



OPEN ACCESS

EDITED BY

Jianan Zhao,
Temple University, Philadelphia, PA,
United States

REVIEWED BY

Mirko Basen,
University of Rostock, Germany
Tom Zaplana,
Muséum National d'Histoire
Naturelle, France

*CORRESPONDENCE

Maddalena Rossi
✉ maddalena.rossi@unimore.it

†These authors have contributed equally
to this work

RECEIVED 05 December 2025

REVISED 20 March 2026

ACCEPTED 20 March 2026

PUBLISHED 10 April 2026

CITATION

Sola L, Candeliere F, Busi E, Raimondi S,
Amaretti A and Rossi M (2026) A genomic
atlas of gut clostridia: phylogeny,
butyrate, and propionate production.
Front. Microbiol. 17:1761627.
doi: 10.3389/fmicb.2026.1761627

COPYRIGHT

© 2026 Sola, Candeliere, Busi, Raimondi,
Amaretti and Rossi. This is an
open-access article distributed under the
terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which does
not comply with these terms.

A genomic atlas of gut clostridia: phylogeny, butyrate, and propionate production

Laura Sola^{1†}, Francesco Candeliere^{1†}, Enrico Busi¹,
Stefano Raimondi^{1,2}, Alberto Amaretti^{1,2} and Maddalena Rossi^{1,2*}

¹Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy, ²Biogest-Siteia,
University of Modena and Reggio Emilia, Reggio Emilia, Italy

Introduction: *Clostridia* is a major microbial class in the human gut, crucial for fermenting undigested carbohydrates and proteins, which produce short-chain fatty acids essential for gut health and immune balance. This study revised the taxonomic classification and phylogeny of all the species of intestinal *Clostridia* catalogued in the Unified Human Gastrointestinal Genome database using a whole-genome approach and assessed butyrate and propionate producing species.

Methods: A total of 1,897 *Clostridia* species, including those with recognised binomial nomenclature and those lacking formal taxonomic classification, were retrieved and reclassified using GTDB-Tk. Their phylogeny was determined by identifying, concatenating, and aligning the 120 ubiquitous single-copy proteins defined in the GTDB. Average amino acid identity (AAI), percentage of conserved proteins (POCP), and phylogenetic relationships were used to organize the species into genera and families. The presence of enzymes belonging to the biosynthetic pathways for butyrate and propionate production was investigated in all genomes with the tool GapSeq.

Results: Reclassification of the genomes resulted in 404 recognised species and 1,493 species lacking formal taxonomic classification. *Oscillospirales* and *Lachnospirales* encompassed most of the species. The pathways leading to butyrate and propionate production were analyzed in their entirety, revealing 519 species as potential butyrate producers, 257 as potential propionate producers and 77 capable of producing both. To assess the abundance of each species, 151 faecal metagenomes of healthy subjects were profiled, indicating that butyrate producing *Clostridia* accounted on average for 28.0% of each microbiome.

Conclusions: This study offers a comprehensive overview of intestinal *Clostridia* diversity, emphasising their role in gut ecosystems and their potential for butyrate and propionate production.

KEYWORDS

butyrate, *Clostridia*, gut microbiome, metagenomics, propionate

1 Introduction

Clostridia are a major class of bacteria within the phylum *Bacillota* that densely inhabits the human gut. They are known for their remarkable metabolic diversity, particularly in how they utilise substrates and employ fermentative pathways. The continuous identification of new anaerobic Gramme-positive bacteria within the class *Clostridia* and its subordinate levels has, over time, led to significant confusion in species classification and relationships.

Several major studies have attempted to resolve this issue by proposing taxonomic revisions, particularly within the genus *Clostridium*, through in-depth analyses of shared genetic and phenotypic characteristics (Haas and Blanchard, 2020). The complexity and inconsistencies of *Clostridia* became especially evident with the advent of molecular techniques, which offered a more precise understanding of their evolutionary connexions. The key study by Collins et al. (1994) presented a significant taxonomic revision of the genus *Clostridium* based on rRNA phylogenetic analysis. This work revealed that *Clostridium* was polyphyletic and proposed classifying the species into clusters that better reflect their evolutionary relationships. A later, thorough analysis of the faecal metagenome from 51 healthy individuals using MetaPhlAn3 identified 77 taxonomic units within the *Clostridia* class, each assigned a recognised binomial nomenclature, further underscoring the urgent need for taxonomic revisions to delineate genera more clearly (Candeliere et al., 2023). The growing interest in rectifying *Clostridia* taxonomy was also evident in several recent publications, ranging from the reclassification of the *Peptostreptococcaceae* (Bello et al., 2024) and specific *Clostridium* and *Butyrivibrio* species (Lawson et al., 2023; Fatahi-Bafghi, 2024) to the phylogenomic mapping of solventogenic *Clostridia* (Jensen et al., 2024). Interest in intestinal *Clostridia* stems primarily from their predominance in the gut and their crucial role in degrading and fermenting complex carbohydrates that escape digestion in the upper gastrointestinal tract. Once these carbohydrates reach the colon, their fermentation produces short-chain fatty acids (SCFAs) such as acetate, propionate, and butyrate. These SCFAs serve as a vital energy source for enterocytes and are crucial for maintaining gut function and immune homeostasis (Sivaprakasam et al., 2016).

Among the SCFAs, butyrate is especially beneficial for gut health. It is the primary energy source for intestinal epithelial cells and provides a first line of cellular defence against pathogens (Hodgkinson et al., 2023; Jacobson et al., 2018). Butyrate also strengthens the intestinal barrier, mitigates proinflammatory signalling, and modulates T-cell responses through epigenetic regulation via the inhibition of histone deacetylases (HDACs) (Lamas et al., 2019; Bach Knudsen et al., 2018). Recent studies also suggest that propionate's cellular mechanisms differ from those of butyrate, indicating specific metabolic roles (Kilner et al., 2012). Besides its function as an energy source, propionate provides cardiovascular protection by reducing hepatic lipid synthesis and blood cholesterol levels. It exerts several other beneficial physiological effects, including anti-lipogenic, cholesterol-lowering, anti-inflammatory, and anti-carcinogenic actions (Hosseini et al., 2011; Vinolo et al., 2011). Another important property is its ability to enhance satiety and modulate appetite through the activation of free fatty acid receptors FFA2 and FFA3, an effect of increasing interest given the global rise in obesity (Arora et al., 2011; Han et al., 2024).

Butyrate can be produced by colonic bacteria either as the end product of carbohydrate fermentation, through the metabolism of lactate and acetate generated by primary fermenters, or as final product of few acetogens which use C1 compounds through the Wood-Ljungdahl pathway (Vital et al., 2014; Litty and Müller, 2021; Trischler et al., 2022). By utilising these fermentation products, butyrate-producing bacteria help

prevent their accumulation, thus stabilising the gut environment. Four main pathways in the gut lead to butyrate production, starting from acetyl-CoA, glutarate, succinate semialdehyde, and lysine (Vital et al., 2014) (Figure 1A). The butyrate fermentation from acetyl-CoA is the most prevalent among gut microbes, especially *Clostridia*, involving the condensation of two acetyl-CoA molecules into acetoacetyl-CoA by thiolase and its subsequent reduction. Crucially, all four pathways converge with the reduction of crotonyl-CoA to butyryl-CoA. The final step, the conversion of butyryl-CoA into butyrate, is predominantly mediated by the enzyme butyryl-CoA:acetate-CoA transferase, encoded by the *but* gene. A less common reaction involves butyrate kinase, encoded by the *buk* gene (Vital et al., 2014; Louis et al., 2004).

Three main biochemical pathways contribute to propionate production: succinate, acrylate, and propanediol pathways (Reichardt et al., 2014) (Figure 1B). In the first pathway, succinate serves as the substrate for propionyl-CoA, being mainly sourced from the breakdown of sugars and proteins, though it is also a product of autotrophic growth in acetogens. (Reichardt et al., 2014; Shin et al., 2016; Fernández-Veledo and Vendrell, 2019). The acrylate pathway utilises lactate, converting it via lactoyl-CoA and acryloyl-CoA intermediates, with the key reaction catalysed by lactoyl-CoA dehydratase encoded by the *lcdA* gene. The propanediol pathway processes deoxy sugars into 1,2-propanediol, which is then metabolised to propionyl-CoA. All three pathways ultimately converge at the formation of propionyl-CoA, which is subsequently converted to propionate.

The cross-feeding of butyrogenic and propionigenic bacteria on SCFAs produced by other intestinal bacteria is also well documented (Louis and Flint, 2017; Duncan et al., 2004). Acetate utilisation can occur in bacteria via two distinct pathways: one involving acetate kinase and phosphate acetyltransferase, and another mediated by AMP-forming acetyl-CoA synthetase (Sun et al., 2020), whereas lactate utilisation involves D- and L-lactate dehydrogenases using NAD⁺ or other electron acceptors, together with pyruvate-ferredoxin oxidoreductase (Duncan et al., 2004).

Proteins also serve as a significant carbon and energy source for some colonic *Clostridia*, which partake in the breakdown and fermentation of undigested proteins and peptides (Amaretti et al., 2019; Raimondi et al., 2021). This process primarily produces carbon dioxide, ammonia, and organic acids, with some amino acids also contributing to butyrate or propionate formation (Raimondi et al., 2021; Barker, 1981; Riedel et al., 2017). For instance, lysine fermentation by the commensal *Intestinimonas* and related taxa results in butyrate production (Bui et al., 2015, 2020). Similarly, *Clostridioides difficile* was reported to produce butyrate from glutamate-derived succinate semialdehyde, while alanine fermentation in *Clostridium propionicum* led to propionate production (Buckel, 2001; Gregory et al., 2021).

