



## OPEN ACCESS

## EDITED BY

Matthias Boll,  
University of Freiburg, Germany

## REVIEWED BY

Wolfgang Buckel,  
University of Marburg, Germany  
Tobias Goris,  
German Institute of Human Nutrition  
Potsdam-Rehbruecke (DIfE), Germany

## \*CORRESPONDENCE

Evgenii N. Frolov  
✉ evgenii\_frolov\_89@gmail.ru

RECEIVED 13 March 2023

ACCEPTED 24 April 2023

PUBLISHED 12 May 2023

## CITATION

Frolov EN, Elcheninov AG, Gololobova AV,  
Toshchakov SV, Novikov AA, Lebedinsky AV  
and Kublanov IV (2023) Obligate autotrophy  
at the thermodynamic limit of life in a new  
acetogenic bacterium.  
*Front. Microbiol.* 14:1185739.  
doi: 10.3389/fmicb.2023.1185739

## COPYRIGHT

© 2023 Frolov, Elcheninov, Gololobova,  
Toshchakov, Novikov, Lebedinsky and  
Kublanov. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in this  
journal is cited, in accordance with accepted  
academic practice. No use, distribution or  
reproduction is permitted which does not  
comply with these terms.

# Obligate autotrophy at the thermodynamic limit of life in a new acetogenic bacterium

Evgenii N. Frolov<sup>1\*</sup>, Alexander G. Elcheninov<sup>1</sup>,  
Alexandra V. Gololobova<sup>1</sup>, Stepan V. Toshchakov<sup>2</sup>,  
Andrei A. Novikov<sup>3</sup>, Alexander V. Lebedinsky<sup>1</sup> and  
Ilya V. Kublanov<sup>1</sup>

<sup>1</sup>Federal Research Center of Biotechnology, Winogradsky Institute of Microbiology, Russian Academy of Sciences, Moscow, Russia, <sup>2</sup>Kurchatov Center for Genome Research, National Research Center "Kurchatov Institute", Moscow, Russia, <sup>3</sup>Department of Physical and Colloid Chemistry, Gubkin University, Moscow, Russia

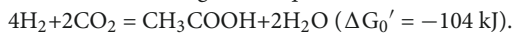
One of the important current issues of bioenergetics is the establishment of the thermodynamic limits of life. There is still no final understanding of what is the minimum value of the energy yield of a reaction that is sufficient to be used by an organism (the so-called "biological quantum of energy"). A reasonable model for determination of the minimal energy yield would be microorganisms capable of living on low-energy substrates, such as acetogenic prokaryotes. The most prominent metabolic feature of acetogens is autotrophic growth with molecular hydrogen and carbon dioxide as the substrates, which is hardly competitive in environments. Most probably, that is why only facultative autotrophic acetogens have been known so far. Here, we describe the first obligately autotrophic acetogenic bacterium *Aceticella autotrophica* gen. nov., sp. nov., strain 3443-3Ac<sup>T</sup>. Phylogenetically, the new genus falls into a monophyletic group of heterotrophic bacteria of the genera *Thermoanaerobacterium*, *Thermoanaerobacter*, and *Caldanaerobacter* (hereinafter referred to as TTC group), where the sole acetogenic representative has so far been the facultatively autotrophic *Thermoanaerobacter kivui*. *A. autotrophica* and *T. kivui* both are acetogens employing energy-converting hydrogenase (Ech-acetogens) that are likely to have inherited the acetogenesis capacity vertically from common ancestor. However, their acetogenic machineries have undergone different adjustments by gene replacements due to horizontal gene transfers from different donors. Obligate autotrophy of *A. autotrophica* is associated with the lack of many sugar transport systems and carbohydrate catabolism enzymes that are present in other TTC group representatives, including *T. kivui*.

## KEYWORDS

acetogenesis, obligate autotrophy, acetogenic bacterium, the Wood-Ljungdahl pathway, Ech complex, electron-bifurcating hydrogenase, *Aceticella*

## 1. Introduction

Acetogenic bacteria are a specialized group of strictly anaerobic microorganisms that are able to produce acetate from two molecules of carbon dioxide (CO<sub>2</sub>) via the Wood–Ljungdahl pathway (WLP) during anaerobic respiration. Their lithotrophic growth occurs according to the equation:



Under environmental conditions, the free energy change of the reaction allows for the synthesis of only a fraction of an ATP mole, and acetogens are thus a paradigm for microbial life under extreme energy limitation (Schuchmann and Müller, 2014).

Currently, over 100 acetogens, representing 28 bacterial genera, have been isolated from diverse habitats (Drake et al., 2008; Allen et al., 2010; Zhilina et al., 2012; Poehlein et al., 2013; Gerritsen et al., 2014; Lawson et al., 2016; Mao et al., 2021). While most acetogenic isolates are in the phylum Firmicutes, they also include Spirochaetes like *Treponema azotonutricum* and *Treponema primitia*, Desulfobacterota like *Desulfotignum phosphitoxidans*, and Acidobacteria like *Holophaga foetida* (hereafter, unless otherwise indicated, taxonomy is according to Bergey's Manual of Systematics of Archaea and Bacteria).<sup>1</sup> Among Archaea, only for *Methanosarcina acetivorans* and *Archaeoglobus fulgidus* and only with the low-potential CO as the electron donor was the energy conservation via acetogenesis reported (Rother and Metcalf, 2004; Henstra et al., 2007; Schöne et al., 2022). In addition, acetogenesis has been suggested for some uncultivated Bacteria like 'Candidatus Adiatrix intracellularis' (Ikeda-Ohtsubo et al., 2016) and Archaea, e.g., members of the phyla Bathyarchaeota and Lokiarchaeota and anaerobic methanotrophic archaea ANME-2a (Evans et al., 2015; He et al., 2016; Orsi et al., 2020; Yang et al., 2020; Loh et al., 2021; Deb and Das, 2022).

The most prominent metabolic feature of acetogens is chemolithoautotrophic growth with H<sub>2</sub> as the energy source and HCO<sub>3</sub><sup>-</sup>/CO<sub>2</sub> as the carbon source and electron acceptor. However, acetogens have to grow under highly competitive conditions, since, in anaerobic environments methanogens and sulfate-reducing prokaryotes are usually the dominant hydrogenotrophs, outcompeting acetogens due to the greater energy yield (Ragsdale and Pierce, 2008) and a lower H<sub>2</sub> threshold (Le Van et al., 1998). Most probably, that is why only facultative autotrophic acetogens have been known so far. Heterotrophic capabilities of acetogens are quite flexible, they are known to utilize a vast variety of organic substrates like sugars, alcohols, organic acids, aldehydes, and aromatic compounds. Acetogens can also use a variety of alternative electron acceptors, like nitrate, nitrite, thiosulfate, sulfate, Fe(III), perchlorate, dimethylsulfoxide, fumarate, and others (Ragsdale and Pierce, 2008). The ability of acetogenic bacteria to form acetate from many organic substrates trophically links them to methanogens that can use acetate as a substrate. This makes acetogens an essential part of anaerobic food webs (Schuchmann and Müller, 2014). Metabolic flexibility is believed to be a key capability of acetogens, explaining the almost ubiquitous distribution of acetogenic bacteria in anoxic environments: soils, freshwater and marine sediments, subsurface samples, alkaline and high-salt environments, hot springs, sludge, and the intestinal tracts of many animals, including

humans and termites (Ragsdale and Pierce, 2008; Schuchmann and Müller, 2014). Thus, acetogens play an important part in the global carbon cycle, although they never represent a numerically significant part of microbial communities.

The catabolism of acetogens is usually regarded as modular. The first module consists of electron-donating reactions (the oxidative part of acetogenesis), the second module includes the reductive reactions of acetogenesis (the Wood–Ljungdahl pathway), while the third module connects redox cofactor recycling to energy conservation and comprises a membrane-bound respiration system involving Rnf (ferredoxin: NAD<sup>+</sup> oxidoreductase termed after *Rhodobacter* nitrogen fixation complex) or Ech (energy-converting hydrogenase) complex and ATP synthase (Schuchmann and Müller, 2014, 2016; Basen and Müller, 2017). In the first module, organic or inorganic substrates are oxidized by donating electrons to ferredoxin (Fd) and NAD<sup>+</sup>. It is important to note that the largest thermodynamic barrier in acetogenesis is the reduction of CO<sub>2</sub> to CO in the carbonyl branch of WLP. This redox couple has a very low standard redox potential (E<sub>0</sub>' = -520 mV, Thauer et al., 1977). Therefore acetogens must employ low potential ferredoxin, and there are ferredoxins that are known to have redox potentials of -500 mV and even lower (Bar-Even, 2013). During chemoorganoheterotrophic growth (Kuhns et al., 2020) a few enzymes such as tungsten-containing aldehyde:ferredoxin oxidoreductase (Heider et al., 1995) or pyruvate:ferredoxin oxidoreductase (Blamey and Adams, 1993) can replenish the low potential reduced ferredoxin, oxidized in the WLP carbonyl branch. However, during chemolithoautotrophic growth with hydrogen, low potential ferredoxin has to be reduced with hydrogen as the electron donor, but this reaction is endergonic (Schuchmann and Müller, 2014) since the standard redox potential of the couple H<sup>+</sup>/H<sub>2</sub> is -414 mV (Thauer et al., 1977). This energy barrier is overcome by means of flavin-based electron bifurcation, where the endergonic reduction of low potential ferredoxin is coupled to the exergonic reduction of NAD<sup>+</sup> (Buckel and Thauer, 2018; Müller et al., 2018). In acetogens, this reaction is catalyzed by the soluble electron-bifurcating hydrogenase (Schuchmann and Müller, 2012, 2014). Ferredoxin and NAD<sup>+</sup> are reduced simultaneously in a 1/1 stoichiometry in an overall exergonic reaction ( $\Delta G_0' = -11 \text{ kJ per mole}$ , Herrmann et al., 2008; Schuchmann and Müller, 2012; Buckel and Thauer, 2013; Wang et al., 2013). Therefore, the oxidative module of acetogenesis yields NADH and reduced ferredoxin. Part of this ferredoxin is consumed in the second, reductive acetogenesis module (specifically, in the WLP carbonyl branch), and the remaining ferredoxin is used to generate a chemiosmotic gradient in the third module of acetogenesis (Schuchmann and Müller, 2014).

It is a unique property of acetogenic microorganisms to use the WLP not only as the carbon fixation pathway in anabolism but also as the terminal electron sink. How the WLP can be coupled to the synthesis of ATP has been thoroughly studied in *Acetobacterium woodii*, which has a simple respiratory chain consisting of a ferredoxin: NAD<sup>+</sup> oxidoreductase (Rnf complex) and a Na<sup>+</sup>-dependent F<sub>1</sub>F<sub>0</sub>-ATP synthase (Fritz and Müller, 2007; Biegel and Müller, 2010; Hess et al., 2013). The Rnf complex generates a transmembrane electrochemical Na<sup>+</sup> gradient, which fuels a Na<sup>+</sup>-dependent ATP synthase. Another model organism for the study of acetogenesis, *Clostridium ljungdahlii*, also contains an Rnf complex, but the chemiosmotic ATP synthesis in this organism

<sup>1</sup> <https://onlinelibrary.wiley.com/doi/book/10.1002/9781118960608>

seems to be driven by protons (Köpke et al., 2010; Tremblay et al., 2012). However, the Rnf complex is not present in every acetogen, and in the case of its absence its role is accomplished by the energy-converting hydrogenases (Ech complexes), detected in the Rnf-free acetogens like *Thermoanaerobacter kivui*, *Thermacetogenium phaeum* and acetogenic representatives of the genus *Moorella* (Hess et al., 2014; Schuchmann and Müller, 2014; Basen and Müller, 2017; Schoelmerich and Müller, 2019). Energy conservation by a hydrogenase-dependent chemiosmotic mechanism has been studied in detail for the thermophilic acetogenic bacterium *Thermoanaerobacter kivui* (Hess et al., 2014; Schoelmerich and Müller, 2019). The operation of the Ech complex of *T. kivui* leads to the translocation of  $H^+$  and  $Na^+$  across the membrane, and the  $H^+$  gradient is harnessed for energy conservation by the  $H^+$ -dependent  $F_1F_0$ -ATP synthase. Thus, acetogenic bacteria have been classified into “Rnf-acetogens” and “Ech-acetogens” (Schuchmann and Müller, 2014).

In this work we isolated the first obligately autotrophic acetogenic bacterium *Aceticella autotrophica* gen. nov., sp. nov., strain 3443-3Ac<sup>T</sup>, and described its metabolic properties, focusing on its energy metabolism and the likely causes for inability to grow heterotrophically.

## 2. Materials and methods

### 2.1. Enrichment, isolation, and physiological studies

Strain 3443-3Ac<sup>T</sup> was isolated from a sample of sediments collected from the terrestrial hot spring Kaskadny at East Thermal Field, Uzon Caldera, Kamchatka, Russia (N54° 30.026' E160° 00.374', elevation 658 m) in August 2015. The thermal fluid parameters were 58°C, pH 3.1 and Eh +90 mV. The indigenous microbial community of the hot spring Kaskadny was described previously (Merkel et al., 2017). Samples were taken anaerobically, placed into tightly stoppered bottles and transported to the laboratory. An enrichment culture was initiated by inoculating the sample [10% (w/v)] into anaerobically prepared sterilized (by autoclaving at 121°C for 1 h) liquid medium of the following composition (g L<sup>-1</sup>): NH<sub>4</sub>Cl, 0.33; KCl, 0.33; MgCl<sub>2</sub> · 2H<sub>2</sub>O, 0.33; CaCl<sub>2</sub> · 6H<sub>2</sub>O, 0.33; KH<sub>2</sub>PO<sub>4</sub>, 0.33; trace element solution (Kevbrin and Zavarzin, 1992), 1 ml L<sup>-1</sup>; vitamin solution (Wolin et al., 1963), 1 ml L<sup>-1</sup>. Sodium sulfide (0.5 g L<sup>-1</sup>) was used as a reducing agent. Resazurin was added as a redox indicator in a concentration of 1.0 mg L<sup>-1</sup>. To adjust the pH of the medium to 6.0, 2N NaOH was used. The medium was dispensed in 5-ml aliquots into 17-ml Hungate tubes; the head space was filled with H<sub>2</sub>/CO<sub>2</sub> (4:1, 152 kPa). After 4 days of incubation at 50°C, a pronounced production of acetate accompanied by microbial growth was observed. The culture was purified using the dilution-to-extinction technique, and the resulting isolate was designated as strain 3443-3Ac<sup>T</sup>.

Cell growth and cell morphology were monitored using an Olympus CX-41 phase-contrast microscope. Negatively stained cells and fine structure of the cells were studied using a JEM-100 electron microscope as described previously (Bonch-Osmolovskaya et al., 1990).

Growth experiments were performed in triplicates using Hungate tubes with the medium of the same composition as used for the pure culture isolation. To determine optimal growth conditions, strain 3443-3Ac<sup>T</sup> was cultivated anaerobically at various temperatures (22–65°C) and pH values 4.3–7.5. The pH in the medium was adjusted with solutions of HCl (3N) or NaOH (2N). To test the reaction of the new isolate to NaCl concentration, it was anaerobically cultivated at NaCl concentrations from 0 to 30 g L<sup>-1</sup> with intervals of 5 g L<sup>-1</sup>. To determine strain 3443-3Ac<sup>T</sup> Na<sup>+</sup>-dependence sodium-deficient medium was prepared by replacing sodium sulfide with 3 mM cysteine-HCl.

The utilization of substrates other than H<sub>2</sub> was tested with 100% CO<sub>2</sub>-filled headspace at 50°C and pH 6.0. Growth experiments for determination of CO consumption were performed exactly as described for *Thermoanaerobacter kivui* by Weghoff and Müller (2016). The utilization of CO was tested with 10, 20, 30, 40, 50, 60, 70, and 100% CO in CO<sub>2</sub>-filled headspace. Electron acceptors were tested in the medium under H<sub>2</sub>/CO<sub>2</sub> atmosphere or in the medium containing glucose, acetate, pyruvate, ethanol, methanol, or formate under CO<sub>2</sub>-filled headspace at 50°C and pH 6.0. Soluble substrates and electron acceptors were added from sterile anaerobic stock solutions before inoculation. Insoluble substrates and electron acceptors were added directly into the tubes with liquid medium prior to sterilization. Utilization of substrates or electron acceptors was monitored by increase in optical density due to microbial growth and by decrease of electron donor or acceptor concentration in the medium during growth.

