



Genome-Based Comparison of All Species of the Genus *Moorella*, and Status of the Species *Moorella thermoacetica* and *Moorella thermoautotrophica*

Stephanie Redl^{1†}, Anja Poehlein^{2†}, Carola Esser³, Frank R. Bengelsdorf³, Torbjørn Ø. Jensen¹, Christian B. Jendresen¹, Brian J. Tindall⁴, Rolf Daniel², Peter Dürre^{3*} and Alex T. Nielsen¹

¹ Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark, ² Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Georg-August University, Göttingen, Germany, ³ Institut für Mikrobiologie und Biotechnologie, Universität Ulm, Ulm, Germany, ⁴ Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Brunswick, Germany

OPEN ACCESS

Edited by:

Martin G. Klotz,
Washington State University,
United States

Reviewed by:

Charles Lovell,
University of South Carolina,
United States
Alexander V. Lebedinsky,
Winogradsky Institute of Microbiology
(RAS), Russia
Aharon Oren,
Hebrew University of Jerusalem, Israel

*Correspondence:

Peter Dürre
peter.duerre@uni-ulm.de

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Microbial Physiology and Metabolism,
a section of the journal
Frontiers in Microbiology

Received: 22 August 2019

Accepted: 19 December 2019

Published: 17 January 2020

Citation:

Redl S, Poehlein A, Esser C,
Bengelsdorf FR, Jensen TØ,
Jendresen CB, Tindall BJ, Daniel R,
Dürre P and Nielsen AT (2020)
Genome-Based Comparison of All
Species of the Genus *Moorella*,
and Status of the Species *Moorella*
thermoacetica and *Moorella*
thermoautotrophica.
Front. Microbiol. 10:3070.
doi: 10.3389/fmicb.2019.03070

Fermentation of gases provides a promising opportunity for the production of biochemicals from renewable resources, which has resulted in a growing interest in acetogenic bacteria. Thermophilic organisms provide potential advantages for the fermentation of, e.g., syngas into for example volatile compounds, and the thermophiles *Moorella thermoacetica* and *Moorella thermoautotrophica* have become model organisms of acetogenic metabolism. The justification for the recognition of the closely related species *M. thermoautotrophica* has, however, recently been disputed. In order to expand knowledge on the genus, we have here genome sequenced a total of 12 different *M. thermoacetica* and *M. thermoautotrophica* strains. From the sequencing results, it became clear that *M. thermoautotrophica* DSM 1974^T consists of at least two different strains. Two different strains were isolated in Lyngby and Ulm from a DSM 1974^T culture obtained from the DSMZ (Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Brunswick, Germany). Phylogenetic analysis revealed a close relationship between all the sequenced genomes, suggesting that the two strains detected in the type strain of the species *M. thermoautotrophica* could not be distinguished at the species level from *M. thermoacetica*. Despite genetic similarities, differences in genomic features were observed between the strains. Differences in compounds that can serve as carbon and energy sources for selected strains were also identified. On the contrary, strain DSM 21394, currently still named *M. thermoacetica*, obviously represents a new *Moorella* species. In addition, based on genome analysis and comparison *M. glycerini* NMP, *M. stamsii* DSM 26217^T, and *M. perchloratireducens* An10 cannot be distinguished at the species level. Thus, this comprehensive analysis provides a significantly increased knowledge of the genetic diversity of *Moorella* strains.

Keywords: anaerobic, thermophile, acetogen, gas fermentation, syngas fermentation, phylogenetic analysis, *Moorella*, *Moorella thermoacetica*

Abbreviations: ANI, average nucleotide identity; ATCC, American Type Culture Collection; CRISPR: clustered regularly interspaced short palindromic repeats; DSM(Z), Deutsche Sammlung von Mikroorganismen (und Zellkulturen); MLSA, multi locus sequence analysis; OG, orthologous groups.

INTRODUCTION

Interest from the research community and industry in acetogenic bacteria has grown within recent years due to their potential to produce valuable compounds from syngas (Latif et al., 2014). Thermophilic acetogens are of significance, since their use would reduce gas cooling requirements, allow for cost-efficient recovery of products with relatively low boiling point (Henstra et al., 2007; Redl et al., 2017), and decrease the risk of contamination.

A well-studied syngas-fermenting thermophile is *Moorella thermoacetica*. The species was isolated from horse feces in 1942 and named *Clostridium thermoaceticum* (Fontaine et al., 1942). The taxonomy of the genus *Clostridium* was restructured in 1994 and *C. thermoaceticum* was transferred to a new genus *Moorella* as *M. thermoacetica* (Collins et al., 1994). Several strains originating from the cultures isolated by Fontaine et al. (1942) are deposited in strain collections. The type strain DSM 521^T and the strain ATCC 39073 have primarily served to elucidate the primary metabolism of *M. thermoacetica* (synonym *C. thermoaceticum*): they were used in experiments to study carbohydrate utilization (Andreesen et al., 1973), the acetate kinase (Schaupp and Ljungdahl, 1974), cytochromes and menaquinones (Gottwald et al., 1975), the formate dehydrogenase (Ljungdahl and Andreesen, 1977), and the utilization of CO (Diekert and Thauer, 1978). The genome of the non-type strain ATCC 39073 was sequenced in 2008 (Pierce et al., 2008) and the genome sequence of the type strain DSM 521^T followed in 2015 (Poehlein et al., 2015). A spore sample of the original *M. thermoacetica* strain isolated in 1942 was deposited by Kerby and Zeikus (1983) as a second representative of the type strain (DSM 2955^T) in the DSMZ (Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Brunswick, Germany). It was shown to utilize H₂/CO₂ as substrate and was also adapted to growth on CO (Kerby and Zeikus, 1983). The ability to utilize gaseous substrates was not shown for ATCC 39073 and DSM 521^T until 1990 (Daniel et al., 1990). Another *M. thermoacetica* strain (Y72) with higher transformation efficiency than ATCC 39073 was described and its draft genome published in 2014 (Tsukahara et al., 2014).

Wiegel et al. (1981) described the isolation of strains closely related to the already known *C. thermoaceticum* (*M. thermoacetica*) strains. The novel strains were shown to grow chemolithotrophically on H₂/CO₂ and chemoheterotrophically on several carbon sources. At that time, the aforementioned strains of *C. thermoaceticum* (*M. thermoacetica*) were not known to utilize H₂/CO₂ and CO. Furthermore, Wiegel et al. (1981) described differences in the cell shape in comparison to *M. thermoacetica*. In addition to *C. aceticum* and *Acetobacterium woodii*, this new strain was the third species known to grow autotrophically using H₂ and CO₂ while producing acetate. Therefore, a new species was proposed and a strain isolated from a Yellowstone hot spring (strain JW 701/3) was deposited as *Clostridium thermoautotrophicum* DSM 1974^T (Wiegel et al., 1981). *C. thermoautotrophicum* was later re-classified as *Moorella thermoautotrophica* in the extensive study of Collins et al. (1994). In addition to *M. thermoautotrophica* DSM 1974^T, which is

the designated type strain, a second *M. thermoautotrophica* strain, DSM 7417, is available. This strain (DSM 7417) was first described in Rijssel et al. (1992) when it appeared as a contamination in a continuous culture. The authors based their decision to place the newly described strain in the species of *M. thermoautotrophica* instead of *M. thermoacetica* mainly on observations regarding the cell shape (Rijssel et al., 1992). Recently, Kimura et al. (2016) requested an opinion regarding the taxonomic status of *M. thermoautotrophica*. Based on DNA–DNA hybridization experiments and 16S rRNA gene sequence analysis, Kimura et al. (2016) concluded that the species *M. thermoautotrophica* should be reclassified as *M. thermoacetica*. Over time, phenotypic differences between *M. thermoacetica* and *M. thermoautotrophica* were described, but often with partly conflicting results (Cato et al., 1986; Das et al., 1989; Yamamoto et al., 1998; Carlier and Bedora-Faure, 2006).

Here, we report that *M. thermoautotrophica* DSM 1974^T is a mixed culture of at least two strains, which we isolated. We sequenced the genome of those two strains as well as the genome of DSM 7417 and nine other *M. thermoacetica* strains, thereby considerably adding to the genomic information of this group of bacteria. We compared the genomes of the strains with the genome of the *M. thermoacetica* strain ATCC 39073 (Pierce et al., 2008) and the type strains DSM 2955^T (Bengelsdorf et al., 2015) and DSM 521^T (Poehlein et al., 2015). In addition, we performed genome comparison with all other genomes of the genus *Moorella*. Furthermore, differences in carbon utilization of the aforementioned strains were characterized. Based on this study, we conclude that the classification of the two strains isolated from DSM 1974^T as a separate species, *M. thermoautotrophica*, is not justified and that based on the data collected both strains should be reclassified as strains of the species *M. thermoacetica*. However, a problem arises due to the fact that the designated type strain deposited in the DSMZ, as DSM 1974^T, appears to be a mixture of two strains. The implications of these findings within the context of the rules of the International Code of Nomenclature (Parker et al., 2019) together with the content of the recent Request for an Opinion of Kimura et al. (2016) are discussed.

MATERIALS AND METHODS

Strains

The strains DSM 521^T, DSM 2955^T, DSM 7417, DSM 21394, DSM 11768, DSM 12797, DSM 12993, DSM 6867, and DSM 11254^T were purchased from DSMZ (Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Brunswick, Germany). The strains isolated from the culture of DSM 1974^T obtained from the DSMZ were deposited at the DSMZ with the numbers DSM 103284 (DSM 1974-Ulm) and DSM 103132 (DSM 1974-HH). Strain ATCC 39073 was purchased from the ATCC (Manassas, VA, United States) and was maintained by a series of transfers (here labeled as ATCC 39073-HH). Prior to extracting DNA for genome sequencing, a single colony was isolated on solid medium.