The present study characterised all the intestinal *Clostridia* catalogued in the Unified Human Gastrointestinal Genome database (UHGG; Almeida et al., 2021), which comprises 4,744 representative species from the human gut microbiome, to resolve taxonomic misclassifications and refine phylogenetic boundaries within this class. The manageable subset of intestinal *Clostridia* in UHGG facilitated a systematic genome-wide analysis combining phylogenomic and genome-based criteria,

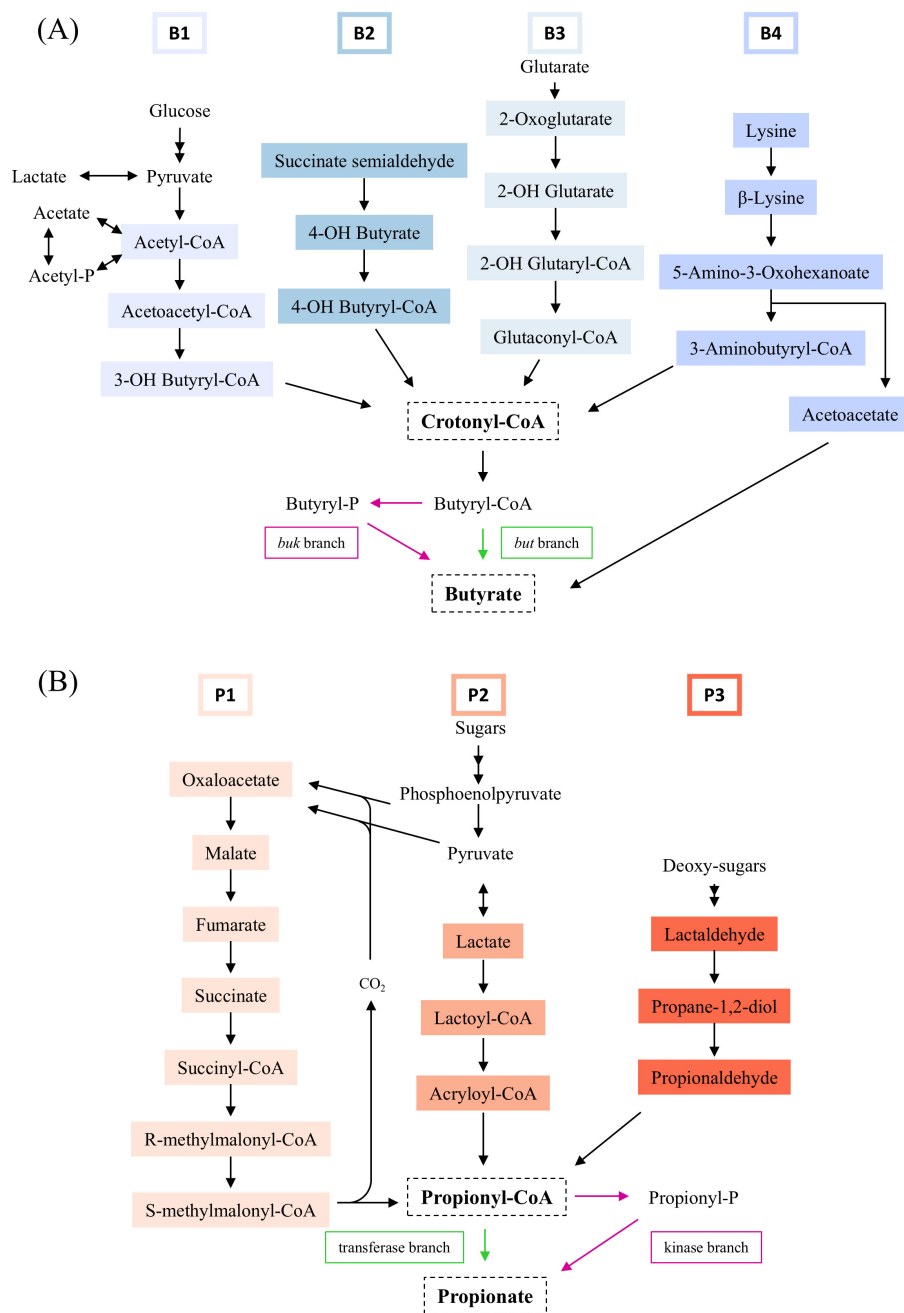


FIGURE 1

Panel A. Pathways for butyrate formation: B1, acetyl-CoA (thiolase); B2, succinate semialdehyde; B3, glutamate; B4, lysine. **Panel B.** Pathways for propionate formation: P1, succinate; P2, acrylate; P3, propanediol.

providing phylogenomic insight into hundreds of gut-associated *Clostridia* species. Given the significant role of intestinal *Clostridia* as SCFA producers, the main metabolic pathways related to butyrate and propionate production were investigated. The aim was to integrate existing knowledge of previously characterised butyrate-/propionate-producing species (Pryde et al., 2002; Louis and Flint, 2009; Louis et al., 2010; Louis and Flint, 2017; Magnúsdóttir et al., 2017; Zou et al., 2019) with poorly characterised and uncultivated ones, including metagenome-assembled genomes (MAGs). Finally, to estimate the potential impact of the clostridial community on intestinal butyrate and propionate levels, the abundances of *Clostridia* were

assessed through the analysis of 151 faecal metagenomes from healthy subjects.

2 Materials and methods

2.1 Genomes retrieval

Clostridia genomes ($n = 1,897$) were retrieved from the UHGG catalogue v2.0.1 (http://ftp.ebi.ac.uk/pub/databases/metagenomics/mgnify_genomes/human-gut/v2.0.1/species_catalogue) and their taxonomic attribution was updated using

GTDB-Tk v2.3.2 (classify_wf function) (Parks et al., 2022; Chaumeil et al., 2022) to align with GTDB taxonomy release R214. Representative genomes were subsequently obtained from GTDB (R214) when GTDB-Tk assignment allowed unambiguous matching ($n = 1,706$). For genomes with truncated or incomplete taxonomy ($n = 191$), UHGG v2.0.1 representatives were retained. GTDB genomes were prioritised due to their superior quality in terms of completeness and lower contamination levels. Among the 1,897 genomes, 404 were identified as representative species with established binomial nomenclature.

2.2 Phylogenomic analysis

To reconstruct the phylogeny of intestinal *Clostridia*, we identified, concatenated, and aligned the 120 ubiquitous single-copy proteins defined in the GTDB, using tools from GTDB-Tk v2.3.2 (Chaumeil et al., 2022). Phylogenetic trees were inferred using the PROTGAMMAAUTO model implemented in RAxML v8.2.12 (Stamatakis, 2014), rooted with *Bacillus licheniformis* ATCC 14580 (GCF_000011645.1), and visualised using iTOL v7.2.2 (Letunic and Bork, 2021). Bootstrapping was conducted with 1,000 replicates to assess tree robustness.

2.3 Genome-based relatedness indices

To assess genomic relationships among species with established nomenclature, the 404 representative species genomes were annotated using Prokka v1.14.5 (Seemann, 2014) and subjected to pairwise comparisons. The pairwise amino acid identity (AAI) was calculated using the EzAAI pipeline (Kim et al., 2021). The percentage of conserved proteins (POCP) was determined using the formula: $[(C1 + C2)/(T1 + T2)] \times 100$, where C1 and C2 denote the number of conserved proteins shared between the two genomes being compared, and T1 and T2 represent the total number of proteins in the respective genomes (Qin et al., 2014). Conserved proteins were identified using Proteinortho v6.3.0 (Lechner et al., 2011), applying thresholds of 40% sequence identity and a minimum alignment length of 50%. Heatmaps were produced for AAI and POCP results, and it was assessed consistency with the current taxonomic attribution.

2.4 Screening for short-chain fatty acids producers

Butyrate and propionate biosynthetic pathways were screened across all 1,897 genomes using GapSeq v1.4.0 (Zimmermann et al., 2021) with default parameters. For butyrate, four pathways converging to crotonyl-CoA (acetyl-CoA, glutarate, lysine, and succinate semialdehyde) were evaluated, along with terminal steps via butyryl-CoA:acetate CoA-transferase or butyrate kinase. For propionate, three pathways (succinate, acrylate, and propanediol) leading to propionyl-CoA were screened, along with terminal CoA-transferase or kinase reactions.

2.5 Metagenomic samples

Clostridia species were analysed in a dataset of 151 metagenomes derived from faecal samples of healthy subjects from different geographic provenance (Supplementary Table 1). These metagenomes are publicly available in NCBI Sequence Read Archive (SRA), with the accession numbers PRJNA557323, PRJNA504891, PRJEB27308, PRJNA529124, PRJNA485056, PRJNA278393, PRJNA375935, PRJDB4176, PRJNA328899, PRJEB17784, PRJNA268964, and PRJEB7369 (Supplementary Table 1). The datasets were generated using whole-genome shotgun sequencing on Illumina paired-end platforms, yielding between 9.2×10^6 and 1.0×10^{10} reads per sample, with read lengths ranging from 100 to 150 base pairs.

2.6 Composition analysis

The microbial composition of the 151 metagenomes was determined using the Kraken2 taxonomic sequence classifier (Wood et al., 2019), followed by Bracken analysis (Lu et al., 2017). We used a custom Kraken2 database based on the UHGG collection (Almeida et al., 2021) to achieve higher resolution in microbial profiling. Bracken was employed to refine Kraken2's initial classifications by re-estimating the number of reads assigned to each species in a sample. All analyses were conducted using the default parameters for k-mer length, minimizer length, and minimizer spacing.

3 Results

3.1 Phylogenomic resolution of intestinal *Clostridia*

The UHGG database encompassed genomes ascribed to 1,897 species of *Clostridia*, among which 246 with recognized binomial nomenclature (hereinafter indicated as recognized species, RS). Reclassification according to GTDB R214 returned 404 RS. The remaining 1,493 *Clostridia* genomes were not ascribed to any recognized species (no species, hereinafter indicated as NS) (Supplementary Spreadsheet 1). The representative genomes obtained following reclassification ($n = 1,706$ from GTDB R214; $n = 191$ from UHGG) exhibited high completeness, with 1,170 genomes (61.7%) showing >90% completeness and an additional 574 genomes (30.3%) exceeding 70% completeness. Notably, among the 404 RS genomes used for comparative analyses, 375 (92.8%) showed >90% completeness, with all remaining genomes exceeding 70%, ensuring high reliability for genome-based relatedness assessments and biochemical pathways reconstruction (Supplementary Spreadsheet 1).

