5. Discussion

Plasmids are an important part of the bacterial genome and this thesis has shown their prevalence in the gut microbiome. The aims, as stated in the introduction, were to quantify the enrichment of plasmids in the culture collection and to investigate the distribution and functionality of AMR and mobility genes on the putative plasmids. This project successfully identified the distribution of putative plasmids and the prevalence of AMR genes in the plasmid-containing isolates, as well as the presence of mobility genes in a subset of plasmids. Investigating the spread of plasmids proved to be challenging; however, the results indicate that it may simply be a case of needing to find the right compatible pair.

5.1 Distribution of Plasmids in the Human Gut Microbiome Culture Collection

Plasmid distribution is predicted to follow a bimodal distribution with two peaks either side of the local minimum, usually around 20kb, as the standard population of plasmids will have large numbers of small mobilisable elements and bigger conjugative elements (Similie C. *et al.,* 2010). The distribution of putative plasmids in the gut microbiome doesn't follow the predicted bimodal model: the majority of the genomes present in the culture collections are predicted to contain only one plasmid, and these plasmids are predicted to be small, high-coverage elements. This may be a result of bias in the screening method, but alternatively this may reflect selection pressures in the niche.

The first explanation for observing primarily small, high-coverage plasmids is that this may be an artefact of the methods used to elucidate the plasmids. Antipov *et al. (*2016) identified that plasmidSPAdes struggles with the identification of plasmids with a coverage that is close to the median chromosomal coverage; since larger plasmids tend to be lower copy number they are more likely to have coverage similar to that of the chromosome. This suggests that plasmidSPAdes is not identifying the true number of plasmids across all size ranges in the gut microbiota. However, Jones and Marchesi (2007) observed similar results using the TRACA system to experimentally isolate plasmids from bacteria cultured from human stool samples. This could support the finding of primarily small elements as a true result, however, their experimental method may also be biased to isolate small elements over larger ones.

Arrendondo-Alonso *et al.* (2016) compared plasmidSPAdes to other completely automated platforms for computational plasmid isolation, giving them recall and precision scores. PlasmidSPAdes was compared to cBar (Zhou F. and Xu Y., 2010), Recycler (Rozov R. *et al.*, 2017), and PlasmidFinder (Carattoli *et al.,* 2014). PlasmidSPAdes had the highest recall score and was one of the software's able to identify novel plasmids. The study also identified the challenges with assembling large plasmids; plasmidSPAdes was flagged as being unable to assemble elements with their correct size,

however the median difference from correct size was 77bp. Ultimately, when compared to other completely automated methods of isolation plasmidSPAdes is the best for its purpose. Ideally plasmid isolation would make use of several automated platforms in conjunction with expert manual trimming, this however introduces high variability in predictions as trimming is dependent on operator expertise (Arrendondo-Alonso S. *et al.,* 2016, bioRxiv).

If the results presented reflect a genuine distribution pattern then this may serve to illustrate the cost of carriage mentioned in the introduction. It is proposed that there is a limit to the amount of extra DNA that can be carried, particularly in the absence of selective pressure. Considering that all the donors contributing to the culture collection are healthy and have not taken antibiotics for at least 6 months, it is likely there is only weak selection pressures in these donors' guts. For example, there is no primary need to maintain plasmids containing high carriage elements like multi-drug resistant cassettes. This point is further demonstrated by the fact that the most frequently identified resistance genes are tetracycline resistances; tetracycline is prevalent in the environment, primarily due to agriculture, and the donors would often be exposed to it (Popowska M. *et al.,* 2012).

