Universiteit
Leiden
The Netherlands

Guardians of the gut: harnessing bioinformatics to study the gut microbiome and faecal microbiota transplantation in intestinal disorders

Nooij, S.

Citation
Nooij, S. (2025, October 10). *Guardians of the gut: harnessing bioinformatics to study the gut microbiome and faecal microbiota transplantation in intestinal disorders*. Retrieved from https://hdl.handle.net/1887/4262800

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/1887/4262800

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


サム

Chapter 1

General introduction

The human microbiome

The microbiome has enjoyed great interest in recent years. The term 'microbiome' describes a community of microorganisms living in a specific environment. This comprises bacteria, viruses, fungi, and other microscopic eukaryotes, but bacteria are often the primary focus. Microbiomes can exist anywhere, from environments like soil to the bodies of animals. Of all microbiomes, the human microbiome is perhaps the most captivating and best known. The human body harbours roughly as many bacterial cells as human cells[1], and over a hundred times more microbial genes than are encoded in the human genome[2]. Bacteria are spread across multiple body sites that each offer their own environmental conditions to which microbial communities have adapted (Figure 1). For example, researchers have identified different microbiomes on the skin, in the airways or respiratory tract, and the gut or gastrointestinal tract[3,4]. The gastrointestinal tract encompasses a large surface area covered with a vast diversity of microorganisms[5], and has been the subject of extensive study[6]. In fact, among human microbiome studies, the gut microbiome is the best-studied body site.

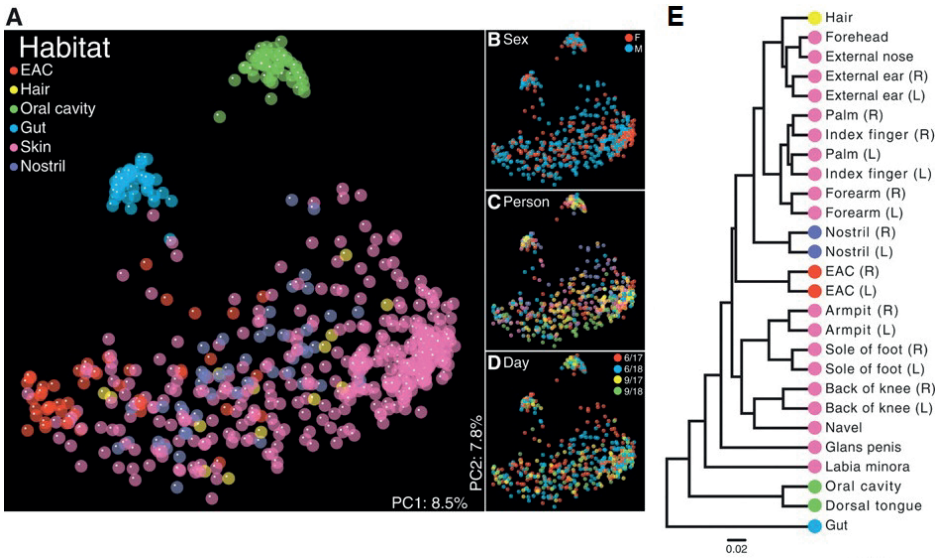

Figure 1. The human microbiome differs per body site. Clustering of microbiome samples based on taxonomic composition reveals that body site or habitat explains the variation most clearly. A) Microbial communities coloured by body site. B-D) the same communities coloured by sex, person or sampling day. E) Hierarchical clustering dendrogram shows that body sites that are closer to one another are more similar in terms of microbial composition. EAC, external auditory canal; L, left; R, right. Figure adapted from reference[4]. Reprinted with permission from AAAS.

