Heterologous gene expression in the human gut bacteria *Eubacterium rectale* and *Roseburia inulinivorans* by means of conjugative plasmids.

Paul O. Sheridan[1,2], Jennifer C. Martin[1], Nigel P. Minton[3], Harry J. Flint[1], Paul W. O'Toole[2], Karen P. Scott[1,#]

[1] Rowett Institute of Nutrition and Health, University of Aberdeen, Foresterhill, Aberdeen AB25 2ZD, UK

[2]School of Microbiology & APC Microbiome Ireland, University College Cork, Cork, Ireland

[3] Centre for Biomolecular Sciences, University Park, Nottingham, NG7 2RD, UK

[#]Corresponding author: Karen P. Scott, k.scott@abdn.ac.uk

Running title: Expression vectors for *Roseburia* and *Eubacterium rectale*

Abstract.

Commensal butyrate-producing bacteria in the Firmicutes phylum are abundant in the human intestine and are important for maintaining health. However, understanding of the metabolism and host interaction of these bacteria is limited by the lack of genetic modification techniques. Here we establish a protocol enabling the transfer of autonomously-replicating shuttle vectors by conjugative plasmid transfer from an *Escherichia coli* donor into representatives of an important sub-group of strictly anaerobic human colonic Firmicutes. Five different plasmid shuttle vectors were tested, each carrying a different origin of replication from Gram-positive bacteria. Plasmid pMTL83151 (pCB102 replicon) were successfully transferred into two strains of *Eubacterium rectale*, while pMTL83151 and pMTL82151 (pBP1 replicon) were transferred into *Roseburia inulinivorans* A2-194. Plasmids that carried a

Streptococcus bovis JB1 glycoside hydrolase family 16 β-(1,3-1,4)-glucanase gene were constructed and conjugated into *Roseburia inulinivorans* A2-194 and *Eubacterium rectale* T1-815, resulting in successful heterologous expression of this introduced enzymatic activity in these two strains of butyrate-producing Firmicutes.

Introduction

The human intestinal microbiota produces short chain fatty acids (SCFA) as end products of anaerobic fermentation. These SCFA are involved in a wide variety of health promoting functions. It has been suggested that intestinal bacteria that produce butyrate as their main fermentation end product have potential as novel, health-promoting probiotics [1, 2]. This prediction seems reasonable, considering several observations. Firstly, butyrate reinforces the colonic defence barrier by stimulating tight junction formation [3], antimicrobial secretion [4], and mucin synthesis [5]. Secondly, butyrate regulates macrophage and dendritic cell differentiation, maturation and function in a manner that promotes tolerance to the intestinal microbiota [6, 7]. These properties make butyrate, or butyrate-producing bacteria, of potential interest for treatment of conditions such as ulcerative colitis (UC) and Crohn's disease, which are characterised by colonic barrier damage and inflammation [8].

Species of the genera *Roseburia, Eubacterium* and *Faecalibacterium* are the most abundant butyrate producing bacteria in the human colonic microbiota [9]. Genome sequences of multiple members of these genera are now available, but the lack of genetic modification techniques has limited our ability to determine the functions of specific genes.

Heterologous gene expression systems have been extensively used in molecular microbiology to determine the impact of environmental stimuli on individual steps within biochemical pathways and also for the expression of a variety of proteins of commercial and

biomedical interest [10]. Although several well studied bacterial and *in vitro* (or cell-free) expression systems are available, these are often insufficient for expression and characterisation of specific proteins [11, 12]. Furthermore, the expression of certain proteins can be toxic in some bacterial hosts, while use of the wrong expression host can result in low or no expression, and/or unfolded or misfolded proteins [13]. It would therefore be advantageous to create expression vectors for key members of the human gut microbiota, to facilitate investigating the function of the vast array of candidate host interaction factors and metabolic enzymes in these bacteria that have been uncovered by recent microbiome studies [14].

In previous work, the conjugative transposons Tn*1545* from *Eubacterium cellulosolvens* and Tn*K10* from *Clostridium saccharolyticum* K10 were transferred into *R. inulinivorans* A2-194 [15], demonstrating conjugation as a means of introducing exogenous DNA into these species. However, in order to be maintained in the progeny of a recipient, a transposon must both transfer into the recipient strain and insert into the recipient chromosome. Both of these steps are low frequency occurrences whose combined probability is the conjugation frequency multiplied by the frequency of chromosomal insertion. The ability of autonomously-replicating plasmid vectors to stably replicate in the recipient greatly mitigates this "bottlenecking effect".

Although no autonomously replicating plasmids have yet been isolated from any member of *Roseburia*/*E. rectale* group, a series of modular plasmids have recently been developed for members of the *Clostridium* genus. These shuttle plasmids/vectors are each composed of four modular sections, with several variations available and can replicate autonomously in a range of clostridial species [16]. Since they replicate in *E. coli*, they can be readily isolated and manipulated, and are then transferable via *E. coli*-*Clostridium* mating.

Our aim here was to establish protocols by which a shuttle vector could be transferred by conjugation from a donor strain into the bacterium of interest, leading to stable plasmid maintenance. Firstly, relevant indigenous antibiotic resistance genes and restriction modification systems (RMS) were identified in candidate butyrate-producing bacteria, using both *in silico* and *in vitro* methods. Protocols for the conjugative transfer of an autonomously-replicating plasmid vector into strains of biological interest were then established. Optimisation of the protocols permitted conjugative transfer of plasmid vectors into *Roseburia inulinivorans* A2-194 and *Eubacterium rectale* T1-815 and enabled the heterologous expression of a β-(1,3-1,4)-glucanase gene from *Streptococcus bovis* JB1 in these dominant butyrate-producing bacteria. The β-(1,3-1,4)-glucanase gene was chosen as this readily demonstrable enzymatic activity is not naturally found in the target bacteria.

Materials and methods

Bacterial strains, plasmids, primers and growth conditions

The strains and plasmids used in this study are described in Table 1. All primers used in this study are listed in Table S1. Anaerobic strains were cultured in the anaerobic media M2GSC [17], YCFAGSC [18] or AMM (anaerobic mating medium, this work). AMM consisted of (per 100 ml) Casitone (1.0 g), yeast extract (0.25 g), $NaHCO_3$ (0.4 g), cysteine (0.1 g), K_2HPO_4 (0.045 g), KH_2PO_4 (0.045 g), NaCl (0.09 g), $(NH_4)_2SO_4$ (0.09 g), $MgSO_4$ $7H_2O$ (0.009 g), $CaCl_2$ (0.009 g), resazurin (0.1 mg), hemin (1 mg), biotin (1 μg), cobalamin (1 μg), p-aminobenzoic acid (3 μg), folic acid (5 μg), pyridoxamine (15 μg) and acetate (5 mM). Broths of these media were divided into 7.5 ml aliquots in Hungate tubes, sealed with butyl rubber septa (Bellco Glass) and agars (with the addition of 2% agar) were divided into 100ml aliquots in Wheaton bottles, with all dispensing carried out under anaerobic conditions using 100% CO_2. *Roseburia* species, *E. rectale* and *Faecalibacterium prausnitzii* cultures were

inoculated using the anaerobic methods described by Bryant, 1972 [19] and incubated anaerobically without agitation at 37 °C. Agar plate work was carried out in Concept Plus Anaerobic Workstation, Ruskinn Technology with a gas mix of CO_2: N_2: H_2, in a ratio of 7:11:2 respectively. *Escherichia coli* strains were cultured aerobically in L-broth and on L-agar at 37 °C, including, where appropriate, 10 μg/ml chloramphenicol.

***In silico* analysis of restriction modification systems (RMS)**

Bioinformatic annotation of the RMS of selected species required the merging of lists of putative restriction-associated proteins from two databases: NCBI and REBASE (http://www.ncbi.nlm.nih.gov and http://rebase.neb.com/rebase/rebase.ftp.html). The physical locations of the sequences encoding the putative proteins in the genomes of *R. intestinalis* L1-82, *R. inulinivorans* A2-194, *E. rectale* A1-86 and *F. prausnitzii* A2-165 were determined and these genomic regions were subject to manual curation in Artemis [20]. The most likely candidates for RMS were then determined based on co-localisation of predicted restriction endonucleases and methylase encoding genes

Preparation of protein extracts for restriction analysis

Cultures of selected butyrate-producing strains were grown in 7.5 ml of M2GSC broth for 48 h at 37 °C. These cultures were transferred to 15 ml plastic tubes and centrifuged (6000 *g*, 10 min, 4 °C). The supernatants were decanted and the pellets were resuspended in 5 ml Sonication Buffer (Tris-HCl (0.315 g), β-mercaptoethanol (0.195 ml) and distilled water (up to 100 ml), pH 8). The solutions were then sonicated with the Sanyo, Soniprep 150 in six 30 sec bursts (22μm amplitude) – cooling for 30 sec on ice between bursts. Streptomycin sulphate was added to the solutions to a final concentration of 2% as a means of removing nucleic acids. These solutions were incubated on ice for 30 min and centrifuged (11,000 *g*, 10 min, 4 °C).

The resulting supernatants were transferred to new 15 ml tubes and polyethylene glycol (PEG) 6000 was added to a final concentration of 10%. These mixes were incubated on ice for 30 min and centrifuged (11,000 *g*, 10 min, 4 °C). The resulting pellets were resuspended in 1 ml of phosphate buffer solution (pH 7) and transferred to 1.5 ml tubes, and PEG 6000 was again added to a final concentration of 10% and incubated on ice for 30 min. These were centrifuged (11,000 *g*, 10 min, 4 °C) and the final pellets were resuspended in 30 µl of phosphate buffer solution. These protein extracts were frozen at -20 °C. All steps after culturing were performed on ice when possible.

During restriction analysis, protein extracts (2 µl) were incubated with purified plasmid DNA (200 ng) in 15 µl of water and NEBuffer 3.1 (NEB) for 90 min at 37 °C prior to separation by agarose gel electrophoresis. This buffer (NEBuffer 3.1) was chosen because in prior tests it facilitated lambda DNA degradation in a similar efficient way with each of the protein extracts.

Development of selective plating conditions for mating experiments

Bacterial strains were tested for resistance to the antibiotics chloramphenicol, thiamphenicol (an analogue of chloramphenicol), erythromycin, tetracycline and rifampicin. Bacteria were spread on YCFAGSC agar plates supplemented with different concentrations of each antibiotic (5, 10 and 25 µg/ml) and growth monitored after 120 h of anaerobic incubation at 37 °C.

Overnight cultures of all *Roseburia* and *Eubacterium rectale* strains listed in Table 1, and *E. coli* CA434 (donor strain) were streaked on duplicate, well-dried M2GSC and YCFAGSC agar plates and incubated either anaerobically or aerobically for 48 h at 37 °C. Additional tests with YCFAGSC containing a reduced SCFA concentration (Table 3) were included.

Conjugation protocol optimised for the *Roseburia inulinivorans* and *Eubacterium rectale*

The *E. coli* donor strain CA434 was first transformed with the relevant plasmid by electroporation using standard procedures [21]. The resulting transformant was grown overnight in 40 ml of LB supplemented with 10 μg/ml of chloramphenicol. The overnight culture was then centrifuged (1200 *g*, 10 min, 20 °C), the supernatant decanted and the pellet gently resuspended in 20 ml of phosphate-buffered saline (PBS, pH 7.4). These washed cells were again centrifuged (1200 *g*, 10 min, 20 °C) and transferred into the anaerobic workstation. The supernatant was decanted and the pellet was gently resuspended in 1 ml of an overnight culture of the recipient strain, grown in M2GSC. 100 μl of this mix was then dotted onto an agar plate of AMM or M2GSC and incubated anaerobically for 48 h. The resulting bacterial growth was scraped off the plate and resuspended in 500 μl of anaerobic-PBS (reduced by boiling, followed by addition of L-cysteine and bubbling with CO_2 gas). 50 μl of this cell suspension was spread onto YCFAGSC agar plates. The high short-chain fatty concentration present in YCFAGCS prohibits the growth of *E. coli* (this work and [22]), and the addition of chloramphenicol (5 or 7.5 μg/ml) selected for plasmid uptake by the recipients. Colonies that grew were purified as single colonies following streaking on fresh YCFAGSC plates supplemented with chloramphenicol.

