



Medium-Chain Fatty Acid Synthesis by “*Candidatus Weimeria bifida*” gen. nov., sp. nov., and “*Candidatus Pseudoramibacter fermentans*” sp. nov.

 Matthew J. Scarborough,^{a,b} Kevin S. Myers,^a  Timothy J. Donohue,^{a,c}  Daniel R. Noguera^{a,b}

^aThe Great Lakes Bioenergy Research Center, University of Wisconsin—Madison, Madison, Wisconsin, USA

^bThe Department of Civil and Environmental Engineering, University of Wisconsin—Madison, Madison, Wisconsin, USA

^cThe Department of Bacteriology, University of Wisconsin—Madison, Madison, Wisconsin, USA

ABSTRACT Chain elongation is emerging as a bioprocess to produce valuable medium-chain fatty acids (MCFA; 6 to 8 carbons in length) from organic waste streams by harnessing the metabolism of anaerobic microbiomes. Although our understanding of chain elongation physiology is still evolving, the reverse β -oxidation pathway has been identified as a key metabolic function to elongate the intermediate products of fermentation to MCFA. Here, we describe two uncultured chain-elongating microorganisms that were enriched in an anaerobic microbiome transforming the residues from a lignocellulosic biorefining process. Based on a multi-omic analysis, we describe “*Candidatus Weimeria bifida*” gen. nov., sp. nov., and “*Candidatus Pseudoramibacter fermentans*” sp. nov., both predicted to produce MCFA but using different substrates. The analysis of a time series metatranscriptomic data set suggests that “*Ca. Weimeria bifida*” is an effective xylose utilizer since both the pentose phosphate pathway and the bifid shunt are active. Furthermore, the metatranscriptomic data suggest that energy conservation during MCFA production in this organism is essential and occurs via the creation of an ion motive force using both the RNF complex and an energy-conserving hydrogenase. For “*Ca. Pseudoramibacter fermentans*,” predicted to produce MCFA from lactate, the metatranscriptomic analysis reveals the activity of an electron-confurcating lactate dehydrogenase, energy conservation via the RNF complex, H_2 production for redox balance, and glycerol utilization. A thermodynamic analysis also suggests the possibility of glycerol being a substrate for MCFA production by “*Ca. Pseudoramibacter fermentans*.” In total, this work reveals unknown characteristics of MCFA production in two novel organisms.

IMPORTANCE Chain elongation by medium-chain fatty acid (MCFA)-producing microbiomes offers an opportunity to produce valuable chemicals from organic streams that would otherwise be considered waste. However, the physiology and energetics of chain elongation are only beginning to be studied, and many of these organisms remain uncultured. We analyzed MCFA production by two uncultured organisms that were identified as the main MCFA producers in a microbial community enriched from an anaerobic digester; this characterization, which is based on meta-multi-omic analysis, complements the knowledge that has been acquired from pure-culture studies. The analysis revealed previously unreported features of the metabolism of MCFA-producing organisms.

KEYWORDS Ech complex, *Firmicutes*, MCFA, *Pseudoramibacter*, RNF complex, “*Ca. Weimeria*,” chain elongation, hydrogenase

Citation Scarborough MJ, Myers KS, Donohue TJ, Noguera DR. 2020. Medium-chain fatty acid synthesis by “*Candidatus Weimeria bifida*” gen. nov., sp. nov., and “*Candidatus Pseudoramibacter fermentans*” sp. nov. *Appl Environ Microbiol* 86:e02242-19. <https://doi.org/10.1128/AEM.02242-19>.

Editor Harold L. Drake, University of Bayreuth

Copyright © 2020 Scarborough et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Daniel R. Noguera, dnoguera@wisc.edu.

Received 30 September 2019

Accepted 25 October 2019

Accepted manuscript posted online 8 November 2019

Published 21 January 2020

Chain elongation has been proposed as a microbial process to produce valuable chemicals from complex waste feedstocks (1). This bioprocess relies on the combined metabolism of an anaerobic microbiome to hydrolyze complex organic substrates, ferment the hydrolyzed products to small organic intermediates (2 to 3 carbon molecules), and elongate these fermentation products to medium-chain fatty acids (MCFA; 6 to 8 carbon molecules) through reverse β -oxidation (1). MCFA are an attractive product due to their high value, relatively low solubility in water, and potential to offset fossil fuel demands for petrochemicals and other products (2, 3). Bioreactors performing chain elongation also provide model systems for studying the metabolic contributions of uncultured organisms to this process. To date, all but one of the known chain-elongating bacteria belong to the *Clostridia* class within the *Firmicutes* phylum (1). These chain-elongating bacteria primarily use lactate (4, 5), ethanol (6), or carbohydrates (7) to drive MCFA production.

We recently described a chain-elongating microbiome that produced sufficient hexanoic and octanoic acids from lignocellulosic biorefining residues to reduce the minimum selling price of ethanol produced in a biorefinery (2). Using a combination of metagenomics and metatranscriptomics, we characterized this microbiome as having a small set of high-abundance organisms (8), with two populations within the *Clostridia* class performing chain elongation. One high-abundance member of this microbiome (LCO1) (8) belonged to the *Lachnospiraceae* family and was predicted to produce MCFA from xylose and other pentoses; the second one (EUB1) (8) corresponded to the *Eubacteriaceae* family and was predicted to produce MCFA from lactate.

Here, we combined multi-omic approaches to further analyze the genomic and metabolic features of these two predicted MCFA-producing organisms, which remain uncultured. A time series gene expression analysis showed that transcripts encoding proteins predicted to be involved in reverse β -oxidation are among the most abundant transcripts after the lignocellulosic biorefinery residues are fed to the microbial community. Our analysis also reveals that both organisms contain transcripts that encode a proton-translocating energy conserving hydrogenase, suggesting contributions of previously unreported metabolic networks to MCFA production. Based on these new results, we conclude that LCO1 represents a novel genus within the *Lachnospiraceae* family and propose the name of "*Candidatus Weimeria bifida*" gen. nov., sp. nov. Our data also predict that EUB1 represents a new species within the *Pseudoramibacter* genus, and we propose the name "*Candidatus Pseudoramibacter fermentans*" sp. nov. to represent this new species. In total, the objective of this work is to provide a greater understanding of MCFA-producing organisms in the context of a chain-elongating microbiome.

RESULTS AND DISCUSSION

Refinement of MAGs. We previously reported the construction of draft metagenome-assembled genomes (MAGs) from an MCFA-producing microbiome fed with lignocellulosic biorefinery residues, in which LCO1 and EUB1 represented the abundant chain-elongating microorganisms (8). These draft MAGs were constructed using DNA samples from the first 120 days of reactor operation. To improve the quality of these MAGs, we obtained Illumina and PacBio sequencing reads from the same microbiome at additional times during a 378-day operational period. We coassembled 244 million Illumina Hi-seq (2 by 250 bp) reads from five time points (days 96, 120, 168, 252, and 378) (see Fig. S1 in the supplemental material) into 24,000 contigs. Contigs were binned into MAGs; the MAGs with relative abundance greater than 1% were then gap filled with PacBio reads from the day 378 sample. This analysis resulted in an overall improvement in MAG quality with respect to completeness, contamination, and number of scaffolds (Table S1). The two most abundant MAGs derived from this analysis (Table 1), LCO1.1 and EUB1.1 (accounting for ~80% of the recovered DNA sequences) (Data File S1), provide improved predictions for the genetic makeup of the organisms previously denoted as LCO1 and EUB1 (8), respectively.

TABLE 1 Summary of metagenome-assembled genomes from an MCFA-producing bioreactor

Organism	Previous name ^a	Relative abundance (%) ^b	Completeness (%) ^c	Contamination (%) ^c	Genome size (Mbp) ^c	No. of scaffolds ^c
"Ca. Weimeria bifida"	LCO1	75.3	96.9 (95.4)	0.5 (0.0)	2.39 (2.10)	10 (44)
"Ca. Pseudoramibacter fermentans"	EUB1	4.7	99.2 (97.8)	0.2 (0.2)	2.29 (2.00)	29 (35)

^aName reported in Scarborough et al. (8).

^bRelative abundance based on DNA read mapping normalized to genome size for the day 252 sample. This date corresponds to the date the metatranscriptomic experiment was initiated.

^cNumbers in parentheses indicate values for metagenome-assembled genomes reported in Scarborough et al. (8).

Phylogenetic classification of LCO1.1 and EUB1.1. The Genome Taxonomy Database (GTDB) tool kit (9) was used to provide a taxonomic classification of LCO1.1 and EUB1.1 (Fig. 1). The reconstructed LCO1.1 genome, which accounted for ~75% of the mapped DNA reads (Table 1), clusters with a group of genomes identified in the GTDB

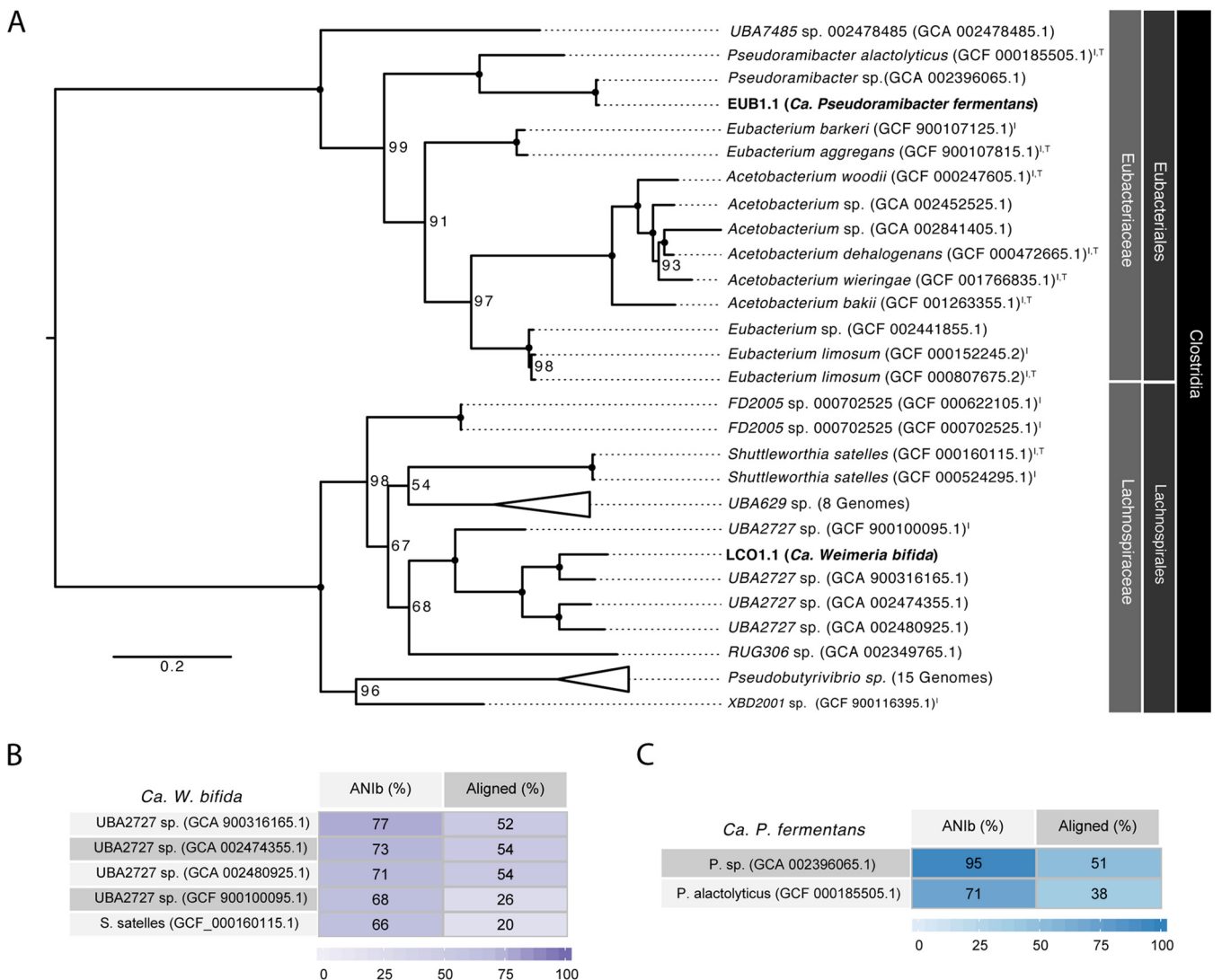


FIG 1 Phylogenetic analysis of two metagenome-assembled genomes (shown in bold text) predicted to perform chain elongation in an anaerobic microbiome converting lignocellulosic residues to medium-chain fatty acids. (A) A phylogenetic tree was created based on concatenated amino acid sequences for 120 single-copy marker genes from the GTDB using RAxML. Bootstrap support values are indicated at the nodes. Closed circles indicate bootstrap support values of 100. Organisms that have representative isolates are indicated by an "I" superscript, and organisms representing NCBI type strains are indicated with a "T" superscript. Genome identifiers are provided in parentheses. (B) BLAST average nucleotide identity (ANiB) comparison for LCO1.1 ("Ca. Weimeria bifida") to related genomes. (C) ANiB comparison for EUB1.1 ("Ca. Pseudoramibacter fermentans") to related genomes.