Ways of studying the microbiome

The discovery of microorganisms, microscopic organisms invisible to the naked eye, started with the refined microscopes that Antoni van Leeuwenhoek constructed around 1670[7]. It is worth mentioning that Jan Swammerdam and Robert Hooke were also building microscopes at that time, but their magnification was much less than Van Leeuwenhoek's[7]. Up to the present day, microscopes remain an important instrument to study microorganisms and indeed numerous kinds and species, from cyanobacteria to selenomonads, have been identified with microscopes[7,8]. In modern history, the development of DNA sequencing techniques has revolutionised the way we can detect and characterise any type of organism. DNA sequencing was pioneered in the 1960s and -70s by Wu, Gilbert and Maxam[9], after which new techniques were developed and combined, giving rise to the method known as Sanger sequencing (Figure 2)[10,11]. This discovery was deemed so important that Frederick Sanger and Walter Gilbert received a Nobel Prize in Chemistry in 1980. With Sanger sequencing, one could read DNA fragments as a sequence of four nucleotides: A, C, G and T. By collecting multiple reads it was then possible to elucidate increasingly complex DNA molecules. The first completely resolved genome was of bacteriophage lambda (48.5 kilobasepairs (kbp))[12]. The technique was further refined and optimised[13], facilitating increasingly large genome sequencing projects spanning from *Haemophilus influenza* (2 Mbp, 1995)[14], to *Saccharomyces cerevisiae* (12 Mbp, 1996)[15], *Caenorhabditis elegans* (100 Mbp, 1998)[16], and finally, in 2001 the first draft sequence of the human genome was published (*Homo sapiens*, 3.3 Gbp)[17,18]. Concurrently, DNA sequencing methods were further improved and a new generation was created. These so-called next-generation sequencing (NGS) or second generation methods allow for rapid high-throughput sequencing of many DNA molecules, thereby enabling an entirely new field of science. Now, there is a third generation, enabling real-time single-molecule sequencing (Figure 2). Not only have these techniques facilitated the study of complete human genomes (genomics), but it has also enabled researchers to study DNA from multiple organisms in an environment (metagenomics)[19]. This in turn led to the launch of the Human Microbiome Project in 2007[3,20]. The term 'microbiome' is used here to refer to all the genes identified using NGS, although it has been used to mean different things[21]. The microbial species from which the genes derive are commonly referred to as 'microbiota'. NGS methods now yield enormous amounts of data, with a single experiment often generating millions of DNA reads. To process all these data, a new type of biologist was needed: the computational biologist or bioinformatician.