Cultivation

Strains were cultivated in 50-ml serum bottles (50% filled) closed with butyl rubber stoppers (bottles and stoppers: Ochs, Germany) containing a magnetic stirring bar and medium with the following composition (in g/l) [13]: KH_2PO_4 (0.5); NH_4Cl (0.4); NaCl (0.4); NaHCO_3 (3.5); yeast extract (0.5); 1% trace element solution was added to the medium. The trace element solution was prepared with 2 g/l nitrilotriacetic acid; the pH adjusted to 6.0 with KOH, and the following compounds added (in mg/l): $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ (1000); $\text{Fe}(\text{SO}_4)_2(\text{NH}_4)_2 \cdot 6 \text{H}_2\text{O}$ (800); $\text{CoCl}_2 \cdot 6 \text{H}_2\text{O}$ (200); $\text{ZnSO}_4 \cdot 7 \text{H}_2\text{O}$ (200); $\text{CuCl}_2 \cdot 2 \text{H}_2\text{O}$ (20); $\text{NiCl}_2 \cdot 6 \text{H}_2\text{O}$ (20); $\text{Na}_2\text{MoO}_4 \cdot 2 \text{H}_2\text{O}$ (20); Na_2SeO_4 (20); Na_2WO_4 (20) mg. The pH of the culture medium was adjusted to 6.5, flushed with $\text{N}_2:\text{CO}_2$ (80:20) and autoclaved at 140°C for 40 min. Solid medium contained 1% GelzanTM and the medium was sterilized at 120°C for 20 min. The following sterile stock solutions were added after autoclaving: CaCl_2 (50 mg/l final), MgCl_2 (330 mg/l final), vitamin solution (1%), cysteine-HCl (1 mM final). The vitamin solution contained (mg/l): biotin (2); folic acid (2); pyridoxine-HCl (10); thiamine HCl (5); riboflavin (5); nicotinic acid (5); calcium D-(+)-pantothenate (5); vitamin B₁₂ (0.5); *p*-aminobenzoic acid (5); thioctic acid (5). The medium was pre-warmed before inoculation. The strains were cultivated at 60°C with stirring at 350 rpm. Fructose as carbon and energy source was added at a final concentration of 60 mM to the medium. The headspace was pressurized with $\text{N}_2:\text{CO}_2$ (80:20) to 3 bar. When gases served as carbon and energy sources, the headspace was flushed for several minutes before inoculation with the gas mixture, and the headspace pressurized to 3 bar after inoculation. $\text{H}_2:\text{CO}_2$ (80:20) served as gaseous substrates. Strain DSM 103132 was isolated from DSM 1974^T using the medium described above solidified with 1% GelzanTM and using 60 mM fructose as the substrate. Strain DSM 103284 was isolated from DSM 1974^T using the DSMZ medium 135, the solid medium contained 1.5% agar. In both cases, single colonies were picked and used for further cultivation.

Extraction of Genomic DNA

Cultures in mid-exponential phase were sampled, the cells were spun down, and DNA was extracted using the Wizard[®] Genomic DNA Purification Kit (Promega, Madison, WI, United States) and the MasterPureTM Gram Positive DNA Purification Kit (Epicentre, Madison, WI, United States) according to the manufacturer's protocol. DNA was quantified using the Qubit dsDNA HS Assay Kit with the Qubit 2.0 fluorometer (Thermo Fisher Scientific, Waltham, MA, United States).

Genome Sequencing

ATCC 39073-HH and DSM 103132 were sequenced using a PacBio RSII instrument (Pacific Biosciences, Menlo Park, CA, United States). SMRTbellsTM libraries were constructed and sequenced following the recommended Pacific Biosciences template preparation protocol. Following SMRTbellTM construction, v2 primers and P4 polymerase were annealed and enzyme bound complexes attached to magnetic beads for loading. Each SMRTbellTM library was loaded onto a SMRT

cell and sequenced on the PacBio RSII. The average reference coverage was above 500 for both strains, resulting from 129,760 and 134,994 reads of ATCC 39073-HH and DSM 103132, respectively, with an average read length of approximately 12,000 bp. Isolated DNA from all remaining strains was used to generate Illumina shotgun sequencing libraries. Sequencing was performed by employing a MiSeq system using MiSeq Reagent Kit v3 (600 cycles), as recommended by the manufacturer (Illumina, San Diego, CA, United States), resulting in 2×300 bp paired end reads. Strain DSM 103284 was sequenced with the Genome Analyzer IIx (Illumina, San Diego, CA, United States) resulting in 2×112 bp paired end reads. Quality filtering of the raw reads was done using Trimmomatic version 0.32 (Bolger et al., 2014). The *de novo* assembly was performed with the SPAdes genome assembler software (Bankevich et al., 2012). The assembly was validated and the read coverage determined with QualiMap (García-Alcalde et al., 2012). For scaffolding the contigs of strain DSM 103284, we used the Move Contigs tool of the Mauve Genome Alignment Software (Darling et al., 2010). Additionally, contigs that could not be ordered with Mauve were examined via Gene Ortholog Neighborhoods based on bidirectional best hits implemented at the IMG-ER (Integrated Microbial Genomes-Expert Review) system (Markowitz et al., 2013). For contig ordering, the genomes of *M. thermoacetica* DSM 521^T (CP012369) and DSM 2955^T (CP012370) were used as references. Sequence gaps were closed by PCR-based techniques and primer walking with conventional Sanger sequencing, using BigDye 3.0 chemistry on an ABI3730XL capillary sequencer (Applied Biosystems, Life Technologies GmbH, Darmstadt, Germany), and employing the Gap4 (v.4.11) software of the Staden Package (Staden et al., 1999). *M. glycerini* DSM 11254^T has been sequenced using a combined approach with Illumina short read and Oxford Nanopore long read technology. Therefore, high molecular weight DNA (HWD) was isolated with the MasterPure Complete DNA & RNA Purification Kit (Biozym, Hessisch Oldendorf, Germany) as recommended by the manufacturer. Quality of isolated DNA was initially checked by agarose gel electrophoresis and validated on an Agilent Bioanalyzer 2100 using an Agilent DNA 12000 Kit as recommended by the manufacturer (Agilent Technologies, Waldbronn, Germany). Concentration and purity of the isolated DNA was first checked with a Nanodrop ND-1000 (PeqLab Erlangen, Germany), and exact concentration was determined using the Qubit[®] dsDNA HS Assay Kit as recommended by the manufacturer (Life Technologies GmbH, Darmstadt, Germany). Illumina shotgun libraries were prepared using the Nextera XT DNA Sample Preparation Kit and subsequently sequenced on a MiSeq system with the reagent kit v3 with 600 cycles (Illumina, San Diego, CA, United States) as recommended by the manufacturer resulting in 1,694,377 paired end reads. For Nanopore sequencing, 1.5 μg HWD was used for library preparation using the Ligation Sequencing Kit 1D (SQK-LSK109) and the Native Barcode Expansion Kit (EXP-NBD104) as recommended by the manufacturer. Sequencing was performed on a MinION device Mk1B using a SpotON Flow Cell R9.4.1 as recommended by the manufacturer for 72 h. This resulted in 162,721 reads with a mean read length of 4,155 bp.

Unicycler v0.4.8 (Wick et al., 2017) was used with default settings to perform a hybrid assembly.

Genome Annotation

The genomes were annotated using the Prokka automatic annotation software (Seemann, 2014). Protein coding, rRNA, and tRNA sequences were annotated using Prodigal (Hyatt et al., 2010), RNAmmer (Lagesen et al., 2007), and Aragorn (Laslett and Canback, 2004) against databases using BLAST (Camacho et al., 2009) and HMMER (Finn et al., 2011). Prediction of non-coding RNAs and CRISPR repeats were done by infernal (Nawrocki and Eddy, 2013) and MinCED¹ based on CRISPR Recognition Tool (Bland et al., 2007). Signal peptides were searched using SignalP (Petersen et al., 2011). Protein coding genes were analyzed for COG (Tatusov et al., 1997; Galperin et al., 2014) functional annotation using Batch CD-search tool (Marchler-Bauer and Bryant, 2004). No secondary metabolite clusters were predicted in an analysis using antiSMASH 2.0 (Medema et al., 2011). The genome sequences and annotation data of all *M. thermoacetica*

strains have been deposited in DDBJ/ENA/GenBank, for detailed information see **Table 1**.

Genome Analysis

For MLSA and gene content analysis, total protein sequences from the 24 genomes were extracted from the corresponding GenBank files using cds_extractor.pl v0.6² and used for downstream analysis with an in-house pipeline at the Göttingen Genomics Laboratory. In detail, proteinortho version 4.25 (default specification: blast = blastp v2.2.24, E-value = 1e-10, alg.-conn. = 0.1, coverage = 0.5, percent_identity = 50, adaptive_similarity = 0.95, inc_pairs = 1, inc_singles = 1, selfblast = 1, unambiguous = 0) (Lechner et al., 2011) was used to generate clusters of orthologs groups, inparalogs were removed. MUSCLE (Edgar, 2004) was employed to align the remaining sequences and poorly aligned positions were automatically filtered from the alignments using Gblocks (Castresana, 2000). A maximum likelihood tree from 1,177 orthologs groups was inferred with 500 bootstraps with RAXML

¹<https://github.com/ctSkennerton/minced>

²<https://github.com/aleimba/bac-genomics-scripts>

TABLE 1 | Genome features of *Moorella* species.

Organism	Accession number	Size [bp]	GC-content [%]	Coding percentage [%]	CDS	Genes	rRNA	tRNA	Contigs	References
<i>M. glycerini</i> DSM 11254 ^T	CP046244 (chromosome) CP046245 (plasmid)	3,559,463	54.74	88.26	3,509	3,564	3	52	2	This study
<i>M. glycerini</i> NMP	CELZ01000000	3,577,805	53.80	88.48	3,636	3,697	4	57	73	Liebensteiner et al., 2015
<i>M. humiferrea</i> DSM 23265 ^T	PVXM00000000	2,628,568	53.52	89.51	2,668	2,721	3	49	63	Poehlein et al., 2018b
<i>M. mulderi</i> DSM 14980 ^T	LTBC00000000	3,307,499	54.54	87.26	3,042	3,099	3	53	72	Castillo Villamizar and Poehlein, 2016
<i>M. perchloratireducens</i> An10	Gp0011525 ^a	3,307,499	53.84	88.69	3,349	3,423	3	52	133	Markowitz et al., 2013
<i>M. stamsii</i> DSM 26217 ^T	PVXL00000000	3,328,173	53.81	87.87	3,306	3,358	3	49	82	Poehlein et al., 2018a
<i>M. thermoacetica</i> ATCC 31490	VCDV00000000	2,616,798	55.81	87.94	2,621	2,676	3	51	26	This study
<i>M. thermoacetica</i> ATCC 33924 ^b	VCDY00000000	2,914,842	55.12	87.81	2,959	3,020	3	57	47	This study
<i>M. thermoacetica</i> ATCC 35608	VCDW00000000	2,611,625	55.83	87.96	2,605	2,661	3	52	26	This study
<i>M. thermoacetica</i> ATCC 39073	CP000232	2,628,784	55.79	86.39	2,465	2,634	3	51	1	Pierce et al., 2008
<i>M. thermoacetica</i> ATCC 49707	VCDX00000000	2,616,845	55.83	87.92	2,619	2,676	4	52	28	This study
<i>M. thermoacetica</i> DSM 103132	CP017019	2,976,077	55.10 ^c	87.94	3,026	3,091	6	58	1	This study
<i>M. thermoacetica</i> DSM 103284	CP017237	2,560,375	55.94 ^c	87.94	2,525	2,58	3	51	1	This study
<i>M. thermoacetica</i> DSM 11768	MIHH00000000	2,851,436	55.66	86.40	2,806	2,865	3	55	92	This study
<i>M. thermoacetica</i> DSM 12797	MIIF00000000	2,746,010	55.54	86.51	2,716	2,774	3	54	83	This study
<i>M. thermoacetica</i> DSM 12993	MDDD00000000	2,648,948	55.74	87.10	2,6	2,659	3	55	40	This study
<i>M. thermoacetica</i> DSM 21394	MDDC00000000	2,567,468	56.95	85.80	2,499	2,559	3	56	30	This study
<i>M. thermoacetica</i> DSM 2955 ^T	CP012370	2,623,349	55.81	88.17	2,624	2,68	3	52	1	Bengelsdorf et al., 2015
<i>M. thermoacetica</i> DSM 512 ^T	CP012369	2,527,564	55.95 ^d	88.06	2,553	2,609	3	52	1	Poehlein et al., 2015
<i>M. thermoacetica</i> DSM 6867	Mddb00000000	2,617,097	55.83	87.80	2,62	2,676	3	52	42	This study
<i>M. thermoautotrophica</i> DSM 7417	MDDE00000000	2,585,122	55.87	87.59	2,558	2,62	6	55	26	This study
<i>M. thermoacetica</i> Y72	BARR00000000	2,603,418	55.89	88.05	2,629	2,715	3	51	95	Tsukahara et al., 2014

