data[28]. In this thesis we opted for the mOTUs tool, as it uses single-copy marker genes for taxonomic profiling and thereby allows (as one of the very few, if not the only tool) for accurate estimation of bacterial cell numbers[20]. For functional profiling, we used the golden standard databases for metabolism (Kyoto Encyclopedia of Genes and Genomes, KEGG) and carbohydrate-active enzymes (CAZy)[22, 23]. Lastly, for resistome profiling we opted for the MEGARes 2.0 database, as it has manually curated hierarchical annotation from antimicrobial resistance genes to antimicrobial resistance mechanisms which greatly facilitate interpretation of results[24].

After data pre-processing, when matrices of e.g. bacterial species or metabolites are obtained, these matrices generally contain hundreds to thousands of features. This requires the use of specialized statistical software such as R[26] and advanced statistical techniques which can deal with the 'curse of dimensionality', whereby more features than samples are present. It needs to be emphasized however that integration of multiple -omics techniques should not be the endpoint of a microbiome study, but that findings should be taken back into the wet lab. Before taking findings back into the wet lab, it is important to be as confident as possible about computational findings, and ideally these would be confirmed by re-using data from previously conducted studies on a similar topic. This is currently often hampered by the use of different methods between research groups, which by itself can induce large variation in outcomes.

Technical opportunities and challenges for the (gut) microbiome field

Standardization of sample processing methods

Performing a clinical microbiome study typically involves multiple steps including sample collection, sample processing and choice of DNA extraction method and sequencing method[27, 29]. The use of different methods at each step in the workflow of a microbiome study complicates comparing results from different studies, as these technical factors affect the obtained profiles[29]. Research consortia have been set up to identify an optimal workflow for processing fecal samples, but this has not led to its widespread adoption across the research community[29]. This is unfortunate, as this would allow for more efficient re-use of (sequence) data from studies. Re-use of data becomes

crucial when researchers perform meta-analyses to identify robust disease-associated microbial signatures. While meta-analyses have been highly informative despite technical differences between studies[30], such an approach may fail when the disease or variable of interest is associated with more subtle changes in the microbiome. In such cases, technical variation may overshadow the biological signal. Standardization of sample processing methods would facilitate meta-analyses to allow for identification of both prominent and subtle disease-associated microbial signatures.

Biological samples with low bacterial biomass: contamination versus biological signal

In recent years, the concept of contamination (the occurrence of sequence reads in a sample which belonged to a microbe not originally present in the sample) has gained recognition[31]. When conducting a microbiome study using biological samples with a low bacterial biomass, like tumor tissue or urine, contamination can pose huge challenges. Nowadays, an increasing number of researchers is including positive and negative controls into their microbiome studies, which is an encouraging trend. At the Center for Microbiome Analyses and Therapeutics, we always include positive controls in the form of mock communities (both cell-based and DNA-based) and negative controls in the form of blank DNA extractions and blank samples for sequencing. For low-biomass samples, it may also be important to include negative controls during sample collection, although it should be noted that this is not always feasible. The inclusion of such controls in other studies has, amongst others, led to debunking of the claim of the existence of both a placental microbiome and a brain microbiome[32, 33]. In table 1 studies that investigated low-biomass samples but did not include appropriate controls are listed (which, importantly, does not necessarily mean that results are not valid)[34].

Table 1: An overview of ten studies which did not report the use of appropriate controls, thereby making it impossible to properly judge the reported results. As can be seen in the last column, a variety of low-bacterial biomass samples is studied for containing a potential microbiome.