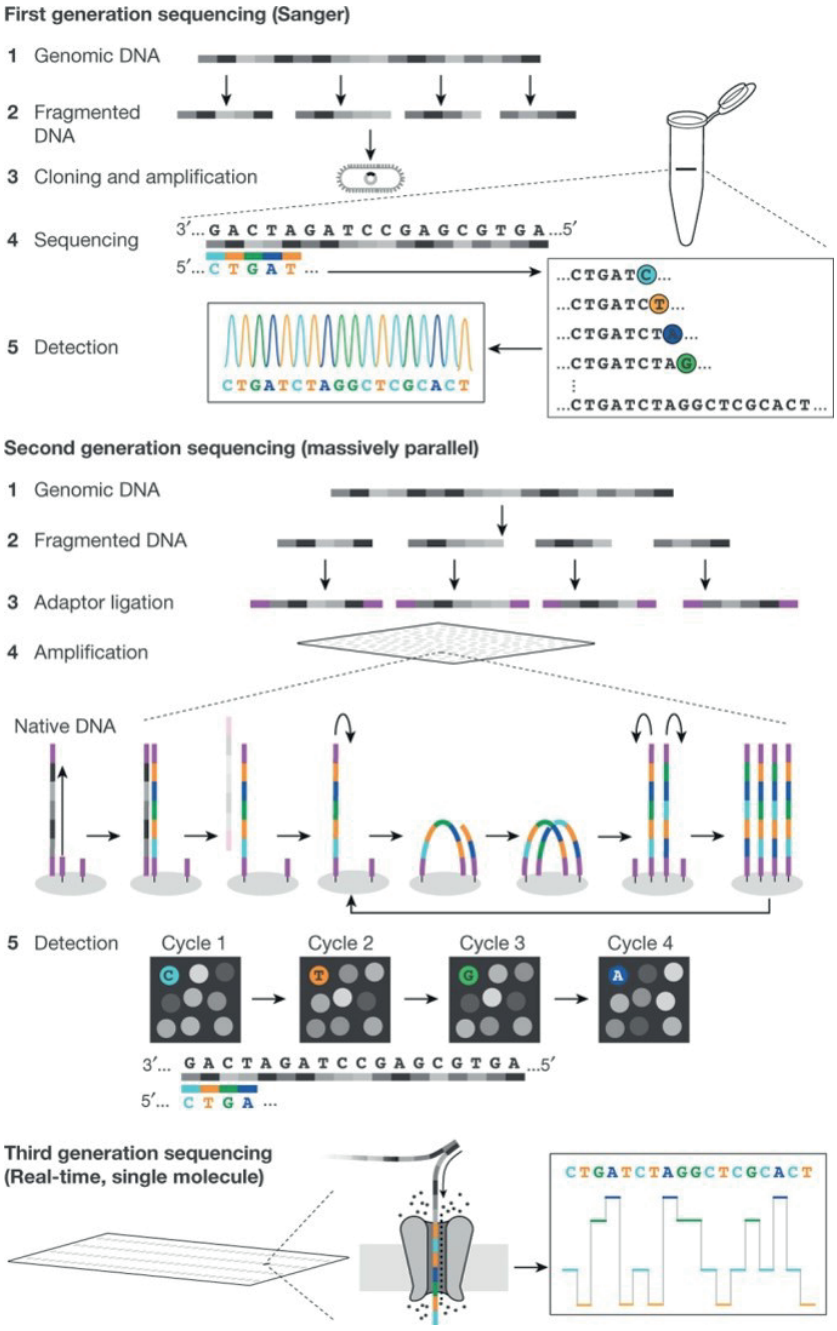

Figure 2. DNA sequencing technologies. Developments of DNA sequencing techniques are listed and illustrated in chronological order from top to bottom. The second generation is generally referred to as next-generation sequencing or NGS. Figure from reference[9].

The computational biologist or bioinformatician

A computational biologist, or bioinformatician, combines skills and knowledge from both biology and computer science[22]. There is some confusion and controversy concerning the terms 'computational biologist' and 'bioinformatician'. However, since there is considerable overlap between the tasks of each I consider them equivalent for the purpose of this text. Bioinformatics as a field has been around since the 1960s, before DNA could be sequenced[23]. Starting out with protein sequence analysis, bioinformatics has expanded to include DNA and many other types of biological data, while benefitting from rapid improvements in computer technology. The advancements in high-throughput biological methods have given rise to biological 'big data'. Due to the high-throughput at which biological data are now generated, the results can no longer be processed and interpreted manually and need increasingly powerful computers and automated workflows. This is the work of a bioinformatician. A bioinformatician needs to understand the biology behind the experiment to formulate the relevant research questions and have a working knowledge of computer science, statistics and/or mathematics to convert those questions into functional computer algorithms, pipelines or tools[24]. The output from computational analyses (often referred to as 'dry lab' work) is typically hypothesis-generating, for instance correlating a specific bacterium or gene to a disease. Consequently, the hypothesis needs to be tested and validated in the so-called wet lab. Therefore, bioinformatics experiments are frequently embedded within larger projects, bringing together collaborators from biomedical backgrounds to computational experts. A good example is the Human Microbiome Project[25], which has run for over a decade and spawned numerous subsidiary studies as well as a large second phase: the Integrative Human Microbiome Project[26]. The Human Microbiome Project was also the largest microbiome project in terms of data generated at the time (Figure 3), underscoring the growing need for data management and computational tools in biology. In a sense, bioinformaticians act as an interpreter in a multidisciplinary team of life science specialists translating between data science and biology. Thus, they must be flexible and able to communicate with people of diverse backgrounds.

Figure 3. Timeline of microbiome projects and size of generated data. Each circle represents a microbiome project and their area and y-coordinate indicate data size in basepairs. The x-coordinate corresponds to the year the study was published and colours separate human (red), animal (black) and environment-associated (green) projects. Figure from reference[25].

Human gut microbiota in health and disease

The advent of second-generation DNA sequencing (NGS) paved the way for a whole new field of microbiology. Researchers were no longer limited to study culturable species, but whole communities were now within reach. Among possible focus areas, the human microbiota has intrigued us for a long time. Early microbiome studies established that microorganisms played key roles in infections, leading to hypotheses as to their involvement in other diseases. The use of metagenomics enabled researchers to compare microbial communities between healthy people and people with various diseases, such as ulcerative colitis and Crohn's disease (UC and CD, together known as inflammatory bowel diseases or IBD), colorectal cancer, obesity, and Parkinson's disease to identify differences in microbiota associated with these conditions[2,27-33]. An important caveat here is that correlation does not imply causation[33]. The presence of certain bacteria in a patient may indicate the cause of disease, or the disease itself could create an environment in which the bacteria can thrive. In this second scenario, the bacteria merely benefit accidentally from conditions brought about by the disease. This was exemplified by a study on type 2 diabetes mellitus in which the researchers found a correlation with the microbiome[34]. Another group later attributed this change to the drug used[35]. Nonetheless, these disease association studies have elucidated the involvement of the microbiota in a number of diseases in which no role for microorganisms was suspected. This has led to the generation of new hypotheses and

further studies into possible mechanisms by which bacteria contribute or exacerbate disease. For instance, discovering the association of *Fusobacterium nucleatum* with colorectal cancer motivated researcher to quickly identify molecular mechanisms by which the bacterium is involved in tumorigenesis[29,36,37].