Strain 3443-3Ac<sup>T</sup> was deposited in the DSMZ (German Collection of Microorganisms and Cell Cultures) and VKM (All-Russian Collection of Microorganisms) under accession numbers DSM 108286 and VKM B-3415, respectively.

### 2.2. Analytical methods

Sulfide was measured colorimetrically with N,N-dimethyl-*p*-phenylenediamine (Trüper and Schlegel, 1964). Sulfate, nitrate, nitrite, perchlorate, and chloride were analyzed with a Stayer liquid chromatograph (Aquilon) equipped with an IonPack AS4-ASC column (Dionex) and conductivity detector; the eluent was bicarbonate (1.36 mM)/carbonate (1.44 mM) at a flow rate of 1.5 ml min<sup>-1</sup>. H<sub>2</sub> and CO<sub>2</sub> were analyzed with Chromatec Crystal 5000.2 gas chromatograph (Chromatec) equipped with a NaX zeolite 60/80 mesh 3 m × 2 mm column (Chromatec) for H<sub>2</sub> and an Hayesep Q 80/100 mesh 3 m × 2 mm column (Chromatec) for CO<sub>2</sub>. Acetate, formate, propionate, butyrate, methanol, ethanol, propanol, isopropanol, butanol, isobutanol were assayed using two methods. The first method was the same gas chromatograph with the flame ionization detector (FID) and capillary column Optima FFAPplus 0.25 μm × 0.32 mm × 30 m (Macherey-Nagel) with argon as the carrier gas. Separation was carried out with temperature programming. For gas chromatography, the samples (0.2 ml) were pre-treated by centrifugation at 12600 × g for 2 min, followed by acidification of clear supernatants with 5 M formic acid to the pH of 2.0. The detection limit of the method was 0.2 mM. The second method for the determination of volatile fatty acids and alcohols was HPLC on a Stayer gradient chromatograph (JSC Aquilon, Russia) equipped with a 7.8 mm × 300 mm Resex



ROA column. The eluent was 0.5 mM HNO<sub>3</sub>; the flow rate was 1.5 mL/min. A conductometer was used as the detector.

Cellular fatty acid (CFA) profiles were determined by GC-MS (Thermo Scientific Trace GC Ultra DSQ II, HP-5MS column, E<sub>1</sub> 70 eV) of methyl ester derivatives prepared from 10 mg of freeze-dried biomass treated by anhydrous HCl/MeOH (Slobodkina et al., 2020). Polar lipids were determined by 2D-TLC as described earlier (Slobodkina et al., 2013). Respiratory lipoquinones were extracted with cold acetone from cells disrupted by grinding in liquid N<sub>2</sub> and further separated by thin-layer chromatography (TLC). The excised bands were analyzed by tandem mass spectrometry (LCQ ADVANTAGE MAX) and the compounds were identified by their ionized masses.

### 2.3. Activity assays in cell extract

Cells of strain 3443-3Ac<sup>T</sup> were suspended in 50 mM Tris-HCl buffer (pH 7.5) containing 120 mM KCl. The cells lysate was obtained using the ultrasonic disintegrator Soniprep 150 Plus (150 W, 23 kHz). The lysate was centrifuged for 30 min (12,100 g; 4°C), and the supernatant (cell extract) was stored in anoxic conditions (flushed with sterile 100% CO<sub>2</sub>) at +4°C until use (not more than 3 days). The protein concentration, measured using Qubit Protein Assay Kit (Thermo Fisher Scientific, USA), was 1.87 mg ml<sup>-1</sup>.

Glucose and cellobiose consumption by strain 3443-3Ac<sup>T</sup> cell extract was determined using a DNS assay (Miller, 1959). D-glucose and D-cellobiose solutions at various (50–500 µg/ml) concentrations were used to plot a calibration curve for DNS assay. The reaction mixtures (700 µl) contained 70 µl of 10 mM glucose or cellobiose in a TrisHCl buffer (50 mM, pH 7.5) and 630 µl of cell extract (the final concentration of sugars in solution was 1 mM). 1 mM glucose or cellobiose solution in TrisHCl buffer as well as cell extract without added sugars were used as controls. All controls and experiments were set in three replications. The reaction mixtures were incubated at the strain 3443-3Ac<sup>T</sup> optimal growth temperature (50°C) for 9 h followed by determination of reducing sugars formation by DNS assay (Miller, 1959).

Hexokinase activity was measured spectrophotometrically using a coupled assay with glucose 6-phosphate dehydrogenase. The reaction rate was followed by the production of NADPH at 340 nm. The assay mixture (2.7 ml) contained 50 mM TrisHCl, pH 8.5; 25 mM NaCl; 2 mM MgCl<sub>2</sub>; 2 mM ATP; 3 mM glucose; 0.8 mM NADP<sup>+</sup>; and 1 U glucose-6-phosphate dehydrogenase from baker's yeast (*S. cerevisiae*) type VII, 3.2 M ammonium sulfate suspension (Cáceres et al., 2003). The reaction was started by the addition of strain 3443-3Ac<sup>T</sup> cell extract (300 µl). The same reaction mixtures (3 ml) without the cell extract or substrate were used as the control assays.

### 2.4. Genome sequencing and assembly

The extraction of genomic DNA was performed as described by Park (2007). The genomic DNA was sonicated with a CovarisS2 device with parameters adjusted to obtain 400 bp fragments. Then DNA fragments were used for the preparation of fragment libraries using the NEBNext<sup>®</sup> Ultra<sup>™</sup> II DNA Library Prep

Kit for Illumina<sup>®</sup> (E7645, NEB) according to the manufacturer's instructions. The library was sequenced on the Illumina HiSeq2500. As a result, 6,593,610 single end reads were obtained (in total 1,648.4 Mb). For nanopore sequencing, the DNA was re-purified using MagAttract HMW DNA Kit (Qiagen, Germany) to enrich high molecular weight DNA fraction and remove co-purified contaminants. Genomic library for nanopore sequencing was prepared with Rapid Barcoding Kit (SQK-RBK004, Oxford Nanopore Technologies, UK). Sequencing was performed with FLO-MIN-106D flow cell (R9.4.1). Basecalling was performed using Guppy basecaller v2.3.5 with flipflop model and resulted in 128,379 reads with a median read length of 611 nucleotides (total of about 167.4 Mb).

Primary *de novo* assembly was obtained by Canu assembler v. 1.9, using only long nanopore reads (Koren et al., 2017). Polishing of the primary assembly was performed both with raw Nanopore reads by Nanopolish (Loman et al., 2015) and correction of the polished assembly with Illumina reads by Pilon (Walker et al., 2014). As a result an assembly consisting of 1 linear contig of approximately 2.24 Mbp total length was obtained. For the assembly improvement, hybrid *de novo* assembly was made by SPAdes assembler version 3.13.0 (Prjibelski et al., 2020) using both long and short reads, resulting in 2.3 Mbp total assembly with N50 of 1,678,036 bp. Manual curation and visualization of two assemblies, performed with CLC genomic workbench 3.10.0 (Qiagen, Germany), allowed one circular chromosome of 2,267,618 bp length to be obtained. Start of the chromosome was set to origin of replication, predicted by DoriC 10.0 web server (Luo and Gao, 2019). Genome sequence, as well as related project information and sample details, were deposited in NCBI database under accession numbers CP060096, PRJNA647162, and SAMN15577649, respectively. The genome sequence was also deposited in IMG with genome ID 2860381621.

### 2.5. Genome annotation

Primary genome annotation was performed by Prokaryotic Genome Annotation Pipeline (PGAP, Tatusova et al., 2016) during the process of genome submission to NCBI submission portal. Additional annotations aimed to improve predictions of protein function were performed using IMG/MER System (Chen et al., 2017) and Rapid Annotation using Subsystem Technology web server (Brettin et al., 2015) and manual curation of key metabolic genes. Refining of the automated annotations and other predictions were done manually as described by Toshchakov et al. (2015). Prediction of mobile elements was performed with *isescan* package (Xie and Tang, 2017), identification and prediction of phage-related genes was done with PHASTER server (Arndt et al., 2016). Prediction of genomic islands (GI) was performed by IslandViewer 4 web tool (Bertelli et al., 2017). Additional annotation of defense systems and biosynthetic gene clusters was performed by defense-finder tool (Tesson et al., 2022) and antiSMASH (Blin et al., 2019) web server, respectively.

Gene gain and loss that occurred during the evolution of *Thermoanaerobacterium*, *Thermoanaerobacter*, *Caldanaerobacter* phylogenetic cluster was analyzed using Count (Csurös, 2010). Short version of phylogenetic tree based on bac120 (see below)

was used as a reference tree. All the proteins involved in the gene gain/loss analysis were listed as their COG identifiers, assigned by the IMG during annotation. Rates in Count were optimized with default parameters.

## 2.6. Phylogenetic and phylogenomic analyses

Phylogenetic analysis based on the bac120 set of conserved proteins (Parks et al., 2017) was performed as follows. All available genomes of representatives of Bergey's-based order *Thermoanaerobacterales* and a large set of reference genomes affiliated to GTDB-based p\_Firmicutes\_A phylum (containing closest relatives of strain 3443-3Ac<sup>T</sup>: *Thermoanaerobacterium*, *Thermoanaerobacter*, *Caldanaerobacter* spp.) and to GTDB-based p\_Firmicutes phylum (as an outgroup), as well as the genome of *Melioribacter roseus* P3M-2<sup>T</sup> (as a root) were downloaded from Genbank. The protein sequences were identified and aligned with the GTDB-tk v.1.3.0 (Parks et al., 2018). The resulting alignment was treated using the Gblock v.0.91b with the gentlest parameters and half gap elimination (Castresana, 2000). The phylogenetic analysis was performed in the RAXML v. 8.2.12 (Stamatakis, 2014) with the PROTGAMMAILG model of amino

acid substitution. Local support values were 1,000 rapid bootstrap replications.

For studies of the evolutionary origin of acetogenic metabolism in our isolate, we constructed phylogenetic trees of proteins involved in acetogenesis. As a source of reference proteins, we chose representative genomes of GTDB rs202 (47,894 genomes). GTDB places emphasis on inclusion of as much high-quality sequence data as possible and on systematization, i.e., provision of full taxonomic assignments from species to domains using uniform approaches of genome-based phylogenetic classification (Parks et al., 2018, 2020). The GTDB taxonomy is in reasonable accord with traditional taxonomy accepted by microbiologists, and the two taxonomies have been gradually getting closer. GTDB representative genomes represent all groups of prokaryotes, both cultured and yet uncultured, that the GTDB team considers to be species based on the phylogenetic criteria applied (one genome for each GTDB species). Thus, GTDB representative genomes present a compact and reasonably adequate sampling of prokaryotic diversity.

Representative genomes of GTDB rs202 (47,894 genomes) were downloaded from GTDB repository<sup>2</sup> to create a local database for tblastn from NCBI-BLAST+ package (Camacho et al., 2009),

<sup>2</sup> <https://data.gtdb.ecogenomic.org/>

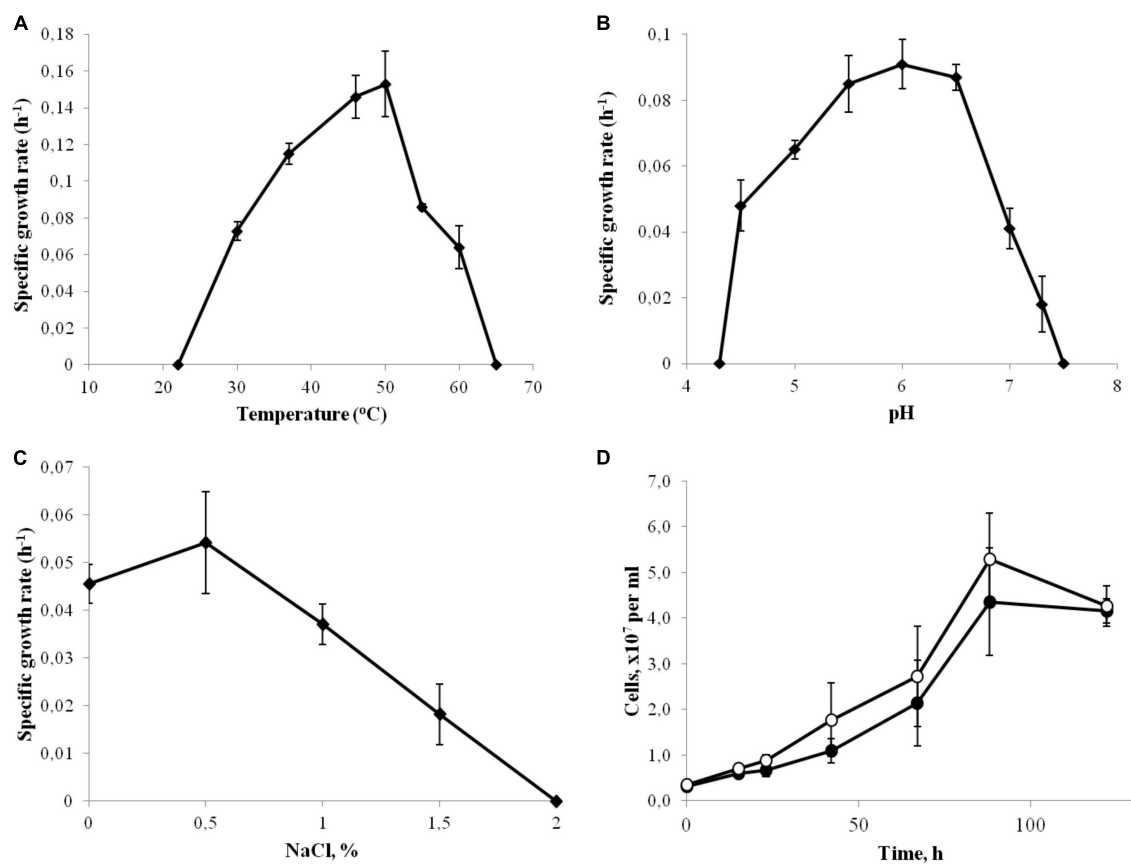


FIGURE 1

Effect of temperature (A), pH (B) and NaCl concentrations (C) on the growth rate of strain 3443-3Ac<sup>T</sup>. (D) Effect of Na<sup>+</sup> on the growth of strain 3443-3Ac<sup>T</sup>; cultures grown on H<sub>2</sub> and CO<sub>2</sub> were passaged in Na<sup>+</sup>-enriched (●) and Na<sup>+</sup>-deficient (○) minimal medium, fifth passage is shown. Error bars indicate standard deviations of biological replicates ( $n = 3$ ).

operated at default values of most parameters (the default for *e*-value is 10). For the *-num\_alignments* and *-num\_descriptions* options, the value of 50,000 was set.

The queries for *tblastn* were the proteins of our strain 3443-3Ac<sup>T</sup> and/or the proteins of *T.kivui*. The *RnfD* and *RnfC* queries were Awo\_c22050 and Awo\_c22060 from *Acetobacterium woodii*. For convenience of further analysis of gene co-occurrence and co-location, the queries were usually concatenated (up to 17 subqueries separated by strings of 100 “x” letters). The cut-off values for identification of *tblastn* hits were usually 25% identity with a particular subquery at 80% coverage. In case of the catalytic subunit of energy-converting hydrogenase (EchE), the identity cut-off value was 28% and there was an additional requirement for the presence of two CxxC nickel binding motifs to distinguish the EchE catalytic subunits from homologous subunits of membrane-bound oxidoreductases (Vignais and Billoud, 2007). Hit sequences extracted from *tblastn* output with home-made software (see below) were checked for validity of gene identification by submitting them to HydDB (Søndergaard et al., 2016) for EchE (9 of 4,275 sequences had to be discarded) or HMMSCAN (Potter et al., 2018). An additional confirmation of the validity of gene identification was the expected clustering of certain hits in the genome (although particular organization of the gene clusters could vary).

The analysis of *tblastn* output (namely, the complete set of alignments provided by *tblastn* was used) was performed by a home-made software, whose earlier version was successfully used by us to analyze genomic contexts of CO dehydrogenase genes in microbial genomes (Teichtmann et al., 2012). The software provided a report on gene co-location (within a user specified distance, 15 kb by default), and also on a tighter co-location of genes (0.5 kb). The software also extracted hit sequences directly from the *tblastn* output and printed them to files. The names of the extracted sequences were formulated so as to include information on the host genome GTDB taxonomic string and the genomic context revealed. These sequences were used to construct Neighbor Joining (NJ) and Maximum Likelihood (ML) trees. The protein sequences extracted from *tblastn* alignments somewhat varied in their completeness (default query coverage requirement was 80%); however, the variable terminal regions that escaped analysis would have been anyway neglected in the process of tree construction, where the parameters that we used included deletion of alignment positions with coverage below 95%.