Authors	Year	Journal	Investigated location
Aagaard et al.	2014	*Science*	Placenta
Schierwagen et al.	2019	*Gut*	Blood
Al Alam et al.	2020	*The American Journal of Respiratory and Critical Care Medicine*	Fetal lung tissue
Branton et al.	2013	*PloS One*	Brain
Gosiewksi et al.	2017	*European Journal of Clinical Microbiology & Infectious Diseases*	Blood
Willis et al.	2020	*Scientific Reports*	Eye tears
Hieken et al.	2016	*Scientific Reports*	Breast tissue
Borewicz et al.	2013	*FEMS Microbiology Letters*	Bronchoalveolar lavage fluid
Cavarretta et al.	2017	*European Urology*	Prostate tissue
Fouts et al.	2012	*Journal of Translational Medicine*	Urine

Another debated topic is the existence of a tumor microbiome and its function, which potentially has strong clinical relevance[35]. A recent and very extensive study has looked into the tumor microbiome at several locations in the human body[36]. While this paper took all possible precautions during sample collection, processing and data analysis to avoid and exclude contaminations, it is not unlikely that some contamination signal has ended up in their final data[36]. It must be noted here that it remains unclear for now what to do with information provided by positive and negative controls for potentially correcting microbiota profiles. Nevertheless, important conclusions can be drawn from these controls. For positive controls, it can be judged whether different steps in the workflow (e.g. DNA extraction and sequencing) can induce technical variation. Negative controls are especially valuable for interpretation of low-bacterial biomass samples, as negative control profiles can be compared with those of the low-biomass samples. In case these are highly similar, this suggests that the microbiota profile of the low-bacterial biomass sample is not reflecting a biological profile, but may rather be a result of contamination. Some methods have been developed to 'clean' potential contaminants from microbiome data based on control data, but no consensus has been reached in the scientific community on how to exactly deal with contamination in low-biomass samples[37-39]. Therefore, at this point, controls mainly serve to verify whether DNA extraction, sequencing and bioinformatic processing have been conducted successfully. In conclusion, it remains highly challenging to separate contamination signals from biological signals in samples with a low bacterial biomass and an important future challenge of the microbiome field is to discover what represents real biology in these cases.

State-of-the-art computational methods to profile microbiomes

Traditionally, the first step of a microbiome study after obtaining sequence data involves the accurate identification and estimation of relative abundance (taxonomic profiling) of the microorganisms in a sample. The most often applied technique for this purpose is 16S rRNA gene amplicon sequencing. This usually involves amplifying a short hypervariable region of the 16S rRNA gene and hereby profiles the bacterial fraction of the microbiota and provides accurate identification up to genus level. Advantages of this method are the relatively low costs and lower complexity as compared to metagenomic sequencing. Species level classification through 16S rRNA gene sequencing could possibly be achieved by the advent of long-read sequencing techniques, but this is not commonly implemented yet[40]. However, the most often used technique for obtaining species level resolution in a microbiome is metagenomic shotgun sequencing.

Metagenomics allows for deep resolution (accurate classification of bacterial species, and sometimes strains) and for insight into the functional potential of the microbiome. By sequencing all DNA in a sample, information is also obtained about other microorganisms than bacteria, although in feces, this is usually only a minor fraction of

reads as compared to bacteria. Therefore, tools for taxonomic profiling of metagenomes are currently largely focused on profiling the bacterial fraction of the microbiome.

Many different tools exist for taxonomic profiling of metagenomes, but they can broadly be divided into assembly-based methods (assembling short reads into larger contigs and classifying these larger contigs to a reference database) and read-based methods (assign reads to taxa by using e.g. specific marker genes)[28]. The selection of specific marker genes for taxonomic profiling is not trivial, but ideally they are universal single-copy markers and phylogenetically informative. A major advantage of single-copy markers is that no correction for genome size of each microbe is required and a closer value to the 'real' relative abundance of (bacterial) cell counts can obtained. As for 16S rRNA gene sequencing, the advent of long-read sequencing techniques may become an important tool in metagenomics, as it may allow for achieving circular bacterial genomes using assembly-based methods[41]. Functional profiling of metagenomes is a more complex and computationally intensive task than taxonomic profiling, as one needs to take all reads into account and can only focus on a subset of genes when specific functionalities are searched for. One method for functional profiling is mapping reads to a specific gene catalog relevant for the sample under investigation, for example the Integrated Gene Catalog for the gut microbiome[42], although other options exist[6, 43]. These genes can then be grouped into more informative functional groups, for example KEGG orthology (KO) groups or into carbohydrate-active enzyme (CAZymes) groups. In the context of colorectal cancer (CRC), functional profiling allowed for detecting a shift from carbohydrate degradation in a healthy microbiome towards amino acid degradation in CRC[30]. After obtaining taxonomic and functional profiles, statistics should be performed on the obtained matrices to answer the relevant research question and to link the microbiome to health or disease.

