**A novel *Thermotoga* strain TFO isolated from a Californian petroleum reservoir
phylogenetically related to *Thermotoga petrophila* and *T. naphthophila*, two
thermophilic anaerobic isolates from a Japanese reservoir: taxonomic and genomic
considerations**

Zara M Summers^{1*}, Hassiba Belahbib², Nathalie Pradel², Manon Bartoli², Pooja Mishra¹,
Christian Tamburini², Alain Dolla², Bernard Ollivier², Fabrice Armougom^{2*}

1. ExxonMobil Research and Engineering Company, 1545 Route 22 East, Annandale
New Jersey 08801, United States

2. Aix Marseille Univ., Université de Toulon, CNRS, IRD, MIO UM 110, 13288,
Marseille, France

*** Corresponding authors**

Fabrice Armougom

Aix Marseille Univ., Université de Toulon, CNRS, IRD, MIO UM 110, 13288, Marseille,
France

Email: fabrice.armougom@mio.osuphytheas.fr

Zara M Summers

ExxonMobil Research and Engineering Company, 1545 Route 22 East, Annandale New
Jersey 08801, United States

Email: zarath.m.summers@exxonmobil.com

Abstract

Hot oil reservoirs harbor diverse microbial communities, with many of them inhabiting thermophilic or hyperthermophilic fermentative *Thermotoga* species. A new *Thermotoga* sp. strain TFO was isolated from an Californian offshore oil reservoir which is phylogenetically related to thermophilic species *T. petrophila* RKU-1^T and *T. naphthophila* RKU-10^T, isolated from the Kubiki oil reservoir in Japan. The average nucleotide identity and DNA-DNA hybridization measures provide evidence that the novel strain TFO is closely related to *T. naphthophila* RKU-10^T, *T. petrophila* RKU-1^T and can not be differentiated at the species level. In the light of these results, the reclassification of *T. naphthophila* RKU-10 and strain TFO as heterotypic synonyms of *T. petrophila* is proposed. A pangenomic survey of closely related species revealed 55 TFO strain-specific proteins, many of which being linked to glycosyltransferases and mobile genetic elements such as recombinases, transposases and prophage, which can contribute to genome evolution and plasticity, promoting bacterial diversification and adaptation to environmental changes. The discovery of a TFO-specific transport system dctPQM, encoding a tripartite ATP-independent periplasmic transporter (TRAP), has to be highlighted. The presence of this TRAP system assumes that it could assist in anaerobic n-alkane degradation by addition of fumarate dicarboxylic acid, suggesting a niche-specific gene pool which correlates with the oil reservoir that *T. petrophila* TFO inhabits. Finally, *T. naphthophila* RKU-10, *T. petrophila* RKU-1^T, *T. petrophila* TFO form a distinct phylogenetic lineage with different geographic origins, share the same type of ecological niche including the burial history of fields. Theses findings might support the indigenous character of this species in oil reservoirs.

Keywords: *Thermotoga*, Taxonomy, Classification, Genomics , Oil reservoirs, Anaerobes

Introduction

A wide range of anaerobic microorganisms with diverse physiological and metabolic abilities are known to inhabit oil reservoirs where they are subjected to drastic *in situ* physicochemical conditions (e.g. high temperature and/or high hydrostatic/lithostatic pressure) [45]. Most of these inhabitants are methanogenic archaea, sulfate-, nitrate- and iron-reducing prokaryotes, homoacetogens, as well as fermentative bacteria [45]. Among the latter, anaerobic fermentative members of the phylum *Thermotogae* are recognized to be widespread in both marine and terrestrial subsurface ecosystems [38,39,48,56].

The phylum *Thermotogae*, which forms an early branching lineage within the *Bacteria* domain, consists of 65 cultivated species disseminated among 13 genera [25]. The originality of the phylum *Thermotogae* stems from the physiological features of its bacterial representatives, which are mostly thermophilic ($T_{\text{opt}} \geq 70\text{ }^{\circ}\text{C}$) and hyper-thermophilic ($T_{\text{opt}} \geq 80^{\circ}\text{C}$) [14], with the exception of some mesophiles ($T_{\text{opt}} 37\text{-}40\text{ }^{\circ}\text{C}$) pertaining to the genus *Mesotoga* [17,43] and moderate thermo-acidophiles (optimal growth at pH 3-5, $T_{\text{opt}} \geq 55\text{ }^{\circ}\text{C}$) of the genera *Athalassotoga* and *Mesoaciditoga* [24,50]. Advances in species isolation combined with the use modern taxonomy practices (i.e genomic signatures, Overall Genome Relatedness Indices) continue to reshape the classification of the phylum *Thermotogae* [2,3]. The phylum is currently divided into four orders (*Thermotogales*, *Kosmotogales*, *Petrotogales*, *Mesoaciditogales*) and five families (*Thermotogaceae*, *Fervidobacteriaceae*, *Kosmotogaceae*, *Petrotogaceae*, *Mesoaciditogaceae*).

Only members of the families *Thermotogaceae*, *Kosmotogaceae*, and *Petrotogaceae* have thus far been observed by either culture-dependent or independent approaches from oilfield environments. Within the family *Thermotogaceae*, they include two representatives of the genus *Thermotoga*, *T. petrophila* RKU-1 [56] and *T. naphthophila* RKU-10 [56], which have been isolated from the Japanese Kubiki oil reservoir.

Herein, we report the phenotypic and genomic characterization of a new *Thermotoga* strain (TFO), isolated from an offshore Californian oil reservoir, which is phylogenetically and genomically related to *Thermotoga petrophila* RKU-1 and *Thermotoga naphthophila* RKU-10.

Methods

Isolation

Pescado Monterey reservoirs have temperature ranges of (85-104 °C). Pescado Monterey depth averages 2,225 m (300 m-700 m water depth, 1500 m overburden), resulting in an average pressure of 22.25 MPa (~3200 psi). The produced fluids consisting of liquid hydrocarbons and production water were collected from a production well into sterile and anaerobic 156 mL serum bottles, and were chilled in the dark with ice and transported to laboratory within 48 hours of collection. Enrichment of thermophiles was performed by inoculating the collected production fluids from the Pescado offshore oil reservoir. Within 7 days, turbidity due to cell growth was observed. Initially, the enrichments consisted of a mixed community, and the sheathed “toga” like hyperthermophilic bacterial strain was purified by inoculating gelrite solidified medium in the bottom of an anaerobic 156 mL serum bottle, and incubating them at 65 °C. After about 7 days of incubation, a few singles colonies (<1 mm in diameter) were observed in the solid medium and were picked with sterile glass Pasteur pipettes in an anaerobic chamber and transferred into liquid medium, and the isolation was repeated three times to ensure the purity of the isolates. The isolated strains were then further purified using dilution to extinction cultivation in a series of tubes containing liquid medium. The isolated strain we discuss in this manuscript is an obligatory anaerobic heterotrophic rod-shaped bacterium with a “toga”, a sheath like structure with ballooning at both ends, which is characteristic of members of the phylum *Thermotogae*. The isolated strain TFO have been deposited for conformity at the Leibniz Institute DSMZ-

German Collection of Microorganisms and Cell Cultures GmbH. The collection number assigned to strain TFO is DSM 111359.