The phylogenetic trees of proteins involved in acetogenesis were constructed as follows. For “total” trees, all sequences of the target proteins extracted from the output of *tblastn* against the database of GTDB rs202 representative genomes were aligned with MAFFT version 7 (Katoh et al., 2019), and NJ and/or ML trees were constructed with MEGA6 (Tamura et al., 2013). For smaller trees, 25 best hits of *A.autotrophica*'s proteins and 25 best hits of *T.kivui*'s proteins were taken after *tblastn* in representative genomes of GTDB rs202. Homologous proteins of TTC group representatives from the same database were added in cases where the same enzymatic activity could be assumed. Redundant proteins were discarded with CD-HIT (Fu et al., 2012) at 100% cut-off value. The ML trees with 100 bootstrap replicates were constructed with MEGA6 after alignment with built-in ClustalW.

Since GTDB taxonomy is firmly phylogeny-grounded, the inclusion of the GTDB taxonomic string in sequence names in

trees allows one to do without juxtaposition of a phylogenetic tree constructed for a particular protein with a tree showing the phylogeny of the core genes of the host organism (since the information on the host phylogeny is provided by the sequence name). Conclusions on vertical inheritance or HGT events can be drawn without such juxtaposition.

## 3. Results and discussion

### 3.1. Physiological characterization of strain 3443-3Ac<sup>T</sup>

Strain 3443-3Ac<sup>T</sup> was isolated from a sample of sediments collected from the terrestrial hot spring Kaskadny at East Thermal Field, Uzon Caldera, Kamchatka, Russia (N54° 30.026' E160° 00.374', elevation 658 m). The isolation was performed in anaerobic medium in the presence of H<sub>2</sub> as the energy source and HCO<sub>3</sub><sup>-</sup>/CO<sub>2</sub> as the carbon source and electron acceptor.

Cells of strain 3443-3Ac<sup>T</sup> were rod-shaped, 0.5–0.7 μm in diameter and 1.5–5 μm in length; occurred singly, in pairs, or in chains depending on the growth phase and growth conditions. No motility was observed. The cells examined in negatively stained specimens (Supplementary Figure 1A) did not exhibit flagella. Strain 3443-3Ac<sup>T</sup> formed round terminal endospores after prolonged (10 days) incubation at 50°C. The percentage of cells with spores, though, was no more than 1–2%. Gram-stain reaction was positive and ultrathin sections of the cells (Supplementary Figure 1B) revealed a Gram-positive cell wall type.

Strain 3443-3Ac<sup>T</sup> was a strictly anaerobic bacterium growing at low-potential redox conditions: no growth occurred in a sodium sulfide-free medium. The temperature range for growth was 30–60°C, with an optimum at 46–50°C (Figure 1A). No growth was detected at 65°C or above, as well as at 22°C or below after incubation for 10 days. The pH range of growth was 4.5–7.3, with an optimum at pH 6.0 (Figure 1B). Growth was not observed at and below pH 4.3 and at and above pH 7.5. The isolate grew at NaCl concentrations of up to 1.5% NaCl (Figure 1C), but optimal growth was observed on the basal medium, where the concentrations of Na<sup>+</sup> and Cl<sup>-</sup> were 0.1 and 0.06%, respectively. No growth was observed at and above 2.0% NaCl (w/v). In order to check for Na<sup>+</sup> dependence of strain 3443-3Ac<sup>T</sup> during autotrophic growth, cells were cultivated in five consecutive passages in a Na<sup>+</sup>-depleted minimal medium (the contaminating amount of Na<sup>+</sup> was less than 100 μM) or a Na<sup>+</sup>-enriched medium containing 70 mM NaCl. The growth rate was similar in all passages indicating that Na<sup>+</sup> concentration did not affect growth. The growth rate and yield in the fifth passage is shown on Figure 1D. Strain 3443-3Ac<sup>T</sup> was able to grow in a vitamin-free medium; however, the growth yield and rate were about 30% lower in this case. The doubling time under the optimal growth conditions (60°C, pH 6.0, no additional NaCl) was 4.5 h.

Strain 3443-3Ac<sup>T</sup> grew chemolithoautotrophically, optimally with H<sub>2</sub> as the energy source and HCO<sub>3</sub><sup>-</sup>/CO<sub>2</sub> as the carbon source and electron acceptor (Figure 2A). Acetate was the only product of the growth. The isolate also used formate (25 mM) as the energy source, but the growth yield was lower compared to growth on H<sub>2</sub> (Figure 2B). The main end products during growth with formate were acetate and CO<sub>2</sub>, but small amounts of

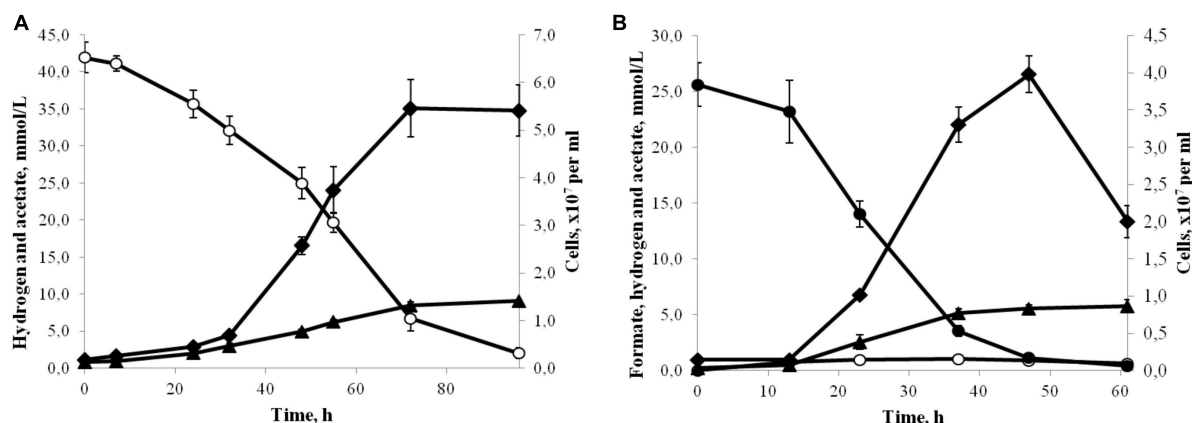


FIGURE 2

(A) Growth (◆) and production of acetate (▲) by strain 3443-3Ac<sup>T</sup> in medium under H<sub>2</sub>(○)/CO<sub>2</sub> at 50°C and pH 6.0. (B) Growth (◆) and production of acetate (▲) and H<sub>2</sub>(○) by strain 3443-3Ac<sup>T</sup> in medium with formate (●) at 50°C and pH 6.0. Error bars indicate standard deviations of biological replicates ( $n = 3$ ).

hydrogen were also formed. Formate was converted to acetate (and CO<sub>2</sub>) with a ratio of 4.2:1. The hydrogen yield reached 1 mmol L<sup>-1</sup> culture in the exponential phase of growth, but decreased to 0.5 mmol L<sup>-1</sup> in the stationary phase. Growth experiments for determination of CO consumption were performed exactly as described for *Thermoanaerobacter kivui* by Weghoff and Müller (2016). The cells were pre-grown on H<sub>2</sub> and CO<sub>2</sub> and transferred to the medium with 10% CO in the gas phase (the gas phase contained CO<sub>2</sub> and did not contain H<sub>2</sub>). By increasing the CO concentration in 10% increments, it was possible to adapt strain 3443-3Ac<sup>T</sup> to grow on up to 50% CO (the gas phase contained also 50% CO<sub>2</sub>). Acetate was the main product of the growth on CO, but small amounts of hydrogen were also formed. Strain 3443-3Ac<sup>T</sup> was unable to utilize yeast extract, beef extract, peptone, pectin, dextrin, starch, cellulose, xylan (0.2, 1.0, or 2.0 g L<sup>-1</sup> each), glucose, fructose, galactose, mannose, arabinose, rhamnose, xylose, ribose, trehalose, cellobiose, sucrose, lactose, maltose, raffinose, sorbitol, mannitol (0.5, 2.0 or 5.0 g L<sup>-1</sup> each), acetate, lactate, pyruvate, malate, propionate, butyrate, fumarate, succinate, citrate, ethanol, propanol, glycerol, methanol (5, 20, or 40 mM each), 2-methoxyphenol, 3,4-dimethoxybenzoate, or 2-methoxybenzoate (1, 5, or 10 mM each). With H<sub>2</sub>/CO<sub>2</sub> (80:20) or formate (20 mM), strain 3443-3Ac<sup>T</sup> did not use nitrate, sulfate, perchlorate, fumarate, Fe (III) citrate (10 mM each), nitrite, sulfite (both 5 mM), or O<sub>2</sub> (2 or 5%) as the electron acceptors. With H<sub>2</sub>/CO<sub>2</sub> (80:20) or formate (20 mM) sulfide formation was observed on medium with thiosulfate (10 mM) or elemental sulfur (10.0 g L<sup>-1</sup>), but acetate was the main product of growth, and growth rate and cell yield were the same as in the absence of these acceptors. No growth was observed on glucose (1.0 g L<sup>-1</sup>) in the presence of thiosulfate (10 mM) or elemental sulfur (10.0 g L<sup>-1</sup>).

To confirm the obligatorily autotrophic nature of strain 3443-3Ac<sup>T</sup>, we measured the consumption of glucose and cellobiose in its cell extracts (Supplementary Table 1). Glucose and cellobiose were chosen because both sugars can be viewed as starting points of glycolysis [substrates for hexokinase or a phosphotransferase system (PTS)], and because all of the genes encoding glycolysis/gluconeogenesis enzymes, including

hexokinase, and a gene of cellulose-phosphorylase (GH94) domain-containing protein but not PTS genes were found in the genome of strain 3443-3Ac<sup>T</sup> (see Supplementary Table 2 and section “3.6. The likely causes for the inability of strain 3443-3Ac<sup>T</sup> to grow heterotrophically”). Moreover, both glucose and cellobiose are used by all of the closest relatives of strain 3443-3Ac<sup>T</sup> [representatives of the genera *Thermoanaerobacterium*, *Thermoanaerobacter*, *Caldanaerobacter*, *Calorimonas*, and *Caldanaerobius* (see section “3.2. Phylogenetic position and taxonomic status of strain 3443-3Ac<sup>T</sup>)”]. No consumption of glucose or cellobiose by the cell extract of strain 3443-3Ac<sup>T</sup> was observed. The lack of cellobiose consumption is consistent with the results of strain 3443-3Ac<sup>T</sup> genome analysis, which showed almost total absence of PTS genes, as well as of beta-glucosidase encoding genes (see section “3.6. The likely causes for the inability of strain 3443-3Ac<sup>T</sup> to grow heterotrophically”). The presence of a hexokinase gene (Supplementary Table 2), prompted us to look for the respective activity in strain 3443-3Ac<sup>T</sup> cell extract, and this activity was not detected. Altogether, the results of the tests with cell extracts support our conclusion about the obligatorily autotrophic nature of strain 3443-3Ac<sup>T</sup>.

Analysis of cellular fatty acids revealed iso-C<sub>15:0</sub> (52.8% of the total fatty acids) and C<sub>16:0</sub> (40.4%) as the major components. The following fatty acids were also detected: iso-C<sub>17:0</sub> (4.8%), C<sub>18:0</sub> (0.9%), iso-C<sub>16:0</sub> (0.8%), and anteiso-C<sub>17:0</sub> (0.4%) (Supplementary Table 3). Polar lipids of strain 3443-3Ac<sup>T</sup> included four unidentified phospholipids and an unknown aminophospholipid (possibly, similar to phosphatidylethanolamine, but with larger R<sub>f1</sub> and lower R<sub>f2</sub> values). Thin-layer chromatography coupled to mass spectrometry (TLC-MS) revealed no respiratory quinones.

### 3.2. Phylogenetic position and taxonomic status of strain 3443-3Ac<sup>T</sup>

The genome of strain 3443-3Ac<sup>T</sup> (see below) contains four almost identical (99.8–100%) 16S rRNA genes. BLAST search revealed 93–94% sequence identity to homologous genes of



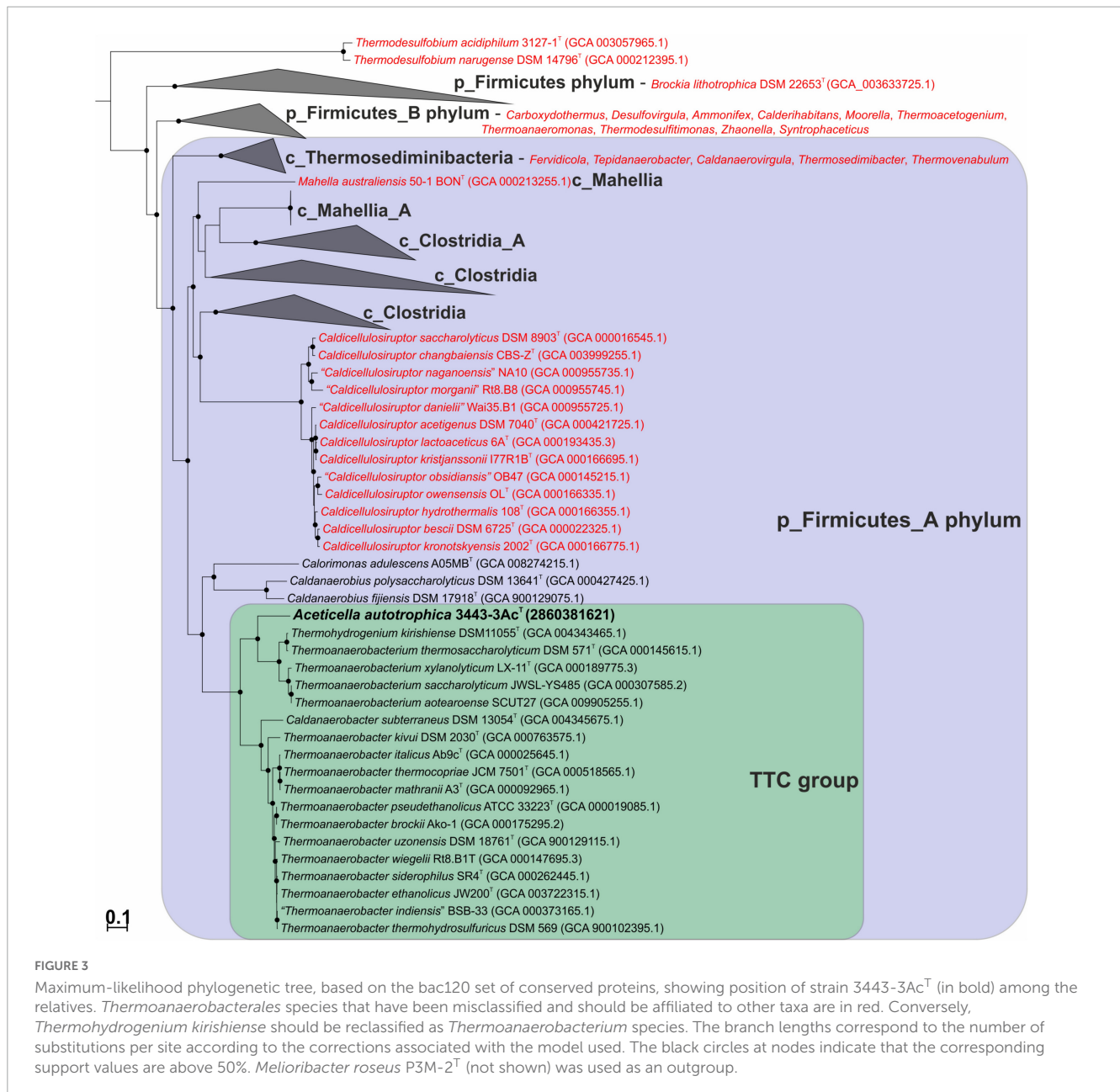


FIGURE 3

Maximum-likelihood phylogenetic tree, based on the bac120 set of conserved proteins, showing position of strain 3443-3Ac<sup>T</sup> (in bold) among the relatives. *Thermoanaerobacteriales* species that have been misclassified and should be affiliated to other taxa are in red. Conversely, *Thermohydrogenium kirishiense* should be reclassified as *Thermoanaerobacterium* species. The branch lengths correspond to the number of substitutions per site according to the corrections associated with the model used. The black circles at nodes indicate that the corresponding support values are above 50%. *Melioribacter roseus* P3M-2<sup>T</sup> (not shown) was used as an outgroup.

representatives of the genus *Thermoanaerobacterium*, and 86–88% sequence identity to those of *Thermoanaerobacter* spp. and *Caldanaerobacter* spp.