Verification of putative transconjugants

Aerobic growth experiments involved streaking of putative transconjugants onto M2GSC agar plates and incubating aerobically at 37 °C for 48 h. Since the obligately anaerobic recipient strains cannot grow in the presence of oxygen, any aerobic growth was attributed to persistent *E. coli* cells. Stocks were made of all cultures able to grow anaerobically and not aerobically. The 16S rRNA gene was amplified directly from bacterial pellets by polymerase chain reaction (PCR) using primers FD1 and RP2 (Table S1). The resulting amplicon was then purified using the Wizard SV Gel and PCR Clean-Up System, following the manufacturer's

instructions, and sequenced using 519R and 926F primers. Sequence quality was checked manually using Chromas Lite software, and the bacterial identity confirmed by BLASTn, querying each sequence against the NCBI 16S rRNA gene database. An alignment of the different sequences is shown (Fig. S1). The presence of the plasmid in putative transconjugants was confirmed by amplifying a nucleotide sequence common to all of the modular plasmids but absent in the recipient's chromosome, using the primers PS#MTL-for and PS#MTL-rev (Table S1). Plasmids were shown to be autonomously-replicating in transconjugants by Southern blotting. Genomic DNA of putative transconjugants was digested with the restriction enzyme HindIII prior to Southern blotting. Southern blotting was performed with DIG High Primer DNA Labelling and Detection Starter Kit II (Roche Diagnostics), following the manufacturer's instructions, using a probe that was specific to a region common to all of the plasmids, but not present in the recipient chromosome.

Coomassie staining of the total protein complement involved centrifuging 7.5 ml overnight cultures (800 *g*, 10 min, 20 °C). The resulting pellets were resuspended in 2 ml of 50 mM sodium phosphate buffer and centrifuged (800 *g*, 10 min, 20 °C). These pellets were resuspended in 200 µl of 50 mM sodium phosphate buffer. 6 µl of these solutions were added to each well of a 10% SDS-PAGE gel, and visualised by staining with Coomassie blue stain followed by destaining with "destain solution" (45.4% methanol, 9.2% acetic acid, 45.4% water).

Plasmid stability determination

Plasmid stability was determined by continuously subculturing transconjugants anaerobically in M2GSC broth lacking antibiotic, as described previously [23]. Briefly, overnight cultures of two transconjugants, isolated from independent conjugation experiments, were diluted to OD_{650} 0.1 in fresh M2GCS broth (lacking antibiotic). The broths were incubated

for 12 h, and used either to inoculate a fresh broth to an OD_{650} 0.1 or to create serial dilutions in reduced PBS (10^{-1} – 10^{-8}). The inoculated fresh broth was incubated for 12 h and either re-inoculated or a dilution series made This was repeated four times. At each stage, the PBS dilutions were plated (50 µl) onto YCFAGSC plates within 1 h of dilution and the plates were incubated for 48 h in the anaerobic workstation. Colonies that grew on these plates were picked in duplicate onto YCFAGSC plates with or without antibiotic. The proportion of bacteria still harbouring the plasmid was calculated by dividing the number of colonies that grew on the antibiotic plates by the cell numbers on the plates lacking antibiotic. All cultures steps were carried out anaerobically.

Growth curves of strains in M2GSC broth lacking antibiotics were used to estimate the number of generation times the bacteria had gone through in each 12 h growth period. Instability was calculated as percentage plasmid loss per generation in the absence of antibiotic selection, using the formula $x = 1 - R^{1/N}$, where x = segregational instability, R = fraction of bacteria still possessing plasmid and N = number of generations.

Construction of the pMTL3β-glu expression vector

An amplicon containing the β-(1,3-1,4)-glucanase gene (with its native promotor) from pL1Hc (Ekinci *et al.* 1997) was generated using the primers PS#Bglu-BamHI-for and PS#Bglu-HindIII-rev, which possess 5'-end restriction sites for BamHI and HindIII, respectively. The PCR was performed in a 50 µl of reaction mix: Bioline Taq Polymerase and buffer, with 2.5 mM $MgCl_2$ and 200 µM of each primer. The PCR involved a hot start (94 °C, 5 min), followed by 28 cycles of denaturation (95 °C, 30 s), annealing (52 °C, 30 s) and elongation (72 °C, 2 min), and followed by a final elongation step (72 °C, 8 min). The product of this reaction was purified from a 1% agarose gel using the QIAquick Gel Extraction Kit, following manufacturer's instructions. The amplicon and the plasmid pMTL83151 were double-digested

separately in 30 µl reactions containing Promega Buffer E and 10 ng each of Promega BamHI and HindIII restriction enzymes, incubated overnight at 37 °C. The double-digest of the plasmid was then dephosphorylated with calf intestinal alkaline phosphatase (CIAP) for 30 min at room temperature, to prevent self-ligation. Both digests were then cleaned using Wizard SV Gel and PCR Clean-Up System and the products were ligated with Promega T4 Ligase and buffer, incubating overnight at 4 °C. The ligation product was then transformed into XL1-Blue Competent Cells (Stratagene), following manufacturer's instructions, and transformants selected based on chloramphenicol resistance. Cloning was confirmed by restriction analysis. The plasmids containing the insert, designated pMTL3β-glu, were purified using QIAprep Spin Miniprep Kit.

Testing functional activity of a heterologously expressed protein

The pMTL3β-glu expression vector was electroporated into *E. coli* donor strain CA434 using standard procedures [21]. It was then transferred from *E. coli* CA434 into *R. inulinivorans* A2-194 and *E. rectale* T1-815 following the optimised conjugation protocol, described above. Transconjugants were streaked on to fresh M2GSC agar plates and incubated anaerobically at 37 °C for 24 h. An overlay solution was prepared anaerobically using 0.1 % Glucagel from PolyCell Technologies (Glucagel contains 78.2 % β-glucan) and 0.8 % agarose in 50mM sodium phosphate buffer (pH 7) and divided into 4 ml aliquots in Hungate tubes. These were then poured over the pre-grown transconjugant-containing agar plates and again incubated overnight anaerobically. Finally the plates were stained with Congo red (1 mg/ml) for 30 min. The Congo red was then decanted and the plates destained using 1M NaCl for 30 min before observing on a light box to visualise clear zone formation.

Quantitative enzymatic assay

Overnight cultures (7 ml in YCFAGSC) were centrifuged (1200 *g*, 10 min, 4 °C) and the resulting pellets were washed twice in 3 ml PBS (pH7.4) and re-centrifuged, before resuspending in 400 µl of 50 mM sodium phosphate buffer (pH 6.5) containing 2mM dithiothreitol (DTT). They were then sonicated in an ice bath with 4 x 30 second bursts or until lysis was visible. Sonicated extracts (25 µl) were incubated aerobically for 2 h at 37 °C in 50 mM sodium phosphate buffer (pH 6.5) containing 2 mM DTT with 1% Glucagel as substrate and the protein concentration in each sample was measured using the method of Lowry. Enzyme activity was determined by measuring the release of reducing sugars [24], as described previously [25]. One unit of enzyme activity is equivalent to the release of 1 µmol glucose min^{-1} $(mg\ protein)^{-1}$.

Results

Antibiotic resistance profiles of candidate recipient strains

All the potential recipient bacteria were susceptible tochloramphenicol, thiamphenicol and rifampicin at the tested concentrations. *R. intestinalis* L1-82 grew in the presence of 10 µg/ml tetracycline and *R. inulinivorans* A2-194 was resistant to erythromycin (Table 2). Bioinformatic analysis of the *R. inulinivorans* A2-194 genome identified a putative macrolide-specific ABC-type efflux carrier (GenBank: CRL37109.1) that may be responsible for the erythromycin resistance phenotype of the strain, and analysis of the *R. intestinalis* L1-82 genome identified chromosomally adjacent putative *tet*O and *tet*(40) genes (WP_006858004 and WP_044999308, respectively). The putative *tet*(40) product is 100 % identical to the experimentally validated *tet*(40) protein from *Clostridium cf. saccharolyticum* K10 (CBK76340.1) [26], making it a likely contributor to tetracycline resistance in *R. intestinalis* L1-82. A rifampicin resistant *R. inulinivorans* A2-194 mutant strain ($A2\text{-}194^{R}$) was obtained

by selecting for the generation of spontaneous mutations that conferred rifampicin resistance (this work).

Chloramphenicol resistance was thus chosen as the preferred marker incorporated in the shuttle vector to select transconjugants as none of the wild type strains could grow on YCFAGSC supplemented with 5 µg/ml of this antibiotic (Table 2).

Developing optimal selective conditions for mating experiments

Different media were tested to distinguish the growth of the donor *E. coli* CA434, and recipient *Roseburia* and *Eubacterium rectale* strains. All of the strains grew to large (> 1 mm) colonies on anaerobic M2GSC incubated anaerobically, while only *E. coli* CA434 grew when incubated aerobically (Table 3). *E. coli* CA434 did not grow on YCFAGSC incubated anaerobically (even after 120 h), but could grow aerobically or if the SCFA concentration was reduced (Table 3). This is assumed to reflect inhibition of anaerobic *E. coli* growth by the SCFA present in the medium [22]. While M2GSC medium was suitable for matings between *E. coli* and the recipient strains, the rumen fluid component can be difficult for many labs to source and has batch-to-batch variations. Therefore, a defined mating medium, AMM, was designed in the course of this work (see Materials & Methods). Anaerobic YCFAGSC supplemented with chloramphenicol was used as the selective medium for transconjugants as it prevented growth of the *E. coli* donor strain and of the wild-type recipient strains, but permitted growth of chloramphenicol resistant transconjugants.

Conjugative transfer of autonomously-replicating plasmid to *E. rectale* A1-86

Plasmid screening of the strains listed in Table 1 did not reveal any small endogenous plasmids that might be developed as vectors, so we decided to consider existing Gram-positive vectors. The plasmid pMTL960 was initially chosen as candidate shuttle vector as it possesses

the origin of replication of the plasmid pCD6 from *Clostridium difficile* – a species belonging to the same order (Clostridiales) as the *Roseburia* genus. This plasmid can be conjugated from *E. coli* into a range of different *Clostridium* species and replicates autonomously within the recipient [23]. However, no transconjugants were obtained following matings between an *E. coli* donor harbouring pMTL960 and *R. inulinivorans* A2-194. *In silico* analysis of the genome sequences of *R. inulinivorans* A2-194 (ACFY01000000), *R. intestinalis* L1-82 (ABYJ00000000.2), *E. rectale* A1-86 (NC_021010.1) and *F. prausnitzii* A2-165 (NZ_ACOP00000000.2) revealed that they possessed a variety of predicted restriction modification systems (RMS) that may hinder the uptake of exogenous DNA (Table 4; Table S3). Therefore, an improved assay for *in vitro* restriction activity (based on [27] was established for the restriction analysis of *E.coli-Clostridium* plasmid shuttle vectors [16, 23]. Plasmids pMTL82151, pMTL83151, pMTL84151 and pMTL85151, which are identical to each other except for their Gram-positive origin of replication (Table 1), were incubated with protein extracts from representative putative recipients. The plasmids were apparently not restricted by *E. rectale* A1-86 (Fig. 1, lanes A), as the smallest visible band is the same size as that of the undigested plasmid. Faint unrestricted bands are also present for *R. inulinivorans* A2-194 incubations with plasmids pMTL84151 and pMTL83151 while pMTL82151 appears to be more degraded (Fig. 1 lanes E). The remaining combinations of plasmids and bacterial protein extracts indicate restriction activity, evidenced by the appearance of specific smaller bands or the complete disappearance of distinct plasmid bands on the agarose gels.

