

Distribution, organization and expression of genes concerned with anaerobic lactate utilization in human intestinal bacteria

Paul O. Sheridan¹, Petra Louis¹, Eleni Tsompanidou², Sophie Shaw³, Hermie J. Harmsen², Sylvia H. Duncan¹, Harry J. Flint¹† and Alan W. Walker¹*,†

Abstract

Lactate accumulation in the human gut is linked to a range of deleterious health impacts. However, lactate is consumed and converted to the beneficial short-chain fatty acids butyrate and propionate by indigenous lactate-utilizing bacteria. To better understand the underlying genetic basis for lactate utilization, transcriptomic analyses were performed for two prominent lactate-utilizing species from the human gut, *Anaerobutyricum soehngenii* and *Coprococcus catus*, during growth on lactate, hexose sugar or hexose plus lactate. In *A. soehngenii* L2-7 six genes of the lactate utilization (*lct*) cluster, including NAD-independent D-lactate dehydrogenase (D-iLDH), were co-ordinately upregulated during growth on equimolar D- and L-lactate (DL-lactate). Upregulated genes included an acyl-CoA dehydrogenase related to butyryl-CoA dehydrogenase, which may play a role in transferring reducing equivalents between reduction of crotonyl-CoA and oxidation of lactate. Genes upregulated in *C. catus* GD/7 included a six-gene cluster (*lap*) encoding propionyl CoA-transferase, a putative lactoyl-CoA epimerase, lactoyl-CoA dehydratase and lactate permease, and two unlinked acyl-CoA dehydrogenase genes that are candidates for acryloyl-CoA reductase. A D-iLDH homologue in *C. catus* is encoded by a separate, partial *lct*, gene cluster, but not upregulated on lactate. While *C. catus* converts three mols of DL-lactate via the acrylate pathway to two mols propionate and one mol acetate, some of the acetate can be re-used with additional lactate to produce butyrate. A key regulatory difference is that while glucose partially repressed *lct* cluster expression in *A. soehngenii*, there was no repression of lactate-utilization genes by fructose in the non-glucose utilizer *C. catus*. This suggests that these species could occupy different ecological niches for lactate utilization in the gut, which may be important factors to consider when developing lactate-utilizing bacteria as novel candidate probiotics.

DATA SUMMARY

Novel draft genomes generated for this study have been made available from GenBank (<https://www.ncbi.nlm.nih.gov/bioproject/>) under BioProject number PRJNA701799. RNA-seq data have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-10136. Further details of additional existing genomic data that were analysed in this project are given in Tables 1 and S2 (available in the online version of this article), and at <https://github.com/SheridanPO/Lactate-utilizing-bacteria>. Supplementary Material can be found in Figshare: [https://doi.org/10.6084/m9.figshare.16510344.v1\[1\]](https://doi.org/10.6084/m9.figshare.16510344.v1[1]).

INTRODUCTION

The human large intestinal microbiota is dominated by obligately anaerobic bacteria, whose growth is largely dependent on the supply of complex carbohydrates and proteins available for fermentation. These substrates are mostly fermented by resident gut

Received 20 April 2021; Accepted 02 November 2021; Published 25 January 2022

Author affiliations: ¹Gut Health Group, Rowett Institute, University of Aberdeen, Foresterhill, AB25 2ZD Aberdeen, UK; ²Medical Microbiology and Infection Prevention, University of Groningen, University Medical Center Groningen, Hanzplein 1, 9713 GZ Groningen, The Netherlands; ³Centre for Genome-Enabled Biology and Medicine, 23 St. Machar Drive, AB24 3RY Aberdeen, UK.

*Correspondence: Alan W. Walker, alan.walker@abdn.ac.uk

Keywords: Human gut microbiota; lactate-utilizing bacteria; anaerobic metabolism; upregulation by lactate; transcriptomics.

Repositories: Novel draft genomes that were generated for this study are available from GenBank under BioProject PRJNA701799. RNA-seq data have been deposited in the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-10136.

†These authors contributed equally to this work

Data statement: All supporting data, code and protocols have been provided within the article or through supplementary data files. Eighteen supplementary tables and one supplementary figure are available with the online version of this article.

000739 © 2022 The Authors



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Impact Statement

Lactate can be produced as a fermentation by-product by many gut bacteria but has the potential to perturb intestinal microbial communities by lowering luminal pH, and its accumulation has been linked to a range of deleterious health outcomes. Fortunately, in healthy individuals, lactate tends not to accumulate as it is consumed by cross-feeding lactate-utilizing bacteria, which can convert it into the beneficial short-chain fatty acids butyrate and propionate. Lactate-utilizing gut bacteria are therefore promising candidates for potential development as novel probiotics. However, lactate utilizers are taxonomically diverse, and the genetic repertoire that underpins the utilization of lactate by these specialized gut bacteria is not fully understood. In this study, we used transcriptomics to compare gene-expression profiles of *Anaerobutyricum soehngenii* and *Coprococcus catus*, two prominent lactate-utilizing species in the human gut, during growth on lactate alone, sugar alone or sugar plus lactate. The results revealed strong upregulation of key, but distinct, gene clusters that appear to be responsible for lactate utilization by these, and other, gut bacterial species. Our results therefore increase mechanistic comprehension of different lactate utilization pathways used by gut bacteria, which may help to inform selection of optimal lactate-utilizing species for development as novel therapeutics against colonic microbiota perturbations.

anaerobes to short-chain fatty acids (SCFAs) and gases, and in the healthy colon the major faecal SCFAs detected are acetate, propionate and butyrate [2]. Lactate is another fermentation product of many anaerobic bacteria that colonize the mammalian gut, sometimes as a major metabolite, as in *Lactobacillus* and *Bifidobacterium* spp. [3]. However, as the pKa (around 3.8) for lactic acid is lower than that of other fermentation acids such as acetate (4.76), accumulation of lactate has the potential to dramatically change the gut environment and gut microbiota composition via reduction in prevailing pH [4].

Lactate may therefore help to inhibit the growth of some pathogenic bacteria with poor tolerance for lower pH [5] and a gut microbiota dominated by lactate-producing bacteria is found in many healthy pre-weaned infants [6]. On the other hand, there are many known deleterious impacts of lactate accumulation in the gut. Indeed, the phenomenon of lactic acidosis, driven by excessive production of lactate following fermentation of easily digestible dietary carbohydrates, is well known in ruminants and is a major problem in animal husbandry [7, 8]. Similarly, in humans, accumulation of toxic D-lactate is life-threatening in short bowel syndrome [9] and lactate accumulation is also associated with severe colitis [10]. Additionally, lactate can be utilized as a carbon and energy source by certain intestinal pathogens such as *Salmonella* [11] and *Campylobacter* [12], compounding the potential health detriments of lactate accumulation in the gut.

However, in the colon of healthy adult humans, as well as in the non-acidotic rumen, lactate does not accumulate, due the activities of cross-feeding, lactate-utilizing bacteria [13]. Colonic lactate concentrations are therefore a balance between bacterial production during fermentation and utilization by lactate-utilizing bacteria, which can use lactate to form acetate, propionate or butyrate [14]. Recent research has linked the populations and activities of these lactate-utilizing bacteria with overall fermentation patterns and productivity [15], and a low abundance of commensal lactate utilizers has also been proposed to make gut microbial communities inherently less stable and more prone to lactate-induced perturbations [16]. As such, lactate-utilizing bacteria have been proposed as promising candidates for development as novel probiotics [16–19].

Table 1. New and existing strains used for *in vitro* cultivation studies in this work, and their corresponding genomes. 'N/A' indicates no genome available for this strain, which was used for *in vitro* work only. The full list of strains used for genomic-based analyses is shown in Table S2

Species	Strain	Donor	Genome accession	Contigs	CDS no.	Genome reference	Strain reference
<i>Anaerobutyricum soehngenii</i>	L2-7	1	GCA_900209925	1	3005	[71]	[72]
	SL6/1/1	2	JAFIQP000000000	498	2900	This study	[73]
	HTF-83D	3	JAFIQO000000000	262	2809	This study	This study
<i>Anaerobutyricum hallii</i>	SM6/1	2	JAFIQQ000000000	341	3202	This study	[73]
	DSM3353	4	GCA_900209925	175	2829	Unpublished	[74]
<i>Anaerostipes hadrus</i>	HTF-920	7	N/A	N/A	N/A	N/A	This study
	SS2/1	2	GCA_000154545	34	3022	Unpublished	[73]
	SSC/2	2	GCA_000210695	120	3170	Unpublished	[73]
<i>Anaerostipes caccae</i>	DSM14662	8	GCA_000154305	26	3460	Unpublished	[72]
<i>Coprococcus catus</i>	GD/7	9	GCA_000210555	1	3261	Unpublished	[73]

The ability to utilize lactate as an energy source for growth appears to be limited to relatively few bacterial species among the human intestinal microbiota, although these species are taxonomically diverse and can utilize lactate in different ways. Selective isolation on DL lactate-containing media resulted in the recovery of *Lachnospiraceae* species subsequently identified as *Eubacterium hallii* (since reclassified as the two species *Anaerobutyricum hallii* and *Anaerobutyricum soehngeni* [20]), *Anaerostipes hadrus* and *Anaerostipes caccae* [21, 22]. These isolates produce butyrate from lactate, with net consumption of acetate, suggesting that they initially convert lactate into pyruvate, which is then routed into the butyrate pathway [22]. The lactate to pyruvate conversion is energetically unfavourable and recent evidence in *Acetobacterium woodii* indicates that it is dependent on electron confurcation [23]. Alternative routes for lactate utilization that are known to be important in the rumen [24] result in propionate formation. These include the acrylate pathway found in the *Negativicutes* species *Megasphaera elsdenii* and in at least one member of the *Lachnospiraceae*, *Coprococcus catus* [25], and the succinate pathway that is found among the *Negativicutes* in *Selenomonas ruminantium* and *Veillonella* spp. [25]. Less advantageously, lactate is also a co-substrate for sulfate-reducing bacteria that use it to form acetate and sulfide, the latter of which may be genotoxic [26].

Here we examine the distribution, organization and regulation of genes involved in lactate utilization among dominant representatives of the human intestinal microbiota, based on new isolates of lactate-utilizing species and on newly available genome sequences. In particular, we identify two gene clusters whose transcription is upregulated during growth with lactate as carbon and energy source. One of these corresponds to the predicted lactate utilization locus, the *lct* gene cluster, of *A. woodii* [23] that was recently identified from proteomic analysis in *A. soehngeni* [27], while the second encodes activities involved in the acrylate pathway of *C. catus*. This investigation reveals differences in the regulation, phylogenetic distribution and metabolic function of these two clusters, and provides novel insights into the mechanistic basis of different lactate-utilization strategies used by human gut bacteria.

METHODS

Isolation of new strains of lactate-utilizing bacteria

New strains of *A. soehngeni* (HTF-83D) and *A. hadrus* (HTF-920, HTF-146, HTF-370 and HTF-412) were isolated from human faecal samples by dilution and culturing of single colonies on the anaerobic medium YCFA [28] supplemented with glucose. Taxonomic identification of strains was initially carried out by 16S rRNA gene sequencing and BLASTn querying each sequence against the NCBI 16S rRNA gene reference database. The new strains that were putatively assigned to the *Anaerobutyricum* genus based on 16S rRNA gene sequences were definitively classified as either *A. hallii* or *A. soehngeni* by comparing the average nucleotide identity (ANI) and average amino acid identity (AAI) of the new isolate draft genomes with genome sequences from the type strains of *A. hallii* (DSM3353^T) and *A. soehngeni* (L2-7^T) using nucmer [29] for ANI and CompareM (<https://github.com/dparks1134/CompareM>) for AAI (Table S1).

Bacterial strains, growth conditions and genomes

The bacterial strains and genomes used in this work, including both new and previously isolated strains, are described in Tables 1 and S2. Routine culturing of bacterial strains was in anaerobic M2GSC medium [30] in 7.5 ml aliquots in Hungate tubes, sealed with butyl rubber septa (Bellco Glass). Growth experiments were carried out in basal YCFA medium [28] supplemented with 35 mM lactate and/or 11 mM glucose (or fructose for *C. catus* since this species is incapable of growing on glucose as sole carbon source). All cultures were incubated anaerobically without agitation at 37 °C, using the anaerobic methods described previously by Bryant [31].