Sometimes these studies gave conflicting results. Certain bacteria may be disease-associated in one study and associated with good health in another study. These discrepancies may arise from differences in study setup, study population (e.g. geographic location, diet, or other potential confounders[38]), and low statistical power. Another possible explanation can be found in the biology of the bacterium in question. In metagenomics studies, bacteria are often classified to a certain taxonomic level, like genus or species. These taxonomic groups, however, encompass a spectrum of different strains of bacteria. The genomes within a bacterial species are 95% identical[39]. Within the remaining 5% there is ample opportunity for differences that cause fundamental shifts in lifestyle, making one strain pathogenic and the other an innocent symbiont or commensal. It is thought that the exact genetic composition, including the presence or absence of certain genes, and interactions with the environment – in this case its host: the human body – determines the bacterium's functions and the possible implications in disease[2,33]. Conflicting results can be clarified through meta-analyses, which enable more robust statistical evaluation of disease-microbiome associations. This approach is also exemplified in **chapter 7**.

Disturbances to microbiota as marker for disease

For several diseases it is well-established that the microbiota play a fundamental role in disease aetiology. Currently the best example is an infection by the gut pathogen *Clostridioides difficile*, which paradoxically often occurs after antibiotic treatment[40]. The mechanism that underlies this process is that the antibiotics deplete part of the resident microbiota, creating a niche – freeing up space as it were – in which survivors may thrive. *C. difficile* produces spores, which are very resilient and generally unaffected by antibiotics. After antibiotic treatment, these spores germinate and quickly occupy the cleared niche. It then produces toxins which both harm the person in which *C. difficile* resides (causing inflammation), and suppresses other bacteria in the gut[40]. This alteration of the gut microbiota composition is termed 'dysbiosis', describing a deviation from the normal microbial balance and is usually related to disease[41-43]. Even though there is no standard way to describe dysbiosis, several parameters have been used to indicate a deviation in gut microbiota. Consequently, the overall microbiota composition has become an indicator of health. This thesis deals with two microbiota-associated diseases: recurrent *C. difficile* infection (CDI) in **chapters 2 and 3** and IBD in **chapters 4, 5, 6 and 7**.

Inflammatory bowel disease is a chronic condition that is characterised by episodes of gastrointestinal inflammation[44]. IBD encompasses Crohn's disease (CD) and ulcerative colitis (UC). CD typically affects the distal gastrointestinal tract, particularly from the end of the small intestine (ileum) to the colon, but it can involve any part of the intestine in a discontinuous pattern[44]. UC typically affects the distal colon and often includes the rectum, but inflammation can extend proximally to involve larger portions or the entire colon in an uninterrupted pattern[44]. IBD often manifests in early adult life and has a genetic component[45]. It is thought to result from an aberrant inflammatory immune response to commensal gut microbes[45]. Hence, the gut microbiota plays a critical role in the pathogenesis of IBD and it is hypothesised that microbiota modulation could alleviate symptoms in affected patients.

Microbiota therapy: faecal microbiota transplantation (to prevent CDI relapse)

After observing an alteration in gut microbiota, the next step in medical science is to understand how the microbiome can be modulated in such a way that it contributes to disease amelioration or cure. Microbiota-based therapies were already used long before the discovery of microorganisms. Using faecal material to treat diarrhoeal diseases was first described in China in the 4th century, and has supposedly been used throughout Eurasia and North Africa up to the 20th century[46], although the use of camel faeces in North Africa is contested[47]. In recent history, the method has been reintroduced and termed faecal microbiota transplantation (FMT). Initially pioneered in the Netherlands in 2013, this treatment proved highly effective against recurrent *C. difficile* infection and has quickly become the new therapeutic approach[48]. Following this success, FMT has also been attempted with different diseases, with varying results[46,49].