with the name UBA2727 within the *Lachnospiraceae* family, in the order *Clostridia*, phylum *Firmicutes*. This cluster contains three MAGs obtained from rumen samples (GTDB accession no. GCA_002474355.1, GCA_002480925.1, and GCA_900316165.1) (9), as well as a genome of an isolate from the Hungate 1000 project (named *Lachnospiraceae* bacterium C10 in the NCBI database; GTDB accession no. GCF_900100095.1) (10) (Data File S2). A comparative analysis of LCO1.1 with the four representatives of the UBA2727 cluster (Fig. 1B) shows average nucleotide identity (ANI) values greater than 70% with the three MAGs and 68% with the Hungate 1000 project isolate. The most closely related type strain is *Shuttleworthia satelles*, an organism isolated from the human oral cavity and described as a carbohydrate consumer and butyrate producer (11). Given the low ANI value (66%) obtained when the LCO1.1 genome is compared to that of *S. satelles*, we propose here the name “*Candidatus Weimeria*” to define the organisms comprised of the UBA2727 cluster plus the LCO1.1 MAG. Further, we propose the epithet “*bifida*” to describe a species within this new genus, represented by LCO1.1.

The EUB1.1 MAG, which accounted for ~5% of the mapped DNA reads (Table 1), clustered in the *Pseudoramibacter* genus, within the *Eubacteriaceae* family, also in the order *Clostridia*. This genus currently contains only one known species, defined by the isolate *Pseudoramibacter alactolyticus* (originally named *Ramibacterium alactolyticum* [12]), which was recovered from a human lung abscess and has been shown to produce short- and medium-chain fatty acids (12). The ANI of EUB1.1 to *P. alactolyticus* is 71%, with an alignment factor of 38% (Fig. 1C). The GTDB also contained a related MAG that was obtained from a sheep rumen metagenomic analysis (9) and identified as *Pseudoramibacter* sp. (GCA_002396065.1). This MAG had an ANI of 95% and an alignment factor of 51% to EUB1.1 (Fig. 1C). Its completion was estimated to be only 57%, which likely contributes to its low alignment factor to EUB1.1. A comparative analysis of these three *Pseudoramibacter* genomes (Data File S3) shows that among their similarities, they all contain genes typically associated with reverse β -oxidation, consistent with their predicted ability to produce MCFA. Based on the clustering of EUB1.1 within the *Pseudoramibacter* genus, on the obtained ANI values, and on the comparative analysis that follows, we propose here a new species within the *Pseudoramibacter* genus, represented by EUB1.1 and *Pseudoramibacter* sp., and identified with the epithet “*fermentans*.”

Multi-omic analysis of “*Ca. Weimeria bifida*” and “*Ca. Pseudoramibacter fermentans*.” Previous analyses showed that “*Ca. Weimeria bifida*” and “*Ca. Pseudoramibacter fermentans*” expressed genes related to MCFA production in the bioreactor at a single time point collected after 96 days of bioreactor operation (8). To further evaluate MCFA production by “*Ca. Weimeria bifida*” and “*Ca. Pseudoramibacter fermentans*,” we analyzed transcript abundance for 36 h after the bioreactor received a transient load of lignocellulosic biorefinery residue (Fig. S1 and Data Files S4 and S5). This experiment was performed after the reactor had been operating continuously for 252 days. Concurrent analysis of the medium over this time period showed that xylose and glycerol were consumed whereas lactate transiently accumulated in the reactor (Fig. S2). We used this multi-omic analysis to investigate substrate utilization, the enzymes involved in converting substrates into intermediates and end products, the potential for MCFA production via the reverse β -oxidation pathway, and the predicted energy-conserving features of the predominant MAGs in this anaerobic microbiome.

(i) Chain elongation in “*Ca. Weimeria bifida*.” Four enzymes are known or predicted to comprise the reverse β -oxidation cycle (Fig. 2). In the first step, acyl-coenzyme A (CoA) acetyltransferase (ACAT) condenses an acetyl-CoA with an acyl-CoA; the product of this reaction is reduced by 3-hydroxy-acyl-CoA dehydrogenase (HAD), followed by dehydration by 2-enoyl-CoA with enoyl-CoA dehydratase (ECoAH) and reduction by an acyl-CoA dehydrogenase (ACD) to form an elongated acyl-CoA. In some organisms, the last dehydrogenation reaction is catalyzed by an electron-bifurcating energy-conserving enzyme whereby the enoyl-CoA reduction with NADH is

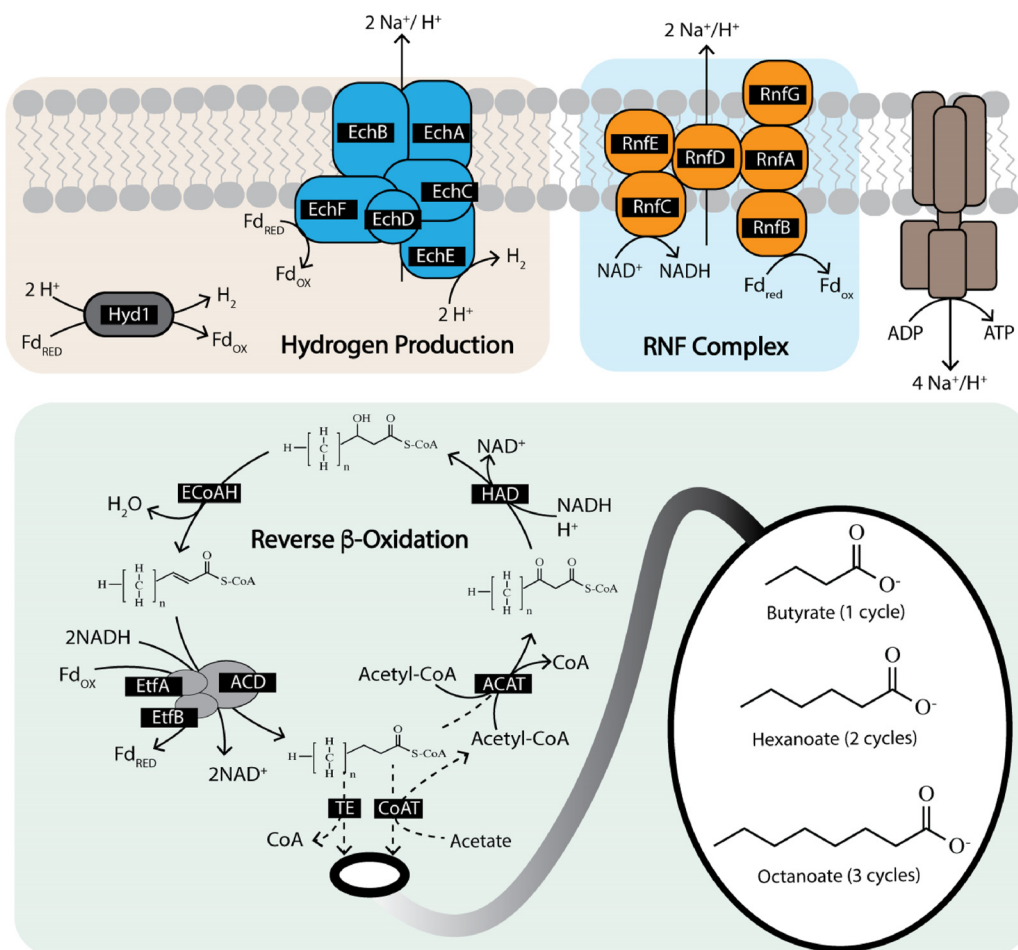


FIG 2 Metabolic pathways involved in chain elongation. Reverse β -oxidation is a four-step process using acyl-CoA acetyltransferase (ACAT), 3-hydroxy-acyl-CoA dehydrogenase (HAD), enoyl-CoA hydratase (ECoAH), and acyl-CoA dehydrogenase (ACD). The reduction of enoyl-CoA with NADH can be combined with the reduction of ferredoxin through an electron-bifurcating acyl-CoA dehydrogenase complex containing EtfA and EtfB. The terminal enzyme of reverse β -oxidation can be either a CoA transferase (CoAT) or thioesterase (TE). Reverse β -oxidation is coupled with proton-translocating enzymes to generate ATP with reduced ferredoxin. This figure is partly adapted from previous work (8).

paired with the reduction of ferredoxin through an ACD complex containing the electron transfer flavoproteins EtfA and EtfB (13). The "Ca. Weimeria bifida" genome has a gene cluster encoding ACAT, HAD, ACD, EtfA, and EtfB (Fig. 3A), while a gene predicted to encode ECoAH is located on a different region of the genome. The abundance of transcripts encoding reverse β -oxidation enzymes was at or above the 90th percentile in all samples analyzed (Fig. 4A). Pairwise comparisons of reverse β -oxidation transcript abundance show high correlations for genes in the cluster encoding ACAT, HAD, ACD, EtfA, and EtfB over the course of this analysis but low correlations with the ECoAH transcript (Fig. S3), suggesting that this later gene is not coexpressed with those predicted to be involved in the reverse β -oxidation pathway of "Ca. Weimeria bifida."

Producing reduced ferredoxin during reverse β -oxidation constitutes a potential energy-conserving process (Fig. 2) because its subsequent oxidation can be used to create an ion motive force by either the RNF complex (RnfABCDEG), as proposed for the MCFA-producing *Clostridium kluyveri* (14), or the energy-conserving hydrogenase complex Ech, as described for *Clostridium thermocellum* (15). In "Ca. Weimeria bifida", the genes encoding the subunits of the RNF complex are found in a single operon arranged as *rnfCDGEAB*, which is consistent with the gene organization in *C. kluyveri*, *Clostridium difficile*, and other *Clostridia* species (16). Genes encoding individual subunits of the Ech