5.2 Genomic Inference of Biological Functions: Plasmid Classification and AMR Distribution

Plasmid types exist for *Enterobacteriaceae, Pseudomonas*, *Staphylococcus*, and replicon typing strategies have been built around the replicons studied in these bacteria. The putative plasmid sequences were BLAST (blastn) against the PlasmidFinder database in an attempt to assign them to an incompatibility group. This strategy was not successful, yielding only 19 hits to the 240 plasmid-containing genomes in the culture collection. Similar strategies have been previously used to attempt to type Firmicute plasmids; for example, the work by Shintani *et al*. However, the plasmids classified were primarily of class *Bacilli,* with hits to only 4 out of the 90 plasmids from the *Clostridia* class in their dataset (Shintani M. *et al.,* 2015). The PF15101 *rep* gene domain was most frequently annotated in the putative plasmid sequences and the Pfam entry for this domain contains over 800 sequences. Using this extensive number of banked *rep* sequences I attempted to type the culture collection putative plasmids by aligning them to these reference sequences and looking at the incompatibility (Inc) groups of the plasmids they aligned to. However, the top BLASTx hits for the plasmids predicted in this study were to other uncharacterised plasmids. There is a large gap in the description of incompatibility groups and there is great scope for expanding our current knowledge of incompatibility groups to allow scientists to better understand all these novel extrachromosomal elements. This point is further illustrated by the phylogenies in Chapter 3: the trees are very low resolution, containing long shallow branches where presumably many genomes have not aligned to the classic *rep* genes. Despite current typing tools clearly only serving a small population of bacteria, aligning these genes does reveal distinct clusters. Classification of these groupings would allow for more educated experimental design around plasmid incompatibility and the ability to replicate in hosts.

The results of the AMR screen with ARIBA indicate a strong presence of AMR in the culture collection. Validating these predictions revealed them to be mostly accurate, with just 7 of the 44

tested predictions showing a different result; this may suggest the presence of non-functional genes, or genes that require induction. The gut microbiome is considered a reservoir for AMR and may contribute to the spread of AMR (Penders J. *et al.,* 2013). Mobility sequences are prevalent in the gut culture collection with 90 out of the 240 plasmid genomes containing plasmids with conjugation (*tra*) or *mob* sequences. In addition, 75% of these 90 are predicted to contain an AMR gene, further supporting the reservoir hypothesis. This may be an underestimation of the abundance of AMR in the gut as these predictions make use of known genetic variants to predict the presence of resistance; however, this would not indicate the presence of novel resistance genes or mechanisms of resistance. Novel resistance genes could be investigated by isolating the plasmids and transforming them into resistance free strains. These transformed bacteria can then be screened against a panel of antibiotics to observe if the addition of the plasmid confers any resistances that were not predicted during the *in silico* screen; this would be particularly useful for observing mechanisms such as novel efflux pumps.

The chord diagrams in Chapter 3 highlight potential broad range plasmids. The second chord diagram in particular identified plasmids in *Blautia* species that are shared with a *Klebsiella* strain. One of these elements is a small element containing tetracycline resistance, a toxin anti toxin system and plasmid recombination enzyme (Pre). The toxin anti toxin genes will help maintain the plasmid in the population; and the Pre protein is a mobilisable element, allowing the plasmid to be co-mobilised alongside conjugation competent plasmids (Priebe S. D. and Lacks S. A., 1989). In Chapter 4 plasmids have been demonstrated to be mobilisable, a plasmid from a health associated Firmicute was mobilised to Enterobacteriaceae species. The annotation of this plasmid revealed that it contained genes for citrate lyase, a component of the fatty acid synthesis chain; and genes for sorbitol transporters, in addition to conjugation genes and resistance genes. Citrate also known as E330 is a preservative found in many foods and household goods; and sorbitol is a sugar alternative found in sugar-free chewing gum and mints. Finding the genes for these processes on a mobile cassette that can move to a broad range of hosts, indicates potential adaptation to regularly consumed dietary nutrients. This is just one example of potentially many broad host range plasmids in the gut microbiota. The phylogenies in Chapter 3 display possible instances of plasmid sharing between the Firmicutes and all other phyla in the culture collection. These predictions can be used to set up more conjugation pairs and observe other instances of plasmid sharing *in vitro.*

A large portion of the Firmicutes are important health associated bacteria, this includes members of the genus *Blautia* which are important in the processing of polysaccharides (Eren A. M., *et al*., 2015) *Klebsiella* is a well-known pathogen and a major concern for the spread of AMR. It is on the WHO list of pathogens for which new antibiotics are critically needed; and recently a woman died from her infection with a strain of *Klebsiella* resistant to all available antibiotics in the United States (WHO, 2016; Chen L., *et al.,* 2016). The spread of plasmids is concerning not only because of the direct spread of resistance genes, but also the spread of genes related to adaptation to living in the gut. This means pathogens can more effectively colonise the gut and spend more time there, which can result

in chronic opportunistic infections as well as exposure to and acquisition of other extrachromosomal elements. One proposed theory regarding the formation of megaplasmids postulates that they are formed by the fusion of mini replicons (Zheng J., *et al.,* 2013). A plasmid with a broad host range *rep* is an ideal candidate for modular assembly of a megaplasmid through the addition of other functional cassettes; and the varying selection pressures of the gut may facilitate the assembly of plasmids with broad functionality.