Culture medium composition

Strain TFO was cultivated in medium containing (per liter): 10 g NaCl, 1 g KCl, 0.2 g CaCl₂·2H₂O, 0.3 g KH₂PO₄, 0.3 g K₂HPO₄, 1 g NH₄Cl, 3 g Na₂S₂O₃, 1 g yeast extract, cysteine-HCl 0.5 g, resazurin (0.1 %) 1 ml and trace minerals [1] 10 ml. The pH was adjusted to 7 with a 10 M KOH solution. The medium was boiled under a stream of O₂-free N₂ gas and cooled at room temperature. Aliquots of 5 ml were dispensed into Hungate tubes, stoppered with butyl rubber stoppers, and degassed under N₂-CO₂ (80/20, vol/vol) and subsequently sterilized by autoclaving at 120 °C for 20 min. Prior to culture inoculation, 0.2 ml of 10 % (w/vol) NaHCO₃, 0.1 ml of 150 g L⁻¹ MgCl₂·6H₂O, 0.1 ml of 2 % Na₂S and 20 mM of glucose from sterile stock solutions were injected into the tubes. The Hungate technique was used throughout the study [20]. A 0.5 ml aliquot of the sample was inoculated into the Hungate tubes that were subsequently incubated at 70 °C. Growth experiments were performed in duplicate.

Determination of the optimal temperature, NaCl and pH conditions for growth

The temperature, NaCl, and pH ranges for growth were determined using the above culture medium. To study the NaCl tolerance, the required quantity of NaCl was added to the tubes (from 0 g/L to 80 g/L). The pH of the medium was adjusted by injecting into the Hungate tubes aliquots of anaerobic stock solutions of 10% NaHCO₃ or 8% Na₂CO₃ (to test pH between 5 and 9.5) and checked after autoclaving. The strain was sub-cultured at least twice under the same experimental conditions before growth rates were determined.

Substrate utilization assays

The various substrates were added at 20 mM concentration to the culture medium without glucose, and in the presence or absence of sodium thiosulfate as potential electron acceptor to

be used by members of the phylum *Thermotogae* [10]. Complex organic substrates (casamino acids, yeast extract, tryptone, and peptone) were tested at a final concentration of 2 g/L, in the minimal culture medium described previously without yeast extract. To test electron acceptor utilization, beside sodium thiosulfate (20 mM), sodium sulfate (20 mM), elemental sulfur (10 g/L), nitrate (20 mM), or nitrite (2 mM) were added at 20 mM final concentration in the culture medium in the absence of thiosulfate. Growth was determined by OD_{600 nm} measurements and microscopy observations, after 24 h of culture incubation at 70 °C. Experiments were performed in duplicate. Substrate consumption and organic acid production were measured by HPLC (using the UltiMate 3000, Thermo Scientific). Thiosulfate consumption was measured by ion chromatography, using the Metrohm 761 Compact IC. H₂S production was measured using the Cord-Ruwisch method [7].

Lipid analysis

Cultures of the three strains were grown-up and harvested at the end of exponential phase and sent to DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany) for fatty acids and polar acids analyses.

Light and scanning electron microscopy

Cells of isolate TFO were analyzed by using the Scanning Electron Microscope (SEM) protocol as described by Murtey et al. [42] and modified as follows. Cells on glass slides were briefly rinsed three times in cacodylate buffer and transferred to freshly prepared alcian blue-lysine added in glutaraldehyde/paraformaldehyde fixative for 2 hours at 4 °C. The glass slides were then transferred to cacodylate buffered glutaraldehyde fixative for 24 hours. The glass slides were then dehydrated in a series of ethanol concentrations for 20 min in each of the following concentrations: 25%, 50%, 70%, 75%, 85%, 90%, 95% and two (20 min) exchanges in anhydrous 100% ethanol followed by chemical fixation in hexamethyldisilazine. Samples were then air dried for more than 3 hours with desiccant before visualizing under Zeiss Crossbeam 540 SEM.

Spectroscopic DNA–DNA hybridization

DNA-DNA Hybridization (DDH) studies were performed using isolate TFO (DSM 111359), *Thermotoga petrophila* RKU-1 (DSM 13995) and *Thermotoga naphthophila* RKU-10 (DSM 13996). The DDH studies were carried out by the DSMZ. Cells were disrupted using a Constant Systems TS 0.75 KW (IUL Instruments, Germany). DNA in the crude lysate was purified by chromatography on hydroxyapatite as described by Cashion and colleagues [5]. DNA-DNA hybridization was carried out as described by De Ley et al. [33] under consideration of the modifications described by Huss et al. [22] using a model Cary 100 Bio UV/VIS-spectrophotometer equipped with a Peltier-thermostatted 6x6 multicell changer and a temperature controller with in-situ temperature probe (Varian).

Genome sequencing

Genomic DNA of the isolate TFO was extracted and purified following manufacturer's instructions (Masterpure DNA isolation kit, Epicenter). Ten mL of cell culture was spun down at 10,000 g for extraction and purification. The Genomic DNA of isolate TFO was sequenced at Synthetic Genomics Inc. (La Jolla, CA, USA) on two SMRT cells on a PacBio RS II (Pacific Biosciences, Menlo Park, CA, USA). The genome was assembled from 373,255,583 post-filtered bases comprising 36,575 reads with an average read length of 10,205 bases.

Core genome-based phylogeny

A core genome phylogeny was inferred using the genomes of strain TFO and its closely related cultivated *Thermotoga* species, including *T. petrophila* RKU-1, *T. naphthophila* RKU-10, *T. neapolitana* NS-E and *T. maritima* MSB8. Uncultivated *Thermotogae* spp. were also included (complete and incomplete genomes). The genomes were retrieved from the NCBI ftp (<ftp://ftp.ncbi.nlm.nih.gov/>), and the core genome identification was performed by the Roary pipeline v3.10.2 [44]. The core gene alignment was generated by the Roary

pipeline with a 50% sequence identity cut-off. The phylogenetic tree was constructed from the 492 core-genes with the BioNJ distance-based phylogeny reconstruction algorithm, implemented in Seaview v4.7 [15]. The bootstrap branch supports were computed by the 'Transfer Bootstrap Expectation' method (TBE) with 1000 replicates [32].

DOE-JGI genome submission and annotation of *Thermotoga* sp. strain TFO

The genome sequence of *Thermotoga* sp. strain TFO was fully annotated using the DOE-JGI Microbial Genome Annotation Pipeline [21] (MGAP v4.16.2). The MGAP pipeline uses Prodigal V2.6.3 [23] for gene calling, the annotation algorithm is HMMER 3.1b2 [40] and Rfam 12.0 is the database [26]. The functional annotation of genes was performed using COG 2003 and Pfam v30 [12]. The annotated genome of *Thermotoga* sp. strain TFO (under the recommended species name *Thermotoga petrophila* TFO) is available at the DOE-JGI site <https://genome.jgi.doe.gov> under the GOLD the project ID Gp0307586.

Comparative genomics

The genomes of *T. petrophila* RKU-1, *T. naphthophila* RKU-10, *T. neapolitana* NS-E, *T. maritima* MSB8 and strain TFO were used for pan-genomic analysis. The gene prediction and annotation were performed using the Prokka package [53]. The clustering of the orthologs for pan-genomic analysis was carried out by Roary [44], applying a minimum of sequence identity and coverage of 50%, as recommended by Tettelin and colleagues [57]. The singleton proteins (strain-specific proteins) did not belong to any orthologous cluster. The specific COG (Cluster of Orthologous Gene) functional annotation was locally performed by rpsblast with the predicted proteins against COG profiles included in the Conserved Domain Database (CDD) at the NCBI (<ftp://ftp.ncbi.nlm.nih.gov/pub/mmdb/cdd>). The COG annotation was selected for E-value below 10^{-5} . Pfam information for TFO-specific proteins were retrieved (E-value $< 10^{-5}$) from local rpsblast using Pfam profiles of the CDD

database. The prophage prediction was investigated using PHAST

(<http://phast.wishartlab.com/>).

The Average Nucleotide Identity (ANI) values were calculated using JSpecies software [51] and as described in [2]. Venn diagram was designed with the VennDiagram R package version 1.7 and the “draw.quintuple.diagram” function. The Figure 6 was drawn by Circos version 0.69 [31] and using the results of pairwise ortholog detection. The contigs of strain TFO were reorder following the results provided by “move contig” function of MAUVE [9] and using *T. petrophila* RKU-1 as reference genome.

