

RESEARCH PAPER



Novel insights into carbohydrate utilisation, antimicrobial resistance, and sporulation potential in *Roseburia intestinalis* isolates across diverse geographical locations

Indrani Mukhopadhyay^{a,b}, Jennifer C. Martin^a, Sophie Shaw^{c,d}, Martin Gutierrez-Torrejon^a, Nikoleta Boteva^a, Aileen J. McKinley^e, Silvia W. Gratz^a, and Karen P. Scott^a

^aGut Microbiology Group, Rowett Institute, University of Aberdeen, Aberdeen, UK; ^bMicrobiology and Immunity, Institute of Medical Sciences, University of Aberdeen, Aberdeen, UK; ^cCentre for Genome Enabled Biology and Medicine, University of Aberdeen, Aberdeen, UK; ^dAll Wales Medical Genomics Service, Institute of Medical Genetics, University Hospital of Wales, Heath Park, Cardiff, UK; ^eDepartment of Surgery, Aberdeen Royal Infirmary Foresterhill, Aberdeen, UK

ABSTRACT

Roseburia intestinalis is one of the most abundant and important butyrate-producing human gut anaerobic bacteria that plays an important role in maintaining health and is a potential next-generation probiotic. We investigated the pangenome of 16 distinct strains, isolated over several decades, identifying local and time-specific adaptations. More than 50% of the genes in each individual strain were assigned to the core genome, and 77% of the cloud genes were unique to individual strains, revealing the high level of genome conservation. Co-carriage of the same enzymes involved in carbohydrate binding and degradation in all strains highlighted major pathways in carbohydrate utilization and reveal the importance of xylan, starch and mannose as key growth substrates. A single strain had adapted to use rhamnose as a sole growth substrate, the first time this has been reported. The ubiquitous presence of motility and sporulation gene clusters demonstrates the importance of these phenotypes for gut survival and acquisition of this bacterium. More than half the strains contained functional, potentially transferable, tetracycline resistance genes. This study advances our understanding of the importance of *R. intestinalis* within the gut ecosystem by elucidating conserved metabolic characteristics among different strains, isolated from different locations. This information will help to devise dietary strategies to increase the abundance of this species providing health benefits.

ARTICLE HISTORY

Received 28 July 2024
Revised 11 February 2025
Accepted 21 February 2025

KEYWORDS





Roseburia intestinalis;
pangenome; butyrate;
carbohydrate utilization;
CAZyme; antimicrobial
resistance; sporulation


Introduction

The gastrointestinal lumen is populated with heterogeneous species of bacteria, which have diverse functions that determine health and disease in the human host.¹ In the colon, the vast majority of these bacteria belong to the Bacillota (previously Firmicutes) and Bacteroidetes phyla.^{2,3} One of the key functions of these bacteria is to convert complex carbohydrates that cannot be digested by host enzymes into bacterial metabolites, including the short chain fatty acids (SCFAs) acetate, butyrate, propionate and succinate.^{4,5} These SCFAs have a profound effect on ‘colonic health’ as they not only constitute one of the major energy sources for colonocytes but also exhibit distinct anti-inflammatory and anti-carcinogenic effects.^{6–8}

They also play a critical role in the maintenance of barrier function in the colon.⁹ Butyrate, in particular, has an immunomodulatory function and is a known histone deacetylase inhibitor, which may explain its protective role in inflammation and carcinogenesis, respectively.¹⁰

Most of the butyrate producing bacteria are Bacillota with *Roseburia* spp., *Agathobacter rectalis* (previously *Eubacterium rectale*), *Anaerobutyricum hallii* (previously *Eubacterium hallii*), *Anaerostipes* spp. and *Faecalibacterium* spp. being the prominent species.^{11,12} Of these, *Roseburia* spp. and *A. rectalis* are believed to comprise nearly 24% of all bacteria in the healthy human colon.^{5,13} The importance of these critical bacteria has been

CONTACT Karen Scott  k.scott@abdn.ac.uk  Gut Microbiology Group, Rowett Institute, University of Aberdeen, Aberdeen AB25 2ZD, UK; Indrani Mukhopadhyay  indrani.mukhopadhyay@abdn.ac.uk  Microbiology and Immunity, Institute of Medical Sciences, University of Aberdeen, Aberdeen AB25 2ZD, UK

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/19490976.2025.2473516>.

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

highlighted in patients with inflammatory bowel disease where they are typically reduced in the mucosa and feces of affected patients.¹⁴ Conversely, restitution of these two species parallels clinical remission in patients with ulcerative colitis who were treated with fecal microbiota transplantation, underlining their true anti-inflammatory potential.¹⁵ On account of the prodigious production of butyrate by *Roseburia* spp. and *A. rectalis*, they are prime candidates as probiotic strains to revert a dysbiotic gut ecosystem.¹⁶

Roseburia spp. have been documented to have numerous health benefits in conditions ranging from inflammatory bowel diseases, colorectal cancer and atherosclerosis.^{17–20} *Roseburia intestinalis* was first identified by our group in human fecal samples as a novel saccharolytic and butyrate producing species.²¹ *R. intestinalis* has since been found to ameliorate intestinal inflammation by its selective effect on T regulatory (Treg) cells and upregulating expression of anti-inflammatory cytokines like transforming growth factor- β (TGF- β) and down-regulating pro-inflammatory cytokines like thymic stromal lymphopoietin (TSLP) and interleukin 17 (IL-17).^{22,23} *R. intestinalis* has also ameliorated inflammatory pathways in mice with experimental colitis and reversed 27-Hydroxycholesterol-induced learning and memory impairment in mice.^{24,25} This bacterial species is depleted in patients with Crohn's disease, highlighting its importance in the maintenance of gut homeostasis in humans.²²

Genomic assessment of *Roseburia* spp. together with *A. rectalis*, a closely related species, has shown multiple shared genotypic and phenotypic traits, including butyrate pathway genes and flagellar motility.^{12,26} Similar concordance was found in the genes encoding oligo-saccharide transporters and regulatory elements in the genomes of these two bacterial species, forming distinct Gram-positive polysaccharide utilization loci (gpPULs).²⁷ Comparative genomic assessment of core and pangenomes of different species within the genus *Roseburia* has uncovered significant differences in both degradative and biosynthetic capabilities of different species, especially with respect to carbohydrate utilization and co-factor synthesis for vitamin

metabolism and utilization.²⁸ Such subtle differences may determine the ability of a particular bacterial species to survive in selective ecological niches in the human colon and also determine its interaction with the other resident bacteria in the lumen. The sporulation capacity of *Roseburia* spp. is poorly characterized, yet this is an important survival and colonization mechanism. The prototype *R. intestinalis* strain L1–82 has been recognized for its potential probiotic properties and safety for use in humans.²⁹ Our study aims to advance knowledge of this important bacterium and specifically focus on the pangenome assessment of *R. intestinalis* strains isolated from healthy human subjects from different geographical regions, corroborating genotypic findings with phenotypic analysis to establish the functional diversity. This will help elucidate species-specific biosynthetic and metabolic pathways of *R. intestinalis*, underpinning its potential health benefits and identifying potential dietary routes to enhance prevalence within an individual.

Methods

Bacterial isolates and growth medium

Two *Roseburia intestinalis* strains (PC335 and PC352) were newly isolated from two healthy adult volunteers using anaerobic culture methods after approval from the Ethics Committee of North of Scotland Research Ethics Service (Reference 17/NS/0112) and Rowett Human Studies Ethical Review Panel, as previously described.³⁰ *R. intestinalis* strain PC335 was isolated from a fecal sample, and PC352 was isolated from a rectal biopsy sample. *Roseburia intestinalis* type strain L1–82 isolated from infant feces,³¹ strain M50/1 isolated from adult feces³² (both by our group at the Rowett Institute), and strain XB6B4 isolated from adult feces in the Unite de Microbiologie, INRAE, France, by Annick Bernalier-Donadille³³ were included for comparison. These five strains were routinely maintained in complex, broad-range YCFAGSC bacteriological medium³⁴ under anaerobic conditions in Hungate tubes (7.5 ml aliquots).

Genomic DNA extraction

Bacterial DNA was isolated from fresh overnight cultures of *R. intestinalis* PC335 and PC352 using Promega Wizard® Genomic DNA Purification Kit (Promega) following the manufacturer's recommendations. The DNA preparation was quantified using the Qubit™ DNA HS Assay Kit (Life Technologies, Thermo Fisher Scientific Inc.) and Qubit™ 3.0 Fluorometer (Life Technologies, Thermo Fisher Scientific Inc.). The integrity of the DNA was confirmed for all samples using the Agilent 2200 TapeStation system.

Whole genome sequencing, assembly, and annotation

DNA extracted from *R. intestinalis* PC335 and PC352 was sequenced at the Centre for Genome-Enabled Biology and Medicine (CGEBM) at the University of Aberdeen. DNA samples were sequenced first on a single Illumina MiSeq v2 Micro flowcell (Illumina, San Diego, CA, USA) using the NextEra XT sequencing kit producing 2 × 150 bp paired end reads according to manufacturer's protocol. Isolates were further sequenced on MinION RevD flowcells (FLO-MIN107) (Oxford NanoPore Technologies, Oxford, UK) for 72 h using the SQK-LSK108 sequencing kit following manufacturer's instructions. The QC-filtered Illumina and MinION sequence data were then subjected to hybrid assembly using UNICYCLER 0.4.5³⁵ assembler which combines both data types and attempts to circularize contigs. Assembly

quality statistics were obtained using QUAST 3.1,³⁶ and per-contig coverage by Illumina and MinION reads was calculated using QUALIMAP 2.1.1.³⁷ Genome completeness was examined by searching the genomes for a set of 148 bacterial universal single-copy genes and quantifying the proportion of genes that had unique mapping locations, using BUSCO 3.0.0.³⁸ All assemblies were processed through the PROKKA 1.11³⁹ pipeline, predicting genes and tRNAs/tmRNAs/rRNAs and annotating gene predictions from the full Swissport and Pfam databases.

Pangenome analysis to identify the core and accessory genomes

Sixteen *R. intestinalis* genomes were included in the analysis, with 14 genomes available from the NCBI genome database and two (PC335 and PC352) available from this current study. All available *R. intestinalis* genomes in the NCIMB genome database at the time of the study (in 2021) were considered for the pangenome analysis. The sequence of a further strain (43_16) was not included in the comparative analysis due to evidence that multiple genetic regions were missing within the published assembly, which would have affected assessment of the core genome. All NCBI genomes were downloaded directly in fast format. The list of final 16 genomes analyzed for this study is given in Table 1.

To ensure standardization across assemblies, all genome assemblies were re-annotated from scratch

Table 1. List of *Roseburia intestinalis* genomes used in this study and annotation statistics.

