



OPEN

## Identification of mucin degraders of the human gut microbiota

Stefano Raimondi<sup>1,3</sup>, Eliana Musmeci<sup>1,3</sup>, Francesco Candelieri<sup>1</sup>, Alberto Amaretti<sup>1,2</sup> & Maddalena Rossi<sup>1,2</sup>✉

Mucins are large glycoproteins consisting of approximately 80% of hetero-oligosaccharides. Gut mucin degraders of healthy subjects were investigated, through a culture dependent and independent approach. The faeces of five healthy adults were subjected to three steps of anaerobic enrichment in a medium with sole mucins as carbon and nitrogen sources. The bacterial community was compared before and after the enrichment by 16S rRNA gene profiling. Bacteria capable of fermenting sugars, such as *Anaerotruncus*, *Holdemania*, and Enterococcaceae likely took advantage of the carbohydrate chains. *Escherichia coli* and Enterobacteriaceae, Peptococcales, the Coriobacteriale *Eggerthella*, and a variety of Clostridia such as Oscillospiraceae, *Anaerotruncus*, and *Lachnospirillum*, significantly increased and likely participated to the degradation of the protein backbone of mucin. The affinity of *E. coli* and Enterobacteriaceae for mucin may facilitate the access to the gut mucosa, promoting gut barrier damage and triggering systemic inflammatory responses. Only three species of strict anaerobes able to grow on mucin were isolated from the enrichments of five different microbiota: *Clostridium disporicum*, *Clostridium tertium*, and *Paraoclostridium benzoelyticum*. The limited number of species isolated confirms that in the gut the degradation of these glycoproteins results from cooperation and cross-feeding among several species exhibiting different metabolic capabilities.

Mucus is a complex gel barrier that covers the wet epithelial surfaces throughout the body, including the gastrointestinal tract, and offers protection against exogenous and endogenous aggressive agents<sup>1,2</sup>. It exerts a variety of functions such as lubrication, hydration, chemical protection, sensing, nutrient reservoir, and barrier against pathogen invasion. A continuous turnover, consisting in a dynamic balance between biosynthesis, secretion, and degradation of its structural components, is crucial for these functions<sup>3</sup>.

Mucins are major component of mucus, made up by glycoproteins with high level of glycosylation. Residues of galactose, N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), fucose, and sialic acid, with relatively small amounts of mannose, constitute the oligosaccharides which represent approx. 80% of mucin mass<sup>4</sup>. The protein core is mainly organized in tandem repeated regions rich in serine, threonine, and proline. Serine and threonine are the site where an O-glycosidic bound links the protein core to GalNAc, the first moiety of the glycan chain<sup>5,6</sup>.

Among all the sites of human body inhabited by a resident microbial community, the colon hosts the most complex and concentrated microbiota<sup>7</sup>. In this site, mucins behave as decoy molecules that avoid the interaction of bacterial adhesins with receptors of the colonic epithelium. In fact, the gut lumen is coated by a two-layered mucus system, with an inner dense stratum firmly attached to the epithelium and an outer one, looser and unattached<sup>8</sup>. The inner layer is thick and stratified and prevents gut bacteria from reaching the epithelial cell surface. It is progressively converted into a laxer and expanded coating through the lytic action of proteases and glycosidases of both the host and the commensal bacteria. The outer mucus layer is the colonization site of the resident commensal microbiota, which differs from that of the digesta-associated and fecal content in terms of relative abundance of the different taxa<sup>9</sup>.

Mucus layers are the frontline of the interaction between the host and the gut bacteria, thus a balanced and symbiotic relationship which benefits both the former and the latter greatly relies on mucus structure<sup>2</sup>. Mucin plays a pivotal role in the selection of bacteria colonizing the mucus by supplying carbon and nitrogen substrates and exposing O-glycan chains that serve as attachment sites for colonization<sup>4,8</sup>. In turn, commensal bacteria feeding on and adhering to mucin limit the penetration of pathogens in the outer layer of the mucus<sup>10</sup>. These bacteria tightly interact with the host and modulate mucin gene expression, glycosylation, and secretion, thus affecting mucus and epithelial homeostasis<sup>3,11</sup>.

<sup>1</sup>Department of Life Sciences, University of Modena and Reggio Emilia, 41125 Modena, Italy. <sup>2</sup>Biogest-Siteia, University of Modena and Reggio Emilia, 42122 Reggio Emilia, Italy. <sup>3</sup>These authors contributed equally: Stefano Raimondi and Eliana Musmeci. ✉email: maddalena.rossi@unimore.it

Sample	Anaerobic mucinolytic bacteria Log (cfu/mL)	Enterobacteriaceae Log (cfu/mL)	Isolated strains	Ref. seq ID
EC 1	7.8 ± 0.9	7.2 ± 0.5	<i>Clostridium disporicum</i>	WC0700 NR_026491.1
			<i>Clostridium disporicum</i>	WC0701 NR_026491.1
EC 2	8.4 ± 0.7	7.3 ± 0.4	<i>Clostridium disporicum</i>	WC0702 NR_026491.1
			<i>Clostridium disporicum</i>	WC0703 NR_026491.1
EC 3	8.5 ± 0.5	7.9 ± 0.5	<i>Clostridium disporicum</i>	WC0704 NR_026491.1
EC 4	8.3 ± 0.6	7.6 ± 0.6	<i>Paraclostridium benzoelyticum</i>	WC0705 NR_148815.1
			<i>Clostridium disporicum</i>	WC0706 NR_026491.1
EC 5	8.4 ± 0.5	8.0 ± 0.3	<i>Paraclostridium benzoelyticum</i>	WC0707 NR_148815.1
			<i>Paraclostridium benzoelyticum</i>	WC0708 NR_148815.1
			<i>Clostridium tertium</i>	WC0709 NR_113325.1

**Table 1.** Cultivable mucin degraders in EC samples. The viable counts of anaerobic mucinolytic bacteria and Enterobacteriaceae, respectively counted onto plates of MM medium and MacConkey, and the taxonomic attribution of the predominant anaerobic is reported. Counts are reported as mean Log (cfu/mL) ± standard deviation, n = 3. For both the media, no significant difference was observed among samples (ANOVA, P > 0.05). Cultivable anaerobic mucinolytic bacteria were significantly more numerous than Enterobacteriaceae (paired samples t-test, P < 0.05). Taxonomic attribution was obtained by alignment of the V1–V4 regions of 16S rRNA gene with the NCBI refseq\_rna database using BLAST. The accession number of best hits and the sequence identity (%) obtained with MUSCLE are reported.

This study aimed to investigate mucin degraders of the gut, to unveil how mucins influence selection of resident bacteria and how specific bacteria shape the mucus layer. A medium containing mucin as sole carbon and nitrogen source (hereinafter referred to as MM medium) was utilized in strictly anaerobic enrichment cultures of human gut microbiota, inoculated with fresh feces of 5 healthy adults. The taxa that took advantage of mucins as growth substrate were identified with a 16S rRNA gene metataxonomic analysis, comparing the bacterial community of the fecal slurries (hereinafter referred to as FS) and the cultures after three enrichment steps (hereinafter referred to as EC). The bacteria able to grow using mucin as sole carbon and nitrogen source were isolated and taxonomically characterized from EC samples.

