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

Nooij, S.


サム

Chapter 8

General Discussion

Summary of main findings

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

Clostridioides difficile is a Gram-positive bacterial pathogen that causes gastrointestinal infections and forms spores that are resistant to antibiotic treatment. Besides, *C. difficile* thrives in an antibiotic-depleted gut microbial environment, leading to a high incidence of recurrent infections. Recurrence rates after first infection are 15-30%, and 40-65% after one or more recurrences[1]. Faecal microbiota transplantation (FMT) has proven to be an effective method to restore the gut microbiota and prevent further recurrences of *C. difficile* infections (CDI). Furthermore, FMT serves as an interesting model for studying the transfer of human gut microbiota. This first part of the thesis leverages FMT data from multiple recurrent CDI patients to study effects on other potentially harmful bacteria.

Chapter 2 focuses on a recently discovered microbial risk factor associated with the development of colorectal cancer: colibactin-producing (*pks*$^+$) *Escherichia coli*. We screened stool metagenomic data from FMT recipients and their respective healthy donors to assess the presence of *E. coli* and the *pks* operon. The objective was to determine both the presence and relative abundance of *E. coli* in general, and the subpopulation capable of producing colibactin. Thereby we also assessed the effect of FMT on this putative carcinogenic bacterium. We found that *pks*$^+$ *E. coli* was present in donors and FMT-treated patients, and decreased in patients after FMT, which was correlated with the absence of *pks*$^+$ *E. coli* in the donor. We found no evidence supporting the transmission of *pks*$^+$ *E. coli* from donor to patient. Consequently, we conclude that FMT affects putatively carcinogenic *E. coli* by reducing their presence and abundance, particularly when the donor is free of the *pks*$^+$ *E. coli* variant. Our results suggest that FMT is more likely to mitigate the risk of colonisation by carcinogenic bacteria rather than elevate the risk through bacterial transfer. This proof of concept may inspire the development of microbiota-based therapies aiming to prevent the development of colorectal cancer.

In **chapter 3**, we studied more FMT sample triads (comprising donor, patient before FMT, and patient after FMT) along with long-term follow-up samples. Through a combination of traditional bacterial culture, whole-genome sequencing (WGS) and metagenomic sequencing, we extensively characterised antibiotic resistance in this cohort and evaluated the effect that FMT exerted on antibiotic-resistant bacteria. Our results indicated that FMT reduced the prevalence of multidrug-resistant (MDR) bacteria in patients. WGS combined with metagenomics suggested that MDR bacteria may persist post-FMT, although their relative abundances were significantly reduced, making them harder to detect. Further analysis of the metagenomic data showed that the relative abundance of antibiotic resistance genes (ARG) decreased after FMT,

while the richness of these genes remained unchanged. Computational predictions identified which ARG-containing contigs were plasmid-derived and we found their quantity was not significantly affected by FMT. In summary, we hypothesise that FMT may effectively reduce MDR bacteria and ARG abundance, potentially decreasing the risk of infection. However, it appears that FMT may not completely eradicate antibiotic-resistant bacteria from the gut microbiome. Thus, FMT could serve as an additional strategy for combatting the spread of antibiotic resistance within a patient population that is at increased risk of bacterial infections.

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

Ulcerative colitis (UC) is a form of inflammatory bowel disease characterised by chronic inflammation of the colon, which is associated with alterations in gut microbial composition and function. A randomised clinical trial was conducted to assess the safety of FMT in these patients, and data were collected to evaluate bacterial colonisation of the recipients' gastrointestinal tracts, both with and without anti-inflammatory pretreatment. **Chapters 4-6** describe analyses of this dataset and the changes in gut microbiota following repeated FMT.

Chapter 4 describes the clinical trial and bacterial engraftment using a combination of microbiota diversity metrics. We hypothesised that anti-inflammatory pretreatment would facilitate colonisation of foreign bacteria. However, our findings indicated that the donor had a more pronounced effect on engraftment. Therefore, we conclude that donor selection and an improved understanding of how to identify optimal donors may be crucial for the success of FMT in UC.

Chapter 5 builds upon the findings of **chapter 4** by delving deeper into microbial ecology. The main objective was to identify changes in the microbial communities correlated with clinical remission. Using a computational modelling approach, we defined clusters of bacterial profiles. Among these clusters, we found one associated with poorer clinical outcomes, specifically a failure to achieve remission. Conversely, a cluster characterised by high relative abundances of *Ruminococcaceae* and *Lachnospiraceae* was associated with treatment success. Using these broad microbiota characteristics, we hypothesised that it may be possible to predict treatment success of FMT early after the procedure.

Chapter 6 continues the investigation from **chapter 4 and 5**, with a combination of engraftment analyses with microbiota dynamics of the bacteria present in the patients before FMT. This adapted methodology addresses a slightly different question: how do the dynamics of donor and patient species correlate with FMT success? We observed that patient species present at high relative abundances often persisted after FMT. In terms of donor-derived species, moderate and stable colonisation of donor species

was correlated with treatment success, while initial high engraftment followed by loss of donor species was indicative of failure to achieve or maintain remission. Taken together, we predict that FMT treatment success in UC depends on resilience of the recipient's resident microbiota, combined with the ability to stably incorporate a moderate amount of healthy donor-derived microbiota.

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

Chapter 7 revolves around the gut bacterium *Ruminococcus gnavus*, which is strongly associated with inflammatory bowel disease, specifically Crohn's disease. At the same time, it has been described to be present in around 90% of healthy adults. Our objective was to re-evaluate these correlations and elucidate the genomic basis of the contrasting host-microbe interactions. We discovered that *R. gnavus* is more prevalent and abundant in conditions such as IBD, type-2 diabetes, hypertension and atherosclerotic cardiovascular disease compared to healthy people. This is also true for Westernised compared to non-Westernised societies and more in infants and young children compared to adults. Based on a large collection of complete and draft genomes, we found that bacterial motility may be overrepresented in infant-derived strains, and isolates derived from healthy people are phylogenetically and functionally different from isolates from Crohn's disease patients. We conclude that there may be distinct subspecies of *R. gnavus* that co-evolved and adapted to environments or lifestyles. Our work shows that within a species there may be several genomically and phenotypically distinct variants. Therefore, more nuance is warranted when attributing disease to bacterial species or describing host-microbe relations in general.

Finally, **chapter 7.2** is a spin-off of **chapter 7.1,** in which we describe complete genomes derived from presumed laboratory contaminants. Our analysis revealed that the sequence data from supposed *R. gnavus* isolates contained DNA from different species, prompting us to computationally separate the sequences belonging to these various species. This resulted in the identification of seventeen *Streptococcus* genomes, one genome from *Bacteroides fragilis* and one *Staphylococcus capitis*. The latter two genomes have not been described in the publication; their genomes have also been deposited in the European Nucleotide Archive public repository. Although the description of these bacteria does not fit the scope of **chapter 7**, we recognise that their genome information could be valuable to the scientific community. As strong supporters of open science and the FAIR data principles, we have made these genomes publicly available for further research.

A broader view of FMT and gut microbiota

The further discussion covers the following topics, which extend from the work presented here to the broader developments in the scientific field:

Unpredictability of FMT: for better or worse

FMT is a promising therapy for various diseases in which the gastrointestinal microbiota plays a role. For example, it can be used to 1) prevent recurrence of *C. difficile* infections[2,3], 2) induce remission in IBD[4,5], and 3) help relieve symptoms of immune checkpoint inhibitor-induced colitis[6,7] and gastrointestinal acute graft-versus-host disease[8-10]. However, as with any therapeutic intervention there is a risk of side-effects[11,12]. The most reported adverse events relate to the procedure itself and the diseases being treated, including increased stool frequency, abdominal pain, nausea, and diarrhoea[13]. The most severe described adverse event is the death of an immunocompromised FMT recipient caused by antibiotic-resistant bacteraemia[14]. The bacterial pathogen was obtained from the donor, who had not been screened for antibiotic-resistant bacteria. Additionally, there have been reports of self-administered FMT from family members providing relief from gastrointestinal symptoms but also resulting in the emergence of a seemingly unrelated condition of acne[15]. Conversely, previous studies, including **chapters 2 and 3** of this thesis, have shown that FMT can also yield beneficial side-effects[16-22]. Next to resolving CDI recurrences, FMT can modulate the gut microbiota by reducing antibiotic-resistant bacteria and putative carcinogenic bacteria. This underscores the therapeutic potential of FMT, while also

stressing the need for rigorous donor screening, and safe, standardised administration methods[23,24]. Furthermore, the effects of FMT are usually studied on the timescale of weeks, while some effect may only arise after years' time. Therefore, the use of registries for both FMT donors and recipients would be useful for monitoring and evaluating these potential long-term effects. FMT is a versatile treatment for various human microbiome-related health issues that merits further investigation. It should only be administered in a professional and controlled setting, with careful monitoring of outcomes.

FMT for IBD: inducing and maintaining remission remains challenging

FMT has been repeatedly tested for its potential to treat IBD, particularly ulcerative colitis, with variable outcomes[5,25]. FMT is capable of inducing remission in UC patients, but success rates are significantly lower compared to those for multiple recurrent CDI, with a mean clinical remission rate of 42% for IBD[25], compared to a 92% clinical resolution rate for CDI[26]. To compare, placebo treatment achieved a 22.6% mean remission in the IBD trials[25]. Next to a different aetiology, the difference in treatment success may partly be explained by the antibiotic pretreatment used for FMT in multiple recurrent CDI (rCDI). The rationale of antibiotic pretreatment is that reducing the patients' existing gut microbiota may facilitate stable introduction of the healthy donor microbiota[27]. While antibiotics are standard therapy for rCDI, they are not typically employed in the treatment of IBD. This is but one of the main factors complicating one-to-one comparisons between these two conditions and FMT.