The 1,897 species of *Clostridia* belonged to 19 different orders. The alignment of 120 ubiquitous single-copy proteins yielded the phylogenetic tree reported in Supplementary Figure 1 and collapsed at the order level in Figure 2. *Eubacteriales* separated at node 1 from lineage A that gave rise to all the other 18 orders. In correspondence of node 2, lineage A split into lineage

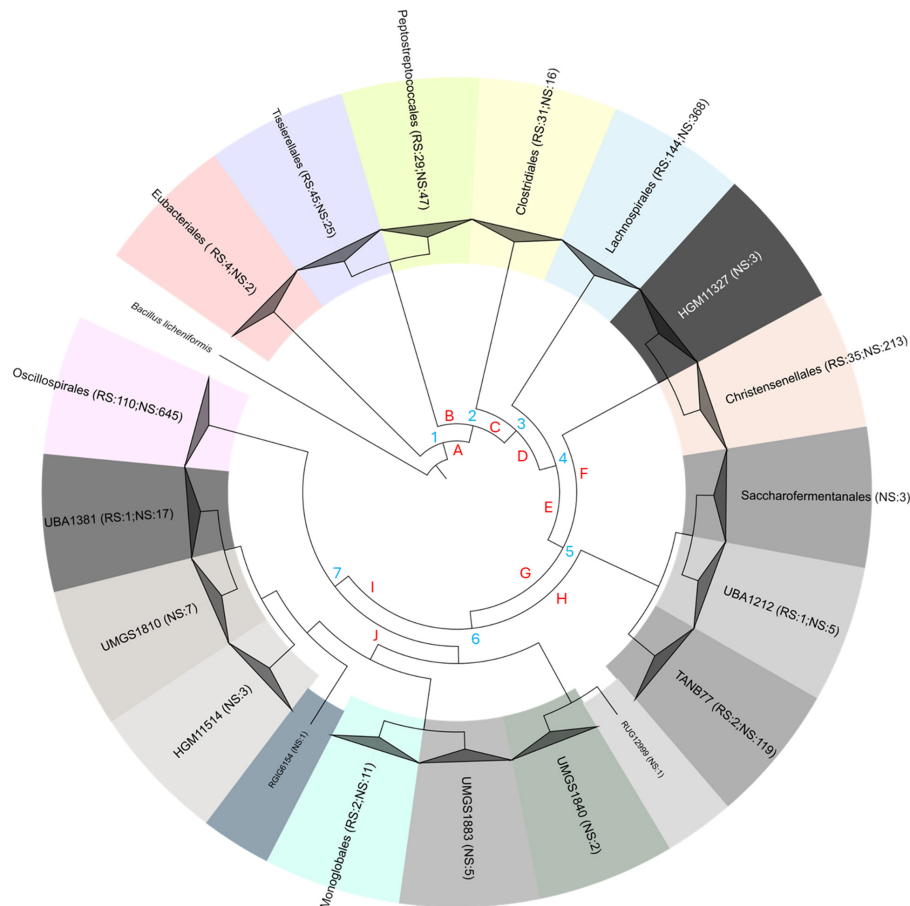


FIGURE 2

Phylogenetic tree of the 1,897 *Clostridia* retrieved from UHGG database, computed from the alignment of 120 ubiquitous single-copy proteins. Branch length is ignored, and leaves are collapsed to the order level for clarity of visualization. Lineages are in red, nodes in cyan. For each order, the number of NS and RS is given in brackets.

B that generated the orders *Peptostreptococcales* and *Tissierellales*, separated from a major one (C), which evolved in the other orders. Lineage C split at node 3 into *Clostridiales* and into branch D, which in turn split at node 4 into *Lachnospirales* from branch E. Taxa bipartition of branch E (node 5) led to branch F (orders HGM11327 and *Christensenellales*) and to branch G. In branch G, node 6 evolved to branch H (*Saccharofermentanales*, UBA1212, and TANB77) and to branch I. In branch I, node 7 gave rise to *Oscillospirales* and to branch J, which led to the orders RUG12999, UMG1840, UMG1883, *Monoglobales*, RGIG6154, HGM11514, UMG1810, and UBA1381.

Most species of *Clostridia* belonged to the orders *Oscillospirales* (755) and *Lachnospirales* (512). Other species mainly belonged to *Christensenellales* (248), *Peptostreptococcales* (76), *Clostridiales* (47), *Monoglobales* (13), and *Eubacteriales* (6). Most of the species attributed to *Oscillospirales* were NS (85.4%, [Supplementary Table 2](#)), and consistently, NS were abundant within the main families *Acutalibacteraceae*, *Oscillospiraceae*, and *Ruminococcaceae* (85.9, 90.8, and 74.3%, respectively). The majority of *Lachnospirales* were ascribed to *Lachnospiraceae* (477/512) and represented by NS (71.9%). On the other hand, *Clostridiaceae* (*Clostridiales*), *Peptoniphilaceae* (*Tissierellales*), and

Peptostreptococcaceae (*Peptostreptococcales*) mainly contained RS. Most of the species of *Christensenellales* were NS (85.9%). The richest family herein observed, *Borkfalkiaceae*, comprised 60 species with only 10 RS. The other main families (> 20 species; CAG-74, UBA1242, CAG-138, and CAG-917) commonly included NS. The order TANB77 evolved in 3 families and mostly included NS (119/121). Its largest family, CAG-508, encompassed 109 species, including the 2 RS *Merdicola faecigallinarum* and *Scatovivens faecipullorum*. Eight orders (UMGS1840, UMG1883, HGM11514, RUG12999, UMG1810, HGM11327, RGIG6154, and *Saccharofermentanales*) did not encompass any RS ([Figure 2](#), [Supplementary Table 2](#)).

3.2 Families and genera delineation

Taxonomic investigation focused on 404 representative species (RS). These RS were assigned to 11 GTDB orders and subjected to pairwise AAI comparisons to evaluate similarity within families and to delineate genera.

Among the 404 RS analyzed, the most represented order was *Lachnospirales* with 144 RS, followed by *Oscillospirales* with 110 RS; *Tissierellales* and *Christensenellales* included 45 and 35 RS,

respectively, while *Clostridiales* and *Peptostreptococcales* comprised 31 and 29 RS. The remaining 8 RS were distributed across minor orders with fewer than 5 representatives. Overall, the 404 RS were assigned to 35 GTDB families, 8 of which contained more than 10 RS: *Lachnospiraceae* (138 RS, order *Lachnospirales*); *Ruminococcaceae*, *Acetivibacteraceae*, and *Oscillospiraceae* (44, 31, and 22 RS, respectively, order *Oscillospirales*); *Peptoniphilaceae* (42 RS, order *Tissierellales*); *Clostridiaceae* (31 RS, order *Clostridiales*); *Anaerovoracaceae* and *Peptostreptococcaceae* (14 RS each, order *Peptostreptococcales*). The lowest AAI value (minAAI) between RS was calculated for each family (Supplementary Spreadsheet 2). Among all intestinal clostridial families, *Peptoniphilaceae* exhibited the lowest minAAI (43.9%). Other families with more than 10 RS showed minAAI values up to 55.3%, as observed in *Peptostreptococcaceae*.

Genera were delineated considering the thresholds of 65% and 50% for AAI and POCP, respectively, according to GTDB genera nomenclature. The 404 RS belonged to 200 GTDB genera, 123 of which were singletons. Of the 77 genera with more than one species, 51 showed both minAAI and minPOCP values above the thresholds (Table 1, A), 12 met the minAAI threshold but not the minPOCP threshold (Table 1, B), 6 genera met the POCP threshold but not the AAI (Table 1, C), while 8 genera did not meet either threshold (Table 1, D).

In some genera (e.g. *Anaerotruncus*, *Dysosmobacter*, *Butyrivibrio*, etc.), the under-threshold values were attributable to a single species showing lower similarity with one or more members of its genus (Supplementary Table 3, Supplementary Spreadsheet 2). In other genera (e.g. *Blautia*, *Blautia* A, *Clostridium*, *Clostridium* J, *Eisenbergiella*, *Enterocloster*, *Ezakiella*, *Lachnoclostridium* A, *Peptoniphilus* A, *Sarcina*, etc.) the exclusion of certain species and/or the subdivision of the nominal genus into two or three distinct groups would result in clusters that meet the established thresholds for genus delineation (Supplementary Table 3). Additionally, the merging of some species belonging to different nominal genera could give rise to 4 putative genera exceeding both thresholds (Supplementary Table 3).

3.3 Profiling of clostridial community in gut microbiomes

The abundance and prevalence of the 1,897 intestinal *Clostridia* species within the human gut microbiota was evaluated through the analysis of 151 faecal samples from healthy subjects (Supplementary Table 1). The profiling carried out with Kraken2 and Bracken identified 4,602 GTDB bacterial species, including all the 1,897 species of *Clostridia*, each of them being identified in no fewer than 10 samples. Of the 404 RS, 119 were detected in all 151 metagenomes. The abundance of *Clostridia* ranged between 9.7 and 88.1% of total bacteria, with a mean of 44.3%. On average, the 1,493 NS accounted for 56.1% of the total relative abundance of *Clostridia*.

The orders *Oscillospirales* and *Lachnospirales*, which encompassed most *Clostridia* species, were quantitatively the most abundant (on average 22.6 and 18.3%, respectively). The 50 most abundant species of *Clostridia* generally

belonged to *Lachnospirales* (26) and *Oscillospirales* (23) and were mostly RS (33) (Figure 3). *Agathobacter rectalis*, *Faecalibacterium longum*, *Faecalibacterium duncaniae*, and *Agathobacter faecis* accounted on average for 1.55, 0.80, 0.78, and 0.76%, respectively, and reached remarkably high levels in single samples (13.5, 14.4, 8.5, and 21.6%, respectively) (Supplementary Spreadsheet 1). Other frequently occurring species that reached high abundance in individual samples were *Ruminococcus E* sp003526955 (max 13.9%), *Mediterraneibacter torques* (max 11.1%), *Butyrivibrio A* crossotus (max 11.1%), and *Blautia A* wexlerae (max 10.2%). Other dominant species, with a mean abundance ranging between 0.4 and 0.7%, were *Faecalibacterium prausnitzii*, ER4 sp000765235 (the most abundant NS), *Gemmiger quicibialis*, *Acetatifactor intestinalis*, *Faecalibacterium prausnitzii* D, and *Fusicatenibacter saccharivorans*.