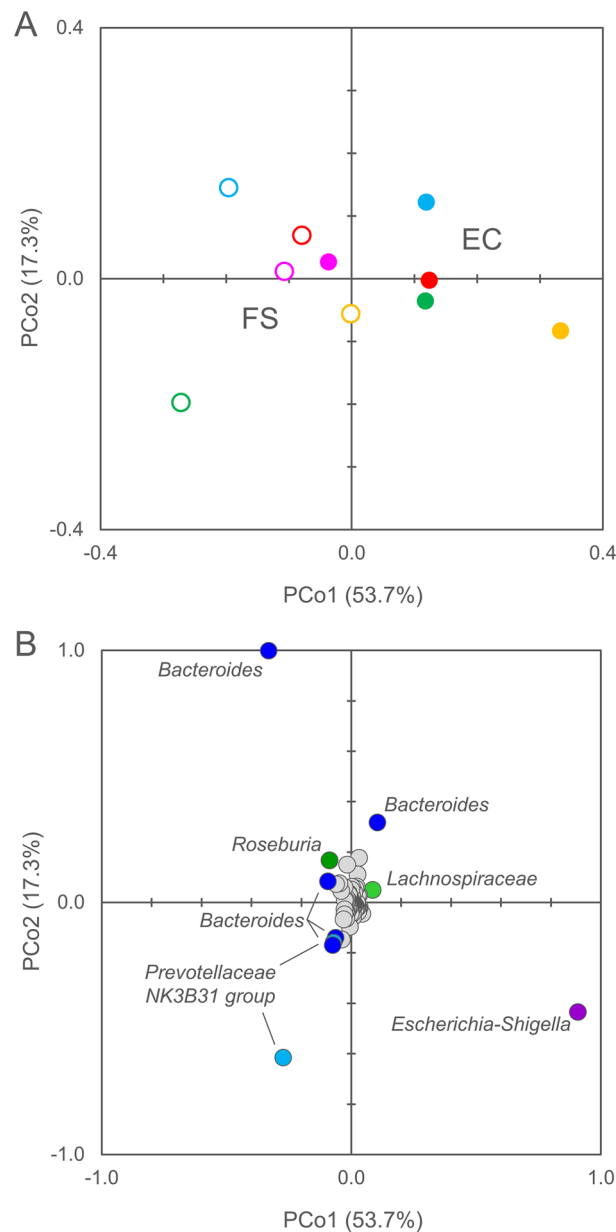
## Results

**Cultivable bacteria enriched on mucin.** Intestinal mucin degrading bacteria were enumerated and isolated from EC samples, i.e. after three enrichment steps in MM broth. To discriminate anaerobic mucin degraders from mucinolytic Enterobacteriaceae, the enumeration was carried out with MM plates in anaerobiosis and MacConkey plates in aerobiosis (Table 1), and the isolates from MM plates were confirmed to be strict anaerobes. Strictly anaerobic mucin degraders were found in the magnitude of 10<sup>8</sup> cfu/mL (mean ± SD = 8.3 ± 0.3 Log<sub>10</sub> cfu/mL), significantly more abundant (P < 0.05) than mucinolytic Enterobacteriaceae, that ranged between 10<sup>7</sup> and 10<sup>8</sup> cfu/mL (mean ± SD = 7.6 ± 0.4 Log<sub>10</sub> cfu/mL). The viable counts of both anaerobic mucin degraders and mucinolytic Enterobacteriaceae did not differ among the EC samples (ANOVA, P > 0.05).

Based on RAPD-PCR fingerprinting, the anaerobic isolates clustered in 1–3 distinct biotypes *per* sample, for a total of 10 strains (Table 1). The partial sequences of 16S rRNA genes showed identity > 98.6% to the following species: *Clostridium disporicum* (6 biotypes), *Clostridium tertium* (1), and *Paraclostridium benzoelyticum* (3). These strains were not able to grow in MM plates without mucin.

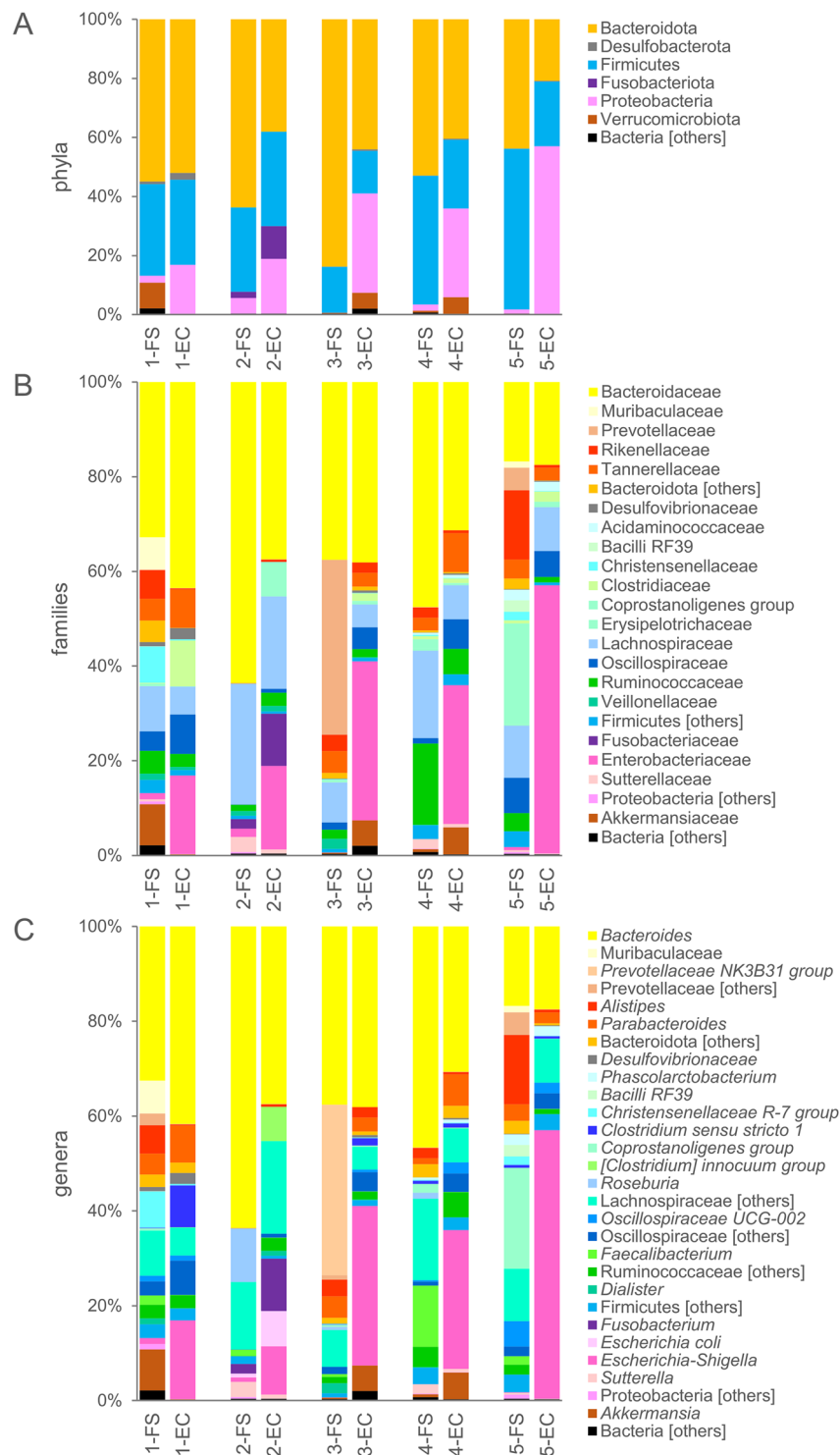
**Microbiota of enrichment cultures grown on mucin.** The metataxonomic survey of 16S rRNA gene in FS and EC samples yielded a total 415,051 sequences (31,153–55,162 *per* sample). The reads were dereplicated into 522 ASVs hitting a reference sequence in Silva database, and collapsed at the 7th level of taxonomic annotation into 150 OTUs (Supplementary Table 1).