It has been hypothesised that a gut microbiota that is more similar to the donors', achieved by high bacterial engraftment, should lead to a healthier and less inflamed gastrointestinal tract in IBD patients. To stimulate engraftment, FMT for IBD might benefit from similar pretreatment as rCDI, as successfully demonstrated in previous studies[28]. Contrastingly, **chapter 4** illustrates that anti-inflammatory pretreatment, which is commonly used for symptom management in IBD, does not enhance colonisation of donor-derived bacteria. An alternative strategy to increase engraftment is to optimally match the donor and recipient microbiota, as suggested in **chapter 6** and other studies[29,30]. Matching may be done based on gut microbiota enterotypes or microbiota distance between donor and recipient[30], or based on recipient microbiota composition and dysbiosis[29]. The methodologies to calculate these parameters are not standardised, may be dataset-specific, and depend on code and instructions provided by the authors. Using microbiota parameters for matching donors to patients requires more thorough preparation to gain detailed knowledge of both donor and patient microbiota, thereby increasing the resources and expertise needed to administer FMT to IBD patients effectively. Nonetheless, how to achieve maximum bacterial transfer and colonisation, and whether that is beneficial to the recipient's health, remains to be established.

Furthermore, IBD is a chronic disease that is unlikely to be cured by a single intervention. The disease itself and the proposed treatment mechanisms are entirely different from CDI and no immunosuppressive therapy achieves response rates as high as those for FMT in rCDI. FMT may be used to reach a state of remission, but to maintain remission repeated FMTs[28], or alternative microbiota modulating therapies such as prebiotics are probably necessary[31]. On the other hand, if IBD is conceptualised as a chronic overreaction of the immune system, like an allergy, a possible treatment approach may also be inspired by modern allergy immunotherapies. Several approaches have been proposed that target the intestinal immune system and trigger anti-inflammatory pathways using small molecules or biological agents[32]. Combining the immune system-targeting approach with techniques from vaccine development[33], one could envision re-training of the immune system by administering microbial surface molecules which may be engineered to manipulate the immune response. Or looking further, it might be possible to design an mRNA vaccine that helps protect against pro-inflammatory bacteria, similar to what has been demonstrated for *C. difficile*[34]. These lines of therapeutic research may complement one another and necessitate a deep understanding of both the gut microbiota and immune system. Which therapy or combination of therapies is most efficacious is probably patient-dependent, with a variety of host-microbe characteristics being relevant. More trials and multidisciplinary research into the interactions between the immune system and gut microbiota are needed to assess effectiveness of the different treatment options for IBD.

Finding the balance

The goal of FMT is to restore a functional gut microbiome, a system in which interactions between human and microbiota are in balance: a state of homeostasis[35-37]. It is thought that finding a proper match between donor and patient microbiota is necessary to re-establish a balanced gut microbiota in the recipient patient. To find this match, researchers have experimented using machine learning models to predict optimal donor-recipient combinations based on microbiota composition, pretreatment and clinical outcome measures[27,29,30,38]. By using such high-dimensional and rich data, solutions may be found that go beyond our current understanding of gut microbiome biology.

The optimal complementary donor microbiota for a patient may depend on the disease and other factors. In the case of rCDI, the main goal is preventing outgrowth and toxin production of *C. difficile*. Other diseases may likewise be treated by inhibiting specific deleterious species, or by shifting the microbiome's collective metabolism, such as with the drug levodopa in Parkinson's patients[39-41]. As for other factors that influence engraftment and thereby possibly treatment success, the microbiota themselves need to adapt to their new environment: the recipient's colon. The colons of donor and recipient will be different in various aspects relevant to the microbiome, including their physiology and immunology[42-44]: they are different ecosystems. Furthermore,

the compatibility rate of microbial communities – that is, the extent to which two microbiotas can co-exist – may also be influenced by lifestyle factors such as diet or physical exercise. These factors have not yet been extensively studied in connection with FMT, because the main disease for which it is used, rCDI, has had high success rates irrespective of the exact microbiota composition[26]. In fact, it is striking that although it is known that diet and exercise impact the gut microbiota[45-47], there appear to be few guidelines to provide recipients with advice or support as to what to eat and do after FMT. Would it be beneficial for recipients if they adopt their donor's lifestyle? For how long, and what aspects matter most? A 2020 survey among mostly gastroenterologists indicated that most FMT experts find it important to consider the diet of both donors and recipients[48]. Moreover, there have been studies on the combined effect of diet and FMT[49-51], and the Australian Centre for Digestive Diseases has published dietary requirements for stool donors prior to donation[52]. Taken together, there appear to be opportunities to include more disciplines into FMT research and thereby further optimise the treatment. However, added complexity should not impose barriers on stool banks, donors, treating physicians and patients.

Monitoring of long-term FMT outcomes

FMT is not a frequently applied therapy, and its variable composition necessitates meticulous monitoring of outcomes. This monitoring is primarily conducted by the treating physician and recorded by the stool banks and registries of the FMT providing centre. Stool samples may be collected in the weeks following FMT to evaluate microbial composition. If the FMT indication pertains a non-chronic illness and the patient reports no further complaints, follow-up with the physician typically ends. Consequently, should new complaints arise, and the patient turns to another physician, a potential link to the FMT may be overlooked. This complicates the long-term evaluation of FMT and identification of potential deleterious side-effects, such as carcinogenesis or infection with MDR bacteria. To accommodate longer-term monitoring of FMT, initiatives for maintaining registries are developed at both national[53,54], and international[55] levels. These registries document the relevant patient history leading to FMT and any subsequent events that can be linked to the FMT, i.e., the introduction of donor-derived bacteria that cause harm over an extended time period. Examples include colorectal carcinogenesis and multidrug-resistant bacterial infections as discussed in **chapters 2 and 3**. As more patients are recorded in the registries, the likelihood of identifying rare occurrences increases, helping FMT researchers enhance FMT safety. Thus, these registries serve as an invaluable tool for advancing our understanding of the long-term effects of FMT.

Controlling FMT quality

Quality control for FMT has two primary aims: 1) to prevent deleterious side-effects (risk management), and 2) achieve the best possible results (efficacy). Both demand a thorough understanding of the faecal microbiota, which can be estimated using

different methods. Donor screening is mostly used to mitigate risks associated with FMT, but can also help inform about the potential efficacy. The NDFB employs a rigorous donor screening protocol, utilising questionnaires addressing risk factors for transmissible diseases and conditions associated with a perturbed microbiome. This primary screening excludes approximately 70% of volunteers that applied to become donors[13,23]. Exclusion criteria include gastrointestinal complaints, recent antibiotic use, travel to countries with endemic gastrointestinal pathogens or MDR bacteria, and a family history of IBD[23]. To predict treatment efficacy, microbial profiling may be used. A Chinese research group has published guidelines that incorporate 16S rRNA gene sequencing for microbial profiling and a machine learning model that uses a random forest approach to match patients with donors based on predefined microbial clusters, so-called enterotypes[30]. Although these methods effectively maintain high success rates for rCDI with minimal side-effects, they can only prevent transfer of known harmful microorganisms. Microbial profiling before FMT may not be necessary for the effectiveness against rCDI, as success rates have been consistently high (~90%) regardless of bacterial composition[26]. It is not technically and financially feasible to determine the exact composition of donor faeces and classify all bacteria, fungi, viruses, metabolites, and food particles. Besides, if the specific microorganisms and metabolites responsible for the therapeutic effects were fully understood, it would be more practical to isolate them and compose a conventional medicine administered as a capsule. Such a drug is referred to as a live biotherapeutic product (LBP; figure 1) and currently under development by several companies[56,57]. LBPs are discussed in more detail below. To conclude, there are several established ways of controlling FMT quality and when knowledge regarding the mechanisms of action accumulates for all the target diseases, FMT will likely be superseded by human-designed drugs.

Next generation of microbiota therapies

The future of microbiota therapy is anticipated to be centred around LBPs (Figure 1). LBPs consist of live microbial organisms selected or developed to treat or prevent human diseases[56]. The primary advantage of LBPs is that their composition is known and stable, unlike donor stool, which facilitates regulation as a drug and should provide more robust and predictable outcomes. This should also make LBPs safer than FMT. However, LBPs as treatment require a deeper understanding of the human gut microbiome obtained through research, and requires significant investments in trials for regulatory approval. Therefore, LBPs are initially quite expensive, posing challenges in competing with FMT. Nevertheless, given their advantages LBPs are likely to eventually replace FMT for routine treatments in the long term.

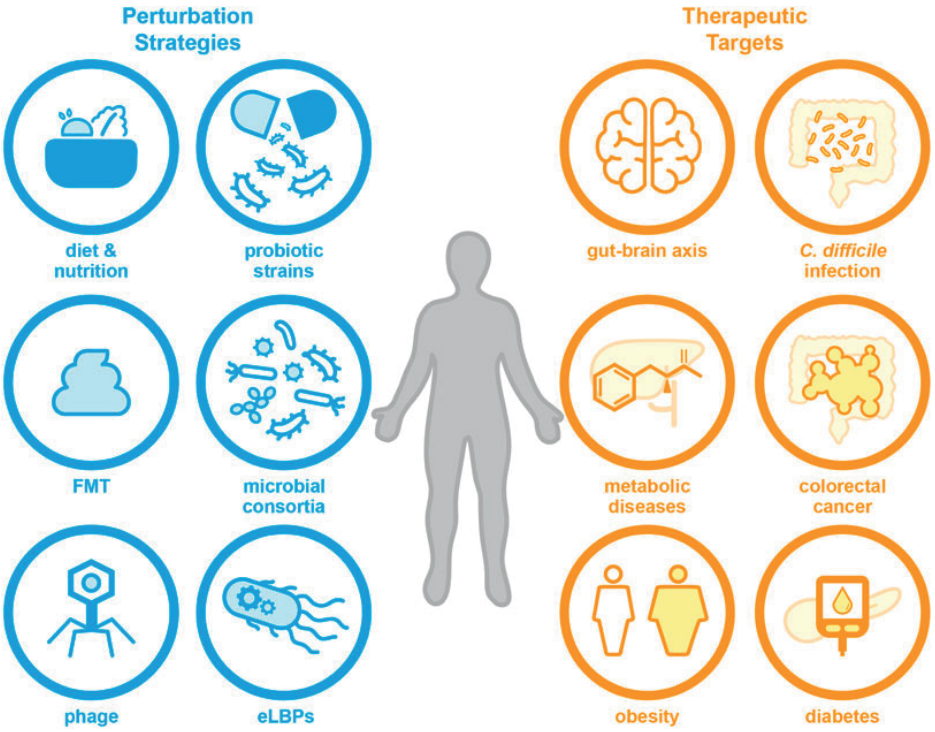

Figure 1. Techniques and aims of microbiome therapies. On the left side, different intervention strategies are listed in blue. Probiotic bacterial strains, microbial consortia and engineered bacteria are examples of live biotherapeutic products. The right-side lists in orange targets for which microbiota therapies are developed or have been tested successfully. FMT: faecal microbiota transplantation, eLBP: engineered live biotherapeutic product. Figure from reference[58].