DNA isolation, genome sequencing and annotation

Genomic DNA was extracted with the Ultraclean Microbial DNA isolation kit (MoBio Laboratories, Carlsbad, CA, USA). The DNA concentration and purity were measured using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and the Qubit double-stranded DNA (dsDNA) HS and BR assay kits (Life Technologies, Carlsbad, CA, USA). One nanogram of bacterial DNA was used for library preparation. The DNA library was prepared using the Nextera XT library preparation kit with the Nextera XT v2 index kit (Illumina, San Diego, CA, USA). The library fragment length was aimed at fragments with a median size of 575 bases and was assessed with the Genomic DNA ScreenTape assay with the 2200 Tape-Station system (Agilent Technologies, Waldbronn, Germany). Subsequently, the library was sequenced on an Illumina MiSeq sequencer, using a 2×250 (500v2) cartridge, with the MiSeq reagent kit v2 generating 250 bp paired-end reads, (Illumina, San Diego, CA, USA). MiSeq data were processed with MiSeq control software v2.4.0.4 and MiSeq Reporter v2.4 (Illumina, San Diego, CA, USA). The resulting fastq files were filtered using TrimGalore v0.4.0 (<https://github.com/FelixKrueger/TrimGalore>), removing one nucleotide off the 3' end (--trim1) and removing both pairs of the paired end reads if one read did not pass filtering (--paired). Filtered fastq files were assembled using the IDBA_UD assembler v1.1.3 [32] using standard settings. Newly generated genome sequences are available from GenBank (BioProject PRJNA701799), with individual accession numbers shown in Tables 1 and S2. Additional sequences were retrieved from GenBank (accession numbers in Table S2). New genome sequences were annotated using Prokka

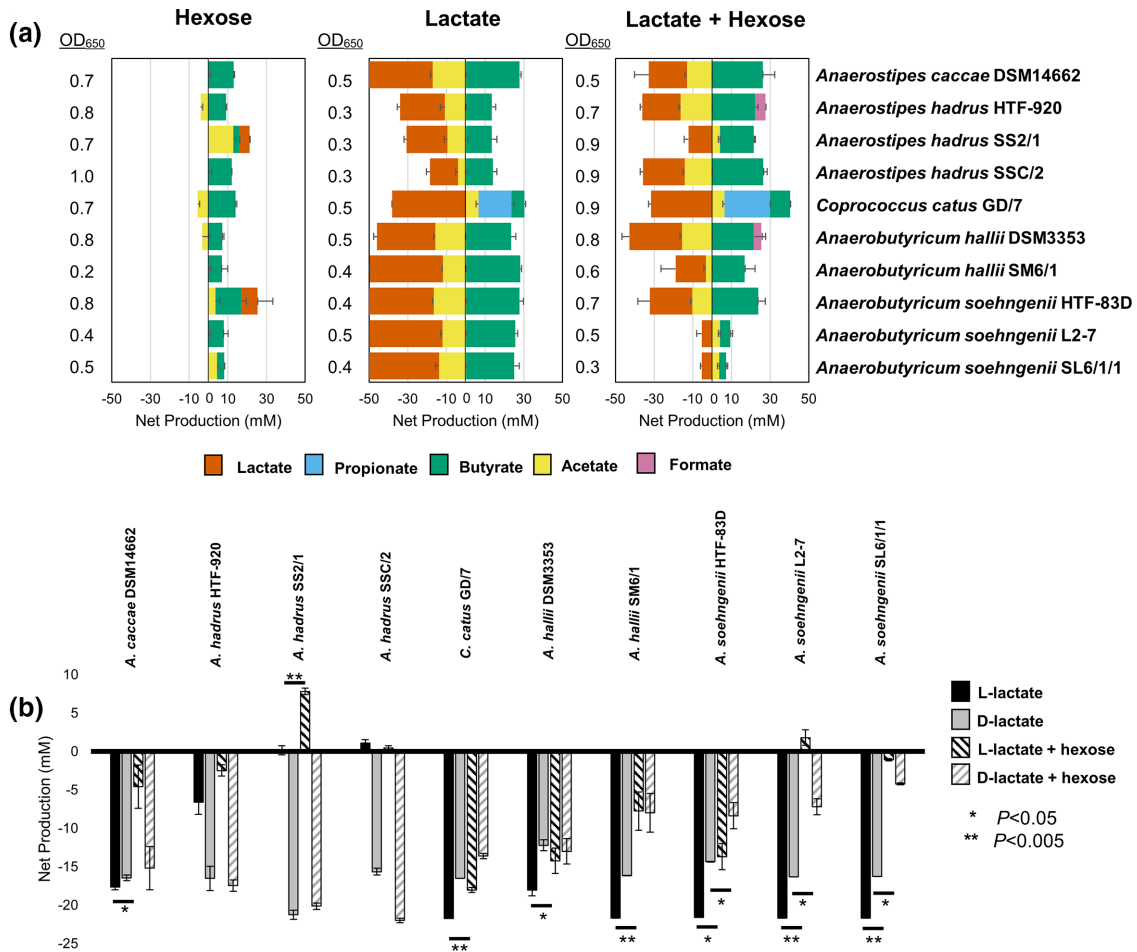


Fig. 1. Fermentation profiles of lactate-utilizing bacteria grown on sugar, on lactate, or a combination of both. (a) Net production (to the right of the figures), or consumption (to the left of the figures) of various short-chain fatty acids by lactate-utilizing bacteria grown on either hexose sugar alone, lactate alone or a combination of both. Cultures were incubated anaerobically for 24 h in media supplemented with 11 mM glucose, 35 mM lactate or both. Glucose was substituted with 11 mM fructose for *C. catus* GD/7. Iso-butyrate, valerate, iso-valerate, and succinate were not detected in any of the samples. (b) Net consumption or production of L- or D-lactate by the tested strains, with or without simultaneous exposure to hexose sugars in the culture medium. Error bars are the SEM of three biological replicates. There was no remaining glucose or fructose in any culture after 24 h, with the exception of *A. soehngenii* L2-7 (2 mM, SEM 2.0) and SL6/1/1 (6 mM, SEM 3.2) grown with a combination of glucose and lactate. Expanded data are shown in Table S3.

v1.11 [33] and publicly available sequences were reannotated in the same manner to ensure uniformity between annotation methods. Genome completeness and contamination was estimated using CheckM v1.1.3 [34] (Table S2).

Functional annotation of lactate-utilization loci and phylogenetic analyses

Lactate-utilization loci were searched for using the hidden Markov model (HMM) profiles of conserved domains found in three-component L-lactate hydrogenase (PF02754, PF02589, PF13183, PF11870), FAD-dependent D-lactate dehydrogenase (PF01565, PF02913, PF09330, PF12838, PF02754), FMN-dependent L-lactate dehydrogenase (PF01070, PF00173), NAD-dependent D-lactate dehydrogenase (PF00389, PF02826), NAD-dependent L-lactate dehydrogenase (PF02866, PF00056), lactate permease (PF02652), lactate racemase (PF09861) and ETF (PF00766, PF01012) using hmmsearch in HMMER3 [35]. Butyrate and propionate producing pathways were detected by BLASTP against UniProt [36] and KEGG [37] with a threshold of 60% identity using the marker genes butyrate kinase and butyryl-CoA:acetate CoA-transferase as indicators of butyrate production, and propanediol utilization protein (PduP), lactoyl CoA dehydratase alpha-subunit (LcdA) and methyl-malonyl-CoA decarboxylase alpha-subunit (MmdA) as indicators of propionate production through the propanediol, acrylate and succinate pathways, respectively [25]. Rnf complexes were detected by querying against amino acid sequences of the experimentally proven Rnf complexes of *Clostridium ljungdahlii* and *A. woodii* [38, 39].

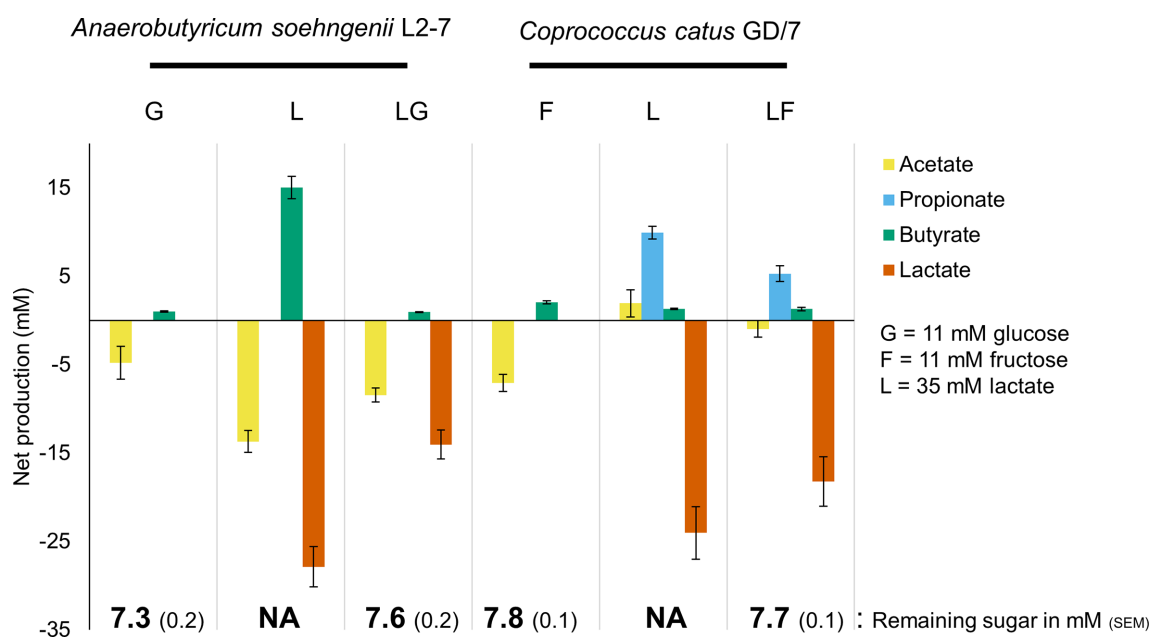


Fig. 2. Fermentation end products of *A. soehngenii* L2-7 and *C. catus* GD/7 during growth on lactate, a monosaccharide alone or on both in combination. Cultures were harvested for RNA extraction at an OD_{650} reading of 0.3. Iso-butyrate, valerate, iso-valerate and succinate were not detected in any of the samples. Error bars are the SEM of eight biological replicates.

All genes containing the lactate permease Pfam [40] domain PF02652 in UniProt (June 2020) [36] were downloaded and clustered into groups of >70% similarity using CD-HIT v4.8.1 [41]. Representatives of each cluster were aligned using MAFFT L-INS-I v7.407 [42]. The acyl-CoA dehydrogenase in the *lct* cluster (L2-7_01909) was used as a BLASTP query against *C. catus* GD/7 and the *A. hadrus* genomes, homologues were aligned using MAFFT L-INS-i [42] and a HMM profile was created from the alignment using hmmbuild in HMMER3 [35]. This profile was then queried against the genomes of known lactate utilizers using hmmsearch (-T 80) to detect divergent members of the acyl-CoA dehydrogenase protein superfamily. Matching sequences were aligned using MAFFT L-INS-i v7.407 [42] and the alignment was manually inspected for likely mis-annotations (none were found). For both the lactate permease and acyl-CoA dehydrogenase alignments, spurious sequences and poorly aligned regions were removed with trimal v1.4.1 (automated1, resoverlap 0.55 and seqoverlap 60) [43] and maximum-likelihood trees were constructed with IQ-TREE v1.6.11 [44] using the best fitting protein model predicted in ModelFinder (from IQ-TREE v1.6.11) [45]. Branch supports were computed using the SH-aLRT test [46] for the lactate permease tree and ultrafast bootstraps for the acyl-CoA dehydrogenase tree. Both trees were rooted using minimal ancestor deviation [47]. The tree figures were generated using iTOL [48].