The analysis of microbiota alpha-diversity revealed that the richness (Chao1) and the evenness (Shannon and Pielou) tended to decrease with the enrichment but did not reach statistical significance (P > 0.05) (Supplementary Fig. 1). According to Weighted UniFrac computation of beta-diversity, FS and EC samples clustered in distinct groups (ANOSIM, PERMANOVA, P < 0.05) in a PCoA plot, with the former laying at lower values of PCo1 than the latter (Fig. 1A). *Escherichia-Shigella* presented the most positive contribution to PCo1, followed by ASVs of *Bacteroides* and of Lachnospiraceae, while other ASVs of *Bacteroides*, Prevotellaceae NK3B31, and *Roseburia* negatively weighted (Fig. 1B). Unweighted UniFrac yielded more dispersed groups (ANOSIM, PERMANOVA, P > 0.05), with EC samples laying at lower PCo2 compared to the corresponding FS ones (Supplementary Fig. 2A). The ASVs that negatively contributed to PCo2 included *Clostridium sensu stricto I* and other *Bacteroides*, while Prevotellaceae NK3B31 and *Eubacterium coprostanoligenes* were positive contributors (Supplementary Fig. 2B). Weighted UniFrac computation takes into account abundances of taxa, unlike Unweighted UniFrac that considers the mere presence or absence. Thus, the taxa that emerged from the former presented relevant changes in terms of relative abundance, while those that emerged from the latter likely appeared or disappeared over the enrichment steps.



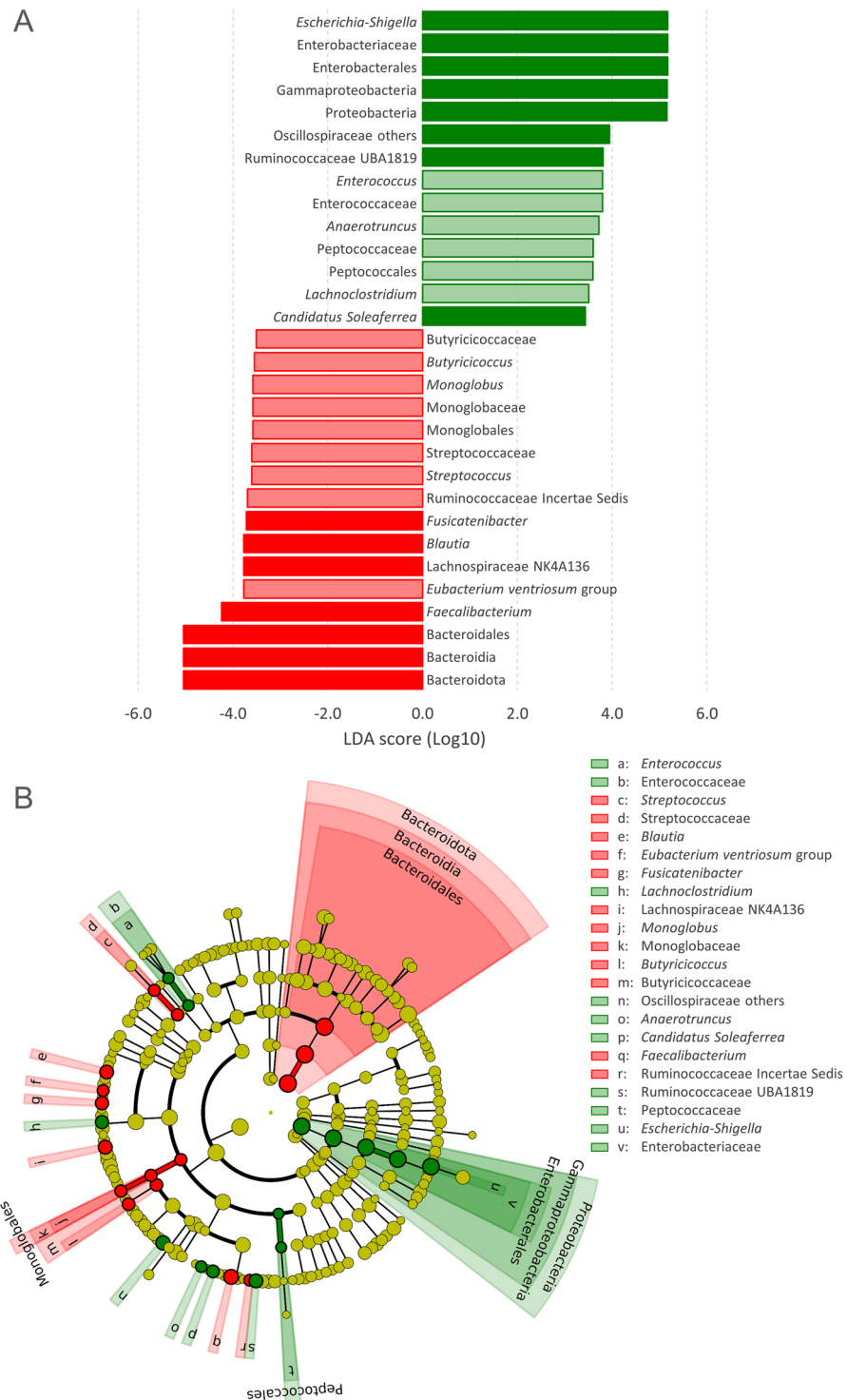
**Figure 1.** Beta diversity analysis of the microbiota in FS and EC samples. **(A)** PCo1-PCo2 visualization of distances computed with Weighted Unifrac, that considers phylogenetic relationships among bacteria. Symbols: FS, empty circle; EC, full circle; different colors correspond to different subjects (1, fuchsia; 2, cyan; 3, green; 4, red; 5, yellow). **(B)** PCo1-PCo2 visualization of the contribution of each ASV. The 10 ASVs exerting the highest effect on PCo1 are labelled and colored according to their phylum (Firmicutes, green; Bacteroidetes, blue; Proteobacteria, purple).

The microbiota of FS samples encompassed 125 of the 150 identified OTUs. It was dominated by Bacteroidota (59.5%), with remarkably high amounts of *Bacteroides* (39.3%) and Prevotellaceae (8.4%), followed by Firmicutes (35.3%), especially Lachnospiraceae (15.4%), and Proteobacteria (2.2%; Fig. 2). The microbiota of EC samples encompassed 101 OTUs, 25 of which became detectable because of the enrichment steps, as they lay below the limit of detection in FS samples. EC were dominated by Bacteroidota (39.3%) and Proteobacteria (30.8%). *Escherichia-Shigella* (with Enterobacteriaceae, Enterobacteriales, Gammaproteobacteria, and Proteobacteria) presented a significant positive differential abundance in EC compared to FS (Fig. 3) and represented one of the dominant genera in EC samples (mean = 30.5%), with remarkably high abundance in sample EC-5 (56.7%). Several taxa of Firmicutes, such as *Enterococcus* and many Clostridia (e.g., Oscillospiraceae, Peptococcaceae, Ruminococcaceae UBA1819, *Soleaferrea*, *Lachnoclostridium*, and *Anaerotruncus*), presented a significant increase in EC samples, but generally remained below a mean abundance of 1%.