Companies have developed LBPs using various methods and with different aims of resolving gastrointestinal complaints. For the treatment of rCDI there are at least five products from three different companies: RBX2660[59], RBX7455[60], VOWST (formerly SER-109)[61], VE303[62,63], MET-2[64], and NTCD-M3[65]. These can be broadly categorised in three groups. The first two (RBX2660 and RBX7455 from Ferring Pharmaceuticals) are derived directly from donor stool and are thus practically identical to 'traditional' FMT. RBX2660 is a suspension for rectal administration and RBX7455 is a freeze-dried (lyophilised) product that is administered as oral capsule[59,60]. Next, MET-2 (from NuBiyota), VE303 (from Vedanta Biosciences) and VOWST (from Seres Therapeutics) are encapsulated consortia of well-characterised bacteria. RBX2660 and VOWST have passed phase 3 clinical trials and have been approved for use against rCDI by the U.S. Food & Drug Administration (FDA[66,67]). RBX7455 and MET-2 have successfully completed a phase 1 trial and VE303 is now in a phase 3 trial. NTCD-M3 (from Destiny Pharma) is a non-toxigenic strain of *C. difficile*, which has been administered through oral liquid formulation in a phase 2 trial. Furthermore, LBPs for treating IBD are under development (GUT-103 and GUT-108[68], and VE202[69-71]). These also consist of well-

defined bacterial consortia. VE202 (From Vedanta Biosciences) is now at stage 2 clinical trials, while GUT-103 and GUT-108 (from Gusto Global) have yet to be tested in a phase 1 trial. Finally, LBPs have been described that consist of a single bacterial strain, which are often genetically engineered to have a specific effect in the gut microbiome[56]. This includes NTCD-M3[65], *E. coli* CWG308[72], *E. coli* Nissle 1917[73-75], and ADS024[76]. As an example application, a strain of *E. coli* Nissle 1917 engineered to overproduce microcin I47 has been found to be effective against multidrug-resistant bacteria[56,73]. It is important to note that this *E. coli* strain also poses a health risk as discussed in **chapter 2**: *E. coli* Nissle 1917 produces colibactin and may contribute to carcinogenesis[77,78]. This illustrates the need for proper assessment and regulation of probiotics and especially the use of genetically modified organisms. This poses challenges for their development and market release, particularly in Europe, where approval processes are generally more stringent. To summarise, the development of LBPs is costly and time-consuming. Until LBPs are approved for use and readily available, FMT remains a valuable experimental treatment for gut microbiota-related conditions.

Deeper understanding of host-microbiota interactions

When studying the microbiota, it is common to refer to individual members as species. However, as shown in **chapters 2, 3 and 7**, a microbial species is a group of possibly divergent organisms under one umbrella name. Evolution has produced a spectrum of different phenotypes, and pathogenicity is not necessarily linked to species (Figure 2). A comprehensive understanding of the microbiota requires consideration of this variation. Several research groups have studied intraspecies variants from metagenomics data in the context of FMT[79-88]. They have used different bioinformatics approaches to differentiate variants based on genomic features, which are summarised below. One group of methods relies on pre-built databases to identify and quantify variants of known species:

- *StrainPhlAn* relies on a custom database of species-specific markers to identify the dominant strain for each species and compare their consensus sequence between samples[82]. StrainPhlAn has been developed further together with MetaPhlAn and is now at version 4 with an extended database to capture a greater diversity of prokaryotes[89].
- *Strainer* utilises genomes of cultured isolates to create a database of unique k-mers for each strain and then matches metagenomic reads to these k-mers to quantify strains[80].
- *StrainPanDA* uses a pangenome-based approach to reconstruct gene content variation per strain[84].
- *SameStr* uses species-specific marker genes to identify single-nucleotide variants but is flexible with databases and can analyse multiple strains per species[85].
- *SNV-FEAST* introduces a new single-nucleotide variant signature identification method based on read mapping to a custom bacterial reference genome database and provides relative abundances of variants in source and sink samples[87].

- Having a pre-built database facilitates the application of these tools to datasets of well-studied microbiomes. Other tools make use of the metagenomic data or user-defined references to call variants of any gene or genome, including:
- *Strain Finder* uses a machine learning approach to estimate the optimal number of strains from reference genome alignments, and then calculates frequencies of these strains across metagenomes[79].
- *inStrain* relies on paired-end reads mapped to any reference sequence, which can be a metagenome-assembled genome from the same dataset and calls single-nucleotide variants to distinguish strains in metagenomes[81].
- *STRONG* uses *de novo* assembly of metagenomic reads to identify strains based on assembly graphs[83].
- *StrainGE* specialises in low-abundance variants and uses read mapping to reference genomes, which may be supplied separately by the user[86].
- *ChronoStrain* focuses on low-abundant taxa and models variant abundances over time with uncertainty quantification using a read-based approach and a custom database of user-specified reference marker sequences[88].

These tools require additional user input but can handle any microbiome regardless of reference databases. While these tools allow for identification of intraspecies variants across samples, and some enable studying lowly abundant variants, only StrainPanDA explicitly incorporates functional information, which is essential for elucidating host-microbe interactions.

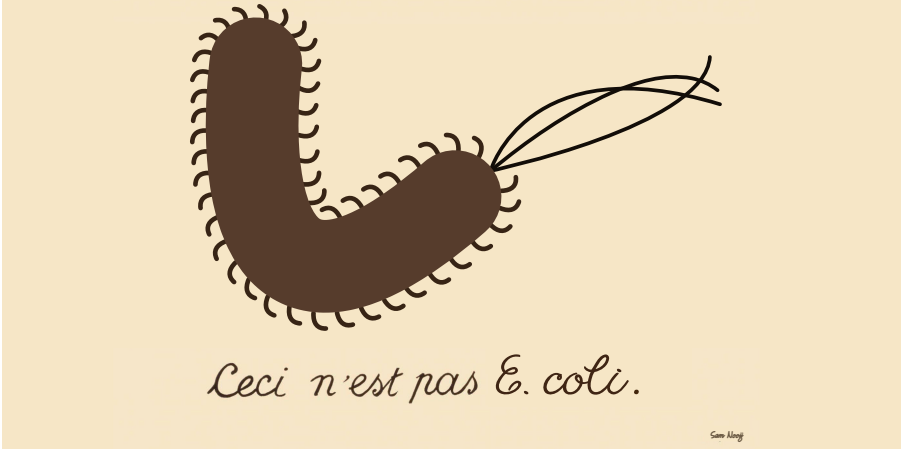

Figure 2. This is not *E. coli*. What appears to be a harmless commensal may actually be a source of infection or harbinger of cancer. A detailed scrutiny of the genome and phenotype is necessary to accurately assess the implications of bacteria and the potential effects on its host. After René Magritte, La Trahison des Images (1929).

The next step in understanding gut bacterial function is to examine the genomes' encoded functionality. This may be partially achieved by using metagenomics

sequencing, which allows for identification of complete genes, representing the functional potential. For example, the recently published tool microSLAM takes into account population structure within metagenomes and the presence of genes to assess associations between species and their genes with diseases[90]. However, to really understand what is going on in the gut microbiome, one should sample not only the DNA but also look at mRNA, proteins, and metabolites. The combination of these techniques in high-throughput fashion is known as 'multi-omics' and discussed further below.

Gut bacteria as double-edged sword

Continuing the discussion of intraspecies variation and differences in pathogenicity; numerous gut bacteria have been characterised as either beneficial or detrimental to the health of their human host (Figure 2). Examples of such ambivalent species include *Clostridioides difficile*[91], *Escherichia coli*[92], *Akkermansia muciniphila*[93], *Prevotella* spp.[94-96], and *Ruminococcus gnavus*[97]. Accurate diagnosis, treatment, and prevention of disorders require methods to distinguish harmful or pathogenic bacteria from commensals. Also, whether or not a bacterium is seen as pathogen depends on host conditions such as immune response. The same bacterium may be harmless when in one person and induce severe symptoms in another, for example with immunocompromised patients, and this may change over time. Therefore, we need to acknowledge that 'pathogen' and 'commensal' describe behavioural traits rather than fixed entities. Proper context is essential to assess these behaviours. For example, *C. difficile* may reside in the gut unnoticed by its host and cause disease upon depletion of the gut microbiota by antibiotics. Or particular variants of *E. coli* may be used as probiotic[98], while other (shiga toxin-producing) strains are severe pathogens[99]. A bacterium's capabilities are primarily dictated by its genetic makeup, which can be determined through genomic analysis. Exceptions such as horizontal gene transfer aside, the genome determines the range of possible behaviours or lifestyles. Therefore, genomics and pangenome analyses are valuable tools for initial screening of potential phenotypes.