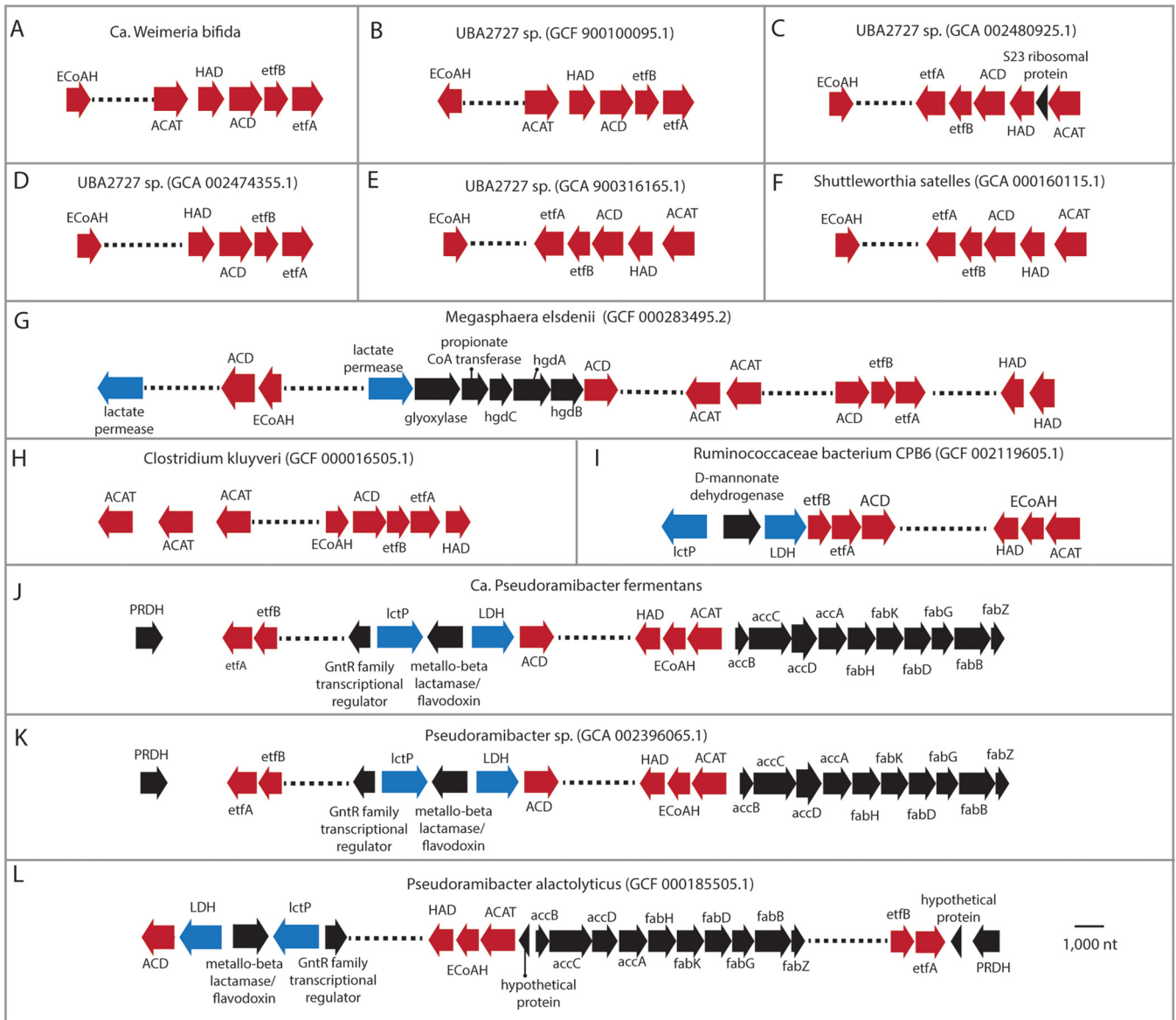


FIG 3 Organization of selected genes in organisms related to “*Ca. Weimeria bifida*,” “*Ca. Pseudoramibacter fermentans*,” and known chain-elongating organisms. Reverse β -oxidation genes are shown in red, and genes involved in lactate utilization, if present, are shown in blue. Reverse β -oxidation genes include acyl-CoA acetyltransferase (ACAT), 3-hydroxy-acyl-CoA dehydrogenase (HAD), enoyl-CoA hydratase (ECoAH), acyl-CoA dehydrogenase (ACD), electron transfer flavoprotein A (*etfA*), and electron transfer flavoprotein B (*etfB*). Lactate utilization genes include lactate permease (*lctP*) and lactate dehydrogenase (LDH). Other genes included in the figure are prephenate dehydrogenase (PRDH) (J, K, and L), 2-hydroxyglutaryl-CoA dehydratase (*hgd*) involved in propionate production in *Megasphaera elsdenii* (G), and acetyl-CoA carboxylase (*acc*) and fatty acid biosynthesis (*fab*) genes when they are adjacent to reverse β -oxidation genes (J, K, and L). nt, nucleotides.

complex are also found in a single operon in “*Ca. Weimeria bifida*” (*echABCDEF*) (Data File S2) and are adjacent to genes encoding HypB, HypC, and HypD, which are involved in hydrogenase assembly (17, 18). Predicted *echABCDEF* gene clusters were not found in genomes of the other known chain-elongating organisms analyzed in this study (Data File S6), but they have been found in a diverse range of other bacteria and archaea (19).

Time series transcriptomic analysis showed that “*Ca. Weimeria bifida*” produced high levels of transcripts encoding the RNF and the Ech complexes when MCFA were produced after addition of lignocellulosic biorefinery residues (Fig. 4). Transcripts encoding subunits of the RNF complex (RnfABCDEFGF) were present above the 90th percentile at several time points after addition of lignocellulosic biorefinery residues

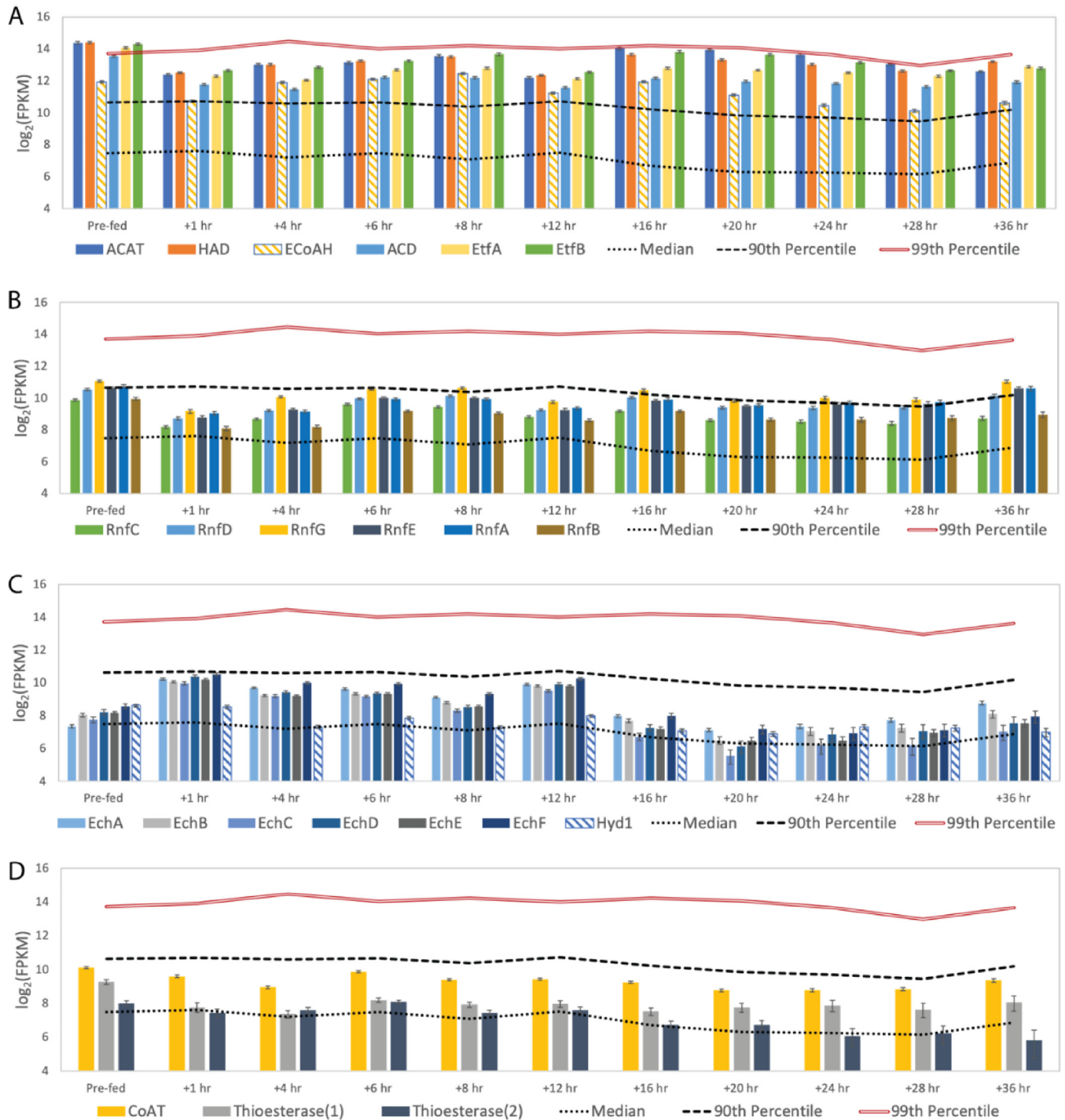


FIG 4 Transcript abundance for predicted chain elongation transcripts by "Ca. Weimeria bifida." "Prefed" represents the time point prior to feeding the reactor. Error bars show 95% confidence levels determined with CuffLinks. Transcript abundance data are shown for transcripts encoding enzymes used for reverse β -oxidation (A), the RNF complex (B), hydrogenases (C), and terminal enzymes of reverse β -oxidation (D). The roles of the designated enzymes in chain elongation are provided in Fig. 2. FPKM, fragments per kilobase per million.

and found to be at or above median expression levels throughout the study (Fig. 4B). Transcripts for genes encoding subunits of the Ech hydrogenase complex followed a different pattern, with high-level abundance only for the first 12 h after addition of lignocellulosic biorefinery residues (Fig. 4C). We also found that transcripts of Ech complex genes were more abundant than those encoding a putative periplasmic

ferredoxin hydrogenase (Hyd1), the only other hydrogenase predicted to be present in the "*Ca. Weimeria bifida*" genome. Thus, the multi-omic data support a role for both the RNF and Ech complexes during MCFA production, likely by conserving energy via generation of an ion motive force (Fig. 2).

The reverse β -oxidation cycle is also predicted to require either a CoA transferase (CoAT) or a thioesterase to remove the CoA from the terminal acyl-CoA molecule, thereby releasing the corresponding acid (Fig. 2). During the course of this experiment, "*Ca. Weimeria bifida*" expressed genes encoding one predicted CoAT and two predicted thioesterases. Transcripts of all three genes were at or above median levels throughout the time course of this experiment (Fig. 4D). The genes encoding putative thioesterase enzymes and CoAT are not located near other predicted reverse β -oxidation genes in the "*Ca. Weimeria bifida*" genome (Data File S2), but the abundance of transcripts encoding one of the predicted thioesterase enzymes (TE.1) (Fig. S3) is more closely correlated with other reverse β -oxidation pathway transcripts than the other predicted thioesterase (TE.2) (Fig. S3). TE.1 was annotated as a thioesterase superfamily protein (Data File S2), and a Basic Local Alignment Search Tool (BLAST) analysis of amino acid sequences (Data File S7) places it in the hotdog fold family of thioesterase proteins, some of which have been shown to cleave CoA from medium- and long-chain acyl-CoA molecules (20). Thus, this analysis indicates that CoAT and TE.1 may participate in the terminal step of reverse β -oxidation in "*Ca. Weimeria bifida*."

(ii) "*Ca. Weimeria bifida*" is predicted to use multiple routes to consume xylose as a source of carbon for MCFA production. Previous studies (8) and the revised genome provided in this work predict that "*Ca. Weimeria bifida*" can metabolize xylose and other pentoses (Fig. 5A) as the main source of energy and carbon when grown in an MCFA-producing microbiome (8). In this time series experiment, we found that transcripts encoding an ABC transporter predicted to transport multiple sugars were present at or above the 99th percentile at several time points (Fig. 5B), suggesting a role for this protein in sugar uptake.

To investigate potential routes for sugar utilization by "*Ca. Weimeria bifida*," we compared transcript abundances of genes encoding enzymes predicted to function in the pentose phosphate and phosphoketolase pathways after the addition of lignocellulosic biorefinery residues (Fig. 5B). Patterns of transcript abundance indicate that genes encoding enzymes for both pathways are expressed above median expression levels, suggesting that both pathways are used for pentose utilization and its subsequent conversion to intermediates that then enter the reverse β -oxidation pathway to produce MCFA.

(iii) Predicted conservation of metabolic functions between "*Ca. Weimeria bifida*" and related organisms. We found that all of the UBA2727 genomes, *Shuttleworthia satelles*, and the LCO1.1 genome contained genes needed for the reverse β -oxidation pathway (Data File S2). In the UBA2727, *S. satelles*, and "*Ca. Weimeria bifida*" genomes, the genes encoding ACAT, HAD, ACD, EtfA, and EtfB are clustered (Fig. 3A to F), whereas ECoAH is located on a different region of the genome. The presence of genes encoding each of these enzymes and their transcript abundance patterns in "*Ca. Weimeria bifida*" during MCFA production suggest that reverse β -oxidation and the electron-bifurcating ACD reaction are conserved and key features of chain elongation in each of these organisms. Genome comparisons revealed that the colocation of genes encoding ACD, EtfA, and EtfB is a common, although not essential, topology among known MCFA-producing organisms. Examples of other MCFA producers with this gene arrangement include *Megasphaera elsdenii* (Fig. 3G), *C. kluveri* (Fig. 3H), and the bacterium *Ruminococcaceae* strain CPB6 (Fig. 3I).