Lack of golden standards for statistical analysis

After having processed raw sequencing data, researchers are faced with the challenge of analyzing complex microbiome data. This usually involves, among others, testing differences in relative abundance of microbial taxa between groups or associating clinical variables with microbiota composition. However, there are no clear guidelines or golden standards for performing such analyses. For example, for a relatively common procedure such as differential abundance testing, many different tests are available and expert opinions differ about which tests are optimal[44, 45]. It is probably not possible to define one optimal test for differential abundance testing, as it is likely that the ideal test will depend on the dataset under study. To define an optimal test, one should have simulated data where a ground truth is known (is a taxon differentially abundant or not). However, the question here is how to define a ground truth, as in when is a taxon defined to be differentially abundant? While for standard differential abundance analysis

a wide variety of tools is available, this is not the case for all analyses that researchers wish to perform. Currently, microbiome studies are moving towards longitudinal data collection and towards causality instead of correlations. However, tools for longitudinal analysis specifically adapted to microbiome data (e.g. taking into account zero-inflation) are scarce and currently available ones are probably insufficient to capture the full complexity of the dynamics of e.g. the gut microbiome. Some tools (e.g. MetaLonDA and MetaDprof) are available which test for differences in microbial taxa over time between different groups[46, 47], but there are no such tools available that also allow for incorporation of covariates, which would be an important next step in development of statistical methods for the microbiome field.

Microbiome-mediated colonization resistance

The gut microbiome has a myriad of functions which are important for maintaining human health, and among these functions is providing colonization resistance against incoming, potentially pathogenic, microorganisms[8]. The notion that the gut microbiota can defend against enteric bacterial pathogens is far from new. For example, a paper from 1962 described that when mice are given streptomycin prior to oral administration of *Salmonella enteridis*, the resistance against this pathogen became 100,000 fold lower, with less than 10 *S. enteridis* cells being able to cause an infection in 63% of mice. In contrast, when no prior antibiotics were administered, a dose of approximately one million *S. enteridis* cells was required to infect the same percentage of mice[48].

I have previously defined colonization resistance as the ability of the microbiome to prevent colonization by exogenous microorganisms[8]. While in literature this mostly refers to incoming bacterial pathogens[49, 50], in my opinion colonization resistance to incoming commensal bacteria or other microorganisms such as viruses, fungi and even parasitic worms should also be considered. Gut microbiome-mediated colonization resistance can be conferred through several mechanisms, including nutrient competition and production of antimicrobial compounds[8]. However, the complete set of mechanisms through which microbiome-mediated colonization resistance is conferred is not completely clear yet and it is very likely that required mechanisms are different against different (pathogenic) microorganisms. It is critical to make a distinction between asymptomatic colonization by a potentially pathogenic microorganism and actual infection with enteropathogenic microorganisms whereby the pathogen causes intestinal disease. It has been shown that colonization of pathogenic bacteria often precedes overt infection[51]. Therefore, this stage might be the ideal period for intervention to prevent infection, especially in vulnerable populations such as hospitalized patients and nursing home residents. These vulnerable populations usually receive a wide array of medication, including antibiotics, which are

able to profoundly impact the gut microbiome and decrease colonization resistance[8, 52]. This concept is especially well established in the case of *Clostridioides difficile*, as *C. difficile* infection (CDI) is often observed after an antibiotic treatment. The altered microbial environment through antibiotic administrations can allow *C. difficile* to outgrow and cause infection[53]. In a landmark paper from 2013 it was shown that restoring the gut microbiome through fecal microbiota transplantation (FMT) is highly effective for curing recurrent CDI, with cure rates up to 89%[54, 55]. This publication has paved the way for development of microbiome-based therapeutics. However, before development of such products can start, fundamental research is necessary for generating insight into which commensal microorganisms can provide protection against enteropathogens.