Results and discussion

Morphological and phenotypic properties of isolate TFO

Strain TFO (DSM 111359) is a heterotrophic strict anaerobe and a Gram-negative rod-shaped bacterium. The cell morphology of TFO isolate was 3 µm long and 0.5 µm wide, exhibiting the characteristic sheath-like structure (called “toga”) of *Thermotogae* representatives (**Figure 1**). The growth range temperature varied from 55 °C to 85 °C with an optimum at 75 °C. No growth was observed below 50 °C or above 90 °C (**Figure 2, A**). The isolate grew in the absence of NaCl and tolerated it up to 60 g/L (**Figure 2, B**). This differs from both *T. petrophila* RKU-1 and *T. naphthophila* RKU-10 as they require NaCl for growth [56] and should therefore be considered as slight halophiles. In contrast, strain TFO should be considered as halotolerant [56]. The optimum pH for growth was 7 and no growth was observed at below pH 5.5 or or above pH 9 (**Figure 2, C**) similarly to other *Thermotoga* representatives (**Table 1**).

Strain TFO was unable to grow autotrophically under H₂/CO₂ (80%/20%) atmosphere with or without acetate (2 mM) as carbon source. The strain could grow heterotrophically on the following sugars: glucose, sucrose, lactose, maltose, fructose, galactose, xylose, mannose, ribose, arabinose, and raffinose with the exception of sorbose and cellulose. Yeast extract

(1g/L) was required for growth. No growth was observed on yeast extract, tryptone, peptone, casamino acids or organic acids including propionate, butyrate, lactate, and acetate when provided as sole energy source (**Table 1**). Contrary to its two closest phylogenetically relatives *T. naphthophila* RKU-10 and *T. petrophila* RKU-1 (**Figure 3**) [56], strain TFO was able to grow on xylose (**Table 1**). However, to be confident with this peculiar phenotypic feature of strain TFO, the growth of *T. naphthophila* RKU-10 and *T. petrophila* RKU-1 on xylose was tested as part of this study and we observed that both strains were also able to grow on this substrate. Additionally, a survey of the gene content of all three strains revealed the presence of shared xylose isomerase, the key enzyme required in the xylose catabolism pathway [54]. Acetate, CO₂ and H₂ were produced by strain TFO when grown on sugars. It was also observed that higher amounts of H₂ was produced when the cells were grown in the absence of thiosulfate compared to growth in the presence of thiosulfate, suggesting that thiosulfate can be used as electron acceptor as confirmed by sulfide production in the presence of this sulfur compound in the culture medium. This is a common physiological trait shared by several *Thermotoga* species [11,48]. The produced acetate/consumed glucose ratio was similar in the presence or absence of thiosulfate. Beside thiosulfate, in the presence of 20 mM glucose, strain TFO was also able to reduce elemental sulfur to sulfide. Addition of either thiosulfate or elemental sulfur in the culture medium enhanced the growth yields as compared to control in the absence of them. On the contrary, the presence of these sulfur compounds in the culture medium of *T. naphthophila* RKU-10 and *T. petrophila* RKU-1 decreases final growth yields and growth rates, even if both strains are recognized as thiosulfate and elemental sulfur reducers [56]. When cultivated under elemental sulfur, strain TFO produced higher amounts of H₂S (4-times more) than when grown under comparable thiosulfate reduction conditions. Sulfate was not used as electron acceptor by strain TFO, and the addition of either nitrate (20 m) or nitrite (2 mM) completely inhibited growth.

Polar lipids

Glycolipids, phospholipids and phosphoglycolipids were identified as polar lipids present for the three strains (data not shown). The profiles of fatty acids were similar (some variations observed for C_{10:0}/C_{14:0}) for the three strains cultivated under same growth conditions (**Table 2**). The major fatty acids were C_{16:0} (> 70%), C_{14:0} (6.98 to 17.7%), C_{10:0} (3.34 to 13.66%), and C_{18:0} (2.0 to 2.84%) for the three strains, as for other *Thermotoga* species [41].

Phylogenetic placement of strain TFO

The 16S rRNA genes do not provide a confident taxonomic identification since the nucleotide sequence of the gene from the novel isolate TFO and those of representatives within the genus *Thermotoga* shared more than 99% of sequence identity. Since the phylogenetic inference relies on 16S rRNA genes, we cannot therefore resolve the species relatedness within the genus *Thermotoga* as previously shown by Belahbib and colleagues [2]. The relationships of strain TFO with the related *Thermotoga* species were thereby explored by a core-genome based phylogeny (utilizing 492 shared genes) that resulted in a well-supported clade (**Figure 3**). Strain TFO forms a lineage with *T. naphthophila* RKU-10 and *T. petrophila* RKU-1 and the uncultivated *T. naphthophila* 46_26.

Species delineation by ANI and DDH measures

In addition to the phylogenetic placement of strain TFO within the genus *Thermotoga*, species delineation was determined by Overall Genome Relatedness Index (OGRI) [6], particularly by the recommended ANI measure [30]. The delineation of inter-species to intra-species relationships is signified by the ANI interval of 95-96% [29], and the ANI measures of the strain TFO with *T. naphthophila* RKU-10 and *T. petrophila* RKU-1 reached the boundary of intra-species relationship with ANI values of 95.80 % and 96.08%, respectively. Strikingly, the ANI value between the species *T. naphthophila* RKU-10 and *T. petrophila* RKU-1 was slightly over the boundary of intra-species relationship (ANI = 96.38%). Overall,

the ANI clustering highlighted species delineation ambiguities within the genus *Thermotoga* (**Figure 4**), exemplified by intra-species relationships for the uncultivated *Thermotoga* sp. Xyl 54, *Thermotoga* sp. Cell2, *Thermotoga* sp. TBGT1765 and *Thermotoga* sp. TBGT1766 (ANI values > 98%).

Considering the ambiguous ANI results for species delineation, complementary DDH experiments were carried out in duplicate with *T. petrophila* RKU-1, *T. naphthophila* RKU-10 and the strain TFO. The results indicated DNA–DNA homology of 83.15% (81.3; 85.0) between the strain TFO and *T. petrophila* RKU-1, and DNA-DNA homology of 74.7% (75.3; 74.1) with *T. naphthophila* RKU-10. Likewise, DNA reassociation between *T. petrophila* RKU-1 and *T. naphthophila* RKU-10 was 77.65% (75.9; 79.4), which is in disagreement with the average value (15%) reported in the original work of Takahata and colleagues [56]. All DDH values of pairwise species comparisons were above 70%, providing evidence that *T. naphthophila* RKU-10, *T. petrophila* RKU-1 and strain TFO could not be differentiated at the species level and, therefore, should be considered as a single species. Based upon ANI and DDH results, *T. petrophila* RKU-1, *T. naphthophila* RKU-10 and *Thermotoga* sp. TFO should be considered as heterotypic synonyms. According to rules 38, 42 and 24b of the bacteriological code [46] the name *Thermotoga petrophila* RKU-1 has priority and should be used for the unified taxon. For that reason, *Thermotoga petrophila* TFO is proposed as species name for the TFO isolate.

Biogeography and geology

Strain TFO forms a distinct lineage with *T. naphthophila* RKU-10 and *T. petrophila* RKU-1 and the uncultivated *T. naphthophila* 46_26 (**Figure 3**). While strain TFO originates from an offshore Californian oil reservoir (Pescado Field), the strains RKU-10, RKU-1 and the uncultivated strain 26_46 originate from Japan (Kubiki Field) and Alaska oil reservoirs, respectively. The global presence of particular bacterial lineage in geographically distant oil

reservoirs could be an indicator of their indigenous nature [35]. Additionally, the Late Miocene Monterey Formation, a siliceous mudstone, is both source and reservoir for the Pescado Field oil [8] (TFO isolate). Reservoir depths are at ~2225 m with temperatures ranging from ~85 to 104 °C. Low gravity, high sulfur crude oil was generated locally. The reservoirs of the Kubiki Field oils are primarily in the Late Miocene Teradomari Formation [27] at shallower depths (of ~1100 m) and lower temperatures (~50-58 °C) [56]. The source of the lighter, low sulfur crude oil in the Kubiki Field (*T. petrophila* and *T. naphthophila*) is not certain, but is also believed to be of Miocene age [58]. It suggests that the burial history of both fields is similar in that the source strata were rapidly buried at similar geologic times and that the reservoir rocks are currently at their maximum temperatures.