While two routes of pentose utilization (the pentose phosphate and phosphoketolase pathways) are predicted to be present in the "*Ca. Weimeria bifida*" genome, none of the UBA2727 genomes nor the *S. satelles* genome contains a predicted phosphoketolase gene. This indicates that the use of multiple pentose consumption pathways may be a unique feature of "*Ca. Weimeria bifida*" not found in other members of the UBA2727 cluster or the nearest type strain. All of the UBA2727 genomes and the

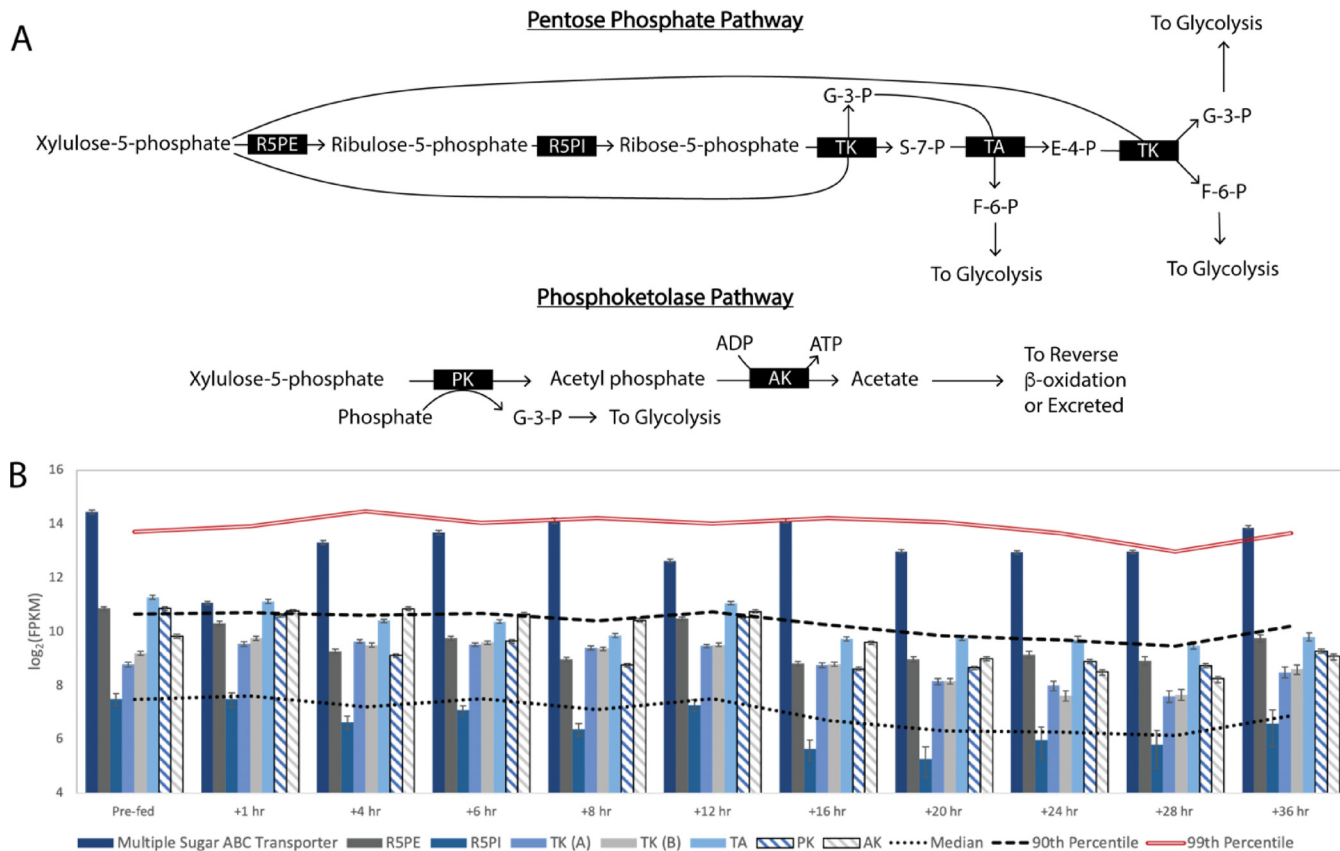


FIG 5 Predicted routes for xylose utilization by "Ca. Weimeria bifida." (A) After xylose activation to xylulose-5-phosphate, "Ca. Weimeria bifida" can consume pentoses via the pentose phosphate pathway or the phosphoketolase pathway. Key enzymes involved in xylose consumption include ribulose 5-phosphate epimerase (R5PE), ribose-5-phosphate isomerase (R5PI), transketolase (TK), transaldolase (TA), transketolase (TK), phosphoketolase (PK), and acetate kinase (AK). (B) Transcript abundance for 36 h after providing the bioreactor with conversion residue. "Prefed" represents the time point prior to feeding the reactor. Error bars show 95% confidence levels determined with CuffLinks.

"Ca. Weimeria bifida" genome contain genes for acetate production, including phosphotransacetylase and acetate kinase, indicating that all of the species in this genus may produce acetate as a product of anaerobic sugar metabolism.

The putative phosphoketolase from "Ca. Weimeria bifida" is similar to enzymes predicted to be present in other firmicutes and betaproteobacteria, including three species of *Megasphaera* (Fig. S4), a genus containing known MCFA producers. The phosphoketolase pathway (Fig. 5A), termed the "bifid shunt" (21) in a *Bifidobacterium* (22), provides an alternative to glycolysis and the pentose phosphate pathway for sugar utilization. In this pathway, the phosphoketolase enzyme splits xylulose-5-phosphate into glyceraldehyde-3-phosphate and acetyl-phosphate (23). Energy can then be conserved through the phosphorylation of ADP and production of acetate by acetate kinase (Fig. 5A). Overall, the phosphoketolase pathway can produce more ATP than the pentose phosphate pathway per mole of xylose (Fig. S5), and it directs carbon to acetate in addition to producing glycolysis intermediates (Fig. 5A).

(iv) Chain elongation by "Ca. Pseudoramibacter fermentans." We previously predicted that "Ca. Pseudoramibacter fermentans" consumed lactate and produced MCFA when this microbiome was supplied with lignocellulosic biorefinery residues (8). In this study, we found that transcripts encoding enzymes predicted to function in the reverse β-oxidation pathway were among the most abundant in "Ca. Pseudoramibacter fermentans" after the addition of lignocellulosic biorefinery residues (Fig. 6A). However, in the revised "Ca. Pseudoramibacter fermentans" genome we found that the gene encoding the putative ECoAH protein is located near genes predicted to encode enzymes in the reverse β-oxidation cycle (Fig. 3), unlike the genomes of "Ca. Weimeria

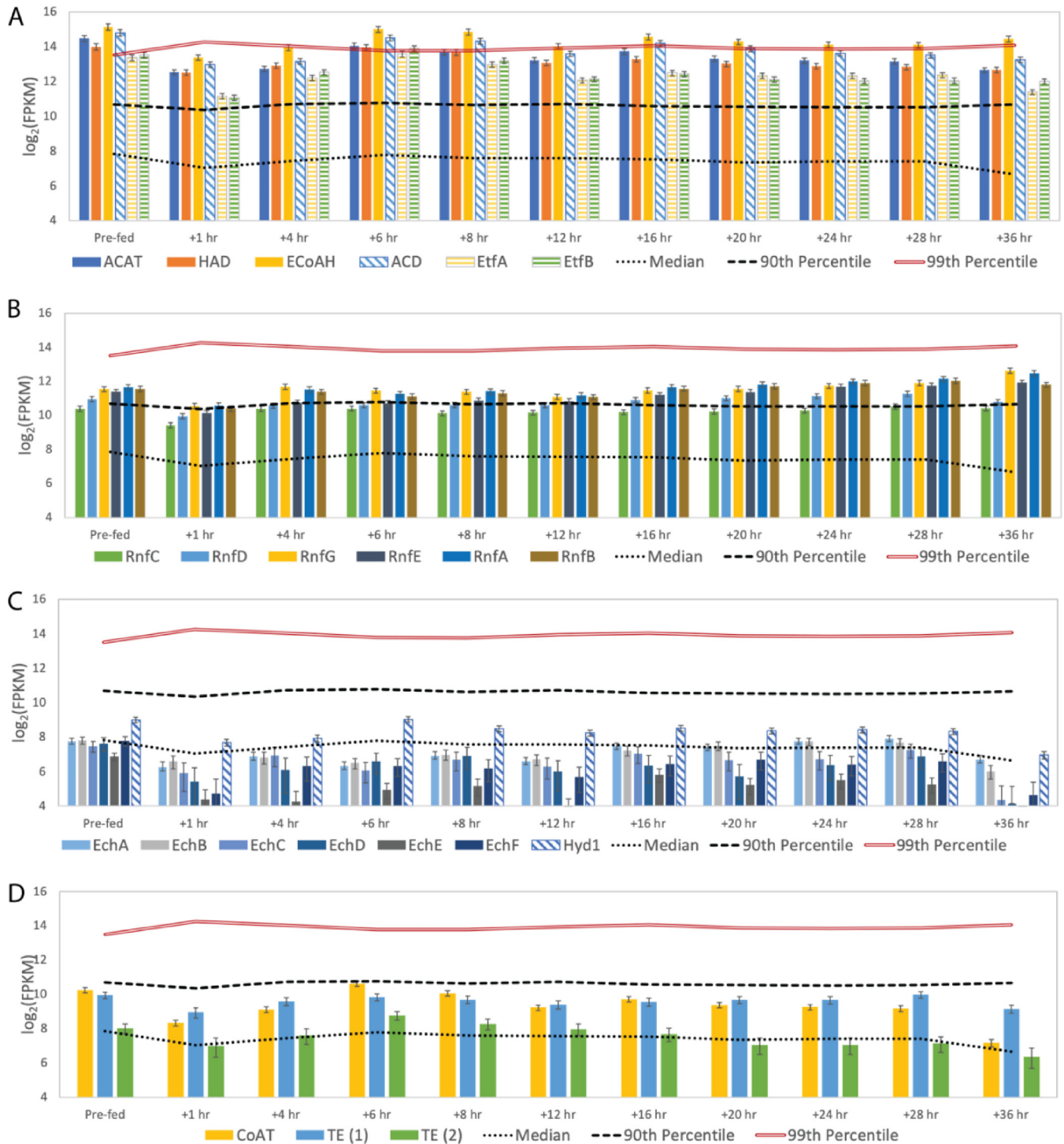


FIG 6 Transcript abundance for chain elongation transcripts by “*Ca. Pseudoramibacter fermentans*.” “Prefed” represents the time point prior to feeding the reactor. Error bars show 95% confidence levels determined with CuffLinks. Transcript abundance data are shown for genes encoding enzymes used for reverse β -oxidation (A), the RNF complex (B), hydrogenases (C), and terminal enzymes of reverse β -oxidation (D). The roles of enzymes in chain elongation and enzyme definitions are shown in Fig. 2.

bifida” and related organisms. In addition, we found that the abundance of the transcript encoding the EcoAH protein correlates well with abundances of “*Ca. Pseudoramibacter fermentans*” genes predicted to encode other enzymes in the reverse β -oxidation cycle (Fig. S6). In “*Ca. Pseudoramibacter fermentans*,” the genes encoding the predicted EtfA and EtfB proteins are not located next to those encoding the ACD

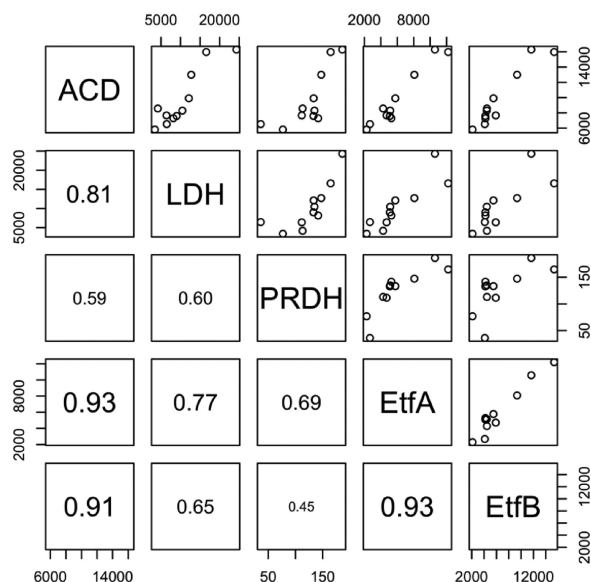


FIG 7 Pairwise linear correlation analyses for expression of “*Ca. Pseudoramibacter fermentans*” genes encoding lactate dehydrogenase (LDH), acetyl-CoA dehydrogenase (ACD), prephenate dehydrogenase (PRDH), electron transfer flavoprotein A (EtfA), and electron transfer flavoprotein B (EtfB). Values on the x and y axes represent the number of fragments per kilobase per million. Numbers within the plots are coefficients of determination.

enzyme. Instead, the *etfAB* genes are in a cluster with a gene that encodes a homologue of a putative prephenate dehydrogenase (PRDH), an enzyme that catalyzes oxidative decarboxylation in the shikimate pathway for tyrosine biosynthesis (24) (Fig. 3). Pairwise gene expression analysis of transcripts encoding the ACD, PRDH, EtfA, and EtfB proteins (Fig. 7) showed that ACD had strong correlations with EtfA ($r^2 = 0.93$) and EtfB ($r^2 = 0.91$), suggesting coregulation of ACD and the electron transport flavoproteins and supporting a role for an electron-bifurcating ACD in the reverse β -oxidation cycle of “*Ca. Pseudoramibacter fermentans*,” which we also predict for “*Ca. Weimeria bifida*.”