3.4 Reconstruction of butyrate pathways

3.4.1 Key reactions from crotonyl-CoA to butyrate: the but and buk branches

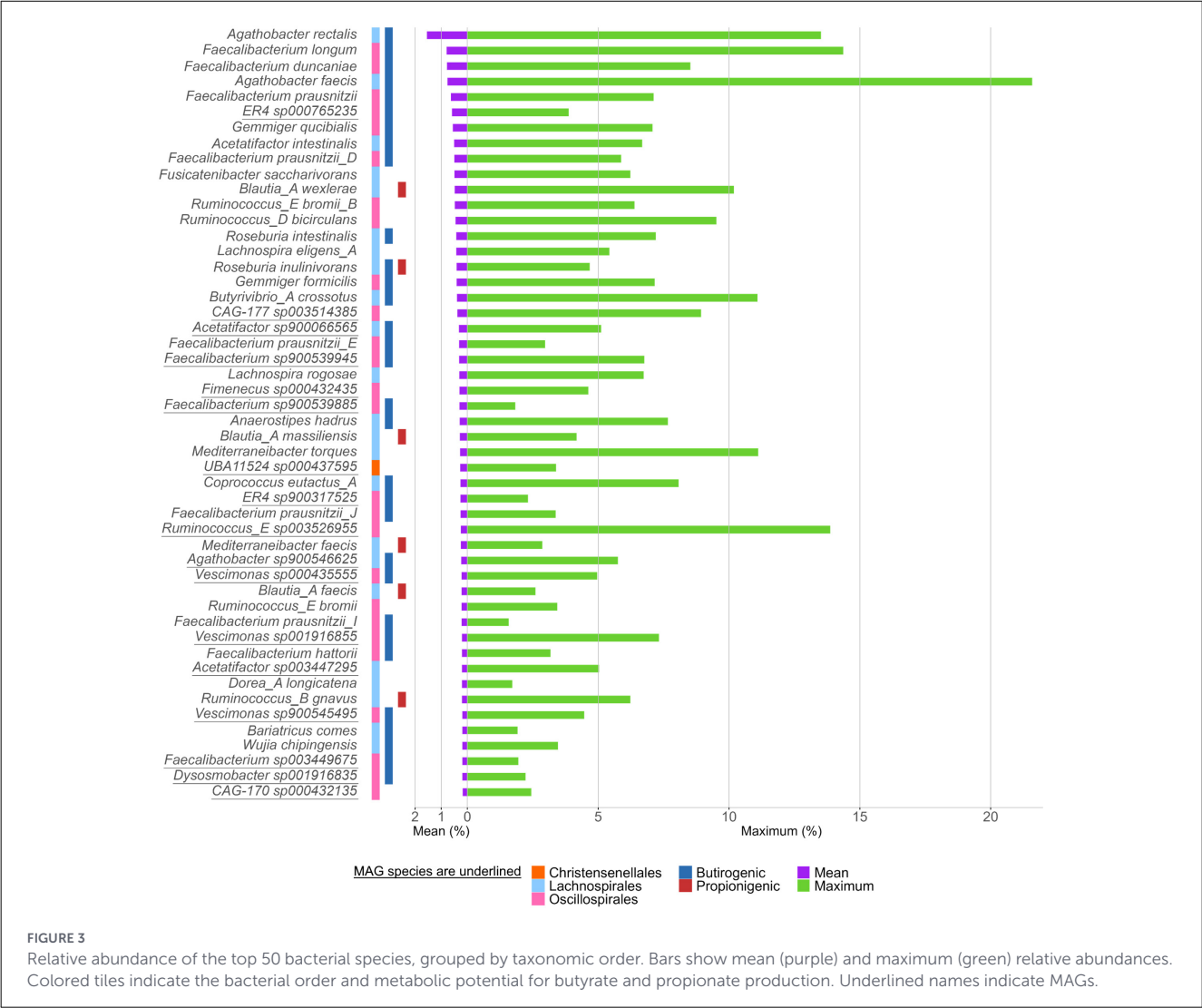
We reconstructed the metabolic pathways for butyrate production across the *Clostridia* class by screening for genes encoding butyryl-CoA dehydrogenase (which converts crotonyl-CoA to butyryl-CoA) and the downstream enzymes that complete butyrate synthesis. Two alternative routes exist: the butyryl-CoA:acetate-CoA transferase pathway (*but* branch) and the phosphate butyryltransferase/butyrate kinase pathway (*buk* branch). Among 771 genomes encoding butyryl-CoA dehydrogenase, 687 genomes (234 RS) possessed complete pathways to butyrate: 605 species encoded the *but* branch, 268 the *buk* branch, and 187 encoded both pathways. However, the distribution of these pathways varied widely across orders of *Clostridia* (Supplementary Spreadsheet 3).

Nearly all *Clostridiales* possessed both branches, whereas the vast majority of *Christensenellales* lacked them entirely, except for the species of *Christensenella*, *Avichristensenella*, *Scybalosoma*, and a few *Limiplasma*, which encoded the *but* branch.

Approximately half of *Lachnospirales* species harbored *but* or *buk* branches. The *but* branch was generally complete in most or all species of *Anaerotruncus*, *Agathobacter*, *Alectryococcobium*, *Anaerobutyricum*, *Anaerostipes*, *Anaerocolumna*, *Butyrivibrio*, *Clostridium* Q, *Coprococcus* A, *Eubacterium* F, *Copromonas*, *Frisingicoccus*, *Lachnoanaerobaculum*, *Mediterraneibacter* A, *Metalachnospira*, *Murimonas*, *Pararoseburia*, *Roseburia*, *Pseudoroseburia*, *Scybalocola*, *Dorea* A, and in some species of *Blautia*, *Lacrimispora*, *Oribacterium*, and *Ventrisoma*. The *buk* branch characterized all or most species of *Acetatifactor*, *Butyrivibrio*, *Eubacterium* G, *Wujia*, and some *Extibacter*. Both branches were present in species of *Coprococcus*, *Eisenbergiella*, *Eubacterium* I, *Hungatella*, *Scatomonas*, and various other genera. The genus *Enterocloster* was remarkably heterogeneous, comprising NS encoding *but*, *buk*, or both.

TABLE 1 Delineation of nominal genera according to AAI and POCP thresholds.

Group	Taxa	minAAI	minPOCP
A	<i>Agathobacter</i> , <i>Alectryocaccommicrobium</i> , <i>Anaerobutyricum</i> , <i>Anaerostipes</i> , <i>Anaerotignum</i> , <i>Baileyella</i> , <i>Borkfalkia</i> , <i>Christensenella</i> , <i>Clostridioides</i> , <i>Clostridium</i> E, <i>Clostridium</i> F, <i>Clostridium</i> G, <i>Coproccoccus</i> , <i>Coproccoccus</i> A, <i>Diplocloster</i> , <i>Dorea</i> A, <i>Egerieicola</i> , <i>Eubacterium</i> , <i>Eubacterium</i> I, <i>Eutepia</i> , <i>Extibacter</i> , <i>Faecalibacterium</i> , <i>Fenollaria</i> , <i>Fimimonas</i> , <i>Finegoldia</i> , <i>Gallimonas</i> , <i>Gemmiger</i> , <i>Hungatella</i> , <i>Intestinimonas</i> , <i>Kallipyga</i> , <i>Lachnoanaerobaculum</i> , <i>Lachnoclostridium</i> B, <i>Lachnospira</i> , <i>Lentihominibacter</i> , <i>Murdochella</i> , <i>Muricomes</i> , <i>Negativibacillus</i> , <i>Oliverpabstia</i> , <i>Paraclostridium</i> , <i>Parvimonas</i> , <i>Peptostreptococcus</i> , <i>Pseudoflavonifractor</i> , <i>Roseburia</i> , <i>Ruminococcus</i> E, <i>Scatavimonas</i> , <i>Scatomorpha</i> , <i>Sellimonas</i> , <i>Stomatobaculum</i> , <i>Terrisporobacter</i> , <i>Vescimonas</i> , <i>Zhenpiania</i>	YES	YES
B	<i>Acutalibacter</i> , <i>Agathobaculum</i> , <i>Anaerotruncus</i> , <i>Bariatricus</i> , <i>Blautia</i> , <i>Clostridium</i> Q, <i>Dysosmobacter</i> , <i>Eisenbergiella</i> , <i>Faecalimonas</i> , <i>Fournierella</i> , <i>Limiplasma</i> , <i>Mediterraneibacter</i>	YES	NO
C	<i>Anaerococcus</i> , <i>Butyribacter</i> , <i>Eubacterium</i> R, <i>Ezakiella</i> , <i>Massilioclostridium</i> , <i>Peptoniphilus</i> A	NO	YES
D	<i>Blautia</i> A, <i>Clostridium</i> , <i>Clostridium</i> J, <i>Enterocloster</i> , <i>Lachnoclostridium</i> A, <i>Marvinbryantia</i> , <i>Ruthenibacterium</i> , <i>Sarcina</i>	NO	NO



Within *Oscillospirales*, *but* and *buk* branches were less common but still widespread among known butyrate producers. The *but* branch characterized most species of *Agathobaculum*, *Butyricoccus*, *Dysosmobacter*, *Eutepia*, *Intestinimonas*, *Onthomonas*, *Pseudoflavonifractor*, *Scatomorpha*, *Pseudoscilispira*, *Faecalibacterium*, *Pygmaibacter*, and some *Anaerotruncus*, *Angelikisella*, and *Limivacinus*. The *buk* branch was present in most *Gemmiger* and in some

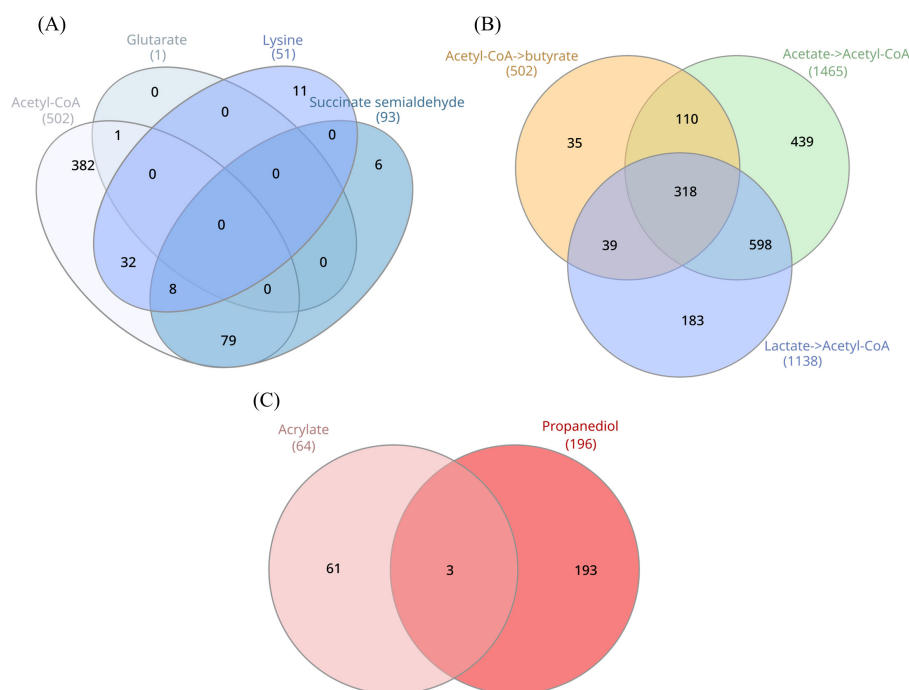


FIGURE 4

Distribution of predicted complete metabolic pathways in clostridial species (RS + NS). Pathways leading to crotonyl-CoA and then to butyrate (A). Acetyl-CoA to butyrate, acetate to acetyl-CoA, and lactate to acetyl-CoA pathways (B). Propionyl-CoA to propionate pathways (C).