<i>R. intestinalis</i> strain	Assembly Identifier	Contigs	Assembly Level	Genome Size (bp)	Coding Sequences	rRNA	tRNA	tmRNA
2789STDY5834960	14207_7#83	101	Scaffold	4322432	3895	2	56	1
AF31-21AC	ASM347551v1	79	Scaffold	4142092	3861	2	52	1
AF36-10AT	ASM347488v1	81	Scaffold	4303042	3911	2	63	1
AM22-21LB	ASM347104v1	68	Scaffold	4112416	3701	2	49	1
AM37-1AC	ASM346772v1	60	Scaffold	4253056	3939	3	59	1
AM43-11	ASM346703v1	91	Scaffold	4337927	3996	3	56	1
BIOML-A1	ASM971834v1	79	Contig	4348274	3977	2	60	1
BSD2780061689_150309_G12	ASM1555478v1	157	Scaffold	4171472	3726	2	56	1
J1101437_171009_C5	ASM1555578v1	95	Scaffold	4570840	4235	2	57	1
L1-82	<i>Roseburia intestinalis</i> strain L1-82	1	Complete	4493348	4067	15	69	1
M50/1	ASM20999v1	1	Chromosome	4143550	3794	2	51	1
MSK.17.84	ASM1330044v1	161	Contig	4374175	4063	2	63	1
PC335	This study	11	Contig	4411266	4042	14	77	1
PC352	This study	16	Contig	4422834	4134	13	59	1
SNUG30017	ASM975952v1	79	Contig	3902409	3518	4	72	1
XB6B4	ASM21065v1	1	Chromosome	4286292	3881	2	62	1

using Prokka (version 1.13.7)³⁹ with the most recent Pfam database⁴⁰ provided. Pangenome analysis was then carried out using Roary (version 3.12.0)⁴¹ classifying genes as core, soft-core, shell, and cloud genes according to their varying presence among the genomes analyzed. Thus, genes present in >99% strains (all 16 strains) were designated as core genes; genes present in 95% to 99% (15) strains were classified as soft-core genes; shell genes were present in 15% to 95% (3–15) strains, and genes present in less than 15% of the strains analyzed (less than 2 genomes) were assigned as cloud genes. The core genome alignment produced by Roary was used for phylogenetic analysis of the relatedness of the strains using RAxML (version 8.2.12)⁴² using the model GTRGAMMA and 100 bootstraps. A circular map of the 16 *R. intestinalis* genomes was constructed using the Proksee web server (<https://proksee.ca/>).⁴³

Carbohydrate active enzymes (CAZymes) were identified in the pangenome using dbCAN2⁴⁴ using search tools HMMER (e-value < 1e-15, coverage > 0.35), DIAMOND (e-value < 1e-102), Hotpep (Frequency > 2.6, Hits > 6) and CGCFinder (Distance ≤ 2, signature genes = CAZyme + TC). These parameters are based on established benchmarks enabling accurate functional annotation. All genes with a consistent positive identification by two or more tools were used for further analysis. Predictions of gpPULs were achieved by further analyses of genes indicated using >2 tools to be GH2, GH3, GH13, GH23, GH43, or GH78. Up to 15 genes up and downstream of each predicted GH were identified, or up to the limits of the contig, and the PROKKA gene annotations were tabulated. gpPULs containing specific glycoside hydrolase genes alongside genes encoding carbohydrate transport enzymes and regulatory genes were then manually curated.

To identify potential antimicrobial resistance genes (AMR) across the pangenome, the consensus sequence of all 11,814 genes were first converted to protein sequences using transeq (part of EMBOSS package; version 6.6.0),⁴⁵ and AMR genes were identified by comparison to the Comprehensive Antibiotic Resistance Database (CARD) with RGI (version 5.1.0),⁴⁶ using embedded detection thresholds. Further, to identify potential mobile genetic elements (MGEs)

across the pangenome, the translated protein sequences of all genes were compared to the ISFinder database⁴⁷ using BLASTP (version 2.9.0)⁴⁸ with an e-value cut-off of 1e-05 and a percent identity cut-off of 70%. Custom Python scripts were then used to identify the top hit per gene based on the smallest e-value (with the highest percentage identity used when e-values were identical).

Functional annotation of genomes

The consensus gene sequence for all of the *R. intestinalis* genes in the core, shell and cloud gene sets were functionally categorized by comparison to the Clusters of Orthologous Groups (COG) database (version COG2020)⁴⁹ using BLASTX (version 2.9.0)⁴⁸ with an e-value cutoff of 1e-06. Custom python scripts were used to filter all hits with a percentage identity >70% and to then select the “top hit” for each gene based on the smallest e-value (when e-values were equal, highest percentage identity was used to determine top hit). These top hits were then matched to the COG identifier and functional group using custom python scripts.

Functional annotation, homology and synteny analysis of *tet(O)* regions

Functional annotation of genes surrounding *tet(O)* in all strains that contained it was performed with BLASTx with standard settings, and functional identification was accepted with >95% cover and sequence homology as cutoffs. Homology and synteny of the *tet(O)* containing regions was performed by aligning each region to its closest relative with BLASTn with standard settings and visualizing the resulting hit table (text) with the Artemis Comparison Tool.⁵⁰

Growth assays on different carbohydrate substrates

The five *R. intestinalis* strains that were available within the Rowett (PC335, PC352, XB6B4, M50/1 and L1–82) were tested to confirm specific phenotypes. Sterile, polystyrene, flat-bottom, 96-well tissue culture plates (Costar) were used. Growth kinetics in the 96-well microplates was monitored (as turbidity) with a computer-controlled plate

reader (BioTek EPOCH 2) running Gen5 Microplate data collection and analysis software (version 3.10) and a 650-nm filter. Growth of the seed cultures was monitored on a Dynamica Halo-Vis 20 spectrophotometer by measuring the optical density at 650 nm (OD_{650}) of 7.5 ml samples.

Short-chain and branched chain fatty acid (SCFA and BCFA) quantification

SCFA/BCFA formation was assessed in culture supernatants by gas chromatography (GC) following standard procedures.⁵¹ Bacterial cultures were grown in triplicate on YCFAGSC overnight and OD_{650} readings were taken to confirm equivalent growth. OD_{650} readings ranged from 2.12 to 2.46, with a mean of 2.26 and a median of 2.23. The SCFA/BCFA concentrations (mM) were calculated in duplicate from standard curves based on peak size and retention time. Concentrations measured in a media-only control are subtracted from concentrations measured in individual bacterial cultures to give a net concentration that is a balance of production and consumption. An internal standard of 2-ethyl butyric acid and an external standard (a mixture of volatile fatty acids and salts [SCFA: 30 mM Acetic acid, 20 mM Propionic acid, 20 mM Butyric acid, 5 mM Valeric acid; BCFA: 5 mM iso-Valeric acid, 5 mM Iso-Butyric acid] plus salt solutions of 10 mM sodium formate; 10 mM Lithium lactate and 10 mM sodium succinate) was included in each GC run.

Antibiotic susceptibility

Antibiotic susceptibility testing was conducted using fresh overnight bacterial cultures grown in YCFAGSC medium. Isolates were screened for possible resistance against a selection of 11 antibiotics as per EFSA guidelines⁵² using a slightly modified version of the agar disc diffusion method.⁵³ The turbidity of each growing bacterial suspension was adjusted by dilution with sterile media to match the turbidity standard of 0.5 McFarland units. Petri plates (9 cm diameter), containing 20 ml of YCFAGSC agar (2%), were overlaid with 10 ml of M2GSC soft agar (1%) containing 50 μ l of the diluted bacterial culture.

Antibiotic discs (Oxoid) were placed on the inoculated plates using the Oxoid disc dispenser. Following 48 h anaerobic incubation at 37°C, inhibition zones around the discs were measured. The tests were conducted and results interpreted according to EUCAST guidelines.⁵⁴

Tests to ascertain the minimum inhibitory concentrations of tetracycline and erythromycin were performed in the same way, but antibiotic-containing discs were replaced with ETEST strips (BioMérieux) impregnated with 0.016–256 μ g/ml tetracycline or erythromycin.

Detection of spores using transmission electron microscopy (TEM)

R. intestinalis strains (L1–82, M50/1, PC335 and PC352) were inoculated into fresh YCFAGSC liquid medium from frozen stocks and grown for 72 h. These cultures were subjected to both heat-shock (80°C, 20 min) and oxygen shock by subsequent exposure to air (30 min). Cultures were then pelleted by centrifugation and treated with primary and secondary fixatives following previously published methods.⁵⁵ Sporulation images were documented using a JEOL 1400 plus; TEM and AMT UltraVUE camera.

Hanging drop microscopy

Motility of the bacterial strains was detected using hanging drop microscopy as described previously.⁵⁶ Briefly, a small drop of bacterial suspension was placed onto a sterile cover slip. The cover slip was then carefully inverted over the well of a concave microscope slide, creating a sealed, humid chamber. This setup allowed observation of bacterial movement in a free-floating state under a light microscope using oil immersion.

Results

***Roseburia intestinalis* pangenome analysis**

Pangenome analysis to determine the over-lapping and strain-specific genes of the 16 selected genomes of *R. intestinalis* identified 11,814 unique protein-coding genes across all genomes. The number of coding sequences in an individual

genome ranged from 3518 to 4235 (Table 1). The gene accumulation curve showed that the number of genes assigned to the core genome decreased progressively with the addition of new strains, while the total pangenome increased (Figure 1(a)). The change in both curves slowed down because the pan-genome of *R. intestinalis* was in an open state, indicating that more unique genes would be added along with the addition of more new strains. The gene frequency plot depicting the

distribution of genes across the strains is shown in Figure 1(b). The pangenome analysis revealed that a core genome containing 2,123 genes was present in all 16 *R. intestinalis* strains and in addition, 3,134 shell genes and 6,547 cloud genes were identified (Figure 1(c)). Interestingly, 5,056 of the cloud genes were unique to one genome among these strains (Table 2). No soft-core genes were identified, and the core and shell genes accounted for 84% of the pangenome, indicating that the genomes of the

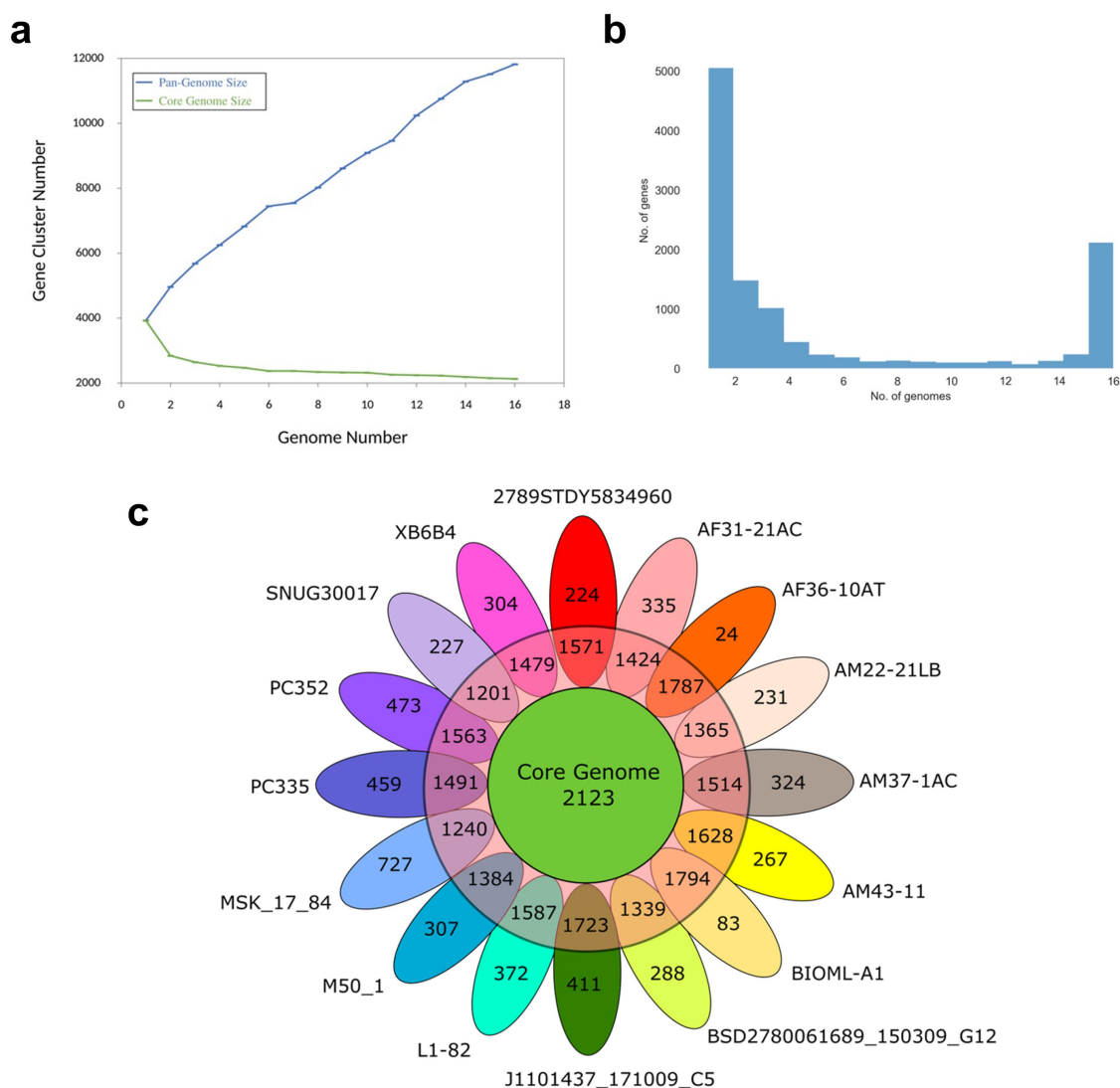


Figure 1. Pangenome analyses of 16 *Roseburia intestinalis* strains. a) pangenome accumulation curves. The blue line denotes the number of unique genes discovered with the sequential addition of new genomes. The green line denotes the number of core genes discovered with the sequential addition of new genomes. b) gene frequency plot of *R. intestinalis* genomes. c) genomic diversity of 16 *R. intestinalis* strains shown in a flower plot. Each strain is shown as a petal. The number of core genes (present in >99% strains) is shown in the center (green circle). Overlapping regions in the red circle show the shell genes (present in 15% to 95% of strains) conserved only within several strains. The numbers in non-overlapping portions show the number of unique strain-specific genes. The strain name is located beside each petal.