The revised genome sequence of “*Ca. Pseudoramibacter fermentans*” also contained complete sets of genes for subunits of the RNF and the Ech complexes, a characteristic also shared with “*Ca. Weimeria bifida*.” The transcript abundances of “*Ca. Pseudoramibacter fermentans*” genes encoding subunits of the RNF complex were at or near the 90th percentile throughout the time series analyzed in this experiment (Fig. 6B), whereas the transcript abundances for the genes predicted to encode the “*Ca. Pseudoramibacter fermentans*” Ech hydrogenase were below the median at most time points after the addition of lignocellulosic biorefinery residues (Fig. 6C). Furthermore, the other hydrogenase predicted to be present in “*Ca. Pseudoramibacter fermentans*,” Hyd1, had transcript abundances that were higher than those of genes encoding the Ech hydrogenase complex. These observations suggest that energy conservation in “*Ca. Pseudoramibacter fermentans*” primarily occurs via generation of an ion motive force by the RNF complex. However, hydrogen production via Hyd1 could be important to maintain redox balance during chain elongation by “*Ca. Pseudoramibacter fermentans*.”

Hydrolysis of the elongated acyl-CoA molecule relies on a CoAT or a thioesterase (Fig. 2). The “*Ca. Pseudoramibacter fermentans*” genome encodes a CoAT and two predicted thioesterases, each of which was expressed at or above median level compared to levels of other transcripts analyzed during this experiment (Fig. 6D). The abundances of the transcripts encoding CoAT and one of the predicted thioesterases (TE.2) (Fig. S6) have higher correlations with those of predicted enzymes in the reverse β -oxidation pathway than the transcripts of the other predicted thioesterase (TE.1) (Fig. S6). As was the case with TE.1 in “*Ca. Weimeria bifida*,” an amino acid sequence analysis

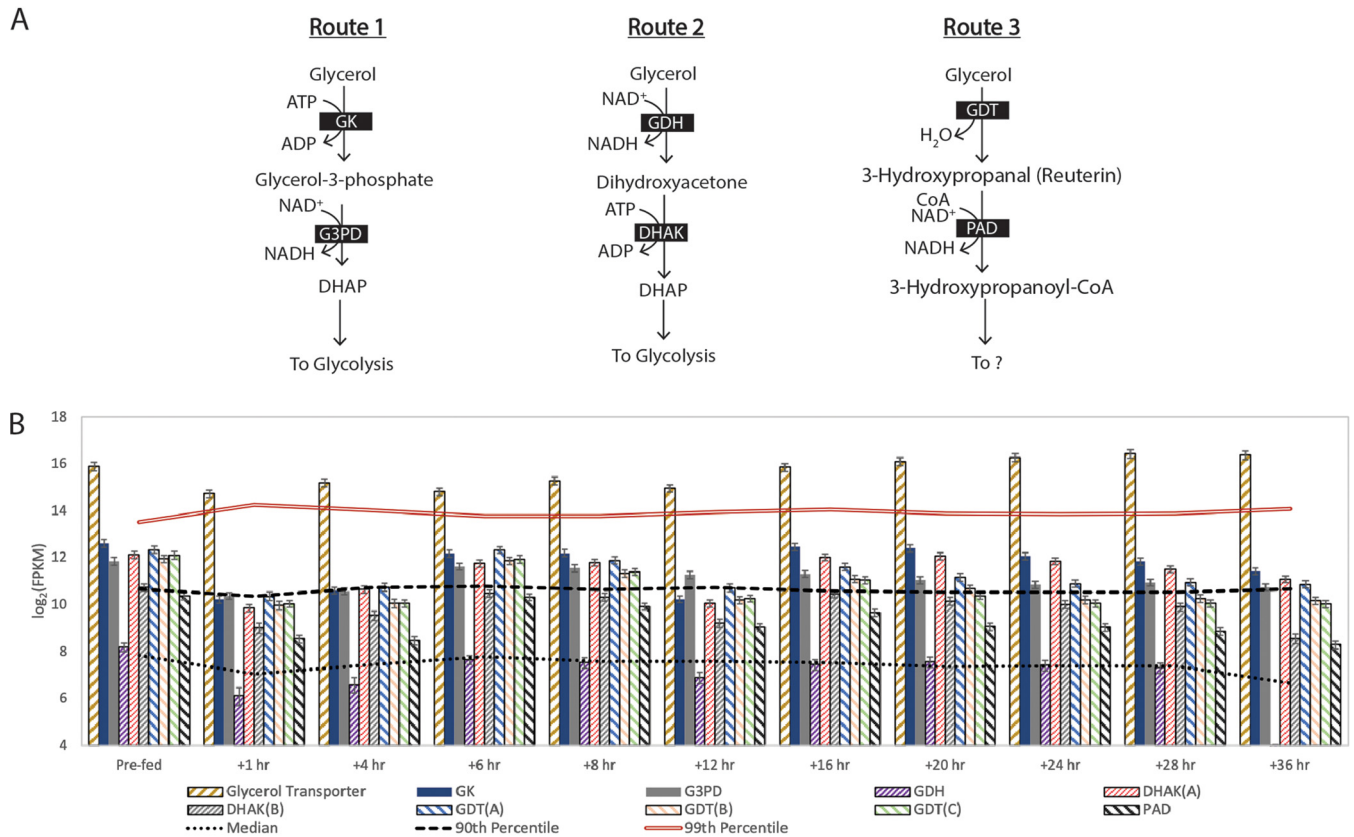


FIG 8 Glycerol utilization by “*Ca. Pseudoramibacter fermentans*.” (A) “*Ca. Pseudoramibacter fermentans*” contains genes for three glycerol utilization routes. Key enzymes in these pathways include glycerol kinase (GK), glycerol-3-phosphate dehydrogenase (G3PD), glycerol dehydrogenase (GDH), dihydroxyacetone kinase (DHAK), glycerol dehydratase (GDT), and propionaldehyde dehydrogenase (PAD). (B) Transcript abundance for glycerol transport and utilization. Subunits of a single enzyme complex are indicated in parentheses. “Prefed” represents the time point prior to feeding the reactor. Error bars show 95% confidence levels determined with CuffLinks.

of the TE.2 from “*Ca. Pseudoramibacter fermentans*” showed that it belongs to the hotdog fold superfamily of thioesterases. Specifically, the TE.2 “*Ca. Pseudoramibacter fermentans*” thioesterase is predicted to be a Paal family enzyme, which includes the medium-chain acyl-CoA thioesterase II of *Escherichia coli* (25). Thus, the transcriptomic analysis suggests that both CoAT and TE.2 may participate in acyl chain release during MCFA synthesis in “*Ca. Pseudoramibacter fermentans*.”

(v) “*Ca. Pseudoramibacter fermentans*” is predicted to use multiple routes for glycerol metabolism. Glycerol is known to be a significant carbon source in the lignocellulosic biorefinery residues used in this and earlier studies (8). Analysis of the reactor medium showed that all of the glycerol was removed within the first 6 h (Fig. S2), suggesting that it is a favored carbon source for one or more of the microbes in this microbiome. The gene most highly expressed by “*Ca. Pseudoramibacter fermentans*” at several time points after the addition of lignocellulosic biorefinery residues encodes a predicted glycerol transporter (Fig. 8B), suggesting that glycerol is rapidly transported by this organism.

To assess how “*Ca. Pseudoramibacter fermentans*” metabolizes glycerol, we monitored transcript abundances of genes predicted to be involved in this process. The revised “*Ca. Pseudoramibacter fermentans*” genome also predicts that this organism contains enzymes to metabolize glycerol, and the time series transcriptomics data showed that this organism expressed genes for three putative glycerol conversion pathways (Fig. 8). The first route predicted to be active in “*Ca. Pseudoramibacter fermentans*” (Fig. 8A, route 1) uses an ATP-dependent glycerol kinase, in which the resultant glycerol-phosphate is oxidized to produce dihydroxy-acetone phosphate

(DHAP). The second predicted route in “*Ca. Pseudoramibacter fermentans*” (Fig. 8A, route 2) uses glycerol dehydrogenase to oxidize glycerol to dihydroxyacetone before it is phosphorylated to DHAP. In a third predicted route in “*Ca. Pseudoramibacter fermentans*” (Fig. 8A, route 3), glycerol is converted to 3-hydroxypropanal (reuterin), a compound proposed to inhibit growth of other microbes by inducing oxidative stress (26). All of the above predicted glycerol utilization genes, with the exception of glycerol dehydrogenase, were expressed above median levels throughout the 36-h time series (Fig. 8B). Thus, it is possible that multiple, or even all three, pathways (Fig. 8A) play a role in glycerol metabolism in “*Ca. Pseudoramibacter fermentans*.”

In addition to expressing genes encoding pathways for multiple glycerol utilization routes, “*Ca. Pseudoramibacter fermentans*” contains a gene cluster encoding predicted polyhedral body proteins involved in propanediol utilization along with genes encoding the multiple subunits of a glycerol dehydratase (Fig. 9). Thus, it is possible that these predicted protein microcompartments may help protect “*Ca. Pseudoramibacter fermentans*” from a toxin like reuterin or from other toxins that are potential products of glycerol metabolism by one or more routes in this organism.

Overall, our data provide new evidence that “*Ca. Pseudoramibacter fermentans*” directs glycerol to central carbon metabolism and potentially produces reuterin, a toxic compound that may impact the growth and abundance of other members in this microbiome. One possible outcome of glycerol metabolism to intermediates that can be further transformed via glycolysis (Fig. 8) is the conversion of glycerol to MCFA. A free-energy analysis of glycerol conversion to butyrate, hexanoate, and octanoate shows that all of these reactions are energetically feasible when hydrogen is produced to maintain internal redox balance (Table 2).

(vi) “*Ca. Pseudoramibacter fermentans*” is predicted to use an electron-confering lactate dehydrogenase for MCFA production. We previously proposed that one member of this MCFA-producing microbiome consumes lactate produced by other community members (8). The “*Ca. Pseudoramibacter fermentans*” genome assembled in this study contains a gene cluster encoding a lactate transporter (lactate permease), lactate dehydrogenase (LDH), and ACD (Fig. 3J). All three of these genes are expressed above the 90th percentile at all time points after the addition of lignocellulosic biorefinery residue (Fig. 10), supporting their role in MCFA production from lactate.