Following conjugation, chloramphenicol resistant transconjugants of *E. rectale* A1-86 were obtained containing the pMTL83151 plasmid. These transconjugants (coded as follows: EAM3- ***E***. *rectale* **A**1-86 harbouring p**M**TL8**3**151) were confirmed to be derived from *E. rectale* A1-86 by aerobic growth tests (*E. rectale* A1-86 does not grow aerobically), Gram-staining, 16S rRNA gene sequencing and SDS PAGE analysis (Fig. 2A). Specific PCR primers

amplified plasmid sequences only in transconjugants (Fig. 2B, Fig. S2). Transferred plasmids were shown to be autonomously-replicating rather than chromosomally integrated in *E. rectale* A1-86 (Fig. 2C) by Southern blotting.

Extending the mating protocol to additional strains

The mating protocol was optimised to achieve a transfer frequency of 1.8 x 10^{-6} transconjugants per potential *E. rectale* A1-86 recipient. Specifically, increasing the mating time from 24 h to 48 h increased transfer frequency 10-fold and changing the donor: recipient ratio from 5:1 to 40:1 increased transfer frequency 5-fold. This optimised protocol was then used in attempts to conjugate pMTL82151, pMTL83151, pMTL84151 and pMTL85151 into the range of strictly anaerobic bacterial strains indicated in Table 1. Putative transconjugants were obtained for transfers of p**M**TL8**3**151 into ***E***. *rectale* **T**1-815 (named ETM3; transfer frequency of 2.3 x 10^{-7} per potential recipient), and p**M**TL8**2**151 and p**M**TL8**3**151 into ***R***. *inulinivorans* **A**2-194 (transfer frequency of 6.1 x 10^{-8} (RAM2) and 1.33 x 10^{-6} (RAM3) per potential recipient). The identity of these transconjugants was confirmed by Gram-staining, 16S rRNA gene sequencing, testing for aerotolerance and by PCR with primers specific for the plasmid (Figure S2). The plasmids were shown to be autonomously-replicating by Southern blotting (Fig. 2C and 2D, Fig. S3), because plasmid DNA extracted from transconjugants co-migrated with purified plasmids from *E. coli*.

The use of a nitrocellulose filter increased the transfer efficiency of pMTL83151 into *R. inulinivorans* A2-194 when a 5:1 donor: recipient ratio was used (5-fold), but did not improve transfer efficiency when a 40:1 donor: recipient ratio was used. Additionally, mating on a nitrocellulose filter appeared to completely prevent conjugation of pMTL83151 into *E. rectale* T1-815. Thus, nitrocellulose filter mating was not used in the optimised mating protocol, but rather the donor/recipient cell mixture was spotted directly onto the centre of an

agar plate. Furthermore, mating on M2GSC and AMM resulted in similar transfer frequencies (for conjugation of pMTL83151 into *E. rectale* T1-815 and *R. inulinivorans* A2-194), meaning that they could be used interchangeably in the mating protocol. Compared to YCFA, AMM possesses a lower concentration of acetate (5 mM), a short chain fatty acid that has been shown to hinder the growth of *E. coli* in pure culture [28] and thus might hinder the growth of the *E. coli* donor during mating. The optimized mating protocol is shown schematically in Fig. 3.

Determining plasmid stability of pMTL83151 in *E. rectale* A1-86

The maintenance of pMTL83151 in *E. rectale* A1-86 in the absence of chloramphenicol selection pressure was calculated. Two *E. rectale* A1-86 transconjugants, EAM3(1) and EAM3(2 were grown in sequential sub-cultures, re-inoculating every 12 h (~ 8 generations). The proportion of bacteria still chloramphenicol resistant, and thus still harbouring pMTL83151, was calculated at each sub-inoculation point and in the final culture after five sub-inoculations (~40 generations) (Table S2). The first sub-inoculation was excluded from stability calculations as chloramphenicol was transferred into this culture from the starting culture during sub-inoculation, whereas chloramphenicol concentrations in subsequent sub-inoculations were deemed too low to select for plasmid persistence. .

The rate of plasmid loss was calculated to be 2% and 5% plasmid loss per generation in the absence of antibiotic selection for transconjugants EAM3(1) and EAM3(2), respectively. In practice, this means that for transconjugants grown from lag phase (OD_{650} 0.1) to stationary phase (OD_{650} 1.0) in a Hungate tube culture (7.5 ml) in the absence of antibiotic selection, over 65 % of the stationary phase bacteria are predicted to retain pMTL83151.

Heterologous gene expression

The extracellular β-(1,3-1,4)-glucanase from *S. bovis* JB1 can be expressed from shuttle vectors in a diverse range of bacteria (*Lactococcus lactis* IL2661, *Enterococcus faecalis* JH2-SS and *E. coli* DH5α) [29]. This expression, which is readily detected by Congo red staining of agar plates containing β-glucan, was controlled by the native *S. bovis* promoter of the β-(1,3-1,4)-glucanase gene and enabled the recipient bacteria to hydrolyse barley β-glucan. Since the genomes of *R. inulinivorans* A2-194 and *E. rectale* T1-815 do not encode a GH16 β-(1,3-1,4) glucanase (29), we chose this enzyme for a proof-of-principle heterologous gene expression test.

The shuttle plasmid pL1Hc [29] had been constructed by cloning a *S. bovis* JB1 chromosomal fragment isolated by endonuclease restriction of the genomic DNA into the plasmid pUC18. pL1Hc therefore contains flanking DNA likely to be non-essential to gene function, including a putative stress response gene. For a more targeted approach, PCR primers containing restriction sites were designed to amplify only the β-(1,3-1,4)-glucanase gene and its regulatory regions from pL1Hc. The resulting amplicon was cloned into pMTL83151, producing the plasmid pMTL3β-glu (Fig. 4). The purified plasmid was then electroporated into *E. coli* strain CA434 and conjugated into *R. inulinivorans* A2-194 and *E. rectale* T1-815.

Activity of the heterologously expressed β-(1,3-1,4)-glucanase in the recombinant strains was assessed by clear zone formation on agar plates overlaid with barley β-(1,3-1,4)-glucan. In *E. coli* CA434, the strain lacking pMTL3β-glu produced no clear zones around single colonies, whereas distinct clear zones were observed around single colonies of the recombinant strain (Fig. S4). Similarly, for *R. inulinivorans* A2-194 and *E. rectale* T1-815, the wild-type strains produced no clear zones around single colonies, whereas distinct clear zones were observed around single colonies of the recombinant strains RABglu and ETBglu (Fig. 5a).

Quantification of β-glucanase activity in the protein extracts by reducing sugar assay revealed that *R. inulinivorans* harbouring pMTL3β-glu (RABglu) possessed twice the β-glucanase activity of the wild-type strain A2-194. *E. rectale* T1-815 harbouring the pMTL3β-glu (ETBglu) possessed over 4-times the β-glucanase activity of the wild-type strain T1-815 (Fig.5b). It is possible that the protein is exported from *R. inulinivorans* cells more efficiently than from *E. rectale*, explaining the activity differences observed in clear zone formation compared to protein extracts.

Discussion:

The importance of butyrate-producing bacteria in maintaining intestinal health is now widely recognised. In recent years, our understanding of the way these bacteria have adapted to the human intestine by utilising dietary and host derived polysaccharides for energy [30-34] and by modulating host immunity via flagella [35] has increased. However, a more complete understanding of the interactions between these bacteria, and with the human host has been limited by a lack of techniques for gene modification.

The plasmid pMTL83151, which possesses the replication region from pCB102 (from *C. butyricum*), was capable of autonomous replication in *E. rectale* A1-86, *E. rectale* T1-815 and *R. inulinivorans* A2-194 while pMTL82151 was capable of replicating in *R. inulinivorans* A2-194. The latter plasmid possesses the origin of replication from pBP1, isolated from *C. botulinum*. The transfer frequency of this plasmid into *R. inulinivorans* was 100-1000–fold lower than that of pMTL83151. However, it appears to have a higher copy number, as a plasmid was visible by gel electrophoresis of isolated DNA from *R. inulinivorans* A2-194 possessing pMTL82151, but not pMTL83151 (data not shown). As all the pMTL80000 series plasmids were identical apart from the Gram-positive origins of replication, it is likely that our inability to introduce by conjugation the plasmids pMTL84151 and pMTL85151 into *R. inulinivorans*

A2-194, *E. rectale* A1-86 and *E. rectale* T1-815 was due to their inability to replicate in the recipient cell. It is also possible that recipient restriction enzymes may have inactivated the plasmid origins of replication.

Expression of the active *S. bovis* β-glucanase in *E. rectale* and *R. inulinivorans* resulted in activity of the β-glucanase enzyme against β-glucan, although it did not enable the strains to grow utilising β-glucan as a sole source of energy (data not shown). This suggests that although the transconjugants were able to degrade β-glucan, these strains are not equipped to import β-glucan or any resulting degradation products and use them for energy. However, β-glucanase expression may offer interesting possibilities in the future as a reporter gene for the analysis of promoter activity in these bacteria.

In this work, we have taken the first steps towards genetic analysis of the butyrate-producing species *Roseburia inulinivorans* and *Eubacterium rectale* which comprise at least 7% of the human intestinal microbiota [36]. The natural progression of this work will involve using these genetic manipulation tools to interrupt chromosomal genes to establish their functionality, or introducing new genes conferring novel abilities on the host bacterium. It will be possible to include different selectable marker genes or alternative origins of replication from native plasmids within shuttle vectors to expand the use of these genetic manipulation tools to other related Firmicute bacteria. *R. intestinalis* L1-82 contains tetracycline resistance genes conferring resistance to 10µg/ml tetracycline (this work). Introducing this gene onto a multicopy plasmid would provide an alternative selectable marker. The development of suicide vectors that cannot themselves replicate in these bacteria, but which contain selectable markers that can be used to drive homologous recombination and chromosomal integration of homologous and heterologous genes, or for the insertional inactivation of existing genes, are crucial to investigating gene function in the mixed ecosystem.

This work represents a crucial first step towards future studies analysing gene expression, regulation, function and microbe-host interactions in this important, but little studied, group of human colonic anaerobic bacteria.

Acknowledgements:

The Rowett Institute (University of Aberdeen) receives financial support from the Scottish Government Rural and Environmental Sciences and Analytical Services (RESAS). POS was a PhD student supported by the Scottish Government (RESAS) and the Science Foundation Ireland, through a centre award (12/RC/2273) to APC Microbiome Ireland, Cork, Ireland.

Tables.

Table 1. Bacterial strains and plasmids.