Genomic flexibility and horizontal gene transfer

A major driver of bacterial evolution and associated genomic diversity is horizontal gene transfer (HGT). Bacteria may take up DNA from their environment through various mechanisms, which may then be incorporated into their chromosome. Evidence of HGT can be found in most bacterial genomes[100], and over half of a bacterium's genes may be mobilisable[101]. These mobile genes may be clinically relevant and encode antibiotic resistance[102], virulence factors[103], and genotoxins like colibactin in *E. coli*[104,105]. As a result, bacterial strains can rapidly adapt their functions when the opportunity arises. This underscores the importance of detailed characterisation methods, and several bioinformatics approaches have been devised to study HGT[101]. Below, I briefly summarise those relevant to bacterial genomics and metagenomics data.

Extrachromosomal mobile genetic elements such as plasmids can be recovered from metagenomes, although the short-read lengths (100-300 bp) typically used pose significant challenges to their accurate reconstruction. The presence of sequence repeats and sequences shared across different genomes, as well as variable read depths make *de novo* assembly particularly challenging, often resulting in fragmented contigs[101]. Nonetheless, identification of mobile element-derived contigs is possible through specific markers by aligning them to a database or using machine learning approaches[106-109]. This method was also applied in **chapter 3**.

Long-read sequencing, as used in **chapter 7**, facilitates reconstruction of extrachromosomal elements. PacBio's single-molecule, real-time platform, and Oxford Nanopore's platform generate sequence reads of 10,000 bp and longer. This is sufficient to bridge complex insertions and may enable recovery of full plasmids[110]. However, if not combined with culture isolation, linking a complete plasmid to its host organism remains difficult. To associate mobile genetic elements with their host, single-cell sequencing[111,112], and proximity ligation (hi-C) are attractive approaches[102,113-115]. Both can be combined with metagenomics to examine mobile elements in a large number of bacteria. These techniques will likely see increased use in the future, and uncover many more cases of HGT.

Beyond microbial genomics: what do gut bacteria do?

Back to the topic of omics methods. This thesis relies on genomics data, with a focus on bacterial DNA. Genomic methods are ideal for microbial profiling, that is, identifying which bacteria are present, essentially answering the question 'who is there?' Whole-genome sequencing techniques can infer gene presence by employing gene prediction tools and known gene databases, and answer 'what can they do?' To delve deeper into the processes of the microbiome and interactions with the host one needs different methods. These include metatranscriptomics to measure and identify RNA, metaproteomics to measure the enzymes that are produced and metabolomics to measure the products of metabolic activity (Figure 3). This additional information is invaluable for understanding the human gut microbiome to the level of molecular pathways and for designing effective interventions. Combining several of these methods, called multi-omics, is the most powerful approach in decoding the gut microbiota and requires specialised computational tools for integration and interpretation[116,117].

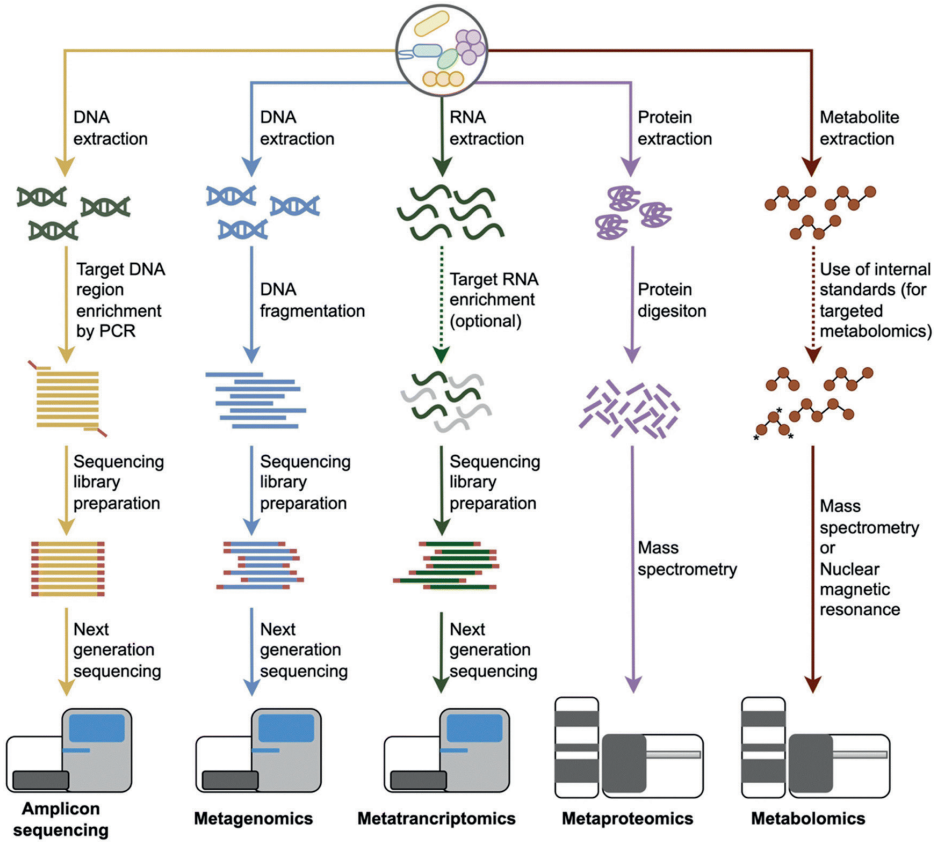

Figure 3. Sample preparation for multi-omics experiments. DNA, RNA, proteins, and metabolites maybe extracted from microbial communities and measured using indicated experimental workflows. Afterwards, data may be integrated using computational methods (not shown). Figure from reference[116].

Obstacles in high-throughput biological science

Modern high-throughput biological omics methods have proven to be very powerful in finding host-microbe interactions and formulating new hypotheses[118-121]. However, they are not without limitations. One major limitation is the inevitability of introducing sampling bias during sample collection, storage, and biochemical processing[122,123]. Despite being theoretically random and relatively unbiased, different experiments have shown that omics methods tend to introduce subtle biases, meaning the resulting data may not be entirely random or fully representative[124]. This also pertains to the identifiable taxa from metagenomic data: while metagenomics does not actively select for specific organisms, sampling kits and processing methods optimised for stool samples tend to predominantly yield bacterial DNA. The taxonomic biases are partly attributable to DNA extraction and library preparation protocols[125-127], and rely on computational databases that tend to reflect well-studied pathogens better than other

microorganisms and may contain biases or errors in the related taxonomy[128]. Therefore, to get a comprehensive view of the microbiome, it is still necessary to undertake multiple experiments and computational analyses. It may be more worthwhile to understand each method's inherent limitations, find the optimal approach for a given research question and report openly on methodological choices[122].

Metagenomics tends to rely on species-rank classifications, which may not be detailed enough to discriminate between disease-associated bacteria and harmless intraspecies variants. Fortunately, commonly used databases and classification methods are constantly improving[89,129], alleviating this problem. Also, the problem can be circumvented by combining the analyses with methods that include gene identification or similar functional inference, like employed in **chapters 2 and 3**.

A significant limitation of metagenomics includes the complexity of generating metagenome-assembled genomes (MAGs) from deeply sequenced, short-read metagenomes. MAGs depend on *de novo* assembly and binning tools to reconstruct as much of an organism's genome as can be predicted based on genetic information and specialised algorithms. However, this approach is liable to underestimate the genome due to 1) missing information because of (quasi-)random sequencing, 2) assembly artifacts or failure to correctly piece together all sequencing reads, and 3) binning errors, where the algorithm may incorrectly predict which assembled contigs belong together[130]. Within a metagenomic analysis, MAGs may be the best possible descriptors of the bacterium from which they derive, but they are typically less complete and more error-prone than genomes from cultured isolates. This we illustrated in **chapter 7**, where we show differences in length, completeness and GC content between MAGs and isolate genomes of *R. gnavus*, even though MAGs were predicted to be of high quality. Therefore, MAGs suffice for some research questions, but complete genomes from isolates are preferred when available.

An important obstacle to the adoption of multi-omics experiments is the high cost[117]. Multi-omics experiments are labour-intensive, require specialised equipment and consumables, and are also data-intensive, necessitating specialised computational expertise. As a result, only larger and well-funded laboratories have the resources to conduct such studies. Besides, depending on the research question, a simpler, more cost-effective experimental approach may be more appropriate. For example, when the possible targets are known detection by PCR may suffice and is faster and cheaper. High-throughput multi-omics experiments have been instrumental to advance our knowledge of the human microbiome, but are not the answer to all questions, nor can they fully replace traditional methods such as bacterial culture.

The proof is in the microbiological pudding

Bacterial culture has been the gold standard to identify and characterise bacteria. Next to facilitating whole-genome sequencing, cultured bacteria allow further experimentation, for example to study interactions with human tissue and the immune system. This enables researchers to validate hypotheses generated using computational experiments, helping to decode the complexities of the human microbiome. Advances in bacterial culturing and the development of high-throughput methods, known as culturomics (Figure 4), have greatly expanded the ability to isolate and study a broader range of previously unculturable gut microbiota[131]. Technologies that have accelerated culturing of many bacteria to become high-throughput include: 1) image recognition software and artificial intelligence (AI) models to identify different colonies on a plate, 2) robot systems that automatically isolate the bacteria, 3) taxonomic identification systems using DNA sequencing or matrix-assisted laser desorption/ionisation-time-of-flight mass spectrometry (MALDI-TOF). Combined with developments of artificial human gut tissue systems[132-136], this provides excellent opportunities for studying host-microbe interactions.