Genomic features of *T. petrophila* TFO

The genome properties of strain TFO and its closely related cultivated species are summarized in **Table 3**. The genome was assembled in eight scaffolds with a total count of 1,885,773 base pairs (bp), slightly larger than the complete *Thermotoga* genomes (*T. neapolitana* NS-E, *T. maritima* MSB8, *T. petrophila* RKU-1, *T. naphthophila* RKU-10). The GC content of the *Thermotoga* representatives ranges 46-47%, which is in agreement with the 46.09% of strain TFO. The genome contained a total of 1966 predicted genes consisting of 52 RNA genes and 1914 proteins. The functional annotation indicated that 76.80% of the proteins were assigned to functional COG categories, with the major designations being general function (8.88%), carbohydrates (8.78%) and amino acid transport and metabolism (8.10%).

Specificities of *T. petrophila* TFO

The pan-genome of cultivated *Thermotoga* species included a core-genome (essential genes for shared across all species) of 1583 orthologous proteins (**Figure 5**). The *Thermotoga* core-genome represents a large fraction of all predicted proteins of each species including 83%,

320 87%, 88%, 83% and 86% for strain TFO, *T. naphthophila* RKU-10, *T. petrophila* RKU-1, *T.*
321 *neapolitana* NS-E and *T. maritima* MSB8, respectively. Considering only the phylogenetic
322 closest neighbors, the orfome of *T. petrophila* and *T. naphthophila* shared 94% and 92% of
323 orthologs with that of TFO strain, respectively. These orthologs were highly conserved at the
324 amino acid identity level since almost all exhibited a score above 90% (**Figure 6**). These
325 three closely related species also displayed conservation of genome synteny and high genome
326 compactness (**Figure 6**). Finally, the number of species/strain-specific proteins (non-essential
327 genes) for strain TFO, *T. petrophila* RKU-1, *T. naphthophila* RKU-10, *T. neapolitana* NS-E
328 and *T. maritima* MSB8 was 55, 37, 56, 131, and 54, respectively. Most of species/strain-
329 specific genes are often acquired through horizontal gene transfer and are assumed to encode
330 adaptive function [36].

331 In this study, among the 55 strain-specific proteins of *T. petrophila*. TFO (**Table 4, Figure**
332 **6**), eight were glycosyltransferases (mainly in the same cluster of genes), which are known to
333 be involved in the biosynthesis of di- oligo- and polysaccharides, protein glycosylation [52],
334 as well as cell wall peptidoglycan synthesis [37]. Likewise, this gene cluster harbors an
335 oligosaccharide-flippase transporter that can promote translocation of lipid-linked
336 oligosaccharides for polysaccharide synthesis and cell wall construction [18].

337 Mobile genetic elements including insertion elements, transposases or bacteriophages are of
338 particular importance for genome diversification. In this way, an incomplete 38-kb prophage
339 (44.37% GC content) was predicted in the TFO genome and localized in the region 1.357.190
340 bp (attL site) to 1.395.092 bp (attR site). Although this prophage region was not detected in
341 the other genomes, only 14 (of the 23 predicted proteins in this region) were TFO-specific
342 proteins, of which nine were prophage-related and five were bacteria-related proteins.
343 Moreover, two transposases of the ISCc3 family (integrase and OrfB) and two
344 recombinases/resolvases were identified solely in *T. petrophila* TFO. The transposases were

345 localized next to homologous proteins to DEAD/DEAH helicases which are important
346 players of processes involving RNA molecules including ribosome biogenesis, transcription,
347 RNA turnover and RNA-protein folding [49].

348 In addition, among the 55 strain-specific proteins of *T. petrophila* TFO, a cluster of eleven
349 proteins (**Table 4**, from NJCICCOL_00759 to NJCICCOL_00770) were mostly associated to
350 bacterial transport systems, including an ABC type transporter of sugar (substrate specificity
351 unknown), and the dctPQM system which encodes a tripartite ATP-independent periplasmic
352 transporter (TRAP) [13]. The TRAP proteins of the strain TFO were found to be
353 phylogenetically related (not shown) to the uncultivated *Thermotoga naphthophila* 46_26
354 (Biosample: SAMN03445142) which originates from the metagenome of the Kuparuk oil
355 field in the Alaskan Arctic North Slope [19]. TRAP systems can transport a variety of
356 substrates under different contexts, including the transport of solutes that could play an
357 important role in the piezophilic lifestyle [47], and particularly in the import of C4-
358 dicarboxylic including succinate, fumarate, butyrate [13]. These organic acids can be the
359 result of the transformation of alkanes, as recently shown by members of candidate phylum
360 *Atribacteria* that ferment short-chain n-alkanes into carboxylic acids [34]. Furthermore, the
361 extracellular uptake of C4-dicarboxylic fumarate could play a key role in the activation of the
362 anaerobic degradation of alkanes, naphthalene and toluene from oil-rich environments [4,28].
363 The homologues of enzymes for the fumarate addition mechanism in anaerobic degradation
364 of alkanes [4] were not unambiguously identified in the genome of strain TFO. Indeed, the
365 key enzyme alkylsuccinate synthase AssA, MassD and pflD genes [28,55] shared at best 32%
366 of sequence identity with a predicted hypothetical protein from the genome of strain TFO,
367 which is an ambiguous score for homolog definition [57]. One has to keep in mind that genes
368 could be missing due to the incompleteness of the genome of *T. petrophila* TFO (eight

scaffolds). As a consequence, the potential alkylsuccinate synthase activity as well as TRAP functional specificity/activity of strain TFO should be experimentally evaluated.

Conclusions

Herein, a novel *Thermotoga* strain (TFO) isolated from a Californian petroleum reservoir was phenotypically and genomically characterized. In the light of modern taxonomy practice with the use of core genome phylogeny, whole genome relatedness indices and DDH measures, our results recommend considering *T. petrophila* RKU-1, *T. naphthophila* RKU-10 and *Thermotoga* TFO as heterotypic synonyms. Based upon Takahata and colleagues' work [56], the name *Thermotoga petrophila* has priority and should be used for the unified taxon. For that reason, *Thermotoga petrophila* TFO is proposed as species name for this oil reservoir isolate.

These three strains accompanied by the uncultivated *Thermotoga naphthophila* 26_46 belong to the same phylogenetic lineage and share the same type of ecological niche with similar burial history (oil field of cultivated species) from distinct geographic regions (Japan, California and Alaska). These considerations could be in favor of the indigenous character of such species in oil reservoirs [35]. The conservation of the genome synteny (**Figure 6**) of the three oilfield strains and the very high level of homologous proteins suggest that they are slow evolving organisms. Nonetheless, the TFO-specific proteins, including mobile genetic elements, could promote bacterial diversification or specialization. These genes can have clear relationships to the local environment and tagged as "niche-specific gene pool" that could encode for degradation of new type of nutrient source [16]. The hypothetical proteins, orphans (protein with no hits), sugar and TRAP transporters that belong to the TFO-specific proteins could thereby bring new metabolic functions associated to its niche which requires further study.

Figure legend

Figure 1: Scanning electron microscopy of the isolate TFO

Morphological features of *T. petrophila* TFO revealed by electron microscopy. Arrows indicate the toga (tg) and cell (c). Bar, 1 μ m.