Strain/plasmid	Relative characteristics	Source/Reference
Strains		
***Eubacterium rectale* A1-86**	Butyrate producing, strict anaerobes of the *Lachnospiraceae* family**.**	(Barcenilla et al. 2000)
***Eubacterium rectale* T1-815**		(Barcenilla et al. 2000)
***Eubacterium rectale* M104/1**		(Louis et al. 2004)
***Eubacterium rectale* L2-21**		(Barcenilla et al. 2000)
***Roseburia inulinivorans* A2-194**		(Duncan et al. 2006)
***Roseburia inulinivorans* A2-194 Rif^R**		This work
***Roseburia inulinivorans* L1-83**		(Barcenilla et al. 2000)
***Roseburia faecis* M72/1**		(Duncan et al. 2006)
***Roseburia intestinalis* L1-82**		(Duncan et al. 2006)
***Roseburia hominis* A2-183**		(Duncan et al. 2006)
***Eubacterium rectale* EAM3**	*E. rectale* A1-86 harbouring pMTL83151	This work
***Eubacterium rectale* ETM3**	*E. rectale* T1-815 harbouring pMTL83151	This work
***Eubacterium rectale* ETBglu**	*E. rectale* T1-815 harbouring pMTL3β-glu	This work
***Roseburia inulinivorans* RAM2**	*R. inulinivorans* A2-194 harbouring pMTL82151	This work
***Roseburia inulinivorans* RAM3**	*R. inulinivorans* A2-194 harbouring pMTL83151	This work
***Roseburia inulinivorans* RABglu**	*R. inulinivorans* A2-194 harbouring pMTL3β-glu	This work
***Faecalibacterium prausnitzii* A2-165**	Non-motile, butyrate producing, strict anaerobes of the *Ruminococcaceae* family.	(Duncan et al. 2002)

***Escherichia coli* CA434**	Conjugative donor. Genotype: HB101 (*thi-1* hsdS20 ($r-_B$, $m-_B$) *supE44 recAB ara-14 leuB5proA2 lacY1 galK rpsL20* (str^R) *xyl-5 mtl-1*) carrying R701 (Tra+, Mob+ conjugative plasmid)	(Williams et al. 1990)
***Escherichia coli* XL1-Blue**	Commercial competent cells (Stratagene)	
Plasmids*		
pMTL82151	ColE1, *cat*P, *tra*J, MCS and pBP1	(Heap et al. 2010b)
pMTL83151	ColE1, *cat*P, *tra*J, MCS and pCB102	(Heap et al. 2009)
pMTL84151	ColE1, *cat*P, *tra*J, MCS and pCD6	(Heap et al. 2009)
pMTL85151	ColE1, *cat*P, *tra*J, MCS and pIM13	(Heap et al. 2009)
pMTL960	ColE1, *cat*P, *tra*J, MCS and pCD6; Non- modular structure	(Purdy et al. 2002)
pMTL3β-glu	pMTL83151 modular plasmid with β-(1,3-1,4)-glucanase gene of *Streptococcus bovis* JB1 expressed from multiple cloning site.	This work

*pBP1, pCB102, pCD6 and pIM13 (replicons of these plasmids). ColE1 (Gram-negative replicon), *cat*P (chloramphenicol resistance gene), *tra*J (origin of transfer) and MCS (multiple cloning site). A2-194 Rif^R was created by selecting spontaneous mutations that conferred rifampicin resistance to A2-194. Overnight cultures were streaked on plates containing 100µg/ml rifampicin and incubated anaerobically at 37 °C for 120 h. Single colonies were subsequently purified and the resistance levels checked.

Table 2. Antibiotic sensitivity testing of strains studied.

Species	Strain	Cm5	Tm5	Erm10	Erm25	Tet10	Tet25	Rif10	Rif25
E. rectale	A1-86	-	-	-	-	-	-	-	-
	M104/1	-	-	-	-	-	-	-	-
	T1-815	-	-	-	-	-	-	-	-
	L2-21	-	-	-	-	-	-	-	-
R. inulinivorans	A2-194	-	-	+	+	-	-	-	-
	A2-194 Rif^R*	-	-	+	+	-	-	+	+
	L1-83	-	-	-	-	-	-	-	-
R. hominis	A2-183	-	-	-	-	-	-	-	-
R. faecis	M72/1	-	-	-	-	-	-	-	-
R. intestinalis	L1-82	-	-	-	-	+	-	-	-
F. prausnitzii	A2-165	-	-	-	-	-	-	-	-

Growth of colonies (+) and no growth (-) on YCFAGSC plates supplemented with antibiotic at levels indicated. Chloramphenicol (Cm), thiamphenicol (Tm),

erythromycin (Erm), tetracycline (Tet) and rifampicin (Rif). 5, 10 and 25 indicate 5 μg/ml, 10 μg/ml and 25 μg/ml, respectively.

Table 3. Growth data for strains and media tested for mating and selection of transconjugants.

		Aerobic		Anaerobic					
Strain		**M2**	**Y**	**M2**	**AMM**	**Y**	**Y(0.25)FA**	**AMMcm5**	**Ycm5**
E. rectale	**A1-86**	-	-	+	+	+	+	+	-
	M104/1	-	-	+	+	+	+	+	-
	T1-815	-	-	+	+	+	+	+	-
	L2-21	-	-	+	+	+	+	+	-
R. inulinivorans	**A2-194**	-	-	+	+	+	+	+	-
	L1-83	-	-	+	+	+	+	-	-
R. faecis	**M72/1**	-	-	+	+	+	+	-	-
R. intestinalis	**L1-82**	-	-	+	+	+	+	-	-
R. hominis	**A2-183**	-	-	+	+	+	+	+	-
F. prausnitzii	**A2-165**	-	-	+	+	+	+	+	-
E. coli	**CA434**	+	+	+	+	-	+	+	-

Growth of colonies (+) and no growth (-). Anaerobic mating medium (**AMM**), YCFAGSC (**Y**), M2GSC (**M2**), YCFAGSC with only one quarter of the normal concentration of each short-chain fatty acid (**Y(0.25)FA**). 5 μg/ml of chloramphenicol (**cm5**).

Table 4. Summary of *in silico* prediction of restriction-modification systems.

Strain	**Restriction-modification systems**
***E. rectale* A1-86**	1 Type I, 2 Type II
***R. inulinivorans* A2-194**	1 Type I
***R. intestinalis* L1-82**	1 Type I, 1 Type II
***F. prausnitzii* A2-165**	4 Type I, 2 Type II, 1 Type III

Detailed description in Table S3.

Figures.

Figure 1. Restriction profile of plasmid DNA incubated with protein extracts. Protein extracts from: A (*Eubacterium rectale* A1-86), B (*Faecalibacterium prausnitzii* A2-165), C (*Roseburia faecis* M72/1), D (*Roseburia intestinalis* L1-82), E (*Roseburia inulinivorans* A2-194) and + (no protein extract) were incubated with each of the four plasmids shown at 37 °C for 90 minutes. The restriction enzyme buffer used for each extract was the one shown to generate the clearest restriction bands by λ DNA restriction. The size reference used (left hand lane) was the Promega 1 Kb DNA ladder.

Figure 2. Verification of transconjugants. A. Coomassie staining of cell lysates reveals that transconjugants (EAM3, 1 and 2) have the same SDS protein profile as *E. rectale* A1-86, different to lane 1 containing the *E. coli* donor. **B.** PCR screening for plasmid reveals its presence in the transconjugants and not in wild-type A1-86 recipient (The complete PCR gel image is shown in Fig. S2). **C, D.** Southern blotting of plasmids shows that they have not inserted into the recipient chromosome in: **C.** ***E****ubacterium rectale* **A**1-86 harbouring pMTL8**3**151 (**EAM3**), ***E****. rectale* **T**1-815 harbouring p**M**TL8**3**151 (**ETM3**) and ***R****oseburia inulinivorans* **A**2-194 harbouring pMTL8**3**151 (**RAM3**) or in **D.** ***R****oseburia inulinivorans* **A**2-194 harbouring p**M**TL8**2**151 (**RAM2**).

Figure 3. Diagrammatic representation of the optimised conjugation protocol. Chloramphenicol (Cm). Media as described in Table 3 and Materials and Methods section.

Figure 4. Construction of the pMTL3β-glu expression vector. The β-(1,3-1,4)-glucanase gene from *S. bovis* JB1 was amplified from pL1Hc using primers possessing 5'-end restriction sites for BamHI and HindIII. This restriction-flanked amplicon and pMTL83151 were restricted with BamHI and HindIII in separate reactions, purified and ligated together with T4 ligase. Dashed lines in pL1Hc indicate where the cloned fragment attaches to the multiple

cloning site of the pUC18 backbone. *rep*H is the replicon of pBP1 (replication in Gram-positive bacteria), *cat*P is the chloramphenicol resistance gene, *ColE1* RNAII is the replicon of ColE1 (replication in Gram-negative bacteria) and *tra*J is the origin of transfer.

Figure 5. β-(1,3-1,4)-glucanase activity A) Activity plates show barley β-glucan degradation after overnight anaerobic incubation with wild-type *R. inulinivorans* A2-194 and *E. rectale* T1-815 or pMTL3β-glu-harbouring transconjugants. Clear zones surrounding bacterial colonies illustrate areas in which the barley β-glucan has been hydrolysed. Transconjugant codes are: ***E****ubacterium rectale* **T**1-815 harbouring pMTL3**β-glu** (**ETBglu**) and ***R****oseburia inulinivorans* **A**2-194 harbouring pMTL3**β-glu** (**RABglu**). **B)** Barley β-glucan degradation after two hour aerobic incubation of 1% Glucagel with sonicated extracts from *Roseburia inulinivorans* A2-194 or *R. inulinivorans* RABglu and *Eubacterium rectale* T1-815 or *E. rectale* ETBglu. Bars represent standard error of triplicate incubations, and the experiment was repeated three times.

References.

[1] F. Van Immerseel, R. Ducatelle, M. De Vos, N. Boon, T. Van De Wiele, K. Verbeke, P. Rutgeerts, B. Sas, P. Louis, H.J. Flint, Butyric acid-producing anaerobic bacteria as a novel probiotic treatment approach for inflammatory bowel disease, J. Med. Microbiol. 59 (2010) 141-143.

[2] J. Durack, S.V. Lynch, The gut microbiome: Relationships with disease and opportunities for therapy, J. Exp. Med. 216 (2019) 20-40.

[3] L. Peng, Z.R. Li, R.S. Green, I.R. Holzman, J. Lin, Butyrate enhances the intestinal barrier by facilitating tight junction assembly via activation of AMP-activated protein kinase in Caco-2 cell monolayers, J. Nutr. 139 (2009) 1619-1625.

[4] J. Schauber, R.A. Dorschner, K. Yamasaki, B. Brouha, R.L. Gallo, Control of the innate epithelial antimicrobial response is cell-type specific and dependent on relevant microenvironmental stimuli, Immunology. 118 (2006) 509-519.

[5] I.A. Finnie, A.D. Dwarakanath, B.A. Taylor, J.M. Rhodes, Colonic mucin synthesis is increased by sodium butyrate, Gut. 36 (1995) 93-99.

[6] A.L. Millard, P.M. Mertes, D. Ittelet, F. Villard, P. Jeannesson, J. Bernard, Butyrate affects differentiation, maturation and function of human monocyte-derived dendritic cells and macrophages, Clin. Exp. Immunol. 130 (2002) 245-255.

[7] P.V. Chang, L. Hao, S. Offermanns, R. Medzhitov, The microbial metabolite butyrate regulates intestinal macrophage function via histone deacetylase inhibition, Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 2247-2252.

[8] J.H. Cho, C. Abraham, Inflammatory bowel disease genetics: Nod2, Annu. Rev. Med. 58 (2007) 401-416.

[9] A. Barcenilla, S.E. Pryde, J.C. Martin, S.H. Duncan, C.S. Stewart, C. Henderson, H.J. Flint, Phylogenetic relationships of butyrate-producing bacteria from the human gut, Appl. Environ. Microbiol. 66 (2000) 1654-1661.

[10] K. Terpe, Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems, Appl. Microbiol. Biotechnol. 72 (2006) 211-222.

[11] F. Bernaudat, A. Frelet-Barrand, N. Pochon, S. Dementin, P. Hivin, S. Boutigny, J. Rioux, D. Salvi, D. Seigneurin-Berny, P. Richaud, Heterologous expression of membrane proteins: choosing the appropriate host, PloS one. 6 (2011) e29191.