In the latest version (as of April 2022) of *Bergey's Manual of Systematics of Archaea and Bacteria* (BMSAB) (see text footnote 3), the genus *Thermoanaerobacterium* is assigned to *Thermoanaerobacteriales* Incertae Sedis Family III, whereas the genera *Thermoanaerobacter* and *Caldanaerobacter* are in the family *Thermoanaerobacteraceae* within the order *Thermoanaerobacteriales*, which in addition include 8 and 12 other genera, respectively. The development of phylogenomic analyses based on ribosomal and other conserved protein sequences opens new windows into microbial taxonomy. In the GTDB (Release 07-RS207) bacterial taxonomy, which is based on concatenated alignment of 120 ubiquitous, single-copy proteins, covering hundreds of thousands genomes (Parks et al., 2018, 2020), only

three genera *Thermoanaerobacterium*, *Thermoanaerobacter*, and *Caldanaerobacter* comprise the family *f\_Thermoanaerobacteraceae* within the order *o\_Thermoanaerobacteriales*. The structure of higher taxa in GTDB is strongly different from that in the more traditional BMSAB taxonomy, and, whereas the GTDB *o\_Thermoanaerobacteriales* is within class *c\_Thermoanaerobacteria* in phylum *p\_Firmicutes\_A*, most of the genera assigned to order *Thermoanaerobacteriales* in BMSAB fall into other higher taxa in GTDB, such as class *c\_Thermosediminibacteria* and various classes within phylum *p\_Firmicutes\_B*.

We constructed a phylogenetic tree based on the bac120 set of conserved proteins for a collection of bacteria that included our strain 3443-3Ac<sup>T</sup>, all of the BMSAB order *Thermoanaerobacteriales* representatives with available genomes, and representatives of the GTDB phyla *p\_Firmicutes* and *p\_Firmicutes\_A* (see Methods



for further details). The tree that we constructed (Figure 3) agreed on the whole with the GTDB phylogenetic taxonomy and showed that our strain 3443-3Ac<sup>T</sup> belonged to a monophyletic clade that included *Thermoanaerobacterium*, *Thermoanaerobacter* and *Caldanaerobacter* representatives (hereafter, TTC group, corresponding to the GTDB family f\_Thermoanaerobacteraceae), as well as representatives of the genera *Calorimonas* and *Caldanaerobius*. Strain 3443-3Ac<sup>T</sup> formed a separate genus-level branch within the TTC cluster.

Thus, based on phenotypic (Table 1), genotypic and phylogenetic characteristics of our isolate, strain 3443-3Ac<sup>T</sup> is considered to represent a new species of a new genus, for which the name *Aceticella autotrophica* gen. nov., sp. nov. is proposed.

### 3.3. General genome properties, genome mobility, and genomic islands

Hybrid sequencing and *de novo* genome assembly allowed the full sequence of the 3443-3Ac<sup>T</sup> circular chromosome to be obtained. The chromosome length was 2,267,618 nucleotides. *In silico* determined G+C content of the genomic DNA of strain 3443-3Ac<sup>T</sup> was 33.0 mol%. Automatic gene calling and annotation resulted in the prediction of 2234 total genes including 2093 protein-coding, 72 RNA genes, and 69 pseudogenes.

The analysis of COG distribution in the proteome of strain 3443-3Ac<sup>T</sup> with comparison to TTC group representatives, performed with STAMP software (Parks et al., 2014), showed remote positioning of strain 3443-3Ac<sup>T</sup> relative to other TTC group genera (Figure 4), reflecting its metabolic peculiarities.

Analysis of genome mobility-related features showed that strain 3443-3Ac<sup>T</sup> can be characterized by the largest proportion of IS-element-related genes among TTC group representatives taken for comparative analysis (Supplementary Figure 2). Analysis of phage-related genes showed at least two incomplete prophage regions of 6.9 and 11.4 kb, possessing 75 and 62% of phage-associated genes, respectively (Supplementary Figure 3).

In accordance with the high level of predicted genome mobility, analysis of laterally acquired genomic regions revealed 13 genomic islands (GI), spanning over 330 kb (Supplementary Figure 3 and Supplementary material 2). The COG functional categories distribution showed that these laterally transferred loci were specifically enriched with mobile elements and genes involved in defense mechanisms (Supplementary material 3). Specifically, GI2 and GI3 possessed two Cas operons of types IB and IIIB, respectively. GI4 includes two restriction-modification systems of classes I and II (ACETAC\_01575-ACETAC\_01585 and ACETAC\_01520-ACETAC\_01530, respectively). GI5 possesses a gene for homodimeric lactococcal bacteriocin - lactococcin 972 (Martínez et al., 1996). GI6 codes for a four-gene Wadjet system (Doron et al., 2018), GI7 possesses phosphorothioation-sensing bacterial defense system SSpBCDE (Xiong et al., 2020). The large GI10 includes several genes for bacteriocin-processing C39 peptidases (Kanonenberg et al., 2013), as well as bacteriocin genes and s toxin/antitoxin systems (Supplementary material 2).

Search for CRISPR-Cas genomic immunity modules performed with cctyper web server (Russel et al., 2020) discovered five Cas operons; however, only one of them (IB type, ACETAC\_00735-

ACETAC\_00770), possessing full adaptation and interference modules, as well as adjacent CRISPR repeat, seems to be functional (Supplementary Figure 4 and Supplementary Table 4). Others were either heterogeneous, located in the genomic islands (see above), or disrupted by IS-element mediated rearrangements.

### 3.4. Genomic determinants of acetogenesis in strain 3443-3Ac<sup>T</sup>

The genome of strain 3443-3Ac<sup>T</sup> encoded all the proteins (Supplementary Table 5) needed for chemolithoautotrophic acetogenic growth with hydrogen (Figure 5). The electron-donating reactions (the first module of acetogenesis) are implemented by the soluble electron-bifurcating [FeFe]-hydrogenase (HydABC complex, ACETAC\_10965-ACETAC\_10975), which is homologous (47–58% amino acid sequence identity of subunits) to the biochemically characterized bifurcating hydrogenases of the acetogenic bacteria *Acetobacterium woodii* and *Moorella thermoacetica* (Schuchmann and Müller, 2012; Wang et al., 2013), and still more similar (79–84% identity of subunits) to the HydABC of the acetogenic *Thermoanaerobacter kivui*, for which the properties have been deduced based on the homology with the aforementioned proteins from *A. woodii* and *M. thermoacetica* (Hess et al., 2014). Apart from the HydABC complex and HydA2, a component of the hydrogen-dependent CO<sub>2</sub> reductase (see below), there are no other uptake hydrogenases encoded in the genome of strain 3443-3Ac<sup>T</sup>. As for H<sub>2</sub>-producing hydrogenases, the genome contains genes for Ni-containing membrane-bound energy-converting hydrogenase, which generates ion-motive force in the course of reduction of protons to molecular hydrogen (see below).

Electrons coming from molecular hydrogen are transferred by the electron-bifurcating hydrogenase to the universal electron carriers NAD<sup>+</sup> and ferredoxin. A portion of the reduced ferredoxin is used in the reductive reactions of Wood-Ljungdahl pathway (the second module of acetogenesis), and the remaining portion is used to generate a chemiosmotic gradient (the third module). One CO<sub>2</sub> molecule is reduced to a methyl group in the methyl branch of the WLP and another CO<sub>2</sub> molecule is reduced to CO in the carbonyl branch. The first step in the methyl branch is the reduction of CO<sub>2</sub> to formate. We predict that in strain 3443-3Ac<sup>T</sup> this reaction is carried out by a hydrogen-dependent CO<sub>2</sub> reductase (HDCR), an enzyme discovered in *Acetobacterium woodii* (Schuchmann and Müller, 2013) and then biochemically characterized also in *Thermoanaerobacter kivui* (Schwarz et al., 2018). This enzyme complex was shown to catalyze the reduction of CO<sub>2</sub> with electrons coming directly from molecular hydrogen. The complex includes a hydrogenase subunit and a formate dehydrogenase subunit, which are connected by two electron-transferring subunits (Schuchmann and Müller, 2013; Schwarz et al., 2018). In both *Acetobacterium woodii* and *Thermoanaerobacter kivui*, the four genes encoding subunits of HDCR are located within one gene cluster. Both gene clusters also include a gene annotated as *fdhD*, encoding a protein with unknown but probably regulatory function. In the genome of strain 3443-3Ac<sup>T</sup> the genes that we predict to encode HDCR subunits are not located in a single gene cluster. Genome analysis revealed the presence

TABLE 1 The principal characteristics distinguishing the genus *Aceticella* gen. nov. from the genera *Thermoanaerobacterium* [data from Yamamoto et al. (1998), Romano et al. (2010), Shang et al. (2013), Onyenwoke and Wiegel (2015), and Lopez et al. (2017)], *Thermoanaerobacter* [data from Wagner et al. (2008), Tomas et al. (2013), and Onyenwoke and Wiegel (2015)], and *Caldanaerobacter* [data from Kim et al. (2001), Kozina et al. (2010), and Fardeau et al. (2015)].

Property	<i>Aceticella</i> (1)	<i>Thermoanaerobacterium</i> (9)	<i>Thermoanaerobacter</i> (16)	<i>Caldanaerobacter</i> (2)
Isolation source	Hot springs	Geothermal sites, hot springs, canned food, dung samples	Geothermal sites, hydrothermal vent, hot springs, water and mud samples from thermal spas, geothermally heated oil samples, extraction juices of beet sugar factories, mesobiotic mud and soil, sugar refinery, household dust, compost of camel feces, spoiled meat, thermophilic reactor fed with household solid waste from a biogas plant	Hot spring, submarine hydrothermal vent, geothermal water, oil well
Cell morphology	Rod-shaped cells occur singly, in pairs or in chains depending on the growth phase and growth conditions.	Rod-shaped cells occur singly, in pairs or in chains depending on the growth phase and growth conditions	Rod-shaped cells occur singly, in pairs or in chains depending on the growth phase and growth conditions. Asymmetric cell division leads to formation of coccoid cells. Stationary growth phase cells are pleomorphic, and some form filaments.	Rod-shaped cells occur singly, in pairs or in chains depending on the growth phase and growth conditions. Cells sometimes branch.
Cell wall	Gram-positive cell-wall structure and Gram-stain- positive	Cells have Gram-positive cell-wall structure, but are Gram-stain-variable	Cells have Gram-positive cell-wall structure, but are Gram-stain-variable	Cells have Gram-positive cell-wall structure, but are Gram-stain-variable
Motility	-	+	V	V
Spores	+	V	V	V
Growth temperature (°C)				
Range	30–60	35–75	30–85	40–85
Optimum	50	50–60	55–71	65–75
Growth pH				
Range	4.5–7.3	3.8 – 8.5	4.0–9.9	4.5–9.2
Optimum	6.0	5.2 –7.8	5.8–8.5	6.5–7.5
NaCl for growth (% w/v)				
Range	0–1.5	0–3.0	0–4.5	0–4.0
Optimum	0	0.1–2.5	0.5–1	0–2.5
Chemoorganoheterotrophic growth	-	+	+	+
Chemolithoheterotrophic growth	-	nd	V	V
Chemolithoautotrophic growth	+	-	V*	-
Substrate utilization				
Glucose	-	+	+	+
Fructose	-	V	+	+
Galactose	-	V	V	+
Mannose	-	+	+	+
Xylose	-	+	V	V
Cellobiose	-	+	V	+
Lactose	-	V	V	+
Maltose	-	+	V	+
Sucrose	-	V	V	V
Pectin	-	V	V	nd
Starch	-	V	V	+
Xylan	-	+	V	V
Cellulose	-	-	V	nd

(Continued)

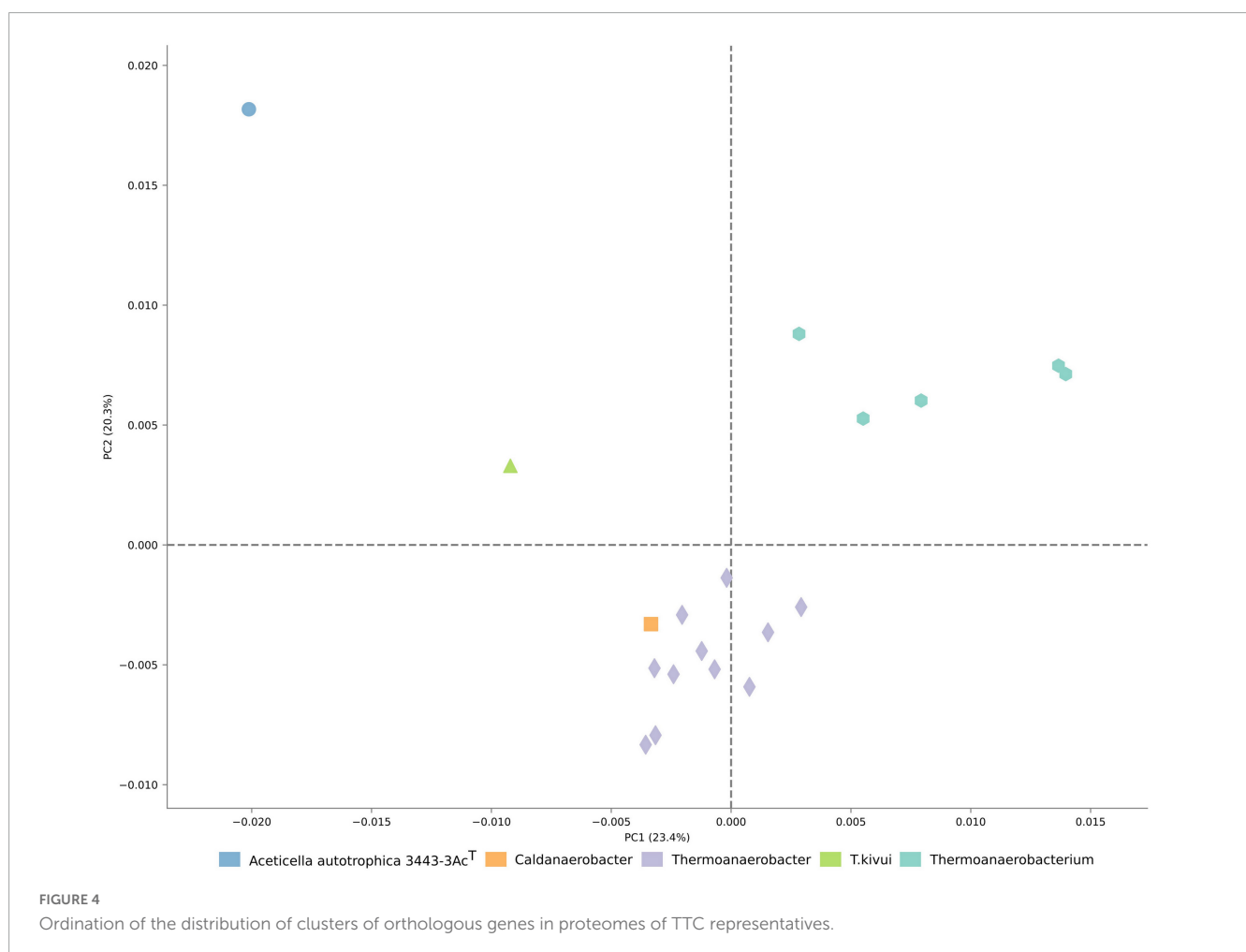
TABLE 1 (Continued)

Property	<i>Aceticella</i> (1)	<i>Thermoanaerobacterium</i> (9)	<i>Thermoanaerobacter</i> (16)	<i>Caldanaerobacter</i> (2)
Fermentation end products from hexoses	NA	Ethanol, acetate, lactate, H <sub>2</sub> and CO <sub>2</sub> (In some instances, butyrate and butanol are also formed)	Ethanol, acetate, lactate, H <sub>2</sub> and CO <sub>2</sub> (In some instances, butyrate, isobutyrate, propionate and succinate are also formed)	Acetate, H <sub>2</sub> , CO <sub>2</sub> , L-alanine, ethanol and lactate (In some instances, propionate, isobutyrate and isovalerate are also formed)
Gelatin hydrolysis	-	V	V	-
Electron acceptors				
Sulfite	-	V	V	V
Sulfur	+	V	V	V
Thiosulfate	+	+	+	+
Nitrate	-	V	V	V
Acetogenesis capacity	+	-	V**	-
DNA G+C content (mol%)	33.0	29–36.1	30–38	33–41

Numbers of species are indicated in parenthesis. +, positive; -, negative; V, characteristic varies among different species; nd, not determined; NA, not applicable.