The amino acid sequence of the predicted “*Ca. Pseudoramibacter fermentans*” LDH in this gene cluster is most closely related to predicted flavin adenine dinucleotide (FAD)-binding oxidoreductase amino acid sequences from several related organisms, including other *Firmicutes* and organisms in the *Fusobacterium* phylum (Data File S8). The “*Ca. Pseudoramibacter fermentans*” LDH clusters with amino acid sequences of proteins from organisms that consume lactate anaerobically (Fig. 11), including two bacteria known to convert lactate to MCFA, *Ruminococcaceae* CPB6 and *M. elsdenii* (5, 7), and with the LDH from *Acetobacterium woodii* that uses electron confurcation to couple lactate and ferredoxin oxidation with NAD⁺ reduction to overcome the thermodynamic bottleneck of lactate oxidation (27). Pairwise correlation analyses of transcript abundance (Fig. 7) for the “*Ca. Pseudoramibacter fermentans*” LDH with EtfA ($r^2 = 0.77$) and EtfB ($r^2 = 0.65$) suggest a role for this dehydrogenase and the electron transfer flavoproteins in anaerobic lactate oxidation. Therefore, we hypothesize that in “*Ca. Pseudoramibacter fermentans*,” the electron transfer flavoproteins EtfA and EtfB have a role both in electron bifurcation by ACD in the reverse β -oxidation cycle and in electron confurcation with LDH.

(vii) Predicted conservation of metabolic functions between “*Ca. Pseudoramibacter fermentans*” and related organisms. The organization of genes encoding the complete reverse β -oxidation pathway in “*Ca. Pseudoramibacter fermentans*” is similar to that found in the other two genomes available for *Pseudoramibacter* organisms (Fig. 3). In all three cases, the gene encoding ECoAH clusters with other genes in the reverse β -oxidation pathway, while the *etfAB* genes are found in a separate region of the genome. In all three organisms, the reverse β -oxidation genes are located near

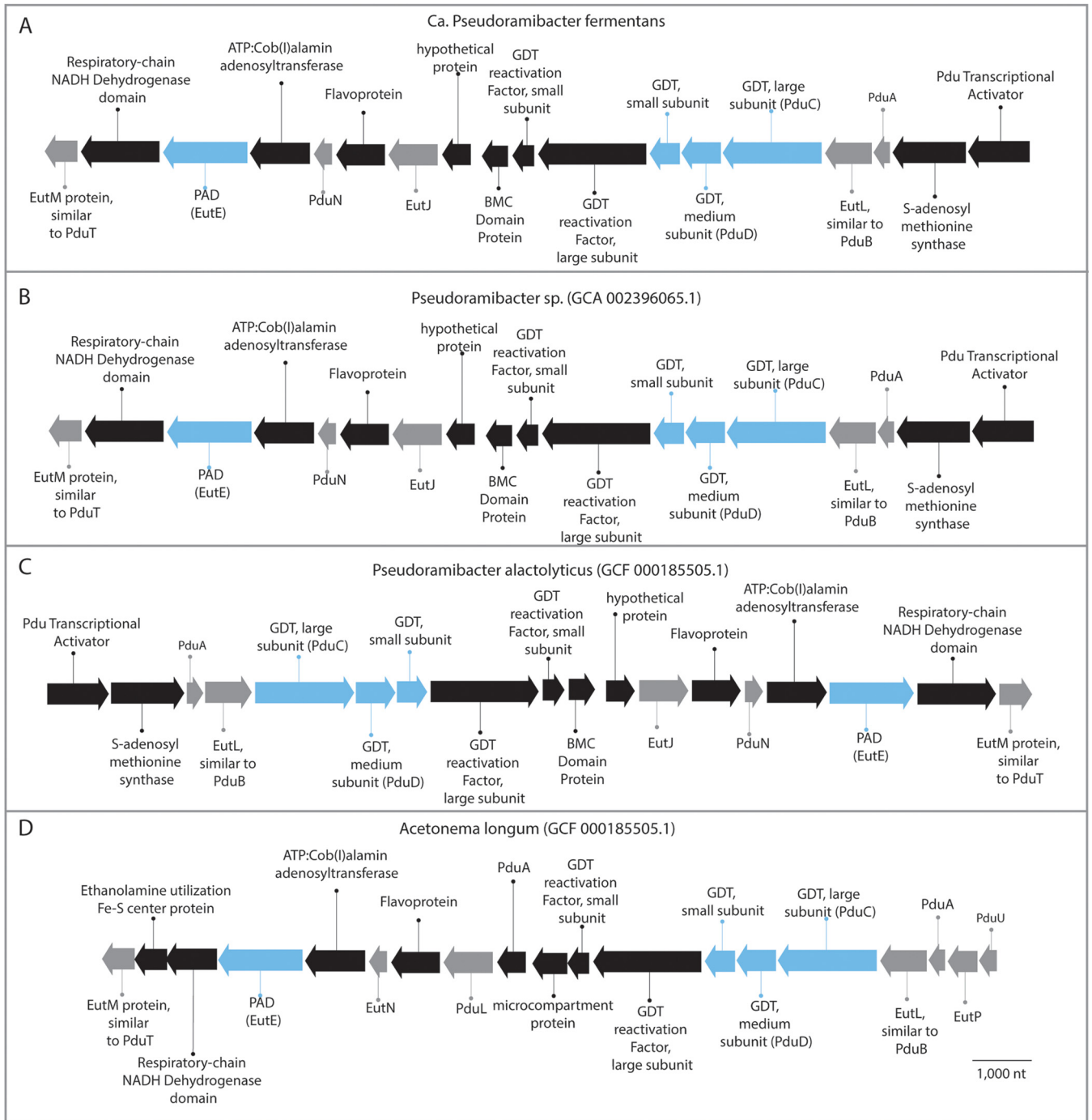


FIG 9 Glycerol utilization gene clusters in “*Ca. Pseudoramibacter fermentans*” (A), related *Pseudoramibacter* species (B and C), and *Acetoneama longum* (D). Blue shading indicates a gene predicted to encode an enzyme involved in route 3 (Fig. 8A), and gray shading indicates a gene predicted to encode a microcompartment protein.

the genes encoding fatty acid biosynthesis (Fig. 3). While other investigators have suggested a potential role for fatty acid biosynthesis genes in MCFA production (28), our data do not provide support for this hypothesis since transcripts for genes encoding enzymes involved in reverse β -oxidation were orders of magnitude more abundant than fatty acid biosynthesis genes (Fig. S7).

The ability of “*Ca. Pseudoramibacter fermentans*” to metabolize glycerol appears to be common in the *Pseudoramibacter* genus as genes for the three proposed routes of

TABLE 2 Thermodynamics of glycerol conversion to butyrate, hexanoate, and octanoate^a

Description reaction	Overall reaction	ΔG° per mol glycerol (kJ mol ⁻¹)
Conversion to butyrate	$2 \text{ C}_3\text{H}_8\text{O}_3 \rightarrow 1 \text{ C}_4\text{H}_7\text{O}_2^- + 2 \text{ CO}_2 + 1 \text{ H}^+ + 4 \text{ H}_2$	-188
Conversion to hexanoate	$3 \text{ C}_3\text{H}_8\text{O}_3 \rightarrow 1 \text{ C}_6\text{H}_{11}\text{O}_2^- + 3 \text{ CO}_2 + 1 \text{ H}_2\text{O} + 1 \text{ H}^+ + 5 \text{ H}_2$	-306
Conversion to octanoate	$4 \text{ C}_3\text{H}_8\text{O}_3 \rightarrow 1 \text{ C}_8\text{H}_{15}\text{O}_2^- + 4 \text{ CO}_2 + 2 \text{ H}_2\text{O} + 1 \text{ H}^+ + 6 \text{ H}_2$	-423

^aFree energies of formation used to calculate ΔG° are provided in Table S2 in the supplemental material.

glycerol metabolism in "Ca. *Pseudoramibacter fermentans*" are also found in the other two *Pseudoramibacter* genomes. This is also the case for the presence of genes predicted to produce protein microcompartments in a cluster with genes encoding glycerol dehydratase (Fig. 9). The production of polyhedral protein microcompartments thought to encase a diol dehydratase when 1,2-propanediol is consumed has been described in other organisms, such as *Salmonella enterica* (a gammaproteobacterium) and *Acetoneema longum* (Fig. 9D), a member of the *Negativicutes* class within the *Firmicutes* phylum (29, 30). In addition to the two other *Pseudoramibacter* genomes which contain glycerol gene clusters identical to those of "Ca. *Pseudoramibacter fermentans*" (Fig. 9B and C), a similar glycerol utilization cluster was identified in *A. longum* (Fig. 9D).

Concluding remarks. Our results reveal several previously unexplored metabolic and energetic features of chain-elongating bacteria. Multi-omic analysis of "Ca. *Weimeria bifida*" suggests that this organism may have a previously undescribed ability to use both the pentose phosphate and the phosphoketolase pathways for pentose consumption and production of acetyl-CoA that is needed for MCFA synthesis. Further, both "Ca. *Weimeria bifida*" and "Ca. *Pseudoramibacter fermentans*" may use multiple hydrogenases, including a proton-translocating energy-conserving hydrogenase (EchABCDEF), to support MCFA production. Although both chain elongators contained genes for the RNF complex, the transcriptomic evidence suggests that this complex may be more important for generating an ion motive force in "Ca. *Weimeria bifida*." Our data also predict that "Ca. *Pseudoramibacter fermentans*" uses several routes to consume glycerol as a carbon source, with potentially toxic intermediates sequestered in protein microcompartments, and a thermodynamic analysis supports the ability of "Ca. *Pseudoramibacter fermentans*" to produce MCFA from this substrate. Finally, our data implicate an electron-confurcating LDH in providing carbon skeletons needed to support MCFA production from lactate by "Ca. *Pseudoramibacter fermentans*." Further work is necessary to elucidate the implications of the genomic features uncovered in this study on the bioenergetics of MCFA production, to assess whether similar processes are involved in the production of MCFA from other substrates, and to develop a better understand-

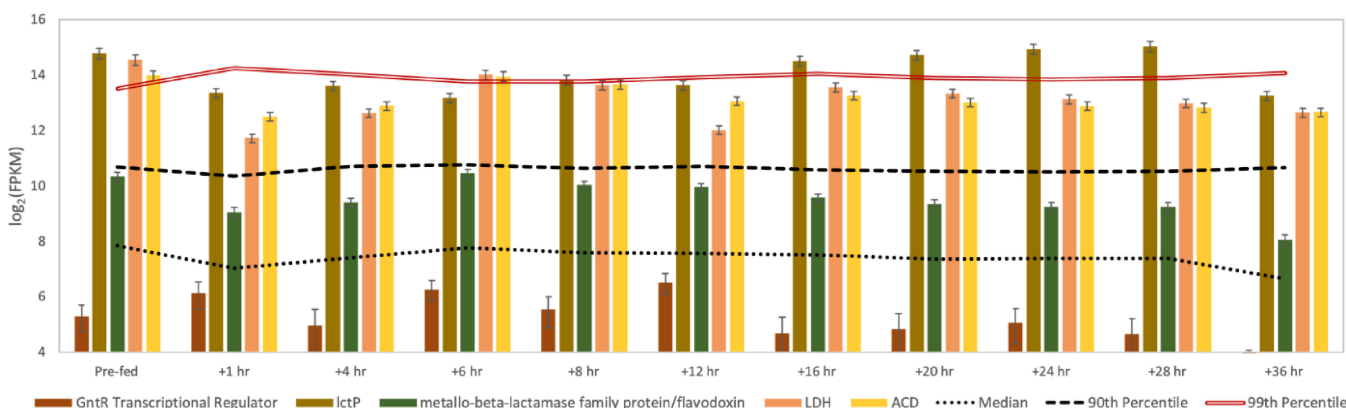


FIG 10 Transcript abundance for a lactate utilization gene cluster in "Ca. *Pseudoramibacter fermentans*." Genes include a transcriptional regulator, lactate permease (*lctP*), a potential flavodoxin, lactate dehydrogenase (LDH), and acetyl-CoA dehydrogenase (ACD). "Prefed" represents the time point prior to feeding the reactor. Error bars show 95% confidence levels determined with CuffLinks.