Differential gene-expression analysis

Fresh YCFA broth supplemented with 11 mM monosaccharide (glucose for *A. soehngenii* and fructose for *C. catus*), 35 mM lactate or both 11 mM monosaccharide and 35 mM lactate were inoculated 1:75 with an overnight grown culture of either *A. soehngenii* L2-7 or *C. catus* GD/7. These cultures were grown to OD_{650} 0.3 (all cultures were in exponential growth phase at this OD_{650} reading) before cultures were centrifuged to cell pellets (5 min, 1200 g, 20 °C). The supernatant was transferred to a new tube for SCFA and monosaccharide analysis, and the cell pellet was resuspended in 500 μ l of RNeasy Lysis Buffer (Qiagen), left at room temperature for 5 min and frozen at -70 °C. Total RNA was extracted using the RNeasy PowerMicrobiome Kit (QIAGEN), following the manufacturer's instructions. This included a bead beating step for lysing Gram-positive cell walls. Ribosomal RNA was depleted using the Illumina Ribo-Zero rRNA removal kit. Libraries were prepared using the Illumina TruSeq stranded mRNA library kit and sequenced at the Centre for Genome-Enabled Biology and Medicine (CGEBM) at The University of Aberdeen using a High Output 1x75 kit on the Illumina NextSeq platform producing 75 bp single end reads. Reads were combined from two runs producing between 14237482 and 16008520 reads per sample. These transcriptomic datasets are available from the ArrayExpress database at EMBL-EBI under accession number E-MTAB-10136. Quality of raw reads was assessed using FastQC v0.11.3 [49] and sequences were filtered using TrimGalore v0.4.0. Between 14141677 and 15921037 sequences remained per sample, which, on average, was 99% of the raw reads. The appropriate reference genomes for *A. soehngenii* L2-7 and *C. catus* GD/7 (Table S2) were indexed in preparation for alignment using the HISAT2 build function (version 2.1.0) [50]. Filtered reads were then aligned to the appropriate reference genome using HISAT2 v2.1.0 [50] with default settings. Alignments were then converted to BAM format and sorted using SAMtools v1.2 [51]. Reads aligned to gene regions were counted using featureCounts (subread package version

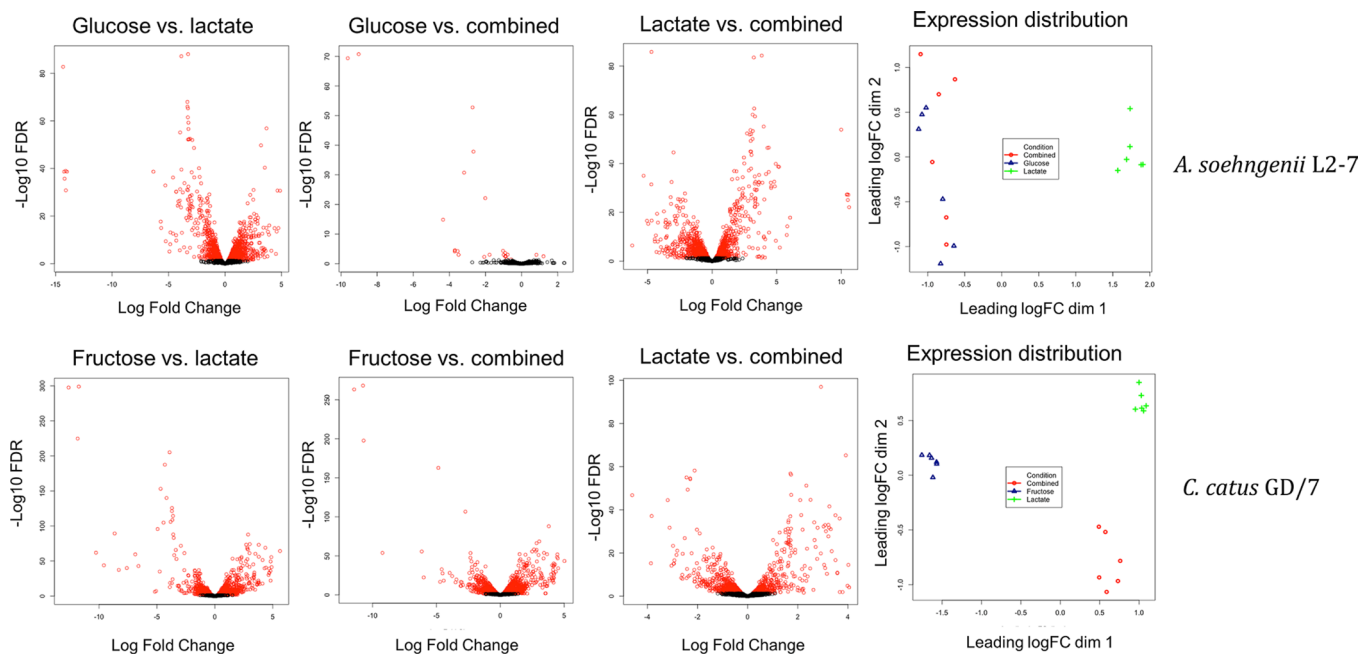


Fig. 3. Volcano plots of log fold gene-expression change vs. $-\log_{10}$ of the false discovery rate (FDR) value and multidimensional scaling plots showing the gene-expression distribution of samples during growth on lactate, monosaccharide or a combination of the two. Genes with a significant FDR value <0.05 in the volcano plots are highlighted in red. A complementary numerical overview is given in Table S5. 'Combined' indicates cultures incubated with both lactate and hexose.

5.0-pl) [52] with settings to split multi-mapping reads as a fraction between aligned regions. Genes with significant differential expression ($FDR < 0.05$) between conditions were identified using edgeR v3.16.5 [53], utilizing the GLM function, as more than two growth conditions were used in this study.

Short-chain fatty acid, monosaccharide and alcohol analysis

SCFA concentrations were measured by gas chromatography (GC) as described previously [54]. In brief, following derivatization of the samples using N-tert-butyltrimethylsilyl-N-methyltrifluoroacetamide, the samples were analysed using a Hewlett Packard gas chromatograph (GC) fitted with a silica capillary column using helium as the carrier gas. L-lactate concentration was assessed by Lactate Reagent (Trinity biotech) using a Konelab 30 chemistry analyser. D-lactate was calculated as the difference between total lactate and L-lactate. Methanol, ethanol, propanol, butanol and pentanol concentrations were also measured by GC using a ZB WAX column. Glucose and fructose concentration were assessed using the glucose hexokinase and fructose hexokinase assays from Konelab, Clinical Diagnostics Finland.

RESULTS

Lactate utilization in human colonic bacteria

Some anaerobic colonic bacteria are capable of utilizing lactate as a carbon and energy source. This process can generate the SCFAs butyrate, propionate or acetate depending on the species [22, 25]. Previous studies have been based on only a small number of available strains, but the isolation and genome sequencing here of new strains of lactate-utilizing species allowed for more in-depth analysis. Each strain was grown in culture provided with glucose (or fructose in the case of *C. catus* GD/7), DL-lactate, or a combination of the two as energy sources (Fig. 1). All nine strains identified as either *Anaerostipes* or *Anaerobutyricum* spp. showed the ability to consume lactate and acetate with production mainly of butyrate, as previously reported [22]. This contrasts with production mainly of propionate from lactate by *C. catus* (Fig. 1). The *A. hadrus* strains tested showed little or no ability to utilize L-lactate, whereas the other species utilized both isomers (Fig. 1b). Lactate utilization was repressed to varying extents by the presence of hexose sugar in the growth medium for the different *Anaerobutyricum* and *Anaerostipes* strains, with utilization of both lactate stereoisomers particularly repressed by the presence of glucose for *A. soehngenii* strains L2-7 and SL6/1/1, but no repression of D-lactate utilization in *A. hadrus* strains (Fig. 1b). Interestingly, although propionate was the main SCFA produced by *C. catus* GD/7, when grown on lactate some butyrate and acetate production was also observed, indicating the presence of these two SCFAs as intermediate- or end-products of lactate utilization in this organism.

Table 2. Selected genes with high expression differences during growth on glucose (G), fructose (F), lactate (L) or a combination of substrates (LG/LF). All transcripts are listed that exhibited >5 log2-fold changes in expression, together with expression changes for linked genes within the same gene cluster. Changes of less than fivefold are listed for the Rnf and Hnd genes clusters in *A. soehngenii* because of their metabolic relevance. All listed genes have an FDR value of <0.05. NoDiff=No significant difference in expression. Expanded to all genes in Tables S6–S11

		log2-fold change		
<i>A. soehngenii</i> L2-7		G v L	G v LG	L v LG
L2-7_01905	Lactate permease (<i>lctE</i>)	-14.3	-4.4	10.0
L2-7_01906	D-iLDH (<i>lctD</i>)	-14.2	-3.7	10.5
L2-7_01907	ETF-B (<i>lctB</i>)	-14.1	-3.5	10.6
L2-7_01908	ETF-A (<i>lctC</i>)	-14.2	-3.7	10.5
L2-7_01909	Acyl-CoA dehydrogenase (<i>lctG</i>)	-14.1	-3.7	10.4
L2-7_01910	Lactate racemase (<i>lctF</i>)	-14.0	-3.5	10.5
L2-7_01911	LutR transcriptional regulator (<i>lctA</i>)	-3.0	NoDiff	2.5
L2-7_02157	Aquaporin family protein	-6.3	-9.6	-3.3
L2-7_02158	Racemase	-5.3	-9.0	-3.7
L2-7_02594	Putative cyclic lactone autoinducer peptide	-5.2	NoDiff	5.8
L2-7_02595	Hypothetical protein	-4.9	NoDiff	4.9
L2-7_02596	RNA polymerase subunit sigma-70	-5.7	NoDiff	6.0
L2-7_02597	Helix-turn-helix domain-containing protein	-5.7	NoDiff	5.8
L2-7_00821	Hnd complex, subunit A	-4.33	NoDiff	5.12
L2-7_00822	Hnd complex, subunit C	-4.07	NoDiff	4.56
L2-7_00823	Hnd complex, subunit D	-3.98	NoDiff	4.38
L2-7_00377	Rnf complex, subunit C	-2.90	NoDiff	2.91
L2-7_00378	Rnf complex, subunit D	-3.08	NoDiff	2.98
L2-7_00379	Rnf complex, subunit G	-3.20	NoDiff	3.13
L2-7_00380	Rnf complex, subunit E	-3.33	NoDiff	3.25
L2-7_00381	Rnf complex, subunit A	-3.27	NoDiff	3.22
L2-7_00382	Rnf complex, subunit B	-3.24	NoDiff	3.15
		log2-fold change		
<i>C. catus</i> GD/7		F v L	F v LF	L v LF
GD-7_01094	Propionyl CoA transferase (<i>lapA</i>)	-10.3	-9.2	1
GD-7_01095	Lactoyl CoA epimerase (<i>lapB</i>)	-11.8	-10.7	1.1
GD-7_01096	Lactoyl CoA dehydratase subunit (<i>lapC</i>)	-12.6	-11.4	1.2
GD-7_01097	Lactoyl CoA dehydratase subunit (<i>lapD</i>)	-13.1	-11.9	1.2
GD-7_01098	Lactoyl CoA dehydratase subunit (<i>lapE</i>)	-13	-11.9	1.2
GD-7_01099	Lactate permease (<i>lapF</i>)	-13	-11.9	1.2
GD-7_00065	3-hydroxybutyryl-CoA dehydratase	-8.6	-6.2	2.5
GD-7_00066	H+/gluconate symporter and related permeases	-6.9	-4.6	2.3
GD-7_00760	Amidohydrolase family protein	-7.6	-4.4	3.2
GD-7_00761	MFS transporter	-8.3	-4.7	3.6
GD-7_01087	Acyl-CoA dehydrogenase	-9.6	-6	3.6
GD-7_01088	ABC transporter substrate-binding protein	-6.6	-3.5	3.1
GD-7_01551	Acyl-CoA dehydrogenase	-11.7	-10.8	1