Figure 2: Effects of temperature (A), NaCl concentration (B) and pH (C) on the growth rates of *T. petrophila* TFO

Strain TFO was grown as described in the Methods section and the growth rates were calculated from a linear regression along the logarithmic part of the growth curves.

Figure 3: Core genome-based phylogeny of the *T. petrophila* TFO within the *Thermotogaceae* family

Phylogenetic tree of 17 *Thermotoga* species deduced from 492 core-genes with the BioNJ distance-based phylogeny reconstruction algorithm. The bootstrap branch supports were computed by the 'Transfer Bootstrap Expectation' method (TBE) with 1000 replicates. The tree is rooted with *Pseudothermotoga elfii* and *P. lettingae*. Species names in bold are cultivated type species. Asterisk (*) means species isolated from oil reservoir.

Figure 4: Whole-genome relatedness within the *Thermotogaceae* family

Hierarchical clustering from Average Nucleotide Identity measures using 22 *Thermotogaceae* genomes including the TFO isolate. Value scale is depicted from red to blue colors for ANI genome pairwise comparison. The value scale ranges from 0 to 100% according to sequence identity level. The vertical red line at the value 95-96 indicates the ANI species demarcation threshold.

Figure 5: Pangenome of *Thermotoga* cultivated species

Five-set Venn diagram of the pangenome of five cultivated *Thermotoga* species based on the prediction of orthologous proteins. Values on diagram represent the numbers of orthologs shared between the species or unique proteins for one species. The core-genome (black bold

value, center), variable-genome (black values, small size) and strain-specific (colored values) sizes are represented.

Figure 6. Ortholog relationships and strainTFO-specific proteins

Visualization of the ortholog relationships between strain TFO and *T. petrophila* RKU-1 and *T. naphthophila* RKU-10. From outside to inside. Size labels is 10^4 base pairs. Genomes are labelled and correspond to circle segments. The ORF positions for strain TFO, *T. naphthophila*, *T. petrophila* are colored light blue, deep red and light green inside the circle segments, respectively. The ORF distribution shows the compactness of genomes since spaces between ORFs are sparsely present (white color). TFO-specific ORFs are colored in black inside the circle segment, and mainly organized in short clusters. Red links, orthologs between strain TFO and *T. naphthophila* sharing more than 90% of sequence identity. Blue links, orthologs between strain TFO and *T. petrophila* sharing more than 90% of sequence identity. Yellow links are orthologs sharing less than 90% of sequence identity. The link directions (uncrossed) highlight the conservation of genome synteny between the species.

Table 1. Characteristics of strain TFO with its closely related species within the genus *Thermotoga*

Characteristic	Strain TFO	<i>Thermotoga naphthophila</i> RKU-10 (Takahata et al, 2001)	<i>Thermotoga petrophila</i> RKU-1 (Takahata et al, 2001)	<i>Thermotoga neapolitana</i> DMS4359 (Jannasch et al, 1988)	<i>Thermotoga maritima</i> MSB8 (Huber et al, 1986)
Origin	Californian produced water	Oil reservoir, Japan	Oil reservoir, Japan	Geothermal hot spring, Italy	Geothermal hot spring, Italy
Temperature for growth (°C)					
Range	55-85	48-86	47-88	55-90	55-90
Optimum	75	80	80	80	80
pH for growth					
Range	6.0-8.5	5.4-9.0	5.2-9.0	5.5-9.0	5.5-9
Optimum	7.0	7.0	7.0	7.0	6.5
Salinity (%)					
Range	0-6	0.1-6	0.1-5.5	0.25-6	0.25-3.75
Optimum	0	1	1	2	2.7
Optimal doubling time (min)	45 min	59 min	54 min	45 min	75 min
Oxygen tolerance (%)	<0	<0	<0	<15	<0
Electron acceptors	Thiosulfate, Elemental sulfur	Thiosulfate Elemental sulfur	Thiosulfate Elemental sulfur	Thiosulfate, Elemental sulfur (Ravot et al, 1995)	Thiosulfate, Elemental sulfur (Ravot et al, 1995)
Utilization of					
Glucose	+	+	+	+	+
Sucrose	+	+	+	+	+
Galactose	+	+	+	+	+
Lactose	+	+	+	+	+
Fructose	+	+	+	+	+
Arabinose	+	+	+	ND	ND
Sorbose	-	ND	ND	ND	ND
Raffinose	+	ND	ND	ND	ND
Mannose	+	ND	ND	ND	ND
Maltose	+	+	+	+	+
Ribose	+	+	+	+	+
Cellulose	-	-	+	ND	ND
Xylose	+	-	-	+	+
Casamino acids	-	-	-	-	-
Tryptone	-	ND	ND	+	+
Peptone	-	+	+	ND	ND
Yeast Extract	-	-	-	ND	+
Glycerol	+	+	+	-	+
Mannitol	-	+	-	-	-
Acetate	-	-	-	-	-
Lactate	-	-	-	-	-
Propionate	-	-	-	-	-
H ₂ CO ₂	-	-	-	-	-
G+C DNA (mol %)	46.0	46.1	46.6	41	46

+, growth; - no growth; ND: not determined.

Table 2. Cellular fatty acid contents of strains TFO, *T. petrophila* RKU-1, and *T. naphthophila* RKU-10.

Fatty acid	Strain TFO	<i>T. petrophila</i> RKU-1	<i>T. naphthophila</i> RKU-10
C _{10:0}	7.30	13.66	3.34
C _{12:0}	0.77	0.48	0.53
C _{14:0}	15.29	6.98	17.70
C _{15:0}	0.38	-	0.16
C _{16:0}	73.40	75.67	74.58
C _{16:0} N alcohol	-	-	0.42
C _{16:0} 3OH	0.46	0.37	0.43
C _{17:0}	0.40	-	-
C _{18:0}	2.00	2.54	2.84
C _{18:1} w9c	-	0.29	-

Values correspond to percentage (%).

Table 3. Genome features of *T. petrophila* TFO and closely cultivated species

Attribute	Isolate TFO	<i>T. naphthophila</i>	<i>T. petrophila</i>	<i>T. neapolitana</i>	<i>T. maritima</i>
Genome size (bp)	1,885,773	1,809,823	1,823,511	1,884,562	1,869,644
DNA coding (bp)	1,791,612	1,731,557	1,744,574	1809617	1,790,184
GC%	46.09	46.14	46.10	46.94	46.25
Total genes	1966	1867	1865	1988	1909
Protein coding	1914	1815	1810	1937	1858
RNA genes	52	52	55	51	51
Pseudo genes	–	47	25	–	–
Genes in internal clusters	340	117	124	126	144
Genes with function prediction	1604	1447	1456	1580	1553
Genes assigned to COGs	1451	1420	1426	1439	1504
Genes with Pfam domains	1674	1370	1605	1662	1657
Genes with signal peptides	36	29	36	27	37
Transmembrane helices	460	442	432	497	461
CRISPR repeats	6	8	7	7	6