\*Chemolithoautotrophic growth was shown only for *Thermoanaerobacter kivui* (H<sub>2</sub>+CO<sub>2</sub> and CO+CO<sub>2</sub>).

\*\*Acetogenesis capacity was shown only for *Thermoanaerobacter kivui*.



of one gene for selenium-free formate dehydrogenase, *fdhF1* (ACETAC\_00565), and two genes for selenium-containing formate dehydrogenases: *fdhF2* (ACETAC\_08370/ACETAC\_08380)

and *fdhF3* (ACETAC\_10785). Of these, the *fdhF2* gene is disrupted by a transposon, whereas *fdhF1* and *fdhF3* encode proteins 76–77% identical to the only formate dehydrogenase

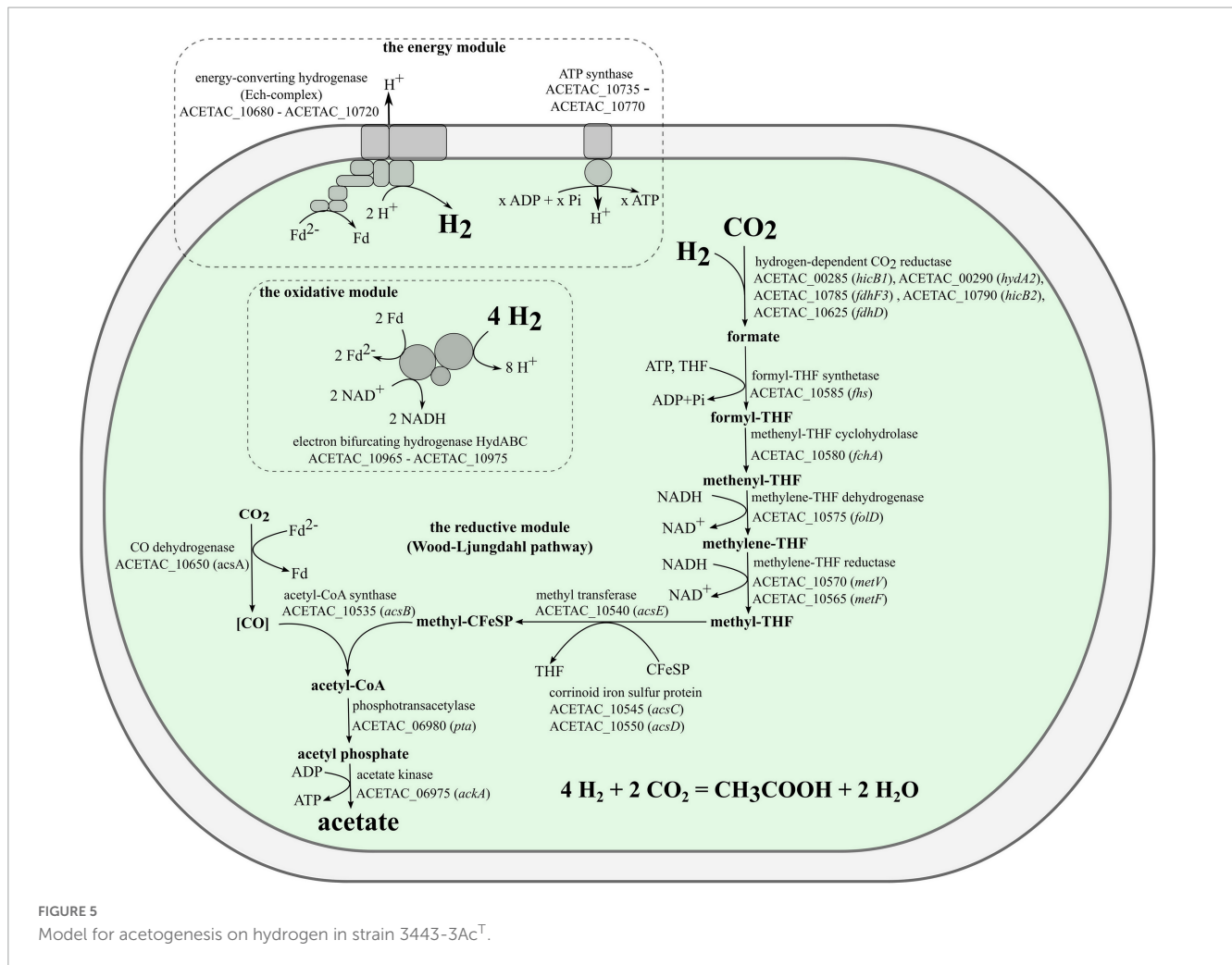


FIGURE 5  
Model for acetogenesis on hydrogen in strain 3443-3Ac<sup>T</sup>.

of *Thermoanaerobacter kivui*. A gene encoding an electron transfer protein *hycB2* (ACETAC\_10790) precedes the *fdhF3* gene. The hydrogenase module of HDCR is encoded by the hydrogenase subunit gene *hyaA2* (ACETAC\_00290) and the nearby electron-transferring subunit gene *hycB1* (ACETAC\_00285). The aforementioned *fdhD* gene of the unknown relevance to HDCR is also present in strain 3443-3Ac<sup>T</sup> genome (ACETAC\_10625), although apart from the apparently relevant genes. Despite the weakly pronounced clustering of the HDCR genes homologs in strain 3443-3Ac<sup>T</sup> genome, we predict the HDCR function based on the high identity values of the respective proteins with the subunits of the biochemically characterized HDCR of *Thermoanaerobacter kivui* (73–82%, Supplementary Table 5).

Other genes encoding functions of the WLP (with notable exception of the CO dehydrogenase gene *acsA*) are located in the 3443-3Ac<sup>T</sup> genome in one gene cluster (hereafter, the WLP gene cluster, Figure 6 and Supplementary Table 5), *acsA* being located 10.5 kb upstream. Very similar gene clusters occur in the genomes of many representatives of the class *c*\_Clostridia orders *o*\_Clostridiales and *o*\_Peptostreptococcales and of the class *c*\_Thermosediminibacteria (GTDB taxonomy), but in the genomes of those bacteria the cluster includes the CO dehydrogenase gene *acsA* (Figure 6). Most similar to our strain's WLP gene cluster in terms of its organization is the gene cluster of

*Thermoanaerobacter kivui*, which, apart from strain 3443-3Ac<sup>T</sup>, is the only TTC group member that has WLP genomic determinants. In the *Thermoanaerobacter kivui*'s genome, *acsA* (one of the two *Thermoanaerobacter kivui*'s CO dehydrogenase genes) also occurs beyond the WLP gene cluster 16.8 kb upstream (Figure 6). The genome of strain 3443-3Ac<sup>T</sup> also contains two CO dehydrogenase genes, and their phylogenies (see below) suggests that it is ACETAC\_10650 that is involved in the WLP in strain 3443-3Ac<sup>T</sup> (and thus should be termed *acsA*). The ACETAC\_02795 gene is located far away from the WLP gene cluster and encodes a CO dehydrogenase with a 96% amino acid sequence identity with the biochemically characterized monofunctional CO dehydrogenase *CooS* of *Thermoanaerobacter kivui* (Schwarz et al., 2020; Jain et al., 2021). Thus, the ACETAC\_02795 product is apparently involved in CO oxidation during growth on that substrate.

At the final acetogenesis stage, phosphotransacetylase (*pta* - ACETAC\_06980) and acetate kinase (*ackA* - ACETAC\_06975) catalyze the conversion of acetyl-CoA to acetate with ATP synthesis from ADP + Pi. However, this substrate-level phosphorylation reaction does not solve the net problem of energy conservation since the ATP molecule is spent at the stage of formate activation by formyl-tetrahydrofolate (THF) synthetase (*fhs* - ACETAC\_10585) in the methyl branch of WLP.



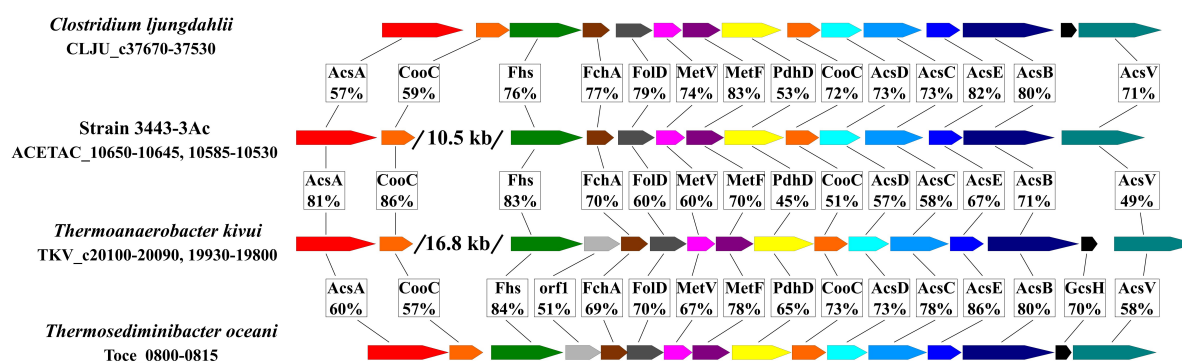


FIGURE 6

Comparison of the WLP gene cluster of strain 3443-3Ac<sup>T</sup> with the gene clusters in *Clostridium ljungdahlii*, *Thermoanaerobacter kivui* and *Thermosediminibacter oceani* in terms of gene order and amino acid sequence identity of the encoded proteins. AcsA, CO dehydrogenase; CooC, maturation factor of the CO dehydrogenase; Fhs, formyl-THF synthetase; FchA, methenyl-THF cyclohydrolase; FoLD, methylene-THF dehydrogenase; MetV and MetF, methylene-THF reductase; PdhD, dihydroliipoamide dehydrogenase; AcsD and AcsC, corrinoid iron sulfur protein; AcsE, methyl transferase; AcsB, acetyl-CoA synthase; GcsH, protein H of a glycine cleavage system; AcsV, acetyl-CoA synthase corrinoid activation protein.

For energy conservation, acetogenic bacteria use a two-modular respiration system comprising either Rnf complex or Ech complex generating the ion-motive force (IMF) and an ATP synthase using IMF for ATP synthesis. Inspection of the genome showed that strain 3443-3Ac<sup>T</sup> does not possess an Rnf complex, but has a gene cluster that we predict to encode subunits of a Ni-containing energy-converting hydrogenase (Ech complex, ACETAC\_10680 - ACETAC\_10720). This gene cluster comprises 9 genes (Supplementary Table 5) and is highly similar (in terms of encoded proteins sequence identity, gene order and phylogeny) to one of the two Ech-encoding gene clusters found in the genome of *Thermoanaerobacter kivui*, namely, to that encoded by TKV\_c01230 - TKV\_c01310 (Schoelmerich and Müller, 2019). The identity of homologous subunits of these Ech complexes of strain 3443-3Ac<sup>T</sup> and *Thermoanaerobacter kivui* is 70–92%. An interesting point, supporting the vertical inheritance of the cluster, is the presence in both of these gene clusters of genes encoding proteins (ACETAC\_10695, 74% identical to TKV\_c01280, and ACETAC\_10690, 80% identical to TKV\_c01290) whose homologs have never been reported to occur as subunits in any hydrogenases studied and that have never been noticed to occur in any other gene clusters recognized to encode hydrogenases.

The Ech complex of strain 3443-3Ac<sup>T</sup> is apparently the only candidate for the mechanism of coupling electron transfer with ion translocation in this bacterium: the genome of 3443-3Ac<sup>T</sup> does not encode an Rnf complex or a NADH-quinone oxidoreductase. Moreover, neither quinones were found by TLC-MS nor quinone biosynthetic pathways were revealed during the genome analysis. Genes encoding the subunits of a F<sub>1</sub>F<sub>0</sub> ATP synthase occur in the genome of strain 3443-3Ac<sup>T</sup> as a single gene cluster (ACETAC\_10735 - ACETAC\_10770) located upstream of the ech gene cluster (Supplementary Table 5). The ATP synthase subunit *c* of strain 3443-3Ac<sup>T</sup> (encoded by *atpE* - ACETAC\_10765), like that of *Thermoanaerobacter kivui* (Hess et al., 2014), does not exhibit the Q...ES/T Na<sup>+</sup> binding motif and features the E55 proton-binding site. Thus, we predict that strain 3443-3Ac<sup>T</sup> uses a respiration system comprising an Ech complex and an H<sup>+</sup>-dependent ATP synthase for energy conservation.

### 3.5. Genomic determinants of formate and carbon monoxide metabolism in strain 3443-3Ac<sup>T</sup>

The first step in formate metabolism is its uptake, but the genome of strain 3443-3Ac<sup>T</sup> lacked genes encoding known formate transporters (PF01226, Form\_Nir\_trans). Recently it has been shown that formate transporter is not essential for growth of *A. woodii* on a high formate concentration (50–500 mM) at pH 7.0. *A. woodii* mutant with deleted formate transporter was capable to grow under these conditions because part of the formate was in the form of uncharged formic acid (Moon et al., 2021), which can diffuse across membranes (Falke et al., 2010). At the optimal concentration of 200 mM and pH 7.0 (pK<sub>a</sub>formic acid = 3.75), 0.1 mM of the substrate was in the form of formic acid, and 199.9 mM was present as formate (Moon et al., 2021). Regrettably, growth of *A. woodii* mutant with deleted formate transporter gene has not been tested at formate concentrations below 50 mM. Meanwhile, strain 3443-3Ac<sup>T</sup> grew on 25 mM formate at lower pH (pH 5.8) and 50°C, which implies that 0.2 mM of the substrate was in the form of formic acid, allowing us to assume that the uptake of formate in strain 3443-3Ac<sup>T</sup> is indeed due to diffusion of formic acid through the membrane.

Strain 3443-3Ac<sup>T</sup> converted formate to acetate (and CO<sub>2</sub>, which was not measured) with a ratio of 4.2:1.0 (Figure 2B). This is comparable to the ratio of 4.4:1.0 obtained for *A. woodii* (Moon et al., 2021). We suppose that the pathway of acetogenesis from formate in strain 3443-3Ac<sup>T</sup> is similar to the pathway postulated for *A. woodii* (Moon et al., 2021) with an important distinction that the Ech complex operates instead of the Rnf complex. Of four molecules of formate metabolized, three are oxidized to CO<sub>2</sub> and H<sub>2</sub> by HDCR, whereas one molecule of formate is converted to formyl-THF, reduced further to methyl-THF in the methyl branch of the WLP (Supplementary Figure 5). In the carbonyl branch, CO<sub>2</sub> is reduced to CO, which is bound to the methyl group by the CODH/ACS complex, generating acetyl-CoA, further converted to acetate. Balance of reducing equivalents is achieved by the interplay of HDCR, bifurcating hydrogenase HydABC, and Ech complex

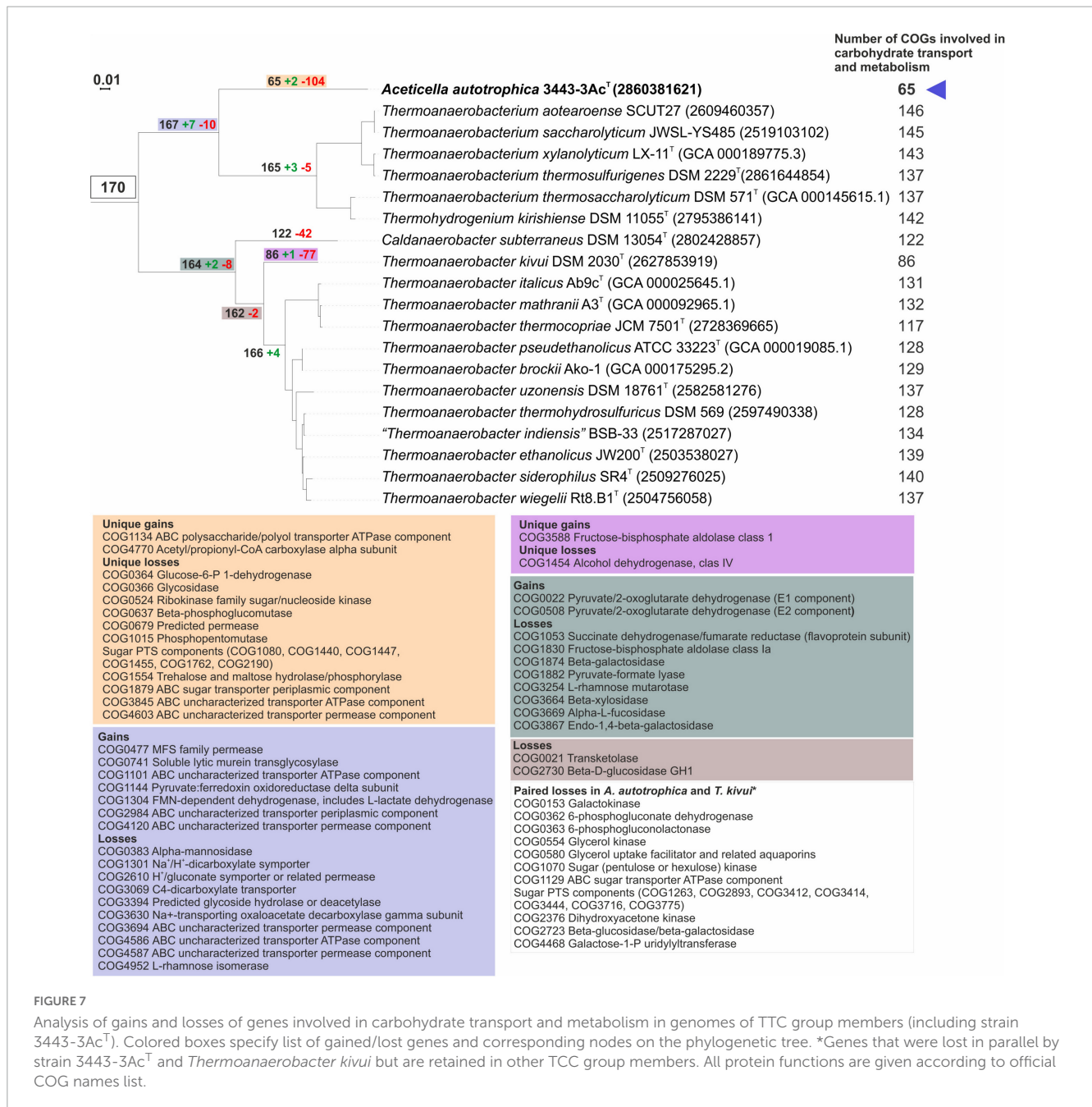


FIGURE 7

Analysis of gains and losses of genes involved in carbohydrate transport and metabolism in genomes of TTC group members (including strain 3443-3Ac<sup>T</sup>). Colored boxes specify list of gained/lost genes and corresponding nodes on the phylogenetic tree. \*Genes that were lost in parallel by strain 3443-3Ac<sup>T</sup> and *Thermoanaerobacter kivui* but are retained in other TCC group members. All protein functions are given according to official COG names list.