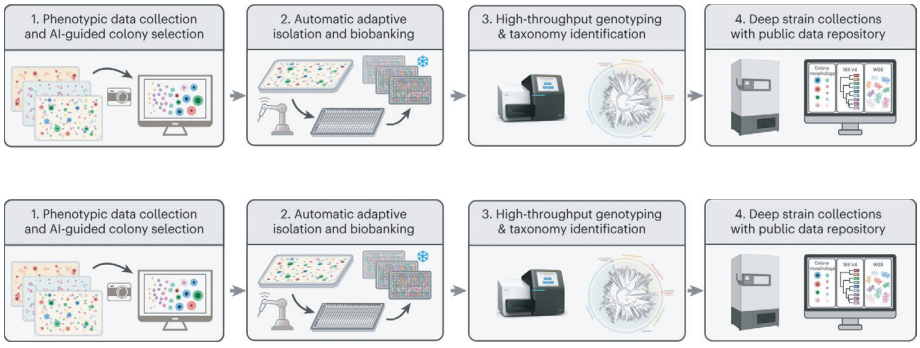

Figure 4. A state-of-the-art culturomics workflow. 1: The process starts by dividing a biological (microbiome) sample and growing the microbiota on different media. Colonies are selected based on phenotype, supported by Artificial Intelligence. 2: Selected colonies are automatically isolated using a robot and stored separately in a biobank. 3: Strains in the biobank are identified using high-throughput genotyping methods such as 16S rRNA sequencing. 4: Isolates are stored in the freezer and their information is made public to facilitate their findability and accessibility. Figure adapted from reference[137].

Developments in computational biology and open data

Omics experiments generate a lot of data which require computers to process and interpret to produce biologically meaningful results. This in turn brings its own set of opportunities and challenges. One major challenge is the need for specialised researchers with expertise in both the biological context and the know-how of the computational methods. This requires specialised training, which is fortunately already integrated in Dutch universities where all biology programmes offer courses with some

form of bioinformatics. The Leiden University Medical Center is taking a next step and will include AI training into its medical programme[138].

Within the broader context of science, a compelling upside to computation compared to traditional wet-lab work is that computers facilitate full reproduction of experiments. Theoretically, rerunning a computer program is simpler than reproducing manual laboratory protocols. Reproducibility has been underappreciated in scientific research, even though it is crucial for validating results and strengthening the foundation of derivative works[139]. This is especially urgent in an era of fake news and a growing distrust in science[140]. Various initiatives aim to improve science's credibility and promote a more transparent evaluation process. For example, there is a broad movement of FAIR data: findable, accessible, interoperable, and reusable[141]. These principles are widely adopted, with many funding bodies demanding data to be made publicly available in grant applications, and scientific journals requiring data sharing as a condition for publication of manuscripts. On a similar note, the FAIR principles have been adapted to software[142], stimulating the publication of research software and thus improving reproducibility. Furthermore, there is a growing interest and appreciation of manuscripts preprints, accelerating knowledge dissemination[143]. These initiatives, and more, can give a boost to the credibility of science[144], while also stimulating different talents within science, a development I sincerely hope will be broadly adopted and valued in the scientific community.

Outlook: solving microbiota in health and disease questions

Connecting the pieces laid out here in the discussion, I envision further developments of host-microbe interaction studies and Open Science, opening avenues for collaborations between researchers of diverse expertise. If bacteriologists, medical microbiologists, gastroenterologists, immunologists, biochemists, bioinformaticians, statisticians and ecologists put their heads together and integrate their knowledge we can learn how the human gut microbiome functions and behaves under varying conditions. This approach should also include fungi and viruses as integral components of the microbiome[121,145-147]. A broader perspective is already exemplified in the national and international One Health movement[148,149], and is further solidified through initiatives such as the recently funded Holomicrobiome[150]. These initiatives aim to unravel all the intricacies of the ecosystems inside and around us, providing insights into the biological processes leading to disease. Understanding the biological processes helps us prevent or control diseases and improve treatment options. This may take the shape of new FMT strategies or live biotherapeutic products and lead to a new era of disease control and preventive medicine. Also, the timing of treatment may shift to prevent symptoms in people whose microbiota indicate an increased risk of developing disease. In this future view, fewer people develop disease, and those who do receive optimal therapies tailored to their needs.

References

1. Song, J. H. & Kim, Y. S. Recurrent Clostridium difficile Infection: Risk Factors, Treatment, and Prevention. *Gut Liver* **13**, 16-24 (2019). https://doi.org/10.5009/gnl18071
2. Minkoff, N. Z. *et al.* Fecal microbiota transplantation for the treatment of recurrent Clostridioides difficile (Clostridium difficile). *Cochrane Database Systematic Reviews* **4**, CD013871 (2023). https://doi.org/10.1002/14651858.CD013871.pub2
3. 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
4. Moayyedi, P. *et al.* Fecal Microbiota Transplantation Induces Remission in Patients With Active Ulcerative Colitis in a Randomized Controlled Trial. *Gastroenterology* **149**, 102-109 e106 (2015). https://doi.org/10.1053/j.gastro.2015.04.001
5. Lopetuso, L. R. *et al.* The first international Rome consensus conference on gut microbiota and faecal microbiota transplantation in inflammatory bowel disease. *Gut* **72**, 1642-1650 (2023). https://doi.org/10.1136/gutjnl-2023-329948
6. Groenewegen, B., Terveer, E. M., Joosse, A., Barnhoorn, M. C. & Zwittink, R. D. Fecal Microbiota Transplantation for Immune Checkpoint Inhibitor-Induced Colitis Is Safe and Contributes to Recovery: Two Case Reports. *Journal of Immunotherapy* **46**, 216-220 (2023). https://doi.org/10.1097/CJI.0000000000000474
7. Wang, Y. *et al.* Fecal microbiota transplantation for refractory immune checkpoint inhibitor-associated colitis. *Nature Medicine* **24**, 1804-1808 (2018). https://doi.org/10.1038/s41591-018-0238-9
8. Malard, F. *et al.* Pooled allogeneic faecal microbiota MaaT013 for steroid-resistant gastrointestinal acute graft-versus-host disease: a single-arm, multicentre phase 2 trial. *EClinicalMedicine* **62**, 102111 (2023). https://doi.org/10.1016/j.eclinm.2023.102111
9. van Lier, Y. F. *et al.* Donor fecal microbiota transplantation ameliorates intestinal graft-versus-host disease in allogeneic hematopoietic cell transplant recipients. *Science Translational Medicine* **12** (2020). https://doi.org/10.1126/scitranslmed.aaz8926
10. Youngster, I. *et al.* Fecal microbiota transplantation in capsules for the treatment of steroid refractory and steroid dependent acute graft vs. host disease: a pilot study. *Bone Marrow Transplant* **59**, 409-416 (2024). https://doi.org/10.1038/s41409-024-02198-2
11. Dailey, F. E., Turse, E. P., Daglilar, E. & Tahan, V. The dirty aspects of fecal microbiota transplantation: a review of its adverse effects and complications. *Current Opinion in Pharmacology* **49**, 29-33 (2019). https://doi.org/10.1016/j.coph.2019.04.008
12. Hallowell, H. A., Gao, A. L. & Suez, J. Double-edged sword: impact of fecal microbiome transplants on the gut resistome. *Current Opinion in Gastroenterology* **39**, 16-22 (2023). https://doi.org/10.1097/MOG.0000000000000894
13. Terveer, E. M. *et al.* Faecal microbiota transplantation for Clostridioides difficile infection: Four years' experience of the Netherlands Donor Feces Bank. *United European Gastroenterology Journal* **8**, 1236-1247 (2020). https://doi.org/10.1177/2050640620957765
14. DeFilipp, Z. *et al.* Drug-Resistant E. coli Bacteremia Transmitted by Fecal Microbiota Transplant. *New England Journal of Medicine* **381**, 2043-2050 (2019). https://doi.org/10.1056/NEJMoa1910437

15. Schewitz, K. in *Business Insider* https://www.businessinsider.com/woman-diy-poop-transplants-got-brother-acne-fecal-microbiota-transplant-2024-2025 (Insider Inc, A woman gave herself poop transplants using her brother's feces to treat debilitating IBS. Then she started getting acne just like him, 2024).

16. Drewes, J. L. *et al.* Transmission and clearance of potential procarcinogenic bacteria during fecal microbiota transplantation for recurrent Clostridioides difficile. *JCI Insight* **4** (2019). https://doi.org/10.1172/jci.insight.130848

17. Jouhten, H., Mattila, E., Arkkila, P. & Satokari, R. Reduction of Antibiotic Resistance Genes in Intestinal Microbiota of Patients With Recurrent Clostridium difficile Infection After Fecal Microbiota Transplantation. *Clinical Infectious Diseases* **63**, 710-711 (2016). https://doi.org/10.1093/cid/ciw390

18. Manges, A. R., Steiner, T. S. & Wright, A. J. Fecal microbiota transplantation for the intestinal decolonization of extensively antimicrobial-resistant opportunistic pathogens: a review. *Infectious Diseases* **48**, 587-592 (2016). https://doi.org/10.1080/23744235.2016.1177199

19. Leung, V., Vincent, C., Edens, T. J., Miller, M. & Manges, A. R. Antimicrobial Resistance Gene Acquisition and Depletion Following Fecal Microbiota Transplantation for Recurrent Clostridium difficile Infection. *Clinical Infectious Diseases* **66**, 456-457 (2018). https://doi.org/10.1093/cid/cix821

20. Millan, B. *et al.* Fecal Microbial Transplants Reduce Antibiotic-resistant Genes in Patients With Recurrent Clostridium difficile Infection. *Clinical Infectious Diseases* **62**, 1479-1486 (2016). https://doi.org/10.1093/cid/ciw185

21. Hourigan, S. K. *et al.* Fecal Transplant in Children With Clostridioides difficile Gives Sustained Reduction in Antimicrobial Resistance and Potential Pathogen Burden. *Open Forum Infectious Diseases* **6**, ofz379 (2019). https://doi.org/10.1093/ofid/ofz379

22. Bajaj, J. S. *et al.* Fecal Microbiota Transplant in Cirrhosis Reduces Gut Microbial Antibiotic Resistance Genes: Analysis of Two Trials. *Hepatology Communications* **5**, 258-271 (2021). https://doi.org/10.1002/hep4.1639