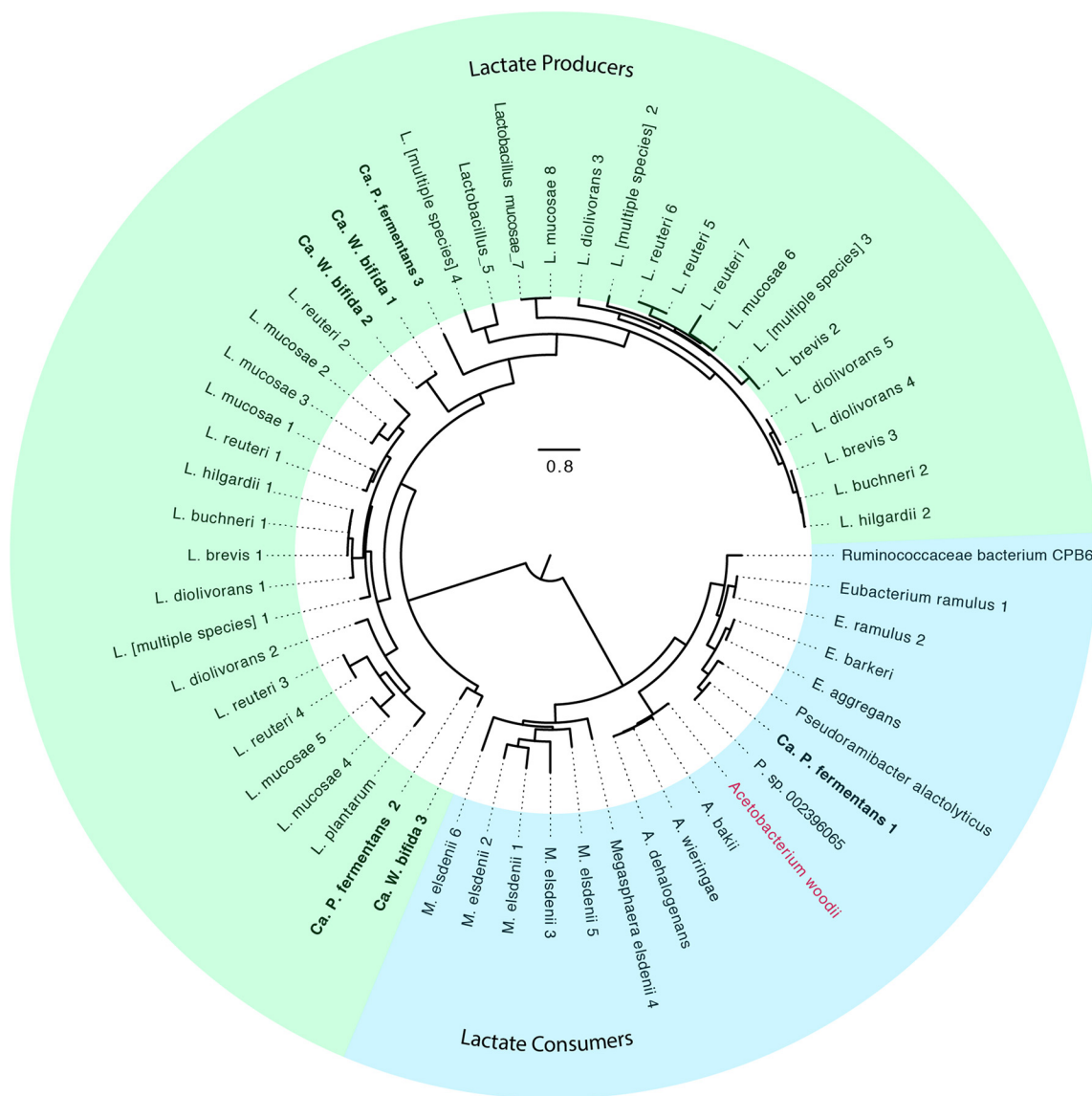


FIG 11 A maximum-likelihood phylogenetic tree of lactate dehydrogenase amino acid sequences among different members of the *Firmicutes* phylum. Anaerobic lactate consumers contain lactate dehydrogenase and form a distinct cluster from lactate producers, such as those found among the lactobacilli.

ing of microbial pathways for production of additional valuable products from renewable organic materials.

Summary descriptions of novel organisms. The formal proposals of the new *Candidatus* genus "*Candidatus Weimeria*" and new species "*Candidatus Weimeria bifida*" and "*Candidatus Pseudoramibacter fermentans*" are given in Table 3.

MATERIALS AND METHODS

Bioreactor operation. We operated a bioreactor containing 150 ml of liquid. Lignocellulosic biorefinery residues, prepared as described previously (2), were added into the bioreactor, and reactor liquid was pumped out of the bioreactor every hour to maintain a residence time in the reactor of 6 days. The pH of the reactor was controlled at 5.5 by adding 5 M KOH through a pump attached to a pH controller. The temperature of the bioreactor was maintained at 35°C using a water bath. For the 36-h time series experiment, 28 ml of liquid was removed from the reactor, and 28 ml of lignocellulosic biorefinery residues was added one time, bringing the starting liquid volume to 150 ml.

Reactor sampling. Prior to feeding the reactor with 28 ml of lignocellulosic biorefinery residues, samples were collected for metagenomic and metatranscriptomic analyses. Samples for DNA sequence analysis were collected in 2-ml centrifuge tubes and centrifuged at 10,000 × g for 10 min. After the supernatant was decanted, cell pellets were stored at -80°C until DNA was extracted. Samples for RNA

TABLE 3 Protologues for “*Ca. Weimeria bifida*” and “*Ca. Pseudoramibacter fermentans*”

Characteristic or parameter	Value for “ <i>Candidatus Weimeria bifida</i> ”	Value for “ <i>Candidatus Pseudoramibacter fermentans</i> ”
Taxon no.	NA ^a	NA
Species name	“ <i>Candidatus Weimeria bifida</i> ”	“ <i>Candidatus Pseudoramibacter fermentans</i> ”
Genus name	“ <i>Candidatus Weimeria</i> ”	<i>Pseudoramibacter</i>
Specific epithet	<i>bifida</i>	<i>fermentans</i>
Genus etymology	Wei.mer'ia. N.L. fem. n., named after Paul J. Weimer, a pioneer in using rumen microbes to produce valuable chemicals	Pseu.do.ra.mi.bac'ter. Gr. adj. <i>pseudês</i> , false; L. masc. n. <i>ramus</i> , a branch; N.L. masc. n. <i>bacter</i> , rod, staff; N.L. masc. n. <i>Pseudoramibacter</i> , false branching rod
Type species of the genus	“ <i>Candidatus Weimeria bifida</i> ”	<i>Pseudoramibacter alactolyticus</i> (GCA_000185505.1)
Taxon number of the type species	NA	NA
Genus status	gen. nov.	Existing, validly published name
Species etymology	bif'i.da. L. fem. adj. <i>bifida</i> , divided into two parts, as in the bifid shunt	fer.men'tans. L. part. adj. <i>fermentans</i> , fermenting
Species status	sp. nov.	sp. nov.
Designation of the type MAG	VOGC00000000	VOGB00000000
MAG/SAG accession number	VOGC01000000	VOGB01000000
Genome status	Draft	Draft
Genome size (bp)	2,394,983	2,285,558
GC mol%	45.9	50.2
Country of origin	USA	USA
Region of origin	Wisconsin	Wisconsin
Source of sample	Bioreactor	Bioreactor
Sampling date(s)	Multiple	Multiple
Geographic location	University of Wisconsin—Madison	University of Wisconsin—Madison
Latitude	43.035801 N	43.035801 N
Longitude	89.359081 W	89.359081 W
Depth	NA	NA
Altitude	NA	NA
Temperature of the sample (°C)	35	35
pH of the sample	5.5	5.5
Relationship to O ₂	Anaerobe	Anaerobe
Energy metabolism	Heterotroph	Heterotroph
Assembly (no. of samples)	5	5
Sequencing technology	Illumina HiSeq 2000 and PacBio	Illumina HiSeq 2000 and PacBio
Binning software used	Anvio, version 5	Anvio, version 5
Assembly software used	Spades, version 3.12	Spades, version 3.12
Habitat	Anaerobic bioreactor	Anaerobic bioreactor
Miscellaneous, extraordinary features relevant for the description	When a bioreactor community, seeded from acid digester sludge collected at the Nine Springs Wastewater Treatment Plant (Madison, WI), is fed lignocellulosic biorefinery residues, “ <i>Ca. Weimeria bifida</i> ” is predicted to use the phosphoketolase pathway (bifid shunt) to degrade xylose and produce butyrate and medium-chain fatty acids via reverse β -oxidation. “ <i>Ca. Weimeria bifida</i> ” is predicted to maintain redox balance by using a ferredoxin-dependent hydrogenase or an energy-conserving Ech hydrogenase to produce hydrogen. While other organisms within the proposed “ <i>Ca. Weimeria</i> ” genus contain genes for pentose degradation via the oxidative pentose phosphate pathway, “ <i>Ca. Weimeria bifida</i> ” is unique in containing genes for the bifid shunt.	When lignocellulosic biorefinery residues are fed to a bioreactor community that was seeded from acid digester sludge collected at the Nine Springs Wastewater Treatment Plant (Madison, WI), “ <i>Ca. Pseudoramibacter fermentans</i> ” is predicted to produce medium-chain fatty acids via reverse β -oxidation, using lactate and glycerol as the organic substrates. “ <i>Ca. Pseudoramibacter fermentans</i> ” contains genes for multiple glycerol utilization routes, including conversion of glycerol to reuterin. Lactate utilization is predicted to involve an electron-confurcating lactate dehydrogenase. In addition, an electron-bifurcating acyl-CoA dehydrogenase is predicted to be used for energy conservation, with reduced ferredoxin produced by this enzyme used by the RNF complex for the creation of an ion motive force to support ATP production. “ <i>Ca. Pseudoramibacter fermentans</i> ” is predicted to produce hydrogen via a ferredoxin-dependent hydrogenase as a mechanism for balancing internal redox conditions.

^aNA, not available.

were collected in 2-ml centrifuge tubes and centrifuged at 10,000 × *g* for 1.5 min. After the supernatant was decanted, samples for RNA were flash frozen in liquid nitrogen and then stored at –80°C until RNA was extracted. Three samples for RNA extraction and sequencing were collected as a control prior to adding the conversion residue. At each time point after the addition of lignocellulosic biorefinery residues, one sample was collected for RNA extraction, and the supernatant from this sample was used

for high-performance liquid chromatography (HPLC) and gas chromatography mass spectrometry (GC-MS) analyses, as described previously (2).

DNA sequencing. DNA was extracted from biomass samples using a phenol-chloroform extraction method described previously (8). For all metagenomic samples (days 96, 120, 168, 252, and 378), sequencing was performed with an Illumina HiSeq 2500 sequencer to generate 2- by 250-bp reads. For the day 378 sample, PacBio sequencing and library preparation were performed by the Joint Genome Institute. The sequence library was prepared using a PacBio 10-kb low-input library preparation, and libraries were sequenced with a PacBio Sequel.

Metagenomic analysis. Quality checking and trimming of Illumina DNA reads were performed using the program Sickle (31). The resulting reads from days 96, 120, 168, 252, and 378 were coassembled using metaSPAdes, version 3.12.00 (32). The assembled contigs were binned using Anvio, version 5 (33). After binning, draft genomes were gap filled with the PacBio reads from day 378 using PBSuite, version 15.8.24 (34). Quality checking was performed on the gap-filled draft genomes using CheckM (35). Draft genomes were annotated with MetaPathways, version 2.5 (36). Taxonomic assignments based on single-copy marker genes were made with the GTDB tool kit (9), and phylogenetic trees were constructed using RAxML (37). Average nucleotide identities were calculated with JSpecies using the ANIb algorithm (38).

Metatranscriptomic analysis. RNA was extracted, and cDNA was synthesized and sequenced as described previously (8). RNA was collected for 11 time points starting on day 252 of reactor operations: before feeding the conversion residue to the bioreactor ("prefed") and at various time points for 36 h after feeding (+1, +4, +6, +8, +12, +16, +20, +24, +28, and +36 h). Triplicate samples were taken for the prefed time point and underwent separate extraction, rRNA depletion, cDNA synthesis, library preparation, and sequencing. Quality checking and trimming of raw reads were performed with Sickle (31), and rRNA reads were removed with SortMeRNA (39). Reads from each time point were mapped to open reading frames in the MAGs with CuffLinks (40). Confidence intervals (95%) were calculated using CuffLinks using the triplicate sample for the prefed time point to estimate variance for each transcript.