Fournierella, *Pseudobutyricoccus*, *Anaerotruncus*, and *Angelikisella*. Both branches were detected in most *Vescimonas*, *Lawsonibacter*, *Flavonifractor*, *Ventrusia*, and in some *Enterenecus*, *Dysosmobacter*.

In *Peptostreptococcales*, the majority of species possessed *but* and/or *buk*, except for a few genera (*Romboutsia*, *Peptacetobacter*, *Filifactor*) and a few NS. Most members of *Clostridioides*, *Paraclostridium*, *Peptostreptococcus*, and related taxa encoded both branches. The *but* branch was complete in *Eubacterium T*, *Fimisoma*, *Lentihominibacter*, and *Criibacterium*, while *Intestinibacter* encoded only the *buk* branch.

Saccharofermentanales, a small order with only three NS, included two species predicted to harbor both *but* and *buk*.

Most *Tissierellales* were predicted to produce butyrate, with the main exceptions being *Ezakiella*, *Finegoldia*, *Fenollaria*, *Neofamilia*, and *Parvimonas*. The *but* branch was found to be complete in *Anaerotruncus*, *Anaerococcus*, *Kallipiga*, *Murdochiella*, *Peptoniphilus A*, *Peptoniphilus B*, *Peptoniphilus C*, and *Peptoniphilus E*. Both branches were detected in *Tissierella*, *Sporanaerobacter*, *Anaerosalibacter*, *Sedimentibacter*, and several other NS.

Neither branch was detected in *Monoglobales* or in other unclassified clostridial lineages (e.g., TANB77, UBA121).

Notably, several species encoded *but* or *buk* genes but lacked butyryl-CoA dehydrogenase or phosphate butyryltransferase. Specifically, 193 species possessed *but* without butyryl-CoA dehydrogenase, mainly among *Christensenellales*, *Lachnospirales*, and *Oscillospirales*, while 125 species encoded *buk* and lacked one or both of these upstream enzymes.

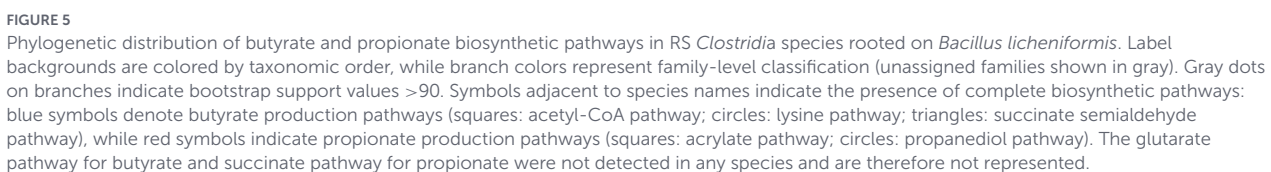
3.4.2 Routes yielding crotonyl-CoA: acetyl-CoA, lysine, succinate semialdehyde, and glutarate pathways

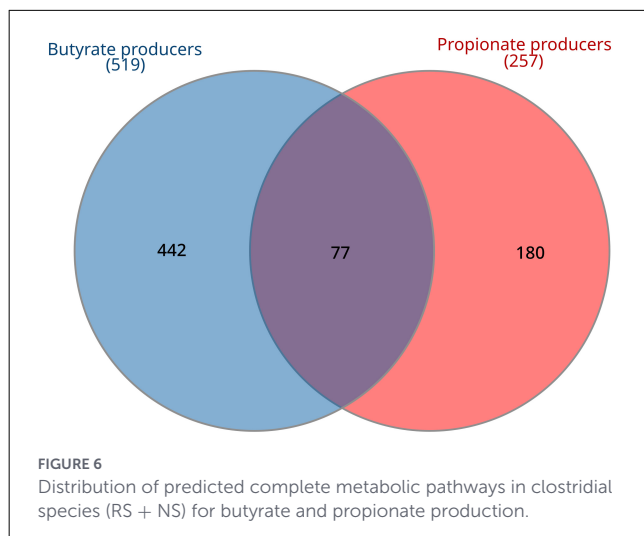
The main metabolic routes yielding crotonyl-CoA, i.e. acetyl-CoA (thiolase), lysine, succinate semialdehyde, and glutarate pathways, were reconstructed in the clostridial genomes. For 519 of the 687 species harboring complete *but* or *buk* branches, at least one complete pathway yielding crotonyl-CoA was identified (Figure 4). The number increased to 569 including those species lacking one metabolic block (Supplementary Figure 2A).

The metabolic route transforming acetyl-CoA into crotonyl-CoA was complete in 521 species, of which 188 were RS (Figure 5). The presence of the genes participating in the transformation of acetyl-CoA to crotonyl-CoA was associated with the presence of complete *buk* and/or *but* branches to yield butyrate in 502 species (Figure 4A). Conversely, 185 species, widespread across all the orders, carrying either the *buk* or *but* branch, showed an incomplete acetyl-CoA-to-crotonyl-CoA pathway, lacking at least one metabolic block.

The pathway transforming lysine into crotonyl-CoA, including *ato*, was complete in 51 species (Figure 4A). The presence of this pathway was always accompanied by a complete *but* branch and in many cases also by the *buk* branch. In 11 species (*Paraclostridium dentum* being the only RS), lysine-to-butyrate conversion was accompanied by interruptions in the acetyl-CoA-to-crotonyl-CoA route.

The succinate semialdehyde pathway was complete in 93 species (Figure 4). RS harboring the complete pathway were found in *Christensenellales* (*Christensenella*, *Scybalosoma*), *Clostridiales*





(*Clostridium*, *Clostridium B*, *Clostridium J*, *Clostridium L*), *Eubacteriales* (*Anaerofustis*), *Lachnospirales* (*Anaerostipes*, *Bariatricus*, *Clostridium Q*, *Enterocloster*, *Ruminococcus G*, *Scatomonas*), *Oscillospirales* (*Agathobaculum*, *Butyricicoccus*, *Butyricicoccus A*, *Dysosmobacter*, *Eutepia*, *Flavonifractor*, *Anaerotruncus*), *Peptostreptococcales* (*Bacilliculturomica*, *Gallibacter*, *Lentihominibacter*, *Zhenpiania*, *Clostridioides*), and *Tissierellales* (*Anaerosalibacter*, *Tissierella*) (Figure 5). Except for rare cases, the succinate semialdehyde pathway co-occurred with a complete *but* branch and, less frequently, *buk*. Most species also encoded acetyl-CoA-to-crotonyl-CoA pathway (Supplementary Spreadsheet 3).

The glutarate pathway was complete only in one NS of *Saccharofermentanales*, while 76 species displayed it with one block missing, usually 2-oxoglutarate reductase (EC 1.1.1.399). These species mainly belonged to *Oscillospirales* (e.g. *Lawsonibacter*, *Oscillibacter*, *Pygmaibacter*, *Clostridium E*), *Lachnospirales* (e.g. *Clostridium Q*, *Copromonas*, *Stomatobaculum*, *Ventrisoma*), and *Tissierellales* (e.g. *Anaerococcus*, *Kallipyga*, *Murdochiella*, *Peptoniphilus A*). With few exceptions, the glutarate pathway co-occurred with a complete *but* branch, in some cases *buk*, and a full acetyl-CoA-to-crotonyl-CoA route.

The potential utilization of acetate and lactate for butyrate production was also investigated, as both routes converge into the thiolase pathway. Enzymes involved in acetate and lactate metabolism were identified in numerous species (1,465 and 1,138, respectively) (Figure 4B), showing substantial overlap with each other and with those predicted to convert acetyl-CoA into butyrate. Of 502 species predicted to convert acetyl-CoA into butyrate, 428 had the genes for acetate metabolism and 357 had those for lactate metabolism, with 318 presenting both routes (Figure 4B).

3.5 Reconstruction of propionate pathways

The metabolic routes yielding propionate were inferred. The ability to convert propionyl-CoA to propionate through

two possible routes was predicted in 1,469 species (361 RS). The route involving phosphate propionyltransferase followed by propionate kinase was complete in 1,423 species, the one involving propionyl-CoA:lactate/acetate transferase was found in 255 species, with 209 species encoding both branches (Supplementary Spreadsheet 3).

The main metabolic routes yielding propionyl-CoA, i.e. succinate, acrylate, and propanediol pathways, were reconstructed. Of the 1,469 species predicted to get propionate from propionyl-CoA, a complete pathway was predicted only in 257 species. The number increased to 659 including those species lacking one metabolic block (Supplementary Figure 2B). Although methylmalonyl-CoA carboxytransferase was widespread across the class *Clostridia* (predicted in 1,458 species), the succinate pathway leading to propionyl-CoA was consistently interrupted at the level of at least three metabolic blocks.

The acrylate pathway leading to propionyl-CoA was complete in 64 species (24 RS), but this number increased to 315 when including those lacking one metabolic block. The missing steps most frequently involved propionate:lactoyl-CoA transferase (115 species) or lactoyl-CoA dehydratase (134 species). The 64 species with a complete pathway were distributed across the orders *Clostridiales* (e.g. *Clostridium*, *Clostridium F*, *Clostridium H*, and *Clostridium X*, *Eubacteriales* (*Garciella*), *Lachnospirales* (e.g. *Anaerotignum*, *Metalachnospira*, *Clostridium Q*, *Coprococcus A*, *Frisingicoccus*, *Ventrisoma*), *Oscillospirales* (e.g. *Clostridium E*, *Agathobaculum*, *Eutepia*, *Intestinimonas*, *Lawsonibacter*, *Scatomorpha*, *Vescimonas*), *Peptostreptococcales* (e.g. *Lentihominibacter*, *Zhenpiania*, *Peptostreptococcus*), and *Tissierellales* (*Anaerococcus*, *Murdochiella*, *Peptoniphilus*, *Sedimentibacter*). All these 64 species were predicted to produce propionate from propionyl-CoA. Notably, 57 out of the 64 species also presented at least one lactate dehydrogenase, enabling the connection between glycolysis and propionate production.

The propanediol pathway leading to propionyl-CoA was complete in 196 species (71 RS), but this number rose to 482 when including those lacking one block, most commonly propanal dehydrogenase (Figure 4C, Supplementary Figure 2B). Species harboring the complete pathway were distributed across *Christensenellales* (e.g. *Borkfalkia*, *Faecaligallichristensenella*, *Christensenella*), *Clostridiales* (e.g. *Clostridium*, *Clostridium G*, *Clostridium X*), *Eubacteriales* (*Eubacterium*), *Lachnospirales* (e.g. *Bianquea*, *Fecimorpha*, *Zhenhengia*, *Anaerobutyricum*, *Blautia A*, *Diplocloster*, *Eisenbergiella*, *Enterocloster*, *Lachnoclostridium*, *Luxibacter*, *Mediterraneibacter*, *Muricomes*, *Robinsonella*, *Roseburia*, *Ruminococcus*, *Scatomonas*, *Schaedlerella*), *Oscillospirales* (e.g. *Clostridium E*, *Flavonifractor*, *Harryflintia*, *Massiliclostridium*), *Peptostreptococcales* (*Bacilliculturomica*, *Emergencia*, *Terrisporobacter*), *Tissierellales* (*Anaerococcus*, *Peptoniphilus A*, *Anaerosalibacter*), and *UBA1381* (*Congzhengia*). The vast majority of these species were predicted to transform propionyl-CoA to propionate.