Table 4. The TFO-specific proteins

Gene name	Prokka Annotation	SI %	cover %	CDD complementary annotation	PHAST annotation
NJCICCOL_00401	Hypothetical protein	100	100	EpsG superfamily	
NJCICCOL_00402	Glycosyl transferase	100	100	GT4_MtfB-like	
NJCICCOL_00403	Glycosyltransferase	100	100	Glyco_trans_1_2	
NJCICCOL_00404	Glycosyltransferase	100	98	Glyco_transf_4	
NJCICCOL_00405	Glycosyltransferase	99	97	stp1 superfamily	
NJCICCOL_00407	Glycosyltransferase	53	100	Glycosyltransferase family A	
NJCICCOL_00408	4Fe-4S dicluster domain-containing protein	63	48	Glycosyltransferase_GTB-type	
NJCICCOL_00409	Oligosaccharide flippase	76	100	MATE_like superfamily	
NJCICCOL_00410	DNA recombinase RMUC	92	100		
NJCICCOL_00411	HP	71	88		
NJCICCOL_00425	Glycosyltransferase family 2	97	100	Glyco_transf_GTA_type	
NJCICCOL_00426	HP	99	100	Glycosyltransferase_GTB-type	
NJCICCOL_00427	Oligosaccharide flippase	99	100	MATE_like superfamily	
NJCICCOL_00593	Integrase	100	100	HTH superfamily regulator	
NJCICCOL_00594	Transposase repeat family ISCc3	100	100	putative transposase OrfB	
NJCICCOL_00759	ABC transporter permease	86	99	AraH	
NJCICCOL_00760	Sugar ABC transporter ATPase	86	99	MglA	
NJCICCOL_0076	ABC transporter substrate binding	75	100	PBP1_ABC_sugar_binding_like_2	
NJCICCOL_00762	Hypothetical protein	60	100	Ureidogly_lyase superfamily	
NJCICCOL_00764	TRAP transporter large permease	93	100	DctQ	
NJCICCOL_00765*	DctQ component; TRAP transporter small permease	100	100		
NJCICCOL_00766*	TRAP transporter solute receptor, DctP family	100	100	Periplasmic_Binding_Protein_Type_2 superfamily	
NJCICCOL_00767*	TRAP transporter solute receptor, DctP family	100	100		
NJCICCOL_00768	4-hydroxy-tetrahydrodipicolinate synthase	79	100	DRE_TIM_metallolyase superfamily	
NJCICCOL_00769	4-hydroxy-tetrahydrodipicolinate synthase	76	100	DapA superfamily	
NJCICCOL_00770	IclR transcriptional regulator	78	98	DNA-binding transcriptional regulator, IclR family	
NJCICCOL_00839	Transposase repeat family ISCc3	100	100	Putative transposase OrfB	
NJCICCOL_00840	Integrase	100	100	transcription regulator	
NJCICCOL_00943	No significant hit	<50	100		
NJCICCOL_01065	Beta-glucosidase B	100	84	beta-D-glucoside glucohydrolase	
NJCICCOL_01406	Recombinase family	50	92	Resolvase Recombinase	Phage

NJCICCOL_01407	DEAD/DEAH box helicase	63	99	Predicted helicase	Phage Helicase
NJCICCOL_01408	HP	58	87	Nuclease homologue	DNAse/Rnase endonuclease Phage
NJCICCOL_01409	No hit				
NJCICCOL_01410	Putative HTH-type transcriptional regulator	65	77	HTH_XRE	Phage transcriptional regulator
NJCICCOL_01411	No significant hit			HTH_XRE	Phage repressor
NJCICCOL_01412	No hit				
NJCICCOL_01413	No significant hit	<50	<40		
NJCICCOL_01414	Hypothetical protein	77	100	MethyltransfD12 superfamily	
NJCICCOL_01415	No significant hit	<50	<50		
NJCICCOL_01416	Phage recombinase RecT	80	99		Phage RecT-like
NJCICCOL_01417	HP	76	100		Phage replication
NJCICCOL_01418	HP	87	100	YdjM,SOS response, DNA damage	Phage hydrolase
NJCICCOL_01419	HP	92	100		
NJCICCOL_01420	HP	80	100	Outer membrane efflux protein	
NJCICCOL_01421	HP	80	100	YjgR/ P-loop_NTPase superfamily	
NJCICCOL_01422	Carbopeptidase regulatory-like domain	80	100	CarboxypepD_reg	
NJCICCOL_01423	no hit				
NJCICCOL_01424	Hypothetical protein	74	95	—	
NJCICCOL_01493					
NJCICCOL_01949	DEAD/DEAH box helicase	63	99	Predicted helicase	
NJCICCOL_01950	recombinase	49	92	Resolvase/recombinase	
NJCICCOL_01951	HP	79	84	—	
NJCICCOL_01952	No hit				
NJCICCOL_01953	carbopeptidase regulatory-like domain	86	97	—	

HP means hypothetical protein. No hit means orphan. No significant hit means that best blast hit is below the threshold 50% of sequence identity. SI corresponds to sequence identity in %. Cover is query sequence cover in %. Asterisk (*) means protein phylogenetically related to *T. naphthophila* 26_46.

REFERENCES

- [1] Balch, W.E., Fox, G.E., Magrum, L.J., Woese, C.R., Wolfe, R.S. (1979) Methanogens: reevaluation of a unique biological group. *Microbiol. Rev.* 43(2), 260–96.
- [2] Belahbib, H., Summers, Z.M., Fardeau, M.L., Joseph, M., Tamburini, C., Dolla, A., Ollivier, B., Armougom, F. (2018) Towards a congruent reclassification and nomenclature of the thermophilic species of the genus *Pseudothermotoga* within the order *Thermotogales*. *Syst. Appl. Microbiol.* 41(6), 555–63.
- [3] Bhandari, V., Gupta, R.S. (2014) Molecular signatures for the phylum (class) *Thermotogae* and a proposal for its division into three orders (*Thermotogales*, *Kosmotogales* ord. nov. and *Petrotogales* ord. nov.) containing four families (*Thermotogaceae*, *Fervidobacteriaceae* fam. nov., *Kosmotoga* fam. nov. and *Petrotogaceae* fam. nov.) and new genus *Pseudothermotoga* gen. nov. with five new combinations. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 105(1), 143–68.
- [4] Bian, X.Y., Mbadinga, S.M., Liu, Y.F., Yang, S.Z., Liu, J.F., Ye, R.Q., Gu, J.D., Mu, B.Z. (2015) Insights into the anaerobic biodegradation pathway of n-Alkanes in oil reservoirs by detection of signature metabolites. *Sci. Rep.* (5), 981.
- [5] Cashion, P., Holder-Franklin, M.A., McCully, J., Franklin, M. (1977) A rapid method for the base ratio determination of bacterial DNA. *Anal. Biochem.* 81(2), 461–6.
- [6] Chun, J., Rainey, F.A. (2014) Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *Int. J. Syst. Evol. Microbiol.* 64(PART 2), 316–24.
- [7] Cord-Ruwisch, R. (1985) A quick method for the determination of dissolved and precipitated sulfides in cultures of sulfate-reducing bacteria. *J. Microbiol. Methods* 4(1), 33–6.
- [8] D. F. Lockman, J.R.S. (1996) The Monterey Formation of the Santa Ynez Unit, Part II: Fractures, Borehole Images, and Production: AAPG Bulletin, Convention and

Exhibition, San Diego, California, p. web.

- [9] Darling, A.C.E., Mau, B., Blattner, F.R., Perna, N.T. (2004) Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14(7), 1394–403.
- [10] Fadhlaoui, K., Ben Hania, W., Armougom, F., Bartoli, M., Fardeau, M.L., Erauso, G., Brasseur, G., Aubert, C., Hamdi, M., Brochier-Armanet, C., Dolla, A., Ollivier, B. (2018) Obligate sugar oxidation in *Mesotoga* spp., phylum *Thermotogae*, in the presence of either elemental sulfur or hydrogenotrophic sulfate-reducers as electron acceptor. *Environ. Microbiol.* 20(1), 281–92.
- [11] Fardeau, M.-L., Ollivier, B., Patel, B.K.C., Magot, M., Thomas, P., Rimbault, A., Rocchiccioli, F., Garcia, J.-L. (1997) *Thermotoga hypogea* sp. nov., a Xylanolytic, Thermophilic Bacterium from an Oil-Producing Well. *Int. J. Syst. Bacteriol.* 47(4), 1013–9.
- [12] Finn, R.D., Coghill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A. (2016) The Pfam protein families database: Towards a more sustainable future. *Nucleic Acids Res.* 44(D1), D279–85.
- [13] Forward, J.A., Behrendt, M.C., Wyborn, N.R., Cross, R., Kelly, D.J. (1997) TRAP transporters: A new family of periplasmic solute transport systems encoded by the *dctPQM* genes of *Rhodobacter capsulatus* and by homologs in diverse gram-negative bacteria. *J. Bacteriol.* 179(17), 5482–5493.
- [14] Frock, A.D., Notey, J.S., Kelly, R.M. (2010) The genus *Thermotoga*: recent developments. *Environ. Technol.* 31(10), 1169–81.
- [15] Gouy, M., Guindon, S., Gascuel, O. (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol.*