To understand the breadth of functions present on the plasmids of the gut microbiota high quality sequences and detailed annotation are required. Long read sequencing of the isolated sequences would contribute to a better understanding of the functions encoded on these plasmids. Long read technologies can help resolve the problems with assembling repetitive regions and using the short read contigs as a reference or scaffold would help correct the errors that occur in long read sequencing, leading to more accurate annotations of functional genes and modification sites such as restriction sites. To investigate these annotations molecular cloning experiments could be used to identify the essential components of these novel plasmids and cassettes. This would allow for the creation of vectors for further genetic manipulation of gut commensals.

5.3 Transformation of Culture Collection Isolates

The difficulties with the experimental transformation may be due to a variety of reasons. The first strategy was to investigate natural competency. Natural competency is a highly coordinated process; in *Bacillus* species, it is controlled by the master regulator gene *comK* in response to a variety of environmental inputs that include nutrient availability and quorum sensing (Hamoen L.W. *et al.,* 2001). I attempted to use starvation to induce DNA uptake but was unsuccessful; this method may not have been the correct or sufficient stimulation to activate competence genes in the bacteria being tested.

The second strategy was to use electroporation. Plasmid transfer was previously tried with ruminuccocal species as described by Cocconcelli *et al.* (1992). The key difference between the protocol used by Cocconcelli and the protocol detailed in Chapter 2 is that the Cocconcelli experiment was carried out entirely under strict anaerobic conditions. During my procedure, some aerobic exposure was necessary, though the time spent in aerobic conditions was kept as short as possible. The bacterial cells did survive the protocol as there was growth on control plates, however the cells plated on selection did not grow. This suggests the stress of having to recover from aerobic exposure and the electrical pulse, in addition to the stress of maintaining a plasmid to survive antibiotic selection may have been too challenging to survival. The other issue highlighted by Cocconcelli *et al. (1992)* was plasmid and host compatibility: they inferred that one of their transformations may have failed as the plasmid could not replicate in the recipient. My candidate recipients were selected to be as phylogenetically close as possible but still may not have been suitable. A final possible source of the failure may have been ineffective competent cells. As previously mentioned, the growth phase of bacteria can affect competency: the growth curves of the potential recipients were plotted (Chapter 4) to observe growing times. Using this data cells were harvested at different time points and a variety of

competency-inducing methods were used, including ice washes and enzymatic cell wall degradation. This points to replication incompatibility being the key reason for the lack of transformants.

The final strategy was to use conjugation, which may also have been affected by plasmid and host compatibility. Another reason for difficulties may be due to the fact that conjugation relies on auxiliary machinery. This is normally encoded on large plasmids, and they are self mobilisable. In contrast, smaller plasmids, such as those identified in the culture collection, tend to contain mobilisable elements which only facilitate mobilisation using the hosts conjugation machinery or in conjunction with larger plasmids encoding their own machinery.

5.4 Applications of Identified Plasmid Sequences

The identification of multiple novel plasmids in commensal bacteria has many implications; one of the immediate applications of the screening is the creation of a database of plasmid sequences. This database can be used to identify extrachromosomal elements, particularly in metagenomic datasets, addressing one of the challenges to metagenomic analysis. As these sequences were isolated from whole genome sequences of pure cultures, this database would make it possible to make more accurate inferences of host-extrachromosomal DNA relationships in metagenomic datasets.

In addition to improving metagenomic data analysis, the database of plasmid sequences has broader applicability. Long read sequencing of these isolated elements, using the short-read sequencing data as a reference, will lead to a more accurate overall sequence assembly and better annotation of the genes. This high-resolution annotation then forms a foundation for future applications in genetic engineering. For example, the completed database will provide a collection of backbones for building vectors tailored to the microbiota. The vast potential applications of engineering the microbiota will be discussed in detail below

5.5 Plasmids and Plasmid Sequence Data as Genetic Tools

The microbiota are responsible for many functions in the gut and as presented in Chapter 1 these functions may be encoded on plasmids. Some functions include aiding in digestion, immunological and neurological signalling, and the regulation of bacterial growth. Digestion is aided through the fermentation of sugars and fibres; processing of short chain fatty acids; and butyrate production. Immunological signals to the gut epithelia include the stimulation of mucus production, the induction of Treg cells, and neurological signalling includes the production of molecules that stimulate the gut-brain axis. The growth of bacterial communities is regulated by the production of bacteriocins and other antimicrobial proteins to induce death, and products to regulate the formation of biofilms (Kali A., 2015).