Thesis aim

The research described in this thesis aims at identifying bacteria with potential antagonistic properties against pathogenic microorganisms and antibiotic resistant bacteria, and to address and contribute to technical challenges and opportunities in the microbiome research field.

Research questions and thesis outline

The research described in this thesis can be divided into three parts. First, we aimed to summarize the current knowledge of microbiome-mediated colonization resistance against enteropathogens (**Chapter 2**) and to provide an overview of opportunities and challenges in development of microbiome therapeutics against such pathogens (**Chapter 3**). In the second part, we focused on method optimization for microbiome research, both for wet-lab and dry-lab procedures (**Chapter 4** and **Chapter 5**). The final part describes changes in the human gut microbiota during infection or asymptomatic colonization by potentially pathogenic enteropathogens, including hookworm (**Chapter 6**), *C. difficile* (**Chapter 7**) and multidrug-resistant bacteria (**Chapters 8 and 9**). We hypothesized that we could identify bacteria or bacterial metabolites that are involved in providing microbiome-mediated colonization resistance against these pathogens. Specific research questions that we aimed to answer in this thesis were the following:

1. What is the current knowledge on microbiome-mediated colonization resistance against enteropathogenic bacteria?
2. What are the current opportunities and challenges in development of microbiome therapeutics against enteropathogenic and antibiotic-resistant bacteria, and how can we translate these into well-designed studies?
3. What is the impact of different DNA extraction procedures and different

bioinformatic pipelines for the obtained microbiota profile? How do positive and negative controls affect interpretation of microbiota profiles for low-bacterial biomass samples?

4. How do we optimize detection of carbohydrate-active enzymes (CAZymes) from (meta)genomic data using Hidden Markov models? What is the difference in CAZyme repertoire in colorectal cancer patients and are these differences independent of the geographical area of the study?
5. How does an infection with the helminth *Necator americanus* affect temporal dynamics of the human gut microbiota?
6. Can we identify bacteria that are associated with protection from asymptomatic colonization by *C. difficile*? Is it possible to understand, based on gut microbiota composition, why some individuals develop *C. difficile* infection but others only remain asymptomatically colonized?
7. Is there a role for microbiome-mediated colonization resistance against asymptomatic gut colonization of MDROs in nursing home residents? Is there spread of MDROs in this nursing home?
8. Is the gut microbiome involved in providing resistance against colonization by extended-spectrum beta-lactamase producing *Escherichia coli* in the general Dutch population?

In more detail:

Chapter 2 aimed to summarize the main mechanisms by which the gut microbiome can provide colonization resistance against enteric bacterial pathogens (nutrient competition, production of antibacterial compounds, maintenance of a healthy mucus layer and bacteriophage deployment). An important research field developing over the last few years is the effect of medication on gut microbiome function, and this chapter therefore also describes the effects of non-antibiotic medication on impacting the ability of the gut microbiome to provide colonization resistance. Lastly, it is explained how eight of the most common enteric bacterial pathogens have developed mechanisms to subvert microbiome-mediated defensive mechanisms, so that they are able to colonize the gut and cause infection.

Chapter 3 was written with the purpose of reviewing the practical aspects for development of live biotherapeutic products (LBP)s to protect against and/or cure bacterial enteric infection or colonization by multi-drug resistant organisms (MDROs). This type of medication offers an excellent alternative to conventional antibiotic therapy, as it does not damage the native microbiota and does not contribute to development of antibiotic resistance. Emergence of highly antibiotic-resistant pathogens are of ever increasing clinical importance, and solutions are urgently required for this, with LBPs being a promising option. For this chapter, we collaborated with experts from Vedanta

Biosciences Inc., a biotech company that developed several LBPs for treatment of inflammatory bowel diseases and CDI that are currently tested in phase II and phase III clinical trials.

Chapter 4 was designed with the rationale to investigate how variations in the workflow of a microbiome study can impact the obtained microbiota profiles. To this end, three different DNA extraction protocols were compared in combination with two bioinformatic pipelines. In addition, we included positive and negative controls in the workflow, an often overlooked matter in microbiome research. We hypothesized that different extraction methods and bioinformatic pipelines would lead to technical variation, but that the biological conclusions would remain the same.

