2. Methods

The Lawley Laboratory Culture collection was created from healthy donor faecal samples as described by Browne *et al.*, and this body of work makes use of the whole genome sequences (WGS) and the pure cultures of human gut microbiota species (Browne H. *et al* 2016).

The workflow is summarised in the flow diagram in Figure 1. WGS data were scanned with plasmidSPAdes for putative plasmid sequences. These sequences were BLAST (blastn) against the NCBI database to filter out well-known and well-characterised plasmids. The remaining plasmids were annotated and their annotations probed for functional genes such as mobility and AMR genes. The presence of the plasmids and their predicted phenotypes were then validated experimentally. In addition, growth curves were plotted, the isolates were 16s sequences validated, and plasmid DNA was extracted for long read sequencing.

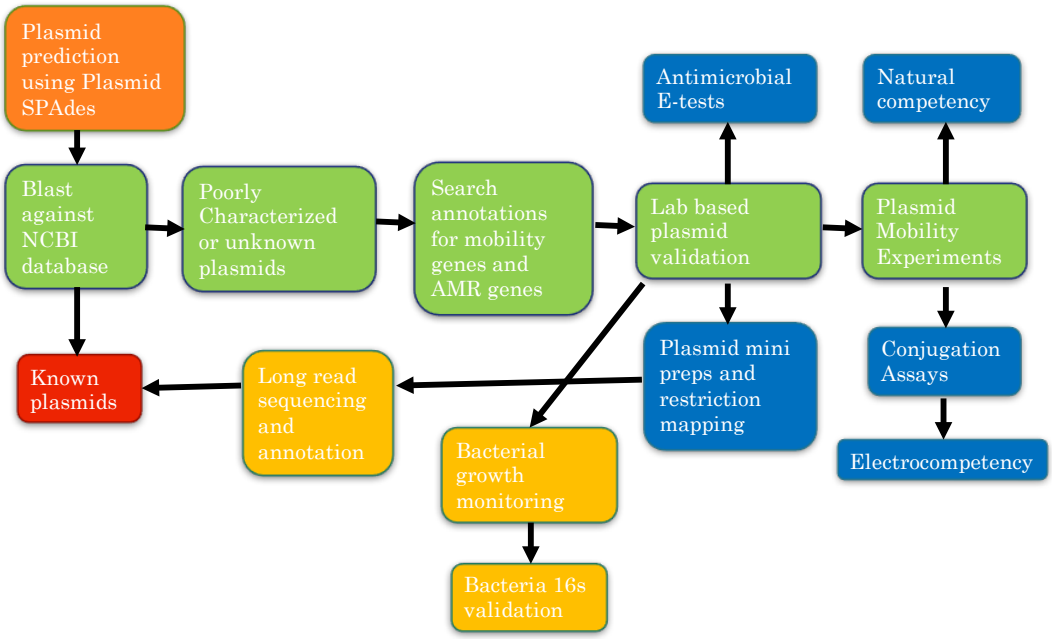

Figure 2.1. Workflow diagram summarising the strategy taken to probe the plasmids of the human gut microbiome.

2.1 Computational Plasmid Isolation, Phenotypic Predictions, and Phylogeny

PlasmidSPAdes, developed by Antipov *et al.* scans WGS data observing the coverage of the contigs (Antipov D. *et al,* 2016). Contigs are called as either chromosomal or plasmid contigs based on their deviation from the median coverage; plasmid contigs are then assembled to provide putative plasmid sequences. PlasmidSPAdes was run with the default parameters (*MaxDeviation* 0.3), as uniform coverage of genome was assumed.

The plasmid sequences were then BLAST (blastn) against the NCBI database using default parameters (*megablast*; *E threshold=10; word size= 28*) to isolate the well-characterised plasmids present in the culture collection.

The plasmid sequences were then annotated using the Sanger Pathogen Informatics Pipeline annotation tool. This tool searches the following databases to annotate sequences.

- RefSeq databases
- UniprotKB (bacteria/virus databases)
- Clusters
- Conserved domain database
- tigrfams
- pfam (A)
- rfam

Antibiotic Resistance Identification by Assembly (ARIBA) developed by Hunt *et al*, identifies resistances genes through alignment and targeted local assembly (Hunt M. *et al.*, 2017, bioRxiv). WGS were scanned for AMR genes using ARIBA with the Comprehensive Antibiotic Resistance Database (CARD- McArthur A.G. *et al,* 2013) and Short Read Sequencing Typing 2, Antimicrobial Resistance Gene Annotation (SRST2_ARGannot- Inouye M. *et al.,* 2014) databases. ARIBA was run using the default parameters as it is engineered to find the best match for a sequence. Putative plasmid sequences were scanned using BLAST against the CARD, SRST2_ARGannot, and ResFinder (Zankari E. *et al.*, 2012) databases at 60%, 70%, and 80% identity, the results were consistent across sensitivities and the results from the 70% identity screen are reported in Chapter 3.