^aGOLD Sequencing Project ID (no accession number available). ^bDNA isolated and sequenced from a single colony. ^cWiegel et al. (1981) stated a GC content of 53–55% in the DNA of strain DSM 1974^T. ^dMatteuzzi et al. (1978) stated a GC content of 54% in the DNA of strain DSM 521^T.

(Stamatakis, 2014). The script PO_2_MLSA.py is available at github³. Visualization of the tree was performed using Dendroscop (Huson and Scornavacca, 2012).

Average Nucleotide Identity (ANIm) analyses were performed using pyani.py⁴. Briefly, nucleotide sequences were extracted from the corresponding GenBank files using seq_format-converter.pl v0.2⁵ and subsequently used to run pyani in ANIm mode (uses MUMmer/NUCmer) to align input sequences. PHASTER (PHAge Search Tool Enhanced Release, Arndt et al., 2016) has been used for the detection of prophage regions. The analysis of genomic islands was performed using IslandViewer 4 (Bertelli et al., 2017).

RESULTS

Strain DSM 1974^T was purchased from DSMZ by our labs (University of Ulm and Technical University of Denmark) separately in 2015. Genome sequencing of the strain in the Göttingen Genomics Laboratory and at the Technical University of Denmark suggested that DSM 1974^T is a mixed culture. After suspecting cross-contamination in our labs, new DSM 1974^T cultures were ordered from DSMZ, however, with the same result. We independently isolated single clones after cultivation of DSM 1974^T on solid medium: DSM 103284 (DSM 1974-Ulm) at the University of Ulm and DSM 103132 (DSM 1974-HH) at the Technical University of Denmark as described in Section “Materials and Methods.” We sequenced the genome of both strains which were derived from the DSM 1974^T culture, as well as the genome of DSM 7417 and the genome of another ATCC 39073 strain, here designated ATCC 39073-HH. In order to determine whether DSM 1974^T is a mixed culture we ordered ATCC 33924^T (that is derived from DSM 1974^T) from the ATCC and sequenced the DNA directly isolated from the freeze-dried culture (data not shown) and from a single colony isolated with the same procedure as for strain DSM 103132. Sequencing results confirmed that ATCC 33924^T = DSM 1974^T deposited at the ATCC is also a mixed culture and the strain isolated from that culture is identical to DSM 103132. The differences between DSM 103132 (isolated in Denmark) and DSM 103284 (isolated in Germany) suggest that slightly different cultivation conditions may favor the selection of different strains from the original mixed culture of DSM 1974^T. In addition, the genomes of 10 different *M. thermoacetica* strains, *M. thermoautotrophica* DSM 7417, and *M. glycerini* DSM 11254^T were sequenced.

Genome Features

Table 1 shows an overview of the *de novo* sequenced genomes of the DSM 1974^T-derived strains (DSM 103284 and DSM 103132) and all other strain sequences in this study compared to the published genomes of type strains DSM 521^T, DSM 2955^T, as well as ATCC 39073, *M. thermoacetica* Y72, *M. glycerini* DSM 11254^T, *M. glycerini* NMP, *M. humiferrea* DSM 23268^T,

M. mulderi DSM 14980^T, *M. perchloratireducens* An10, and *M. stamsii* DSM 26217^T. In order to investigate the phylogeny of the strains, we first compared the 16S rRNA gene sequences of the type strains, ATCC 39073-HH, DSM 103132, and DSM 103284. The sequence similarity between the strains in the 16S rRNA gene region is at least 99.74%, as no more than 3 nucleotide mismatches could be found. In strains DSM 103284 and ATCC 39073-HH, the gene regions are identical. According to Stackebrandt and Goebel (1994), bacteria showing less than 97% similarity in their 16S rRNA gene sequences belong to different species, while additional methods must be taken into consideration when the 16S rRNA similarity values are above 97%. All strains were analyzed with respect to prophages and interestingly none of the strains harbors a complete prophage. In all strains, a different number of incomplete phages (between 1 and 5; for details see **Supplementary Table S2**) were detected. Six strains, *M. thermoacetica* DSM 103132 and DSM 103284, *M. glycerini* DSM 11254^T and NMP, *M. perchloratireducens* An10, and *M. stamsii* DSM 26217^T contain putative phages, marked as “questionable.” These DNA regions show similarity to different *Bacillus* phages or to a Stx2-converting phage (**Supplementary Table S2**). We also checked some completely sequenced strains for the presence of genomic islands and found 9 such regions in strains DSM 103284 and DSM 2955^T as well as 10 genomic islands in strain DSM 521^T and ATCC 39073. Strain DSM 103132 harbors 36 genomic islands in total and one of these regions has a size of 166 kbp (**Supplementary Table S3**). All genomic islands contain mainly hypothetical proteins, transposases, or transcriptional regulators and only a few genes coding for enzymes (for details see **Supplementary Table S3**). With respect to plasmids, a 50-kbp plasmid was found in *M. glycerini* DSM 11254^T. None of the *M. thermoacetica* or *M. thermoautotrophica* strains was found to carry a plasmid. All other *Moorella* species could not be analyzed in detail, as they are draft genomes and there is no evidence for plasmid replication genes in these genomes.

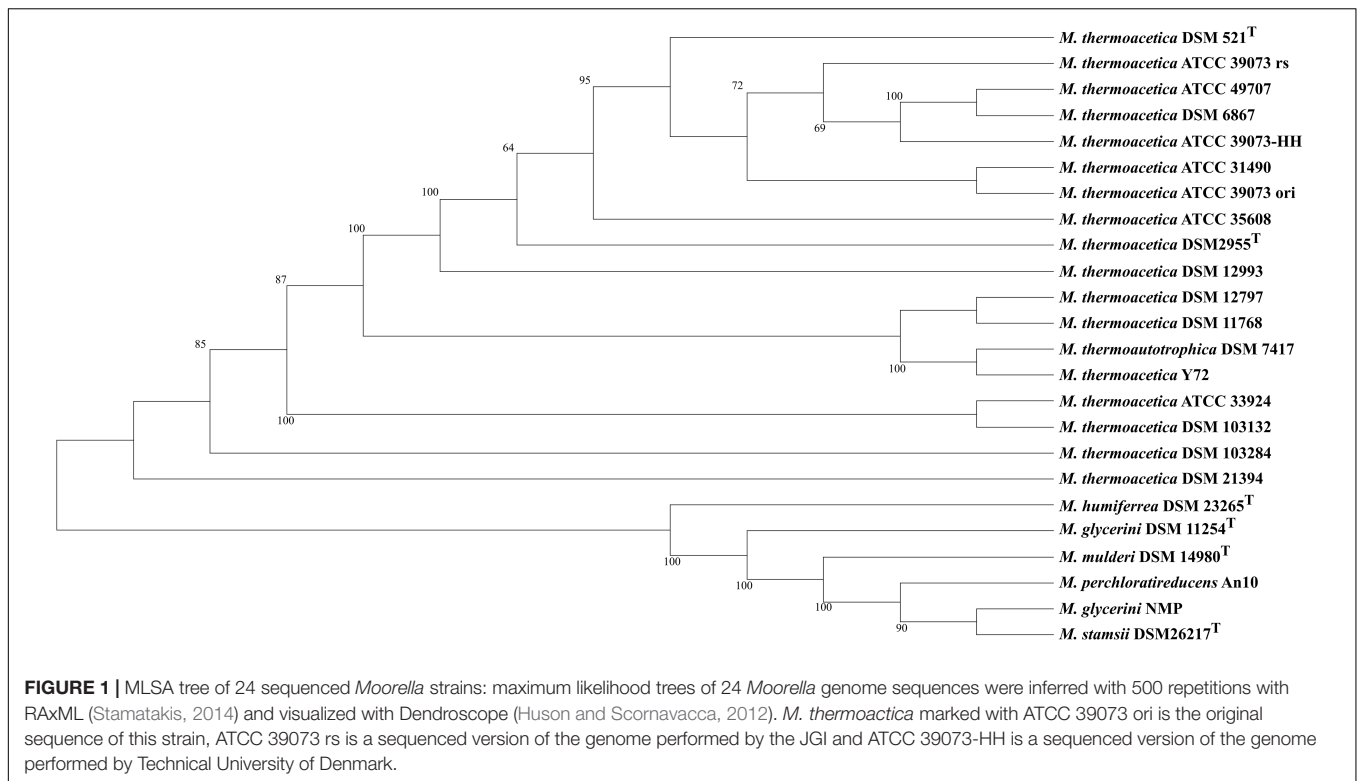
Phylogenetic Analysis

We used MLSA based on the detected core genome (1,177 OGs excluding paralogs) to perform phylogenetic analysis of our strains (**Figure 1**) and an average nucleotide identity analysis (ANIm) (**Figure 2**). The phylogenetic tree yielded two main clades, one consisting of all *M. thermoacetica* and *M. thermoautotrophica* strains and the second of *M. glycerini* DSM 11254^T, *M. glycerini* NMP, *M. humiferrea* DSM 23268^T, *M. mulderi* DSM 14980^T, *M. perchloratireducens* An10, and *M. stamsii* DSM 26217^T. The first main clade shows three distinct subclades, one consisting of the different versions of *M. thermoacetica* ATCC 39079, DSM 521^T, DSM 2955^T, ATCC 49707, ATCC 31490, ATCC 35608^T, DSM 12993 and DSM 6867. It should be noted that DSM 521^T, DSM 2955^T, ATCC 35608^T, and ATCC 49707 are all derived from the same original strain. The second subclade consists of strains DSM 12797, DSM 11786, Y72, and DSM 7417. Strains ATCC 33924 and DSM 103132 form the third subclade. Strain DSM 103284, isolated from the mixed culture of DSM 1974^T and strain DSM 21394 cluster outside of the three subclades. Interestingly, DSM 21394