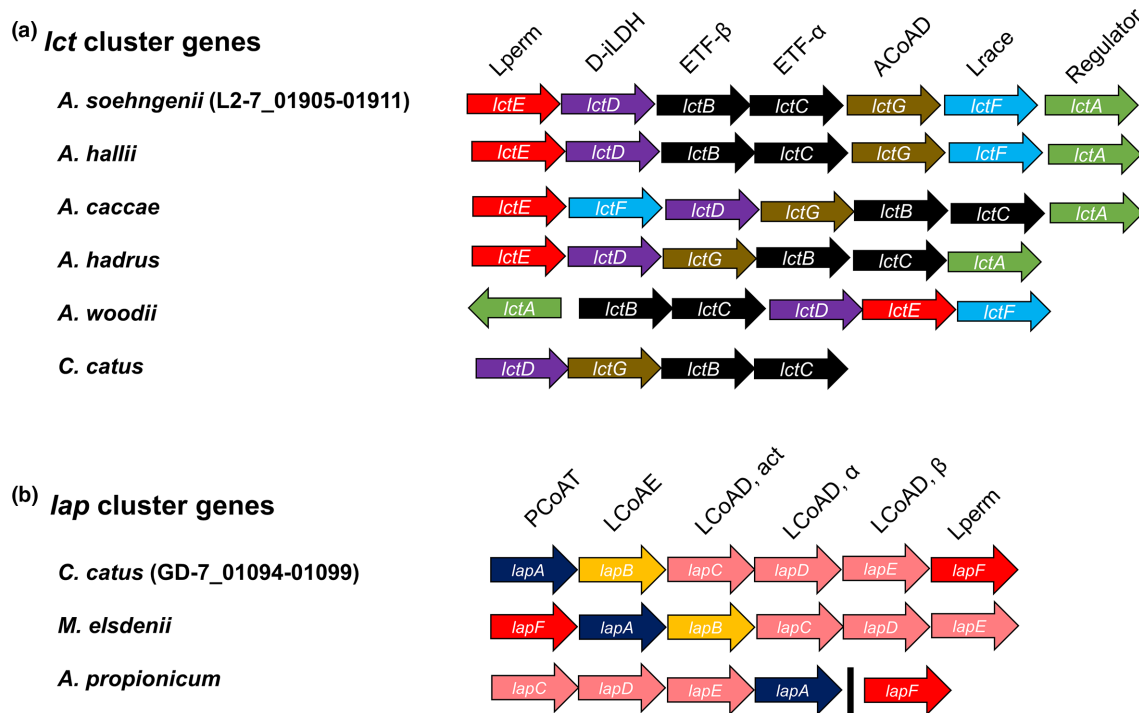


Fig. 4. Organization of lactate utilization loci observed in (a) *A. soehngeni* and (b) *C. catus* and other previously described lactate-utilizing bacteria. The black line in the *A. propionicum* *lap* cluster indicates that *lapF* gene is present in a different region of the genome. Lactate permease (Lperm), NAD-independent D-lactate dehydrogenase (D-iLDH), electron transfer flavoprotein (ETF α and β), acyl-CoA dehydrogenase (ACoAD), lactate racemase (Lracc), lactate operon transcriptional regulator (Regulator) propionyl-CoA transferase (PCoAT), lactoyl CoA epimerase (LCoAE), lactoyl-CoA transferase (LCoAD, act, α and β). Expanded to more lactate-utilizing bacteria genomes in Tables S13 and S14.

Changes in gene expression during growth on DL-lactate

To investigate the genetic control and regulation of lactate utilization, transcriptomic analyses were performed on two prevalent lactate-utilizing species from the human gut, *A. soehngeni* and *C. catus*. This involved assessing the differential expression of genes during growth on a hexose sugar (glucose for *A. soehngeni* and fructose for *C. catus*), on DL-lactate or on a combination of sugar plus DL-lactate (Fig. S1). Growth was carefully monitored in each of eight replicate experiments for each strain to ensure sampling during exponential growth phase (all cells harvested at $OD_{650}=0.3$ in order to keep this consistent between experiments and strains). RNA was extracted from cell pellets and the culture supernatants were analysed for their SCFA profiles and remaining hexose (Fig. 2). In agreement with the results shown in Fig. 1, in the absence of glucose *A. soehngeni* L2-7 had utilized much of the lactate at the time of sampling, but lactate utilization was significantly repressed in the presence of glucose ($P<0.001$). In the absence of hexose sugar, *C. catus* GD/7 had also utilized much of the lactate, but utilization was slightly lower in the presence of fructose ($P<0.01$). Sugar was still present at 65–70% of the initial concentration in the growth media supplemented with it (Fig. 2, Table S4), indicating there was potential for repression of lactate utilization at the point of RNA harvesting.

SCFA production levels were generally as expected, except for a relatively low concentration of butyrate in the *A. soehngeni* cultures incubated with lactate plus glucose compared to those from lactate growth alone (Fig. 2). Although speculative, we have two possible alternative explanations for this finding. It may be that more of the carbon went into cell mass, or was incorporated into other fermentation or intermediate precursor metabolites not measured here (such as CO_2 or propanediol) when both carbon sources were present.

Transcriptomics sequencing was carried out on RNA extractions from six of the replicates per strain (Fig. S1). The results from the experiments with *A. soehngeni* L2-7 showed differences in overall transcriptomic profile were observed during growth on lactate compared to glucose alone or a combination of glucose and lactate (Fig. 3). Conversely, little overall difference in transcriptomic profile was observed between glucose and the combination of glucose and lactate, with only 25 differentially expressed genes as opposed to 1427 between glucose and lactate as a sole carbon source. In *C. catus* GD/7, the overall transcriptomic profile was again strongly influenced by the growth substrate (Fig. 3). In *C. catus* GD/7, however, and in contrast to *A. soehngeni* L2-7, lactate-induced genes were not simply repressed by the presence of monosaccharide (fructose), with all three transcriptome profiles differing from each other.

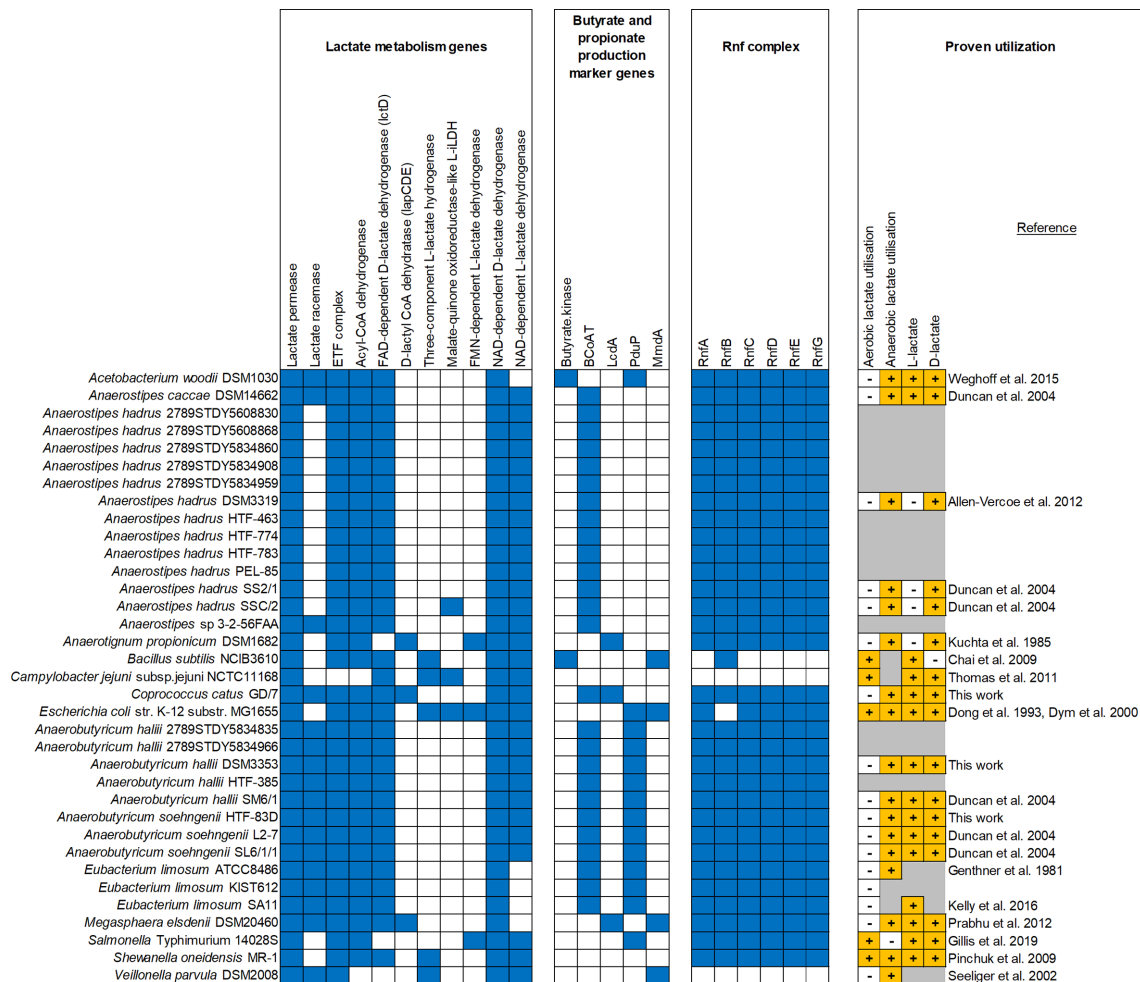


Fig. 5. Distribution of lactate utilization and associated genes in lactate-utilizing bacteria. Presence (blue) and absence (white) of genes was determined as described in Methods. The ability (+) or inability (-) of selected strains to utilize different isomers of lactate during aerobic or anaerobic growth was obtained from the literature [12, 21–23, 57, 59, 67, 75–81]. Certain details could not be recovered for some strains (marked in grey). Full names of genes indicated by short-hand gene nomenclature are as follows: Butyryl-CoA:acetate CoA-transferase (BCoAT), lactoyl-CoA dehydratase, alpha subunit (LcdA), propanediol utilization CoA-dependent propionaldehyde dehydrogenase (PduP), methyl-malonyl-CoA decarboxylase, alpha-subunit (MmdA) and Rnf complex subunits (RnfA-G).

In *A. soehngenii* L2-7, 11 transcripts showed >5 log₂-fold increase in expression during growth on lactate alone compared with growth on glucose alone, and increases were lower when glucose was present alongside lactate (Table 2). Six transcripts (L2-7_01905–01910) that showed the highest coordinate upregulation (log₂ >14) were encoded by the *lct* gene cluster that includes an NAD-independent lactate dehydrogenase (iLDH), lactate permease, ETF electron transfer flavoprotein (alpha and beta subunits), lactate racemase and an acyl-CoA dehydrogenase (Table 2). The linked regulatory gene *lctA* showed a lower amplitude of induced expression. A second cluster of three genes that showed the same behaviour, also at lower amplitude, included an autoinducer, RNA polymerase sigma factor and HTH regulatory protein. Two further transcripts, from a second lactate racemase and closely linked aquaporin genes, were also highly induced by lactate, but were not repressed by glucose. The components of the Rnf complex, which couples the oxidation of reduced ferredoxin to the reduction of NAD [38], were highly induced by lactate and repressed by glucose, as were the components of the Hnd complex, which catalyses the reduction of NADP in the presence of molecular H₂ to yield NADPH [55, 56].

In *C. catus* GD/7, seven genes showed >10 log₂ fold increased expression during growth on lactate alone compared with fructose alone (Table 2). Six of these were encoded by a cluster of linked genes, but this cluster showed no sequence similarity with the *lct* cluster of *A. soehngenii* described above and comprised genes (GD-7_01094–01099) encoding three components of lactoyl-CoA dehydratase, propionyl-CoA transferase, a methyl-malonyl-CoA epimerase homologue and a lactate permease (Table 2). These genes encode most of the activities required for the acrylate pathway to convert lactate into propionate, consistent with previous

evidence using isotopically labelled lactate showing that this bacterium uses the acrylate pathway for lactate utilization [25]. This locus is thus a cluster encoding lactate utilization via the acrylate pathway (*lap*). It seems likely that the methyl-malonyl-CoA epimerase homologue (*lapB*) functions as a lactoyl-CoA epimerase, catalysing the reversible conversion of L-lactoyl-CoA to D-lactoyl-CoA. The upregulated lactate permease gene was one of four encoded at separate locations in the *C. catus* GD/7 genome (Table S12). These six genes, together with two unlinked genes (GD-7_01551, GD-7_01087) encoding acyl-CoA dehydrogenases, were also highly upregulated on lactate when fructose was present. The two acyl-CoA dehydrogenases are candidates for propionyl-CoA dehydrogenase (acrylate reductase) activity.