[12] E.D. Carlson, R. Gan, C.E. Hodgman, M.C. Jewett, Cell-free protein synthesis: applications come of age, Biotechnol. Adv. 30 (2012) 1185-1194.

[13] B. Miroux, J.E. Walker, Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels, J. Mol. Biol. 260 (1996) 289-298.

[14] M.A. Fischbach, Microbiome: focus on causation and mechanism, Cell. 174 (2018) 785-790.

[15] K.P. Scott, J.C. Martin, J. Mrazek, H.J. Flint, Transfer of conjugative elements from rumen and human Firmicutes bacteria to *Roseburia inulinivorans*, Appl. Environ. Microbiol. 74 (2008) 3915-3917.

[16] J.T. Heap, O.J. Pennington, S.T. Cartman, N.P. Minton, A modular system for Clostridium shuttle plasmids, J. Microbiol. Methods. 78 (2009) 79-85.

[17] K. Miyazaki, J.C. Martin, R. Marinsek-Logar, H.J. Flint, Degradation and utilization of xylans by the rumen anaerobe *Prevotella bryantii* (formerly P. ruminicola subsp. brevis) B(1)4, Anaerobe. 3 (1997) 373-381.

[18] M. Lopez-Siles, T.M. Khan, S.H. Duncan, H.J. Harmsen, L.J. Garcia-Gil, H.J. Flint, Cultured representatives of two major phylogroups of human colonic *Faecalibacterium*

573 *prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth, Appl.
574 Environ. Microbiol. 78 (2012) 420-428.

575 [19] M. Bryant, Commentary on the Hungate technique for culture of anaerobic bacteria, Am.
576 J. Clin. Nutr. 25 (1972) 1324-1328.

577 [20] K. Rutherford, J. Parkhill, J. Crook, T. Horsnell, P. Rice, M.A. Rajandream, B. Barrell,
578 Artemis: sequence visualization and annotation, Bioinformatics. 16 (2000) 944-945.

579 [21] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular cloning, Cold spring harbor laboratory
580 press New York, 1989.

581 [22] K.P. Scott, H. Flint, Transfer of plasmids between strains of *Escherichia coli* under
582 rumen conditions, J. Appl. Bacteriol. 78 (1995) 189-193.

583 [23] D. Purdy, T.A. O'Keeffe, M. Elmore, M. Herbert, A. McLeod, M. Bokori-Brown, A.
584 Ostrowski, N.P. Minton, Conjugative transfer of clostridial shuttle vectors from *Escherichia*
585 *coli* to *Clostridium difficile* through circumvention of the restriction barrier, Mol. Microbiol.
586 46 (2002) 439-452.

587 [24] M. Lever, Carbohydrate determination with 4-hydroxybenzoic acid hydrazide
588 (PAHBAH): effect of bismuth on the reaction, Anal. Biochem. 81 (1977) 21-27.

589 [25] H.J. Flint, C.A. McPherson, J. Martin, Expression of two xylanase genes from the rumen
590 cellulolytic bacterium *Ruminococcus flavefaciens* 17 cloned in pUC13, J. Gen. Microbiol.
591 137 (1991) 123-129.

592 [26] K.A. Kazimierczak, M.T. Rincon, A.J. Patterson, J.C. Martin, P. Young, H.J. Flint, K.P.
593 Scott, A new tetracycline efflux gene, *tet*(40), is located in tandem with *tet*(O/32/O) in a
594 human gut firmicute bacterium and in metagenomic library clones, Antimicrob. Agents
595 Chemother. 52 (2008) 4001-4009.

596 [27] H. Flint, A. Thomson, Deoxyribonuclease activity in rumen bacteria, Lett. Appl.
597 Microbiol. 11 (1990) 18-21.

598 [28] A.J. Roe, C. O'Byrne, D. McLaggan, I.R. Booth, Inhibition of *Escherichia coli* growth
599 by acetic acid: a problem with methionine biosynthesis and homocysteine toxicity,
600 Microbiology. 148 (2002) 2215-2222.

601 [29] M.S. Ekinci, S.I. McCrae, H.J. Flint, Isolation and overexpression of a gene encoding an
602 extracellular beta-(1,3-1,4)-glucanase from *Streptococcus bovis* JB1, Appl. Environ.
603 Microbiol. 63 (1997) 3752-3756.

604 [30] K.P. Scott, J.C. Martin, C. Chassard, M. Clerget, J. Potrykus, G. Campbell, C.D. Mayer,
605 P. Young, G. Rucklidge, A.G. Ramsay, H.J. Flint, Substrate-driven gene expression in
606 *Roseburia inulinivorans*: importance of inducible enzymes in the utilization of inulin and
607 starch, Proc. Natl. Acad. Sci. U. S. A. 108 Suppl 1 (2011) 4672-4679.

[31] K.P. Scott, J.C. Martin, G. Campbell, C.D. Mayer, H.J. Flint, Whole-genome transcription profiling reveals genes up-regulated by growth on fucose in the human gut bacterium "Roseburia inulinivorans", J. Bacteriol. 188 (2006) 4340-4349.

[32] D.W. Cockburn, N.I. Orlovsky, M.H. Foley, K.J. Kwiatkowski, C.M. Bahr, M. Maynard, B. Demeler, N.M. Koropatkin, Molecular details of a starch utilization pathway in the human gut symbiont *Eubacterium rectale*, Mol. Microbiol. (2015).

[33] A.G. Ramsay, K.P. Scott, J.C. Martin, M.T. Rincon, H.J. Flint, Cell-associated alpha-amylases of butyrate-producing Firmicute bacteria from the human colon, Microbiology. 152 (2006) 3281-3290.

[34] P.O. Sheridan, J.C. Martin, T.D. Lawley, H.P. Browne, H.M. Harris, A. Bernalier-Donadille, S.H. Duncan, P.W. O'Toole, K.P. Scott, H.J. Flint, Polysaccharide utilisation loci and nutritional specialisation in a dominant group of butyrate-producing human colonic Firmicutes., Microbial Genomics (2015).

[35] B.A. Neville, P.O. Sheridan, H.M. Harris, S. Coughlan, H.J. Flint, S.H. Duncan, I.B. Jeffery, M.J. Claesson, R.P. Ross, K.P. Scott, Pro-inflammatory flagellin proteins of prevalent motile commensal bacteria are variably abundant in the intestinal microbiome of elderly humans, PloS one. 8 (2013) e68919.

[36] A. Zhernakova, A. Kurilshikov, M.J. Bonder, E.F. Tigchelaar, M. Schirmer, T. Vatanen, Z. Mujagic, A.V. Vila, G. Falony, S. Vieira-Silva, J. Wang, F. Imhann, E. Brandsma, S.A. Jankipersadsing, M. Joossens, M.C. Cenit, P. Deelen, M.A. Swertz, LifeLines cohort study, R.K. Weersma, E.J. Feskens, M.G. Netea, D. Gevers, D. Jonkers, L. Franke, Y.S. Aulchenko, C. Huttenhower, J. Raes, M.H. Hofker, R.J. Xavier, C. Wijmenga, J. Fu, Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity, Science. 352 (2016) 565-569.

Table S1. Oligonucleotide primers.

Primer	Sequence (5'-3')	Target	Reference
PS#MTL-for	TATCTATGATACCGTGGTCAAC	pMTL80000 series plasmids	This work
PS#MTL-rev	CTGCTGAAGCCAGTTACC	pMTL80000 series plasmids	This work
PS#Bglu-BamHI-for	GGCCGGATCCGATATTGGCTGCAGTTATT	*S. bovis* JB1 β-glucanase gene with native promoter	This work
PS#Bglu-HindIII-rev	GCGCAAGCTTGGAAACAGCTATGACCATG		This work
FD1	AGAGTTTGATCCTGGCTCAG	Full 16S rRNA gene	[1]
RP2	ACGGCTACCTTGTTACGACTT	Full 16S rRNA gene	[1]
519R	GWATTACCGCGGCKGCTG	16S rRNA gene (universal)	[2]
926F	ACTCAAAGGAATTGACGG	16S rRNA gene (universal)	[3]

Nucleotide code: Guanine (G), adenine (A), thymine (T), cytosine (C), adenine or thymine (W) and guanine or thymine (K)

Table S2. Stability of pMTL83151 in *E. rectale* A1-86 in the absence of antibiotic selection.

Sub-inoculum	R of EAM3(1)	R of EAM3(2)	Generations (N)	Instability of EAM3(1)	Instability of EAM3(2)
1	0.92	0.96	8	-	-
2	0.88	0.50	16	0.01	0.05
3	0.74	0.21	24	0.01	0.06
4	0.60	0.18	32	0.02	0.05
5	0.44	0	40	0.02	-
			Average	**0.02**	**0.05**

E. rectale A1-86 transconjugants EAM3(1) and EAM3(2), harbouring pMTL83151, were isolated from independent conjugations. Number of generations was estimated by incubation time. Stability was calculated using the equation $x = 1 - R^{1/N}$, where x = segregational instability and was taken as the average of sub-inoculums 2-5 for EAM3(1) and 2-4 for EAM3(2). R = fraction of bacteria still possessing plasmid and N = number of generations.

Table S3. Detailed *in silico* prediction of restriction-modification systems.

Strain		Locus tag	Annotation
***R. intestinalis* L1-82**	**Type I**	RINT_03062c	Type 1 RM system, R subunit
		RINT_03061c	Type 1 RM system, M subunit
		RINT_03060c	Type 1 RM system, S subunit
	Type II	RINT_00004c	BspRI-like methylase
		RINT_00005c	BspRI-like restriction endonuclease
		RINT_00287c	Eco57I-like methylase
		RINT_00286c	BsuBI/PstI-like restriction endonuclease
***R. inulinivorans* A2-194**	**Type I**	RINU_01867c	R subunit
		RINU_01866c	M subunit
		RINU_01864c	S subunit
***E. rectale* A1-86**	**Type I**	EUR_27160	M subunit
		EUR_27170	S subunit
		EUR_27190	nucleotidyltransferase substrate binding protein
		EUR_27200	Site-specific recombinase XerD
		EUR_27210	S subunit
		EUR_27220	Predicted AAA-ATPase
		EUR_27230	R subunit
	Type II	EUR_02430	DNA adenine methylase (dam)
		EUR_02440	Hypothetical protein
		EUR_02450	DpnIIB-like methylase
		EUR_02460	5-methylcytosine-specific restriction enzyme
		EUR_02470	DpnII-like restriction endonuclease
		EUR_07860	D12 class N6 adenine-specific DNA methylase
		EUR_07870	MjaII-like restriction endonuclease
***F. prausnitzii* A2-165**	**Type I**	FPRAU_01424	Type 1 RM system, DNA inversion
		FPRAU_00430c	Type 1 RM system, DNA inversion
		FPRAU_02637c	Type 1 RM system, looks like DNA inversion but no inverted repeats
		FPRAU_03233	Type 1 RM system
	Type II	FPRAU_00871	Eco57I-like methylase
		FPRAU_00872	Restriction endonuclease
		FPRAU_01759	NgoMIV-like methylase
		FPRAU_01758	NgoMIV-like restriction endonuclease
	Type III	FPRAU_00185c	Hypothetical protein
		FPRAU_00184c	M subunit
		FPRAU_00183c	R subunit
		FPRAU_00182c	Hypothetical protein
		FPRAU_00181c	TIR-domain containing protein

Note – A similar analysis was not done for *Roseburia faecis* as the genome sequence was not then available