³<https://github.com/jvollme>

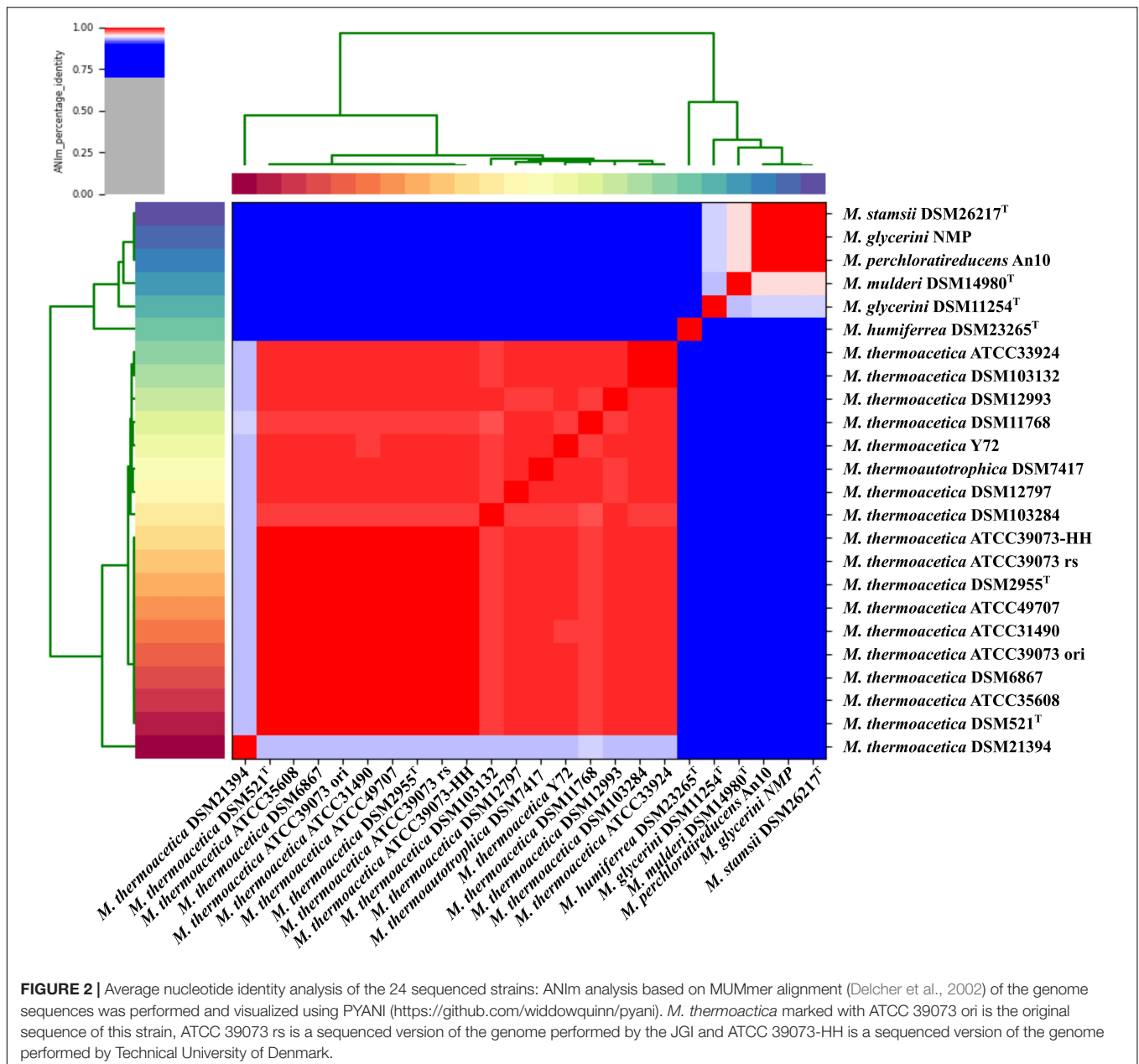
⁴<https://github.com/widdowquinn/pyani>

⁵<https://github.com/aleimba/bac-genomics-scripts>



is the strain with the third highest number of singletons (300 OGs). Whilst MLSA can provide insight into the phylogenetic relationship of organisms, for taxonomic studies there is a requirement for other methods, such as ANI analysis (Richter and Rosselló-Móra, 2009), which is a suitable *in silico* alternative for DNA–DNA hybridization (Goris et al., 2007). We performed an ANI analysis based on MUMmer alignment (ANIm) of the 24 genomes to define species and their complexes (Figure 2). We identified a large cluster comprising all *M. thermoacetica* and *M. thermoautotrophica* strains including DSM 103132 and DSM 103284, which have been both re-isolated from DSM 1974^T as well as DSM 7417. The latter two strains are currently considered to be *M. thermoautotrophica* strains. However, our analysis clearly shows that these strains would be more appropriately classified as *M. thermoacetica* isolates, since we identified ANIm values between 98 and 99% compared to *M. thermoacetica* DSM 512^T and DSM 2955^T, the two independent deposits of the type strain of this species in the DSMZ (Supplementary Table S1). Richter and Rosselló-Móra (2009) proposed a threshold for the species boundary of 95% ANI, making reference to both ANIb and ANIm values. However, careful examination of their original data suggests that ANIb and ANIm do not give the same values and the species boundary for the two may be different. ANIm values of 98–99% are clearly above this threshold, but values of 95–96% need to be taken with caution. Our analysis also revealed that strain DSM 21394 has an ANIm value of 94% (Supplementary Table S1) compared to the other *M. thermoacetica* strains, which is below the threshold for the species boundary and further studies would be needed to

determine whether this strain should also be re-classified. This is also depicted in Figure 2, where all strains belonging to one species are marked in red tones. Interestingly, *M. stamsii* DSM 26217^T, *M. glycerini* NMP, and *M. perchloratireducens* An10 showed an ANIm value of 100% and they should therefore belong to the same species. However, the name *M. perchloratireducens* has not been validly published and *M. glycerini* NMP is not the nomenclatural type of the species so no formal nomenclatural action is required under the International Code of Nomenclature of Prokaryotes (Parker et al., 2019), since the names *M. glycerini* and *M. perchloratireducens* can only be formally considered to be heterotypic synonyms if both are validly published and are the corresponding nomenclatural types. The ANIm value of the type strain of *M. stamsii* to the other two strains is also 100%, indicating that all three should be placed in the same species, i.e., *M. stamsii*, which has been validly published. These results are in contrast to the published viewpoint that *M. stamsii* and *M. perchloratireducens* represent distinct species. It is common practice to determine the 16S rRNA gene sequence of a novel isolate and initially investigate the similarity value to the 16S rRNA gene sequences of other type strains before deciding how to further characterize a strain. In the case of 16S rRNA gene sequence similarity values of 97% and greater it is common practice to determine DNA–DNA hybridization values (which is now being replaced by ANI or digital DNA–DNA hybridization studies) to evaluate whether one is dealing with a new species. Where the 16S rRNA gene sequence similarity values are less than 97%, it is generally assumed that one has a novel species. Key discrepancies in the study of *M. glycerini*, *M. stamsii*, and



M. perchloratireducens are the 16S rRNA gene sequences and the genomic similarity. In the case of *M. glycerini*, the 16S rRNA gene sequence determined in the original study (Slobodkin et al., 1997), U82327, showed a pairwise similarity of 99.3% to the 16S rRNA sequence determined in the genome (CP046244). The 16S rRNA gene sequence determined in the original study of *M. stamsii* (Alves et al., 2013), HF563589, showed a pairwise similarity of 99.3% to the 16S rRNA sequence determined in the genome contig PVXL01000051. When U82327 and HF563589 were compared by Alves et al. (2013), the similarity values were 97%, but comparison of the 16S rRNA gene sequences obtained from the genomes (CP046244 and PVXL01000051) now gives 99.2% similarity and 100% similarity to CELZ01000013. In the

case of DNA–DNA hybridization between these two strains the value was 51.1–53.3% (duplicated measurements). The 16S rRNA gene sequence from the genome of *M. glycerini* DSM 11254^T contains a large deletion that does not occur in U82327 or any of the other PCR-amplified 16S rRNA gene sequences or those determined via genome sequencing of the same strain (**Supplementary Figure S1**). The PCR-amplified 16S rRNA gene from *M. stamsii* (HF563589) also appears to contain numerous additional bases.

In the case of *M. perchloratireducens*, comparison of the 16S rRNA gene sequence determined in the original study, EF060194 (Balk et al., 2008), with that extracted from the genome (Gp0011525) showed 95.1% similarity. While EF060194

showed 97% sequence similarity with the 16S rRNA gene sequence from the genome of *M. thermoacetica* ATCC 39073 (CP00232), comparisons with the 16S rRNA gene sequence from Gp0011525 indicated that the genome-derived sequences showed 95% sequence similarity. In contrast, comparisons between EF060194 (*M. perchloratireducens*) and U82327 (*M. glycerini*)/HF563589 (*M. stamsii*) gave sequence similarity values of 93.9 and 93.1%, respectively. However, comparisons based on the 16S rRNA gene sequences extracted from the genomes Gp0011525 (*M. perchloratireducens*), CP046244 (*M. glycerini*), and PVXL01000051 (*M. stamsii*) gave pairwise similarities of 99.2–100%. No DNA–DNA hybridization studies were carried out by Balk et al. (2008), because they used a 16S rRNA gene sequence “threshold” of 98% 16S rRNA similarity. These results suggest significant discrepancies between the 16S rRNA gene sequence EF060194 obtained by primer amplified sequencing and that determined by genome sequencing that are evident in the alignments (**Supplementary Figure S1**) and are difficult to attribute to experimental error without further confirmatory work. It is interesting to note that of the two deposits of *M. perchloratireducens* An10, ATCC BAA-1531 and JCM 14829 only ATCC BAA-1531 is currently available and is the source strain for the genome Gp0011525. In the case of *M. mulderi* DSM 14980, the genome-derived 16S rRNA gene sequence (LTBC01000042.1) contains a large insert not present in sequence of the original PCR-amplified gene deposited as AF487538.1 (**Supplementary Figure S1**).