Table 2. Pangenome summary statistics per *R. intestinalis* strain.

<i>R. intestinalis</i> strain	No. of core genes	No. of shell genes	No. of cloud genes	No. of unique genes	No. of exclusively absent genes
2789STDY5834960	2123	1303	492	224	9
AF31-21AC	2123	1275	484	335	14
AF36-10AT	2123	1521	290	24	0
AM22-21LB	2123	1217	379	231	24
AM37-1AC	2123	1340	498	324	17
AM43-11	2123	1440	455	267	14
BIOML-A1	2123	1511	366	83	5
BSD2780061689_150309_G12	2123	1231	396	288	7
J1101437_171009_C5	2123	1535	599	411	5
L1-82	2123	1369	590	372	5
M50_1	2123	1267	424	307	39
MSK_17_84	2123	1098	869	727	8
PC335	2123	1293	657	459	10
PC352	2123	1340	696	473	30
SNUG30017	2123	1045	383	227	36
XB6B4	2123	1326	457	304	26

strains were remarkably conserved. The notable differences in the number of cloud genes between strains, ranging from only 24 genes in AF36-10AT to 727 genes in MSK.17.84 (Figure 1(c)), suggests that these two strains may have adapted to different habitats and that strain-specific genes may have been acquired through horizontal gene transfer from other species. Exclusively absent genes, classified as genes present in all other genomes, but not the genome specified were also identified in all but one strain (Table 2).

Phylogeny and genome comparisons

A phylogenetic tree based on the core genome identified up to three distinct clades potentially coinciding with three different geographical locations where the strains were isolated. The first clade consisted of six strains isolated from Asia (AM22-21LB, AM37-1AC, AF36-10AT, SNUG30017, AM43-11 and AF31-21AC), and the second clade consisted of four strains from Europe (M50/1, XB6B4, PC352 and PC335). A possible third clade comprised two strains isolated from America (J1101437_171009_C5 and BSD2780061689_150309_G12) (Figure 2(a)), while a further three strains were not part of any clade. This provides the first evidence of potential geographic stratification of *R. intestinalis* strains on a global scale, with strains isolated in Europe and Asia clustering separately. Detailed comparison of the sixteen *R. intestinalis* genome sequences is illustrated in the form of a circular map in Figure 2(b).

Functional traits of the *R. intestinalis* core and accessory genes

The proportion of genes assigned to specific functional groups differed across the core, shell and cloud genomes (Supplementary Table S1; Supplementary Figure S1). Functional enrichment using COG showed that the most conserved predicted proteins were those involved in primary metabolism and DNA processing functions. Notably, “translation, ribosomal structure and biogenesis” formed the largest portion of core genes (13.38%), with the proportion of this category of genes reduced in the shell (4.63%) and cloud (3.13%) genomes. Likewise, the proportion of genes with the function “energy production and conversion” also decreased from core to shell and cloud genomes. The proportion of “defense mechanisms” genes increased from core, to shell, to cloud genome, and this pattern was also observed for genes involved in “transcription”, “replication, recombination and repair”, and “cell wall/membrane/envelope biogenesis”. Interestingly, none of the genes associated with the “mobilome: prophages, transposons” was found in the core genome, with increasing proportions detected in the shell (4.92%) and cloud (12.53%) genomes.

There were clear differences in the proportion of unique genes within different functional categories across each genome (Supplementary Table S2; Supplementary Figure S2). For instance, genes related to “transcription” formed a relatively high proportion of the unique genes in most genomes,



Figure 2. Genomic relatedness of *Roseburia intestinalis* strains. a) Phylogenetic tree visualized alongside the pangenome gene presence/absence results. The strains originating from Asia are shown in red, the ones from Europe are in blue and the ones from America are in green. b) circular map showing a comparison of the *Roseburia intestinalis* genomes.

but <10% in AF36-10AT, BIOML-A1, and J1101437_171009_C5. The functional category of “exclusively absent” genes varied between each genome (Supplementary Table S3; Supplementary Figure S3).

Carbohydrate utilisation by *R. intestinalis*

Previous studies have reported that *R. intestinalis* can ferment dietary fibers to produce butyrate under anaerobic conditions, specializing in utilizing selective glycans like xylan and β -mannan.^{27,57} However, most of these studies have been carried

out using the type-strain L1–82 and there is limited information available on the carbohydrate utilization pattern of other *R. intestinalis* strains. In the present study, *R. intestinalis* genes were systematically identified against the CAZy database to obtain a comprehensive understanding of the carbohydrate degradative ability. A total of 295 genes within the pangenome contained at least one CAZy domain with 23 genes containing two or more domains. Of these domains, 63.0% were glycoside hydrolases (GH), 23.8% were glycosyl transferases (GT), 8.6% were carbohydrate-binding modules (CBM) and 4.6% were carbohydrate esterases (CE) (Figure 3(a)). No polysaccharide lyases (PL)

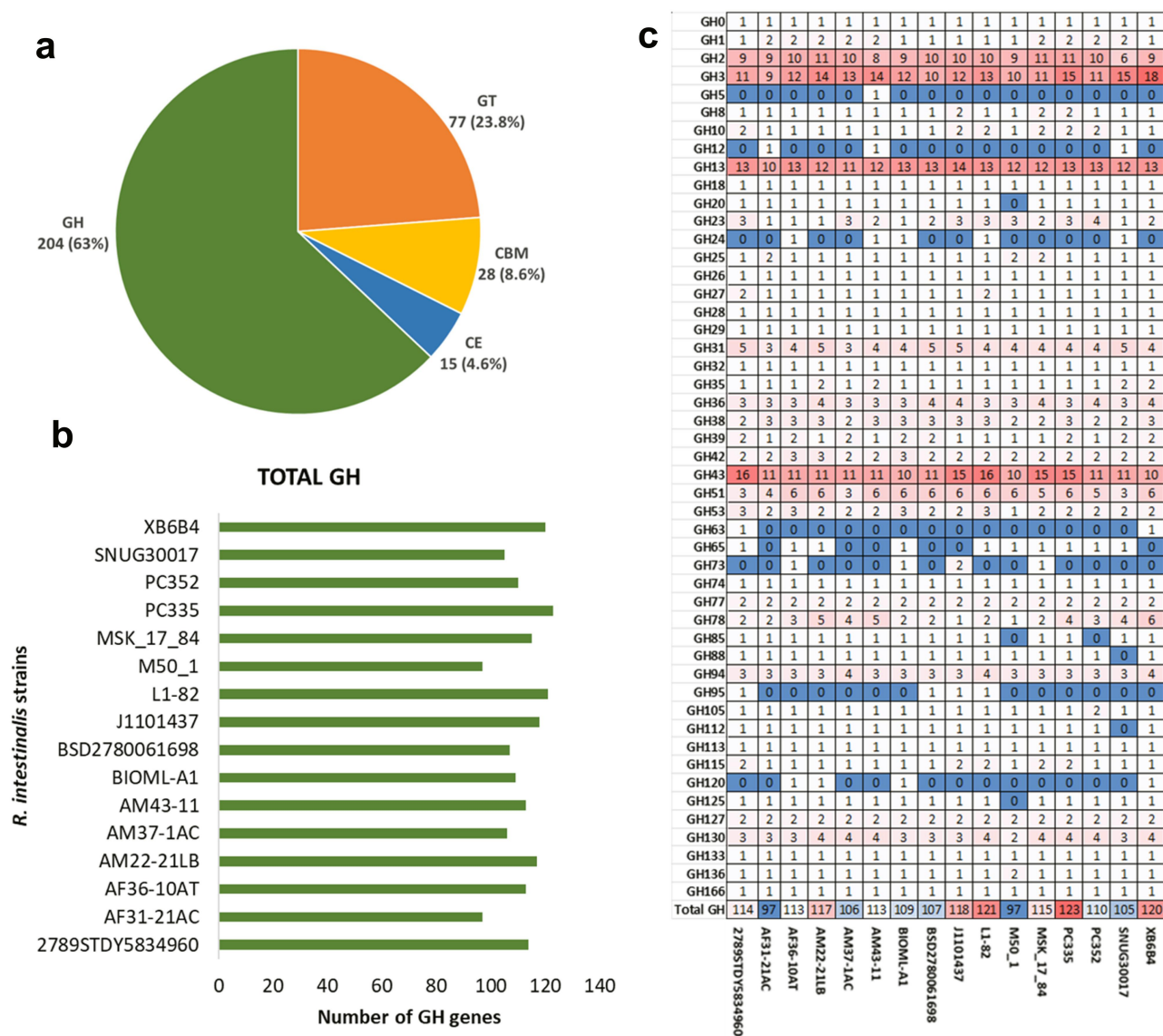


Figure 3. Carbohydrate active enzymes (CAZymes) a) percentage of each category of CAZyme (CBM – carbohydrate-binding module, CE – carbohydrate esterase, GH – glycoside hydrolase, GT – glycosyl transferase) b) total GH genes identified in each *R. intestinalis* genome c) number of GH genes across genomes with red indicating the most and blue indicating the least.

were identified in the *R. intestinalis* pangenome. Ninety-six CAZy domains were present in all strains, indicating considerable inter-strain conservation.

Each individual genome of the 16 human *R. intestinalis* strains encoded between 97 and 123 glycoside hydrolases (GHs) (Figure 3(b), Supplementary Figure S4). Most of these GH domains (over 50%) represented just six CAZyme families; GH3 (12.75%), GH43 (11.76%), GH2 (9.8%), GH13 (6.37%), GH23 (5.39%) and GH78 (5.39%). Among these families, GH3, GH43 (and additionally GH10 and GH51) are mainly involved in xylan utilization; GH13 enzymes with starch utilization; GH2 (and additionally GH31 and GH130) in mannose utilization and GH78 with rhamnose utilization. The presence of multiple copies of these specific gene families (GH2, GH3, GH13, GH43) in all isolates indicates that starch, xylan and mannose represent core substrates for *R. intestinalis* growth. Of the total 204 GH genes identified, 65 were from the core gene set across all 16 genomes, with a further 33 genes present in 10 or more genomes (Figure 3(c)).