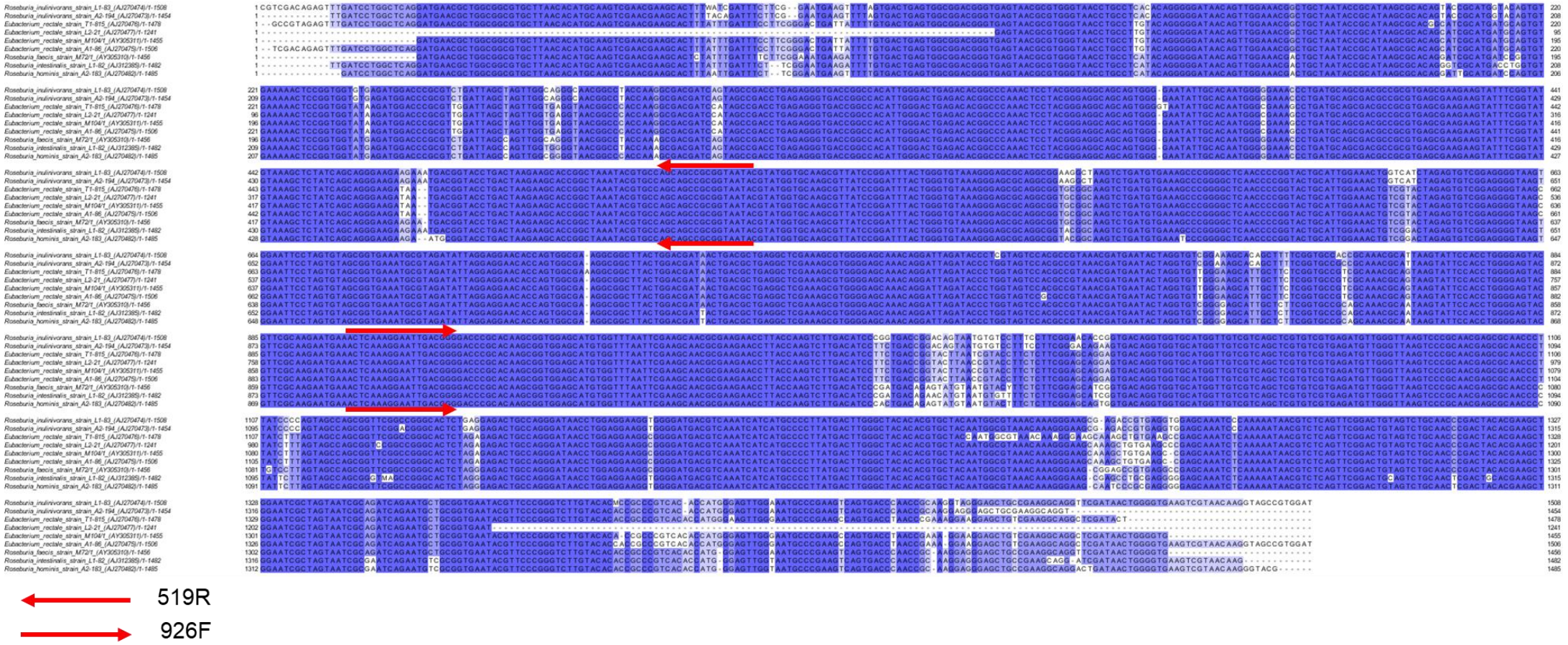


Figure S1. Alignment of 16S rRNA gene sequences from *Roseburia*/*E. rectale* strains. The positions of sequencing primers 519R and 926F are indicated with red arrows.

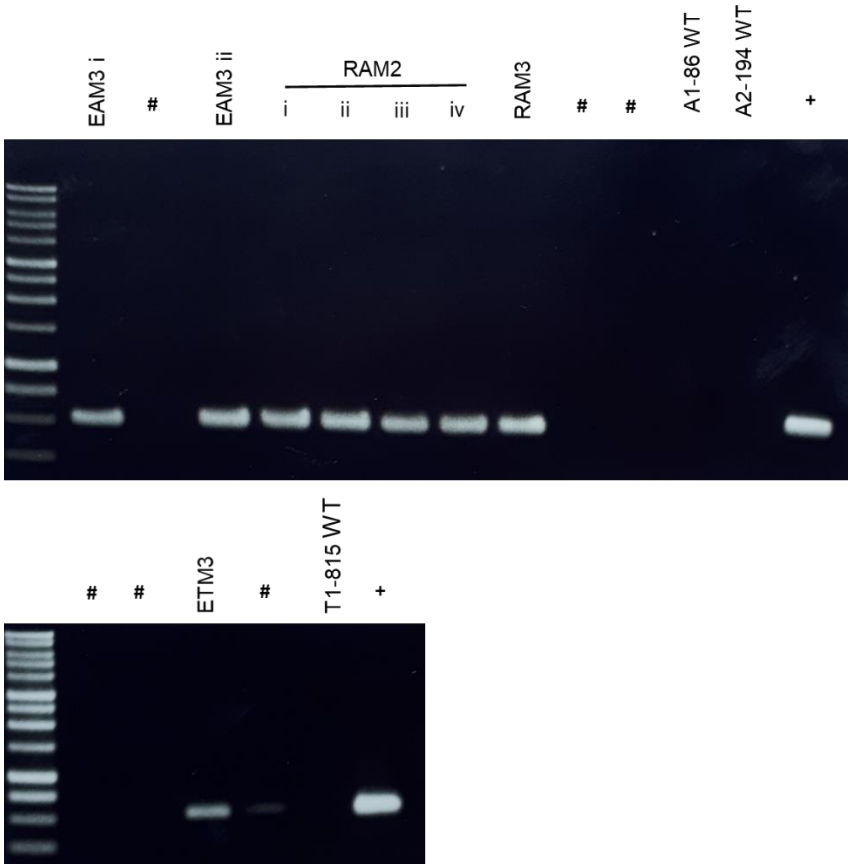


Figure S2. Original gel electrophoresis images of PCR verification of transconjugants. This original gel image contains some negative results that have been labelled #.

The confirmed transconjugants are labelled as described in the main manuscript. The size reference used (left hand lane) was the Promega 1 Kb DNA ladder. The positive control (+) is the plasmid pMTL82151, which is identical to the pMTL83151 in the amplified region.

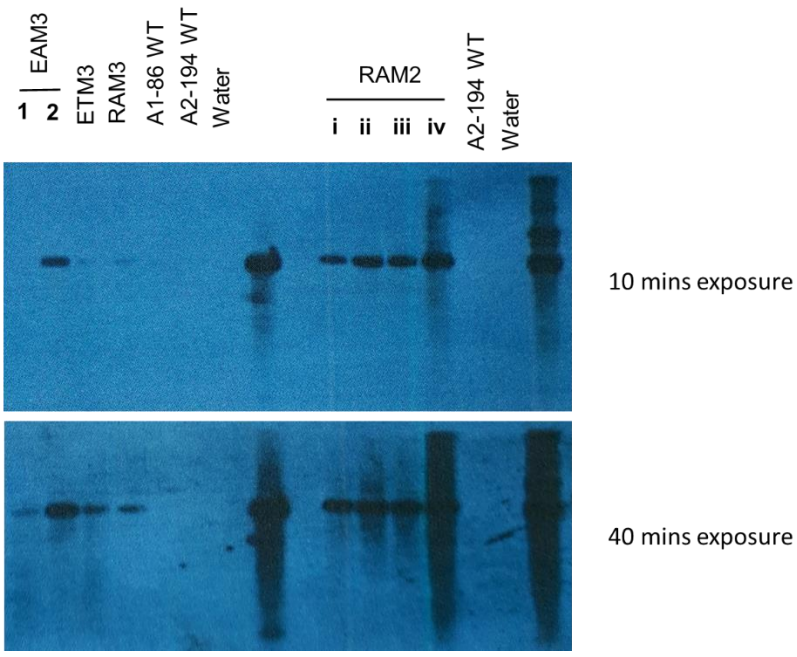


Figure S3. Original southern blot images at 10 mins and 40 mins film exposure.

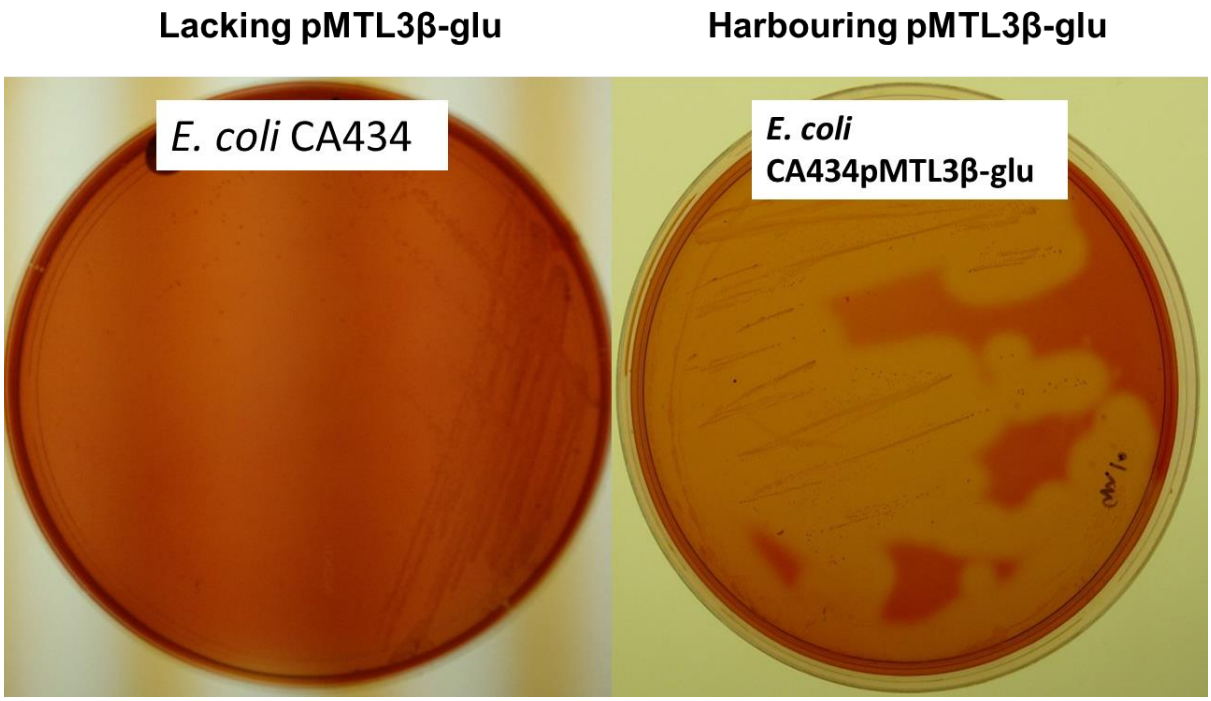


Figure S4. Beta-glucanase activity by conjugation donor *E. coli* CA434 harbouring pMTL3β-glu. Barley β-glucan degradation after overnight aerobic incubation with wild-type or pMTL3β-glu-harbouring transformants of conjugation donor *E. coli* CA434. Clear zones surrounding bacterial colonies are areas in which the barley β-glucan has been hydrolysed.

References:

[1] J. Wood, K.P. Scott, G. Avgustin, C.J. Newbold, H.J. Flint, Estimation of the relative abundance of different *Bacteroides* and *Prevotella* ribotypes in gut samples by restriction enzyme profiling of PCR-amplified 16S rRNA gene sequences, Appl. Environ. Microbiol. 64 (1998) 3683-3689.

[2] S. Turner, K.M. Pryer, V.P. Miao, J.D. Palmer, Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis1, J. Eukaryot. Microbiol. 46 (1999) 327-338.

[3] G. Muyzer, A. Teske, C.O. Wirsen, H.W. Jannasch, Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments, Arch. Microbiol. 164 (1995) 165-172.