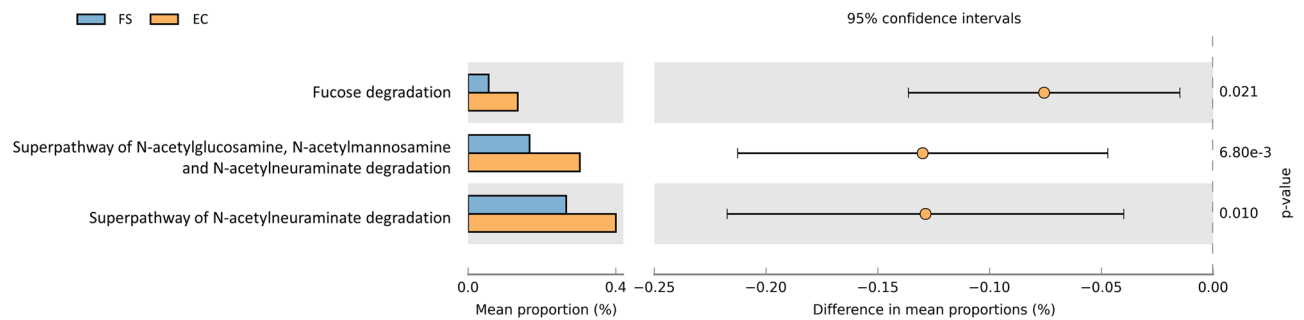


**Figure 2.** Stacked bar-plot representation of microbiota composition before and after the enrichment in MM, with taxonomic features collapsed at the level of phyla (A), families (B), and genera (C). The taxa that remained unclassified at the deeper level or that never occurred with abundance higher than 0.5% are grouped and marked with “[others]”.

A number of taxa were not detected in any FS sample but appeared after the enrichment: *Eggerthella* occurred in 4 EC samples; *Clostridium innocuum*, *Enterococcus durans*, *Erysipelatoclostridium*, and *Soleiferrea* in 3; *Ange-lakisiella* and *Eubacterium nodatum* in 2 EC samples. These taxa were generally a minor portion of the microbiota (<1%), except for *Clostridium innocuum* that reached the 7.3% in EC-2. *Enterorhabdus*, *Gordonibacter*, *Dielma*, *Enterococcus faecalis* and other *Enterococcus* spp., *Catabacter*, *Ruminiclostridium*, Lachnospiraceae FCS020 group,



**Figure 3.** Effect of mucin enrichment on faecal microbiota. The taxonomic biomarkers characterizing FS and EC according to LEfSe are colored in red and green, respectively. **(A)** Logarithmic LDA scores of the taxa exhibiting significant differential abundance ( $P < 0.05$ , logarithmic LDA score  $\geq 2.0$ ) between FS and EC; taxa that were  $> 1\%$  in at least one sample are reported with bars in dark shades. **(B)** Cladogram visualization of the taxonomic biomarkers. The figure was created with the software LEfSe, Galaxy version 1.0, available at <https://huttenhower.sph.harvard.edu/galaxy/>.



**Figure 4.** Pathways related to mucin degradation reconstructed from 16S rRNA profile with PICRUSt2 and presenting a significant ( $P < 0.05$ ) differential abundance between FS and EC. The figure was created with STAMP v2.1.3, available at <https://beikolab.cs.dal.ca/software/STAMP>.

*Hydrogenoanaerobacterium*, *Caproiciproducens*, Ruminococcaceae DTU089, *Peptostreptococcus*, *Peptoniphilus*, *Tissierella*, *Sedimentibacter*, *Citrobacter*, and *Cloacibacillus* were detected in a single EC sample, generally in very low concentration, except for *Dielma* (1%) and *Cloacibacillus* (1.5%).

Bacteroidota, Bacteroidia, and Bacteroidales, significantly decreased as consequence of mucin enrichment, albeit the genus *Bacteroides* remained the most abundant in 4 out of 5 EC samples. The genus *Streptococcus* also diminished, leading to significantly lower abundances also of Streptococcaceae. Numerous taxa of Clostridia decreased during the enrichment steps, such as *Eubacterium ventriosum*, the genera *Faecalibacterium*, *Fusicatenibacter*, *Blautia*, *Monoglobus* (plus Monoglobaceae and Monoglobales), and *Butyricoccus* (and Butyricocccaceae). Furthermore, the abundance of a clade of Lachnospiraceae (NK4A136) and Ruminococcaceae *incertae sedis* significantly diminished. Other taxa detected only in some FS samples decreased below the limit of detection in the EC samples, such as Coriobacteriales *incertae sedis*, other Eggerthellaceae, *Paraprevotella*, Defluviitaleaceae\_UCG-011, *Eubacterium siraeum* group, *Veillonella*, *Butyricimonas*, Erysipelotrichaceae\_UCG-003, *Turicibacter*, *Coprococcus*, *Lachnospira*, *Roseburia*, Anaerovoracaceae Family XIII UCG-001, and *Haemophilus*.

The functions of the fecal cultures' metagenomes were predicted with PICRUSt2. A significant increase (t-test,  $p < 0.05$ ) of pathways related to mucin metabolism was detected, i.e. fucose degradation, superpathway of N-acetylneuraminic acid degradation, and superpathway of GlcNAc, N-acetylmannosamine, and N-acetylneuraminic acid degradation (Fig. 4). In particular, genes encoding N-acetylglucosamine kinase (EC:2.7.1.59), N-acetylglucosamine-6-phosphate deacetylase (EC:3.5.1.25), N-acetylneuraminic epimerase (EC:5.1.3.24), N-acetylglucosamine-6-phosphate 2-epimerase (EC:5.1.3.9), N-acetylmannosamine kinase (EC:2.7.1.60) showed a statistically significant increase in the enriched cultures.

## Discussion

This study aimed to identify and isolate the intestinal bacteria that can have a role in mucin degradation, utilizing three steps of enrichment where the intestinal microbiota of healthy subjects was inoculated in a medium containing mucin as the sole carbon and nitrogen source. Mucin can support the growth of gut microbes with different metabolic capabilities. Both saccharolytic and proteolytic activities are involved in its breakdown, providing the bacterial community with carbon and nitrogen supply and sustaining growth of both sugar and amino acid fermenters. Based on the increase of genes and pathways involved in mucin degradation, the cultures successfully selected bacteria performing mucin breakdown and fermentation.

The 16S rRNA gene profiling highlighted the taxa that proliferated in mucin, likely because of cooperation and crossfeeding. Proteobacteria and taxa related to Enterobacteriaceae encompassed the main biomarkers of growth, in agreement with their major role as protein degraders in gut microbiota. As opportunistic commensals, they may persist in the gut contributing to the mucus homeostasis without interacting with epithelial cells or they may degrade mucins, reach the epithelium, and initiate disease<sup>12–14</sup>. The high affinity of Enterobacteriaceae for mucins confirms the possibility that they may get in contact with the intestinal epithelial barrier and exert this pathogenic action. In fact, these bacteria exhibit an intrinsic resistance to the oxygen released from the host tissue and may concur to disrupt the integrity of the epithelium, promoting a leaky gut and triggering systemic inflammatory response<sup>15–17</sup>.