A total of 519 species harbored complete butyrate production pathways, 257 species possessed complete propionate production pathways, and 77 species could produce both SCFAs. 442 butyrate and 180 propionate producers exhibited exclusive production capabilities for only one of these SCFAs (Figure 6).

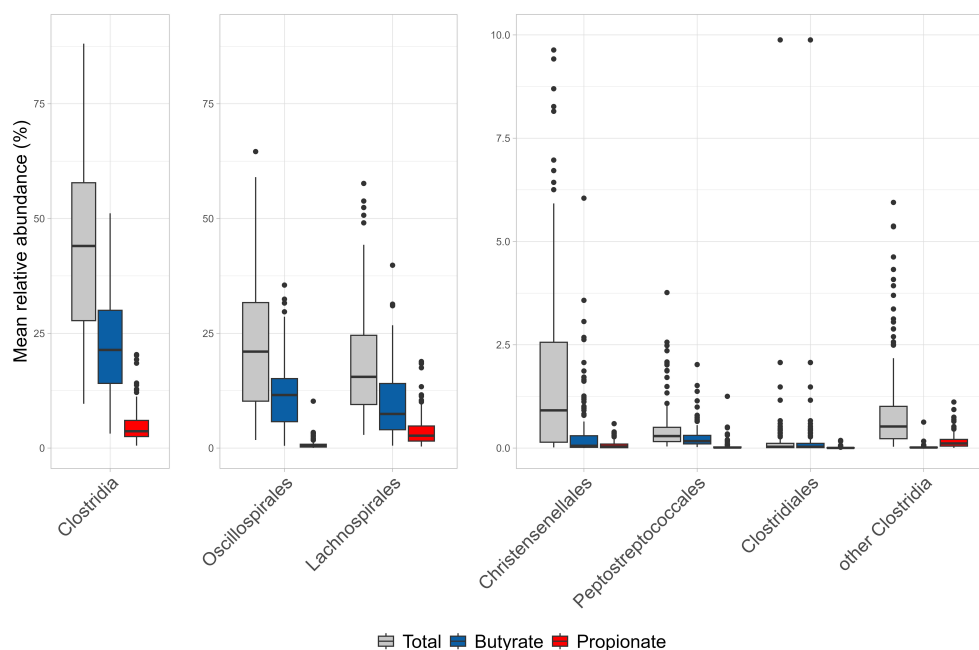


FIGURE 7

Relative abundance and pathway distribution of butyrate and propionate producers across *Clostridia* taxa. Boxplots showing the mean relative abundance (%) across samples for different taxonomic groups. In each sub-panel, gray boxes represent total abundance, blue boxes indicate butyrate producers, and red boxes indicate propionate producers. Whiskers extend to $1.5 \times$ the interquartile range (IQR) beyond the first and third quartiles.

3.6 Profiling of butyrate and propionate producing *Clostridia* in human gut microbiomes

Putative butyrate and propionate producing *Clostridia* were quantified based on the relative abundance of species within the microbiomes.

Clostridia potentially producing butyrate accounted on average for 22.2% (median = 21.4%), ranging from 3.2 to 51.1 % (Figure 7). With regard to the routes generating butyrate, the most represented was the acetyl-CoA pathway (median = 21.2%), while glutarate, lysine and succinate semialdehyde pathways were minor (each with a median < 1.5%). The *but* branch was predominant compared to *buk*, in terms of bacterial abundance (median = 19.9 and 7.0%, respectively) (Supplementary Figure 3).

The orders with the highest abundance of putative butyrate producers were *Oscillospirales* and *Lachnospirales*, where species with at least one complete butyrate pathway accounted on average for 11.9% (median = 11.6%) and 9.5% (median = 7.4%), respectively. The abundance of putative butyrate producers within *Christensenellales*, *Peptostreptococcales*, *Clostridiales*, and cumulatively the other clostridial orders presented a median relative abundance < 0.2%, much lower compared to that of *Oscillospirales* and *Lachnospirales* (Figure 7).

Of the 50 most abundant *Clostridia* species (Figure 3), 31 were potential butyrate producers, collectively accounting for 12.7% of the microbiome and representing more than half of putative butyrate producers (22.2%). As a whole, a minority of butyrogenic *Clostridia* (76/519) accounted for more than 80% of butyrate producers. The most abundant butyrate producers belonged to

the genera *Acetatifactor*, *Agathobacter*, *Faecalibacterium*, *Gemmiger*, and *Roseburia*.

The abundance of putative propionate producer *Clostridia* was lower compared to butyrate producers (Figure 7). Propionigenic *Clostridia* accounted on average for 4.9% (median = 3.7%). Of the two routes that were found complete in *Clostridia*, the propanediol pathway was the most represented (median = 3.0%), while acrylate pathway was the less abundant (median = 0.5%) (Suppl. Figure 3). *Lachnospirales* was the richest order in propionate producers (median = 2.7%), followed by *Oscillospirales* (median = 0.5%). The contribution of *Christensenellales*, *Peptostreptococcales*, *Clostridiales*, and cumulatively the other clostridial orders to the propionigenic community was lower, with abundances < 0.12%.

Among the 50 most abundant *Clostridia*, only 6 *Lachnospiraceae* species were putative propionate producers, all harboring the propanediol pathway: *Blautia A faecis*, *Blautia A. wexlerae*, *Blautia A. massiliensis*, *Mediterraneibacter faecis*, *Roseburia inulinivorans*, and *Ruminococcus B. gnavus*. Only *R. inulinivorans* was also able to produce butyrate.

RS and NS contributed similarly to the mean abundance of putative butyrate producers, accounting for 12.1 and 10.1%, respectively, whereas putative propionate producers accounted for 3.3 and 1.6% for RS and NS, respectively (Suppl. Spreadsheet 1).

4 Discussion

This study aimed to provide a comprehensive overview of all species within the class *Clostridia* identified in the human gut,

including both recognized species (RS) and those detected through metagenomic sequencing without formal species assignment (NS). With the aim of creating a valuable reference for future microbiome research and enabling a more accurate interpretation of *Clostridia* diversity within the human gut microbiome, the analysis encompassed phylogenomic relationships, identification of taxonomic inconsistencies, and metabolic reconstruction of SCFA production pathways, complemented by a quantitative profiling of *Clostridia* across 151 healthy human gut microbiomes.

The genomes of all the intestinal *Clostridia* included in the UHGG catalogue were retrieved and reclassified according to the GTDB taxonomy (R214), which assigned 404 species as RS and 1,493 as NS. We adopted the GTDB framework because it relies on a standardized and phylogenomically consistent approach based exclusively on whole-genome sequences and conserved single-copy marker proteins, ensuring objective and reproducible taxonomic assignments. In contrast, the taxonomy based on List of Prokaryotic names with Standing in Nomenclature (LPSN) also incorporates historical and phenotypic criteria. Some discrepancies still exist, likely due to delays in taxon revision and updating. For example, while GTDB recognizes the orders *Lachnospirales* and *Oscillospirales*, these have not yet been validated in LPSN and remain classified under the order *Eubacteriales*.

The orders *Lachnospirales* and *Oscillospirales*, covering the majority of *Clostridia* species, were mostly composed of NS, as already observed (Almeida et al., 2021). In these and the other major clostridial orders, the genomes of NS were smaller than those of RS (Supplementary Figure 4), supporting the hypothesis that uncultivable *Clostridia* may lack functions related to stress defense and/or exhibit more auxotrophies, which makes their isolation challenging, as suggested by Nayfach et al. (2019).

Genus- and family-level taxonomic assignments are conventionally determined using protein sequence comparisons rather than nucleotide-based metrics (Parks et al., 2018; Zheng et al., 2020). Consistent with this standard framework, AAI and POCP were employed to delineate genera containing at least one RS, applying the widely accepted cutoffs of 65% and 50%, respectively (Konstantinidis et al., 2017; Qin et al., 2014). The use of these established protein-based metrics and validated thresholds minimizes methodological bias and provides an objective basis for assessing existing taxonomic assignments, thereby identifying unresolved issues and inconsistencies in clostridial classification (Riesco and Trujillo, 2024). Of the 77 GTDB genera with multiple species, only 50 met both the minimum AAI and POCP thresholds. Meanwhile, 18 genera failed either AAI or POCP threshold, and 8 failed both (Table 1). Some putative genera could also emerge from merging species with conflicting genus assignments or by refining current genus boundaries. These observations highlight that, despite substantial improvements have already been made in the phylogenomic taxonomy of *Clostridia* (Candeliere et al., 2023; Bello et al., 2024), the delineation of some clostridial genera still needs to be resolved. Otherwise, it should not be excluded that the threshold metrics could be revised to meet the evidence generated from the reconstruction of phylogenomic relationships.

In addition to enabling precise taxonomic and phylogenetic analysis of *Clostridia*, RS genomes and reconstructed MAGs

from global-scale genome-resolved metagenomics enabled the prediction of functional capacities in terms of butyrate and propionate production. Butyrate pathways were predicted in 519 of the 1,897 *Clostridia* species (i.e. 27.3%), that were widespread across all major orders, whereas propionate production was much less common compared to butyrate (257/1,897; 13.5%). Examination of the phylogenetic tree revealed that butyrate and propionate production were concentrated within specific clades that did not always coincide with established genera or families. Putative butyrate-producing *Clostridia* were both more diverse (number of species) and more abundant than propionate producers, highlighting their major role within the intestinal ecosystem (Kircher et al., 2022). In terms of abundance within the gut microbiota of 151 healthy subjects, interindividual variability was substantial for the clostridial content, reflecting on the capability of producing butyrate and propionate.