Fig. 1

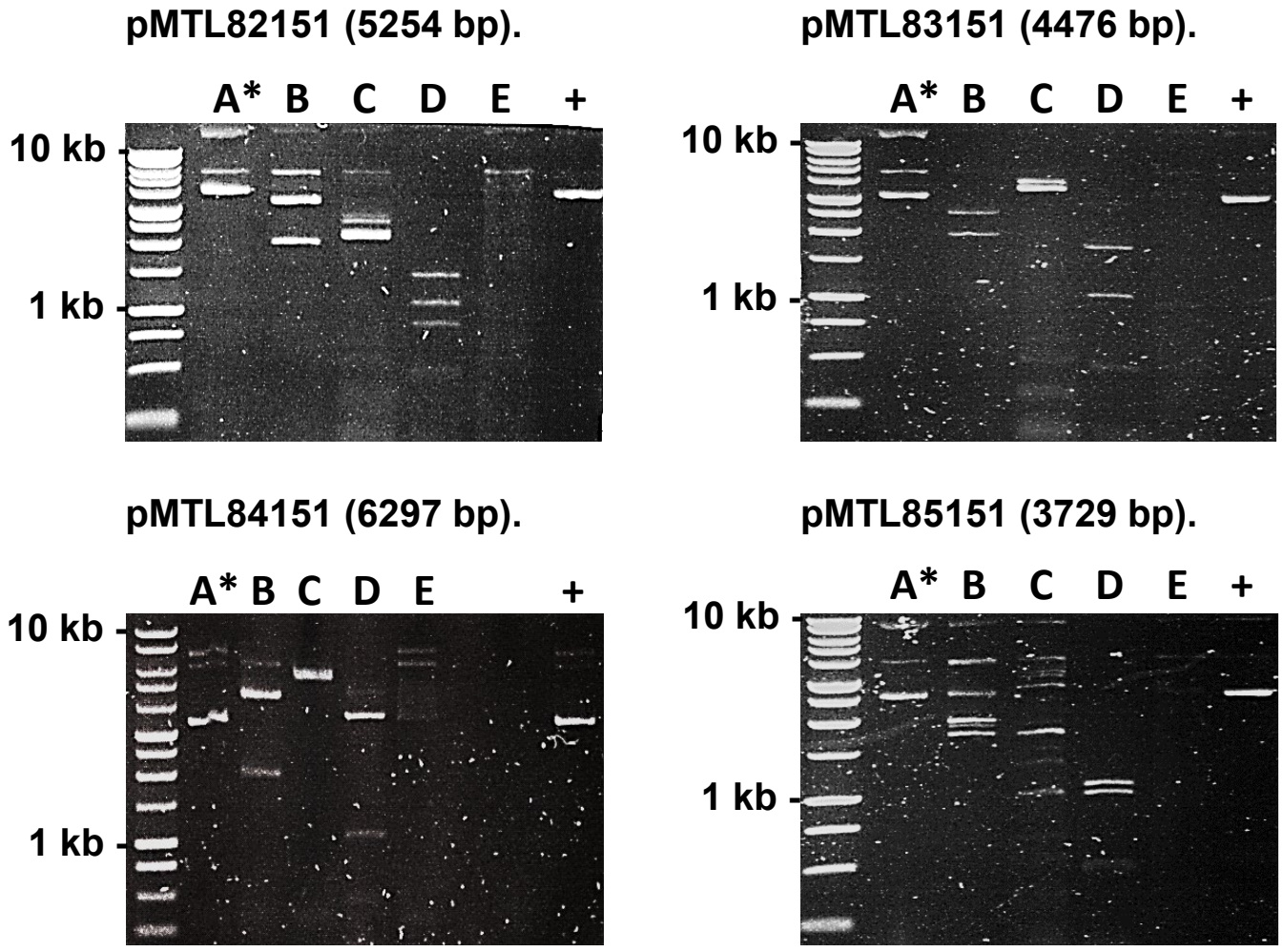

pMTL82151 (5254 bp).
A* B C D E +
10 kb
1 kb
pMTL83151 (4476 bp).
A* B C D E +
10 kb
1 kb
pMTL84151 (6297 bp).
A* B C D E +
10 kb
1 kb
pMTL85151 (3729 bp).
A* B C D E +
10 kb
1 kb

Fig. 2

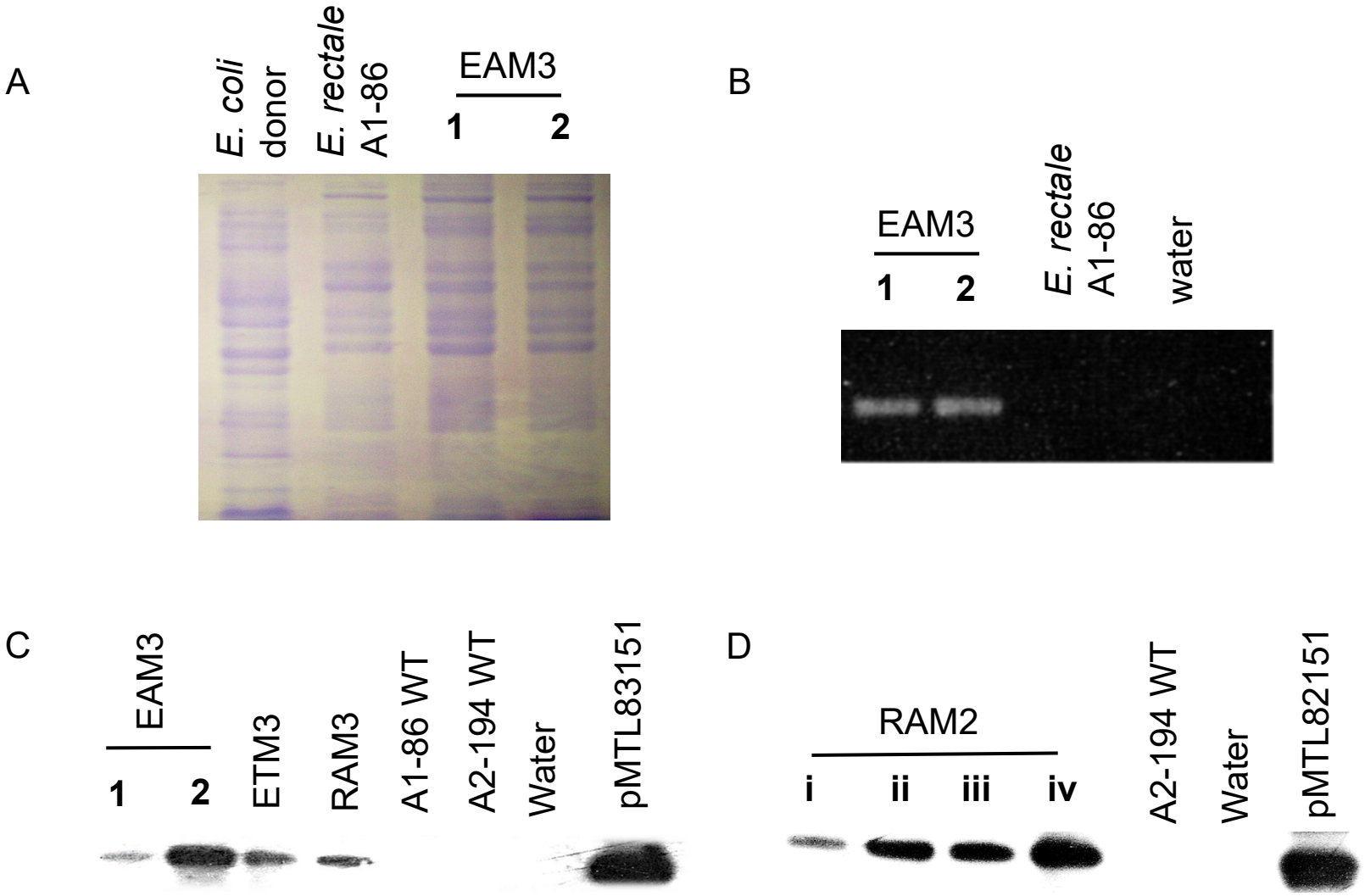
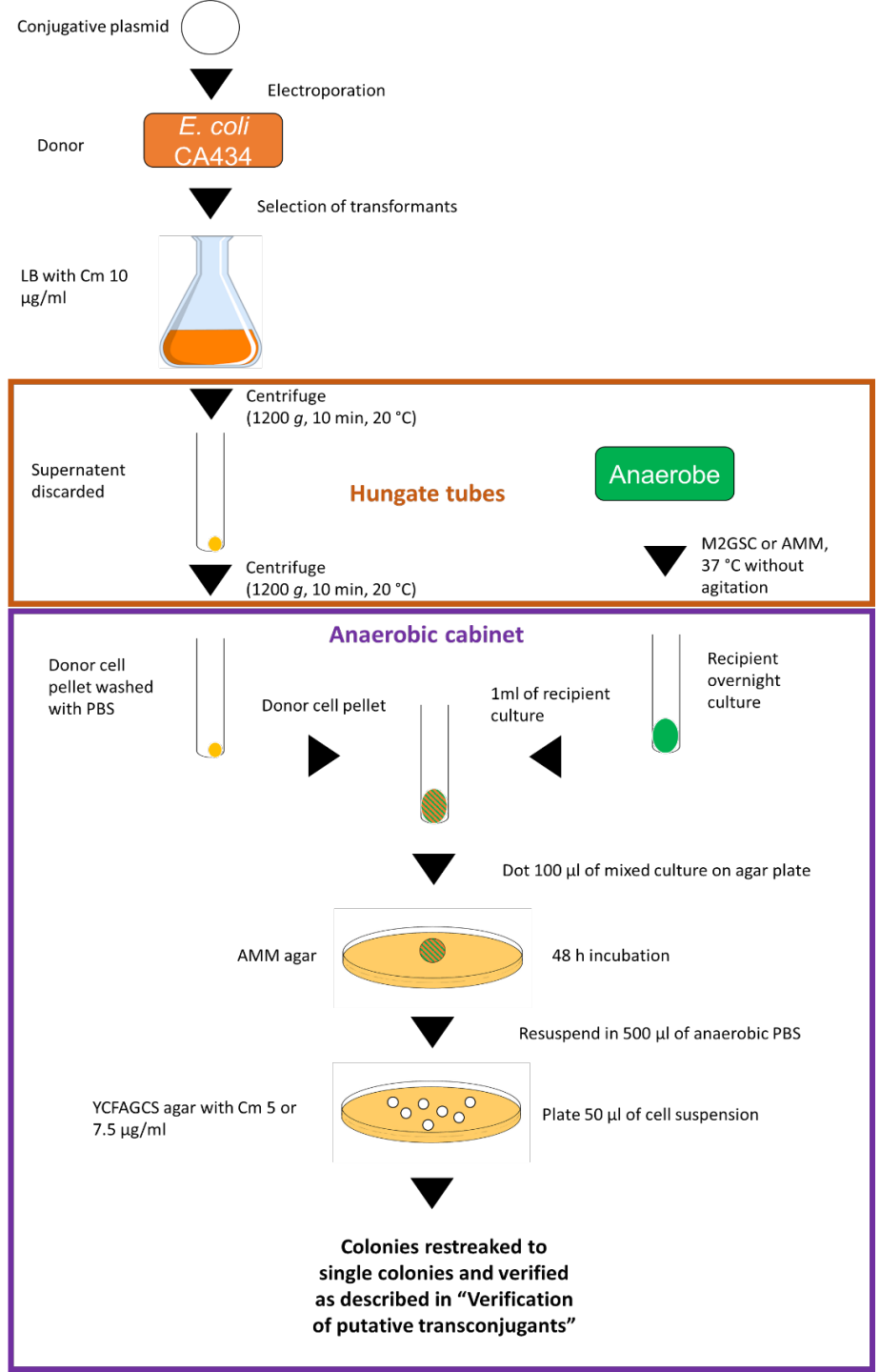