(Supplementary Figure 5): HydABC oxidizes the H<sub>2</sub> produced by HDCR and Ech, thus forming reduced ferredoxin and NADH, and Ech oxidizes part of the reduced ferredoxin with the formation of H<sub>2</sub> and generation of a chemiosmotic gradient.

The key enzyme in CO utilization is carbon monoxide dehydrogenase. The genome of strain 3443-3Ac<sup>T</sup> contains two CO dehydrogenase genes (see above). Based on the high sequence identities of the CO dehydrogenases of strain 3443-3Ac<sup>T</sup> with the biochemically characterized CO dehydrogenases of *Thermoanaerobacter kivui*, we predict that the ACETAC\_10650 gene is the *acsA*, encoding the CO dehydrogenase component of the CODH/ACS complex, whereas the ACETAC\_02795 gene is to be termed *cooS* and it encodes a monofunctional CO dehydrogenase. Thus, CO is oxidized by *cooS*, reducing ferredoxin, reoxidized then

by the Ech complex with the formation of H<sub>2</sub> and generation of chemiosmotic gradient (Supplementary Figure 6). We assume that part of the H<sub>2</sub> formed is used by HDCR to reduce CO<sub>2</sub> to formate, and the other part is oxidized by the bifurcating hydrogenase HydABC to produce NADH, required by the WLP methyl branch. CO and formate are further converted to acetate via the WLP.

The proposed metabolic schemes of strain 3443-3Ac<sup>T</sup> growth on either formate or CO suggest balance of reducing equivalents without production of hydrogen or consumption of exogenous CO<sub>2</sub>. However, in both cases H<sub>2</sub> is an essential intracellular intermediate, and due to its high diffusivity, part of it should be prone to escape the cell. Actually, small hydrogen production was observed both on formate and CO in our experiments, as it was observed during the growth on CO of *T. kivui*

(Weghoff and Müller, 2016), for which an analogous metabolic scheme has been proposed (Schoelmerich and Müller, 2019).

It should be pointed out that the proposed schemes of strain 3443-3Ac<sup>T</sup> metabolism during growth on H<sub>2</sub>+CO<sub>2</sub>, HCOOH, and CO (Figure 5 and Supplementary Figures 5, 6) suggest operation of the same enzymatic machinery, with the distinctions being the involvement of CooS in the growth on CO and the opposite direction of HDCR operation in case of growth on formate.

### 3.6. The likely causes for the inability of strain 3443-3Ac<sup>T</sup> to grow heterotrophically

All representatives of the genera *Thermoanaerobacterium*, *Thermoanaerobacter* and *Caldanaerobacter* (TTC group) are chemoorganoheterotrophic bacteria able to grow by fermenting sugars. Strain 3443-3Ac<sup>T</sup>, is thus an exception, being, according to our knowledge, the first obligately autotrophic representative of the TTC group. Genome analysis showed that genes coding for glycolysis/gluconeogenesis enzymes are present in genomes of all representatives the TTC group, including strain 3443-3Ac<sup>T</sup> (Supplementary Table 2).

To decipher the inability of strain 3443-3Ac<sup>T</sup> to grow on carbohydrates, genes gain and loss analysis within the TTC phylogenetic cluster was performed. Comparison of the detected full COG sets (Supplementary Figure 7) and COG sets involved in carbohydrate transport and metabolism (Figure 7) showed a drastic decrease in the number of carbohydrate metabolism-related genes in strain 3443-3Ac<sup>T</sup> (Figure 7 and Supplementary material 4), while total number of COGs was similar to those in other representatives of the TTC cluster. The single acetogen in the TTC group (apart from strain 3443-3Ac<sup>T</sup>), *Thermoanaerobacter kivui* DSM 2030, had a similar profile of gene losses, but in *T.kivui*'s case the losses do not happen to be critical for the ability to utilize sugars. To reveal these critical losses, we compared unique gains/losses of *Aceticella autotrophica* with those of *T. kivui* among COGs involved in carbohydrate transport and metabolism, and also payed attention to paired losses among these COGs (i.e., losses that were parallel in these two organisms but did not occur in other TTC group members). Overall, strain 3443-3Ac<sup>T</sup> had much more unique losses compared to *T. kivui*, and these losses mainly affected putative sugar-specific or uncharacterized transport systems, glycoside hydrolases and some other sugar-active enzymes. Among paired losses of strain 3443-3Ac<sup>T</sup> and *T. kivui*, there were genes encoding various sugar-specific transporters, such as phosphotransferase system (PTS) transporters; sugar-specific adenosine triphosphate-binding cassette (ABC) transporters; and other permeases. As for the PTS, *T. kivui* genome encoded two complete PTSs consisting of IICB-IIA and IIA-IIC-IIB components while in the genome of strain 3443-3Ac<sup>T</sup> only a gene fragment of a PTS putative EIIB component (ACETAC\_00175) was found, indicating the PTS loss might be one of the losses critical to make 3443-3Ac<sup>T</sup> an obligate autotroph. Interestingly, a gene encoding a cellulose-phosphorylase (GH94) domain-containing protein is present in the genomes of both strains as well as in other TTC members, which may be another indication of an operational PTS in strain 3443-3Ac<sup>T</sup> ancestors. Additional paired

losses included genes encoding enzymes involved in carbohydrate catabolism: sugar kinases, isomerases and transferases required for carbohydrate interconversions before entry into central catabolic pathways. Enzymes of galactose catabolism - galactokinase and galactose-1-phosphate uridylyltransferase - were also among paired losses. Oxidative branch of pentose phosphate pathway has undergone impairment in both organisms, due to the loss of phosphogluconolactonase, decarboxylating 6-phosphogluconate dehydrogenase and, in strain 3443-3Ac<sup>T</sup>, also of glucose-6-P 1-dehydrogenase (a unique loss of the strain).

Thus, we suggest that the inability of strain 3443-3Ac<sup>T</sup> to use sugars as carbon and energy sources is associated with the loss of sugar transport systems and enzymes involved in carbohydrate catabolism.

### 3.7. On the likely evolutionary origin of acetogenic metabolism in the TTC phylogenetic group

Strains 3443-3Ac<sup>T</sup> and *Thermoanaerobacter kivui* DSM 2030<sup>T</sup> (the single strain of the *T. kivui* species) are the only two organisms among members of the TTC group that possess WLP and exhibit acetogenesis capacity. Thus, a question arises as to whether the WLP and the acetogenesis capacity were acquired by strain 3443-3Ac<sup>T</sup> and *Thermoanaerobacter kivui* via horizontal transfer events or were inherited vertically from a common ancestor. The latter scenario implies losses of acetogenesis determinants in other species of the TTC group.

Neither of acetogenesis determinants is acetogenesis-specific. Instead, as discussed in the section "1. Introduction," the acetogenesis capacity requires the presence of a particular combination of enzymes and enzymatic complexes each of which may be involved in other processes. This required combination includes (1) WLP enzymes, (2) an enzymatic complex coupling ion translocation with electron transfer from ferredoxin to a low-potential acceptor such as H<sup>+</sup> or NAD<sup>+</sup> (Ech-hydrogenase or Rnf, respectively), and (3) an enzymatic machinery for the production of reduced ferredoxin, the donor for Ech-hydrogenase or for Rnf. Mere presence of WLP enzymes may mean just the presence of a pathway of autotrophic CO<sub>2</sub> fixation coupled to a mechanism of energy generation different from those employed in acetogenesis or be indicative of acetate oxidation capacity. Nevertheless, it seems appropriate to consider the evolutionary origin of acetogenesis in strain 3443-3Ac<sup>T</sup> beginning with tracing the evolution of WLP determinants, first of all, of the WLP-specific carbon monoxide dehydrogenase/acetyl-CoA synthase (CODH/ACS) enzymatic complex, whose role in the case of acetogenesis is the reduction of CO<sub>2</sub> to CO and further condensation of CO with methyl group, resulting in acetyl-CoA production. This complex occurs in both bacteria and archaea; in bacteria, the encoding genes are termed *acsABCDE*.

To trace the origin of the specific key determinants of WLP in strain 3443-3Ac<sup>T</sup> and *T. kivui*, the only representatives of the TTC group in which these genes have been found, we constructed phylogenetic trees for their AcsABCDE proteins. Four of these five trees (those of AcsBCDE) exhibited similar topologies (Supplementary Figures 8–12 and Supplementary material 5):



the proteins of 3443-3Ac<sup>T</sup> occurred within clades formed by proteins belonging to representatives of the GTDB orders o\_Clostridiales and o\_Peptostreptococcales within class c\_Clostridia, whereas proteins of *T.kivui* were found among proteins of class c\_Thermosediminibacteria members. This suggests that the ancestors of strain 3443-3Ac<sup>T</sup> and *T. kivui* acquired the *acsBCDE* genes independently via HGT from different donors.

As discussed above, in strain 3443-3Ac<sup>T</sup> and *T. kivui*, as well as in many representatives of the class c\_Clostridia orders o\_Clostridiales and o\_Peptostreptococcales and of class c\_Thermosediminibacteria (GTDB taxonomy), the *acsBCDE* genes occur in a common cluster (the WLP cluster) with some accessory genes and with genes responsible for the WLP methyl branch (Figure 6 and Supplementary material 5). Phylogenetic trees that we constructed for the products of these genes (Supplementary Figures 13–20) showed topologies similar to those of the *AcsBCDE* proteins: the proteins of strain 3443-3Ac<sup>T</sup> were among proteins of c\_Clostridia, whereas the proteins of *T.kivui* were among or close to proteins of c\_Thermosediminibacteria. Thus, the WLP gene clusters were likely acquired by (ancestors of) strain 3443-3Ac<sup>T</sup> and *T. kivui* as entire units via independent multigenic transfer events. Certain differences between the topologies of particular trees are likely to result from subsequent gene replacements via HGT across shorter phylogenetic distances. The members of the TTC group seem prone to ready acquisition of genetic information, as evidenced by demonstration of wide occurrence of natural competence among *Thermoanaerobacter* and *Thermoanaerobacterium* species (Shaw et al., 2010; Zeldes et al., 2023).

Given this phylogeny of the specific key determinants of WLP in strain 3443-3Ac<sup>T</sup> and *T. kivui* and lack of WLP in other TTC group members, it seemed natural to infer two independent events of *de novo* emergence of WLP and thus acetogenesis capacity in strain 3443-3Ac<sup>T</sup> and *T. kivui*. However, further genomic analysis revealed facts that prompted us to also consider another scenario for the presence of WLP and acetogenesis capacity in the two strains and to choose this scenario as a more likely one. Namely, the WLP gene clusters currently occurring in the genomes of strain 3443-3Ac<sup>T</sup> and *T. kivui* may have resulted from replacements of a vertically inherited original WLP gene cluster that was functional in an ancestor of TTC. The facts indicating the pre-existence of WLP are as follows:

(1) The strain 3443-3Ac<sup>T</sup> and *T. kivui*'s *acsA* (CO dehydrogenase) genes exhibit unusual location in the genomes (they occur beyond the WLP gene clusters, Figure 6), and the phylogeny of these genes differs from the phylogeny of the WLP gene clusters and complies with vertical inheritance from a common ancestor: in CO-dehydrogenase tree, the ACETAC\_10650 and TKV\_c20100 proteins of strain 3443-3Ac<sup>T</sup> and *T. kivui* appeared as closest relatives, forming a separate lineage between CODHs of c\_Clostridia and c\_Thermosediminibacteria recognized as *AcsAs* due to clustering with *AcsB* in the genomes (Supplementary Figures 21, 22 and Supplementary material 6).

(2) Further support for the idea that WLP as a biochemical phenomenon was acquired by strain 3443-3Ac<sup>T</sup> and *T. kivui* from a common ancestor comes from examination of phylogeny of formate dehydrogenases and of the enzyme complexes that

are obligately required by WLP only when it operates as a part of acetogenesis process. Formate dehydrogenase genes (Supplementary Figure 23) and the gene clusters encoding bifurcating hydrogenase (Supplementary Figure 24) and energy-converting hydrogenase (Supplementary Figure 25) were apparently inherited by strain 3443-3Ac<sup>T</sup> and *T. kivui* vertically from a common ancestor. This conclusion is based on their occurrence in distinctly separated two-member clades in phylogenetic trees and their pairwise identity values in the range of 75–85% (the AAI value between strain 3443-3Ac<sup>T</sup> and *T. kivui* is 67% for complete proteomes and 74% for the 120 conserved proteins used for phylogenetic analysis). Such identity values well agree with vertical inheritance from a common ancestor.

(3) The genomes of strain 3443-3Ac<sup>T</sup> and *T. kivui* encode proteins representing all 18 COGs required for biosynthesis of cobalamin, the cofactor necessary for WLP. The number of COGs of this set represented in other members of the TTC group varied from 7 to 16. This occurrence pattern agrees with the hypothesis about the presence of acetogenesis capacity in the common ancestor of the TTC group and its loss in most of the modern TTC species, with the loss rate varying in lineages.

Our conclusions are in accordance with the results of analysis of the entire evolutionary history of the CODH/ACS-encoding genes performed by Adam et al. (2018), who arrived at the conclusion that these genes were usually inherited as a single unit, their inheritance mostly followed vertical pattern, and their phylogeny can be traced to LUCA. As for Firmicutes sensu Bergey's (corresponding to nine GTDB phyla, from p\_Firmicutes to p\_Firmicutes\_H), their *acsAB* genes were inferred to be monophyletic, but their *acsCDE* genes were demonstrated to have a deviating evolutionary history in a particular Firmicutes lineage corresponding to GTDB phylum Firmicutes\_A, where the *acsCDE* genes do not ascend to LUCA directly but are thought to result from replacement of native (vertically inherited) *acsCDE* genes by *acsCDE* genes coming from *Nitrospirae* (Adam et al., 2018). However, this replacement of particular details of the WLP mechanism cannot prevent one from assuming that, in Firmicutes, the WLP as biochemical phenomenon is monophyletic, was inherited vertically and likely ascends to LUCA. This is however quite a general overview of the evolutionary pattern, and the question is whether within particular lineages of Firmicutes the WLP could have been lost and later reacquired via horizontal transfer(s) of the necessary determinants. Based on the results of our analysis, such "loss and reacquisition" pattern does not anyway seem applicable to the case of our strain and *T. kivui*. We hypothesize that the WLP pathways in strain 3443-3Ac<sup>T</sup> and *T. kivui* have not arisen independently *de novo* via horizontal gene transfers; instead, the capacity was inherited vertically from a common ancestor. Horizontal gene transfers did play a role, but this role consisted in replacement of details of the evolutionarily persisting mechanism, which has been lost by other TTC group members, or, looking at a different angle, has not been inherited by them from the ancestral pangenome.