Chapter 5 describes the development of a novel bioinformatic tool which profiles CAZymes in the human gut from shotgun metagenomic data. To this end, we aimed to extensively optimize settings of Hidden Markov models, annotate the Integrated Gene Catalog with CAZymes and design a novel annotation scheme for substrate specificity. The rationale for designing a novel annotation scheme was that it can be confusing to deal with a large list of different CAZymes, rather than informative functional annotation (e.g. dietary fiber metabolism). Lastly, we applied this tool in metagenomes of colorectal cancer cohorts to identify colorectal cancer-specific CAZyme signatures. With regard to these cohorts, we expected to see a decrease in fiber-degrading CAZymes in the colorectal cancer patients, as epidemiological studies strongly suggest a link between dietary fiber consumption and colorectal cancer development.

For **Chapter 6** we aimed to investigate the effect of a helminth (*Necator americanus*) infection on the bacterial gut microbiota and vice-versa. This helminth is highly prevalent in third-world countries and resides in the duodenum. We used a controlled human infection model, in which human volunteers were infected with this helminth and followed longitudinally. This helminth is highly prevalent in third-world countries and resides in the duodenum. We hypothesized that colonization and infection rates of *N. americanus* would be associated with gut microbiota composition.

Chapter 7 describes a cross-sectional study in which the bacterial gut microbiota of three groups was compared, namely CDI patients, hospitalized patients asymptomatically colonized with *C. difficile* and a control group of hospitalized patients without *C. difficile*. The aim of this study was to investigate whether specific bacterial signatures were associated with resistance against asymptomatic *C. difficile* colonization and against development of CDI. We hypothesized that patients asymptomatically colonized with *C. difficile* would have a different microbiota as compared to patients who were not.

For **Chapter 8** we conducted a point-prevalence study with four time points performed in a Dutch nursing home, where we analyzed microbiota-associated risk factors for asymptomatic MDRO colonization in a cross-sectional and longitudinal manner. In addition, we aimed to identify clinical risk factors for MDRO colonization, to investigate MDRO spread within the nursing home using whole-genome sequencing and we further investigated unexpected findings from 16S rRNA sequencing of the gut microbiota using metagenomic sequencing.

Chapter 9 describes a study which aimed to elucidate whether the microbiome provides resistance against asymptomatic gut colonization by ESBL-producing *E. coli* in adults in the general Dutch population. To this end, we collected paired fecal metagenomics and metabolomics data from individuals, who were, or were not, colonized by this bacterium. This study is unique in the sense that we were able to select samples from a large Dutch population cohort (PIENTER-3). In this way, we could exclude many common confounding factors encountered in gut microbiome research and match colonized individuals to non-colonized individuals on several clinic variables (age, sex, travel history and ethnicity).

Chapter 10 contains the general discussion of the research presented in this thesis, and describes future research directions which are crucial for advancing the microbiome field in the author's opinion.