Among the 50 most abundant *Clostridia* species, most were putative butyrate producers. Moreover, 76 out of 519 species accounted for >80% of total butyrate producers, indicating that butyrate production is a function largely driven by a limited number of dominant taxa, i.e. the genera *Acetatifactor*, *Agathobacter*, *Faecalibacterium*, *Gemmiger*, and *Roseburia*, well-known contributors to colonic butyrate pools and often associated with a healthy gut environment (Singh et al., 2023; Zmora et al., 2019; Mann et al., 2024). On the other hand, within the 50 most abundant taxa, only *Blautia A faecis*, *B. A wexlerae*, *B. A massiliensis*, *Mediterraneibacter faecis*, *Roseburia inulinivorans*, and *Ruminococcus B gnavus* were putative propionate producers, with *R. inulinivorans* also harboring the capacity to produce butyrate. This limited overlap between butyrate and propionate metabolism suggests a clear functional specialization among abundant *Clostridia* species and underscores a marked asymmetry in SCFA metabolism among *Clostridia*, with butyrate production being taxonomically concentrated in this class and functionally dominant, and propionate production distributed across fewer, mostly *Lachnospiraceae*-affiliated, species.

RS and NS similarly contributed to butyrate production within the microbiota herein analyzed, while propionate producers were less evenly distributed. These findings further support the efforts aiming to expand the culturome and to pursue the isolation and characterization of NS. In this context, ER4 sp000765235, the most abundant and prevalent NS, predicted to produce butyrate, was recently isolated and characterized taxonomically as *Hominicoprocola fusiformis* (Afrizal et al., 2022).

The pathways leading to butyrate and propionate production were analyzed in their entirety, encompassing all the clostridial RS and NS, providing for the first time a comprehensive view on the comparative genomics of these pathways in *Clostridia*. The vast majority of putative butyrate producers harbored a complete pathway originating from acetyl-CoA, indicating a direct connection with central carbon metabolism and glycolysis, whereas butyrate production via the glutarate, succinate semialdehyde, and lysine pathways, funneling carbon atoms derived from amino acid metabolism, was much less frequent. More than

a hundred species harboring complete *but* and *buk* branches lacked any pathway providing them with crotonyl-CoA. It should be taken into account that gaps in genome sequencing and assembly, that especially affect NS, may have led to an underestimation of pathway completeness. Another possible cause of the apparent underrepresentation of complete pathways may lie in the GapSeq tool, which prioritizes UniProt-reviewed proteins as queries for homology searches. These reference proteins may originate from bacteria phylogenetically distant from the *Clostridia* class and, consequently, may be less effective in detecting clostridial homologs, if present. For instance, it is unlikely that the glutarate pathway, comprising four metabolic steps, would be present in 78 species, all of them lacking only 2-oxoglutarate reductase. This indicates that the enzyme is not being detected or discarded possibly due to limitations in the query sequence.

Propionate producers were concentrated within specific clades, not necessarily coinciding with homogeneous taxonomic entities (except for *Blautia* A, *Eisenbergiella*, *Clostridium* E, and a few other genera). In agreement with previous studies, the propanediol pathway was the most common, followed by the acrylate one, while the succinate pathway, utilized by *Bacteroides*, the main intestinal propionate producers, was never found to be complete (Reichardt et al., 2014). Butyrate and propionate pathways coexisted in a minority of the species (77). Among them, some species of *Lachnospiraceae* (*C. catus* and *R. inulinivorans*) switch from butyrate to propionate production on different substrates (Reichardt et al., 2014).

Despite the lack of acrylate and propanediol pathways, the widespread occurrence of phosphate propionyltransferase and propionate kinase suggests that these enzymes might function in additional pathways, such as the degradation of several amino acids that yield 2-oxobutanoate and then propionyl-CoA and/or propionyl-phosphate. This observation deserves further investigation.

Enzymes involved in acetate and lactate metabolism were identified in numerous species (1,465 and 1,138, respectively), showing substantial overlap with each other and with those predicted to convert acetyl-CoA into butyrate. With regard to the utilization of acetate and lactate in butyrate production, an important cross-feeding occurring in the intestine, due to the reversible nature of many of these reactions (the direction depending on the energy and the redox status of the cell), it remains to be clarified whether such a great number of species is actually capable of transforming acetate and/or lactate into butyrate or whether acetate, lactate and butyrate are just alternative fermentation end products, toward which the carbon flow diverges at the level of pyruvate.

The widespread presence of both butyrate and propionate pathways across phylogenetically distant orders, as well as their relatively homogeneous occurrence within specific clades, suggests that it is worth considering how these fermentation pathways evolved and were inherited throughout the evolutionary history of the class, ultimately giving rise to their present polyphyletic distribution.

The outcomes of this study can help the scientific community better understand *Clostridia*, key players in all anoxic ecosystems,

including the gut, by deepening awareness of their vast diversity and providing new insights into their phylogenetic relationships. Moreover, the knowledge of taxa involved in butyrate and propionate biosynthesis is expanded, highlighting the role of yet uncharacterized species.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

LS: Formal analysis, Investigation, Writing – original draft. FC: Formal analysis, Investigation, Visualization, Writing – original draft. EB: Formal analysis, Writing – original draft. SR: Conceptualization, Visualization, Writing – original draft. AA: Conceptualization, Visualization, Writing – original draft. MR: Supervision, Writing – original draft.

Funding

The author(s) declared that financial support was received for this work and/or its publication. The research activities have been funded by Piano Nazionale Di Ripresa E Resilienza (Pnrr) - Missione 4 Componente 2, “Dalla ricerca all’impresa” Investimento 1.3, Creazione di “Partenariati estesi alle università, ai centri di ricerca, alle aziende per il finanziamento di progetti di ricerca di base”, finanziato dall’Unione europea - NextGenerationEU” - Progetto identificato con codice PE00000019, Titolo “HEAL ITALIA” - Spoke 5 - CUP E93C22001860006 Avviso MUR DD. 341 del 15.03.2022.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2026.1761627/full#supplementary-material>

References

- Afrizal, A., Hitch, T. C. A., Viehof, A., Treichel, N., Riedel, T., Abt, B., et al. (2022). Anaerobic single-cell dispensing facilitates the cultivation of human gut bacteria. *Environ. Microbiol.* 24, 3861–3881. doi: 10.1111/1462-2920.15935
- Almeida, A., Nayfach, S., Boland, M., Strozzi, F., Beracochea, M., Shi, Z. J., et al. (2021). A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* 39, 105–114. doi: 10.1038/s41587-020-0603-3
- Amaretti, A., Gozzoli, C., Simone, M., Raimondi, S., Righini, L., Pérez-Brocal, V., et al. (2019). Profiling of protein degraders in cultures of human gut microbiota. *Front. Microbiol.* 10:2614. doi: 10.3389/fmicb.2019.02614
- Arora, T., Sharma, R., and Frost, G. (2011). Propionate. anti-obesity and satiety enhancing factor?. *Appetite* 56, 511–515. doi: 10.1016/j.appet.2011.01.016
- Bach Knudsen, K. E., Lærke, H. N., Hedemann, M. S., Nielsen, T. S., Ingerslev, A. K., Gundelund Nielsen, D. S., et al. (2018). Impact of diet-modulated butyrate production on intestinal barrier function and inflammation. *Nutrients* 10:1499. doi: 10.3390/nu10101499
- Barker H. A. (1981). Amino acid degradation by anaerobic bacteria. *Annu. Rev. Biochem.* 50, 23–40. doi: 10.1146/annurev.bi.50.070181.000323
- Bello, S., McQuay, S., Rudra, B., and Gupta, R. S. (2024). Robust demarcation of the family Peptostreptococcaceae and its main genera based on phylogenomic studies and taxon-specific molecular markers. *Int. J. Syst. Evol. Microbiol.* 74:10.1099/ijsem.0.006247. doi: 10.1099/ijsem.0.006247
- Buckel, W. (2001). Unusual enzymes involved in five pathways of glutamate fermentation. *Appl. Microbiol. Biotechnol.* 57, 263–273. doi: 10.1007/s002530100773
- Bui, T. P., Ritari, J., Boeren, S., de Waard, P., Plugge, C. M., and de Vos, W. M. (2015). Production of butyrate from lysine and the amadori product fructosyllysine by a human gut commensal. *Nat. Commun.* 6:10062. doi: 10.1038/ncomms10062
- Bui, T. P. N., Troise, A. D., Nijse, B., Roviello, G. N., Fogliano, V., and de Vos, W. M. (2020). Intestinimonas-like bacteria are important butyrate producers that utilize Ne-fructosyllysine and lysine in formula-fed infants and adults. *J. Funct. Foods* 70:103974. doi: 10.1016/j.jff.2020.103974
- Candeliere, F., Musmeci, E., Amaretti, A., Sola, L., Raimondi, S., and Rossi, M. (2023). Profiling of the intestinal community of *Clostridia*: taxonomy and evolutionary analysis. *Microbiome Res. Rep.* 2:13. doi: 10.20517/mrr.2022.19
- Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2022). GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* 38, 5315–5316. doi: 10.1093/bioinformatics/btac672
- Collins, M. D., Lawson, P. A., Willems, A., Cordoba, J. J., Fernandez-Garayzabal, J., Garcia, P., et al. (1994). The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int. J. Syst. Bacteriol.* 44, 812–826. doi: 10.1099/00207713-44-4-812
- Duncan, S. H., Holtrop, G., Lobley, G. E., Calder, A. G., Stewart, C. S., and Flint, H. J. (2004). Contribution of acetate to butyrate formation by human faecal bacteria. *Br. J. Nutr.* 91, 915–923. doi: 10.1079/BJN20041150
- Fatahi-Bafghi, M. (2024). Reclassification of butyriovibrio crossotus moore et al. 1976 (Approved Lists 1980) into a novel genus as eshraghiella crossota gen. nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* 74:10.1099/ijsem.0.006509. doi: 10.1099/ijsem.0.006509
- Fernández-Veledo, S., and Vendrell, J. (2019). Gut microbiota-derived succinate: friend or foe in human metabolic diseases?. *Rev. Endocr. Metab. Disord.* 20, 439–447. doi: 10.1007/s11154-019-09513-z
- Gregory, A. L., Pensinger, D. A., and Hryckowian, A. J. (2021). A short chain fatty acid-centric view of clostridioides difficile pathogenesis. *PLoS Pathog.* 17:e1009959. doi: 10.1371/journal.ppat.1009959
- Haas, K. N., and Blanchard, J. L. (2020). Reclassification of the clostridium clostridioforme and clostridium sphenoides clades as enterocloster gen. nov. and lacrimispora gen. nov. including reclassification of 15 taxa. *Int. J. Syst. Evol. Microbiol.* 70, 23–34. doi: 10.1099/ijsem.0.003698
- Han, K., Meadows, A. M., Rodman, M. J., Russo, A. C., Sharma, R., Singh, K., et al. (2024). Propionate functions as a feeding state-dependent regulatory metabolite to counter proinflammatory signaling linked to nutrient load and obesity. *J. Leukoc. Biol.* 115, 738–749. doi: 10.1093/jleuko/qiae006
- Hodgkinson, K., El Abbar, F., Dobranowski, P., Manoogian, J., Butcher, J., Figeys, D., et al. (2023). Butyrate's role in human health and the current progress toward its clinical application to treat gastrointestinal disease. *Clin. Nutr.* 42, 61–75. doi: 10.1016/j.clnu.2022.10.024
- Hosseini, E., Grootaert, C., Verstraete, W., and Van de Wiele, T. (2011). Propionate as a health-promoting microbial metabolite in the human gut. *Nutr. Rev.* 69, 245–258. doi: 10.1111/j.1753-4887.2011.00388.x
- Jacobson, A., Lam, L., Rajendram, M., Tamburini, F., Honeycutt, J., Pham, T., et al. (2018). A gut commensal-produced metabolite mediates colonization resistance to salmonella infection. *Cell Host Microbe*, 24, 296–307. doi: 10.1016/j.chom.2018.07.002
- Jensen, R. O., Schulz, F., Roux, S., Klingeman, D. M., Mitchell, W. P., et al. (2024). Phylogenomics and genetic analysis of solvent-producing *Clostridium* species. *Sci Data* 11:432. doi: 10.1038/s41597-024-03210-6
- Kilner, J., Waby, J. S., Chowdry, J., Khan, A. Q., Noirel, J., Wright, P. C., et al. (2012). A proteomic analysis of differential cellular responses to the short-chain fatty acids butyrate, valerate and propionate in colon epithelial cancer cells. *Mol. Biosyst.* 8, 1146–1156. doi: 10.1039/C1MB05219E
- Kim, D., Park, S., and Chun, J. (2021). Introducing EzAAI: a pipeline for high throughput calculations of prokaryotic average amino acid identity. *J. Microbiol.* 59, 476–480. doi: 10.1007/s12275-021-1154-0
- Kircher, B., Woltemate, S., Gutzki, F., Schlüter, D., Geffers, R., Bähre, H., et al. (2022). Predicting butyrate- and propionate-forming bacteria of gut microbiota from sequencing data. *Gut microbes* 14:2149019. doi: 10.1080/19490976.2022.2149019
- Konstantinidis, K. T., Rosselló-Móra, R., and Amann, R. (2017). Uncultivated microbes in need of their own taxonomy. *ISME J.* 11, 2399–2406. doi: 10.1038/ismej.2017.113
- Lamas, A., Regal, P., Vázquez, B., Cepeda, A., and Franco, C. M. (2019). Short chain fatty acids commonly produced by gut microbiota influence *Salmonella enterica* motility, biofilm formation, and gene expression. *Antibiotics* 8:265. doi: 10.3390/antibiotics8040265
- Lawson, P. A., Saavedra Perez, L., and Sankaranarayanan, K. (2023). Reclassification of clostridium cocleatum, clostridium ramosum, clostridium spiroforme and clostridium saccharogumia as thomasclavelia cocleata gen. nov., comb. nov., thomasclavelia ramosa comb. nov., gen. nov., thomasclavelia spiroformis comb. nov. and thomasclavelia saccharogumia comb. nov. *Int. J. Syst. Evol. Microbiol.* 73:10.1099/ijsem.0.005694. doi: 10.1099/ijsem.0.005694
- Lechner, M., Findeiss, S., Steiner, L., Marz, M., Stadler, P. F., and Prohaska, S. J. (2011). Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC bioinformatics* 12:124. doi: 10.1186/1471-2105-12-124
- Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* 49, W293–W296. doi: 10.1093/nar/gkab301
- Litty, D., and Müller, V. (2021). Butyrate production in the acetogen *Eubacterium limosum* is dependent on the carbon and energy source. *Microb. Biotechnol.* 14, 2686–2692. doi: 10.1111/1751-7915.13779