The first step in glycan degradation often requires attachment of glycans to the bacterial cell surface and is typically mediated by carbohydrate-binding modules (CBMs). The most common CBMs identified in the *R. intestinalis* pangenome were associated with glycogen binding (CBM48), xylan binding (CBM6, CBM22, CBM86), starch binding (CBM34 and CBM83), mannan binding (CBM23 and CBM27) and L-rhamnose binding activity (CBM67). Of the CBMs identified, only two (both CBM50, associated with chitin or peptidoglycan degradation) were present as single-domain genes in the shell genome of 14 and 3 strains. All strains harbored single genes containing CBM50, except isolate SNUG30017 where two separate genes contained a CBM50 domain. The remaining 26 CBMs were part of multi-domain genes. In fact, 20/23 dual-domain genes contained a CBM linked to a GH. Two genes each contained four domains: CBM86, CBM22, GH10 and CBM9 that are all associated with xylan utilization. Although these genes were in the shell genome (identified in 12 and 5 isolates), all 16 genomes possessed a version of this multi-domain gene, with only one genome encoding both genes (PC352). Five multi-domain genes were part of

the core genome, present in all 16 isolates. Four of these were involved in starch utilization, combining GH13 with CBM43 (and one with CBM48) while the fifth containing CBM27, CBM23 and GH26 is likely to be involved in mannan utilization. The rhamnose utilization domains CBM67 and GH78 were always detected together comprising part of three clonal genes, one found in three isolates (AM37, L1–82 and XB6B4) and one in each of the single strains PC335 and PC352. The final three multi-domain genes each contained two GT domains. Two of these contained GT8 and GT111 (a β -1,3-galactofuranosyltransferase), while the other one containing GT2 and GT4 is likely to have multiple activities.

In addition to multi-domain genes, *Roseburia* species have been shown to contain loci containing catalytic enzymes alongside genes encoding regulatory and transport functions, known as gpPULs.²⁷ The 16 *R. intestinalis* genomes contained multiple potential gpPULs (Supplementary file 1), some of which contained the highly conserved CAZyme families. Specific gpPULs involved in xylan, mannose and starch utilization were identified, many present in all 16 strains (Figure 4, Supplementary Figure S5). Xylan utilization is highly conserved in *R. intestinalis* and nine specific xylan gpPULs were identified, frequently containing multiple GH43 CAZymes. Xylan gpPUL 1 and gpPUL 7 both contained 6 different GH43 enzymes but were each only present in single strains (L1–82 and PC335 respectively). In contrast, xylan gpPULs 2, 3, 4, 5 and 6 were present (and highly conserved) in most strains (Figure 4(a)). Mannose gpPUL 1, identified in all isolates, contained three GH2 domain enzymes, and an additional six CAZymes. Mannose gpPUL 2 was also found in all 16 isolates, while mannose gpPULs 3 and 4 were present in 7 and 10 isolates, respectively (Figure 4(b)). Starch gpPULs 1, 2 and 5 were also found in all strains, while starch gpPULs 3 and 4 were identified in 8 and 12 strains, respectively (Supplementary Figure S5). Only two gpPULs potentially involved in rhamnose utilization were identified. One of these (rhamnose gpPUL 1) was found in all strains and corresponds to mannose gpPUL 1. In contrast, rhamnose gpPUL 2, which contains two

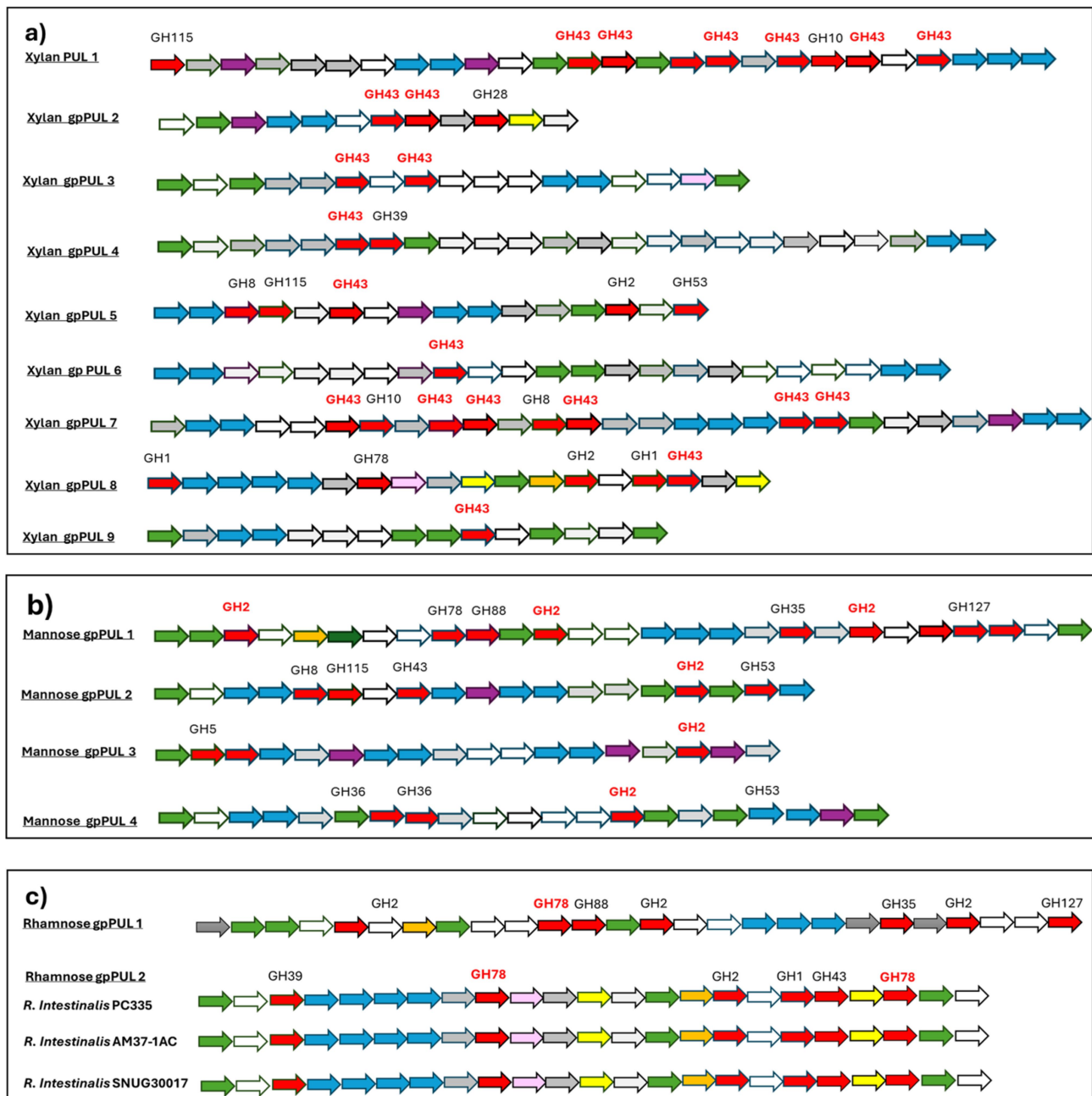


Figure 4. *Roseburia intestinalis* predicted gpPULs. Schematic representation of gpPULs concerned with a) xylan, b) mannose and c) rhamnose utilization. Glycoside hydrolase (GH) genes are colored red, and those associated with utilization of the key substrates further labeled in bold red text. ABC-transporter system component genes are colored blue. Transcriptional regulator genes are colored green. Xylose/rhamnose isomerase genes are colored yellow. Orange color represents the MFS/sugar transport genes, pink color represents the Methyl-accepting chemotaxis protein (MCP) signaling domain protein, and purple are bacterial extracellular solute binding genes. Hypothetical genes are colored gray. Genes with other unrelated functions are colored white.

GH78 enzymes and a further four CAZymes, was only found in three strains (Figure 4(c)). Some of these gpPULs also contained chemotaxis proteins that may facilitate utilization of the specific substrates in the competitive gut environment.

Growth assays to confirm functionality of CAZy genes identified

In order to ensure that the most abundant GH and CBM enzymes identified from the CAZyme analysis were functional, we tested the growth of the five available *R. intestinalis* strains (L1–82, M50/1,

XB6B4, PC335 and PC352) on YCFA media supplemented with single different carbohydrate substrates (Figure 5(a)). All five strains grew very well on the control substrate 0.5% glucose (average maximum OD₆₅₀ 1.0). They were also all able to grow well on 0.5% xylo-oligosaccharide (XOS) (avg. max. OD₆₅₀ 0.8) and to a moderate level on 0.5% Xylan (avg. max. OD₆₅₀ 0.5). Strain-specific differences were noted for the growth on 0.5% mannose with XB6B4, PC352 and L1-82 strains showing the highest growth (avg. max. OD₆₅₀ 0.78), whereas moderate growth was noted for PC335 strain (avg. max. OD₆₅₀ 0.67) and lower growth for M50/1 (avg. max. OD₆₅₀ 0.48). Moreover, there was shorter lag phase observed for XB6B4 strain compared to the other strains. Most strains grew well on 0.5% potato starch (avg. max. OD₆₅₀ 0.9), except PC335 did not grow in repeated tests and PC352 had a long (12–24 h) lag period. In contrast, the strains varied in ability to grow on 0.5% corn starch with the older isolates (L1-82, M50/1, XB6B4) growing very well (avg. max. OD₆₅₀ 0.8) while the newer isolates (PC335 and PC352) did not grow at all. None of the strains were initially able to grow on 0.5% rhamnose. However,

when PC335 was trained through repeated inoculations to grow on rhamnose under laboratory conditions, it became able to grow well (avg. max. OD₆₅₀ 0.6) (Figure 5(b)). The other strains tested (L1-82, PC352, M50/1) could not be trained to grow on rhamnose. Interestingly, of these five strains, only strain PC335 contained the rhamnose gpPUL 2.

Short chain fatty acid (SCFA) analysis

Genome analysis indicates that the Butyryl-CoA: acetate CoA-transferase gene was a single-core gene, present in the genomes of all isolates analyzed. None of the strains contained the butyrate kinase gene consistent with previous work demonstrating the butyrate synthesis pathway via the CoA-transferase gene in *Roseburia* species.¹² Analysis of SCFA production confirmed that all strains produced butyrate, with concentrations ranging from 13.4 mM to 25.3 mM. In pure culture, strains also produced significant amounts of formate (10.5–19.4 mM) and lower amounts of lactate (2.0–6.3 mM) (Figure 5(c)). *R. intestinalis* strains M50/1 and XB6B4 also consumed significant amounts