Wider occurrence of lactate-utilization gene clusters in the genomes of human colonic bacteria

Homologues of the putative lactate utilization genes identified in *A. soehngenii* L2-7 and *C. catus* GD/7 were also detected in other known lactate-utilizing bacteria from the human colonic microbiota and other environments (Fig. 4a, b, Table S13 and S14). Homologues of the upregulated *lct* genes from *A. soehngenii* L2-7 were present in the two species of *Anaerobutyricum*, *Anaerostipes caccae*, *Eubacterium limosum*, *C. catus* and *Megasphaera elsdenii* (Table S13), although as noted above they were upregulated only moderately in *C. catus* (Table S9). The genome for *A. hadrus* DSM3319 lacked the racemase gene, consistent with the observation that it consumed only the D isomer of lactate (Fig. 1b). The *lct* homologues are also present in the *Acetobacterium woodii* gene cluster, but this cluster does not contain the putative acyl-CoA dehydrogenase (ACoAD) gene *lctG*.

Analysis of five other publicly available *C. catus* genomes confirmed the presence of highly similar *lap* cluster genes in all other tested isolates of this species (Table S14). Homologs of the six clustered *lap* cluster genes were also detected in *Megasphaera elsdenii*, another species known to use the acrylate pathway for lactate utilization, although the percentage identity was very low for the lactate permease gene (Table S14). Five of the six *lap* cluster genes, were detected in *Anaerotignum propionicum*, excluding *lapB*. This supports the prediction of *lapB* as a lactoyl-CoA epimerase, as *Anaerotignum propionicum* cannot utilize L-lactate [57].

In most butyrate-producing *Firmicutes* species from the human colon, most or all of the genes encoding the central butyrate pathway are present in a single gene cluster, although the butyryl-CoA acetate:CoA-transferase is separately encoded [58]. This was found to be true for *A. soehngenii* but not for *C. catus* GD/7, where the key functions are encoded in at least three separate locations in the genome (Table S15). While the crotonase and beta-hydroxybutyrate dehydrogenase genes are linked in *C. catus*

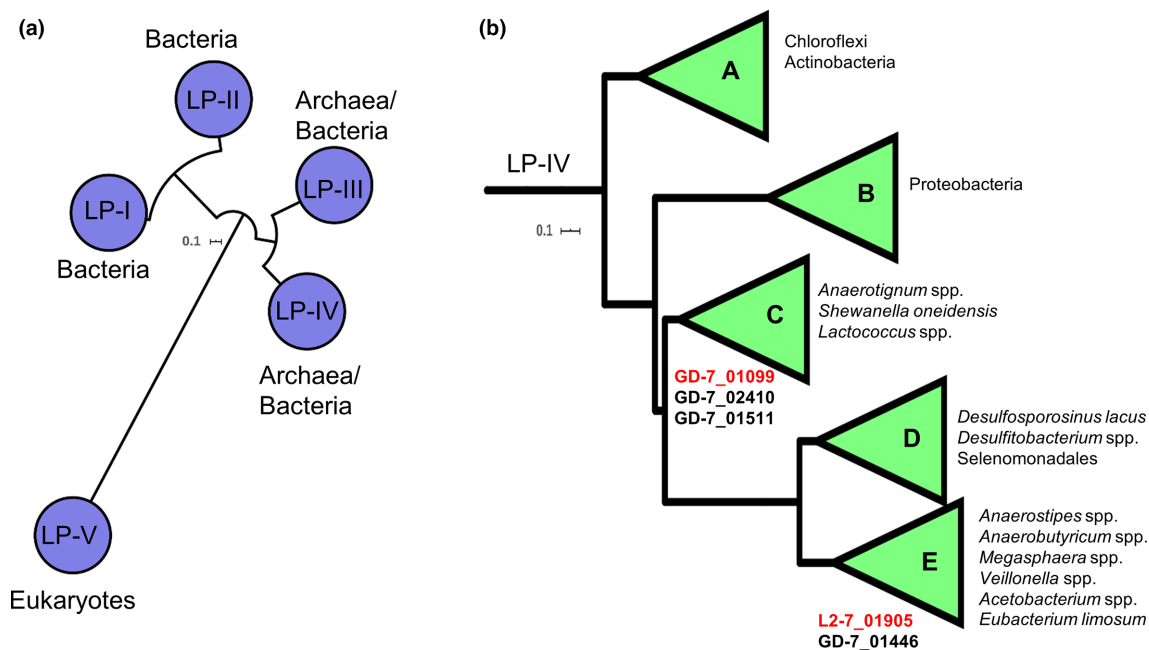


Fig. 6. Phylogeny of lactate permease families. (a) Families of lactate permease, and (b) subfamilies of LP-IV. The conserved lactate permease domain PF02652 was detected in 17653 protein sequences from Uniprot (June 2020). These proteins were formed into 1627 clusters of 70% identity. A maximum-likelihood tree was constructed from the representative sequences of these clusters using the LG+F+R8 model. All branches shown have >70% support of 1000 SH-*alrt* replicates. Proteins were clustered into protein families and subfamilies using the average branch length distances to the leaves in the phylogenetic tree. The tree was rooted using minimal ancestor deviation. Lactate permeases of *C. catus* GD/7 and *A. soehngenii* L2-7 are listed underneath the subfamilies of which they belong to. The upregulated lactate permeases detected in the preceding transcriptomics work in this study are indicated in red font. Representative genes of each family of lactate permease and each subfamily of LP-IV are listed in Table S16 and Table S17, respectively.

GD/7, two candidate thiolase genes are unlinked. Meanwhile, the butyryl-CoA dehydrogenase (BCD) in the *C. catus* GD/7 genome (GD-7_01569) is situated downstream of the iLDH (GD-7_01568) and upstream of two ETF genes (GD-7_01570, 01571) (Fig. 4a, Table S15). These appear to be homologous with the *lctBCFG* genes of *A. soehngenii*, although these *lct* genes were not strongly upregulated by lactate in *C. catus*.

Of the lactate-utilizing bacteria analysed, only *Veillonella parvula*, *Bacillus subtilis* and *Campylobacter jejuni* appear to lack the Rnf complex, and all genomes except *C. jejuni* and *Shewanella oneidensis* encoded genes indicative of butyrate and/or propionate production (Fig. 5). For bacteria that did not encode L-iLDHs, the lack of a lactate racemase corresponded with an inability to utilize L-lactate. As described earlier, the *lct* and *lap* clusters use different mechanisms of lactate utilization and are present in a different range of organisms. Additionally, *Shewanella oneidensis* uses a third mechanism of lactate utilization, involving a very different lactate dehydrogenase [59]. Homologues of this three-component L-lactate dehydrogenase are present in the genomes of *B. subtilis*, *C. jejuni*, *E. coli* and *V. parvula* (Fig. 5). Other types of NAD-independent L- and D-lactate dehydrogenases have also been found in organisms from all three domains of life [60]. Thus, there appear to be several divergent paradigms of lactate utilization in bacteria. However, all of these mechanisms appear to require members of the lactate permease superfamily and most possess some member of the acyl-CoA dehydrogenase superfamily. Therefore, the diversity of these proteins warranted detailed further investigation.

Phylogeny of lactate permease superfamilies

Phylogenetic analysis of 17653 lactate permeases from bacteria, archaea and eukaryotes revealed five families of lactate permease: LP-I and LP-II, which were present exclusively in bacteria; LP-III and LP-IV, which were present in bacteria and archaea; and LP-V, which was present exclusively in eukaryotes (Fig. 6a). The lactate permeases identified in the transcriptomics work presented here belong to LP-IV, which was further divided into five subfamilies (A–E). Three of the four lactate permease genes in the *C. catus* GD/7 genome, including the copy upregulated on lactate, clustered with the *Anaerotignum* and *S. oneidensis* lactate permeases in LP-IV-C (Fig. 6b). The fourth clustered with the lactate-induced *A. soehngenii* lactate permease in LP-IV-E. This subfamily notably also contains the lactate permeases of *Anaerostipes* spp., *Megasphaera* spp., *Veillonella* spp., *Acetobacterium* spp. and *E. limosum*.

Lactate permeases are also notably absent from the genomes of the gut bacteria *Eubacterium rectale* A1-86, *Roseburia intestinalis* L1-82, *Roseburia hominis* A2-183, *Coprococcus* sp. L2-50 and *Faecalibacterium prausnitzii* A2-165, which do not utilize lactate [22], thus indicating a role for lactate permease as a marker gene of lactate utilization, despite disparate mechanisms of lactate oxidation.

Phylogeny of Acyl-CoA dehydrogenase superfamilies

The acyl-CoA dehydrogenase in the *lct* cluster (L2-7_01909) and its homologues in the *C. catus* and *A. hadrus* genomes were used to create a HMM profile to detect divergent members of this protein superfamily in the genomes of other lactate-utilizing bacteria. These genes were found to be ubiquitous in lactate-utilizing bacteria, with many genomes possessing more than one copy (Table 3). Analysis of these genes revealed that they can be divided into a multitude of clades (Fig. 7), many of which, such as the experimentally proven [61] acyl-CoA reductase of *A. woodii*, are the only representatives of their clade in our dataset. One clade (acyl-CoA dehydrogenase family 5) contained the butyryl-CoA dehydrogenase present in the gene cluster responsible for butyrate production from carbohydrates [58]. This clade also contained the lactate induced *lct* cluster acyl-CoA dehydrogenase, indicating that it too is a butyryl-CoA dehydrogenase. The highly lactate-induced acyl-CoA dehydrogenases of *C. catus* (GD-7_01551 and GD-7_01087) form separate clades (acyl-CoA dehydrogenase families 2 and 4, respectively) with genes from several other species. No genes of these clades have been experimentally characterized, complicating functional prediction. Additionally, genes from *C. catus*, *A. soehngenii* and *A. caccae* form a clade (family 1) with the experimentally proven [62] acryloyl reductase of *Anaerotignum propionicum*.

A. soehngenii L2-7 and *C. catus* GD/7 possess four and five acyl-CoA dehydrogenases, respectively (Fig. 7, Table 3). Two of the *A. soehngenii* L2-7 gene products group with butyryl-CoA dehydrogenase (BCD) enzymes, including the BCD in the central butyrate pathway (L2-7_00754) and the upregulated *lctG* product (L2-7_01919). This strongly suggests that the lactate-induced *lctG* gene product is an alternative BCD enzyme that links reduction of crotonyl-CoA to the oxidation of lactate when lactate is being utilized.