23. Terveer, E. M. *et al.* How to: Establish and run a stool bank. *Clinical Microbiology and Infection* **23**, 924-930 (2017). https://doi.org/10.1016/j.cmi.2017.05.015

24. Zhang, S. *et al.* Donor Screening for Fecal Microbiota Transplantation in China: Evaluation of 8483 Candidates. *Gastroenterology* **162**, 966-968 e963 (2022). https://doi.org/10.1053/j.gastro.2021.11.004

25. Narula, N. *et al.* Systematic Review and Meta-analysis: Fecal Microbiota Transplantation for Treatment of Active Ulcerative Colitis. *Inflammatory Bowel Diseases* **23**, 1702-1709 (2017). https://doi.org/10.1097/MIB.0000000000001228

26. Quraishi, M. N. *et al.* Systematic review with meta-analysis: the efficacy of faecal microbiota transplantation for the treatment of recurrent and refractory Clostridium difficile infection. *Alimentary Pharmacology & Therapeutics* **46**, 479-493 (2017). https://doi.org/10.1111/apt.14201

27. Schmidt, T. S. B. *et al.* Drivers and determinants of strain dynamics following fecal microbiota transplantation. *Nature Medicine* **28**, 1902-1912 (2022). https://doi.org/10.1038/s41591-022-01913-0

28. Haifer, C. *et al.* Lyophilised oral faecal microbiota transplantation for ulcerative colitis (LOTUS): a randomised, double-blind, placebo-controlled trial. *The Lancet Gastroenterology and Hepatology* **7**, 141-151 (2022). https://doi.org/10.1016/S2468-1253(21)00400-3

29. Podlesny, D. *et al.* Identification of clinical and ecological determinants of strain engraftment after fecal microbiota transplantation using metagenomics. *Cell Reports Medicine* **3**, 100711 (2022). https://doi.org/10.1016/j.xcrm.2022.100711

30. He, R. *et al.* The interplay of gut microbiota between donors and recipients determines the efficacy of fecal microbiota transplantation. *Gut Microbes* **14**, 2100197 (2022). https://doi.org/10.1080/19490976.2022.2100197

31. Kaplan, G. G. & Ng, S. C. Understanding and Preventing the Global Increase of Inflammatory Bowel Disease. *Gastroenterology* **152**, 313-321.e312 (2017). https://doi.org/10.1053/j.gastro.2016.10.020

32. Neurath, M. F., Sands, B. E. & Rieder, F. Cellular immunotherapies and immune cell depleting therapies in inflammatory bowel diseases: the next magic bullet? *Gut*, gutjnl-2024-332919 (2024). https://doi.org/10.1136/gutjnl-2024-332919

33. Super, M. *et al.* Biomaterial vaccines capturing pathogen-associated molecular patterns protect against bacterial infections and septic shock. *Nature Biomedical Engineering* **6**, 8-18 (2022). https://doi.org/10.1038/s41551-021-00756-3

34. Alameh, M. G. *et al.* A multivalent mRNA-LNP vaccine protects against Clostridioides difficile infection. *Science* **386**, 69-75 (2024). https://doi.org/10.1126/science.adn4955

35. Gupta, V. K. *et al.* A predictive index for health status using species-level gut microbiome profiling. *Nature Communications* **11**, 4635 (2020). https://doi.org/10.1038/s41467-020-18476-8

36. Shi, N., Li, N., Duan, X. & Niu, H. Interaction between the gut microbiome and mucosal immune system. *Military Medical Research* **4**, 14 (2017). https://doi.org/10.1186/s40779-017-0122-9

37. Dixit, K., Chaudhari, D., Dhotre, D., Shouche, Y. & Saroj, S. Restoration of dysbiotic human gut microbiome for homeostasis. *Life Sciences* **278**, 119622 (2021). https://doi.org/10.1016/j.lfs.2021.119622

38. Ianiro, G. *et al.* Variability of strain engraftment and predictability of microbiome composition after fecal microbiota transplantation across different diseases. *Nature Medicine* **28**, 1913-1923 (2022). https://doi.org/10.1038/s41591-022-01964-3

39. Maini Rekdal, V., Bess, E. N., Bisanz, J. E., Turnbaugh, P. J. & Balskus, E. P. Discovery and inhibition of an interspecies gut bacterial pathway for Levodopa metabolism. *Science* **364** (2019). https://doi.org/10.1126/science.aau6323

40. van Kessel, S. P. *et al.* Gut bacterial tyrosine decarboxylases restrict levels of levodopa in the treatment of Parkinson's disease. *Nature Communications* **10**, 310 (2019). https://doi.org/10.1038/s41467-019-08294-y

41. Vendrik, K. E. *et al.* Safety and feasibility of faecal microbiota transplantation for patients with Parkinson's disease: a protocol for a self-controlled interventional donor-FMT pilot study. *BMJ Open* **13**, e071766 (2023). https://doi.org/10.1136/bmjopen-2023-071766

42. Becattini, S. *et al.* Rapid transcriptional and metabolic adaptation of intestinal microbes to host immune activation. *Cell Host & Microbe* **29**, 378-393 e375 (2021). https://doi.org/10.1016/j.chom.2021.01.003

43. Pral, L. P., Fachi, J. L., Correa, R. O., Colonna, M. & Vinolo, M. A. R. Hypoxia and HIF-1 as key regulators of gut microbiota and host interactions. *Trends in Immunology* **42**, 604-621 (2021). https://doi.org/10.1016/j.it.2021.05.004

44. Litvak, Y., Byndloss, M. X. & Bäumler, A. J. Colonocyte metabolism shapes the gut microbiota. *Science* **362**, eaat9076 (2018). https://doi.org/doi:10.1126/science.aat9076

45. Wastyk, H. C. *et al.* Gut-microbiota-targeted diets modulate human immune status. *Cell* **184**, 4137-4153 e4114 (2021). https://doi.org/10.1016/j.cell.2021.06.019

46. Ross, F. C. *et al.* The interplay between diet and the gut microbiome: implications for health and disease. *Nature Reviews Microbiology* **22**, 671-686 (2024). https://doi.org/10.1038/s41579-024-01068-4

47. Clarke, S. F. *et al.* Exercise and associated dietary extremes impact on gut microbial diversity. *Gut* **63**, 1913-1920 (2014). https://doi.org/10.1136/gutjnl-2013-306541

48. Clancy, A. K., Gunaratne, A. W. & Borody, T. J. Dietary Management for Faecal Microbiota Transplant: An International Survey of Clinical and Research Practice, Knowledge and Attitudes. *Frontiers in Nutrition* **8**, 653653 (2021). https://doi.org/10.3389/fnut.2021.653653

49. Kedia, S. *et al.* Faecal microbiota transplantation with anti-inflammatory diet (FMT-AID) followed by anti-inflammatory diet alone is effective in inducing and maintaining remission over 1 year in mild to moderate ulcerative colitis: a randomised controlled trial. *Gut* **71**, 2401-2413 (2022). https://doi.org/10.1136/gutjnl-2022-327811

50. Lai, Z. L. *et al.* Fecal microbiota transplantation confers beneficial metabolic effects of diet and exercise on diet-induced obese mice. *Scientific Reports* **8**, 15625 (2018). https://doi.org/10.1038/s41598-018-33893-y

51. Rinott, E. *et al.* Effects of Diet-Modulated Autologous Fecal Microbiota Transplantation on Weight Regain. *Gastroenterology* **160**, 158-173 e110 (2021). https://doi.org/10.1053/j.gastro.2020.08.041

52. Centre for Digestive Diseases. *The CDD FMT Donor Diet*, <https://centrefordigestivediseases.com/the-cdd-donor-diet/> (2020).

53. Kelly, C. R., Kim, A. M., Laine, L. & Wu, G. D. The AGA's Fecal Microbiota Transplantation National Registry: An Important Step Toward Understanding Risks and Benefits of Microbiota Therapeutics. *Gastroenterology* **152**, 681-684 (2017). https://doi.org/10.1053/j.gastro.2017.01.028

54. Yau, Y. K. *et al.* Long-Term Safety Outcomes of Fecal Microbiota Transplantation: Real-World Data Over 8 Years From the Hong Kong FMT Registry. *Clinical Gastroenterology and Hepatology* **22**, 611-620.e612 (2024). https://doi.org/10.1016/j.cgh.2023.09.001

55. Cammarota, G. *et al.* International consensus conference on stool banking for faecal microbiota transplantation in clinical practice. *Gut* **68**, 2111-2121 (2019). https://doi.org/10.1136/gutjnl-2019-319548

56. Charbonneau, M. R., Isabella, V. M., Li, N. & Kurtz, C. B. Developing a new class of engineered live bacterial therapeutics to treat human diseases. *Nature Communications* **11**, 1738 (2020). https://doi.org/10.1038/s41467-020-15508-1

57. Navalkele, B. & Chopra, T. Clinical Application of Live Biotherapeutic Products in Infectious Diseases. *Frontiers in Microbiomes* **3** (2024). https://doi.org/10.3389/frmbi.2024.1415083

58. Rutter, J. W., Dekker, L., Owen, K. A. & Barnes, C. P. Microbiome engineering: engineered live biotherapeutic products for treating human disease. *Frontiers in Bioengineering and Biotechnology* **10**, 1000873 (2022). https://doi.org/10.3389/fbioe.2022.1000873

59. 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

60. Khanna, S. *et al.* RBX7455, a Non-frozen, Orally Administered Investigational Live Biotherapeutic, Is Safe, Effective, and Shifts Patients' Microbiomes in a Phase 1 Study for Recurrent Clostridioides difficile Infections. *Clinical Infectious Diseases* **73**, e1613-e1620 (2021). https://doi.org/10.1093/cid/ciaa1430

61. 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

62. Louie, T. *et al.* VE303, a Defined Bacterial Consortium, for Prevention of Recurrent Clostridioides difficile Infection: A Randomized Clinical Trial. *JAMA* **329**, 1356-1366 (2023). https://doi.org/10.1001/jama.2023.4314