Enterobacteriaceae, Peptococcales/Peptococcaceae, a variety of Clostridia (such as Oscillospiraceae, *Lachnospira*, and *Anaerotruncus*), and the Coriobacteriales *Eggerthella*, are protein degraders that may have taken part to the breakdown of protein backbone of mucin<sup>18</sup>. The genus *Anaerotruncus* includes species (*Anaerotruncus colihominis* and *Anaerotruncus massiliensis*) that ferment some amino acids and also a small set of carbohydrates, including GlcNAc, and may have participated to the degradation of both the protein and saccharidic portions of mucins<sup>19,20</sup>. *Holdemania*, a Firmicute belonging to the family Erysipelotrichaceae, and Enterococcaceae (Lactobacillales) are saccharolytic bacteria, thus their contribution to mucin degradation seems limited to the carbohydrate chains. Several other taxa, which are still poorly characterized at the physiological level, emerged as participating in mucin catabolism, for some species representing one of the first sources of functional information. Interestingly, the list of bacteria that got enriched in EC includes taxa that are associated to colorectal adenoma and cancer (e.g. *Lachnospira*) and are emerging as uncommon human pathogens (*Eggerthella* and some Enterococcaceae)<sup>21–23</sup>.



Even though the 16S rRNA gene profiling disclosed a complex population of bacteria taking advantage of mucin, only three species of strict anaerobic bacteria (*C. disporicum*, *C. tertium*, and *P. benzoelyticum*) were isolated from the cultures enriched with 5 different microbiotas. *C. disporicum* is a saccharolytic bacterium that presents epimerase activity resulting in the production of ursodeoxycholic acid<sup>24,25</sup>. *C. tertium* is an uncommon pathogen, albeit associated with cases of severe diseases, mostly in immunocompromised and neutropenic patients<sup>26</sup>. *P. benzoelyticum* can ferment some amino acids, including threonine, a major component of mucin backbone<sup>27</sup>. Interestingly, the isolates of *C. disporicum* and *C. tertium* are all designated as *Clostridium* sensu strictu 1 according to SILVA database. The 16S rRNA gene profiling revealed that the OTU *Clostridium* sensu strictu 1 increased in all the samples where it was present, although it did not reach a significant differential abundance, and was composed of four ASVs contributing to FS and EC separation along PCo2 in the plot of Unweighted UniFrac beta-diversity.

Albeit *Clostridium* species harbor few glycosyl hydrolases (GH) involved in mucin degradation and, for instance, genome of *C. sporogenes*, a closely related species to *C. tertium*, has no mucin-related GHs, based on preliminary genomic analysis (data not shown), the genome of *C. tertium* WC 0709 has at least one exo- $\alpha$ -sialidase (EC 3.2.1.18) belonging to the GH33 family, involved in the cleavage of the terminal sialic acid, and a GH3 family  $\beta$ -hexosaminidase (EC 3.2.1.52) that can release GalNac. On the other side *P. bifementans*, closely related to *P. benzoelyticum*, harbors only the galactosidases GH2 and GH20. Generally, commensal clostridia do not appear to harbor GH33 or other mucin-related GHs, whereas two pathogenic clostridia degrade mucin glycans: *C. perfingens* and *C. septicum*<sup>28</sup>. The capability of *C. disporicum*, *C. tertium*, and *P. benzoelyticum* to dismantle complex mucin glycans deserves deeper investigations and opens new perspectives on the relationship of these species with the host, that can go beyond the commensalism. Incidentally, *C. tertium* is described as an uncommon pathogen associated with various cases of severe diseases<sup>26,29</sup>. The glycans of intestinal mucins, enriched in diverse monosaccharides with different positional and stereochemical linkages and branching, are hydrolyzed by a number of diverse sialidases, hexosaminidases, galactosidases, and fucosidases. The genetic and functional characterization of the enzymes that in *C. disporicum*, *C. tertium*, and *P. benzoelyticum* play key roles in disassembling mucin oligosaccharides is a great challenge.

In the gut, a complex community of saccharolytic and proteolytic bacteria is expected to participate to mucin breakdown, with linkage-specific glycosidases and proteases of primary degraders providing other bacteria with simpler oligosaccharides and peptides, and eventually with their monomeric moieties. It is plausible that several taxa detected by the 16S rRNA gene analysis could have grown on mucins only when co-cultured within a bacterial community where each species contributed with different hydrolytic activities, while they could not grow in pure cultures where they did not benefit from crossfeeding relationships. Only *C. disporicum*, *C. tertium*, and *P. benzoelyticum* gained carbon, energy, and nitrogen from mucin in pure cultures. The isolates of these species deserve deeper investigation, including the genome sequencing and physiology studies, aiming to unveil the key functions that enabled growth on mucin, in particular their set of glycosyl hydrolases and proteases and their amino acid biosynthetic capability.

To date, *Bacteroides* and *Akkermansia muciniphila* have been recognized as main mucin degraders, with novel insights on endo-acting O-glycanases involved in the first steps of glycan breakdown<sup>30</sup>. In this study, the relative abundance of *Bacteroides* did not increase significantly, even if some ASVs of *Bacteroides* heavily contributed to separate samples along both PCo1 or PCo2 in the Weighted UniFrac analysis, confirming the presence of some species greatly associated to mucin degradation. It is plausible that different populations of *Bacteroides*, which could not be discriminated into species by the partial 16S rRNA gene sequencing, behaved differently with respect to mucin breakdown, exerting contrary effects and balancing the counts of the genus. Indeed, a wide number of ASVs attributed to *Bacteroides* presented opposite contributions along the principal coordinates, in agreement with the fact that the genus *Bacteroides* comprises tens of species that supply different and complementary functions to the intestinal community<sup>31</sup>.

One of the key members of the colonic mucus-associated microbiota is *A. muciniphila*. It releases monosaccharides and amino acids during mucin degradation, providing nutrients to other bacteria of the gut, and exerts beneficial effects in various metabolic disorders, being generally negatively correlated with inflammation and metabolic disorders<sup>32</sup>. In this study, a sole ASV ascribed to *Akkermansia* was detected in 3 out of 5 set of samples. In subject 1, it was remarkably abundant in the founding microbiota (8.7%) but was negligible (0.2%) after the enrichment steps. In samples 3 and 4, *Akkermansia* was initially < 0.6% and increased by one magnitude over the enrichment. It is plausible that in sample 1 the high initial load of *Akkermansia* prevented a net increase of the abundance, that was further squeezed by major increase in some other groups.

The absence of *Bacteroides* and *Akkermansia* also within the isolates is noteworthy. These bacteria likely encountered some substrate limitation that hampered growth. In fact, the medium did not contain nitrogen sources other than mucin in order to be as much stringent as possible. Thus, all the amino acids had to be obtained from the protein backbone of mucin, and this might have been too challenging, in the absence of a cooperative crossfeeding. The composition of the medium and the choice of the pH are major drivers of microbial selection and a mucin-based medium with different composition would likely result in the isolation of different bacterial species. Another drawback of the present study was the use of mucin from porcine stomach, that mainly consists of MUC5AC and MUC6 mucins and differs in purity and glycosylation profile from the colonic one MUC2, but is a cheap and easily available substrate, commonly utilized to culture mucin degraders<sup>5,33–35</sup>.