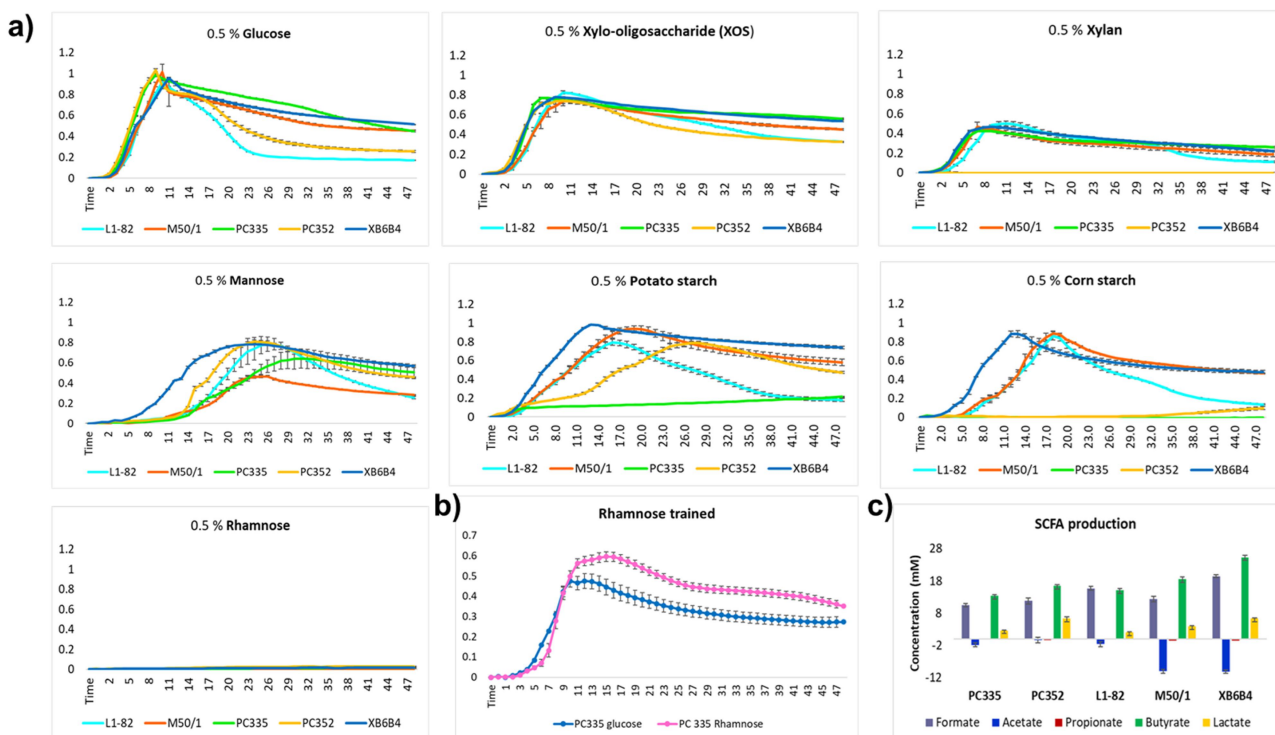


Figure 5. Growth kinetics of five *R. intestinalis* strains on different carbohydrate substrates (0.5%). b) growth of the 'trained' PC335 strain on rhamnose. c) Short Chain Fatty Acid concentrations measured in *R. intestinalis* strains cultured in YCFAGSC media. Net concentrations shown are relative to a media-only control and are a balance of production and consumption.

(10 mm) of acetate. None of the strains produced propionate, which was consistent with the lack of any genes linked to propionate production.

Antibiotic resistance and mobile genetic elements in *R. intestinalis*

Seven genes potentially encoding antimicrobial resistance (AMR) proteins were identified in the *R. intestinalis* pangenome which included the tetracycline-resistance genes *tet*(O) and *tet*(40) and erythromycin resistance gene *ermB* (Supplementary Table S4).

All the tetracycline resistance genes were present in the shell of the pangenome indicating acquisition may have been due to environmental pressure. The *tet*(O) gene was present in 10 of the 16 *R. intestinalis* strains (L1–82, AF31-21AC, AF36-10AT, AM22-21LB, AM43–11, BSD2780061689_150309_G12, BIOML-A1, PC335, PC352 and XB6B4) while the *tet*(40) gene was present in four strains (AF31-21AC, AM22-21LB, AM43–11 and L1–82). Notably, the *tet*(40) genes were located downstream of *tet*(O), and never occurred alone. The *tet*(40) gene was predominantly detected in Asian strains (3 out of 4), with a single occurrence in a European strain L1–82 and none in the American strains. The *erm*(B) gene was found in the shell genome of only three strains (AF31-21AC, AM37-1AC and AM43–11), all Asian isolates and two of which also contained *tet*(O).

The composition of the AMR gene neighborhood can affect the transfer of AMR genes between gut bacteria. For instance, genes that encode conjugative transposons or plasmids can facilitate the transfer of AMR genes between different bacteria through horizontal gene transfer events, while insertion sequences may facilitate their movement within the genome. A total of 48 genes were identified with significant hits to the ISFinder database (Supplementary Table S5), of which 44 were identified as transposases. Thirty-eight of these genes were present in the cloud gene set of either one or two genomes and none in the core genome.

The presence of any possible mobile genetic elements (MGE) that could potentially spread tetracycline resistance from the 10 *R. intestinalis* strains was investigated by specifically analyzing the genetic neighborhood surrounding the *tet*(O) and *tet*(40) genes. The putative functions of the

surrounding genes were identified through homology using BLASTp analysis of 10 open reading frames (ORFs) upstream and downstream of *tet*(O) and/or *tet*(40) genes. A schematic diagram representing these ORFs is presented in Figure 6. The transposon-encoded gene *tnpV* was located immediately upstream of the *tet*(O) gene in 8 of the 10 genomes. A gene encoding the conjugative transfer protein (CTP) was present immediately downstream of *tet*(O) genes in six of these genomes but was absent for the four strains containing the *tet*(40) gene. There was considerable conservation of the genetic neighborhood downstream of *tet*(O) (plus/minus *tet*(40)), particularly in isolates close together in the phylogenetic tree (Figure 2). Many of the genes identified have potential roles in gene mobility (helicases, relaxases and recombinases), suggesting the presence of a mobile genetic element. A Venn diagram (Supplementary Figure S6) was used to visualize the main features distinguishing the genome sequences. In strain AM43–11, the contig sequence containing the *tet*(O) and *tet*(40) genes was truncated immediately upstream of *tet*(O) so the presence of the *tnpV* gene could not be determined.

Antibiotic resistance phenotype confirmation

The antibiotic resistance phenotypes of the five strains in our culture collection were tested to verify the AMR genotypes. Disc assays confirmed that strain M50/1 (no tetracycline resistance gene) was sensitive to tetracycline, while the remaining four strains which all contained the *tet*(O) gene were resistant to tetracycline (30 µg/ml) (Table 3). Tests to determine the minimum inhibitory concentrations (MIC) further confirmed the sensitivity of strain M50/1 that could not grow in tetracycline concentrations exceeding 0.019 µg/ml. The MIC for strains XB6B4, L1–82 and PC352 ranged from 16 to 32 µg/ml while strain PC335 was able to grow up to 128 µg/ml. All but one of the strains tested were sensitive to erythromycin (15 µg/ml) with strain L1–82 showing intermediate sensitivity. This is consistent with the lack of *erm*(B) in these five isolates. MIC tests confirmed the sensitivity of the four isolates to erythromycin, while isolate L1–82,

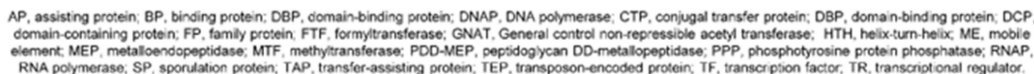


Figure 6. Comparison of the genetic environment of the *tet(O)* gene in *R. intestinalis* strains.

Table 3. Antimicrobial resistance profile of *Roseburia intestinalis* strains.

Antibiotic/Strain	M50/1	L1–82	XB6B4	PC335	PC352
Amp (10 µg/ml)	S	S	S	S	S
Van (30 µg/ml)	S	S	S	S	S
Strep* (10 µg/ml)	IS*	IS*	R*	R*	R*
Kan* (30 µg/ml)	IS*	IS*	IS*	IS*	R*
CN (30 µg/ml)	S	S	S	S	S
Erm (15 µg/ml)	S	IS	S	S	S
Tet (30 µg/ml)	S	R	IS	IS	R
Cip* (10 µg/ml)	IS*	IS*	R*	R*	R*
MTZ (5 µg/ml)	S	S	S	S	S
Chl (30 µg/ml)	S	S	S	S	S
DA (10 µg/ml)	S	S	S	S	S

*Intrinsic resistance reported in anaerobic bacteria.

Phenotypes based on the diameter of zones (mm) indicating sensitive are as follows: Resistance (R, 8–10 mm), intermediate susceptibility (IS, 11–15 mm), and susceptibility (>15 mm). Amp – Ampicillin; Van – Vancomycin; Strep – Streptomycin; Kan – Kanamycin; CN – Gentamycin; Erm – Erythromycin; Tet – Tetracycline; Cip – Ciprofloxacin; MTZ – Metronidazole; Chl – Chloramphenicol; DA – Clindamycin.

although lacking a defined *erm*(B) gene, again showed intermediate sensitivity with some breakthrough growth occurring up to 250 µg/ml. The only other resistances observed were to the antimicrobials for which anaerobic bacteria often have intrinsic resistance, namely Streptomycin, Kanamycin and Ciprofloxacin.

Sporulation capability in *R. intestinalis*

A core set of approximately 60 genes encompassing key genes involved in various stages of spore formation, maturation, and eventual germination governs the complex process of sporulation within the Bacillota phylum in the human gut.^{55,58,59} These genes include, but are not limited to, *spo0A* as a master regulator of sporulation initiation, *sigE*, *sigF*, *sigG* and *sigK* responsible for directing RNA polymerase toward sporulation-specific genes, and *cotH*, *cotY*, and *cotA* which contribute to spore coat assembly. Additionally, genes like *gerA*, *gerB*, and *gerC* associated with germination are vital for the spore's transition back into an active vegetative state.

The sporulation signature genes had significant hits in the *R. intestinalis* pangenome. Of these, 38 sporulation related genes were identified in the core genome across all 16 genomes and another seven were present in all but one strain (Supplementary Table S6). The master regulator *spo0A* gene along with all the key sporulation specific RNA polymerase sigma factors (*sigE*, *sigF*, *sigG* and *sigK*) were within the core genome for all 16 *R. intestinalis* genomes except that strain AM37-

1AC apparently lacked the *spo0A* gene. The majority of the conserved core sporulation genes described in other Bacillota were also present in the *R. intestinalis* genomes.^{55,58,59} A list of sporulation genes identified for the representative *R. intestinalis* strain PC352 is given in Table 4.

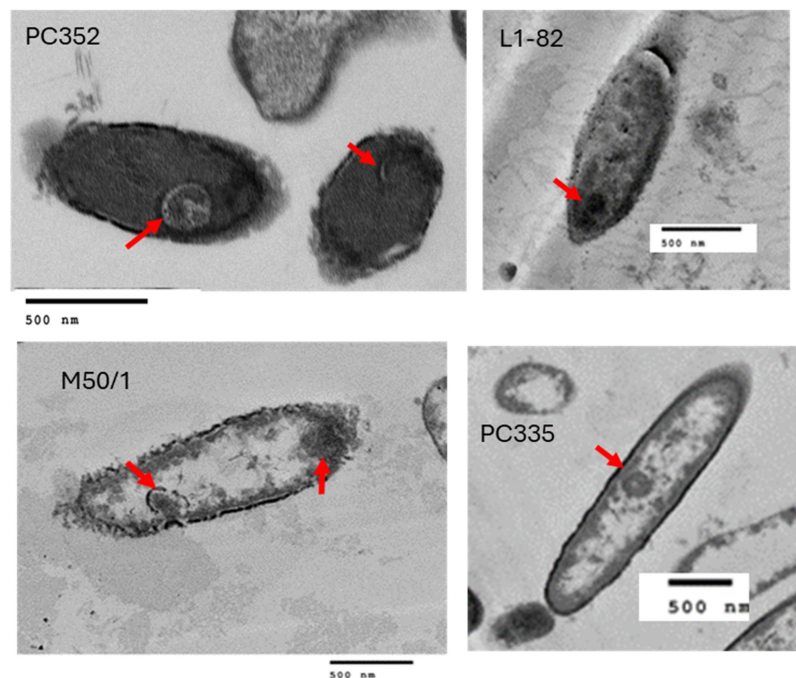
We tested the actual sporulation capability of *R. intestinalis* strains (L1–82, M50/1, PC335 and PC352) in laboratory tests. *R. intestinalis* cultures grown for 72 h were subjected to heat-shock and oxygen exposure to induce stress and attempt to trigger sporulation. Subsequently, cells were fixed and examined by TEM. Spores are clearly visible at the magnifications shown (Figure 7). Further work would be required to ascertain whether strains AM36-10AT and BIOML-A1, which contain the same 59 sporulation genes, also sporulate, and also the minimum number of genes required to enable sporulation.