The mechanism underlying FMT is thought to be a restoration of a healthy gut microbiota. This is done by taking the faeces of thoroughly screened healthy volunteer donors, making a suspension from the faeces and transplanting this into the gastrointestinal tract of the patient. This procedure is preceded by antibiotic treatment to eradicate *C. difficile*, but antibiotic pretreatment has not always been used with FMT for other diseases. A new generation of microbiome-modulating therapies are under development, but none have yet reached the stage where they can replace FMT[50-54]. To reach this goal, we need to advance our understanding of human-microbiome interactions and identify specific bacterial strains with therapeutic potential.

Research questions and thesis outline

This thesis describes computational studies of the human gut microbiota with a focus on FMT. The Leiden University Medical Center (LUMC) hosts a dedicated faeces bank, the Netherlands Donor Feces Bank (NDFB), that supplies ready-to-use faecal suspensions for FMT to treating physicians throughout the country. The NDFB is part of the Medical Microbiology and Infection Prevention subdepartment in the Center for Infectious Diseases (LU-CID), creating a multidisciplinary environment to study the microbiota in relation to FMT. Currently, FMT is routinely used to treat multiple recurrent CDI. Besides this, the NDFB has initiated clinical trials for other indications including UC and collaborates with various partners in scientific studies aimed at elucidating the mechanisms underlying FMT. These studies focus primarily on the gut microbiota, employing metagenomic sequencing of faecal samples from healthy donors and recipients before and after FMT. This approach provides a comprehensive view of the microbial composition and its functional potential, enabling exploration of the microbiota and addressing diverse research questions. The main research question of this thesis is: how does FMT affect the recipients' microbiota? The secondary question is: what makes one species variant commensal and the other disease-associated?

Part 1: Concomitant microbiota impacts after faecal microbiota transplantation for recurrent *Clostridioides difficile* infections

Chapter 2 describes a study on putatively procarcinogenic *Escherichia coli* strains within a cohort of recurrent CDI patients and their respective donors. Previous research had established that this *E. coli* strain produces a toxin, colibactin, capable of inducing DNA damage in the host and potentially contributing to the development of colorectal cancer. Furthermore, it was known that the species *E. coli*, and even this specific procarcinogenic strain, is a common inhabitant of the human gut. We screened our metagenomic data for these specific genotoxin genes to evaluate the prevalence and abundance among FMT recipients and their healthy donors. The questions we asked here are: is colibactin-producing *E. coli* present in recurrent CDI patients and healthy faeces donors, and how does FMT affect this *E. coli* variant?

In **chapter 3**, we studied the effect of FMT on antibiotic resistance genes within the gut microbiota. Antibiotic resistance, particularly in hospital settings, is viewed as one of the most significant public health concerns as it jeopardises the ability to treat bacterial infections[55,56]. This not only increases the morbidity and mortality of bacterial infections, but some surgical procedures may also not be feasible anymore. Alternatives to antibiotics are being explored, and FMT has been proposed to control the spread of antibiotic resistance, and possibly eradicate resistant bacteria. In this chapter we

increased the cohort size by including more patient samples compared to **chapter 2**. We also collected long-term follow-up samples after FMT and used both traditional bacterial culture techniques as well as whole-genome and metagenomic sequencing to study the effect of FMT on antibiotic resistance genes. We identified multidrug resistant (MDR) bacteria from cultures. This was compared to metagenomic sequencing data to estimate the relative abundance of these MDR bacteria, while also screening the metagenome for antibiotic resistance genes. In doing so, we provide answers to the questions: 1) does FMT eradicate MDR bacteria? 2) How does FMT affect antibiotic resistance genes in the gut microbiota?

Part 2: Microbiota alterations after faecal microbiota transplantation for ulcerative colitis

In **chapters 4, 5 and 6** we shift our focus to ulcerative colitis (UC, a type of IBD). It has been found that microbial components contribute to this multifactorial disease in addition to disbalances in the immune system and lifestyle factors such as diet. By altering the microbiota it may be possible to induce remission and temporarily control the disease. The NDFB conducted a small randomised clinical trial with two rationally selected donors to assess the colonisation of donor species in patients while also evaluating safety and feasibility of FMT. **Chapter 4** describes the FECBUD clinical trial, and the colonisation of donor-derived bacteria after FMT, known as engraftment, following pretreatment with the anti-inflammatory drug budesonide. The main question here is: does pretreatment with budesonide increase engraftment after FMT? **Chapter 5** extends the FECBUD analysis, applying ecological models to understand bacterial community processes that correlate with clinical remission of IBD symptoms. This chapter addresses the question: what microbial changes in UC patients after FMT correlate with clinical remission or treatment success? **Chapter 6** present a secondary continuation of the analysis in which principles of ecological population dynamics combined with statistical modelling are applied to the FMT metagenomics data. Here, we sought to answer 1) how do dynamics in donor and patient microbial species' abundance correlate with FMT success?, and 2) how can we recognise treatment success early using microbial ecology parameters?