This study shed light on the bacterial taxa thriving in cultures of human microbiota with mucin as primary substrate for growth. It provided the first information on *C. disporicum*, *C. tertium*, and *P. benzoelyticum* as bacteria able to grow in pure cultures utilizing mucin as sole carbon and nitrogen source, complementing the recognized role of *Bacteroides* and *Akkermansia* in mucin degradation and turnover. Sequencing of the 10 strains is in progress, as well as the formulation of a new mucin-based medium allowing a faster growth with higher yields, in order to better address physiology and metabolism studies. Proteobacteria, including Enterobacteriaceae

and *E. coli*, are among the microbial group that most benefit from mucin as substrate, increasing the risk of the flourishing of opportunistic pathogens or bacteria associated to gut inflammation. Furthermore, metataxonomic analysis identified *Lachnospirillum*, *Eggerthella*, *Anaerotruncus*, *Enterococcaceae*, and *Peptococcaceae* as taxa that took advantage, likely in a cross-feeding relationship, in the mucin-based medium. The limited number of isolated species able to grow on mucin in pure culture confirms that the degradation of these glycoproteins in the gut results from cooperation and cross-feeding among several species exhibiting different metabolic capabilities.

## Materials and methods

**Chemicals and media.** All the chemicals were purchased from Sigma-Aldrich (Merck KGaA, Darmstadt, Germany), unless otherwise stated. Mucin was purified from porcine stomach type II mucin (Sigma-Aldrich), according to the protocol by Miller and Hoskins<sup>36</sup>. Briefly, mucin was precipitated with ice-cold ethanol at the concentration of 60% (v/v) from a clarified solution, obtained by centrifugation (10,000×g for 10 min at 4 °C) of a 25 g/L mucin suspension in 0.1 M NaCl containing 0.02 M phosphate buffer, pH 7.6. The precipitate was dissolved in 0.1 M NaCl, re-precipitated with ethanol (60% v/v), and dissolved in water. The mucin solution was dialyzed with (14 kDa cut-off) against water and finally freeze dried. According to the producer, mucin had a molecular mass of 4000–5500 kDa, with a carbohydrate portion of 80–84%, and a protein portion of 16–20%. Proline, serine, and threonine residues in the peptide backbone comprise approx. 43% of the amino acid composition.

MM medium contained 3.0 g/L purified mucin, 2.0 g/L  $\text{KH}_2\text{PO}_4$ , 4.5 g/L NaCl, 0.5 g/L  $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.045 g/L  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ , 0.005 g/L  $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.01 g/L hemin, 0.05 g/L bile salts (Oxgall, BD Difco, Sparks, MD, USA), 0.6 mg/L resazurin, 2.0 mL/L minerals solution (0.5 g/L EDTA, 0.010 g/L  $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ , 0.003 g/L  $\text{MnCl}_2 \cdot 7\text{H}_2\text{O}$ , 0.03 g/L  $\text{H}_3\text{BO}_3$ , 0.02 g/L  $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$ , 0.001 g/L  $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$ , 0.002 g/L  $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$ , and 0.003 g/L  $\text{NaMoO}_4 \cdot 2\text{H}_2\text{O}$ ), 1.4 mL/L vitamins solution (1.0 g/L menadione, 2.0 g/L biotin, 2.0 g/L calcium pantothenate, 10 g/L nicotinamide, 0.5 g/L cyanocobalamin, 0.5 g/L folic acid, 4 g/L thiamine, and 5 g/L PABA), and 40 mL/L reducing solution (12.5 g/L L-cysteine-HCl and 80 g/L  $\text{NaHCO}_3$ ). The basal medium was autoclaved at 121 °C for 20 min and supplemented with filter sterilized (0.22 µm) minerals, vitamins, and reducing solutions. Soft-agar MM plates, containing 8 g/L agar (BD Difco), were used to count, and purify mucin degraders. MacConkey agar plates (BD Difco) were utilized to count *Escherichia coli* and Enterobacteriaceae and to discriminate them from strictly anaerobic isolates.

**Enrichment cultures.** This study was carried out under the recommendations of the protocol approved with ref. no. 125-15 by the local research ethics committee (Comitato Etico Provinciale, Azienda Policlinico di Modena, Italy), with written informed consent from all subjects following the Declaration of Helsinki. Five healthy adults (3 men and 2 women, aged 25–50 years) who had not taken prebiotics and/or probiotics in the previous 2 weeks or antibiotics for at least 3 months were enrolled for feces collection. Fecal samples were collected fresh, sealed in anaerobic plastic bags (AnaeroGen, Oxoid, Basingstoke, UK), transferred in an anaerobic cabinet (Concept Plus, Ruskinn Technology, Ltd., Bridgend, UK) with an atmosphere of 85%  $\text{N}_2$ , 10%  $\text{CO}_2$ , and 5%  $\text{H}_2$ , and homogenized (10%, w/v) in sterile MM.

The fecal slurries (FS) were inoculated with a syringe (10% v/v) into butyl-rubber stoppered bottles containing 50 mL of sterile anaerobic MM. After an incubation of 72 h at 37 °C, these cultures were utilized to seed (10% v/v) the next passage in MM, repeated to accomplish three enrichment steps. Once grown, the third enrichment culture (EC) was utilized for isolation of mucinolytic bacteria. The viable counts of anaerobic mucinolytic bacteria and Enterobacteriaceae in EC samples were determined in triplicate onto agar plates of MM and MacConkey, respectively. For both the media, one-way ANOVA followed by Tukey's *post-hoc* analysis was utilized to compare the counts of the different subjects. Paired samples t-test was utilized to compare MM and MacConkey counts.

**Isolation of anaerobic mucin degraders.** All the steps of the isolation and purification procedure were carried out in the anaerobic cabinet. The EC was serially diluted (1:10) in anaerobic MM, then 1 mL of each dilution was poured into plates and covered with 20 mL of melted soft-agar MM. The plates were allowed to cool and incubated in anaerobiosis at 37 °C for 72–96 h. Inclusion colonies were picked and checked for aerobic growth at 37 °C in MacConkey agar plates. Colonies unable to grow in aerobiosis were further purified at least three times by serial dilution and isolation in soft-agar MM.

Approximately 10 pure clones per sample were fingerprinted by RAPD-PCR according to Quartieri et al.<sup>37</sup> and clustered into biotypes with a similarity level of 75% using the Pearson correlation coefficient. Single biotypes from each sample were taxonomically characterized by partial sequencing of the 16S rRNA gene (Eurofins Genomics, Ebersberg, Germany), utilizing primers targeting the V1–V4 regions<sup>38</sup>. Sequences were compared to the NCBI refseq\_rna database with the alignment tools BLAST and MUSCLE (<https://www.ncbi.nlm.nih.gov/>; <https://www.ebi.ac.uk/Tools/msa/muscle/>).

**16S rRNA gene profiling.** Total DNA was extracted from 2 mL of FS and EC samples with the QiAmp PowerFecal DNA kit (Qiagen, Hilden, Germany), according to the manufacturer's protocol. The DNA was normalized to 5 ng/µL after quantification with a Qubit 3.0 fluorimeter (Thermo Fisher Scientific, Waltham, MA, USA).