A
E. coli donor
E. rectale A1-86
EAM3
1
2
B
EAM3
1
2
E. rectale A1-86
water
C
EAM3
1
2
ETM3
RAM3
A1-86 WT
A2-194 WT
Water
pMTL83151
D
RAM2
i
ii
iii
iv
A2-194 WT
Water
pMTL82151

Fig. 3


Conjugative plasmid
Electroporation
Donor
E. coli CA434
Selection of transformants
LB with Cm 10 µg/ml
Centrifuge (1200 g, 10 min, 20 °C)
Supernatent discarded
Hungate tubes
Anaerobe
M2GSC or AMM, 37 °C without agitation
Centrifuge (1200 g, 10 min, 20 °C)
Anaerobic cabinet
Donor cell pellet washed with PBS
Donor cell pellet
1ml of recipient culture
Recipient overnight culture
Dot 100 µl of mixed culture on agar plate
AMM agar
48 h incubation
Resuspend in 500 µl of anaerobic PBS
YCFAGCS agar with Cm 5 or 7.5 µg/ml
Plate 50 µl of cell suspension
Colonies restreaked to single colonies and verified as described in "Verification of putative transconjugants"

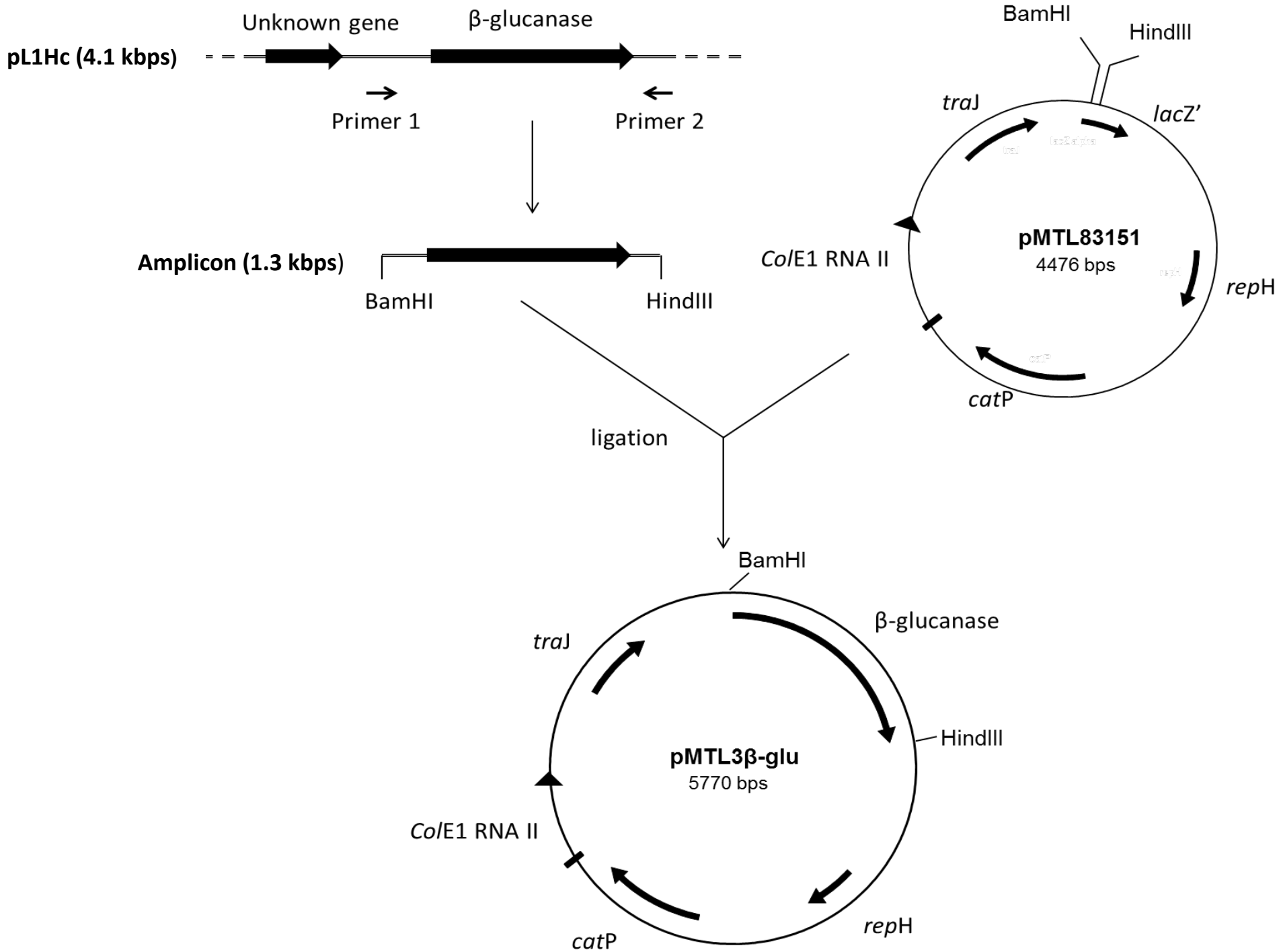
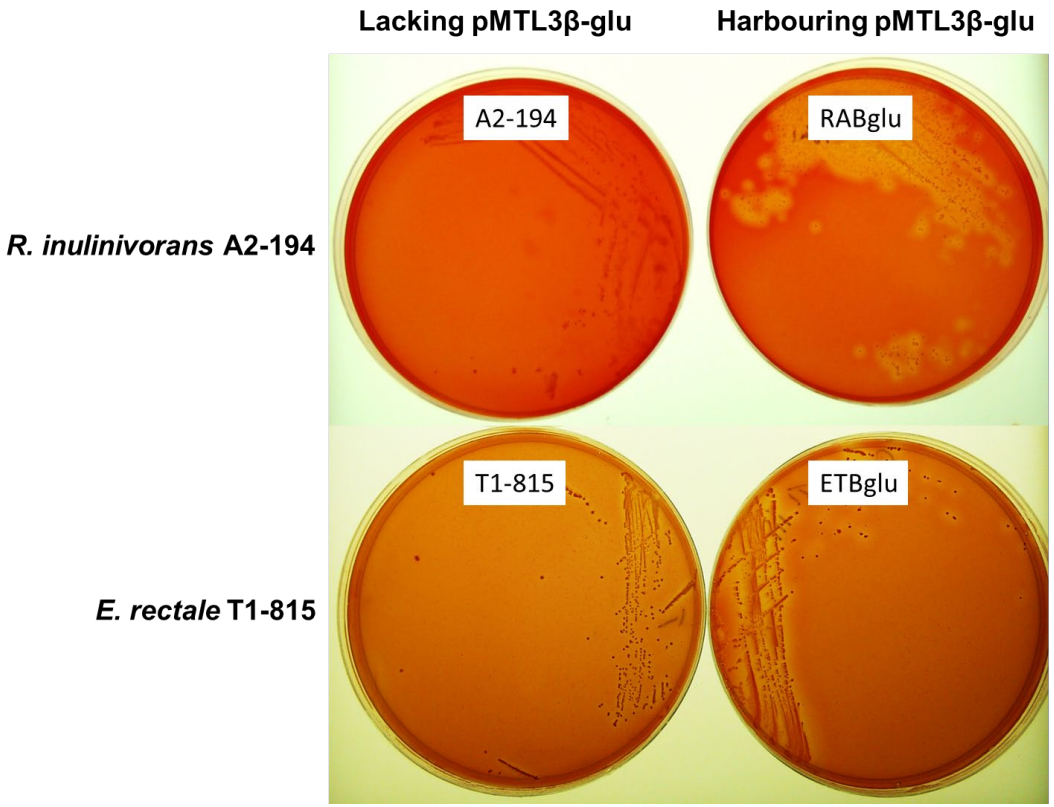

pL1Hc (4.1 kbps)
Unknown gene
β-glucanase
Primer 1
Primer 2
Amplicon (1.3 kbps)
BamHI
HindIII
ligation
BamHI
HindIII
traJ
lacZ'
pMTL83151
4476 bps
ColE1 RNA II
repH
catP
BamHI
β-glucanase
traJ
HindIII
pMTL3β-glu
5770 bps
ColE1 RNA II
repH
catP

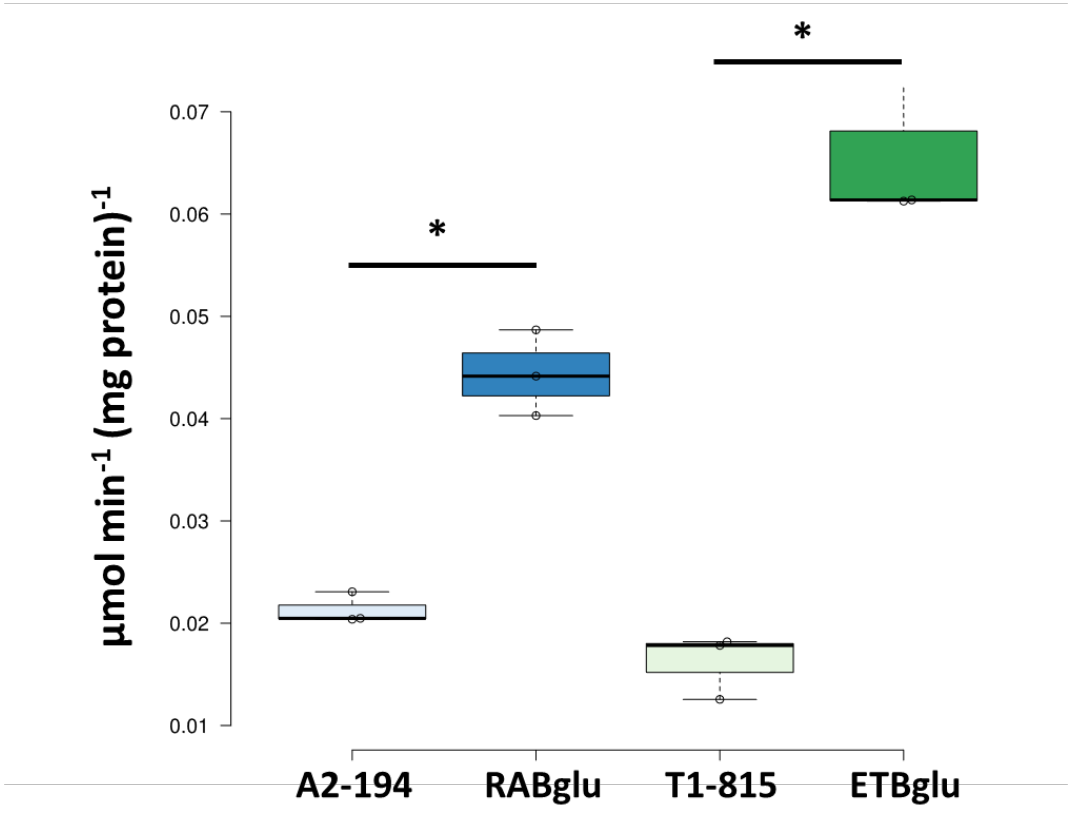
Fig. 4

Fig. 5

A


Lacking pMTL3β-glu
Harbouring pMTL3β-glu
R. inulinivorans A2-194
E. rectale T1-815
A2-194
RABglu
T1-815
ETBglu

B


μmol min-1 (mg protein)-1
0.07
0.06
0.05
0.04
0.03
0.02
0.01
*
*
A2-194
RABglu
T1-815
ETBglu