Part 3: Global distribution and genome biology of gut bacterium *Ruminococcus gnavus*

The final part, **chapter 7**, zooms in on one species of bacteria implicated in IBD in several studies: *Ruminococcus gnavus*. This bacterium has been reported to be a common gut bacterium, present in up to 90% of healthy people at low relative abundances. During flares of Crohn's disease (CD), *R. gnavus* has been seen to bloom.

Moreover, several molecular mechanisms have been identified that may play a role in the interaction between *R. gnavus* and the human immune system and contribute to inflammation. We conducted a global meta-analysis to assess the prevalence and abundance of *R. gnavus* in metagenomes, and collected isolates from healthy people and Crohn's disease patients to compare their genetic make-up. Hereby, our aim was to elucidate genomic differences between isolates from Crohn's patients and healthy people, building foundational knowledge of why *R. gnavus* behaves like an innocent commensal in one person and like a pathogen in others.

During this study, we unexpectedly generated a number of whole-genome sequences of *Streptococcus* bacteria. These sequences emerged from presumed pure cultures of *R. gnavus*. Although we cannot be sure about the exact source, we decided to briefly study and summarise these genomes as they might be a valuable resource to future research. These genomes have been deposited in a public repository and are described in **chapter 7.2**.

Chapter 8 concludes the thesis with a general discussion, summarising results from all chapters and placing them in both scientific and societal context. It also proposes ideas for future research directions.

References

1. Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biology* **14**, e1002533 (2016). https://doi.org/10.1371/journal.pbio.1002533
2. Gilbert, J. A. *et al.* Current understanding of the human microbiome. *Nature Medicine* **24**, 392-400 (2018). https://doi.org/10.1038/nm.4517
3. Human Microbiome Project, C. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207-214 (2012). https://doi.org/10.1038/nature11234
4. Costello, E. K. *et al.* Bacterial community variation in human body habitats across space and time. *Science* **326**, 1694-1697 (2009). https://doi.org/10.1126/science.1177486
5. Marchesi, J. & Shanahan, F. The normal intestinal microbiota. *Current Opinion in Infectious Diseases* **20**, 508-513 (2007). https://doi.org/10.1097/QCO.0b013e3282a56a99
6. Gomaa, E. Z. Human gut microbiota/microbiome in health and diseases: a review. *Antonie Van Leeuwenhoek* **113**, 2019-2040 (2020). https://doi.org/10.1007/s10482-020-01474-7
7. Cocquyt, T. Positioning Van Leeuwenhoek's microscopes in 17th-century microscopic practice. *FEMS Microbiology Letters* **369** (2022). https://doi.org/10.1093/femsle/fnac031
8. Dobell, C. *Antony van Leeuwenhoek and his "little animals"*. (Harcourt, Brace and company, 1932).
9. Shendure, J. *et al.* DNA sequencing at 40: past, present and future. *Nature* **550**, 345-353 (2017). https://doi.org/10.1038/nature24286
10. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences* **74**, 5463-5467 (1977). https://doi.org/10.1073/pnas.74.12.5463
11. Maxam, A. M. & Gilbert, W. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences* **74**, 560-564 (1977). https://doi.org/10.1073/pnas.74.2.560
12. Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F. & Petersen, G. B. Nucleotide sequence of bacteriophage lambda DNA. *Journal of Molecular Biology* **162**, 729-773 (1982). https://doi.org/10.1016/0022-2836(82)90546-0
13. Prober, J. M. *et al.* A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science* **238**, 336-341 (1987). https://doi.org/10.1126/science.2443975
14. Fleischmann, R. D. *et al.* Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science* **269**, 496-512 (1995). https://doi.org/10.1126/science.7542800
15. Goffeau, A. *et al.* Life with 6000 genes. *Science* **274**, 546, 563-547 (1996). https://doi.org/10.1126/science.274.5287.546
16. Consortium, C. e. S. Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* **282**, 2012-2018 (1998). https://doi.org/10.1126/science.282.5396.2012
17. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921 (2001). https://doi.org/10.1038/35057062