Flagella and motility

Flagellum synthesis also requires multiple genes, which have been shown to be present in three motility gene loci in *R. intestinalis* L1–82.⁶⁰ All of the required genes were part of the core genome in all 16 *R. intestinalis* strains demonstrating their ability to construct complete flagella (Supplementary Table S7). The motility of strains PC335 and PC352 was confirmed using hanging drop method and light microscopy of live cells. A complete set of chemotaxis genes was identified in the core genome (Supplementary Table S8),

Table 4. Sporulation gene signatures in *R. intestinalis* and other spore forming bacteria.

Stages of sporulation and function	Sporulation gene signature (Browne et al, 2016, Nature)	Sporulation genes in <i>R. bromii</i> L2-63 (Mukhopadhyaya et al, 2018, Environ Microbiol)	Sporulation genes in <i>R. intestinalis</i> PC352 (This study)
Stage I Initiation of spore formation	<i>sigH</i> , <i>spo0A</i> , <i>parB</i> , <i>sepF</i> ,	<i>sigH</i> , <i>spo0A</i> , <i>spo0J</i> family (<i>parB_1</i> , <i>parB_2</i> , <i>parB_3</i> , <i>parB_4</i> , <i>parB_5</i>), <i>sepF</i>	<i>sigH</i> , <i>sigA</i> , <i>spo0A</i> , <i>spo0J</i> , <i>ParB</i> , <i>whiA</i> , <i>sepF</i>
Stage II Asymmetric cell division	<i>sigE</i> , <i>sigF</i> , <i>mcsA</i> , <i>spolIGA</i> , <i>spolIR</i> , <i>spolp</i> , <i>ykvl</i> ,	<i>sigE</i> , <i>sigF</i> , <i>spolIGA</i> , <i>spolIAA_1</i> , <i>spolIAA_2</i> , <i>spolIR</i> , <i>spolIQ</i> , <i>spolIE</i> , <i>rsbU</i> , <i>spolIP</i> , L2-63_01395	<i>sigE</i> , <i>sigF</i> , <i>spolIE</i> , , <i>spolIGA</i> , <i>spolIP</i> , <i>spolIQ</i> , <i>spolIR</i>
Stage III Engulfment of forespore by mother cell	<i>spolIAB</i> , <i>spolIAC</i> , <i>spolIAD</i> , <i>spolIAE</i> , <i>spolIAG</i> , <i>spolIAH</i> , <i>spolIAA</i> , <i>ftsQ</i> , <i>sigG</i> , CD630_10450, CD630_20350	<i>sigG</i> , <i>spolIAG</i> , <i>spolIAE</i> , <i>spolIAD</i> , <i>spolIAC</i> , <i>spolIAB</i> , <i>spolIE</i> , <i>spolIE</i> , L2-63_01811	<i>sigG</i> , <i>spolIAB</i> , <i>spolIAC</i> , <i>spolIAD</i> , <i>spolIAE</i> , <i>spolIAF</i> , <i>spolIAH</i> , <i>spolID</i>
Stage IV Formation of cortex	<i>sigK</i> , <i>spoIVA</i> , <i>spoIVB</i> , <i>spoIVFB</i> , <i>YlbJ</i> , <i>dacF1</i> , <i>dacF</i> , <i>dapG</i>	<i>sigK</i> , <i>spoIVA</i> , <i>spoIVB</i> , <i>yqfD</i> , <i>YlbJ</i>	<i>sigK</i> , <i>spoIVA</i> , <i>spoIVB</i> , <i>ylbJ</i> , <i>ytfJ</i>
Stage V Formation of coat	<i>spoVAD</i> , <i>spoVE</i> , <i>spoVFA</i> , <i>cotG</i> , <i>cotCB</i> , <i>sipL</i> , <i>soda</i> , <i>spoVT</i> , CD630_34940, CD630_35690	<i>spoVAE</i> , <i>spoVAC</i> , <i>SpoVAD</i> , <i>spoVFA</i> , <i>spoVTytgP</i> , <i>ftsI</i> , <i>ftsW_2</i> , <i>ugtp</i> , <i>cotJA</i> , <i>safa</i>	<i>spoVAA</i> , <i>spoVAD</i> , <i>spoVA</i> , <i>spoVB</i> , <i>spoVD</i> , <i>spoVT</i> , <i>spoVG</i> , <i>cotJB</i> , <i>yabP</i> , <i>yabQ</i> , <i>cotS</i>
Germination of spore	<i>gpr</i> , <i>spmA</i> , <i>spmB</i> , <i>tepA</i>	<i>gpr</i> , <i>gerA</i> , <i>gerC</i> , <i>sleB</i> , <i>cwlJ</i> , <i>spmA</i> , <i>spmB</i> , L2-63_01843	<i>gpr</i> , <i>gerA</i> , <i>sleB</i> , <i>sleC</i> , <i>cwlD</i> , <i>cwlA</i> , <i>cwlK</i> , <i>spmA</i> , <i>spmB</i> , <i>yndE</i> , <i>prkc</i>
Sporulation initiation inhibitor	—	<i>soj_1</i> , <i>soj_2</i> , <i>soj_3</i> , <i>KapD</i>	<i>soj_1</i> , <i>soj_2</i> , <i>soj_3</i> , <i>soj_4</i> , <i>soj_5</i> , <i>soj_6</i>
Genes not associated with any specific stage of sporulation	<i>asrC</i> , <i>dnaA</i> , <i>PHP domain</i> , <i>bltD</i> , <i>pdp</i> , <i>dnaA</i> , <i>ylxY</i> , <i>rplJ</i> , <i>trxB</i> , <i>ybbH</i> , <i>walR</i> , <i>purB</i> , <i>zupT</i> , CD630_25890, 22730, 05720, 01290, 20480, 00550, 21080, 08810, 24430	<i>paiA</i> , <i>yebN</i> , <i>xynD</i> , <i>ytfJ</i> , <i>YlmC</i> , <i>yunB</i> , <i>yhhT</i> , <i>yabP</i> , <i>whiA</i> , <i>yqfc</i> , <i>sspC2</i> , L2-63_00863	<i>sspD</i> , <i>ykvp</i>

**Figure 7.** Transmission electron microscopy images showing spores present (red arrows) in individual cells of four *R. intestinalis* strains.

adjacent to one of the motility loci, indicating the co-dependence of these two bacterial traits.

Discussion

Roseburia intestinalis is an important butyrate producing bacterial species, abundant in the human large intestine, with potential beneficial effects in inflammatory bowel disease, colorectal cancer, atherosclerosis and metabolic syndrome, and also a possible role as a next generation probiotic.^{29,61} This study is the first specific pangenome analysis of *R. intestinalis* using 16 known human isolates of *R. intestinalis* to provide a comprehensive assessment of the metabolic function of this species. The genome analysis showed clustering of strains in two clades from Asia and Europe with a possible further American clade, illustrating potential geographical stratification. This type of stratification has been demonstrated and studied extensively in bacteria such as *Helicobacter pylori* where phenotypic and genotypic differences have been noted among bacterial strains from large datasets from different continents, reflecting evolutionary dynamics, human diet, and human migration.⁶²

The genomes of the 16 *R. intestinalis* strains encoded between 97 and 123 glycoside hydrolases (GHs), more than the 70 previously reported.²⁸ Five of these GHs were conserved across all strains in the core genome and were involved in starch, xylan, and mannose utilization. Growth assays confirmed the functionality of these genes in the five strains tested, and the presence of these GHs is consistent with earlier reports on substrate utilization of these bacteria.^{21,63} The combination of substrate binding and GH domains in multi-domain representatives of these genes in all isolates further indicates the importance of starch, xylan and mannose as key substrates for *R. intestinalis* growth and survival in the large intestine. There was no obvious reason from the genome information for the poorer ability of strains PC335 and PC352 to utilize the potato and corn starch for growth compared to other isolates.

The large number of CAZymes in the genomes, frequently present in multifunctional gpPULs, also highlights the dependence of these bacteria on variable substrate degradation, facilitating cross-feeding of other bacterial strains in the gut lumen

and maintaining a healthy gut ecology.^{12,57} Studies like this, identifying and confirming key growth substrates are also crucial in determining potential prebiotic and dietary approaches to enhance populations of these important resident bacteria *in vivo*.

Strain-specific growth on rhamnose was unique to the most recently isolated PC335 strain. This was also the only cultured strain containing the unique rhamnose gpPUL 2. The other cultured strains (and indeed all the other strains) all contained rhamnose gpPUL 1, which was also identified as mannose gpPUL 1 and may thus be involved in mannose utilization. It is possible that rhamnose gpPUL 2 is present in strain PC335 (and strains AM37-1AC and SNUG30017) that contains two GH78 enzymes, three regulatory proteins and a chemotaxis protein is more specific for rhamnose than the other GH78 enzymes detected during the genome analysis. Growth on rhamnose only occurred after the strain was repeatedly subcultured on this substrate, implying that this is an inducible trait. The related bacterium, *Roseburia inulinivorans*, was previously shown to differentially express genes for fucose utilization, accompanied by a shift to propionate production,⁶⁴ although propionate was not detected following growth of strain PC335 on rhamnose. The presence of these specific CAZy domains highlights the adaptive growth capabilities of these bacteria to target rhamnose, a less common but nutritionally relevant carbohydrate, as part of their metabolic strategies. This novel finding may reflect the current use of rhamnose as a sweetener and anti-obesity agent and indicate bacterial evolution following shifts in human dietary predilections in recent years.⁶⁵

In the set of *R. intestinalis* strains analyzed in this study, tetracycline resistance genes were widespread, identified in 10 out of 16 strains. All these genes were present in the shell of the pangenome, and some were in close proximity to transposon coding gene *tnpV*, and other mobility genes suggesting that they were acquired horizontally from other bacterial species. Tetracycline resistance *tet* (40) genes were located in tandem with *tet*(O) in four strains, including three of the six Asian strains. *Erm*(B) was only identified in three strains, all Asian, two of which also included *tet*(O) and *tet* (40). This higher presence of resistance genes could

reflect a more widespread use of antibiotics in these populations, although this, in addition to determining whether geographical clustering is linked to dietary differences, could only be confirmed if associated metadata was available. A comparative analysis of genes flanking the *tet(W)* resistance genes in various anaerobic gut bacteria including *Roseburia hominis* A2–183 also suggested the presence of conserved sequences potentially comprising a mobile mini-element, indicating the role of horizontal gene transfer in the acquisition of tetracycline resistance genes.⁶⁶

Antibiotic sensitivity testing confirmed that the four strains available were all able to grow in the presence of tetracycline while strain M50/1 (lacking the resistance genes) was sensitive to tetracycline. The observed MIC ranged from 16 to 128 µg/ml with co-carriage of *tet(40)* not appearing to confer additional resistance. A separate study on *R. intestinalis* L1–82 strain also identified *tet(O)* and *tet(40)* resistance genes but, in contrast to our data, these authors deemed the strain sensitive to tetracycline and erythromycin using the disk diffusion method.²⁹ This highlights the challenges in assessing antibiotic resistance in bacterial strains and is especially pertinent as an expert panel have listed *Roseburia* as one of several probiotic candidate species.⁶⁷ The identification of functional tetracycline resistance genes, potentially encoded on mobile genetic elements, in multiple but not all strains indicate the importance of careful evaluation of individual strains for their suitability as probiotic and live-biotherapeutic candidates to minimize the risk of transferring resistance to other resident gut bacterial populations. Understanding the role of commensal bacteria as reservoirs of (potentially transferable) antimicrobial resistance genes provides critical information for global antimicrobial stewardship.