The ANIm values between *M. glycerini* (strain NMP), *M. stamsii*, and *M. perchloratireducens* indicate that they belong to the same species. Although the 16S rRNA gene sequence of the type strains of *M. glycerini* and *M. stamsii* are 99.3%, the AMIm value of 94% indicates that they are different species. In the case of *M. mulderi* DSM 14980 the genome-based 16S rRNA gene sequence similarity to *M. glycerini* DSM 11254 is 98.8% and the AMIm value 93%, indicating that they are different species. When compared to the genome-based 16S rRNA gene sequences of *M. glycerini* (strain NMP), *M. stamsii*, and *M. perchloratireducens* the value is 99.3% and the ANIm value 96%; this would appear to indicate that *M. mulderi* DSM 14980 is a member of the same species as *M. stamsii* DSM 26217. However, the original work of Richter and Rosselló-Móra (2009) indicate that an ANI cut-off of 95% ANIb is equivalent to an ANIm value of 96.5%, indicating that *M. mulderi* DSM 14980 and *M. stamsii* DSM 26217 are not members of the same species. This work also indicates the importance of examining the data beyond simple similarity values, where examination of the individual 16S rRNA gene sequence alignments, the differences in gene content and genome size provide extra valuable detail.

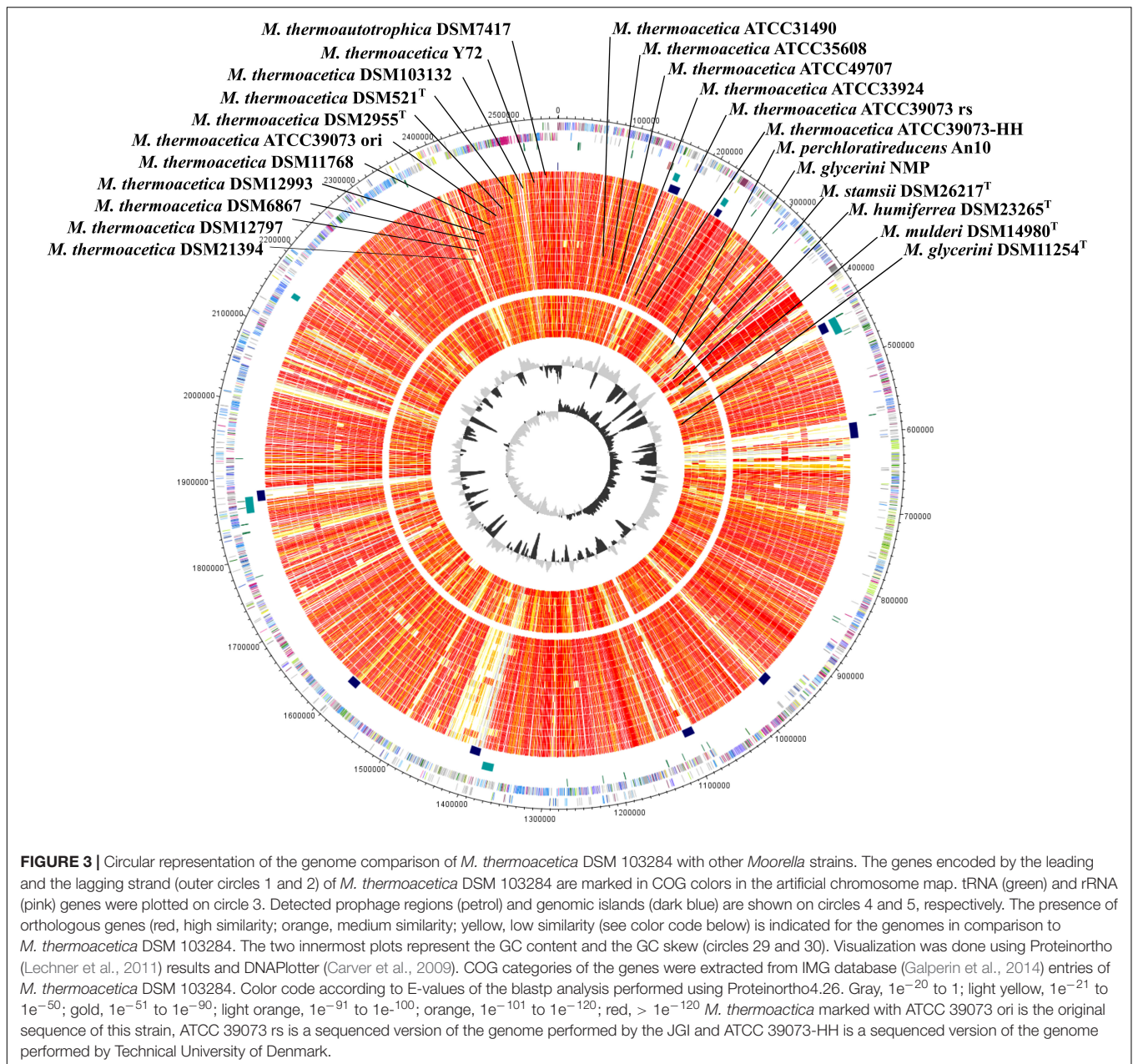
Genome Comparison

Until recently, only the sequence of the non-type strains *M. thermoacetica* ATCC 39079 and *M. thermoacetica* Y72 were publicly available, but many other strains, including the two independently deposited type strains of the species (DSM 521^T and DSM 2955^T), and several other strains are available at the German Collection of Microorganisms and Cell Cultures (DSMZ Brunswick), including strain DSM 1974^T. We sequenced

the genomes of all these strains and performed whole genome comparison of all *M. thermoacetica* strains, and comparison with the genomes of five other species, namely *M. stamsii* DSM 26217^T, *M. humiferrea* DSM 23265^T, *M. glycerini* DSM 11254^T, *M. glycerini* NMP, *M. perchloratireducens* An10, and *M. mulderi* DSM 14980^T (**Figure 3**). All *M. thermoacetica* strains have a comparable genome size of 2.52–2.64 Mb, except the two closely clustering strains DSM 103132 and ATCC 33924, which have larger genomes (2.98 and 2.91 Mb). *M. glycerini* NMP has the largest genome size in our comparison with 3.58 Mb, followed by *M. glycerini* DSM 11254^T with 3.56 Mb. A whole genome comparison based on protein encoding genes revealed a core genome shared by all 24 strains of 1,297 OGs including paralogs and a pan genome of 8,042 OGs (**Figure 4**). The pan genome includes the core and the flexible genome, OGs shared by at least two genomes, but not by all genomes in the comparison. The size of the core genomes is half the size of the complete genome of the *M. thermoacetica* strains, due to the high proportion of *M. thermoacetica* strains in our comparison. We found a broad range of singletons, meaning genome-specific genes, varying between 15 and 275 OGs in the *M. thermoacetica* group. The highest number of singletons (674 OGs) was found in the genome of *M. glycerini* DSM 11254^T. The flexible genome harbors for example a complete gene cluster encoding a pyruvate:ferredoxin oxidoreductase, which is only present in DSM 103284, DSM 11768, DSM 512^T, DSM 2955^T, DSM 12797, and all ATCC 39073 genomes. A cluster encoding an anaerobic dimethylsulfoxide reductase (DSMO reductase) is present in all genomes compared here, except of *M. mulderi* DSM 14980 and *M. thermoacetica* DSM 103132, which has been re-isolated from the mixed culture DSM 1974. We also identified OGs that are specific for the above-mentioned phylogenetic clades. We identified, for example, a gene cluster coding for a carbohydrate-specific ABC transport system, which is exclusively present in the first main clade comprising all *M. thermoacetica* strains, but which is absent in the second main clade consisting of *M. stamsii* DSM 26217^T, *M. humiferrea* DSM 23265^T, *M. glycerini* DSM 11254^T, *M. perchloratireducens* An10, and *M. mulderi* DSM 14980^T. We also identified gene clusters specific for the first main clade, for example a cluster encoding, amongst other genes, a ribose permease, L-rhamnose mutarotase, and a L-fucose isomerase probably involved in rhamnose and fucose metabolism. There are also genome-specific genes. *M. thermoacetica* DSM 103284 for example harbors a hydrogenase gene cluster that could not be identified in any other genome analyzed in this study.

Phenotypical and Physiological Differences Between *M. thermoacetica* and *M. thermoautotrophica* Strains

Several phenotypical and physiological differences between *M. thermoacetica* and *M. thermoautotrophica* strains regarding compounds that can serve as carbon and energy source have been described in the literature. Those results are sometimes contradictory to each other. We therefore tested whether there are differences between the strains regarding carbon source utilization and whether the results can give a hint toward



the identity of strain DSM 1974^T. Our results largely agree with the results reported in the literature. Within the tested strains, only DSM 103132 can utilize arabinose (Table 2). As already published, *M. thermoacetica* (DSM 521^T or ATCC 39073) (Fontaine et al., 1942; Andreesen et al., 1973; Cato et al., 1986) and DSM 1974^T (Wiegel et al., 1981; Cato et al., 1986) are not able to utilize arabinose. DSM 103132 was also found to be the only strain that could utilize formate, but only reaching low optical densities. DSM 1974^T has been reported to utilize formate (Wiegel et al., 1981; Fröstl et al., 1996), like ATCC 39073 (Fröstl et al., 1996). All tested strains were able to grow on fructose and glucose, and these substrates led to the highest cell density, which is in agreement with literature on

M. thermoacetica (DSM 521^T or ATCC 39073) (Fontaine et al., 1942; Andreesen et al., 1973). DSM 521^T was the only strain that did not utilize H₂ + CO₂ as carbon and energy source in our experiments. ATCC 39073 (Daniel et al., 1990; Fröstl et al., 1996) and DSM 1974^T (Wiegel et al., 1981; Savage and Drake, 1986; Fröstl et al., 1996) have been reported to utilize methanol. In agreement with our results, DSM 521^T has been reported not to grow on methanol (Cato et al., 1986). All strains tested in this study grew with pyruvate as energy and carbon source. Interestingly, Cato et al. (1986) indicated that pyruvate does not serve as a growth-supportive substrate for DSM 1974^T (Wiegel et al., 1981). None of the tested strains, except for DSM 103284, could utilize rhamnose, which is in line with DSM 1974^T being

TABLE 2 | Substrate utilization by selected *Moorella* strains.