18. International Human Genome Sequencing, C. Finishing the euchromatic sequence of the human genome. *Nature* **431**, 931-945 (2004). https://doi.org/10.1038/nature03001

19. Venter, J. C. *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**, 66-74 (2004). https://doi.org/10.1126/science.1093857

20. Team, N. I. H. H. M. P. A. A review of 10 years of human microbiome research activities at the US National Institutes of Health, Fiscal Years 2007-2016. *Microbiome* **7**, 31 (2019). https://doi.org/10.1186/s40168-019-0620-y

21. Berg, G. *et al.* Microbiome definition re-visited: old concepts and new challenges. *Microbiome* **8**, 103 (2020). https://doi.org/10.1186/s40168-020-00875-0

22. Berger, B. *et al.* What are the keys to succeeding as a computational biologist in today's research climate? *Cell Systems* **13**, 781-785 (2022). https://doi.org/10.1016/j.cels.2022.09.005

23. Gauthier, J., Vincent, A. T., Charette, S. J. & Derome, N. A brief history of bioinformatics. *Brief Bioinform* **20**, 1981-1996 (2019). https://doi.org/10.1093/bib/bby063

24. Loman, N. & Watson, M. So you want to be a computational biologist? *Nature Biotechnology* **31**, 996-998 (2013). https://doi.org/10.1038/nbt.2740

25. Gevers, D. *et al.* The Human Microbiome Project: a community resource for the healthy human microbiome. *PLoS Biology* **10**, e1001377 (2012). https://doi.org/10.1371/journal.pbio.1001377

26. Integrative, H. M. P. R. N. C. The Integrative Human Microbiome Project. *Nature* **569**, 641-648 (2019). https://doi.org/10.1038/s41586-019-1238-8

27. Frank, D. N. *et al.* Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proceedings of the National Academy of Sciences* **104**, 13780-13785 (2007). https://doi.org/doi:10.1073/pnas.0706625104

28. Gevers, D. *et al.* The Treatment-Naive Microbiome in New-Onset Crohn's Disease. *Cell Host & Microbe* **15**, 382-392 (2014). https://doi.org/https://doi.org/10.1016/j.chom.2014.02.005

29. Kostic, Aleksandar D. *et al.* Fusobacterium nucleatum Potentiates Intestinal Tumorigenesis and Modulates the Tumor-Immune Microenvironment. *Cell Host & Microbe* **14**, 207-215 (2013). https://doi.org/https://doi.org/10.1016/j.chom.2013.07.007

30. Scheperjans, F. *et al.* Gut microbiota are related to Parkinson's disease and clinical phenotype. *Movement Disorders* **30**, 350-358 (2015). https://doi.org/https://doi.org/10.1002/mds.26069

31. Ley, R. E., Turnbaugh, P. J., Klein, S. & Gordon, J. I. Human gut microbes associated with obesity. *Nature* **444**, 1022-1023 (2006). https://doi.org/10.1038/4441022a

32. Turnbaugh, P. J. *et al.* An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027-1031 (2006). https://doi.org/10.1038/nature05414

33. Gilbert, J. A. *et al.* Microbiome-wide association studies link dynamic microbial consortia to disease. *Nature* **535**, 94-103 (2016). https://doi.org/10.1038/nature18850

34. Qin, J. *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55-60 (2012). https://doi.org/10.1038/nature11450

35. Forslund, K. *et al.* Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* **528**, 262-266 (2015). https://doi.org/10.1038/nature15766

36. Castellarin, M. *et al.* Fusobacterium nucleatum infection is prevalent in human colorectal carcinoma. *Genome Research* **22**, 299-306 (2012). https://doi.org/10.1101/gr.126516.111

37. Kostic, A. D. *et al.* Genomic analysis identifies association of Fusobacterium with colorectal carcinoma. *Genome Research* **22**, 292-298 (2012). https://doi.org/10.1101/gr.126573.111

38. Vujkovic-Cvijin, I. *et al.* Host variables confound gut microbiota studies of human disease. *Nature* **587**, 448-454 (2020). https://doi.org/10.1038/s41586-020-2881-9

39. Goris, J. *et al.* DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology* **57**, 81-91 (2007). https://doi.org/10.1099/ijs.0.64483-0

40. Smits, W. K., Lyras, D., Lacy, D. B., Wilcox, M. H. & Kuijper, E. J. Clostridium difficile infection. *Nature Reviews Disease Primers* **2**, 16020 (2016). https://doi.org/10.1038/nrdp.2016.20