References

1. Sender R, Fuchs S, Milo R. Are We Really Vastly Outnumbered? Revisiting the Ratio of Bacterial to Host Cells in Humans. Cell. 2016;164(3):337-40.
2. Berg G, Rybakova D, Fischer D, Cernava T, Vergès MC, Charles T, et al. Microbiome definition re-visited: old concepts and new challenges. Microbiome. 2020;8(1):103.
3. Team NHMPA. A review of 10 years of human microbiome research activities at the US National Institutes of Health, Fiscal Years 2007-2016. Microbiome. 2019;7(1):31.
4. Consortium HMP. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486(7402):207-14.
5. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. Nat Methods. 2015;12(10):902-3.
6. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. Nat Methods. 2018;15(11):962-8.
7. Sommer F, Backhed F. The gut microbiota--masters of host development and physiology. Nat Rev Microbiol. 2013;11(4):227-38.
8. Ducarmon QR, Zwittink RD, Hornung BVH, van Schaik W, Young VB, Kuijper EJ. Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. Microbiol Mol Biol Rev. 2019;83(3).
9. Escherich T. Die darmbakterien des neugeborenen und säuglings. Fortschr Med. 1885;3(515-522):547-54.
10. Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, et al. Environmental genome shotgun sequencing of the Sargasso Sea. Science. 2004;304(5667):66-74.
11. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. 2006;444(7122):1027-31.
12. Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A. 1977;74(11):5088-90.
13. Lane DJ, Pace B, Olsen GJ, Stahl DA, Sogin ML, Pace NR. Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. Proc Natl Acad Sci U S A. 1985;82(20):6955-9.
14. Knight R, Vrbanac A, Taylor BC, Aksenov A, Callewaert C, Debelius J, et al. Best practices for analysing microbiomes. Nat Rev Microbiol. 2018;16(7):410-22.
15. Zhang X, Li L, Butcher J, Stintzi A, Figeys D. Advancing functional and translational microbiome research using meta-omics approaches. Microbiome. 2019;7(1):154.
16. Zierer J, Jackson MA, Kastenmüller G, Mangino M, Long T, Telenti A, et al. The fecal metabolome as a functional readout of the gut microbiome. Nat Genet. 2018;50(6):790-5.
17. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37(8):852-7.
18. Ramiro-Garcia J, Hermes G, Giatsis C, Sipkema D, Zoetendal E, Schaap P, et al. NG-Tax, a highly accurate and validated pipeline for analysis of 16S rRNA amplicons from complex biomes [version 2; peer review: 2 approved, 1 approved with reservations, 1 not approved]. 2018;5(1791).
19. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41(Database issue):D590-6.
20. Milanese A, Mende DR, Paoli L, Salazar G, Ruscheweyh HJ, Cuenca M, et al. Microbial abundance, activity and population genomic profiling with mOTUs2. Nat Commun. 2019;10(1):1014.
21. Coelho LP, Alves R, Monteiro P, Huerta-Cepas J, Freitas AT, Bork P. NG-meta-profiler: fast processing of metagenomes using NGLess, a domain-specific language. Microbiome. 2019;7(1):84.
22. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016;44(D1):D457-62.
23. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014;42(Database issue):D490-5.
24. Doster E, Lakin SM, Dean CJ, Wolfe C, Young JG, Boucher C, et al. MEGARes 2.0: a database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data. Nucleic Acids Res. 2020;48(D1):D561-d9.
25. Verhoeven A, Giera M, Mayboroda OA. KIMBLE: A versatile visual NMR metabolomics workbench in KNIME. Anal Chim Acta. 2018;1044:66-76.
26. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2020. https://www.r-project.org/2020.