This study confirmed the presence of flagellin proteins and mobility phenotype in these strains. Flagella are also immunomodulatory molecules, and purified flagellin proteins from a different *Roseburia* species (*R. inulinivorans* A2–194) and the related bacterium *Eubacterium rectale* A1–86 stimulated secretion of the pro-inflammatory cytokine IL-8.⁶⁸ However, motility is also an important feature for successful colonization and substrate utilization within the gut ecosystem, further emphasized by the co-location of chemotaxis

genes adjacent to the flagellin synthesis genes. Flagellin proteins in *R. inulinivorans* A2–194 were upregulated during growth on starch compared to inulin.⁶⁹

Another important finding of this study was the detection of a sporulation signature in the genomes of all 16 *R. intestinalis* strains with the majority of the genes present in the core genome. This suggests that spore formation is an intrinsic trait of this species and we confirmed the presence of spores in strain four strains by transmission electron microscopy. The *spoOA* gene was not detected in strain AM37-1AC, but this may be due to the relative incompleteness of this genome sequence. Spore formation in this anaerobic species can explain their environmental survival, host-to-host transmission and inheritability within families.⁷⁰ The ability to form spores has significant implications in potential marketing of *R. intestinalis* strains as probiotics, providing alternative delivery options than utilization of viable anaerobic bacteria.

This comprehensive assessment of *R. intestinalis* strains through both genotypic and phenotypic characterization with specific focus on antimicrobial resistance, carbohydrate utilization and sporulation has shown distinct geographical evolution and gene acquisition profiles. Our results demonstrate that together, sporulation and specific dietary components (namely xylan, starch and mannose) are key factors in maintaining populations of these important butyrate producers in the human gut microbiota. This study also underscores the critical importance of validating phenotypic characteristics inferred from genomic information and emphasizes that the mere presence of genes does not guarantee their functionality. Ultimately, this improved understanding of carbohydrate utilization and sporulation in this important group of butyrate producing bacteria has significant implications for both dietary and probiotic interventions aimed at enhancing gut health.

Acknowledgments

We are indebted to our volunteers for providing the faecal and biopsy samples without which this study would not have been possible. We thank the members of the Rowett Gut Health research team for discussions and advice. The authors thank

the Centre for Genome Enabled Biology and Medicine specifically Dr Jin Pu for Illumina and Oxford Nanopore sequencing and useful discussions. Microscopy was performed in the Microscopy and Histology Core Facility at the University of Aberdeen.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by funding from Probi AB [Grant Ref: RG14104]. The Rowett Institute (University of Aberdeen) receives financial support from the Scottish Government Rural and Environmental Sciences and Analytical Services (RESAS).

ORCID

Indrani Mukhopadhyay  <http://orcid.org/0000-0003-2577-518X>

Authors' contributions

KPS, SWG and IM conceived the study. AJM, KPS and SWG collected all volunteer samples. IM and JM performed all sample processing, experiments and analyses. NB did AMR testing and SCFA analysis for the five selected strains. SS conducted pan-genome data analysis. MGT helped with AMR gene analysis and drafted figure for it. IM, KPS and SWG interpreted the data and wrote the manuscript. All authors read and approved the final manuscript.

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAFVC000000000 and JAFVVD000000000. The version described in this paper is version XXXXXX010000000.

Consent to participate

All subjects provided signed informed consent before participation.

Ethics approval

The study was performed in accordance with the principles of the Declaration of Helsinki and the study protocol was approved by the Ethics Committee of North of Scotland Research Ethics Service (Reference 17/NS/0112) and Rowett Human Studies Ethical Review Panel.

References

1. Kobozev I, Reinoso Webb C, Furr KL, Grisham MB. Role of the enteric microbiota in intestinal homeostasis and inflammation. *Free Radic Biol Med*. 2014;68:122–133. doi:10.1016/j.freeradbiomed.2013.11.008.
2. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto J-M, et al. Enterotypes of the human gut microbiome. *Nature*. 2011;473(7346):174–180. doi:10.1038/nature09944.
3. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*. 2010;464(7285):59–65. doi:10.1038/nature08821.
4. Louis P, Hold GL, Flint HJ. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat Rev Microbiol*. 2014;12(10):661–672. doi:10.1038/nrmiicro3344.
5. Louis P, Flint HJ. Formation of propionate and butyrate by the human colonic microbiota. *Environ Microbiol*. 2017;19:29–41. doi:10.1111/1462-2920.13589.
6. Hamer HM, Jonkers D, Venema K, Vanhoutvin S, Troost FJ, Brummer R-J. Review article: the role of butyrate on colonic function. *Aliment Pharmacol Ther*. 2008;27(2):104–119. doi:10.1111/j.1365-2036.2007.03562.x.
7. Singh N, Gurav A, Sivaprakasam S, Brady E, Padia R, Shi H, Thangaraju M, Prasad P, Manicassamy S, Munn D, et al. Activation of Gpr109a, receptor for niacin and the commensal metabolite butyrate, suppresses colonic inflammation and carcinogenesis. *Immunity*. 2014;40(1):128–139. doi:10.1016/j.immuni.2013.12.007.
8. Fung KY, Cosgrove L, Lockett T, Head R, Topping DL. A review of the potential mechanisms for the lowering of colorectal oncogenesis by butyrate. *Br J Nutr*. 2012;108(5):820–831. doi:10.1017/S0007114512001948.
9. Wang HB, Wang PY, Wang X, Wan YL, Liu YC. Butyrate enhances intestinal epithelial barrier function via up-regulation of tight junction protein claudin-1 transcription. *Dig Dis Sci*. 2012;57(12):3126–3135. doi:10.1007/s10620-012-2259-4.
10. Kazemi Sefat NA, Mohammadi MM, Hadjati J, Talebi S, Ajami M, Daneshvar H. Sodium butyrate as a histone deacetylase inhibitor affects toll-like receptor 4 expression in colorectal cancer cell lines. *Immunol Invest*. 2019;48(7):759–769. doi:10.1080/08820139.2019.1595643.
11. Barcenilla A, Pryde SE, Martin JC, Duncan SH, Stewart CS, Henderson C, Flint HJ. Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Appl Environ Microbiol*. 2000;66(4):1654–1661. doi:10.1128/AEM.66.4.1654-1661.2000.

12. Louis P, Flint HJ. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol Lett.* **2009**;294(1):1–8. doi:[10.1111/j.1574-6968.2009.01514.x](https://doi.org/10.1111/j.1574-6968.2009.01514.x).
13. Duncan SH, Aminov RI, Scott KP, Louis P, Stanton TB, Flint HJ. Proposal of *Roseburia faecis* sp. nov. *Roseburia hominis* sp. nov. And *Roseburia inulinivorans* sp. nov. based on isolates from human faeces. *Int J Syst Evol Microbiol.* **2006**;56(10):2437–2441. doi:[10.1099/ijms.0.64098-0](https://doi.org/10.1099/ijms.0.64098-0).
14. Parada Venegas D, De la Fuente MK, Landskron G, González MJ, Quera R, Dijkstra G, Harmsen HJM, Faber KN, Hermoso MA. Short chain fatty acids (SCFAs)-mediated gut epithelial and immune regulation and its relevance for inflammatory bowel diseases. *Front Immunol.* **2019**;10:277. doi:[10.3389/fimmu.2019.00277](https://doi.org/10.3389/fimmu.2019.00277).
15. Paramsothy S, Nielsen S, Kamm MA, Deshpande NP, Faith JJ, Clemente JC, Paramsothy R, Walsh AJ, van den Bogaerde J, Samuel D, et al. Specific bacteria and metabolites associated with response to fecal microbiota transplantation in patients with ulcerative colitis. *Gastroenterology.* **2019**;156(5):1440–1454.e2. doi:[10.1053/j.gastro.2018.12.001](https://doi.org/10.1053/j.gastro.2018.12.001).
16. Tamañai-Shacoory Z, Smida I, Bousarghin L, Loreal O, Meuric V, Fong SB, Bonnaure-Mallet M, Jolivet-Gougeon A. *Roseburia* spp.: a marker of health? *Future Microbiol.* **2017**;12(2):157–170. doi:[10.2217/fmb-2016-0130](https://doi.org/10.2217/fmb-2016-0130).
17. Kumari R, Ahuja V, Paul J. Fluctuations in butyrate-producing bacteria in ulcerative colitis patients of North India. *World J Gastroenterol.* **2013**;19:3404–3414. doi:[10.3748/wjg.v19.i22.3404](https://doi.org/10.3748/wjg.v19.i22.3404).
18. Geng J, Fan H, Tang X, Zhai H, Zhang Z. Diversified pattern of the human colorectal cancer microbiome. *Gut Pathog.* **2013**;5(1):2. doi:[10.1186/1757-4749-5-2](https://doi.org/10.1186/1757-4749-5-2).
19. Kasahara K, Krautkramer KA, Org E, Romano KA, Kerby RL, Vivas EI, Mehrabian M, Denu JM, Bäckhed F, Lusi AJ, et al. Interactions between *Roseburia intestinalis* and diet modulate atherogenesis in a murine model. *Nat Microbiol.* **2018**;3(12):1461–1471. doi:[10.1038/s41564-018-0272-x](https://doi.org/10.1038/s41564-018-0272-x).
20. Kang X, Liu C, Ding Y, Ni Y, Ji F, Lau HCH, Jiang L, Sung JJ, Wong SH, Yu J. *Roseburia intestinalis* generated butyrate boosts anti-PD-1 efficacy in colorectal cancer by activating cytotoxic CD8⁺ T cells. *Gut.* **2023**;72(11):2112–2122. doi:[10.1136/gutjnl-2023-330291](https://doi.org/10.1136/gutjnl-2023-330291).
21. Duncan SH, Hold GL, Barcenilla A, Stewart CS, Flint HJ. *Roseburia intestinalis* sp. nov. a novel saccharolytic, butyrate-producing bacterium from human faeces. *Int J Syst Evol Microbiol.* **2002**;52(5):1615–1620. doi:[10.1099/00207713-52-5-1615](https://doi.org/10.1099/00207713-52-5-1615).
22. Shen Z, Zhu C, Quan Y, Yang J, Yuan W, Yang Z, Wu S, Luo W, Tan B, Wang X. Insights into *Roseburia intestinalis* which alleviates experimental colitis pathology by inducing anti-inflammatory responses. *J Gastro Hepatol.* **2018**;33(10):1751–1760. doi:[10.1111/jgh.14144](https://doi.org/10.1111/jgh.14144).
23. Zhu C, Song K, Shen Z, Quan Y, Tan B, Luo W, Wu S, Tang K, Yang Z, Wang X. *Roseburia intestinalis* inhibits interleukin-17 excretion and promotes regulatory T cells differentiation in colitis. *Mol Med Report.* **2018**;17:7567–7574. doi:[10.3892/mmr.2018.8833](https://doi.org/10.3892/mmr.2018.8833).
24. Luo W, Shen Z, Deng M, Li X, Tan B, Xiao M, Wu S, Yang Z, Zhu C, Tian L, et al. *Roseburia intestinalis* supernatant ameliorates colitis induced in mice by regulating the immune response. *Mol Med Report.* **2019**;20:1007–1016. doi:[10.3892/mmr.2019.10327](https://doi.org/10.3892/mmr.2019.10327).
25. Sun X, Zhou C, Ju M, Feng W, Guo Z, Qi C, Yang K, Xiao R. *Roseburia intestinalis* supplementation could reverse the learning and memory impairment and m6A methylation modification decrease caused by 27-hydroxycholesterol in mice. *Nutrients.* **2024**;16(9):1288. doi:[10.3390/nu16091288](https://doi.org/10.3390/nu16091288).
26. Aminov RI, Walker AW, Duncan SH, Harmsen HJM, Welling GW, Flint HJ. Molecular diversity, cultivation, and improved detection by fluorescent in situ hybridization of a dominant group of human gut bacteria related to *roseburia* spp. Or eubacterium rectale. *Appl Environ Microbiol.* **2006**;72(9):6371–6376. doi:[10.1128/AEM.00701-06](https://doi.org/10.1128/AEM.00701-06).
27. Sheridan PO, Martin JC, Lawley TD, Browne HP, Harris HMB, Bernalier-Donadille A, Duncan SH, O'Toole PW, Scott PK, Flint JH. Polysaccharide utilization loci and nutritional specialization in a dominant group of butyrate-producing human colonic firmicutes. *Microb Genomics.* **2016**;2(2). doi:[10.1099/mgen.0.000043](https://doi.org/10.1099/mgen.0.000043).
28. Hillman ET, Kozik AJ, Hooker CA, Burnett JL, Heo Y, Kiesel VA, Nevins CJ, Oshiro JMKI, Robins MM, Thakkar RD, et al. Comparative genomics of the genus *Roseburia* reveals divergent biosynthetic pathways that may influence colonic competition among species. *Microb Genomics.* **2020**;6(7). doi:[10.1099/mgen.0.000399](https://doi.org/10.1099/mgen.0.000399).
29. Zhang C, Ma K, Nie K, Deng M, Luo W, Wu X, Huang Y, Wang X. Assessment of the safety and probiotic properties of *Roseburia intestinalis*: a potential “next generation probiotic”. *Front Microbiol.* **2022**;13:973046. doi:[10.3389/fmicb.2022.973046](https://doi.org/10.3389/fmicb.2022.973046).
30. Mukhopadhyaya I, Martin JC, Shaw S, McKinley AJ, Gratz SW, Scott KP. Comparison of microbial signatures between paired faecal and rectal biopsy samples from healthy volunteers using next-generation sequencing and culturomics. *Microbiome.* **2022**;10(1):171. doi:[10.1186/s40168-022-01354-4](https://doi.org/10.1186/s40168-022-01354-4).
31. Duncan SH, Hold GL, Barcenilla A, Stewart CS, Flint HJ. *Roseburia intestinalis* sp. nov. a novel saccharolytic, butyrate-producing bacterium from human faeces. *Int J Syst Evol Microbiol.* **2002**;52(5):1615–1620. doi:[10.1099/00207713-52-5-1615](https://doi.org/10.1099/00207713-52-5-1615).
32. Louis P, Duncan SH, McCrae SI, Millar J, Jackson MS, Flint HJ. Restricted distribution of the butyrate kinase