63. 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

64. Kao, D. *et al.* The effect of a microbial ecosystem therapeutic (MET-2) on recurrent Clostridioides difficile infection: a phase 1, open-label, single-group trial. *The Lancet Gastroenterology and Hepatology* **6**, 282-291 (2021). https://doi.org/10.1016/S2468-1253(21)00007-8

65. Gerding, D. N. *et al.* Administration of Spores of Nontoxigenic Clostridium difficile Strain M3 for Prevention of Recurrent C difficile Infection: A Randomized Clinical Trial. *JAMA* **313**, 1719-1727 (2015). https://doi.org/10.1001/jama.2015.3725

66. Kaslow, D. C. Approval Letter - REBYOTA. (US Food & Drug Administration, STN: 125739, 2022).

67. Mendoza, M. K., David C. Approval Letter - VOWST. (US Food & Drug Administration, STN: 125757, 2023).

68. van der Lelie, D. *et al.* Rationally designed bacterial consortia to treat chronic immune-mediated colitis and restore intestinal homeostasis. *Nature Communications* **12**, 3105 (2021). https://doi.org/10.1038/s41467-021-23460-x

69. Atarashi, K. *et al.* Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* **500**, 232-236 (2013). https://doi.org/10.1038/nature12331

70. Furusawa, Y. *et al.* Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* **504**, 446-450 (2013). https://doi.org/10.1038/nature12721

71. Atarashi, K. *et al.* Induction of Colonic Regulatory T Cells by Indigenous *Clostridium* Species. *Science* **331**, 337-341 (2011). https://doi.org/doi:10.1126/science.1198469

72. Paton, A. W. *et al.* Recombinant probiotics for treatment and prevention of enterotoxigenic Escherichia coli diarrhea. *Gastroenterology* **128**, 1219-1228 (2005). https://doi.org/10.1053/j.gastro.2005.01.050

73. Sonnenborn, U. Escherichia coli strain Nissle 1917-from bench to bedside and back: history of a special Escherichia coli strain with probiotic properties. *FEMS Microbiology Letters* **363** (2016). https://doi.org/10.1093/femsle/fnw212

74. Duan, F. & March, J. C. Engineered bacterial communication prevents Vibrio cholerae virulence in an infant mouse model. *Proceedings of the National Academy of Sciences* **107**, 11260-11264 (2010). https://doi.org/10.1073/pnas.1001294107

75. Hwang, I. Y. *et al.* Engineered probiotic Escherichia coli can eliminate and prevent Pseudomonas aeruginosa gut infection in animal models. *Nature Communications* **8**, 15028 (2017). https://doi.org/10.1038/ncomms15028

76. O'Donnell, M. M. *et al.* Identification of ADS024, a newly characterized strain of Bacillus velezensis with direct Clostridiodes difficile killing and toxin degradation bio-activities. *Scientific Reports* **12**, 9283 (2022). https://doi.org/10.1038/s41598-022-13248-4

77. Nougayrede, J. P. *et al.* A Toxic Friend: Genotoxic and Mutagenic Activity of the Probiotic Strain Escherichia coli Nissle 1917. *mSphere* **6**, e0062421 (2021). https://doi.org/10.1128/mSphere.00624-21

78. Olier, M. *et al.* Genotoxicity of Escherichia coli Nissle 1917 strain cannot be dissociated from its probiotic activity. *Gut Microbes* **3**, 501-509 (2012). https://doi.org/10.4161/gmic.21737

79. Smillie, C. S. *et al.* Strain Tracking Reveals the Determinants of Bacterial Engraftment in the Human Gut Following Fecal Microbiota Transplantation. *Cell Host Microbe* **23**, 229-240 e225 (2018). https://doi.org/10.1016/j.chom.2018.01.003

80. Aggarwala, V. *et al.* Precise quantification of bacterial strains after fecal microbiota transplantation delineates long-term engraftment and explains outcomes. *Nature Microbiology* **6**, 1309-1318 (2021). https://doi.org/10.1038/s41564-021-00966-0

81. Olm, M. R. *et al.* inStrain profiles population microdiversity from metagenomic data and sensitively detects shared microbial strains. *Nature Biotechnology* **39**, 727-736 (2021). https://doi.org/10.1038/s41587-020-00797-0

82. Truong, D. T., Tett, A., Pasolli, E., Huttenhower, C. & Segata, N. Microbial strain-level population structure and genetic diversity from metagenomes. *Genome Research* **27**, 626-638 (2017). https://doi.org/10.1101/gr.216242.116

83. Quince, C. *et al.* STRONG: metagenomics strain resolution on assembly graphs. *Genome Biology* **22**, 214 (2021). https://doi.org/10.1186/s13059-021-02419-7

84. Hu, H. *et al.* StrainPanDA: Linked reconstruction of strain composition and gene content profiles via pangenome-based decomposition of metagenomic data. *iMeta* **1**, e41 (2022). https://doi.org/10.1002/imt2.41

85. Podlesny, D. *et al.* Metagenomic strain detection with SameStr: identification of a persisting core gut microbiota transferable by fecal transplantation. *Microbiome* **10**, 53 (2022). https://doi.org/10.1186/s40168-022-01251-w

86. van Dijk, L. R. *et al.* StrainGE: a toolkit to track and characterize low-abundance strains in complex microbial communities. *Genome Biology* **23**, 74 (2022). https://doi.org/10.1186/s13059-022-02630-0

87. Briscoe, L., Halperin, E. & Garud, N. R. SNV-FEAST: microbial source tracking with single nucleotide variants. *Genome Biology* **24**, 101 (2023). https://doi.org/10.1186/s13059-023-02927-8

88. Kim, Y. *et al.* Strain tracking with uncertainty quantification. *bioRxiv*, 2023.2001.2025.525531 (2024). https://doi.org/10.1101/2023.01.25.525531

89. Blanco-Miguez, A. *et al.* Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. *Nature Biotechnology* (2023). https://doi.org/10.1038/s41587-023-01688-w

90. Goldman, M., Zhao, C. & Pollard, K. S. Improved detection of microbiome-disease associations via population structure-aware generalized linear mixed effects models (microSLAM). *bioRxiv*, 2024.2006.2027.600934 (2024). https://doi.org/10.1101/2024.06.27.600934

91. Martin, J. S., Monaghan, T. M. & Wilcox, M. H. Clostridium difficile infection: epidemiology, diagnosis and understanding transmission. *Nature Reviews Gastroenterology & Hepatology* **13**, 206-216 (2016). https://doi.org/10.1038/nrgastro.2016.25

92. Croxen, M. A. *et al.* Recent advances in understanding enteric pathogenic Escherichia coli. *Clinical Microbiology Reviews* **26**, 822-880 (2013). https://doi.org/10.1128/CMR.00022-13

93. Wolter, M. *et al.* Diet-driven differential response of Akkermansia muciniphila modulates pathogen susceptibility. *Molecular Systems Biology* **20**, 596-625 (2024). https://doi.org/10.1038/s44320-024-00036-7

94. Abdelsalam, N. A., Hegazy, S. M. & Aziz, R. K. The curious case of Prevotella copri. *Gut Microbes* **15**, 2249152 (2023). https://doi.org/10.1080/19490976.2023.2249152

95. Blanco-Miguez, A. *et al.* Extension of the Segatella copri complex to 13 species with distinct large extrachromosomal elements and associations with host conditions. *Cell Host & Microbe* **31**, 1804-1819 e1809 (2023). https://doi.org/10.1016/j.chom.2023.09.013

96. Tett, A., Pasolli, E., Masetti, G., Ercolini, D. & Segata, N. Prevotella diversity, niches and interactions with the human host. *Nature Reviews Microbiology* **19**, 585-599 (2021). https://doi.org/10.1038/s41579-021-00559-y

97. Crost, E. H., Coletto, E., Bell, A. & Juge, N. Ruminococcus gnavus: friend or foe for human health. *FEMS Microbiology Reviews* **47** (2023). https://doi.org/10.1093/femsre/fuad014

98. Wassenaar, T. M. Insights from 100 Years of Research with Probiotic E. Coli. *European Journal of Microbiology and Immunology* **6**, 147-161 (2016). https://doi.org/10.1556/1886.2016.00029

99. Kaper, J. B., Nataro, J. P. & Mobley, H. L. Pathogenic Escherichia coli. *Nature Reviews Microbiology* **2**, 123-140 (2004). https://doi.org/10.1038/nrmicro818

100. Arnold, B. J., Huang, I. T. & Hanage, W. P. Horizontal gene transfer and adaptive evolution in bacteria. *Nature Reviews Microbiology* **20**, 206-218 (2022). https://doi.org/10.1038/s41579-021-00650-4

101. Brito, I. L. Examining horizontal gene transfer in microbial communities. *Nature Reviews Microbiology* **19**, 442-453 (2021). https://doi.org/10.1038/s41579-021-00534-7

102. Kent, A. G., Vill, A. C., Shi, Q., Satlin, M. J. & Brito, I. L. Widespread transfer of mobile antibiotic resistance genes within individual gut microbiomes revealed through bacterial Hi-C. *Nature Communications* **11**, 4379 (2020). https://doi.org/10.1038/s41467-020-18164-7

103. Bohm, M. E., Huptas, C., Krey, V. M. & Scherer, S. Massive horizontal gene transfer, strictly vertical inheritance and ancient duplications differentially shape the evolution of Bacillus cereus enterotoxin operons hbl, cytK and nhe. *BMC Evolutionary Biology* **15**, 246 (2015). https://doi.org/10.1186/s12862-015-0529-4

104. Putze, J. *et al.* Genetic structure and distribution of the colibactin genomic island among members of the family Enterobacteriaceae. *Infection and Immunity* **77**, 4696-4703 (2009). https://doi.org/10.1128/IAI.00522-09