		ATCC 39073-HH	DSM 103132	DSM 2955 ^T	DSM 521 ^T	DSM 103284	DSM 7417	Data for original strain DSM 1974 ^T (Wiegel et al., 1981)
Arabinose	60 mM	NG	0.42	NG	NG	NG	NG	NG
Formate	10 mM	NG	0.12	NG	NG	NG	NG	+
Fructose	60 mM	1.65	1.30	1.66	0.67	1.24	1.29	++
Glucose	60 mM	1.50	0.22	1.90	0.59	0.25	0.23	++
H ₂ /CO ₂	30 psi	0.11	0.22	0.45	NG	0.28	0.36	+
Methanol	60 mM	0.22	0.10	0.49	NG	0.50	0.26	+
Pyruvate	60 mM	0.39	0.33	0.28	0.33	0.56	0.27	NG
Rhamnose	60 mM	NG	NG	NG	NG	0.60	NG	NR
Sucrose	60 mM	1.60	0.26	NG	NG	NG	NG	NG
Xylose	60 mM	1.50	NG	0.88	0.44	0.70	0.70	NG

The highest optical densities are reported that were measured for the respective carbon sources. NG, no growth; NR, not reported; +, slow growth or low optical density; ++, fast growth, optical density at 600 nm above 1.0 with 0.5% (wt/vol) carbon source after 4 days.

complicated by the fact that DSM 1974^T, the strain which led to the proposal of the new species *C. thermoautotrophicum* (Wiegel et al., 1981) and was later transferred to the genus *Moorella* as *M. thermoautotrophica* (Collins et al., 1994) was consistently shown by genome sequencing to consist of two different strains. The isolation of two different strains that have subsequently been deposited as DSM 103132 and DSM 103284 confirms these observations. However, taking the original data of Wiegel et al. (1981) and comparing them with the data collected in this study for DSM 103132 and DSM 103284 does not show a large number of significant differences in the physiology of the strains. Based on the current data and taking into consideration the methods originally used by Wiegel et al. (1981) it is not possible to determine whether the original strain of Wiegel, JW 701/3, was a mixture of two different strains of the same species, whether the original strain was a pure culture, but a mixed culture was submitted for deposit (that methods used at the time would not have detected), or whether a second strain was introduced into the culture subsequent to accession to the DSMZ. Cross-contamination of strains is one possible explanation: the spores of *Moorella* species are highly heat-resistant and are not sufficiently inactivated by a standard autoclaving at 121°C (Fontaine et al., 1942). Byrer et al. (2000) for example described the strains JW/DB-2 and JW/DB-4 (ATCC number BAA-48) that show unusually heat-resistant spores. However, given the resolution of methods used at the time, one also cannot exclude with certainty that the original culture did not consist of more than one strain. One interesting aspect is that Wiegel et al. (1981) report that DNA–DNA hybridization supported the recognition of strains JW 701/3 and strain KIVU as members of the same species, but distinct from *C. thermoacetikum* (*M. thermoacetica*). Kimura et al. (2016) have previously reported a similar problem with the designated type strain of *M. thermoautotrophica*. Formulated as a Request for an Opinion, this limits any action that can be taken to a formal ruling by that body. However, their work concentrates largely on the interpretation of 16S rRNA gene sequences that appear to have been obtained by both cloning and the isolation of strains from the culture supplied. Representative partial sequences of the 16S rRNA genes of the seven groups obtained by cloning

and sequencing of the isolates have been deposited as LC133084–LC133087 and designated in the publication as representing OUT-1 to OUT-4 in that order, respectively. Kimura et al. (2016) concentrate on a single 16S rRNA gene sequence deposited as L09168 (from DSM 1974) and do not mention that additional sequences are available, X58353 and X77849. X58353 (strain JW 701/3; 1155 bases, but with numerous Ns) was deposited in 1990 from the University of Kiel and will not be considered further. X77849 was deposited in 1994 from the University of Reading in co-operation with Dr. Hippe (DSMZ curator of the strain at the time) and is derived from DSM 1974 and presumably directly from stocks held in the DSMZ. L09168 was deposited in 1993 from The University of Queensland. A direct alignment of the two sequences L09168 and X77849 indicates that, ignoring a small number of Ns in X77849, the two are not identical making it difficult to conclude whether either of the two can be considered to be a 100% accurate reflection of the original gene sequences from the same strain. Similarly, a comparison with the 16S rRNA sequences from Kimura et al. (2016) also indicate that neither of the two sequences (X77849 and L09168) (Supplementary Figure S2) show 100% similarity with those obtained by Kimura et al. (2016). It should also be remembered that the sequences X77849 and L09168 are only one part of the evidence that were not obtained directly when the type strain was originally described and “verification” of X77849 vs. L09168 does not allow one to conclude that one sequence is “correct” and the other in error. If one were to extend the reasoning of Kimura et al. (2016) to other similar cases one would conclude that given the differences between the 16S rRNA gene sequence obtained by direct amplification and that extracted from the genome of *M. stamsii* that the type strain does not exist. An even more dramatic example is the case of *Alterococcus agarolyticus* (Shieh and Jean, 1998) that started its taxonomic career as an atypical member of the *Enterobacteriaceae* (Shieh and Jean, 1998) under the 16S rRNA gene sequence AF075271.1 (deposited 19th June 1998) that was substituted for by AF075271.2 (deposited 21st August 2002) and is widely accepted as a member of the *Verrucomicrobia*. Under these circumstances, the nomenclatural type currently available certainly does not correspond to the 16S rRNA gene sequence originally deposited as AF075271.1

and one would have to conclude that the type strain no longer exists. However, put in context other data in the original publication clearly indicates that *Alterococcus agarolyticus* was an atypical member of the *Enterobacteriaceae* and that the original 16S rRNA gene sequence AF075271.1 is in error and should have been verified.

In the case of *M. thermoautotrophica*, comparison with the 16S rRNA gene sequence deposited as X77849 and L09168 also needs to be treated with caution if the original source culture (DSM 1974^T) was not a pure culture or where the quality/accuracy of gene sequencing technologies may have changed over the decades. No attempt was made to compare the physiological/biochemical properties of the strains studied by Kimura et al. (2016) with the original work of Wiegel et al. (1981) and relies solely on one older gene sequence (L09168) that is not corroborated by another sequence (X77849) obtained at about the same time from the same source culture, DSM 1974^T. Examining the 16S rRNA sequences deposited by Kimura et al. (2016) (LC133084-LC133087) against L09168, X77849 and those extracted from the genomes derived from subcultures of DSM 1974 and ATCC 33924 (including re-deposits as DSM 103132 and DSM 103284), i.e., CP017019.1 (positions 154745–156300 and 147549–149104), CP017237.1 (positions 144877–146432), and VCDX01000030.1 (positions 1667–112) indicates that toward the end of the single primer amplified partial sequences LC133085 and LC133086 gaps are present that are not otherwise present in any of the other sequences in a region that could be considered to be conserved (**Supplementary Figure S2**). These gaps have, therefore, not been taken into consideration in the analysis here. Kimura et al. (2016) do not provide alignments of sequences in support of their work and make it impossible to determine why they consider “none of the sequences were similar to *M. thermoautotrophica* DSM 1974^T (L09168),” when in fact they show only minimal differences in the alignments presented here. Although alignments are critical steps in the evaluation of sequence-based data (both nucleotide and amino acid based) they are rarely given, contrary to recommendations (Tindall et al., 2010), making the direct verification of the resulting interpretation via this critical step impossible and are therefore included in **Supplementary Figures S1, S2**. The sequence LC133087 appears to belong to a strain having the most similar 16S rRNA sequence to *M. humiferrea* strain 64_FGQ^T (GQ872425) and will not be considered further. In the alignment shown, CP017019.1 (positions 154745–156300), CP017019.1 (positions 147549–149104), VCDX01000030.1 (positions 1667–112), and LC133086.1 have a “T” at position 280 (alignment numbering, **Supplementary Figure S2**) while CP017237.1 (positions 144877–146432), LC133084.1, and LC133085.1 have a “C” at the same position. LC133086.1 differs from CP017019.1 (positions 154745–156300), CP017019.1 (positions 147549–149104), and VCDX01000030.1 (positions 1667–112) in having a “T” position 435 rather than a “C” that is present in all other sequences (**Supplementary Figure S2**). LC133084.1 appears to be identical in the aligned bases to CP017237.1 (positions 144877–146432), but LC133085.1 has an “A” at position 745 rather than a “G” that is present in all other sequences (**Supplementary Figure S2**). Based on these observations, the only organism

recovered in this study and that of Kimura et al. (2016) is that represented by LC133084.1 and CP017237.1 (DSM 103284). While this demonstrates the care that has to be taken in evaluating the interpretation of the data used by Kimura et al. (2016), the major problem that arises centers on the fact that the strains isolated by Kimura et al. (2016) have not been deposited in a culture collection and comparison with the original physiological and biochemical data published by Wiegel et al. (1981) cannot be made. Based on an evaluation of the 16S rRNA sequences determined previously and those determined here it is not possible to conclude that the type strain no longer exists, since it was deposited as DSM 1974 and ATCC 33924 and the 16S rRNA sequences deposited as X77849 and L09168 do not appear to be fully accurate.