DISCUSSION

Lactate can be utilized by many aerotolerant bacteria through its conversion to pyruvate via D- or L-lactate oxidases [60, 63]. This reaction is not available to obligate anaerobes however, and different mechanisms are therefore required in the absence of oxygen. As noted previously, anaerobic conversion of lactate to pyruvate is energetically unfavourable [23] and this ability appears to have a limited phylogenetic distribution. Here, we have identified two different six-gene clusters (*lct* and *lap*) whose transcriptional expression is highly upregulated during growth on lactate as the sole added carbon and energy source. These two gene clusters correspond to two entirely different mechanisms for lactate utilization. The upregulated cluster in *A. soehngenii* corresponds to

Table 3. Acyl-CoA dehydrogenase families of lactate-utilizing bacteria. These enzymes are ubiquitous in lactate-utilizing bacteria, with some genomes possessing multiple genes. Many of these genes form five distinct gene families (Fig. 7). The number of genes of a given family that were upregulated during growth on lactate in *A. soehngenii* L2-7 and *C. catus* GD/7 is shown in parentheses. Black slashes indicate differences in gene copy number between queried genomes of the same species. Expanded to include more lactate-utilizing bacteria genomes in Table S18

Species	Acyl-CoA dehydrogenase families					No. of genomes queried
	1	2	3	4	5	
<i>Coprococcus catus</i>	2	1 (1)	0	1	1 (1)	1
<i>Anaerobutyricum soehngenii</i>	0/1	0	0	0/1	2 (1)	3
<i>Anaerobutyricum hallii</i>	0	0	0	0	2	5
<i>Anaerostipes caccae</i>	1	1	1	2	1	1
<i>Anaerostipes</i> sp. 3-2-56FAA	1	1	1	2	1	1
<i>Anaerostipes hadrus</i>	0	0/1	1	0/1/2	1	12
<i>Eubacterium limosum</i>	0	0	0	0/1	0/2	3
<i>Megasphaera elsdenii</i>	0	0	2	0	0	1
<i>Escherichia coli</i>	0	0	0	1	0	1
<i>Anaerotignum propionicum</i>	1	1	0	0	1	1

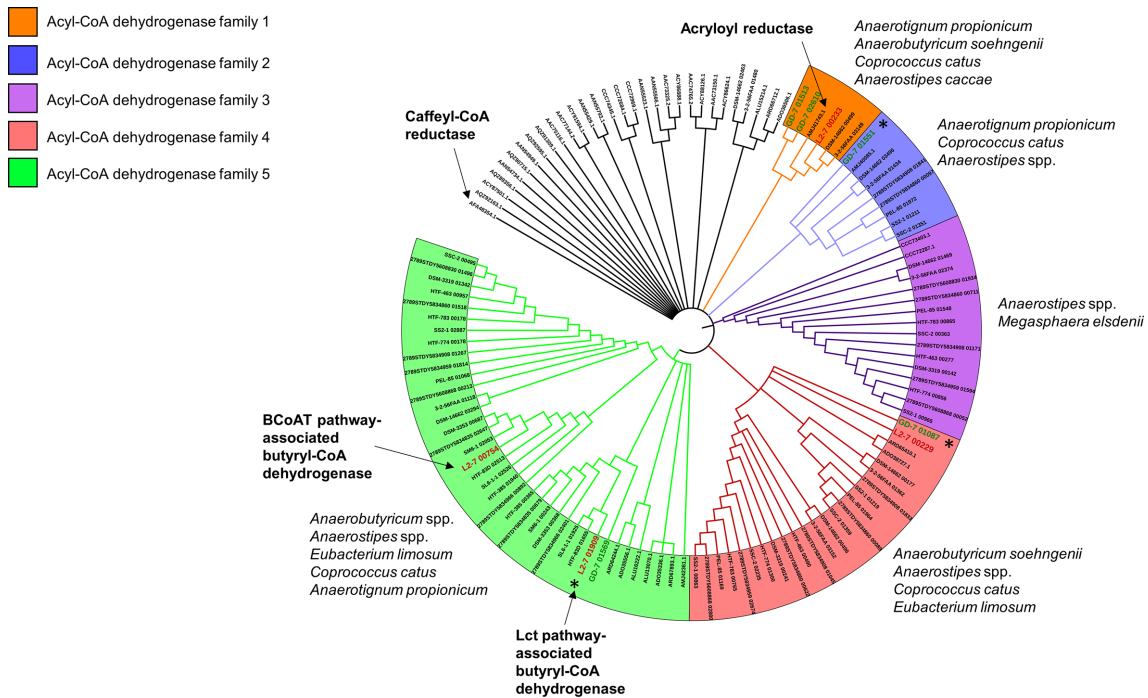


Fig. 7. Phylogeny of Acyl-CoA dehydrogenase genes present in lactate-utilizing bacteria. Branch validation was performed using 1000 ultrafast bootstrap replicates and a hill-climbing nearest-neighbour interchange (NNI) search was performed to reduce the risk of overestimating branch supports. Branches with less than 90% UFBoot support were collapsed. Genes labelled 'Caffeyl-CoA reductase' and 'Acryloyl reductase' were experimentally verified by Bertsch et al. [61] and Hetzel et al. [62], respectively, and the gene labelled 'Butyryl-CoA dehydrogenase' is present in the butyrate production loci as described previously [58]. This maximum-likelihood tree was reconstructed using the LG+G4 model. The tree was rooted using minimal ancestor deviation. *Anaerobutyricum soehngenii* L2-7 and *Coprococcus catus* GD/7 genes are represented in red and green font, respectively. Asterisks indicate genes that were highly upregulated by lactate in the transcriptomics work presented in this study.

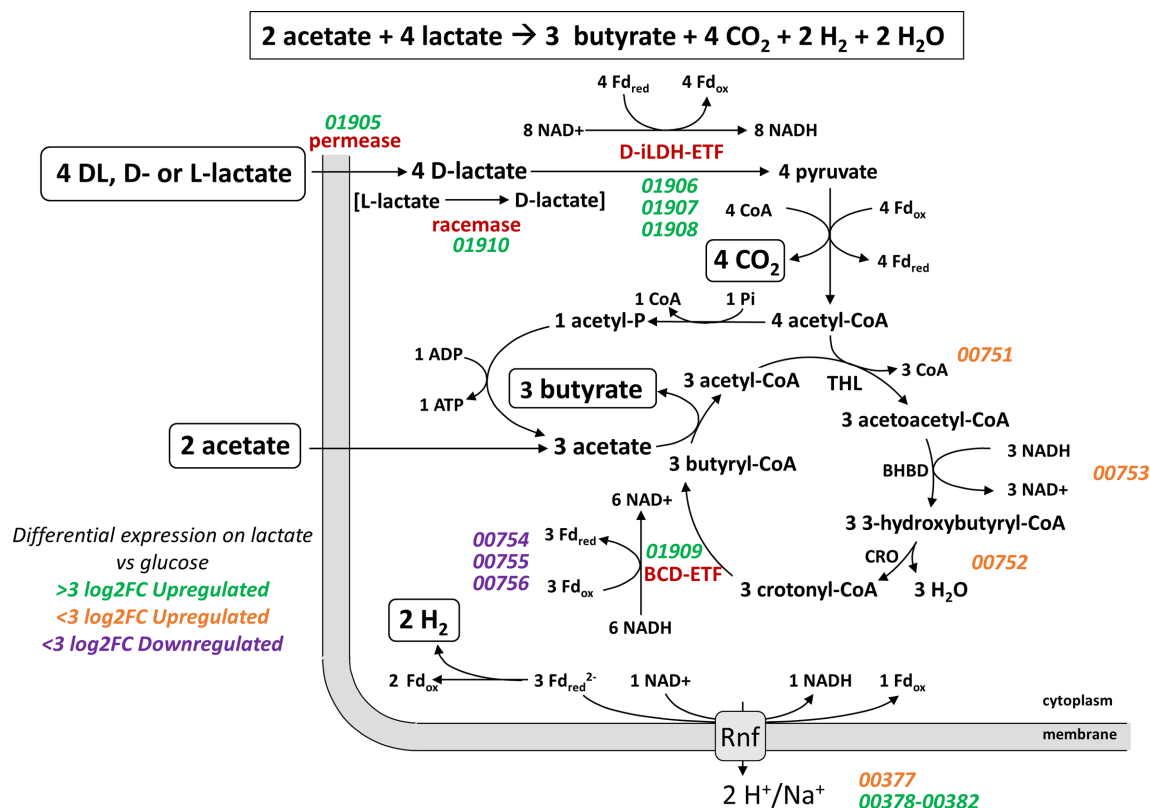


Fig. 8. Proposed mechanisms of butyrate formation from lactate in *A. soehngeni* L2-7. Thiolase (THL), β -hydroxybutyryl-CoA dehydrogenase (BHBD), crotonase (CRO), NAD-independent D-lactate dehydrogenase (D-iLDH), electron transfer flavoprotein (ETF), ferredoxin (Fd), butyryl-CoA dehydrogenase (BCD), adenosine triphosphate (ATP), adenosine diphosphate (ADP), oxidized nicotinamide adenine dinucleotide (NAD⁺), reduced nicotinamide adenine dinucleotide (NADH), inorganic phosphate (Pi), and Coenzyme A (CoA). Locus tags are in italics and are coloured based on log₂ fold change of their expression during growth on lactate versus glucose.

five genes of the *lct* gene cluster identified previously in *Acetobacterium woodii*, which is an acetogen that has been shown to convert lactate to pyruvate via a mechanism involving electron conformation [23]. Products of this *lct* gene cluster were recently detected by proteomic analysis [27], and the cluster reported in the related lactate-utilizing species *A. hallii*, *A. rhamnosivorans* and *Anaerostipes caccae*. Five of the six clustered genes are also present in *Anaerostipes hadrus*, but all strains of this species with sequenced genomes examined here lacked the lactate racemase (*lctF*), thus explaining why they can only utilize D-lactate.

Our analysis confirms that the *lct* gene cluster is present in the newly sequenced strains of *A. soehngeni*, *A. hallii* and *A. hadrus* described here. Notably, however, the *lct* cluster in all these strains encodes an acyl-CoA dehydrogenase specified by the *lctG* gene that is lacking from the *A. woodii* cluster [64]. Since the *lctG* gene product is related to, but distinct from, the BCD gene encoded by the central butyrate pathway cluster (L2-7_00751-00756) [58], it seems likely that it is an alternative butyryl-CoA dehydrogenase (BCD) that acts to reduce crotonyl-CoA when lactate is the growth substrate (Fig. 8). While the conversion of pyruvate to acetyl-CoA can produce enough reduced ferredoxin to balance the oxidation of lactate via the D-iLDH/ETF complex [65], redox balance and net energy production are also linked to the operation of the butyrate cycle (Fig. 8). The upregulation by lactate of an orthologous enzyme catalysing the BCD reaction, encoded by *lctG*, may therefore play a critical role. The exact relationships between the co-ordinately upregulated iLDH, BCD and ETF proteins are not known for *A. soehngeni*, although it has been suggested that in *C. butyricum* the LDH-ETF and BCD-ETF complexes act separately [66].

It has been suggested that the molecular hydrogen producing-Hnd complex is involved in balancing redox states during lactate oxidation in *Desulfovibrio* species [56], and its induction during growth of *A. soehngeni* on lactate indicates a likely mechanism for the lactate oxidation-coupled hydrogen production observed in this species [22]. However, resolution of these mechanisms will require further detailed biochemical investigation of the enzymes' activity and ETF binding of the LctG protein. It would be of particular interest to carry out enzymatic studies to establish that the *lctG* product does indeed have butyryl-CoA dehydrogenase activity. In addition, it would be interesting to carry out affinity studies comparing the