105. Chagneau, C. V. *et al.* The pks island: a bacterial Swiss army knife? Colibactin: beyond DNA damage and cancer. *Trends in Microbiology* **30**, 1146-1159 (2022). https://doi.org/10.1016/j.tim.2022.05.010

106. Jiang, X., Hall, A. B., Xavier, R. J. & Alm, E. J. Comprehensive analysis of chromosomal mobile genetic elements in the gut microbiome reveals phylum-level niche-adaptive gene pools. *PLoS One* **14**, e0223680 (2019). https://doi.org/10.1371/journal.pone.0223680

107. Arredondo-Alonso, S., Willems, R. J., van Schaik, W. & Schurch, A. C. On the (im)possibility of reconstructing plasmids from whole-genome short-read sequencing data. *Microbial Genomics* **3**, e000128 (2017). https://doi.org/10.1099/mgen.0.000128

108. Krawczyk, P. S., Lipinski, L. & Dziembowski, A. PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Research* **46**, e35 (2018). https://doi.org/10.1093/nar/gkx1321

109. Camargo, A. P. *et al.* Identification of mobile genetic elements with geNomad. *Nature Biotechnology* (2023). https://doi.org/10.1038/s41587-023-01953-y

110. Suzuki, Y. *et al.* Long-read metagenomic exploration of extrachromosomal mobile genetic elements in the human gut. *Microbiome* **7**, 119 (2019). https://doi.org/10.1186/s40168-019-0737-z

111. Labonte, J. M. *et al.* Single cell genomics indicates horizontal gene transfer and viral infections in a deep subsurface Firmicutes population. *Frontiers in Microbiology* **6**, 349 (2015). https://doi.org/10.3389/fmicb.2015.00349

112. Brito, I. L. *et al.* Mobile genes in the human microbiome are structured from global to individual scales. *Nature* **535**, 435-439 (2016). https://doi.org/10.1038/nature18927

113. Press, M. O. *et al.* Hi-C deconvolution of a human gut microbiome yields high-quality draft genomes and reveals plasmid-genome interactions. *bioRxiv*, 198713 (2017). https://doi.org/10.1101/198713

114. Yaffe, E. & Relman, D. A. Tracking microbial evolution in the human gut using Hi-C reveals extensive horizontal gene transfer, persistence and adaptation. *Nature Microbiology* **5**, 343-353 (2020). https://doi.org/10.1038/s41564-019-0625-0

115. McCallum, G. E. *et al.* Noise reduction strategies in metagenomic chromosome confirmation capture to link antibiotic resistance genes to microbial hosts. *bioRxiv*, 2022.2011.2005.514866 (2022). https://doi.org/10.1101/2022.11.05.514866

116. Arıkan, M. & Muth, T. Integrated multi-omics analyses of microbial communities: a review of the current state and future directions. *Molecular Omics* **19**, 607-623 (2023). https://doi.org/10.1039/d3mo00089c

117. Sauceda, C. *et al.* Stool multi-omics for the study of host–microbe interactions in inflammatory bowel disease. *Gut Microbes* **14**, 2154092 (2022). https://doi.org/10.1080/19490976.2022.2154092

118. Zepeda-Rivera, M. *et al.* A distinct Fusobacterium nucleatum clade dominates the colorectal cancer niche. *Nature* **628**, 424-432 (2024). https://doi.org/10.1038/s41586-024-07182-w

119. Fukugaiti, M. H. *et al.* High occurrence of Fusobacterium nucleatum and Clostridium difficile in the intestinal microbiota of colorectal carcinoma patients. *Brazilian Journal of Microbiology* **46**, 1135-1140 (2015). https://doi.org/10.1590/S1517-838246420140665

120. Ning, L. *et al.* Microbiome and metabolome features in inflammatory bowel disease via multi-omics integration analyses across cohorts. *Nature Communications* **14**, 7135 (2023). https://doi.org/10.1038/s41467-023-42788-0

121. Liu, N. N. *et al.* Multi-kingdom microbiota analyses identify bacterial-fungal interactions and biomarkers of colorectal cancer across cohorts. *Nature Microbiology* **7**, 238-250 (2022). https://doi.org/10.1038/s41564-021-01030-7

122. Walker, A. W. & Hoyles, L. Human microbiome myths and misconceptions. *Nature Microbiology* **8**, 1392-1396 (2023). https://doi.org/10.1038/s41564-023-01426-7

123. Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J. & Segata, N. Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology* **35**, 833-844 (2017). https://doi.org/10.1038/nbt.3935

124. Rajilic-Stojanovic, M., Smidt, H. & de Vos, W. M. Diversity of the human gastrointestinal tract microbiota revisited. *Environmental Microbiology* **9**, 2125-2136 (2007). https://doi.org/10.1111/j.1462-2920.2007.01369.x

125. Costea, P. I. *et al.* Towards standards for human fecal sample processing in metagenomic studies. *Nature Biotechnology* **35**, 1069-1076 (2017). https://doi.org/10.1038/nbt.3960

126. Szostak, N. *et al.* The standardisation of the approach to metagenomic human gut analysis: from sample collection to microbiome profiling. *Scientific Reports* **12**, 8470 (2022). https://doi.org/10.1038/s41598-022-12037-3

127. Fernandez-Pato, A. *et al.* Choice of DNA extraction method affects stool microbiome recovery and subsequent phenotypic association analyses. *Scientific Reports* **14**, 3911 (2024). https://doi.org/10.1038/s41598-024-54353-w

128. Margos, G. *et al.* Evidence of taxonomic bias in public databases: The example of the genus Borrelia. *Ticks and Tick-borne Diseases* **13**, 101994 (2022). https://doi.org/10.1016/j.ttbdis.2022.101994

129. Ruscheweyh, H. J. *et al.* Cultivation-independent genomes greatly expand taxonomic-profiling capabilities of mOTUs across various environments. *Microbiome* **10**, 212 (2022). https://doi.org/10.1186/s40168-022-01410-z

130. Setubal, J. C. Metagenome-assembled genomes: concepts, analogies, and challenges. *Biophysical Reviews* **13**, 905-909 (2021). https://doi.org/10.1007/s12551-021-00865-y

131. Lagier, J. C. *et al.* Culture of previously uncultured members of the human gut microbiota by culturomics. *Nature Microbiology* **1**, 16203 (2016). https://doi.org/10.1038/nmicrobiol.2016.203

132. Zhang, J. *et al.* An immune-competent human gut microphysiological system enables inflammation-modulation by Faecalibacterium prausnitzii. *NPJ Biofilms Microbiomes* **10**, 31 (2024). https://doi.org/10.1038/s41522-024-00501-z

133. Xiang, Y. *et al.* Gut-on-chip: Recreating human intestine in vitro. *Journal of Tissue Engineering* **11**, 2041731420965318 (2020). https://doi.org/10.1177/2041731420965318

134. Min, S., Kim, S. & Cho, S. W. Gastrointestinal tract modeling using organoids engineered with cellular and microbiota niches. *Experimental Molecular Medicine* **52**, 227-237 (2020). https://doi.org/10.1038/s12276-020-0386-0

135. Schuren F, A. V., Keijser B, Abeln E, der Vossen J, Montijn R. The i-screen: A Versatile Preclinical Platform for Gut Microbiota Studies. *J Prob Health* **7** (2019). https://doi.org/10.35248/2329-8901.19.7.212

136. Williamson, I. A. *et al.* A High-Throughput Organoid Microinjection Platform to Study Gastrointestinal Microbiota and Luminal Physiology. *Cellular and Molecular Gastroenterology and Hepatology* **6**, 301-319 (2018). https://doi.org/10.1016/j.jcmgh.2018.05.004

137. Huang, Y. *et al.* High-throughput microbial culturomics using automation and machine learning. *Nature Biotechnology* **41**, 1424-1433 (2023). https://doi.org/10.1038/s41587-023-01674-2

138. Leiden University. *LUMC first medical programme to include AI in curriculum*, <https://www.universiteitleiden.nl/en/news/2024/08/lumc-first-medical-programme-to-include-ai-in-curriculum> (2024).

139. Sutter, P. M. *Rescuing Science: Restoring Trust in an Age of Doubts*. (Rowman & Littlefield Publishers, 2024).

140. Gerbina, T. V. Science Disinformation: On the Problem of Fake News. *Scientific and Technical Information Processing* **48**, 290-298 (2021). https://doi.org/10.3103/S0147688221040092

141. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18

142. Barker, M. *et al.* Introducing the FAIR Principles for research software. *Scientific Data* **9**, 622 (2022). https://doi.org/10.1038/s41597-022-01710-x

143. ASAPbio. *Biology preprints over time*, <https://asapbio.org/preprint-info/biology-preprints-over-time> (2020).

144. Korbmacher, M. *et al.* The replication crisis has led to positive structural, procedural, and community changes. *Communications Psychology* **1**, 3 (2023). https://doi.org/10.1038/s44271-023-00003-2

145. Haak, B. W. *et al.* Integrative Transkingdom Analysis of the Gut Microbiome in Antibiotic Perturbation and Critical Illness. *mSystems* **6** (2021). https://doi.org/10.1128/mSystems.01148-20

146. Kazemian, N. *et al.* The trans-kingdom battle between donor and recipient gut microbiome influences fecal microbiota transplantation outcome. *Scientific Reports* **10**, 18349 (2020). https://doi.org/10.1038/s41598-020-75162-x

147. Krautkramer, K. A., Fan, J. & Bäckhed, F. Gut microbial metabolites as multi-kingdom intermediates. *Nature Reviews Microbiology* **19**, 77-94 (2021). https://doi.org/10.1038/s41579-020-0438-4

148. Netherlands Center for One Health. *Homepage - NCOH*, <https://ncoh.nl> (2024).

149. World Health Organization. *One Health*, <https://www.who.int/health-topics/one-health> (2024).

150. Holomicrobiome Initiative. *Holomicrobiome Initiative*, <https://holomicrobioom.nl/en/home-en/> (2024).