Data availability. The whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession numbers [VOGB00000000](https://doi.org/10.1093/bioinformatics/btq000) and [VOGC00000000](https://doi.org/10.1093/bioinformatics/btq000). The versions described in this paper are versions VOGB01000000 and VOGC01000000. The raw DNA sequences are available in the National Center for Biotechnology Information Sequence Read Archive (SRA) database under BioProject [PRJNA418244](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA418244) for days 96 and 120 and BioProject [PRJNA535528](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA535528) for days 168, 252, and 378. The raw RNA sequences are available in the SRA database under BioProject [PRJNA535528](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA535528).

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 0.9 MB.

SUPPLEMENTAL FILE 2, XLSX file, 0.02 MB.

SUPPLEMENTAL FILE 3, XLSX file, 0.8 MB.

SUPPLEMENTAL FILE 4, XLSX file, 0.4 MB.

SUPPLEMENTAL FILE 5, XLSX file, 1.9 MB.

SUPPLEMENTAL FILE 6, XLSX file, 2 MB.

SUPPLEMENTAL FILE 7, XLSX file, 0.9 MB.

SUPPLEMENTAL FILE 8, XLSX file, 0.6 MB.

SUPPLEMENTAL FILE 9, XLSX file, 0.02 MB.

ACKNOWLEDGMENTS

This work was funded by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-SC0018409). M.J.S. is supported by the National Science Foundation Graduate Research Fellowship Program under grant DGE-1256259. We thank the University of Wisconsin Biotechnology Center and the U.S. Department of Energy Joint Genome Institute for library preparation and sequencing. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under contract number DE-AC02-05CH11231.

We also thank Pamela Camejo and Francisco Moya for assistance with RNA sequencing and data analyses and Maria Chuvochina for help with the etymology of the proposed names.

REFERENCES

1. Angenent LT, Richter H, Buckel W, Spirito CM, Steinbusch KJJ, Plugge CM, Strik D, Grootscholten TIM, Buisman CJN, Hamelers H. 2016. Chain elongation with reactor microbiomes: open-culture biotechnology to produce biochemicals. *Environ Sci Technol* 50:2796–2810. <https://doi.org/10.1021/acs.est.5b04847>.
2. Scarborough MJ, Lynch G, Dickson M, McGee M, Donohue TJ, Noguera DR. 2018. Increasing the economic value of lignocellulosic stillage through medium-chain fatty acid production. *Biotechnol Biofuels* 11:200. <https://doi.org/10.1186/s13068-018-1193-x>.
3. Sarria S, Kruyer NS, Peralta-Yahya P. 2017. Microbial synthesis of

- medium-chain chemicals from renewables. *Nat Biotechnol* 35: 1158–1166. <https://doi.org/10.1038/nbt.4022>.
4. Prabhu R, Altman E, Eiteman MA. 2012. Lactate and acrylate metabolism by *Megasphaera elsdenii* under batch and steady-state conditions. *Appl Environ Microbiol* 78:8564–8570. <https://doi.org/10.1128/AEM.02443-12>.
 5. Zhu X, Zhou Y, Wang Y, Wu T, Li X, Li D, Tao Y. 2017. Production of high-concentration n-caproic acid from lactate through fermentation using a newly isolated Ruminococcaceae bacterium CPB6. *Biotechnol Biofuels* 10:102. <https://doi.org/10.1186/s13068-017-0788-y>.
 6. Kenealy WR, Cao Y, Weimer PJ. 1995. Production of caproic acid by cocultures of ruminal cellulolytic bacteria and *Clostridium kluyveri* grown on cellulose and ethanol. *Appl Microbiol Biotechnol* 44:507–513. <https://doi.org/10.1007/s002530050590>.
 7. Nelson R, Peterson D, Karp E, Beckham G, Salvachúa D. 2017. Mixed carboxylic acid production by *Megasphaera elsdenii* from glucose and lignocellulosic hydrolysate. *Fermentation* 3:10. <https://doi.org/10.3390/fermentation3010010>.
 8. Scarborough MJ, Lawson CE, Hamilton JJ, Donohue TJ, Noguera DR. 2018. Metatranscriptomic and thermodynamic insights into medium-chain fatty acid production using an anaerobic microbiome. *mSystems* 3:e00221-18. <https://doi.org/10.1128/msystems.00221-18>.
 9. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 36:996–1004. <https://doi.org/10.1038/nbt.4229>.
 10. Seshadri R, Leahy SC, Attwood GT, Teh KH, Lambie SC, Cookson AL, Eloe-Fadrosh EA, Pavlopoulos GA, Hadjithomas M, Varghese NJ, Paez-Espino D, Hungate 1000 Project Collaborators, Perry R, Henderson G, Creevey CJ, Terrapon N, Lapebie P, Drula E, Lombard V, Rubin E, Kyrpides NC, Henrissat B, Woyke T, Ivanova NN, Kelly WJ. 2018. Cultivation and sequencing of rumen microbiome members from the Hungate 1000 Collection. *Nat Biotechnol* 36:359–367. <https://doi.org/10.1038/nbt.4110>.
 11. Downes J, Munson MA, Radford DR, Spratt DA, Wade WG. 2002. Shuttleworthia satelles gen. nov., sp. nov., isolated from the human oral cavity. *Int J Syst Evol Microbiol* 52:1469–1475. <https://doi.org/10.1099/00207713-52-5-1469>.
 12. Holdeman LV, Cato EP, Moore WEC. 1967. Amended description of *Ramibacterium alactolyticum* Prevot and Taffanel with proposal of a neotype strain. *Int J Syst Bacteriol* 17:323–341. <https://doi.org/10.1099/00207713-17-4-323>.
 13. Buckel W, Thauer RK. 2013. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na⁺ translocating ferredoxin oxidation. *Biochim Biophys Acta* 1827:94–113. <https://doi.org/10.1016/j.bbabi.2012.07.002>.
 14. Li F, Hinderberger J, Seedorf H, Zhang J, Buckel W, Thauer RK. 2008. Coupled ferredoxin and crotonyl coenzyme A (CoA) reduction with NADH catalyzed by the butyryl-CoA dehydrogenase/Etf complex from *Clostridium kluyveri*. *J Bacteriol* 190:843–850. <https://doi.org/10.1128/JB.01417-07>.
 15. Wang M, Zhao Q, Li L, Niu K, Li Y, Wang F, Jiang B, Liu K, Jiang Y, Fang X. 2016. Contributing factors in the improvement of cellulose H₂ production in *Clostridium thermocellum*/*Thermoanaerobacterium* cocultures. *Appl Microbiol Biotechnol* 100:8607–8620. <https://doi.org/10.1007/s00253-016-7776-1>.
 16. Biegel E, Schmidt S, Gonzalez JM, Muller V. 2011. Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. *Cell Mol Life Sci* 68:613–634. <https://doi.org/10.1007/s00018-010-0555-8>.
 17. Blokesch M, Bock A. 2002. Maturation of [NiFe]-hydrogenases in *Escherichia coli*: the HypC cycle. *J Mol Biol* 324:287–296. [https://doi.org/10.1016/s0022-2836\(02\)01070-7](https://doi.org/10.1016/s0022-2836(02)01070-7).
 18. Maier T, Jacobi A, Sauter M, Bock A. 1993. The product of the *hypB* gene, which is required for nickel incorporation into hydrogenases, is a novel guanine nucleotide-binding protein. *J Bacteriol* 175:630–635. <https://doi.org/10.1128/jb.175.3.630-635.1993>.
 19. Calusinska M, Happe T, Joris B, Wilmette A. 2010. The surprising diversity of clostridial hydrogenases: a comparative genomic perspective. *Microbiology* 156:1575–1588. <https://doi.org/10.1099/mic.0.032771-0>.
 20. Dillon SC, Bateman A. 2004. The Hotdog fold: wrapping up a superfamily of thioesterases and dehydratases. *BMC Bioinformatics* 5:109. <https://doi.org/10.1186/1471-2105-5-109>.
 21. O'Callaghan A, van Sinderen D. 2016. Bifidobacteria and their role as members of the human gut microbiota. *Front Microbiol* 7:925. <https://doi.org/10.3389/fmicb.2016.00925>.
 22. de Vries W, Stouthamer AH. 1967. Pathway of glucose fermentation in relation to the taxonomy of bifidobacteria. *J Bacteriol* 93:574–576.
 23. Liu L, Zhang L, Tang W, Gu Y, Hua Q, Yang S, Jiang W, Yang C. 2012. Phosphoketolase pathway for xylose catabolism in *Clostridium acetobutylicum* revealed by ¹³C metabolic flux analysis. *J Bacteriol* 194: 5413–5422. <https://doi.org/10.1128/JB.00713-12>.
 24. Ku HK, Do NH, Song JS, Choi S, Yeon SH, Shin MH, Kim KJ, Park SR, Park IY, Kim SK, Lee SJ. 2011. Crystal structure of prephenate dehydrogenase from *Streptococcus mutans*. *Int J Biol Macromol* 49:761–766. <https://doi.org/10.1016/j.ijbiomac.2011.07.009>.
 25. Kunishima N, Asada Y, Sugahara M, Ishijima J, Nodake Y, Sugahara M, Miyano M, Kuramitsu S, Yokoyama S, Sugahara M. 2005. A novel induced-fit reaction mechanism of asymmetric hot dog thioesterase PAA1. *J Mol Biol* 352:212–228. <https://doi.org/10.1016/j.jmb.2005.07.008>.
 26. Talarico TL, Casas IA, Chung TC, Dobrogosz WJ. 1988. Production and isolation of reuterin, a growth inhibitor produced by *Lactobacillus reuteri*. *Antimicrob Agents Chemother* 32:1854–1858. <https://doi.org/10.1128/aac.32.12.1854>.
 27. Weghoff MC, Bertsch J, Muller V. 2015. A novel mode of lactate metabolism in strictly anaerobic bacteria. *Environ Microbiol* 17:670–677. <https://doi.org/10.1111/1462-2920.12493>.
 28. Han W, He P, Shao L, Lu F. 2018. Metabolic interactions of a chain elongation microbiome. *Appl Environ Microbiol* 84:e01614-18. <https://doi.org/10.1128/AEM.01614-18>.
 29. Bobik TA, Havemann GD, Busch RJ, Williams DS, Aldrich HC. 1999. The propanediol utilization (*pdu*) operon of *Salmonella enterica* serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B(12)-dependent 1, 2-propanediol degradation. *J Bacteriol* 181:5967–5975.
 30. Tocheva EI, Matson EG, Cheng SN, Chen WG, Leadbetter JR, Jensen GJ. 2014. Structure and expression of propanediol utilization microcompartments in *Acetone nema longum*. *J Bacteriol* 196:1651–1658. <https://doi.org/10.1128/JB.00049-14>.
 31. Joshi NA, Fass JN. 2011. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (version 1.33). <https://github.com/najoshi/sickle>.
 32. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* 27:824–834. <https://doi.org/10.1101/gr.213959.116>.
 33. Eren AM, Esen OC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319. <https://doi.org/10.7717/peerj.1319>.
 34. English AC, Richards S, Han Y, Wang M, Vee V, Qu J, Qin X, Muzny DM, Reid JG, Worley KC, Gibbs RA. 2012. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS One* 7:e47768. <https://doi.org/10.1371/journal.pone.0047768>.
 35. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
 36. Konwar KM, Hanson NW, Bhatia MP, Kim D, Wu SJ, Hahn AS, Morgan-Lang C, Cheung HK, Hallam SJ. 2015. MetaPathways v2.5: quantitative functional, taxonomic and usability improvements. *Bioinformatics* 31: 3345–3347. <https://doi.org/10.1093/bioinformatics/btv361>.
 37. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>.
 38. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32:929–931. <https://doi.org/10.1093/bioinformatics/btv681>.
 39. Kopylova E, Noe L, Touzet H. 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28:3211–3217. <https://doi.org/10.1093/bioinformatics/bts611>.
 40. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28:511–515. <https://doi.org/10.1038/nbt.1621>.