The Request for an Opinion of Kimura et al. (2016) also misinterprets the wording of Rule 18c and draws incorrect conclusions. Tindall (2016) provided a detailed discussion of the incorrect interpretation of Rule 18c that was also applied by Kimura et al. (2016). Based on the evidence presented by Kimura et al. (2016) and that obtained in this work one cannot conclude that the nomenclatural type no longer exists, but rather there may be an issue with the purity of the culture deposited/currently available. The current study covers the physiological/biochemical properties of strains isolated from DSM 1974^T and expands on the genomic characterization of the strains studied. While it is clear that DSM 103132 and DSM 103284 (both derived from DSM 1974^T, the designated type strain of *M. thermoautotrophica*) are more appropriately considered to be members of the species *M. thermoacetica*, there is a formal nomenclatural issue that also needs to be addressed that requires reference to be made to the International Code of Nomenclature of Prokaryotes (Parker et al., 2019). Typically, the nomenclatural type of a species as defined in Rule 18a is an axenic culture, but there are instances where one component part of a syntrophic co-culture has been named and the co-culture accepted as the nomenclatural type (type strain). However, when mixed cultures or consortia are considered (see Rule 31a and 31b) and these are treated as a “single” biological entity, the names associated with them are not validly published and could be applied to *M. thermoautotrophica*. In the case of DSM 1974^T and ATCC 33924^T, although the strains currently in circulation appear to be a mixed culture, there is no unambiguous evidence that the parent culture, strain JW 701/3, was also a mixed culture. In contrast to the study of Kimura et al. (2016), it has been possible to study in greater detail pure cultures of strains isolated from DSM 1974^T (that is the parent deposit for all other culture collection strains) and subsequently deposited as DSM 103132 and DSM 103284. In both cases, the strains appear to be members of the species *M. thermoacetica*. One possible solution would be to designate one of them as a neotype, although based on the physiological and biochemical data presented here neither of the two strains (DSM 103132 or DSM 103284) can unambiguously be shown to be more similar in its properties than the other to the data originally published by Wiegel et al. (1981). Irrespective of which course of action is taken, it is clear that the culture of DSM 1974^T made available to the current authors contains strains that should be classified in the species *M. thermoacetica* leading to the logical conclusion that

DSM 103132 and DSM 103284 should be assigned to that species. This nomenclatural conclusion is inescapable, irrespective of whether one follows the arguments of Kimura et al. (2016), where the name *M. thermoautotrophica* would eventually be rejected, declared to not have been validly published, or whether one considers the names *M. thermoacetica* (Fontaine et al., 1942; Collins et al., 1994) and *M. thermoautotrophica* (Wiegel et al., 1981; Collins et al., 1994) to be heterotypic synonyms. In the latter case, priority is governed by Rule 23a, 38 and 42 where the dates of valid publication of the epithets are taken into consideration, i.e., *thermoacetica* Fontaine et al. (1942) has priority over *thermoautotrophica* Wiegel et al. (1981). This also leads to the use of the name *M. thermoacetica* (Fontaine et al., 1942; Collins et al., 1994) and recognition of *M. thermoautotrophica* (Wiegel et al., 1981; Collins et al., 1994) as the later heterotypic synonym when their respective nomenclatural types are considered to members of the same taxon. The current authors favor the latter course of action, but the Judicial Commission may also decide otherwise. Also, *M. thermoautotrophica* DSM 7417 should be reclassified as *M. thermoacetica* as well.

In addition to resolving the *M. thermoacetica/thermoautotrophica* problem, this comprehensive analysis of the genus *Moorella* by the study of a significant number of novel genome sequences and knowledge of phenotypic differences led to two other important conclusions. First, strain DSM 21394, currently still named *M. thermoacetica*, clearly does not belong to this species. Reclassification and renaming as a new species are required. Secondly, *M. glycerini* NMP, *M. stamsii* DSM 26217^T, and *M. perchloratireducens* cannot be distinguished at species level. Furthermore, *M. glycerini* NMP has been wrongly assigned as *M. glycerini* as this strain shows an ANIm value of 94% similarity compared to the type strain DSM 11254^T and is clearly a different species despite the high 16S rRNA gene sequence pairwise similarity of 99.7%. Based on the data presented here, *M. glycerini* NMP, *M. stamsii* DSM 26217^T, and *M. perchloratireducens* are all members of the same species. Although reclassification of these three strains may be required, caution needs to be exercised when one considers differences between the data reported here and that previously reported in the literature (Slobodkin et al., 1997; Balk et al., 2008; Alves et al., 2013), especially with regards to the 16S rRNA gene sequences and the genomic similarity inferred from DNA–DNA hybridization experiments vs. *in silico* comparisons.

REFERENCES

- Alves, J. I., van Gelder, A. H., Alves, M. M., Sousa, D. Z., and Plugge, C. M. (2013). *Moorella stamsii* sp. nov., a new anaerobic thermophilic hydrogenogenic carboxydrotroph isolated from digester sludge. *Int. J. Syst. Evol. Microbiol.* 63, 4072–4076. doi: 10.1099/ijss.0.050369-0
- Andreesen, J. R., Schaupp, A., Neurauder, C., Brown, A., and Ljungdahl, L. G. (1973). Fermentation of glucose, fructose, and xylose by *Clostridium thermoacetatum*: effect of metals on growth yield, enzymes, and the synthesis of acetate from CO₂. *J. Bacteriol.* 114, 743–751.
- Arndt, D., Grant, J. R., Marcu, A., Sajed, 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. doi: 10.1093/nar/gkw387

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the IMG, GenBank, NCBI.

AUTHOR CONTRIBUTIONS

SR, AP, FB, TJ, CJ, PD, and AN conceived and designed the experiments. SR, AP, CE, FB, and TJ performed the experiments. SR, AP, CE, FB, TJ, CJ, PD, and AN analyzed the data. SR, AP, FB, TJ, CJ, BT, RD, PD, and AN wrote the manuscript.

FUNDING

This work was supported by The Novo Nordisk Foundation to the Technical University of Denmark (grant number NNF10CC1016517) and a Ph.D. grant to SR from the People Programme (Marie Curie Actions) of the European Union Seventh Framework Programme FP7–People-2012-ITN, under grant agreement number 317058, “BACTORY.” TJ was supported by EU H2020’s Research and Innovation Programme under grant agreement number 731263 “AMBITION.” CJ was supported by a grant from the Novo Nordisk Foundation (grant number NNF15OC0015246). Support was also provided by the ERA-IB 7 Project “OBAC” (FKZ 031B0274) to the groups of RD and PD. Independent Research Fund Denmark under grant number 7026-00025B.

ACKNOWLEDGMENTS

We acknowledge the support for the PacBio sequencing given by Björn Hallström, Science for Life Laboratory (KTH) and conducted by Christian Tellgren-Roth (UGC, NGI). We thank Melanie Heinemann and Sarah Teresa Schüßler for technical support and Richard Egelkamp for help with graphic tools.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2019.03070/full#supplementary-material>