With a wide range of information from the human body presented to the microbiota, the wide range of effects that they can produce, and high compatibility with the host, commensal bacteria are poised to be ideal sensors and effectors in the gut. The advances in techniques of synthetic biology have allowed the manipulation of well-characterised bacterial strains; e.g. developing engineered bacteria

and even isolating key genetic components to build cell-free systems for the production of proteins, metabolites, and therapeutics (Hodgman C.E. and Jewett M.C., 2012). The manipulation of commensal strains has been limited due to difficulties with culturing them for biological and genomic analysis. The development of endogenous and exogenous techniques as discussed in the introduction has led to increased isolation of commensal bacteria and their genetic information. With better understanding of both commensal bacteria and their mobile genetic elements, genetic engineering of these organisms is now ready to be realised. Commensal bacteria provide an advantage over traditional commercial strains by being a stable part of the host biological make up, able to exist in the niche over extended periods of time without causing a disease response, and likely contributing to host health.

One of the main outputs of bacterial engineering is the development of a DNA-based memory device; this would be a device with the ability to recognise input from its surroundings and execute pre-programmed responses. This would allow for a system with real time responses providing early diagnosis and treatment of conditions in the gut. Bacteria could be engineered to assist in the diagnosis of diseases by programming them for use as biomarkers and diagnostic sensors. For example, bacteria have previously been used as biomarkers in the screening of tumours as they can easily colonise tumour cells. These bacteria have been programmed with genes allowing them to produce fluorescence, bioluminescence, magnetic particles, or positrons for tumour imaging; in addition, they can infect in a tumour cell specific fashion enabling the monitoring of metastasis (Bernardes N. *et al.,* 2013). Engineering the native flora to produce these responses as a result of changes in the host that indicate cancer would provide a stable and long-term monitoring system and enable rapid diagnosis.

Beyond signals and diagnosis, bacteria can be programmed to elicit an active response that aides in combating an illness. Bacteria have been programmed to release substances in response to inflammation, neoplastic changes, infection, tissue damage – amongst other signals. Inflammation in the gut is a large concern since it features in many disease processes of this organ, including allergic responses, inflammatory bowel disease, and bowel cancers (Shen N. and Clemente J.C., 2015). Programming gut bacteria could lead to more comfortable lives allowing people to eat foods they previously couldn’t, preventing and soothing painful flare ups, and minimising the amount of inflammatory damage faced by the gut epithelium – a major risk factor for the development of bowel cancer.

Furthermore, commensal bacteria could be used as a next generation strategy for vaccination against infection or on-going prophylaxis against infection, by engineering them to produce small molecules or peptides that stimulate antigen production. This would be particularly useful for the immunocompromised, giving an increased level of protection and allowing infection to be identified and treated faster (Braff D. *et al.,* 2016; Kali A., 2015). Other strategies being developed include using bacteria to produce small molecules that disrupt infection by altering signalling within the community.

Goh *et al* describe a method by which bacteria were programed to secrete cholera auto-inducers- CAI-1 and AI-2, small molecules that are used by *Vibrio cholera* in quorum sensing. By sensing its population density *V. cholera* can determine whether or not to secrete toxins. Goh *et al.* (2012) found that the synthetic bacteria were successful in inducing density-dependent quorum signalling, preventing the secretion of toxins by *V. cholera*. Strategies like this could be used to prevent opportunistic infections colonising niches on the human body.

Other small molecules that bacteria can be used to produce include anti-cancer agents. Cancer treatments have been developed using antibodies to deliver small molecules in a specific fashion. Scientists are now looking to use similar strategies with bacteria (Bernades N. *et al.,* 2013). This strategy would be especially potent to treat bowel cancers since it could be delivered in a highly localised manner using microorganisms already prevalent in the niche. In addition, bacteria may have an advantage over antibody therapy by overcoming several issues including: antibody degradation to their high gastrointestinal instability (Casanova E.B., 2013); the inability of antibodies to penetrate the centre of tumours; and their stimulation of the immune response that can result in cytokine release syndrome or hypersensitivity (Deng R. *et al.,* 2012; Elbakri A. *et al.,* 2010). Bacteria can easily 'infect' all the cells in a tumour and commensal species should not cause an immune response in the gut as the bacteria are familiar to the host.