Replicon typing was attempted using the PlasmidFinder Database (Carattoli *et al.,* 2014); plasmid sequences were BLAST (blastn) against the database at 70% identity. Plasmid sequences were also BLAST (blastx) against the database of sequences for *rep* domain PF01051**,** using an expected threshold of 10.

To build the phylogenies multiple sequence comparison by log-expectation (MUSCLE) was used to align the culture collection sequences to the RepA gene of an IncFII plasmid and to a RepE sequence of an unclassified plasmid, both from the culture collection. Muscle was run with default parameters (*maxiters 16*) to provide the most accurate alignment (Edgar R.C., 2004). FastTree was used to build the phylogenies using the generalised time-reversible model flag; this was selected as it is considered the most reliable general nucleotide evolutionary model (Price M.N. *et al,* 2010; Tavaré S. 1986).

2.2 Bacterial Culture and Growth Curve Plotting

Bacterial culture for aero-tolerant isolates was carried out under standard aerobic conditions (37°C, 5% CO_2). Aero-sensitive species were cultured under anaerobic conditions (37°C, Anaerobic gas

(ANO2) =10% H_2, 10% CO_2, 80% N_2). Bacteria were cultured in YCFA (Yeast extract-Casein hydrolysate-fatty acids) medium broth and plates (composition below).

To observe the growth curves bacteria were inoculated in 200µl of YCFA broth in a flat bottomed 96 well plate (Costar) and the optical density (OD) readings were recorded every 30 min over 24hr by the FLUOstar Omega Microplate Reader (BMG Labtech).

2.3 Growth Medium

YCFA medium (pH 7.45) consisted of (per 100ml):

Bacto Casitone	1.0g
Yeast extract	0.25g
$NaHCO_3$	0.4g
Glucose	0.2g
Maltose	0.2g
Cellobiose	0.2g
Mineral Solution I ($3g\ K_2HPO_4$, 1L $d.H_2O$)	15ml
Mineral Solution II ($3g\ KH_2PO_4$, $6g\ (NH_4)_2SO_4$, 6g NaCl, $0.6g\ MgSO_4$, $0.6g\ CaCl_2$, 1L $d.H_2O$)	15ml
VFA mix (17ml Acetic acid, 6ml Propionic acid, 1ml n-Valeric acid, 1ml Isovaleric acid, 1ml Isobutyric acid, 26ml NaOH(10M))	0.62ml
Haemin Solution (0.28g KOH, 25ml Ethanol 95%, 100mg Haemin, ≤100ml)	1ml
Vitamin solution I (1mg Biotin, 1mg Cobalamin, 3mg PABA (P-amino benzoic acid), 5mg Folic acid, 15mg Pyridoxine, d.H2O ≤100ml)	100µl
Vitamin solution II (5mg Thiamin, 5mg Riboflavin, d.H2O ≤100ml)	100µl (after autoclaving)
Resazurin (0.1%)	0.1ml
L-cysteine	0.1g
$d.H_2O$	≤100ml
Bacto Agar	1.6g

2.4 Species Validation by 16S rRNA gene PCR

The isolates were species validated using PCR of the 16S rRNA gene. PCR was carried out using the standard broad bacterial 16S rRNA gene primers (Sigma):

- Forward-Bact-7F (5'-AGA GTT TGA TYM TGG CTC AG-3')
- Reverse-Bact-1510R (5'-ACG GYT ACC TTG TTA CGA CTT-3')

and the GoTaq® Hot Start Polymerase kit (Promega). The reactions were completed in 25µl reaction volumes, and the master mix composition as follows: 1x Green GoTaq® Flexi Buffer, 25mM $MgCl_2$, 10mM dNTPs,10pM each forward and reverse primers, 1.25 U GoTaq® Polymerase, 2µl 1/10 dilution overnight culture, and sterile water up to 25µl. An *E. coli* sample and a sterile water sample were used as the positive and negative control respectively.

The PCR was run according to the protocol below on the DNA Engine Tetrad (MJ Research) with a constant lid temperature of 100ºC:

- 95ºC - 15mins
- 95ºC - 30secs \
- 58ºC - 30secs → 30 cycles
- 72ºC - 2mins /
- 72ºC - 8mins
- 10ºC - holding temperature.

The PCR product clean up and sequencing were provided by Eurofins Genomics. Sequence files were aligned to the 16S rRNA sequence database of the culture collection to identify the species matches.

2.5 Plasmid Extractions and Digests

Plasmid extraction was conducted using the QIAprep min prep kit (Qiagen) and the AquaPlasmid plasmid preparation buffer (MultiTarget), according to manufacturers' instructions. The Qiagen kit is a column based system and the upper size limit stated by the protocol is 12kb; the AquaPlasmid buffer is a column free application and suitable for plasmids of all sizes including megaplasmids. Extractions on a plasmid containing *E. coli* strain were done in parallel as a technical control.

Restriction sites were predicted using SnapGene (GSL Biotech), single cutters were preferred. Plasmid digests were set up in 50µl reaction volumes with 1U of enzyme (NEB), 1x buffer (NEB), and 500ng of DNA for 2 hours.