- Biol. Evol. 27(2), 221–4.
- [16] Hall, J.P.J., Brockhurst, M.A., Harrison, E. (2017) Sampling the mobile gene pool: Innovation via horizontal gene transfer in bacteria. *Philos. Trans. R. Soc. B Biol. Sci.* 372(1735).
- [17] Ben Hania, W., Postec, A., Aüllo, T., Ranchou-Peyruse, A., Erauso, G., Brochier-Armanet, C., Hamdi, M., Ollivier, B., Saint-Laurent, S., Magot, M., Fardeau, M.L. (2013) *Mesotoga infera* sp. nov., a mesophilic member of the order *Thermotogales*, isolated from an underground gas storage aquifer. *Int. J. Syst. Evol. Microbiol.* 63(PART8), 3003–8.
- [18] Hong, Y., Liu, M.A., Reeves, P.R. (2018) Progress in our understanding of Wzx flippase for translocation of bacterial membrane lipid-linked oligosaccharide. *J. Bacteriol.* 200(1).
- [19] Hu, P., Tom, L., Singh, A., Thomas, B.C., Baker, B.J., Piceno, Y.M., Andersen, G.L., Banfield, J.F. (2016) Genome-resolved metagenomic analysis reveals roles for candidate phyla and other microbial community members in biogeochemical transformations in oil reservoirs. *MBio* 7(1).
- [20] Hungate, R.E. (1969) Chapter IV A Roll Tube Method for Cultivation of Strict Anaerobes. *Methods Microbiol.* 3(PART B), 117–32.
- [21] Huntemann, M., Ivanova, N.N., Mavromatis, K., Tripp, H.J., Paez-Espino, D., Tennessen, K., Palaniappan, K., Szeto, E., Pillay, M., Chen, I.M.A., Pati, A., Nielsen, T., Markowitz, V.M., Kyrpides, N.C. (2016) The standard operating procedure of the DOE-JGI Metagenome Annotation Pipeline (MAP v.4). *Stand. Genomic Sci.* 11(1), 1–6.
- [22] Huss, V.A.R., Festl, H., Schleifer, K.H. (1983) Studies on the Spectrophotometric Determination of DNA Hybridization from Renaturation Rates. *Syst. Appl. Microbiol.*

4(2), 184–92.

- [23] Hyatt, D., Chen, G.L., LoCascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J. (2010) Prodigal: Prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11, 119.
- [24] Itoh, T., Onishi, M., Kato, S., Iino, T., Sakamoto, M., Kudo, T., Takashina, T., Ohkuma, M. (2016) *Athalassotoga saccharophila* gen. nov., sp. nov., isolated from an acidic terrestrial hot spring, and proposal of *Mesoaciditogales* ord. nov. and *Mesoaciditogaceae* fam. nov. in the phylum *Thermotogae*. Int. J. Syst. Evol. Microbiol. 66(2), 1045–51.
- [25] Jiang, L., L'Haridon, S., Jebbar, M., Xu, H., Alain, K., Shao, Z. (2017) Complete genome sequence and whole-genome phylogeny of *Kosmotoga pacifica* type strain SLHLJ1T from an East Pacific hydrothermal sediment. Stand. Genomic Sci. 12(1), 3.
- [26] Kalvari, I., Nawrocki, E.P., Argasinska, J., Quinones-Olvera, N., Finn, R.D., Bateman, A., Petrov, A.I. (2018) Non-Coding RNA Analysis Using the Rfam Database. Curr. Protoc. Bioinforma. 62(1).
- [27] Kawai, K., Totani, S. (1971) Relationship between crude-oil properties and geology in some oil and gas fields in the Niigata basin, Japan. Chem. Geol. 8, 219–48.
- [28] Khelifi, N., Amin Ali, O., Roche, P., Grossi, V., Brochier-Armanet, C., Valette, O., Ollivier, B., Dolla, A., Hirschler-Réa, A. (2014) Anaerobic oxidation of long-chain n-alkanes by the hyperthermophilic sulfate-reducing archaeon, *Archaeoglobus fulgidus*. ISME J. 8(11), 2153–66.
- [29] Kim, M., Oh, H.S., Park, S.C., Chun, J. (2014) Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. Int. J. Syst. Evol. Microbiol. 64(PART 2), 346–51.

- [30] Konstantinidis, K.T., Tiedje, J.M. (2005) Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci. U. S. A.* 102(7), 2567–72.
- [31] Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., Marra, M.A. (2009) Circos: An information aesthetic for comparative genomics. *Genome Res.* 19(9), 1639–45.
- [32] Lemoine, F., Domelevo Entfellner, J.B., Wilkinson, E., Correia, D., Dávila Felipe, M., De Oliveira, T., Gascuel, O. (2018) Renewing Felsenstein’s phylogenetic bootstrap in the era of big data. *Nature* 556(7702), 452–6.
- [33] De Ley, J., Cattoir, H., Reynaerts, A. (1970) The quantitative measurement of DNA hybridization from renaturation rates. *Eur. J. Biochem.* 12(1), 133–42.
- [34] Liu, Y.F., Qi, Z.Z., Shou, L. Bin., Liu, J.F., Yang, S.Z., Gu, J.D., Mu, B.Z. (2019) Anaerobic hydrocarbon degradation in candidate phylum ‘*Atribacteria*’ (JS1) inferred from genomics. *ISME J.* 13(9), 1–14.
- [35] Magot, M. (2005) Indigenous Microbial Communities in Oil Fields. In: Ollivier, B., Magot, M., (Eds.), *Petroleum Microbiology*, ASM Press, Washington, DC, pp. 21–34.
- [36] Martínez-Carranza, E., Barajas, H., Alcaraz, L.D., Servín-González, L., Ponce-Soto, G.Y., Soberón-Chávez, G. (2018) Variability of bacterial essential genes among closely related bacteria: The case of *Escherichia coli*. *Front. Microbiol.* 9, 1059.
- [37] Mesleh, M.F., Rajaratnam, P., Conrad, M., Chandrasekaran, V., Liu, C.M., Pandya, B.A., Hwang, Y.S., Rye, P.T., Muldoon, C., Becker, B., Zuegg, J., Meutermans, W., Moy, T.I. (2016) Targeting Bacterial Cell Wall Peptidoglycan Synthesis by Inhibition of Glycosyltransferase Activity. *Chem. Biol. Drug Des.* 87(2), 190–9.
- [38] Miranda-Tello, E., Fardeau, M.L., Joulain, C., Magot, M., Thomas, P., Tholozan, J.L., Olivier, B. (2007) *Petrotoga halophila* sp. nov., a thermophilic, moderately halophilic, fermentative bacterium isolated from an offshore oil well in Congo. *Int. J. Syst. Evol.*

Microbiol. 57(1), 40–4.