- pathway among butyrate-producing bacteria from the human colon. *J Bacteriol.* **2004**;186(7):2099–2106. doi:10.1128/JB.186.7.2099-2106.2004.
33. Chassard C, Goumy V, Leclerc M, Del'homme C & Bernalier-Donadille A. Characterization of the xylan-degrading microbial community from human faeces. *FEMS Microbiol Ecol.* **2007**;61(1):121–131. doi:10.1111/j.1574-6941.2007.00314.x.
 34. Duncan SH, Hold GL, Harmsen HJ, Stewart CS, Flint HJ. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov. comb. nov. *Int J Syst Evol Microbiol.* **2002**;52(6):2141–2146. doi:10.1099/00207713-52-6-2141.
 35. Wick RR, Judd LM, Gorrie CL, Holt KE, Phillippy AM. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Comput Biol.* **2017**;13(6):e1005595. doi:10.1371/journal.pcbi.1005595.
 36. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics.* **2013**;29(8):1072–1075. doi:10.1093/bioinformatics/btt086.
 37. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics.* **2016**;32(2):292–294. doi:10.1093/bioinformatics/btv566.
 38. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* **2015**;31(19):3210–3212. doi:10.1093/bioinformatics/btv351.
 39. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* **2014**;30(14):2068–2069. doi:10.1093/bioinformatics/btu153.
 40. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**;47(D1):D427–D432. doi:10.1093/nar/gky995.
 41. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics.* **2015**;31(22):3691–3693. doi:10.1093/bioinformatics/btv421.
 42. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* **2014**;30(9):1312–1313. doi:10.1093/bioinformatics/btu033.
 43. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen C-Y, Graham M, Van Domselaar G, Stothard P. Proksee: in-depth characterization and visualization of bacterial genomes. *Nucleic Acids Res.* **2023**;51(W1):W484–W492. doi:10.1093/nar/gkad326.
 44. Zhang H, Yohe T, Huang L, Entwistle S, Wu P, Yang Z, Busk PK, Xu Y, Yin Y. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **2018**;46:W95–W101. doi:10.1093/nar/gky418.
 45. Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends Genet.* **2000**;16(6):276–277. doi:10.1016/S0168-9525(00)02024-2.
 46. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen ALV, Cheng AA, Liu S, et al. CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **2020**;48(D1):D517–D525. doi:10.1093/nar/gkz935.
 47. Siguier P, Pérochon J, Lestrade L, Mahillon J, Chandler M. Isfinder: the reference centre for bacterial insertion sequences. *Nucleic Acids Res.* **2006**;34(90001):D32–D36. doi:10.1093/nar/gkj014.
 48. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinf.* **2009**;10(1):1–9. doi:10.1186/1471-2105-10-421.
 49. Galperin MY, Kristensen DM, Makarova KS, Wolf YI, Koonin EV. Microbial genome analysis: the COG approach. *Briefings Bioinf.* **2019**;20(4):1063–1070. doi:10.1093/bib/bbx117.
 50. Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the artemis comparison tool. *Bioinformatics.* **2005**;21(16):3422–3423. doi:10.1093/bioinformatics/bti553.
 51. Richardson A, Calder A, Stewart C, Smith A. Simultaneous determination of volatile and non-volatile acidic fermentation products of anaerobes by capillary gas chromatography. *Lett Appl Microbiol.* **1989**;9(1):5–8. doi:10.1111/j.1472-765X.1989.tb00278.x.
 52. Additives EPO, Feed POSUIA. Guidance on the assessment of bacterial susceptibility to antimicrobials of human and veterinary importance. *Efsa J.* **2012**;10:2740.
 53. Kirby W, Bauer A, Sherris J, Turck M. Antibiotic sensitivity testing by a standardized disc diffusion method. *Am J Clin Pathol.* **1966**;45(4-ts):493–496. doi:10.1093/ajcp/45.4_ts.493.
 54. Matuschek E, Brown DF, Kahlmeter G. Development of the EUCAST disk diffusion antimicrobial susceptibility testing method and its implementation in routine microbiology laboratories. *Clin Microbiol Infect.* **2014**;20(4):O255–O266. doi:10.1111/1469-0691.12373.
 55. Mukhopadhyaya I, Morais S, Laverde-Gomez J, Sheridan PO, Walker AW, Kelly W, Klieve AV, Ouwerkerk D, Duncan SH, Louis P, et al. Sporulation capability and amylosome conservation among diverse human colonic and rumen isolates of the keystone starch-degrader *Ruminococcus bromii*. *Environ Microbiol.* **2018**;20(1):324–336. doi:10.1111/1462-2920.14000.
 56. Jain A, Jain R, Jain S. Motility Testing – Hanging Drop Method and Stab. In: *Basic Techniques in Biochemistry, Microbiology and Molecular Biology*. Humana, New York, NY: Springer Protocols

- Handbooks; 2020. p. 121–122. doi:[10.1007/978-1-4939-9861-6_34](https://doi.org/10.1007/978-1-4939-9861-6_34).
57. Bhattacharya A, Majtorp L, Birgersson S, Wiemann M, Sreenivas K, Verbrugghe P, Van Aken O, Van Niel E, Ståhlbrand H. Cross-feeding and enzymatic catabolism for mannan-oligosaccharide utilization by the butyrate-producing gut bacterium *roseburia hominis* A2-183. *Microorganisms*. 2022;10(12):2496. doi:[10.3390/microorganisms10122496](https://doi.org/10.3390/microorganisms10122496).
 58. Browne HP, Forster SC, Anonye BO, Kumar N, Neville BA, Stares MD, Goulding D, Lawley TD. Culturing of ‘unculturable’ human microbiota reveals novel taxa and extensive sporulation. *Nature*. 2016;533(7604):543–546. doi:[10.1038/nature17645](https://doi.org/10.1038/nature17645).
 59. Galperin MY, Yutin N, Wolf YI, Vera Alvarez R, Koonin EV, Shank EA. Conservation and evolution of the sporulation gene set in diverse members of the firmicutes. *J Bacteriol*. 2022;204(6):e00079–00022. doi:[10.1128/jb.00079-22](https://doi.org/10.1128/jb.00079-22).
 60. Neville BA. Functional genomics of motile commensal intestinal bacteria. 2013.
 61. Zhang H, Duan Y, Cai F, Cao D, Wang L, Qiao Z, Hong Q, Li N, Zheng Y, Su M, et al. Next-generation probiotics: microflora intervention to human diseases. *Biomed Res Int*. 2022;2022(1):5633403. doi:[10.1155/2022/5633403](https://doi.org/10.1155/2022/5633403).
 62. Kodaman N, Pazos A, Schneider BG, Piazzuelo MB, Mera R, Sobota RS, Sicinschi LA, Shaffer CL, Romero-Gallo J, de Sablet T, et al. Human and *Helicobacter pylori* coevolution shapes the risk of gastric disease. *Proc Natl Acad Sci USA*. 2014;111(4):1455–1460. doi:[10.1073/pnas.1318093111](https://doi.org/10.1073/pnas.1318093111).
 63. La Rosa SL, Leth ML, Michalak L, Hansen ME, Pudlo NA, Glowacki R, Pereira G, Workman CT, Arntzen MØ, Pope PB, et al. The human gut firmicute *Roseburia intestinalis* is a primary degrader of dietary β -mannans. *Nat Commun*. 2019;10(1):1–14. doi:[10.1038/s41467-019-08812-y](https://doi.org/10.1038/s41467-019-08812-y).
 64. Scott KP, Martin JC, Campbell G, Mayer C-D, Flint HJ. Whole-genome transcription profiling reveals genes up-regulated by growth on fucose in the human gut bacterium “*Roseburia inulinivorans*”. *J Bacteriol*. 2006;188(12):4340–4349. doi:[10.1128/JB.00137-06](https://doi.org/10.1128/JB.00137-06).
 65. Lv S, Hu T, Zhang R, Zhou Y, Yu W, Wang Z, Shi C, Lian J, Huang S, Pei G, et al. Rhamnose displays an anti-obesity effect through stimulation of adipose dopamine receptors and thermogenesis. *Diabetes*. 2023;72(3):326–335. doi:[10.2337/db22-0552](https://doi.org/10.2337/db22-0552).
 66. Kazimierczak KA, Flint HJ, Scott KP. Comparative analysis of sequences flanking tet(w) resistance genes in multiple species of gut bacteria. *Antimicrob Agents Chemother*. 2006;50(8):2632–2639. doi:[10.1128/AAC.01587-05](https://doi.org/10.1128/AAC.01587-05).
 67. Hill C, Guarner F, Reid G, Gibson GR, Merenstein DJ, Pot B, Morelli L, Canani RB, Flint HJ, Salminen S, et al. Expert consensus document. The international scientific association for probiotics and prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol*. 2014;11(8):506–514. doi:[10.1038/nrgastro.2014.66](https://doi.org/10.1038/nrgastro.2014.66).
 68. Neville BA, Sheridan PO, Harris HMB, Coughlan S, Flint HJ, Duncan SH, Jeffery IB, Claesson MJ, Ross RP, Scott KP, et al. Pro-inflammatory flagellin proteins of prevalent motile commensal bacteria are variably abundant in the intestinal microbiome of elderly humans. *PLOS ONE*. 2013;8(7):e68919. doi:[10.1371/journal.pone.0068919](https://doi.org/10.1371/journal.pone.0068919).
 69. Scott KP, Martin JC, Chassard C, Clerget M, Potrykus J, Campbell G, Mayer C-D, Young P, Rucklidge G, Ramsay AG, et al. Substrate-driven gene expression in *Roseburia inulinivorans*: importance of inducible enzymes in the utilization of inulin and starch. *Proc Natl Acad Sci USA*. 2011;108(supplement_1):4672–4679. doi:[10.1073/pnas.1000091107](https://doi.org/10.1073/pnas.1000091107).
 70. Attar N. Microbiome: spore formation in the human gut microbiota. *Nat Rev Microbiol*. 2016;14(7):403. doi:[10.1038/nrmicro.2016.75](https://doi.org/10.1038/nrmicro.2016.75).