- Balk, M., van Gelder, T., Weelink, S. A., and Stams, A. J. M. (2008). (Per)chlorate reduction by the thermophilic bacterium *Moorella perchloratireducens* sp. nov., isolated from underground gas storage. *Appl. Environ. Microbiol.* 74, 403–409. doi: 10.1128/AEM.01743-07
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comp. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Bengelsdorf, F. R., Poehlein, A., Esser, C., Schiel-Bengelsdorf, B., Daniel, R., and Dürre, P. (2015). Complete genome sequence of the acetogenic bacterium *Moorella thermoacetica* DSM 2955^T. *Genome Announc.* 3:e1157-15. doi: 10.1128/genomeA.01157-15
- Bertelli, C., Laird, M. R., Williams, K. P., Simon Fraser University Research Computing Group, Lau, B. Y., Hoard, 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
- Bland, C., Ramsey, T. L., Sabree, F., Lowe, M., Brown, K., Kyrpidis, N. C., et al. (2007). CRISPR recognition tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. *BMC Bioinform.* 8:209. doi: 10.1186/1471-2105-8-209
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Byrer, D. E., Rainey, F. A., and Wiegel, J. (2000). Novel strains of *Moorella thermoacetica* form unusually heat-resistant spores. *Arch. Microbiol.* 174, 334–339. doi: 10.1007/s002030000211
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.* 10:1. doi: 10.1186/1471-2105-10-421
- Carlier, J.-P., and Bedora-Faure, M. (2006). Phenotypic and genotypic characterization of some *Moorella* sp. strains isolated from canned foods. *Syst. Appl. Microbiol.* 29, 581–588. doi: 10.1016/j.syapm.2006.01.002
- Carver, T., Thomson, N., Bleasby, A., Berriman, M., and Parkhill, J. (2009). DNAPlotter: circular and linear interactive genome visualization. *Bioinformatics* 25, 119–120. doi: 10.1093/bioinformatics/btn578
- Castillo Villamizar, G. A., and Poelein, A. (2016). Genome sequence of the acetogenic bacterium *Moorella mulderi* DSM 14980^T. *Genome Announc.* 4:e444-16. doi: 10.1128/genomeA.00444-16
- Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* 17, 540–552. doi: 10.1093/oxfordjournals.molbev.a026334
- Cato, E. P., George, W. L., and Finegold, S. M. (1986). Genus *Clostridium*. *Bergey's Manual of Syst. Bacteriol.* 2, 1141–1200.
- Collins, M. D., Lawson, P. A., Willems, A., Cordoba, J. J., Fernandez-Garayzabal, J., Garcia, P., et al. (1994). The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. *Int. J. Syst. Bacteriol.* 44, 812–826. doi: 10.1099/00207173-44-4-812
- Daniel, S. L., Hsu, T., Dean, S. I., and Drake, H. L. (1990). Characterization of the H₂- and CO-dependent chemolithotrophic potentials of the acetogens *Clostridium thermoacetum* and *Acetogenium kivui*. *J. Bacteriol.* 172, 4464–4471. doi: 10.1128/jb.172.8.4464-4471.1990
- Darling, A. E., Mau, B., and Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi: 10.1371/journal.pone.0011147
- Das, A., Hugenholz, J., van Halbeek, H., and Ljungdahl, L. G. (1989). Structure and function of a menaquinone involved in electron transport in membranes of *Clostridium thermoautotrophicum* and *Clostridium thermoacetum*. *J. Bacteriol.* 171, 5823–5829. doi: 10.1128/jb.171.11.5823-5829.1989
- Delcher, A. L., Phillippy, A., Carlton, J., and Salzberg, S. L. (2002). Fast algorithms for large-scale genome alignment and comparison. *Nucleic Acids Res.* 30, 2478–2483. doi: 10.1093/nar/30.11.2478
- Diekert, G. B., and Thauer, R. K. (1978). Carbon monoxide oxidation by *Clostridium thermoacetum* and *Clostridium formicoacetum*. *J. Bacteriol.* 136, 597–606.
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Finn, R. D., Clements, J., and Eddy, S. R. (2011). HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39, W29–W37. doi: 10.1093/nar/gkr367
- Fontaine, F. E., Peterson, W. H., McCoy, E., Johnson, M. J., and Ritter, G. J. (1942). A new type of glucose fermentation by *Clostridium thermoacetum*. *J. Bacteriol.* 43, 701–715.
- Fröstl, J. M., Seifritz, C., and Drake, H. L. (1996). Effect of nitrate on the autotrophic metabolism of the acetogens *Clostridium thermoautotrophicum* and *Clostridium thermoacetum*. *J. Bacteriol.* 178, 4597–4603. doi: 10.1128/jb.178.15.4597-4603.1996
- Galperin, M. Y., Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2014). Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* 43, D261–D269. doi: 10.1093/nar/gku1223
- García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L. M., Götz, S., Tarazona, S., et al. (2012). Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* 28, 2678–2679. doi: 10.1093/bioinformatics/bts503
- Goris, J., Konstantinidis, K. T., Klappenbach, J. A., Coenye, T., Vandamme, P., and Tiedje, J. M. (2007). DNA–DNA hybridization values and their relationship to whole-genome sequence similarities. *Int. J. Syst. Evol. Microbiol.* 57, 81–91. doi: 10.1099/ijms.0.64483-0
- Gottwald, M., Andreesen, J. R., LeGall, J., and Ljungdahl, L. G. (1975). Presence of cytochrome and menaquinone in *Clostridium formicoacetum* and *Clostridium thermoacetum*. *J. Bacteriol.* 122, 325–328.
- Henstra, A. M., Sipma, J., Rinzema, A., and Stams, A. J. M. (2007). Microbiology of synthesis gas fermentation for biofuel production. *Curr. Opin. Biotechnol.* 18, 200–206. doi: 10.1016/j.copbio.2007.03.008
- Huson, D. H., and Scornavacca, C. (2012). Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* 61, 1061–1067. doi: 10.1093/sysbio/sys062
- Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* 11:1. doi: 10.1186/1471-2105-11-119
- Kerby, R., and Zeikus, J. G. (1983). Growth of *Clostridium thermoacetum* on H₂/CO₂ or CO as energy source. *Curr. Microbiol.* 8, 27–30. doi: 10.1007/bf01567310
- Kimura, Z.-I., Hoshino, T., and Murakami, K. (2016). The status of the species *Moorella thermoautotrophica* Wiegel et al., 1981. Request for an Opinion. *Int. J. Syst. Evol. Microbiol.* 66, 3249–3251. doi: 10.1099/ijsem.0.001163
- Lagesen, K., Hallin, P., Rødland, E. A., Stærfield, H.-H., Rognes, T., and Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108. doi: 10.1093/nar/gkm160
- Laslett, D., and Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 32, 11–16. doi: 10.1093/nar/gkh152
- Latif, H., Zeidan, A. A., Nielsen, A. T., and Zengler, K. (2014). Trash to treasure: production of biofuels and commodity chemicals via syngas fermenting microorganisms. *Curr. Opin. Biotechnol.* 27, 79–87. doi: 10.1016/j.copbio.2013.12.001
- Lechner, M., Findeiß, S., Steiner, L., Marz, M., Stadler, P. F., and Prohaska, S. J. (2011). Proteinortho: detection of (co-) orthologs in large-scale analysis. *BMC Bioinform.* 12:1. doi: 10.1186/1471-2105-12-124
- Liebensteiner, M. G., Pinkse, M. W., Nijse, B., Verhaert, P. D., Tsesmetzis, N., Stams, A. J., et al. (2015). Perchlorate and chlorate reduction by the Crenarchaeon *Aeropyrum pernix* and two thermophilic Firmicutes. *Environ. Microbiol. Rep.* 7, 936–945. doi: 10.1111/1758-2229.12335
- Ljungdahl, L. G., and Andreesen, J. R. (1977). Formate dehydrogenase, a selenium-tungsten enzyme from *Clostridium thermoacetum*. *Methods Enzymol.* 53, 360–372. doi: 10.1016/s0076-6879(78)53042-5
- Marcellin, E., Behrendorff, J. B., Nagaraju, S., DeTissera, S., Segovia, S., Palfreyman, R. W., et al. (2016). Low carbon fuels and commodity chemicals from waste gases—systematic approach to understand energy metabolism in a model acetogen. *Green Chem.* 18, 3020–3028. doi: 10.1039/c5gc02708j
- Marchler-Bauer, A., and Bryant, S. H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 32, W327–W331.
- Markowitz, V. M., Chen, I.-M. A., Palaniappan, K., Chu, K., Szeto, E., Pillay, M., et al. (2013). IMG 4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res.* 42, D560–D567. doi: 10.1093/nar/gkt963
- Matteuzzi, D., Hollaus, F., and Biavati, B. (1978). Proposal of neotype for *Clostridium thermohydrosulfuricum* and the merging of *Clostridium tartarivorum* with *Clostridium thermosaccharolyticum*. *Int. J. Syst. Evol. Microbiol.* 28, 528–531. doi: 10.1099/00207713-28-4-528
- Medema, M. H., Blin, K., Cimermancic, P., Jager, V., de, Zakrzewski, P., et al. (2011). antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* 39, W339–W346. doi: 10.1093/nar/gkr466
- Nawrocki, E. P., and Eddy, S. R. (2013). Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29, 2933–2935. doi: 10.1093/bioinformatics/btt509
- Parker, C. T., Tindall, B. J., and Garrity, G. M. (2019). International Code of Nomenclature of Prokaryotes (2008 revision). *Int. J. Syst. Evol. Microbiol.* 69, S1–S111.
- Petersen, T. N., Brunak, S., Heijne, G., von, and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701

- Pierce, E., Xie, G., Barabote, R. D., Saunders, E., Han, C. S., Detter, J. C., et al. (2008). The complete genome sequence of *Moorella thermoacetica* (f. *Clostridium thermoaceticum*). *Environ. Microbiol.* 10, 2550–2573. doi: 10.1111/j.1462-2920.2008.01679.x
- Poehlein, A., Bengelsdorf, F. R., Esser, C., Schiel-Bengelsdorf, B., Daniel, R., and Dürre, P. (2015). Complete genome sequence of the type strain of the acetogenic bacterium *Moorella thermoacetica* DSM 521^T. *Genome Announc.* 3:e1159-15. doi: 10.1128/genomeA.01159-15
- Poehlein, A., Böer, T., Steensen, K., and Daniel, R. (2018a). Draft genome sequence of the hydrogenogenic carboxydrotroph *Moorella stamsii* DSM 26271. *Genome Announc.* 6:e345-18. doi: 10.1128/genomeA.00345-18
- Poehlein, A., Keyl, A., Milsch, J. C., and Daniel, R. (2018b). Draft genome sequence of the thermophilic acetogen *Moorella humiferrea* DSM 23265. *Genome Announc.* 6:e357-18. doi: 10.1128/genomeA.00357-18
- Redl, S., Sukumara, S., Ploeger, T., Wu, L., Jensen, T. O., Nielsen, A. T., et al. (2017). Thermodynamics and economic feasibility of acetone production from syngas using the thermophilic production host *Moorella thermoacetica*. *Biotechnol. Biofuels* 10:150. doi: 10.1186/s13068-017-0827-8
- Richter, M., and Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proc. Natl. Acad. Sci. U.S.A.* 106, 19126–19131. doi: 10.1073/pnas.0906412106
- Rijssel, M., Veen, I., and Hansen, T. A. (1992). A lithotrophic *Clostridium* strain with extremely thermoresistant spores isolated from a pectin-limited continuous culture of *Clostridium thermosaccharolyticum* strain Haren. *FEMS Microbiol. Lett.* 91, 171–175. doi: 10.1016/0378-1097(92)90679-i
- Savage, M. D., and Drake, H. L. (1986). Adaptation of the acetogen *Clostridium thermoautotrophicum* to minimal medium. *J. Bacteriol.* 165, 315–318. doi: 10.1128/jb.165.1.315-318.1986
- Schaupp, A., and Ljungdahl, L. G. (1974). Purification and properties of acetate kinase from *Clostridium thermoaceticum*. *Arch. Microbiol.* 100, 121–129. doi: 10.1007/bf00446312
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Shieh, W. Y., and Jean, W. D. (1998). *Alterococcus agarolyticus*, gen.nov., sp.nov., a halophilic thermophilic bacterium capable of agar degradation. *Can. J. Microbiol.* 44, 637–645. doi: 10.1139/w98-051
- Slobodkin, A., Reysenbach, A. L., Mayer, F., and Wiegel, J. (1997). Isolation and characterization of the homoacetogenic thermophilic bacterium *Moorella glycerini* sp. nov. *Int. J. Syst. Bacteriol.* 47, 969–974. doi: 10.1099/00207713-47-4-969
- Stackebrandt, E., and Goebel, B. M. (1994). Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int. J. Syst. Evol. Microbiol.* 44, 846–849. doi: 10.1099/00207713-44-4-846
- Staden, R., Beal, K. F., and Bonfield, J. K. (1999). The Staden package, 1998. *Methods Mol. Biol.* 132, 115–130. doi: 10.1385/1-59259-192-2:115
- 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
- Tatusov, R. L., Koonin, E. V., and Lipman, D. J. (1997). A genomic perspective on protein families. *Science* 278, 631–637. doi: 10.1126/science.278.5338.631
- Tindall, B. J. (2016). What does rule 18c of the international code of nomenclature of bacteria really say? *Int. J. Syst. Evol. Microbiol.* 66, 3622–3624. doi: 10.1099/ijsem.0001245
- Tindall, B. J., Rosselló-Móra, R., Busse, H. J., Ludwig, W., and Kämpfer, P. (2010). Notes on the characterization of prokaryote strains for taxonomic purposes. *Int. J. Syst. Evol. Microbiol.* 60, 249–266. doi: 10.1099/ijse.0.016949-0
- Tsukahara, K., Kita, A., Nakashimada, Y., Hoshino, T., and Murakami, K. (2014). Genomeguided analysis of transformation efficiency and carbon dioxide assimilation by *Moorella thermoacetica* Y72. *Gene* 535, 150–155. doi: 10.1016/j.gene.2013.11.045
- Wick, R. R., Judd, L. M., Gorrie, C. L., and Holt, K. E. (2017). Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* 13:e1005595. doi: 10.1371/journal.pcbi.1005595
- Wiegel, J., Braun, M., and Gottschalk, G. (1981). *Clostridium thermoautotrophicum* species novum, a thermophile producing acetate from molecular hydrogen and carbon dioxide. *Curr. Microbiol.* 5, 255–260. doi: 10.1007/bf01571158
- 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: 10.1016/s0378-1097(98)00094-9

Conflict of Interest: BT was employed by company Leibniz-Institut DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Brunswick, Germany.

The remaining 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.

Copyright © 2020 Redl, Poehlein, Esser, Bengelsdorf, Jensen, Jendresen, Tindall, Daniel, Dürre and Nielsen. 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.