27. Ducarmon QR, Hornung BVH, Geelen AR, Kuijper EJ, Zwittink RD. Toward Standards in Clinical Microbiota Studies: Comparison of Three DNA Extraction Methods and Two Bioinformatic Pipelines. mSystems. 2020;5(1).
28. Ye SH, Siddle KJ, Park DJ, Sabeti PC. Benchmarking Metagenomics Tools for Taxonomic Classification. Cell. 2019;178(4):779-94.
29. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, et al. Towards standards for human fecal sample processing in metagenomic studies. Nat Biotechnol. 2017;35(11):1069-76.
30. Wirbel J, Pyl PT, Kartal E, Zych K, Kashani A, Milanese A, et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. Nat Med. 2019;25(4):679-89.
31. Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. Trends Microbiol. 2019;27(2):105-17.
32. de Goffau MC, Lager S, Sovio U, Gaccioli F, Cook E, Peacock SJ, et al. Human placenta has no microbiome but can contain potential pathogens. Nature. 2019;572(7769):329-34.
33. Bedarf JR, Beraza N, Khazneh H, Özkurt E, Baker D, Borger V, et al. Much ado about nothing? Off-target amplification can lead to false-positive bacterial brain microbiome detection in healthy and Parkinson's disease individuals. Microbiome. 2021;9(1):75.
34. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. The placenta harbors a unique microbiome. Sci Transl Med. 2014;6(237):237ra65.
35. Zhao K, Hu Y. Microbiome harbored within tumors: a new chance to revisit our understanding of cancer pathogenesis and treatment. Signal Transduct Target Ther. 2020;5(1):136.
36. Nejman D, Livyatan I, Fuks G, Gavert N, Zwang Y, Geller LT, et al. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. Science. 2020;368(6494):973-80.
37. Karstens L, Asquith M, Davin S, Fair D, Gregory WT, Wolfe AJ, et al. Controlling for Contaminants in Low-Biomass 16S rRNA Gene Sequencing Experiments. mSystems. 2019;4(4).
38. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome. 2018;6(1):226.
39. McKnight DT, Huerlimann R, Bower DS, Schwarzkopf L, Alford RA, Zenger KR. microDecon: A highly accurate read□subtraction tool for the post□sequencing removal of contamination in metabarcoding studies. J Environmental DNA. 2019;1(1):14-25.
40. Earl JP, Adappa ND, Krol J, Bhat AS, Balashov S, Ehrlich RL, et al. Species-level bacterial community profiling of the healthy sinonasal microbiome using Pacific Biosciences sequencing of full-length 16S rRNA genes. Microbiome. 2018;6(1):190.
41. Moss EL, Maghini DG, Bhatt AS. Complete, closed bacterial genomes from microbiomes using nanopore sequencing. Nat Biotechnol. 2020;38(6):701-7.
42. Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, et al. An integrated catalog of reference genes in the human gut microbiome. Nat Biotechnol. 2014;32(8):834-41.
43. Niu SY, Yang J, McDermaid A, Zhao J, Kang Y, Ma Q. Bioinformatics tools for quantitative and functional metagenome and metatranscriptome data analysis in microbes. Brief Bioinform. 2018;19(6):1415-29.
44. Lin H, Peddada SD. Analysis of microbial compositions: a review of normalization and differential abundance analysis. NPJ Biofilms Microbiomes. 2020;6(1):60.
45. Calgaro M, Romualdi C, Waldron L, Risso D, Vitulo N. Assessment of statistical methods from single cell, bulk RNA-seq, and metagenomics applied to microbiome data. Genome Biol. 2020;21(1):191.
46. Metwally AA, Yang J, Ascoli C, Dai Y, Finn PW, Perkins DL. MetaLonDA: a flexible R package for identifying time intervals of differentially abundant features in metagenomic longitudinal studies. Microbiome. 2018;6(1):32.
47. Luo D, Ziebell S, An L. An informative approach on differential abundance analysis for time-course metagenomic sequencing data. Bioinformatics. 2017;33(9):1286-92.
48. Bohnhoff M, Miller CP. Enhanced susceptibility to Salmonella infection in streptomycin-treated mice. J Infect Dis. 1962;111:117-27.
49. Sorbara MT, Pamer EG. Interbacterial mechanisms of colonization resistance and the strategies pathogens use to overcome them. Mucosal Immunol. 2019;12(1):1-9.
50. Buffie CG, Pamer EG. Microbiota-mediated colonization resistance against intestinal pathogens. Nat Rev Immunol. 2013;13(11):790-801.
51. Cassone M, Mody L. Colonization with Multi-Drug Resistant Organisms in Nursing Homes: Scope, Importance, and Management. Curr Geriatr Rep. 2015;4(1):87-95.
52. Kim S, Covington A, Pamer EG. The intestinal microbiota: Antibiotics, colonization resistance, and enteric

pathogens. Immunol Rev. 2017;279(1):90-105.
53. Smits WK, Lyras D, Lacy DB, Wilcox MH, Kuijper EJ. Clostridium difficile infection. Nat Rev Dis Primers. 2016;2:16020.
54. van Nood E, Vrieze A, Nieuwdorp M, Fuentes S, Zoetendal EG, de Vos WM, et al. Duodenal infusion of donor feces for recurrent Clostridium difficile. N Engl J Med. 2013;368(5):407-15.
55. Terveer EM, Vendrik KE, Ooijevaar RE, Lingen EV, Boeije-Koppenol E, Nood EV, et al. Faecal microbiota transplantation for Clostridioides difficile infection: Four years' experience of the Netherlands Donor Feces Bank. United European Gastroenterol J. 2020:2050640620957765.