5.6 Strategies to Engineer the Microbiota

The primary barrier to employing any of these therapies is difficulties in engineering commensal bacteria. Strategies and tools are defined for well-characterised bacteria like commercial stains of *E. coli,* but not for commensal species. Chapter 1 introduced us to the main challenges faced when studying commensals including the fact that many commensal organisms remained unculturable. This has been combatted with a variety of strategies including the work Browne *et al.* (2016), which allowed the development of the culture collection utilised in this study. The next hurdle is identifying genetic elements compatible with commensal species that can be used as vectors; the experimental and computational isolation strategies have been developed complement each other, and this thesis has aimed to illustrate that.

The next step is annotation and isolation of genes encoded by the isolated plasmids. Classification of replicons and functional annotation will provide insights into potentially novel functions encoded on plasmids, and isolation of the annotated genes will allow the building of synthetic gene circuits. Synthetic circuits enable the building of precise networks of programmed responses, including failsafe's, creating controlled genetically modified organisms (GMOs).

Three particularly pertinent types of gene circuit include dynamic, logic and communication circuits detailed in Figure 1. Dynamic circuits involve interlinked positive- and negative-feedback loops, which produce an oscillatory response. Figure 1 gives an example of positive arabinose signalling linked with negative *lac* repressor signalling. Each feedback loop is modulated by the abundance of small

molecule inducers and the reciprocal signalling produces oscillations in green fluorescent protein expression (GFP). This dynamic circuit system could be useful for developing bacterial prophylactics: bacteria would produce antimicrobials molecules in response to the positive disease signal, but this response could be switched off in the absence of a condition. Logic circuits make use of 'AND' and 'NOT' gates which can be used to build systems that follow a set of conditional rules. The system in Figure 1 requires the presence of both m1 'AND' m2 to inhibit the production of yellow fluorescent protein (YFP); systems like this could be used to build very specific biomarkers based on cell type as well as intracellular signals. Communication circuits can use quorum sensing as a method of population control: the example in Figure 1 uses the small molecule AHL to signal density; increased density and therefore increased AHL induces expression of a killer gene, reducing population density. Communication signals and circuits of this type are manipulated in the method of inhibiting *V. cholera* toxin secretion described above (Goh YL *et al.,* 2012).

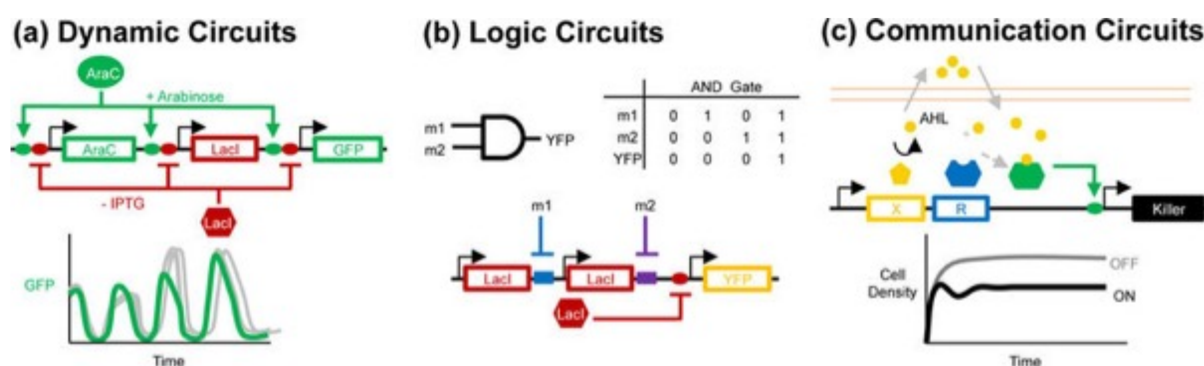

Figure 5.1. *Picture displaying the three key synthetic circuits: dynamic, logic, and communication circuits.* (a) Dynamic circuits employ interlinked positive and negative feedback loops to produce an oscillatory response. This example uses a positive arabinose signalling linked with negative *lac* repressor signalling. Each feedback loop is modulated by the abundance of small molecule inducers and the reciprocal signalling produces oscillations in green fluorescent protein expression (GFP). (b) Logic circuits use "AND" or "NOT" gates to build conditional rules for the system to follow. This example requires the presence of both m1 'AND' m2 to inhibit the production of yellow fluorescent protein (YFP). (c) Communication circuits use quorum sensing for population control. This example shows the small molecule AHL being used to signal density; increased density and therefore increased AHL induces expression of a killer gene, reducing population density. These circuits are assembled to build complex networks of responses, which are then often delivered on plasmid DNA (Chen Y.Y. *et al.*, 2012)