Partial 16S rRNA gene sequences were amplified using Probio\_Uni/Probio\_Rev primers, which targeted the V3 region of the 16S rRNA gene. Amplicons were sequenced using a MiSeq platform (Illumina Inc., San Diego, CA, USA) according to Milani et al.<sup>39</sup>. Raw sequences were analyzed with the QIIME 2.0 pipeline, version 2019.10<sup>40</sup>, with appropriate plugins for trimming (CUTADAPT) and denoising (DADA2) into amplicon sequence variants (ASVs)<sup>41,42</sup>. Taxonomy assignment was carried out with the feature classifier VSEARCH<sup>43</sup>,



with SILVA SSU database release 138 (<https://www.arb-silva.de/download/arb-files/>) as reference and the the similarity threshold set at 0.97. The appropriate QIIME2 plugins were utilized to compute the alpha- (observed taxa, Chao1, Shannon, and Pielou's evenness) and beta-diversity (Unweighted Normalized UniFrac, and Weighted Normalized UniFrac) and to compare them within and between FS and EC samples (i.e., the Kruskal–Wallis test for alpha diversity; ANOSIM and PERMANOVA for beta diversity). Differences were considered significant for  $P < 0.05$ . Principal Coordinate Analysis (PCoA) was computed with QIIME2, based on the beta-diversity distance matrices.

Linear discriminant analysis Effect Size (LEfSe, <http://huttenhower.sph.harvard.edu/galaxy>) algorithm was utilized to discover distinctive taxonomic features characterizing FS and EC samples<sup>44</sup>. Taxa presenting a significant differential abundance ( $P < 0.05$ ) and logarithmic LDA (linear discriminant analysis) score  $> 2$  were considered as microbial biomarkers of FS or EC samples.

The dataset was analyzed with PICRUST2 to predict metagenomes' functions and investigate the presence of any difference between the FS and EC samples in terms of enzymes and pathways related to mucin metabolism<sup>45</sup>. The tool predicted the enzymes and the relative MetaCyc pathways abundances<sup>46</sup>. STAMP was used to visualize the results and conduct statistical analysis<sup>47</sup>.

The 16S rRNA gene sequences datasets generated and analyzed during the current study are available in the NCBI repository with the BioProject ID: PRJNA649741 (<https://www.ncbi.nlm.nih.gov/>).

Received: 26 January 2021; Accepted: 7 May 2021

Published online: 27 May 2021

## References

- Bansil, R. & Turner, B. S. The biology of mucus: Composition, synthesis and organization. *Adv. Drug Deliv. Rev.* **124**, 3–15. <https://doi.org/10.1016/j.addr.2017.09.023> (2018).
- Alemao, C. A. *et al.* Impact of diet and the bacterial microbiome on the mucous barrier and immune disorders. *Allergy* <https://doi.org/10.1111/all.14548> (2020) (Online ahead of print).
- Wagner, C. E., Wheeler, K. M. & Ribbeck, K. Mucins and their role in shaping the functions of mucus barriers. *Annu. Rev. Cell Dev. Biol.* **34**, 189–215. <https://doi.org/10.1146/annurev-cellbio-100617-062818> (2018).
- Tailford, L. E., Crost, E. H., Kavanaugh, D. & Juge, N. Mucin glycan foraging in the human gut microbiome. *Front. Genet.* **6**, 81. <https://doi.org/10.3389/fgene.2015.00081> (2015).
- Corfield, A. P. The interaction of the gut microbiota with the mucus barrier in health and disease in human. *Microorganisms*. **6**, 78. <https://doi.org/10.3390/microorganisms6030078> (2018).
- Bergstrom, K. S. & Xia, L. Mucin-type O-glycans and their roles in intestinal homeostasis. *Glycobiology* **23**, 1026–1037. <https://doi.org/10.1093/glycob/cwt045> (2013).
- Sender, R., Fuchs, S. & Milo, R. Revised estimates for the number of human and bacteria cells in the body. *PLoS Biol.* **14**, e1002533. <https://doi.org/10.1371/journal.pbio.1002533> (2016).
- Johansson, M. E., Larsson, J. M. & Hansson, G. C. The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislation of host-microbial interactions. *Proc. Natl. Acad. Sci.* **108**, 4659–4665. <https://doi.org/10.1073/pnas.1006451107> (2011).
- Donaldson, G. P., Lee, S. M. & Mazmanian, S. K. Gut biogeography of the bacterial microbiota. *Nat. Rev. Microbiol.* **14**, 20–32. <https://doi.org/10.1038/nrmicro3552> (2016).
- Linden, S. K., Sutton, P., Karlsson, N. G., Korolik, V. & McGuckin, M. A. Mucins in the mucosal barrier to infection. *Mucosal Immunol.* **1**, 183–197. <https://doi.org/10.1038/mi.2008.5> (2008).
- Cornick, S., Tawiah, A. & Chadee, K. Roles and regulation of the mucus barrier in the gut. *Tissue Barriers*. **3**, e982426. <https://doi.org/10.4161/21688370.2014.982426> (2015).
- Troge, A. *et al.* More than a marine propeller—the flagellum of the probiotic *Escherichia coli* strain Nissle 1917 is the major adhesin mediating binding to human mucus. *Int. J. Med. Microbiol.* **302**, 304–314. <https://doi.org/10.1016/j.ijmm.2012.09.004> (2012).
- Raimondi, S. *et al.* Antibiotic resistance, virulence factors, phenotyping, and genotyping of *E. coli* isolated from the feces of healthy subjects. *Microorganisms*. **7**, 251. <https://doi.org/10.3390/microorganisms7080251> (2019).
- Shawki, A. & McCole, D. F. Mechanisms of intestinal epithelial barrier dysfunction by adherent-invasive *Escherichia coli*. *Cell. Mol. Gastroenterol. Hepatol.* **3**, 41–50. <https://doi.org/10.1016/j.jcmgh.2016.10.004> (2016).
- Ugalde-Silva, P., Gonzalez-Lugo, O. & Navarro-Garcia, F. Tight junction disruption induced by type 3 secretion system effectors injected by enteropathogenic and enterohemorrhagic *Escherichia coli*. *Front. Cell. Infect. Microbiol.* **6**, 87. <https://doi.org/10.3389/fcimb.2016.00087> (2016).
- Edwards, L. A., Bajaj-Elliott, M., Klein, N. J., Murch, S. H. & Phillips, A. D. Bacterial-epithelial contact is a key determinant of host innate immune responses to enteropathogenic and enteroaggregative *Escherichia coli*. *PLoS ONE* **6**, e27030. <https://doi.org/10.1371/journal.pone.0027030> (2011).
- Litvak, Y., Byndloss, M. X. & Bäuml, A. J. Colonocyte metabolism shapes the gut microbiota. *Science* **362**, eaat9076. <https://doi.org/10.1126/science.aat9076> (2018).
- Amaretti, A. *et al.* Profiling of protein degraders in cultures of human gut microbiota. *Front. Microbiol.* **10**, 2614. <https://doi.org/10.3389/fmicb.2019.02614> (2019).
- Lau, S. K. *et al.* Bacteraemia caused by *Anaerotruncus colihominis* and emended description of the species. *J. Clin. Pathol.* **59**, 748–752. <https://doi.org/10.1136/jcp.2005.031773> (2006).
- Togo, A. H. *et al.* *Anaerotruncus massiliensis* sp. nov., a succinate-producing bacterium isolated from human stool from an obese patient after bariatric surgery. *New Microbes New Infect.* **29**, 100508. <https://doi.org/10.1016/j.nmni.2019.01.004> (2019).
- Liang, J. Q. *et al.* Fecal microbial DNA markers serve for screening colorectal neoplasm in asymptomatic subjects. *J. Gastroenterol. Hepatol.* <https://doi.org/10.1111/jgh.15171> (2020) (Online ahead of print).
- Bo, J. *et al.* *Eggerthella lenta* bloodstream infections: Two cases and review of the literature. *Future Microbiol.* **15**, 981–985. <https://doi.org/10.2217/fmb-2019-0338> (2020).
- García-Solache, M. & Rice, L. B. The *Enterococcus*: A model of adaptability to its environment. *Clin. Microbiol. Rev.* **32**, e00058–e118. <https://doi.org/10.1128/CMR.00058-18> (2019).
- Horn, N. *Clostridium disporicum* sp. nov., a saccharolytic species able to form two spores per cell, isolated from a rat cecum free. *Int. J. Syst. Bacteriol.* **37**, 398–401. <https://doi.org/10.1099/00207713-37-4-398> (1987).
- Tawthep, S. *et al.* Isolation of six novel 7-oxo- or urso-type secondary bile acid-producing bacteria from rat cecal contents. *J. Biosci. Bioeng.* **124**, 514–522. <https://doi.org/10.1016/j.jbiosc.2017.06.002> (2017).