41. Weiss, G. A. & Hennet, T. Mechanisms and consequences of intestinal dysbiosis. *Cellular and Molecular Life Sciences* **74**, 2959-2977 (2017). https://doi.org/10.1007/s00018-017-2509-x

42. Duan, H. *et al.* Antibiotic-induced gut dysbiosis and barrier disruption and the potential protective strategies. *Critical Reviews in Food Science and Nutrition* **62**, 1427-1452 (2022). https://doi.org/10.1080/10408398.2020.1843396

43. Lee, J.-Y., Tsolis, R. M. & Bäumler, A. J. The microbiome and gut homeostasis. *Science* **377**, eabp9960 (2022). https://doi.org/doi:10.1126/science.abp9960

44. Abraham, C. & Cho, J. H. Inflammatory bowel disease. *New England Journal of Medicine* **361**, 2066-2078 (2009). https://doi.org/10.1056/NEJMra0804647

45. Khor, B., Gardet, A. & Xavier, R. J. Genetics and pathogenesis of inflammatory bowel disease. *Nature* **474**, 307-317 (2011). https://doi.org/10.1038/nature10209

46. Stripling, J. & Rodriguez, M. Current Evidence in Delivery and Therapeutic Uses of Fecal Microbiota Transplantation in Human Diseases—Clostridium difficile Disease and Beyond. *The American Journal of the Medical Sciences* **356**, 424-432 (2018). https://doi.org/https://doi.org/10.1016/j.amjms.2018.08.010

47. Koopman, N., van Leeuwen, P., Brul, S. & Seppen, J. History of fecal transplantation; camel feces contains limited amounts of Bacillus subtilis spores and likely has no traditional role in the treatment of dysentery. *PLoS ONE* **17**, e0272607 (2022). https://doi.org/10.1371/journal.pone.0272607

48. van Nood, E. *et al.* Duodenal infusion of donor feces for recurrent Clostridium difficile. *New England Journal of Medicine* **368**, 407-415 (2013). https://doi.org/10.1056/NEJMoa1205037

49. Porcari, S. *et al.* Key determinants of success in fecal microbiota transplantation: From microbiome to clinic. *Cell Host & Microbe* **31**, 712-733 (2023). https://doi.org/https://doi.org/10.1016/j.chom.2023.03.020

50. Feuerstadt, P. *et al.* SER-109, an Oral Microbiome Therapy for Recurrent *Clostridioides difficile* Infection. *New England Journal of Medicine* **386**, 220-229 (2022). https://doi.org/doi:10.1056/NEJMoa2106516

51. Sims, M. D. *et al.* Safety and Tolerability of SER-109 as an Investigational Microbiome Therapeutic in Adults With Recurrent Clostridioides difficile Infection: A Phase 3, Open-Label, Single-Arm Trial. *JAMA Network Open* **6**, e2255758-e2255758 (2023). https://doi.org/10.1001/jamanetworkopen.2022.55758

52. Dsouza, M. *et al.* Colonization of the live biotherapeutic product VE303 and modulation of the microbiota and metabolites in healthy volunteers. *Cell Host & Microbe* **30**, 583-598. e588 (2022). https://doi.org/10.1016/j.chom.2022.03.016

53. Khanna, S. *et al.* Efficacy and Safety of RBX2660 in PUNCH CD3, a Phase III, Randomized, Double-Blind, Placebo-Controlled Trial with a Bayesian Primary Analysis for the Prevention of Recurrent Clostridioides difficile Infection. *Drugs* **82**, 1527-1538 (2022). https://doi.org/10.1007/s40265-022-01797-x

54. Feuerstadt, P. *et al.* Retrospective Analysis of the Safety and Efficacy of Fecal Microbiota, Live-jslm (REBYOTATM) Administered Under Enforcement Discretion to Patients With Clostridioides difficile Infection. *Open Forum Infectious Diseases* **10** (2023). https://doi.org/10.1093/ofid/ofad171

55. Antimicrobial Resistance, C. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet* **399**, 629-655 (2022). https://doi.org/10.1016/S0140-6736(21)02724-0

56. Tillotson, G. S. & Zinner, S. H. Burden of antimicrobial resistance in an era of decreasing susceptibility. *Expert Review of Anti-infective Therapy* **15**, 663-676 (2017). https://doi.org/10.1080/14787210.2017.1337508