2.6 Plasmid Visualisation by Gel Electrophoresis

After isolation and digestion, plasmids were visualised by gel electrophoresis on 0.8% agarose in Tris-acetate EDTA (TAE). Gels were run at 75V for 1 hour; stained in TAE with 0.5 mg/L ethidium bromide for 30 mins; and visualised using the GelDoc-It^2® system (UVP).

2.7 Antimicrobial Sensitivity Testing

Antimicrobial sensitivity levels were tested using Etest strips (Biomérieux) on YCFA plates according to manufacturer's instructions.

2.8 Transformation

Natural competency can be induced by stimuli including DNA damage and starvation; starvation was used to attempt to induce natural competency in commensals (Huddleston J. R., 2014). Overnight bacterial cultures were filtered to collect nutrient deficient or spent media. This spent media was then mixed with fresh YCFA media in a series of ratios decreasing tenfold from 100% to 10% spent media. Bacteria were incubated in these mixes overnight with 0.5, 1.0, or 2.0µg of plasmid DNA. The cultures were plated on selection YCFA plates and growth observed.

Bacterial cells were harvested after 2hr, 6hr, and overnight incubations for the creation of competent cells. Electrocompetent cells were created by washing cells in ice cold 10% glycerol as described by NEB (NEB protocols, [*Accessed Apr 2017*]). Cells were resuspended in either 10% glycerol or SMP buffer (sucrose-magnesium-phosphate: 270 mM sucrose, 1 mM MgCl2, and 5 mM sodium phosphate, pH 6.5), then aliquoted and stored at -80ºC. Protoplasts were created as described by Rattanachaikunsopon and Phumkhachorn (2009). Cells were incubated in digestion buffer (50 mg/L mutanolysin, 0.5 M sorbitol, 0.01 M Tris-hydrochloride, pH 7.0) at 37 ºC for 1hr, washed and

resuspended in transformation buffer (0.5 M sorbitol, 0.02 M maleate, 0.02 M MgCl2, pH 6.5), then aliquoted and stored at -80ºC (Rattanachaikunsopon P. and Phumkhachorn P., 2009)

Electroporation was carried out in 2mm gap cuvettes (Flowgen Bioscience), using a MicroPulser Electroporation System (Bio-Rad). Aliquots of competent cells were thawed on ice, and incubated with 0.5, 1.0, or 2.0µg of plasmid DNA on ice for 20 min. Cells were transferred to a chilled cuvette and pulsed at 25µF and 1.2, 1.5, or 2.5 kV. Glass bead transformation was carried out using acid washed glass beads (Sigma-150-212µm) as described by Rattanachaikunsopon and Phumkhachorn. Protoplasts were thawed and added to the transformation mix (15% w/v Polyethylene glycol (PEG) 6000, 0.3% w/v acid washed glass beads); and the mix was agitated by vortex at the highest speed for 15 seconds (Rattanachaikunsopon P. and Phumkhachorn P., 2009). 1ml of 1% sucrose YCFA was added to the tube immediately after agitation, samples recovered for 3 hours at 37ºC anaerobically, and were plated on selection and control YCFA plates, as detailed below.

	YCFA	**YCFA +Tetracyline**
Electroporated +DNA	Positive	Positive
Electroporated	Positive	Negative
Thawed Competent Cells	Positive	Negative
Agitated +DNA	Positive	Positive
Agitated	Positive	Negative
Digested	Positive	Negative

2.9 Conjugation

Donor and recipient strains were cultured in YCFA broth anaerobically overnight. Culture OD readings were taken using the WPA Biowave (Biochrom) and the strains mixed in a 4:1 ratio, by volume, of donor to recipient. The culture mixes were immobilised onto Mixed Cellulose Ester Membranes (Millipore- 0.45µm pore size) using sterile syringes (BD Plastipak) and filter holders (Millipore). The membrane was placed onto YCFA plates and incubated anaerobically overnight. The membranes were then washed in 1ml of PBS and the bacterial suspension was plated onto selection and control YCFA plates, as detailed below. Conjugations using an E. coli mating pair were done in parallel as a technical control

	Nalidixic Acid	**Tetracycline**	**Nalidixic Acid + Teracycline**
Donor (D)	Negative	Positive	Negative
Recipient (R)	Positive	Negative	Negative
D+R Mix	Positive	Positive	Positive

	Erythromycin	**Tetracycline**	**Erythromycin + Teracycline**
Donor (D)	Negative	Positive	Negative
Recipient (R)	Positive	Negative	Negative
D+R Mix	Positive	Positive	Positive

	YCFA Anaerobic	**YCFA Aerobic**	**YCFA+Tetracycline Anaerobic**	**YCFA+Tetracycline Aerobic**
Donor (D)	Positive	Negative	Positive	Negative
Recipient (R)	Positive	Positive	Negative	Negative
D+R Mix	Positive	Positive	Positive	Positive