- Louis, P., Duncan, S. H., McCrae, S. I., Millar, J., Jackson, M. S., and Flint, H. J. (2004). Restricted distribution of the butyrate kinase pathway among butyrate-producing bacteria from the human colon. *J. Bacteriol.* 186, 2099–2106. doi: 10.1128/JB.186.7.2099-2106.2004
- Louis, P., and Flint, H. J. (2009). Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol. Lett.* 294, 1–8. doi: 10.1111/j.1574-6968.2009.01514.x
- Louis, P., and Flint, H. J. (2017). Formation of propionate and butyrate by the human colonic microbiota. *Environ. Microbiol.* 19, 29–41. doi: 10.1111/1462-2920.13589
- Louis, P., Young, P., Holtrop, G., and Flint, H. J. (2010). Diversity of human colonic butyrate-producing bacteria revealed by analysis of the butyryl-CoA:acetate CoA-transferase gene. *Environ. Microbiol.* 12, 304–314. doi: 10.1111/j.1462-2920.2009.02066.x
- Lu, J., Breitwieser, F. P., Thielen, P., and Salzberg, S. L. (2017). Bracken: estimating species abundance in metagenomics data. *PeerJ. Comput. Sci.* 3:e104. doi: 10.7717/peerj-cs.104
- Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D. A., Bauer, E., Noronha, A., et al. (2017). Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.* 35, 81–89. doi: 10.1038/nbt.3703
- Mann, E. R., Lam, Y. K., and Uhlig, H. H. (2024). Short-chain fatty acids: linking diet, the microbiome and immunity. *Nat. Rev. Immunol.* 24, 577–595. doi: 10.1038/s41577-024-01014-8
- Nayfach, S., Shi, Z. J., Seshadri, R., Pollard, K. S., and Kyrpides, N. C. (2019). New insights from uncultivated genomes of the global human gut microbiome. *Nature* 568, 505–510. doi: 10.1038/s41586-019-1058-x
- Parks, D. H., Chuvochina, M., Rinke, C., Mussig, A. J., Chaumeil, P. A., and Hugenholtz, P. (2022). GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res.* 50, D785–D794. doi: 10.1093/nar/gkab776
- Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* 36, 996–1004. doi: 10.1038/nbt.4229
- Pryde, S. E., Duncan, S. H., Hold, G. L., Stewart, C. S., and Flint, H. J. (2002). The microbiology of butyrate formation in the human colon. *FEMS Microbiol. Lett.* 217, 133–139. doi: 10.1111/j.1574-6968.2002.tb11467.x
- Qin, Q. L., Xie, B. B., Zhang, X. Y., Chen, X. L., Zhou, B. C., Zhou, J., et al. (2014). A proposed genus boundary for the prokaryotes based on genomic insights. *J. Bacteriol.* 196, 2210–2215. doi: 10.1128/JB.01688-14
- Raimondi, S., Calvini, R., Candeliere, F., Leonardi, A., Ulrici, A., Rossi, M., et al. (2021). Multivariate analysis in microbiome description: correlation of human gut protein degraders, metabolites, and predicted metabolic functions. *Front. Microbiol.* 12:723479. doi: 10.3389/fmicb.2021.723479
- Reichardt, N., Duncan, S. H., Young, P., Belenguer, A., McWilliam Leitch, C., Scott, K. P., et al. (2014). Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J.* 8, 1323–1335. doi: 10.1038/ismej.2014.14
- Riedel, T., Wetzel, D., Hofmann, J. D., Plorin, S. P. E. O., Dannheim, H., Berges, M., et al. (2017). High metabolic versatility of different toxigenic and non-toxigenic *Clostridioides difficile* isolates. *Int. J. Med. Microbiol. IJMM* 307, 311–320. doi: 10.1016/j.ijmm.2017.05.007
- Riesco, R., and Trujillo, M. E. (2024). Update on the proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 74:006300. doi: 10.1099/ijsem.0.006300
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Shin, J., Song, Y., Jeong, Y., and Cho, B. K. (2016). Analysis of the Core Genome and Pan-Genome of Autotrophic Acetogenic Bacteria. *Front. Microbiol.* 7:1531. doi: 10.3389/fmicb.2016.01531
- Singh, V., Lee, G., Son, H., Koh, H., Kim, E. S., Unno, T., et al. (2023). Butyrate producers, “the sentinel of gut”: their intestinal significance with and beyond butyrate, and prospective use as microbial therapeutics. *Front. Microbiol.* 13:1103836. doi: 10.3389/fmicb.2022.1103836
- Sivaprakasam, S., Prasad, P. D., and Singh, N. (2016). Benefits of short-chain fatty acids and their receptors in inflammation and carcinogenesis. *Pharmacol. Ther.* 164, 144–151. doi: 10.1016/j.pharmthera.2016.04.007
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Sun, S., Ding, Y., Liu, M., Xian, M., and Zhao, G. (2020). Comparison of glucose, acetate and ethanol as carbon resource for production of poly(3-Hydroxybutyrate) and other acetyl-CoA derivatives. *Front. Bioeng. Biotechnol.* 8:833. doi: 10.3389/fbioe.2020.00833
- Trischler, R., Roth, J., Sorbara, M. T., Schlegel, X., and Müller, V. (2022). A functional Wood-Ljungdahl pathway devoid of a formate dehydrogenase in the gut acetogens *Blautia wexlerae*, *Blautia luti* and beyond. *Environ. Microbiol.* 24, 3111–3123. doi: 10.1111/1462-2920.16029
- Vinolo, M. A., Rodrigues, H. G., Nachbar, R. T., and Curi, R. (2011). Regulation of inflammation by short chain fatty acids. *Nutrients* 3, 858–876. doi: 10.3390/nu3100858
- Vital, M., Howe, A. C., and Tiedje, J. M. (2014). Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. *mBio* 5:e00889. doi: 10.1128/mBio.00889-14
- Wood, D. E., Lu, J., and Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biol.* 20:257. doi: 10.1186/s13059-019-1891-0
- Zheng, J., Wittouck, S., Salvetti, E., Franz, C. M. A. P., Harris, H. M. B., Mattarelli, P., et al. (2020). A taxonomic note on the genus *Lactobacillus*: description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int. J. Syst. Evol. Microbiol.* 70, 2782–2858. doi: 10.1099/ijsem.0.004107
- Zimmermann, J., Kaleta, C., and Waschina, S. (2021). gapseq: informed prediction of bacterial metabolic pathways and reconstruction of accurate metabolic models. *Genome Biol.* 22:81. doi: 10.1186/s13059-021-02295-1
- Zmora, N., Suez, J., and Elinav, E. (2019). You are what you eat: diet, health and the gut microbiota. *Nat. Rev. Gastroenterol. Hepatol.* 16, 35–56. doi: 10.1038/s41575-018-0061-2
- Zou, Y., Xue, W., Luo, G., Deng, Z., Qin, P., Guo, R., et al. (2019). 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat. Biotechnol.* 37, 179–185. doi: 10.1038/s41587-018-0008-8