The final step is the actual engineering process and physical manipulation. As shown in Chapter 4, transformation of commensals is not always straightforward. As we continue to study and understand commensals we will be able to optimise methods of transformation, allowing us to develop *in vitro* and *ex vivo* bacteriotherapies.

For *in vivo* alterations, viral vectors have been shown to be highly efficient gene therapy vehicles, and are the most commonly used therapeutic vector (Wirth T. *et al.,* 2013). Investigation into the gut virome is a growing field and will likely lead to the discovery of new phages and the development of new tools and techniques for probing metagenomic datasets (Ogilvie L.A. and Jones B.V., 2015).

Plasmid delivery will be greatly enhanced by the creation of a database of gut phages, providing a catalogue of gut specific vectors that can be used to alter commensal bacteria. Phage based plasmid delivery would be advantageous as phages are often highly species specific making them great candidates for *in vivo* microbiome engineering. Phage-based strategies have been used as antimicrobial therapies, by delivering antimicrobial gene networks, or by inducing death directly. (Braff D. *et al.,* 2016). Second generation phage strategies include the use of phagemids: non-replicating viral capsules, capable of effectively delivering high copy number synthetic plasmids. Krom *et al.* (2015) developed this strategy for the delivery of non-lytic antimicrobials and to investigate alternatives to traditional antibiotics, finding that it successfully caused lysis free death of peritonitis causing *E. coli* in mice. The modular design of the system would allow for the easy production of phagemids targeting a variety of bacteria, which could induce transient product generation or stable integration of a new segment of DNA. Therefore, the database of plasmids from this study can contribute to a range of templates for designing phagemids to target the gut microbiota; and additional studies on novel plasmids and phages will extend the range of available templates.

5.7 Challenges to Bacteriotherapy

In vitro and *ex vivo* strategies will allow for the large-scale production of engineered microbes, but there are challenges to the delivery of these bacteriotherapies. One of the key challenges to bacteriotherapy is identifying a suitable route of administration. Probiotics are often used to bolster the gut microbiome and are generally administered orally, however the amounts of bacteria that survive the oral route of admission are very variable between manufacturers due to species robustness and delivery vehicle (Govender M. *et al.,* 2014). Probiotics primarily utilise lactic acid bacteria (LAB) as they are known for their tolerance of gastric acid; this includes the recently reported successful and proof-of-principle trial in India that definitively demonstrated the beneficial effects of probiotics (Panigrahi P. *et al.,* 2017)

The current strategy for delivering bacteriotherapies to the gut microbes is faecal microbiota transplantation (FMT); this therapy is primarily used for the treatment of recurring *Clostridium difficile,* however it is being investigated as a treatment for other conditions. This strategy typically involves the delivery of a native complex community of bacteria via enema, colonoscopy, or nasogastric tube, but could be adapted to introduce other bacterial populations including engineered commensals. One of the limitations to the use of this therapy is the public attitude to FMT. While the majority of the research surrounding attitudes has been conducted around recurring *C. difficile* infection (CDI) scenarios, the results are telling and have prompted investigation into the role language and the media can play in changing perceptions surrounding the procedure. Suggesting that moving away from terminology such as “the ick/yuck factor” when talking about the procedure will provide a more positive image and increase receptivity (Chuong K.H. *et al.,* 2015). Patients are reported to be receptive to the treatment if mandated by a physician, and if the contents are made more aesthetically pleasing, i.e. in a pill or clear solution. Many also indicated discomfort with route of delivery, preferring a colonoscopy to nasogastric tube delivery (Park L. *et al.,* 2017).