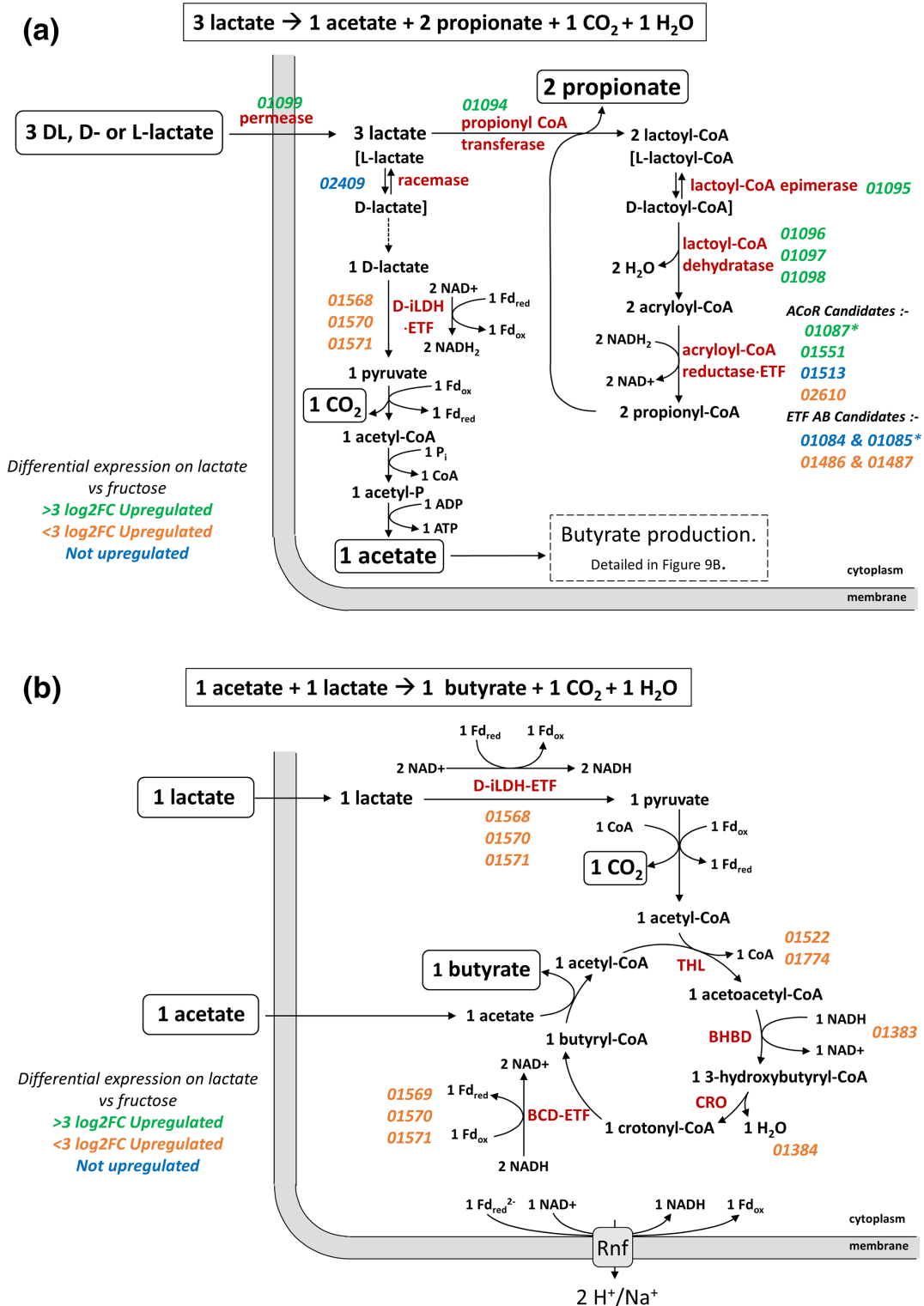


Fig. 9. Proposed mechanisms of (a) propionate and acetate formation from lactate and (b) butyrate formation from acetate and lactate in *C. catus* GD/7. NAD-independent D-lactate dehydrogenase (D-iLDH), electron transfer flavoprotein (ETF), ferredoxin (Fd), thiolase (THL), β -hydroxybutyryl-CoA dehydrogenase (BHBD), crotonase (CRO), butyryl-CoA dehydrogenase (BCD), adenosine triphosphate (ATP), adenosine diphosphate (ADP), oxidized nicotinamide adenine dinucleotide (NAD⁺), reduced nicotinamide adenine dinucleotide (NADH), inorganic phosphate (Pi), and Coenzyme A (CoA). Locus tags are in italics and are coloured based on log₂ fold change of their expression during growth on lactate versus fructose. Asterisks indicate most likely candidate genes.

binding preferences of the LctG enzyme and its BCD homologue for the two pairs of ETFAB proteins that are encoded by the *lct* cluster and by the central butyrate pathway genes.

The *lap* cluster was identified in *C. catus* GD/7. This cluster has no equivalent in *A. soehngenii* and codes for enzymes required for the acrylate pathway for the conversion of lactate to propionate, notably the three subunits of lactoyl-CoA dehydratase, propionyl-CoA transferase and a presumptive lactoyl-CoA epimerase (Fig. 9). Although lactate permeases are present in both the *lap* and *lct* clusters, their sequences are only distantly related to each other. In addition, it seems likely that one of the two unlinked acyl-CoA dehydrogenases that were also upregulated by lactate must correspond to propionyl-CoA dehydrogenase that, along with ETF proteins, carries out the acryloyl-CoA reductase reaction [62].

Operation of the acrylate pathway is normally assumed to result in the conversion of three mol lactate to two mol propionate and one mol of acetate (Fig. 9a) [3]. The acetate arises through conversion of lactate to pyruvate via iLDH, with the reducing equivalents balancing those needed to reduce two acrylate to two propionate. Our results however indicate the formation of some butyrate, in addition to propionate, during growth on lactate. This can be explained readily if we assume that some of the acetate produced is converted along with additional lactate to butyrate (Fig. 9b), as happens in *A. soehngenii*. The finding that the iLDH gene is linked to BCD and ETF genes in *C. catus* GD/7 appears consistent with this proposal. An unexplained feature is the 3-OH butyryl-CoA dehydratase transcript (GD-7_00065) that was upregulated by lactate, although this might possibly be an orthologue of the second β -hydroxybutyryl-CoA dehydrogenase gene (GD-7_01384) that is linked to crotonase in *C. catus* GD/7.

Homologues of *lap* genes were detected in a few other bacterial species, but notably were present in *Megasphaera elsdenii*, which is also known to use the acrylate pathway for lactate utilization. Indeed, *M. elsdenii* is known to convert lactate to butyrate under conditions of carbon-limited growth [67]. It is not known whether the same applies in *C. catus*. *M. elsdenii* also encodes activities involved in the formation of propionyl-CoA from pyruvate via malate and succinate [68]. However, the gene encoding for methyl-malonyl-CoA decarboxylase alpha-subunit (*mmdA*), an indicator of propionate production through the succinate pathway, was not present in the genome of *C. catus* GD/7 (Fig. 5), indicating that *C. catus* does not utilize this pathway.

Two other anaerobic lactate-utilizing species, *Veillonella parvula* and *Anaerotignum propionicum*, showed high levels of sequence identity with several genes of the *lct* cluster, but lacked a close homologue of the α -iLDH. The latter species however also carries homologues for most of the *lap* cluster and utilizes lactate via the acrylate pathway [69], whereas *V. parvula* encodes a three-component L-lactate dehydrogenase, homologous to that of *S. oneidensis* (Fig. 5), and uses the 'randomizing' succinate pathway rather than the acrylate pathway in producing propionate [3].

To summarize, we have shown that two distinct routes for anaerobic lactate fermentation in obligate anaerobes, via pyruvate to butyrate and via acrylate to propionate, both involve major transcriptional upregulation of the key genes involved (the *lct* cluster in *A. soehngenii* and the *lap* cluster in *C. catus*) in response to lactate. There was however a clear difference between the two strains in the impact of hexose sugars on gene expression. Whereas *A. soehngenii* *lct* expression was subject to partial repression by glucose, this was not the case for *C. catus* acrylate genes with fructose. We should note, however, that present and previous results detected no glucose repression in strains of *A. hadrus* (referred to as 'A. coli' SS2/1 in Muñoz-Tamayo *et al.* [70]), which, like *A. soehngenii*, also possess the *lct* cluster and produce butyrate from D-lactate. The lack of repression of lactate utilization genes by hexose in species such as *C. catus* and *A. hadrus* may give them particularly important roles in controlling lactate concentrations within gut communities under conditions when there are significant concentrations of free sugars. For example, in the rumen, *C. catus* and *Megasphaera elsdenii* have been implicated in differences in animal productivity that are associated with lactate utilization on diets high in readily fermentable carbohydrates [15]. In the human colon monosaccharides are generally assumed to be present at very low concentrations, with non-digestible polysaccharides and oligosaccharides being the main available energy sources for bacterial growth. Duncan *et al.* [22] showed that in co-cultures between *Bifidobacterium adolescentis* and *A. soehngenii* L2-7 (then named 'Eubacterium hallii') growing on starch, lactate produced from starch by *B. adolescentis* was converted completely into butyrate. This suggests that bacteria such as *A. soehngenii* can play a very significant role in preventing lactate accumulation where polysaccharides such as resistant starch are also available as energy sources. A more detailed, quantitative, investigation into the impact of different concentrations of hexoses, oligosaccharides and polysaccharides on the lactate utilization rates in a range of dominant gut lactate-utilizing bacteria, preferably done in anaerobic continuous culture, would clearly be desirable in the future. Better understanding of the regulatory responses to these alternative carbon sources may be relevant when selecting strains of lactate-utilizing bacteria intended for use as therapeutic probiotics with the aim of reducing intestinal lactate concentrations and enhancing the stability of the microbial community.

Funding information

P.L., S.H.D., H.J.F., A.W.W. and the Rowett Institute received core financial support from the Scottish Government Rural and Environmental Sciences and Analytical Services (SG-RESAS). The work was also partly funded by Chr. Hansen, a global bioscience company that develops natural ingredient

solutions for the food, nutritional, pharmaceutical and agricultural industries, who supported P.O.S. and E.T. Chr. Hansen exerted no influence on results obtained and presented in this manuscript. This also does not alter our adherence to Microbial Genomics' policies on sharing data and materials.

Acknowledgements

The authors would like to thank Donna Henderson (Rowett Institute, University of Aberdeen) for carrying out gas chromatography analysis of short-chain fatty acids, and also thank the Centre for Genome-Enabled Biology and Medicine (CGEBM) at the University of Aberdeen for carrying out the Illumina-based transcriptomics sequencing. We would also like to acknowledge the support of the Maxwell computer cluster funded by the University of Aberdeen.

Author contributions

P.O.S. carried out laboratory and *in silico* experiments/data analysis. P.L., S.H.D., H.J.F. and A.W.W. also contributed to experimental design, data interpretation and visualization. E.T. and H.J.M. generated essential strain data. S.S. contributed to formal analysis and visualization. P.O.S., P.L., S.H.D., H.J.F. and A.W.W. wrote the manuscript. All co-authors commented on advanced drafts of the manuscript, and contributed edits to the final draft.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