26. Muñoz, M. *et al.* Comparative genomics identifies potential virulence factors in *Clostridium tertium* and *C. paraputrificum*. *Virulence*. **10**, 657–676. <https://doi.org/10.1080/21505594.2019.1637699> (2019).
27. Sasi Jyothsna, T. S., Tushar, L., Sasikala, C. & Ramana, C. V. *Paraclostridium benzoelyticum* gen. nov., sp. nov., isolated from marine sediment and reclassification of *Clostridium bif fermentans* as *Paraclostridium bif fermentans* comb. nov. Proposal of a new genus *Paeniclostridium* gen. nov. to accommodate *Clostridium sordellii* and *Clostridium ghonii*. *Int. J. Syst. Evol. Microbiol.* **66**, 1268–1274. <https://doi.org/10.1099/ijsem.0.000874> (2016).
28. Low, K. E., Smith, S. P., Abbott, D. W. & Boraston, A. B. The glycoconjugate-degrading enzymes of *Clostridium perfringens*: Tailored catalysts for breaching the intestinal mucus barrier. *Glycobiology* <https://doi.org/10.1093/glycob/cwaa050> (2020).
29. Kiu, R. *et al.* Preterm infant-associated *Clostridium tertium*, *Clostridium cadaveris*, and *Clostridium paraputrificum* strains: Genomic and evolutionary insights. *Genome Biol. Evol.* **9**, 2707–2714. <https://doi.org/10.1093/gbe/evx210> (2017).
30. Png, C. W. *et al.* Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *Am. J. Gastroenterol.* **105**, 2420–2428. <https://doi.org/10.1038/ajg.2010.281> (2010).
31. Wexler, H. M. Bacteroides: The good, the bad, and the nitty-gritty. *Clin. Microbiol. Rev.* **20**, 593–621. <https://doi.org/10.1128/CMR.00008-07> (2007).
32. Zhou, J. C. & Zhang, X. W. *Akkermansia muciniphila*: A promising target for the therapy of metabolic syndrome and related diseases. *Chin. J. Nat. Med.* **17**, 835–841. [https://doi.org/10.1016/S1875-5364\(19\)30101-3](https://doi.org/10.1016/S1875-5364(19)30101-3) (2019).
33. Holmén Larsson, J. M., Thomsson, K. A., Rodríguez-Piñeiro, A. M., Karlsson, H. & Hansson, G. C. Studies of mucus in mouse stomach, small intestine, and colon. III. Gastrointestinal Muc5ac and Muc2 mucin O-glycan patterns reveal a regiospecific distribution. *Am. J. Physiol. Gastrointest. Liver Physiol.* **305**, G357–G363. <https://doi.org/10.1152/ajpgi.00048.2013> (2013).
34. Van den Abbeele, P. *et al.* Butyrate-producing *Clostridium* cluster XIVa species specifically colonize mucins in an in vitro gut model. *ISME J.* **7**, 949–961. <https://doi.org/10.1038/ismej.2012.158> (2013).
35. Marczynski, M. *et al.* Structural alterations of mucins are associated with losses in functionality. *Biomacromol* **22**, 1600–1613. <https://doi.org/10.1021/acs.biomac.1c00073> (2021).
36. Miller, R. S. & Hoskins, L. C. Mucin degradation in human colon ecosystems. Fecal population densities of mucin-degrading bacteria estimated by a “most probable number” method. *Gastroenterology* **81**, 759–765 (1981).
37. Quartieri, A. *et al.* Comparison of culture-dependent and independent approaches to characterize fecal bifidobacteria and lactobacilli. *Anaerobe* **38**, 130–137. <https://doi.org/10.1016/j.anaerobe.2015.10.006> (2016).
38. Raimondi, S. *et al.* Microbiota of sliced cooked ham packaged in modified atmosphere throughout the shelf life: Microbiota of sliced cooked ham in MAP. *Int. J. Food Microbiol.* **289**, 200–208. <https://doi.org/10.1016/j.ijfoodmicro.2018.09.017> (2019).
39. Milani, C. *et al.* Assessing the fecal microbiota: An optimized ion torrent 16S rRNA gene-based analysis protocol. *PLoS ONE* **8**, e68739. <https://doi.org/10.1371/journal.pone.0068739> (2013).
40. Bolyen, E. *et al.* Author correction: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 1091. <https://doi.org/10.1038/s41587-019-0252-6> (2019).
41. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–17. <https://doi.org/10.14806/ej.17.1.200> (2011).
42. Callahan, B. J. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods*. **13**, 581–583. <https://doi.org/10.1038/nmeth.3869> (2016).
43. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: A versatile open source tool for metagenomics. *PeerJ* **4**, e2584. <https://doi.org/10.7717/peerj.2584> (2016).
44. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60. <https://doi.org/10.1186/gb-2011-12-6-r60> (2011).
45. Douglas, G. M. *et al.* PICRUSt2 for prediction of metagenome functions. *Nat. Biotechnol.* **38**, 685–688. <https://doi.org/10.1038/s41587-020-0548-6> (2020).
46. Caspi, R. *et al.* The MetaCyc database of metabolic pathways and enzymes—A 2019 update. *Nucleic Acids Res.* **48**, D445–D453. <https://doi.org/10.1093/nar/gkz862> (2020).
47. Parks, D. H., Tyson, G. W., Hugenholtz, P. & Beiko, R. G. STAMP: Statistical analysis of taxonomic and functional profiles. *Bioinformatics* **30**, 3123–3124. <https://doi.org/10.1093/bioinformatics/btu494> (2014).

## Author contributions

S.R. and M.R. conceived the study and supervised experiments; S.R. and E.M. carried out enrichment cultures and isolation; F.C. and A.A. carried out metagenome analysis, bioinformatics, and data presentation; A.A. and M.R. wrote the main manuscript text; all authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-90553-4>.

**Correspondence** and requests for materials should be addressed to M.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021