## 4. Conclusion

Here, we described the first obligately autotrophic acetogenic bacterium *Aceticella autotrophica* gen. nov., sp. nov., strain



3443-3Ac<sup>T</sup>, isolated from a terrestrial hot spring of Uzon Caldera, Kamchatka. *A. autotrophica* is capable of acetogenic growth on H<sub>2</sub>/CO<sub>2</sub>, CO, or formate and is unable to grow heterotrophically, which makes it unique among acetogenic microorganisms. So far only facultative autotrophic acetogens have been known, and their metabolic versatility has been considered as the characteristic trait related to the low competitiveness of autotrophic acetogenesis in environmental conditions, where it occurs at the thermodynamic limit. We show that the inability of *A. autotrophica* to grow heterotrophically is associated with the loss of sugar-specific transport systems and many enzymes involved in carbohydrate catabolism. In ecological and evolutionary terms, the ability to survive in nature via obligately autotrophic acetogenic growth might be due to peculiarities of energy-converting mechanisms. However, our genomic analysis showed that *A. autotrophica*'s energy metabolism apparently involves cooperation of energy-converting hydrogenase, bifurcating hydrogenase, Wood-Ljungdahl pathway enzymes, and proton-dependent ATP synthase, i.e., the mechanism already described recently in so-called Ech-acetogens. Further studies are to show whether *A. autotrophica* has acquired yet unraveled additional mechanisms of energy conversion or peculiarities of enzyme kinetics allowing it to live as an obligatory acetogen.

#### 4.1. Description of *Aceticella* gen. nov.

*Aceticella* (A.ce.ti.cel'la. L.n. *acetum* vinegar; L.n. *cella* cell; N.L. fem. n. *Aceticella* vinegar cell).

Cells are rod-shaped. Cell wall of Gram-positive type. Spore-forming. Moderately thermophilic. Obligate chemolithoautotrophic anaerobe, growing by acetogenesis with H<sub>2</sub>, formate, or CO as the energy source and HCO<sub>3</sub><sup>-</sup>/CO<sub>2</sub> as the electron acceptor. Acetate is the only end product of the acetogenic growth on H<sub>2</sub>. The main end product during acetogenic growth with formate or CO is acetate, but small amounts of hydrogen are also formed. In the presence of thiosulfate or elemental sulfur sulfide formation is observed, but acetate also is the main product of the growth. The type species is *Aceticella autotrophica*.

#### 4.2. Description of *Aceticella autotrophica* sp. nov.

*Aceticella autotrophica* (au.to.tro'phi.ca. Gr. pron. *autos* self; Gr. adj. *trophikos* one who feeds; N.L. fem. adj. *autotrophica* autotrophic).

Cells are rod-shaped, 0.5–0.7 μm in diameter and 1.5–5 μm in length, growing singly, in pairs or in chains depending on the growth phase and growth conditions. Cells are non-motile. Spore-forming. Moderately thermophilic. Growth occurs between 30 and 60°C, with an optimum at 46–50°C. The pH range for growth is 4.5–7.3, with an optimum pH of 6.0. Growth does not occur above a NaCl concentration of 1.5% (w/v). Growth is not Na<sup>+</sup>-dependent. Cells are capable to grow in vitamin-free medium. Obligate chemolithoautotrophic anaerobe, growing by acetogenesis with H<sub>2</sub> as the energy source and HCO<sub>3</sub><sup>-</sup>/CO<sub>2</sub> as the carbon source and the electron acceptor. Acetate is the only end product of the growth on H<sub>2</sub> and HCO<sub>3</sub><sup>-</sup>/CO<sub>2</sub>. Formate or CO can also be used as the energy source. The main end product

during growth with formate or CO is acetate, but small amounts of hydrogen are also formed. No growth occurs with yeast extract, beef extract, peptone, pectin, dextrin, starch, cellulose, xylan, glucose, fructose, galactose, mannose, arabinose, rhamnose, xylose, ribose, trehalose, cellobiose, sucrose, lactose, maltose, raffinose, sorbitol, mannitol, acetate, lactate, pyruvate, malate, propionate, butyrate, fumarate, succinate, citrate, ethanol, propanol, glycerol, methanol, 2-methoxyphenol, 3,4-dimethoxybenzoate, or 2-methoxybenzoate. Nitrate, sulfate, perchlorate, fumarate, Fe (III) citrate, nitrite, sulfite, or O<sub>2</sub> are not used as the electron acceptors. Sulfide formation is observed on the medium with thiosulfate or elemental sulfur, but acetate is the main product of the growth. In the presence of thiosulfate or elemental sulfur, growth rate and cell yield are the same as in the absence of these compounds. The G+C content of genomic DNA is 33 mol %. The major cellular fatty acid is iso-C<sub>15:0</sub> and C<sub>16:0</sub>. Minor components are iso-C<sub>17:0</sub>, C<sub>18:0</sub>, iso-C<sub>16:0</sub>, and anteiso-C<sub>17:0</sub>. Membranes polar lipids include four unidentified phospholipids and unknown aminophospholipid. No respiratory quinones were detected. The type strain is 3443-3Ac<sup>T</sup> (=DSM 108286<sup>T</sup> = VKM B-3415<sup>T</sup>), isolated from a sample of sediments collected from the terrestrial hot spring Kaskadny at East Thermal Field, Uzon Caldera, Kamchatka, Russia (N54° 30.026' E160° 00.374', elevation 658 m).

#### Data availability statement

Strain 3443-3Ac<sup>T</sup> was deposited in the DSMZ (German Collection of Microorganisms and Cell Cultures) and VKM (All-Russian Collection of Microorganisms) under accession numbers: DSM 108286 and VKM B-3415, respectively. Genome sequence, as well as related project information and sample details, were deposited in NCBI database under accession numbers: CP060096, PRJNA647162, and SAMN15577649, respectively. The genome sequence was also deposited in IMG with genome ID 2860381621.

#### Author contributions

EF, AL, and IK: conceptualization. EF, AG, ST, and AE: data curation. IK: funding acquisition. EF, AL, AG, ST, AN, and AE: investigation. EF and IK: project administration and supervision. EF, AG, ST, and AN: resources. EF, AL, ST, AE, and IK: visualization. EF and AL: writing – original draft preparation. All authors contributed to the formal analysis, methodology, validation, and writing – review and editing and article and approved the submitted version.

#### Funding

This work was supported by the grant from the Russian Science Foundation # 21-14-00242 (enrichment, isolation, and physiological studies, analytical methods, genome sequencing and assembly, genome annotation, whole genome analyses). In addition, the work of the AL, AE, and AG on phylogenetic analyses was supported by the Ministry of Science and Higher Education of the Russian Federation.

## Acknowledgments

We are grateful to the staff of Kronotsky Nature Reserve for their assistance in the organization of field studies in the Uzon Caldera. We thank Kseniya Zayulina for her help in measuring glucose and cellobiose consumption of cell extract of strain 3443-3Ac<sup>T</sup> using a DNS assay.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

- Adam, P. S., Borrel, G., and Gribaldo, S. (2018). Evolutionary history of carbon monoxide dehydrogenase/acetyl-CoA synthase, one of the oldest enzymatic complexes. *Proc. Natl. Acad. Sci. U S A.* 115, E1166–E1173.
- Allen, T. D., Caldwell, M. E., Lawson, P. A., Huhnke, R. L., and Tanner, R. S. (2010). *Alkalibaculum bacchi* gen. nov., sp. nov., a CO-oxidizing, ethanol-producing acetogen isolated from livestock-impacted soil. *Int. J. Syst. Evol. Microbiol.* 60, 2483–2489. doi: 10.1099/ijs.0.018507-0
- Arndt, D., Grant, J. R., Marcu, A., Saied, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44, W16–W21.
- Bar-Even, A. (2013). Does acetogenesis really require especially low reduction potential? *Biochim. Biophys. Acta* 1827, 395–400.
- Basen, M., and Müller, V. (2017). "Hot" acetogenesis. *Extremophiles* 21, 15–26.
- Bertelli, C., Laird, M. R., Williams, K. P., Simon Fraser University Research Computing Group, Lau, B. Y., Hoad, G., et al. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* 45, W30–W35. doi: 10.1093/nar/gkx343
- Biegel, E., and Müller, V. (2010). Bacterial Na<sup>+</sup>-translocating ferredoxin:NAD<sup>+</sup> oxidoreductase. *Proc. Natl. Acad. Sci. U S A.* 107, 18138–18142.
- Blamey, J. M., and Adams, M. W. (1993). Purification and characterization of pyruvate ferredoxin oxidoreductase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Biochim. Biophys. Acta* 1161, 19–27. doi: 10.1016/0167-4838(93)90190-3
- Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S. Y., et al. (2019). antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47, W81–W87. doi: 10.1093/nar/gkz310
- Bonch-Osmolovskaya, E. A., Sokolova, T. G., Kostrikin, N. A., and Zavarzin, G. A. (1990). *Desulfurella acetivorans* gen. nov. and sp. nov.—a new thermophilic sulfur-reducing eubacterium. *Arch. Microbiol.* 153, 151–155.
- Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., et al. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* 5:8365. doi: 10.1038/srep08365
- Buckel, W., and Thauer, R. K. (2013). Energy conservation via electron bifurcating ferredoxin reduction and proton/Na<sup>+</sup> translocating ferredoxin oxidation. *Biochim. Biophys. Acta* 1827, 94–113. doi: 10.1016/j.bbabi.2012.07.002
- Buckel, W., and Thauer, R. K. (2018). Flavin-based electron bifurcation, a new mechanism of biological energy coupling. *Chem. Rev.* 118, 3862–3886.
- Cáceres, A. J., Portillo, R., Acosta, H., Rosales, D., Quiñones, W., Avilán, L., et al. (2003). Molecular and biochemical characterization of hexokinase from *Trypanosoma cruzi*. *Mol. Biochem. Parasitol.* 126, 251–262. doi: 10.1016/s0166-6851(02)00294-3
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421.
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552.
- Chen, I. A., Markowitz, V. M., Chu, K., Palaniappan, K., Szeto, E., Pillay, M., et al. (2017). IMG/M: integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res.* 45, D507–D516.
- Csurös, M. (2010). Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics* 26, 1910–1912. doi: 10.1093/bioinformatics/btq315
- Deb, S., and Das, S. K. (2022). Phylogenomic analysis of metagenome-assembled genomes deciphered novel acetogenic nitrogen-fixing *Bathyarchaeota* from hot spring sediments. *Microbiol. Spectr.* 10:e0035222. doi: 10.1128/spectrum.00352-22
- Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., et al. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* 359:eaar4120. doi: 10.1126/science.aar4120
- Drake, H. L., Gößner, A. S., and Daniel, S. L. (2008). Old acetogens, new light. *Ann. N Y Acad. Sci.* 1125, 100–128. doi: 10.1196/annals.1419.016
- Evans, P. N., Parks, D. H., Chadwick, G. L., Robbins, S. J., Orphan, V. J., Golding, S. D., et al. (2015). Methane metabolism in the archaeal phylum *Bathyarchaeota* revealed by genome-centric metagenomics. *Science* 350, 434–438.
- Falke, D., Schulz, K., Doberenz, C., Beyer, L., Lilie, H., Thieme, B., et al. (2010). Unexpected oligomeric structure of the FocA formate channel of *Escherichia coli*: a paradigm for the formate-nitrite transporter family of integral membrane proteins. *FEMS Microbiol. Lett.* 303, 69–75.
- Fardeau, M.-L., Ollivier, B., and Cayol, J. (2015). "Caldanaerobacter," in *Bergey's Manual of Systematics of Archaea and Bacteria*, ed. W. B. Whitman (Chichester: John Wiley & Sons, Ltd). doi: 10.1002/9781118960608.gbm00745
- Fritz, M., and Müller, V. (2007). An intermediate step in the evolution of ATPases—The F<sub>1</sub>F<sub>0</sub>-ATPase from *Acetobacterium woodii* contains F-type and V-type rotor subunits and is capable of ATP synthesis. *FEBS J.* 274, 3421–3428. doi: 10.1111/j.1742-4658.2007.05874.x
- Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28, 3150–3152. doi: 10.1093/bioinformatics/bts565
- Gerritsen, J., Fuentes, S., Grievink, W., van Niftrik, L., Tindall, B. J., Timmerman, H. M., et al. (2014). Characterization of *Romboutsia ilealis* gen. nov., sp. nov., isolated from the gastro-intestinal tract of a rat, and proposal for the reclassification of five closely related members of the genus *Clostridium* into the genera *Romboutsia* gen. nov., *Intestinibacter* gen. nov., *Terrisporobacter* gen. nov. and *Asaccharospora* gen. nov. *Int. J. Syst. Evol. Microbiol.* 64, 1600–1616. doi: 10.1099/ijs.0.059543-0
- He, Y., Li, M., Perumal, V., Feng, X., Fang, J., Xie, J., et al. (2016). Genomic and enzymatic evidence for acetogenesis among multiple lineages of the archaeal phylum *Bathyarchaeota* widespread in marine sediments. *Nat. Microbiol.* 1:16035. doi: 10.1038/nmicrobiol.2016.35
- Heider, J., Ma, K., and Adams, M. W. (1995). Purification, characterization, and metabolic function of tungsten-containing aldehyde ferredoxin oxidoreductase from the hyperthermophilic and proteolytic archaeon *Thermococcus* strain ES-1. *J. Bacteriol.* 177, 4757–4764. doi: 10.1128/jb.177.16.4757-4764.1995
- Henstra, A. M., Dijkema, C., and Stams, A. J. (2007). *Archaeoglobus fulgidus* couples CO oxidation to sulfate reduction and acetogenesis with transient formate accumulation. *Environ. Microbiol.* 9, 1836–1841. doi: 10.1111/j.1462-2920.2007.01306.x
- Herrmann, G., Jayamani, E., Mai, G., and Buckel, W. (2008). Energy conservation via electron-transferring flavoprotein in anaerobic bacteria. *J. Bacteriol.* 190, 784–791. doi: 10.1128/JB.01422-07

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2023.1185739/full#supplementary-material>