- Sheridan PO, Louis P, Tsompanidou E, Shaw S, Harmsen HJ, et al. *Figshare* 2022.
- Flint HJ, Scott KP, Louis P, Duncan SH. The role of the gut microbiota in nutrition and health. *Nat Rev Gastroenterol Hepatol* 2012;9:577–589.
- Gottschalk G. *Bacterial Metabolism*. Verlag: New York, Heidelberg, Berlin: Springer; 1979.
- Duncan SH, Louis P, Thomson JM, Flint HJ. The role of pH in determining the species composition of the human colonic microbiota. *Environ Microbiol* 2009;11:2112–2122.
- McWilliam Leitch EC, Stewart CS. Susceptibility of *Escherichia coli* O157 and non-O157 isolates to lactate. *Lett Appl Microbiol* 2002;35:176–180.
- Stewart CJ, Ajami NJ, O'Brien JL, Hutchinson DS, Smith DP, et al. Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature* 2018;562:583–588.
- Slyter LL. Influence of acidosis on rumen function. *J Anim Sci* 1976;43:910–929.
- Russell JR, Hino T. Regulation of lactate production in *Streptococcus bovis*: A spiraling effect that contributes to rumen acidosis. *J Dairy Sci* 1985;68:1712–1721.
- Kaneko T, Bando Y, Kurihara H, Satomi K, Nonoyama K, et al. Fecal microflora in a patient with short-bowel syndrome and identification of dominant lactobacilli. *J Clin Microbiol* 1997;35:3181–3185.
- Vernia P, Caprilli R, Latella G, Barbetti F, Magliocca FM, et al. Fecal lactate and ulcerative colitis. *Gastroenterology* 1988;95:1564–1568.
- Gillis CC, Hughes ER, Spiga L, Winter MG, Zhu W, et al. Dysbiosis-associated change in host metabolism generates lactate to support salmonella growth. *Cell Host & Microbe* 2018;23:54–64.
- Thomas MT, Shepherd M, Poole RK, van Vliet AHM, Kelly DJ, et al. Two respiratory enzyme systems in *Campylobacter jejuni* NCTC 11168 contribute to growth on L-lactate. *Environ Microbiol* 2011;13:48–61.
- Counotte GH, Lankhorst A, Prins RA. Role of DL-lactic acid as an intermediate in rumen metabolism of dairy cows. *J Anim Sci* 1983;56:1222–1235.
- Belenguer A, Duncan SH, Calder AG, Holtrop G, Louis P, et al. Two routes of metabolic cross-feeding between *Bifidobacterium adolescentis* and butyrate-producing anaerobes from the human gut. *Appl Environ Microbiol* 2006;72:3593–3599.
- Shabat SKB, Sasson G, Doron-Faigenboim A, Durman T, Yaacoby S, et al. Specific microbiome-dependent mechanisms underlie the energy harvest efficiency of ruminants. *ISME J* 2016;10:2958–2972.
- Wang SP, Rubio LA, Duncan SH, Donachie GE, Holtrop G, et al. Pivotal Roles for pH, lactate, and lactate-utilizing bacteria in the stability of a human colonic microbial ecosystem. *mSystems* 2020;5:e00645–20.
- Van Immerseel F, Ducatelle R, De Vos M, Boon N, Van De Wiele T, et al. Butyric acid-producing anaerobic bacteria as a novel probiotic treatment approach for inflammatory bowel disease. *J Med Microbiol* 2010;59:141–143.
- Perraudeau F, McMurdie P, Bullard J, Cheng A, Cutcliffe C, et al. Improvements to postprandial glucose control in subjects with type 2 diabetes: a multicenter, double blind, randomized placebo-controlled trial of a novel probiotic formulation. *BMJ Open Diabetes Res Care* 2020;8:e001319.
- Gilijamse PW, Hartstra AV, Levin E, Wortelboer K, Serlie MJ, et al. Treatment with *Anaerobutyricum soehngeni*: a pilot study of safety and dose-response effects on glucose metabolism in human subjects with metabolic syndrome. *NPJ Biofilms Microbiomes* 2020;6:16.
- Shetty SA, Zuffa S, Bui TPN, Aalvink S, Smidt H, et al. Reclassification of *Eubacterium hallii* as *Anaerobutyricum hallii* gen. nov., comb. nov., and description of *Anaerobutyricum soehngeni* sp. nov., a butyrate and propionate-producing bacterium from infant faeces. *Int J Syst Evol Microbiol* 2018;68:3741–3746.
- Allen-Vercoe E, Daigneault M, White A, Panaccione R, Duncan SH, et al. *Anaerostipes hadrus* comb. nov., a dominant species within the human colonic microbiota; reclassification of *Eubacterium hadrum* Moore et al. *Anaerobe* 1976;18:523–529.
- Duncan SH, Louis P, Flint HJ. Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. *Appl Environ Microbiol* 2004;70:5810–5817.
- Weghoff MC, Bertsch J, Müller V. A novel mode of lactate metabolism in strictly anaerobic bacteria. *Environ Microbiol* 2015;17:670–677.
- Counotte GH, Prins RA, Janssen RH, Debie MJ. Role of *Megasphaera elsdenii* in the Fermentation of dl-[2-C]lactate in the Rumen of Dairy Cattle. *Appl Environ Microbiol* 1981;42:649–655.
- Reichardt N, Duncan SH, Young P, Belenguer A, McWilliam Leitch C, et al. Phylogenetic distribution of three pathways for propionate production within the human gut microbiota. *ISME J* 2014;8:1323–1335.
- Marquet P, Duncan SH, Chassard C, Bernalier-Donadille A, Flint HJ. Lactate has the potential to promote hydrogen sulphide formation in the human colon. *FEMS Microbiol Lett* 2009;299:128–134.
- Shetty SA, Boeren S, Bui TPN, Smidt H, de Vos WM. Unravelling lactate-acetate and sugar conversion into butyrate by intestinal anaerobutyricum and anaerostipes species by comparative proteogenomics. *Environ Microbiol* 2020;22:4863–4875.
- Lopez-Siles M, Khan TM, Duncan SH, Harmsen HJM, Garcia-Gil LJ, et al. Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. *Appl Environ Microbiol* 2012;78:420–428.
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, et al. Versatile and open software for comparing large genomes. *Genome Biol* 2004;5:R12.
- Miyazaki K, Martin JC, Marinsek-Logar R, Flint HJ. Degradation and utilization of xylans by the rumen anaerobe *Prevotella bryantii* (formerly *P. ruminicola* subsp. *brevis*) B(1)4. *Anaerobe* 1997;3:373–381.

31. Bryant MP. Commentary on the Hungate technique for culture of anaerobic bacteria. *Am J Clin Nutr* 1972;25:1324–1328.
32. Peng Y, Leung HCM, Yiu SM, Chin FYL. IDBA-UD: a *de novo* assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 2012;28:1420–1428.
33. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;30:2068–2069.
34. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–1055.
35. Eddy SR. Profile hidden Markov models. *Bioinformatics* 1998;14:755–763.
36. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 2019;47:D506–D515.
37. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res* 2004;32:D277–80.
38. Tremblay PL, Zhang T, Dar SA, Leang C, Lovley DR. The Rnf complex of *Clostridium ljungdahlii* is a proton-translocating ferredoxin:NAD⁺ oxidoreductase essential for autotrophic growth. *mBio* 2012;4:e00406–12.
39. Biegel E, Schmidt S, Müller V. Genetic, immunological and biochemical evidence for a Rnf complex in the acetogen *Acetobacterium woodii*. *Environ Microbiol* 2009;11:1438–1443.
40. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, et al. The Pfam protein families database in 2019. *Nucleic Acids Res* 2019;47:D427–D432.
41. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 2012;28:3150–3152.
42. Katoh K, Toh H. Recent developments in the MAFFT multiple sequence alignment program. *Briefings in Bioinformatics* 2008;9:286–298.
43. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;25:1972–1973.
44. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;32:268–274.
45. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017;14:587–589.
46. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, et al. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* 2010;59:307–321.
47. Tria FDK, Landan G, Dagan T. Phylogenetic rooting using minimal ancestor deviation. *Nat Ecol Evol* 2017;1:1–7.
48. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 2007;23:127–128.
49. Andrews S. FastQC: a quality control tool for high throughput-sequence data. UK: Babraham Institute, 2010. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
50. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 2019;37:907–915.
51. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–2079.
52. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 2014;30:923–930.
53. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;26:139–140.
54. Richardson AJ, Calder AG, Stewart CS, Smith A. Simultaneous determination of volatile and non-volatile acidic fermentation products of anaerobes by capillary gas chromatography. *Lett Appl Microbiol* 1989;9:5–8.
55. de Luca G, de Philip P, Rousset M, Belaich JP, Dermoun Z. The NADP-reducing hydrogenase of *Desulfovibrio fructosovorans*: evidence for a native complex with hydrogen-dependent methylviologen-reducing activity. *Biochem Biophys Res Commun* 1998;248:591–596.
56. Malki S, Saimmaime I, De Luca G, Rousset M, Dermoun Z, et al. Characterization of an operon encoding an NADP-reducing hydrogenase in *Desulfovibrio fructosovorans*. *J Bacteriol* 1995;177:2628–2636.
57. Kuchta RD, Abeles RH. Lactate reduction in *Clostridium propionicum*. Purification and properties of lactyl-CoA dehydratase. *J Biol Chem* 1985;260:13181–13189.
58. Louis P, Flint HJ. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol Lett* 2009;294:1–8.
59. Pinchuk GE, Rodionov DA, Yang C, Li X, Osterman AL, et al. Genomic reconstruction of *Shewanella oneidensis* MR-1 metabolism reveals a previously uncharacterized machinery for lactate utilization. *Proc Natl Acad Sci U S A* 2009;106:2874–2879.
60. Jiang T, Gao C, Ma C, Xu P. Microbial lactate utilization: enzymes, pathogenesis, and regulation. *Trends Microbiol* 2014;22:589–599.
61. Bertsch J, Parthasarathy A, Buckel W, Müller V. An electron-bifurcating caffeyl-CoA reductase. *J Biol Chem* 2013;288:11304–11311.
62. Hetzel M, Brock M, Selmer T, Pierik AJ, Golding BT, et al. Acryloyl-CoA reductase from *Clostridium propionicum*. An enzyme complex of propionyl-CoA dehydrogenase and electron-transferring flavoprotein. *Eur J Biochem* 2003;270:902–910.
63. Gibello A, Collins MD, Domínguez L, Fernández-Garayzábal JF, Richardson PT. Cloning and analysis of the L-lactate utilization genes from *Streptococcus iniae*. *Appl Environ Microbiol* 1999;65:4346–4350.
64. Schoelmerich MC, Katsyv A, Sung W, Mijic V, Wiechmann A, et al. Regulation of lactate metabolism in the acetogenic bacterium *Acetobacterium woodii*. *Environ Microbiol* 2018;20:4587–4595.
65. Müller V, Chowdhury NP, Basen M. Electron bifurcation: a long-hidden energy-coupling mechanism. *Annu Rev Microbiol* 2018;72:331–353.
66. Detman A, Mielecki D, Chojnacka A, Salamon A, Błaszczak MK, et al. Cell factories converting lactate and acetate to butyrate: *Clostridium butyricum* and microbial communities from dark fermentation bioreactors. *Microb Cell Fact* 2019;18:36.
67. Prabhu R, Altman E, Eiteman MA. Lactate and acrylate metabolism by *Megasphaera elsdenii* under batch and steady-state conditions. *Appl Environ Microbiol* 2012;78:8564–8570.
68. Chen L, Shen Y, Wang C, Ding L, Zhao F, et al. *Megasphaera elsdenii* lactate degradation pattern shifts in rumen acidosis models. *Front Microbiol* 2019;10:162.
69. Johns AT. The mechanism of propionic acid formation by *Clostridium propionicum*. *J Gen Microbiol* 1952;6:123–127.
70. Muñoz-Tamayo R, Laroche B, Walter E, Doré J, Duncan SH, et al. Kinetic modelling of lactate utilization and butyrate production by key human colonic bacterial species. *FEMS Microbiol Ecol* 2011;76:615–624.
71. Shetty SA, Ritari J, Paulin L, Smidt H, De Vos WM. Complete genome sequence of *Eubacterium hallii* strain L2-7. *Genome Announc* 2017;5:e01167-17.
72. Barcenilla A, Pryde SE, Martin JC, Duncan SH, Stewart CS, et al. Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Appl Environ Microbiol* 2000;66:1654–1661.
73. Louis P, Duncan SH, McCrae SI, Millar J, Jackson MS, et al. Restricted distribution of the butyrate kinase pathway among

- butyrate-producing bacteria from the human colon. *J Bacteriol* 2004;186:2099–2106.
74. Holdeman LV, Moore WEC. New genus, *Coprococcus*, twelve new species, and emended descriptions of four previously described species of bacteria from human feces. *Int J Syst Bacteriol* 1975;25:98.
 75. Chai Y, Kolter R, Losick R. A widely conserved gene cluster required for lactate utilization in *Bacillus subtilis* and its involvement in biofilm formation. *J Bacteriol* 2009;191:2423–2430.
 76. Dong JM, Taylor JS, Latour DJ, Iuchi S, Lin EC. Three overlapping *lct* genes involved in L-lactate utilization by *Escherichia coli*. *J Bacteriol* 1993;175:6671–6678.
 77. Dym O, Pratt EA, Ho C, Eisenberg D. The crystal structure of D-lactate dehydrogenase, a peripheral membrane respiratory enzyme. *Proc Natl Acad Sci U S A* 2000;97:9413–9418.
 78. Genthner BR, Davis CL, Bryant MP. Features of rumen and sewage sludge strains of *Eubacterium limosum*, a methanol- and H₂-CO₂-utilizing species. *Appl Environ Microbiol* 1981;42:12–19.
 79. Kelly WJ, Henderson G, Pacheco DM, Li D, Reilly K, et al. The complete genome sequence of *Eubacterium limosum* SA11, a metabolically versatile rumen acetogen. *Stand in Genomic Sci* 2016;11:26.
 80. Gillis CC, Winter MG, Chanin RB, Zhu W, Spiga L, et al. Host-derived metabolites modulate transcription of salmonella genes involved in l-lactate utilization during gut colonization. *Infect Immun* 2019;87:e00773-18.
 81. Seeliger S, Janssen PH, Schink B. Energetics and kinetics of lactate fermentation to acetate and propionate via methylmalonyl-CoA or acrylyl-CoA. *FEMS Microbiol Lett* 2002;211:65–70.

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.