- [39] Miranda-Tello, E., Fardeau, M.L., Thomas, P., Ramirez, F., Casalot, L., Cayol, J.L., Garcia, J.L., Ollivier, B. (2004) *Petrotoga mexicana* sp. nov., a novel thermophilic, anaerobic and xylanolytic bacterium isolated from an oil-producing well in the Gulf of Mexico. Int. J. Syst. Evol. Microbiol. 54(1), 169–74.
- [40] Mistry, J., Finn, R.D., Eddy, S.R., Bateman, A., Punta, M. (2013) Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. Nucleic Acids Res. 41(12).
- [41] Mori, K., Yamazoe, A., Hosoyama, A., Ohji, S., Fujita, N., Ishibashi, J.I., Kimura, H., Suzuki, K.I. (2014) *Thermotoga profunda* sp. nov. and *Thermotoga caldifontis* sp. nov., anaerobic thermophilic bacteria isolated from terrestrial hot springs. Int. J. Syst. Evol. Microbiol. 64(PART 6), 2128–36.
- [42] Murtey, M. Das., Ramasamy, P. (2016) Sample Preparations for Scanning Electron Microscopy – Life Sciences. In: Janecek, M., Kral, R., (Eds.), Modern Electron Microscopy in Physical and Life Sciences, In Tech, pp. 161–86.
- [43] Nesbø, C.L., Bradnan, D.M., Adebisuyi, A., Dlutek, M., Petrus, A.K., Foght, J., Doolittle, W.F., Noll, K.M. (2012) *Mesotoga prima* gen. nov., sp. nov., the first described mesophilic species of the *Thermotogales*. Extremophiles 16(3), 387–93.
- [44] Page, A.J., Cummins, C.A., Hunt, M., Wong, V.K., Reuter, S., Holden, M.T.G., Fookes, M., Falush, D., Keane, J.A., Parkhill, J. (2015) Roary: Rapid large-scale prokaryote pan genome analysis. Bioinformatics 31(22), 3691–3.
- [45] Pannekens, M., Kroll, L., Müller, H., Mbow, F.T., Meckenstock, R.U. (2019) Oil reservoirs, an exceptional habitat for microorganisms. N. Biotechnol. 49, 1–9.
- [46] Parker, C.T., Tindall, B.J., Garrity, G.M. (2015) International Code of Nomenclature of Prokaryotes. Int. J. Syst. Evol. Microbiol.

- [47] Pradel, N., Ji, B., Gimenez, G., Talla, E., Lenoble, P., Garel, M., Tamburini, C., Fourquet, P., Lebrun, R., Bertin, P., Denis, Y., Pophillat, M., Barbe, V., Ollivier, B., Dolla, A. (2013) The First Genomic and Proteomic Characterization of a Deep-Sea Sulfate Reducer: Insights into the Piezophilic Lifestyle of *Desulfovibrio piezophilus*. PLoS One 8(1).
- [48] Ravot, G., Magot, M., Fardeau, M.L., Patel, B.K., Prensier, G., Egan, A., Garcia, J.L., Ollivier, B. (1995) *Thermotoga elfii* sp. nov., a novel thermophilic bacterium from an African oil-producing well. Int. J. Syst. Bacteriol. 45(2), 308–14.
- [49] Redder, P., Hausmann, S., Khemici, V., Yasrebi, H., Linder, P. (2015) Bacterial versatility requires DEAD-box RNA helicases. FEMS Microbiol. Rev. 39(3), 392–412.
- [50] Reysenbach, A.L., Liu, Y., Lindgren, A.R., Wagner, I.D., Sislak, C.D., Mets, A., Schouten, S. (2013) *Mesoaciditoga lauensis* gen. nov., sp. nov., a moderately thermoacidophilic member of the order *Thermotogales* from a deep-sea hydrothermal vent. Int. J. Syst. Evol. Microbiol. 63(PART 12), 4724–9.
- [51] Richter, M., Rosselló-Móra, R. (2009) Shifting the genomic gold standard for the prokaryotic species definition. Proc. Natl. Acad. Sci. U. S. A. 106(45), 19126–31.
- [52] Schmid, J., Heider, D., Wendel, N.J., Sperl, N., Sieber, V. (2016) Bacterial glycosyltransferases: Challenges and Opportunities of a Highly Diverse Enzyme Class Toward Tailoring Natural Products. Front. Microbiol. 7, 182.
- [53] Seemann, T. (2014) Prokka: Rapid prokaryotic genome annotation. Bioinformatics 30(14), 2068–9.
- [54] Sekar, R., Shin, H.D., DiChristina, T.J. (2016) Activation of an otherwise silent xylose metabolic pathway in *Shewanella oneidensis*. Appl. Environ. Microbiol. 82(13), 3996–4005.

- [55] Stagars, M.H., Emil Ruff, S., Amann, R., Knittel, K. (2016) High diversity of anaerobic alkane-degrading microbial communities in marine seep sediments based on (1-methylalkyl)succinate synthase genes. *Front. Microbiol.* 6, 1511.
- [56] Takahata, Y., Nishijima, M., Hoaki, T., Maruyama, T. (2001) *Thermotoga petrophila* sp. nov. and *Thermotoga naphthophila* sp. nov., two hyperthermophilic bacteria from the Kubiki oil reservoir in Niigata, Japan. *Int. J. Syst. Evol. Microbiol.* 51(PART 5), 1901–9.
- [57] Tettelin, H., Massignani, V., Cieslewicz, M.J., Donati, C., Medini, D., Ward, N.L., Angiuoli, S. V., Crabtree, J., Jones, A.L., Durkin, A.S., DeBoy, R.T., Davidsen, T.M., Mora, M., Scarselli, M., Margarit y Ros, I., Peterson, J.D., Hauser, C.R., Sundaram, J.P., Nelson, W.C., Madupu, R., Brinkac, L.M., Dodson, R.J., Rosovitz, M.J., Sullivan, S.A., Daugherty, S.C., Haft, D.H., Selengut, J., Gwinn, M.L., Zhou, L., Zafar, N., Khouri, H., Radune, D., Dimitrov, G., Watkins, K., O'Connor, K.J.B., Smith, S., Utterback, T.R., White, O., Rubens, C.E., Grandi, G., Madoff, L.C., Kasper, D.L., Telford, J.L., Wessels, M.R., Rappuoli, R., Fraser, C.M. (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: Implications for the microbial “pan-genome.” *Proc. Natl. Acad. Sci.* 102(39), 13950–5.
- [58] Ujiie, Y., Sherwood, N., Faiz, M., Wilkins, R.W.T. (2004) Thermal maturity and suppressed vitrinite reflectance for Neogene petroleum source rocks of Japan. *Am. Assoc. Pet. Geol. Bull.* 88(10), 1335–56.

Figure2

A

B

C

Figure3

Thermotoga sp. TKOL6
Thermotoga sp. Mc24
T. martima MSB8
Thermotoga sp. 2812B
Thermotoga sp. EMP
T. petrophila RKU-1
T. naphthophila RKU-10
***Thermotoga* sp. TFO**
Thermotoga sp. RQ2
Thermotoga sp. Cell2
Thermotoga sp. Xyl54
Thermotoga sp. TBGT1766
Thermotoga sp. TBGT1765
Thermotoga sp. TBGT1765
T. neapolitana NS-E
Thermotoga sp. RQ7
T. caldifonctis AZM44c09
P. hypogea DSM 11164
P. thermarum DSM 5069
P. profunda AZM34c06
P. lettingae TMO
P. elfii DSM 9442

P. elfii DSM 9442
P. lettingae TMO
P. profunda AZM34c06
P. thermarum DSM 5069
P. hypogea DSM 11164
T. caldifonctis AZM44c09
Thermotoga sp. RQ7
T. neapolitana NS-E
Thermotoga SG1
Thermotoga sp. TBGT1765
Thermotoga sp. TBGT1766
Thermotoga sp. Xyl54
Thermotoga sp. Cell2
Thermotoga sp. RQ2
***Thermotoga* sp. TFO**
T. naphthophila RKU-10
T. petrophila RKU-1
Thermotoga sp. EMP
Thermotoga sp. 2812B
T. martima MSB8
Thermotoga sp. Mc24
Thermotoga sp. TKOL6

Figure5

Figure6

[Click here to access/download](#)

Table

Table1.docx

Click here to access/download

Table

Table2.docx

Click here to access/download
Table
Table3.docx

Click here to access/download

Table

Table4.docx