- Hess, V., Poehlein, A., Weghoff, M. C., Daniel, R., and Müller, V. (2014). A genome-guided analysis of energy conservation in the thermophilic, cytochrome-free acetogenic bacterium *Thermoanaerobacter kivui*. *BMC Genomics* 15:1139. doi: 10.1186/1471-2164-15-1139
- Hess, V., Schuchmann, K., and Müller, V. (2013). The ferredoxin: NAD<sup>+</sup> oxidoreductase (Rnf) from the acetogen *Acetobacterium woodii* requires Na<sup>+</sup> and is reversibly coupled to the membrane potential. *J. Biol. Chem.* 288, 31496–31502.
- Ikeda-Ohtsubo, W., Strassert, J. F., Köhler, T., Mikaelyan, A., Gregor, I., McHardy, A. C., et al. (2016). 'Candidatus Adiatrix intracellularis', an endosymbiont of termite gut flagellates, is the first representative of a deep-branching clade of *Deltaproteobacteria* and a putative homoacetogen. *Environ. Microbiol.* 18, 2548–2564. doi: 10.1111/1462-2920.13234
- Jain, S., Katsyv, A., Basen, M., and Müller, V. (2021). The monofunctional CO dehydrogenase CooS is essential for growth of *Thermoanaerobacter kivui* on carbon monoxide. *Extremophiles* 26:4. doi: 10.1007/s00792-021-01251-y
- Kanonenberg, K., Schwarz, C. K., and Schmitt, L. (2013). Type I secretion systems - a story of appendices. *Res. Microbiol.* 164, 596–604. doi: 10.1016/j.resmic.2013.03.011
- Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108
- Kevbrin, V. V., and Zavarzin, G. A. (1992). Effect of sulfur compounds on the growth of the halophilic homoacetogenic bacterium *Acetohalobium arabaticum*. *Microbiology* 61, 563–567.
- Kim, B. C., Grote, R., Lee, D. W., Antranikian, G., and Pyun, Y. R. (2001). *Thermoanaerobacter yonseiensis* sp. nov. m a novel extremely thermophilic, xylose-utilizing bacterium that grows at up to 85 degrees C. *Int. J. Syst. Evol. Microbiol.* 51, 1539–1548. doi: 10.1099/00207713-51-4-1539
- Köpke, M., Held, C., Hujer, S., Liesegang, H., Wiezer, A., Wollherr, A., et al. (2010). *Clostridium ljungdahlii* represents a microbial production platform based on syngas. *Proc. Natl. Acad. Sci. U S A.* 107, 13087–13092. doi: 10.1073/pnas.1004716107
- Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., and Phillippy, A. M. (2017). Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* 27, 722–736. doi: 10.1101/gr.215087.116
- Kozina, I. V., Kublanov, I. V., Kolganova, T. V., Chernyh, N. A., and Bonch-Osmolovskaya, E. A. (2010). *Caldanaerobacter uzonensis* sp. nov., an anaerobic, thermophilic, heterotrophic bacterium isolated from a hot spring. *Int. J. Syst. Evol. Microbiol.* 60, 1372–1375.
- Kuhns, M., Trifunovic, D., Huber, H., and Müller, V. (2020). The Rnf complex is a Na<sup>+</sup> coupled respiratory enzyme in a fermenting bacterium. *Thermotoga maritima*. *Commun. Biol.* 3:431. doi: 10.1038/s42003-020-01158-y
- Lawson, P. A., Citron, D. M., Tyrrell, K. L., and Finegold, S. M. (2016). Reclassification of *Clostridium difficile* as *clostridioides difficile* (hall and O'Toole 1935) Prévot 1938. *Anaerobe* 40, 95–99. doi: 10.1016/j.anaerobe.2016.06.008
- Le Van, T. D., Robinson, J. A., Ralph, J., Greening, R. C., Smolenski, W. J., Leedale, J. A., et al. (1998). Assessment of reductive acetogenesis with indigenous ruminal bacterial populations and *Acetivomaculum ruminis*. *Appl. Environ. Microbiol.* 64, 3429–3436. doi: 10.1128/AEM.64.9.3429-3436.1998
- Loh, H. Q., Hervé, V., and Brune, A. (2021). Metabolic potential for reductive acetogenesis and a novel energy-converting [NiFe] hydrogenase in *Bathyrarchaea* from termite guts - a genome-centric analysis. *Front. Microbiol.* 11:635786. doi: 10.3389/fmicb.2020.635786
- Loman, N. J., Quick, J., and Simpson, J. T. (2015). A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat. Methods* 12, 733–735. doi: 10.1038/nmeth.3444
- Lopez, G., Canas-Duarte, S. J., Pinzon-Velasco, A. M., Vega-Vela, N. E., Rodriguez, M., Restrepo, S., et al. (2017). Description of a new anaerobic thermophilic bacterium, *Thermoanaerobacterium butyriciformans* sp. nov. *Syst. Appl. Microbiol.* 40, 96–91. doi: 10.1016/j.syapm.2016.11.006
- Luo, H., and Gao, F. (2019). DoriC 10.0: an updated database of replication origins in prokaryotic genomes including chromosomes and plasmids. *Nucleic Acids. Res.* 47, D74–D77.
- Mao, Z., Gräßle, F., Frey, J., Franchini, P., Schleheck, D., Müller, N., et al. (2021). *Phosphitispora fastidiosa* gen. nov. sp. nov., a new dissimilatory phosphite-oxidizing anaerobic bacterium isolated from anaerobic sewage sludge. *Int. J. Syst. Evol. Microbiol.* 71:5142. doi: 10.1099/ijsem.0.005142
- Martínez, B., Suárez, J. E., and Rodríguez, A. (1996). Lactococcin 972: a homodimeric lactococcal bacteriocin whose primary target is not the plasma membrane. *Microbiology* 142, 2393–2398. doi: 10.1099/00221287-142-9-2393
- Merkel, A. Yu, Pimenov, N. V., Rusanov, I. I., Slobodkin, A. I., Slobodkina, G. B., Tarnovetckii, I. Yu, et al. (2017). Microbial diversity and autotrophic activity in Kamchatka hot springs. *Extremophiles* 21, 307–317. doi: 10.1007/s00792-016-0903-1
- Miller, G. L. (1959). Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 31, 426–428.
- Moon, J., Dönig, J., Kramer, S., Poehlein, A., Daniel, R., and Müller, V. (2021). Formate metabolism in the acetogenic bacterium *Acetobacterium woodii*. *Environ. Microbiol.* 23, 4214–4227. doi: 10.1111/1462-2920.15598
- Müller, V., Chowdhury, N. P., and Basen, M. (2018). Electron bifurcation: a long-hidden energy-coupling mechanism. *Annu. Rev. Microbiol.* 72, 331–353. doi: 10.1146/annurev-micro-090816-093440
- Onyenwoke, R. U., and Wiegel, J. (2015). "Thermoanaerobacterium," in *Bergey's Manual of Systematics of Archaea and Bacteria*, ed. W. B. Whitman (Chichester: John Wiley & Sons, Ltd).
- Orsi, W. D., Vuillemin, A., Rodriguez, P., Coskun, ÖK., Gomez-Saez, G. V., Lavik, G., et al. (2020). Metabolic activity analyses demonstrate that Lokiarchaeon exhibits homoacetogenesis in sulfidic marine sediments. *Nat. Microbiol.* 5, 248–255.
- Park, D. (2007). Genomic DNA isolation from different biological materials. *Methods Mol. Biol.* 353, 3–13.
- Parks, D. H., Chuvochina, M., Chaumeil, P.-A., Rinke, C., Mussig, A. J., and Hugenholtz, P. (2020). A complete domain-to-species taxonomy for bacteria and Archaea. *Nat. Biotechnol.* 38, 1079–1086.
- Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* 36, 996–1004. doi: 10.1038/nbt.4229
- Parks, D. H., Rinke, C., Chuvochina, M., Chaumeil, P. A., Woodcroft, B. J., Evans, P. N., et al. (2017). Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* 2, 1533–1542.
- Parks, D. H., Tyson, G. W., Hugenholtz, P., and Beiko, R. G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30, 3123–3124.
- Poehlein, A., Daniel, R., Schink, B., and Semeonova, D. D. (2013). Life based on phosphite: a genome-guided analyses of *Desulfotignum phosphitoxidans*. *BMC Genomics* 14:753.
- Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Res.* 46, W200–W204.
- Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., and Korobeynikov, A. (2020). Using SPAdes de novo assembler. *Curr. Protoc. Bioinformatics* 70:e102.
- Ragsdale, S. W., and Pierce, E. (2008). Acetogenesis and the wood-ljungdahl of CO<sub>2</sub> fixation. *Biochim. Biophys. Acta* 1784, 1873–1898.
- Romano, I., Dipasquale, L., Orlando, P., Lama, L., d'Ippolito, G., Pascual, J., et al. (2010). *Thermoanaerobacterium thermostercus* sp. nov., a new anaerobic thermophilic hydrogen-producing bacterium from buffalo-dung. *Extremophiles* 14, 233–240. doi: 10.1007/s00792-010-0303-x
- Rother, M., and Metcalf, W. W. (2004). Anaerobic growth of *Methanosarcina acetivorans* C2A on carbon monoxide: an unusual way of life for a methanogenic archaeon. *Proc. Natl. Acad. Sci. U S A.* 101, 16929–16934. doi: 10.1073/pnas.0407486101
- Russel, J., Pinilla-Redondo, R., Mayo-Muñoz, D., Shah, S. A., and Sørensen, S. J. (2020). CRISPR-CasTyper: automated identification, annotation, and classification of CRISPR-Cas Loci. *CRISPR J.* 3, 462–469. doi: 10.1089/crispr.2020.0059
- Schoelmerich, M. C., and Müller, V. (2019). Energy conservation by a hydrogenase-dependent chemiosmotic mechanism in an ancient metabolic pathway. *Proc. Natl. Acad. Sci. U S A.* 116, 6329–6334.
- Schöne, C., Poehlein, A., Jehmlich, N., Adlung, N., Daniel, R., von Bergen, M., et al. (2022). Deconstructing *Methanosarcina acetivorans* into an acetogenic archaeon. *Proc. Natl. Acad. Sci. U S A.* 119:e2113853119.
- Schuchmann, K., and Müller, V. (2012). A bacterial electron bifurcating hydrogenase. *J. Biol. Chem.* 287, 31165–31171.
- Schuchmann, K., and Müller, V. (2013). Direct and reversible hydrogenation of CO<sub>2</sub> to formate by a bacterial carbon dioxide reductase. *Science* 342, 1382–1385.
- Schuchmann, K., and Müller, V. (2014). Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* 12, 809–821. doi: 10.1038/nrmicro3365
- Schuchmann, K., and Müller, V. (2016). Energetics and application of heterotrophy in acetogenic bacteria. *Appl. Environ. Microbiol.* 82, 4056–4069. doi: 10.1128/AEM.00882-16
- Schwarz, F. M., Cirus, S., Jain, S., Baum, C., Wiechmann, A., Basen, M., et al. (2020). Revealing formate production from carbon monoxide in wild type and mutants of Rnf- and Ech-containing acetogens, *Acetobacterium woodii* and *Thermoanaerobacter kivui*. *Microb. Biotechnol.* 13, 2044–2056. doi: 10.1111/1751-7915.13663
- Schwarz, F. M., Schuchmann, K., and Müller, V. (2018). Hydrogenation of CO<sub>2</sub> at ambient pressure catalyzed by a highly active thermostable biocatalyst. *Biotechnol. Biofuels* 11:237. doi: 10.1186/s13068-018-1236-3
- Shang, S., Qian, L., Zhang, X., and Li, K. (2013). *Thermoanaerobacterium calidifontis* sp. nov., a novel anaerobic, thermophilic, ethanol-producing bacterium from hot springs in China. *Arch. Microbiol.* 195, 439–445.
- Shaw, A. J., Hogsett, D. A., and Lynd, L. R. (2010). Natural competence in *Thermoanaerobacter* and *Thermoanaerobacterium* species. *Appl. Environ. Microbiol.* 76, 4713–4719.
- Slobodkina, G. B., Merkel, A. Y., Novikov, A. A., Bonch-Osmolovskaya, E. A., and Slobodkin, A. I. (2020). *Pelomicrobium methylotrophicum* gen. nov., sp. nov. a moderately thermophilic, facultatively anaerobic, lithoautotrophic and

- methylotrophic bacterium isolated from a terrestrial mud volcano. *Extremophiles* 24, 177–185. doi: 10.1007/s00792-019-01145-0
- Slobodkina, G. B., Panteleeva, A. N., Kostrikin, N. A., Kopsyn, D. S., Bonch-Osmolovskaya, E. A., and Slobodkin, A. I. (2013). *Tepidibacillus fermentans* gen. nov., sp. nov.: a moderately thermophilic anaerobic and microaerophilic bacterium from an underground gas storage. *Extremophiles* 17, 833–839. doi: 10.1007/s00792-013-0564-2
- Sondergaard, D., Pedersen, C. N., and Greening, C. (2016). HydDB: a web tool for hydrogenase classification and analysis. *Sci. Rep.* 6:34212. doi: 10.1038/srep34212
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033
- Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., Zaslavsky, L., et al. (2016). NCBI prokaryotic genome annotation pipeline. *Nucleic Acids. Res.* 44, 6614–6624. doi: 10.1093/nar/gkw569
- Teichtmann, S. M., Lebedinsky, A. V., Colman, A. S., Sokolova, T. G., Woyke, T., Goodwin, L., et al. (2012). Evidence for horizontal gene transfer of anaerobic carbon monoxide dehydrogenases. *Front. Microbiol.* 3:132. doi: 10.3389/fmicb.2012.00132
- Tesson, F., Hervé, A., Mordret, E., Touchon, M., d'Humières, C., Cury, J., et al. (2022). Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nat. Commun.* 13:2561. doi: 10.1038/s41467-022-30269-9
- Thauer, R. K., Jungermann, K., and Decker, K. (1977). Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.* 41, 100–180. doi: 10.1128/br.41.1.100-180.1977
- Tomas, A. F., Karakashev, D., and Angelidaki, I. (2013). *Thermoanaerobacter pentosaceus* sp. nov., an anaerobic, extremely thermophilic, high ethanol-yielding bacterium isolated from household waste. *Int. J. Syst. Evol. Microbiol.* 63, 2396–23404. doi: 10.1099/ijs.0.045211-0
- Toshchakov, S. V., Kublanov, I. V., Messina, E., Yakimov, M. M., and Golyshin, P. N. (2015). “Hydrocarbon and lipid microbiology protocols. genetic, genomic and system analyses of communities,” in *Genomic Analysis of Pure Cultures and Communities*, eds T. J. McGenity, K. N. Timmis, and B. N. Fernandez (Berlin: Springer-Verlag). doi: 10.1007/8623\_2015\_126
- Tremblay, P. L., Zhang, T., Dar, S. A., Leang, C., and Lovley, D. R. (2012). The Rnf complex of *Clostridium ljungdahlii* is a proton-translocating ferredoxin: NAD<sup>+</sup> oxidoreductase essential for autotrophic growth. *mBio* 4:e00406-12. doi: 10.1128/mBio.00406-12
- Trüper, H. G., and Schlegel, H. G. (1964). Sulfur metabolism in *Thiorhodaceae*. I. quantitative measurements on growing cells of *Chromatium okenii*. *Antonie van Leeuwenhoek* 30, 225–238. doi: 10.1007/BF02046728
- Vignais, P. M., and Billoud, B. (2007). Occurrence, classification, and biological function of hydrogenases: an overview. *Chem. Rev.* 107, 4206–4272. doi: 10.1021/cr050196r
- Wagner, I. D., Zhao, W., Zhang, C. L., Romanek, C. S., Rohbe, M., and Wiegel, J. (2008). *Thermoanaerobacter uzonensis* sp. nov., an anaerobic thermophilic bacterium isolated from a hot spring within the Uzon Caldera, Kamchatka, Far East Russia. *Int. J. Syst. Evol. Microbiol.* 58, 2565–2573. doi: 10.1099/ijs.0.65343-0
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963
- Wang, S., Huang, H., Kahnt, J., and Thauer, R. K. (2013). A reversible electron-bifurcating ferredoxin- and NAD-dependent [FeFe]-hydrogenase (HydABC) in *Moorella thermoacetica*. *J. Bacteriol.* 195, 1267–1275. doi: 10.1128/JB.02158-12
- Weghoff, M. C., and Müller, V. (2016). CO metabolism in the thermophilic acetogen *Thermoanaerobacter kivui*. *Appl. Environ. Microbiol.* 82, 2312–2319. doi: 10.1128/AEM.00122-16
- Wolin, E. A., Wolin, M. J., and Wolfe, R. S. (1963). Formation of methane by bacterial extracts. *J. Biol. Chem.* 238, 2882–2888. doi: 10.1016/S0021-9258(18)67912-8
- Xie, Z., and Tang, H. (2017). ISEScan: automated identification of insertion sequence elements in prokaryotic genomes. *Bioinformatics* 33, 3340–3347. doi: 10.1093/bioinformatics/btx433
- Xiong, X., Wu, G., Wei, Y., Liu, L., Zhang, Y., Su, R., et al. (2020). SspABCD-SspE is a phosphorothioation-sensing bacterial defence system with broad anti-phage activities. *Nat. Microbiol.* 5, 917–928. doi: 10.1038/s41564-020-0700-6
- Yamamoto, K., Murakami, R., and Takamura, Y. (1998). Isoprenoid quinone, cellular fatty acid composition and diaminopimelic acid isomers of newly classified thermophilic anaerobic Gram-positive bacteria. *FEMS Microbiol. Lett.* 161, 351–358. doi: 10.1111/j.1574-6968.1998.tb12968.x
- Yang, S., Lv, Y., Liu, X., Wang, Y., Fan, Q., Yang, Z., et al. (2020). Genomic and enzymatic evidence of acetogenesis by anaerobic methanotrophic archaea. *Nat. Commun.* 11:3941. doi: 10.1038/s41467-020-17860-8
- Zeldes, B., Poehlein, A., Jain, S., Baum, C., Daniel, R., Müller, V., et al. (2023). DNA uptake from a laboratory environment drives unexpected adaptation of a thermophile to a minor medium component. *ISME Commun.* 3:2. doi: 10.1038/s43705-022-00211-7
- Zhilina, T. N., Zavarzina, D. G., Panteleeva, A. N., Osipov, G. A., Kostrikin, N. A., Tourova, T. P., et al. (2012). *Fuchsiella alkaliacetigena* gen. nov., sp. nov., an alkaliphilic, lithoautotrophic homoacetogen from a soda lake. *Int. J. Syst. Evol. Microbiol.* 62, 1666–1673. doi: 10.1099/ijs.0.034363-0