Once delivered the next major issue is the problem of stable engraftment: often therapeutic effects are lost as bacteria are pushed out of the community (Mimee M. *et al.,* 2016). Using engineered commensal species over commercial LAB should help as the bacteria would be part of the native community and not outcompeted. The ideal strategy, particularly for prophylactic bacteriotherapy, may be an *ex vivo* therapy where individuals provide a sample that is then engineered and supplied back to them. This would address another major patient concern, which is the origin and safety of donors.
Another consideration is the biosafety of these engineered bacteria and any side effects they could have in the host and effects after being released in to the environment. Ideally, we should be able to induce the death of the engineered species in a specific manner and using a mechanism that is not likely to induce resistance. One solution may be to include a 'kill switch' in the engineering of the synthetic circuits, that will eliminate either the modified organism or the engineered DNA. (Sonnenberg J. L., 2015). Cailando and Voigt (2015) designed a CRISPR strategy that resulted in the degradation of the introduced synthetic plasmid once engineered organisms left their containment zone and were no longer exposed to specific amino acids; gut specific metabolites could be used to facilitate a similar strategy.

5.8 Non-Health Associated Uses of Engineered Microbiota

Engineering the microbiome can provide benefits other than direct improvement of an individuals' health, including environmental and economic benefits. According to the World Health Organisation (WHO) at least 2 billion people globally are using a water source contaminated with faeces and facing an increased risk of contracting a water borne disease (WHO, 2017). The BioVolt by Cambrian Innovation (MA, USA) is one example of a microbial fuel cell (MFC) being used to clean up water and create energy simultaneously. MFCs consist of bacteria that produce electrons as they respire, producing clean water, and enough energy to sustain themselves plus excess. Engineering strategies are already being utilised in this field with Kirchhofer *et al*., (2017) providing a strategy for engineering to increase their electrical output. Stool bacteria could be engineered to conduct this procedure *in situ,* which would increase the energy output per litre of wastewater. A strategy like this would allow for easy environmental clean-up of water reservoirs, leading to increased population health and providing alternative energy sources.

5.9 Strategies for Investigating Engineered Microbiota

The examples above show that engineering the microbiome can provide solutions to a variety of health and non-health associated problems. However, the development of these solutions requires a variety of techniques in our "lab toolbox". We can now: culture gut bacterial species, scan the genomes for extrachromosomal elements, isolate these elements, and the techniques for manipulating the microbiota are being investigated. The next hurdle will be to develop techniques for *in situ* manipulations and culture. The current models being used include 2D and 3D cell culture, and gnotobiotic mouse models. Intestinal cells from patients, healthy donors, or induced pluripotent stem cells (iPSCs) are either grown in traditional 2D, or in 3D organoids. The organoid models allow for the

study of individual differences, as well as population-level differences, in an anatomically relevant model. Gnotobiotic animals are raised in a germ-free environment and inoculated with the desired community of bacteria. This model has been widely used for studying the microbiome, with a PubMed search for germ-free ‘AND’ microbiome revealing nearly 300 papers published in the last year with over 30 published in high impact journals.

The next generation of *in situ* testing tools may be the organ-on-chip models (Bhatia S.N., and Ingber D.E., 2014). These devices consist of a microchip with microfluidics channels engrafted with cells of the desired organ arranged appropriately within the 3D space. Once again, the channels can be engrafted with cells from patients, healthy donors, or iPSCs allowing different levels of variation to be studied. The channel is irrigated with the appropriate media; and these models can be used to study a wide variety of interactions. Mechanical forces can be modelled with the liquid flow, and interactions between various cells can be studied either by addition to the main channel or culturing in adjacent channels (Ingber D.E. 2016). This allows for a highly modular system where stimuli can be easily added and removed one at a time. This method has been used to culture bacteria with an intestinal epithelium and this system was maintained for several weeks – much longer than is possible with organoid cultures. In addition to bacteria, circulating immune cells can be added. This was used to demonstrate that the presence of an *E. coli* endotoxin only caused an irritable bowel disease (IBD) phenotype when circulating immune cells were present as well (Kim H.J. *et al.,* 2016). This is a highly tractable system that allows for highly flexible real-time investigation, and the use of cell cultures works towards the goals of the 3Rs to limit the use of animal models, which are currently imperative for these types of studies.

5.10 Summary

This body of work aimed to introduce and discuss the biology behind plasmids of the gut microbiome, investigate and validate the presence of plasmids in the microbiota, and provide scope for the future uses of these extrachromosomal elements. The plasmids observed in the microbiota provide insights into the prevalence extrachromosomal DNA and its function the gut, and present a starting point for the development of commensal engineering vectors.