UNIVERSIDADE D
COIMBRA

Luciana Costa de Albuquerque Pinto

Classification and identification of thermophilic organisms isolated from São Pedro do Sul hot spring and revision of the classification of genus *Meiothermus*:

Integration of genomics into prokaryotic taxonomy

Tese no âmbito do doutoramento em Biociências no ramo de especialização em Microbiologia e apresentada ao Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Julho de 2021

Faculdade de Ciências e Tecnologia da Universidade de Coimbra
Departamento de Ciências da Vida

CLASSIFICATION AND IDENTIFICATION OF THERMOPHILIC ORGANISMS ISOLATED FROM SÃO PEDRO DO SUL HOT SPRING AND REVISION OF THE CLASSIFICATION OF GENUS *MEIOTHERMUS*: Integration of Genomics Into Prokaryotic Taxonomy

Luciana Costa de Albuquerque Pinto

Tese no âmbito do doutoramento em Biociências no ramo de especialização em Microbiologia e apresentada ao Departamento de Ciências da Vida da Faculdade de Ciências e Tecnologia da Universidade de Coimbra

Julho de 2021

1 2 9 0

UNIVERSIDADE D
COIMBRA

This research was supported by the European Union's Horizon 2020 Research and Innovation programme under Metafluidics Grant Agreement No 685474. This work was also supported by: FEDER funds through the Operational Programme Competitiveness Factors - COMPETE 2020 and national funds by FCT – Foundation for Science and Technology under the strategic project UID/NEU/04539/2013; Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 and national funds by FCT – Foundation for Science and Technology under the strategic project UIDB/04539/2020; GenomePT project (POCI-01-0145-FEDER-022184), with funds from COMPETE 2020 - Operational Programme for Competitiveness and Internationalisation (POCI), Lisboa, Portugal Regional Operational Programme (Lisboa2020), Algarve Portugal Regional Operational Programme (CRESC Algarve2020), under the PORTUGAL 2020 Partnership Agreement, through the European Regional Development Fund (ERDF), and by Foundation for Science and Technology (FCT).

Dedico este trabalho ao meu Professor Milton Simões da Costa (*in memorian*). Saudade eterna.

Dedico este trabalho ao meu Pai Ciro Craveiro de Albuquerque Pinto (*in memorian*). Saudade eterna.

Dedico este trabalho, com muito amor e gratidão, à minha mãe Maria de Lurdes Costa Lages Pinto.

Agradecimentos

Ao Professor Milton Simões da Costa que esteve sempre presente em todo o meu percurso académico. Esta dissertação representa todo o trabalho e amizade que desenvolvemos durante 25 anos. Sem a sua orientação científica durante todos estes anos a concretização deste trabalho não teria sido possível. Eternamente agradecida por tudo o que tive oportunidade de aprender e vivenciar durante estes anos na sua companhia! Muito obrigada Professor!

À Doutora Conceição Egas por todo o suporte e incentivo que demonstrou durante o difícil ano de 2020. Muito obrigada por toda a disponibilidade e apoio na elaboração desta tese!

Ao Professor António Veríssimo pela disponibilidade e aconselhamentos. Muito obrigada!

Ao Centro de Neurociências e Biologia Celular (CNC) da Universidade de Coimbra que proporcionou as condições necessárias para a realização deste trabalho.

A todos os amigos e colegas que se cruzaram comigo pelo Laboratório de Microbiologia do Departamento de Ciências da Vida da FCTUC no Edifício do Patronato. Cada um à sua maneira e no seu tempo acompanhou o meu percurso na Microbiologia, contribuindo também para a elaboração da presente tese.

A todos os amigos, colegas e familiares que me acompanharam em algum período nesta caminhada. Obrigada por tão bons momentos!

Ao Bruno, pelas longas conversas! Obrigada *primo*!

Ao Pai Ciro e à Mãe Lurdes, agradeço-vos a Vida e todos os ensinamentos!

Cláudia, obrigada por estares sempre aí! Estamos juntas!

Mãe guerreira com coração gigante, obrigada por tudo! Estamos juntas!

Resumo

A vida em ambientes extremos tem sido intensamente estudada ao longo dos anos, com especial atenção para a diversidade dos microrganismos e para os mecanismos moleculares envolvidos na sua adaptação a esses ambientes. Organismos que prosperam sob condições ambientais extremas são designados de extremófilos. Organismos que apresentam crescimento ótimo a temperaturas elevadas são designados termófilos. Estes organismos podem ser encontrados em ambientes geotérmicos continentais, hidrotérmicos marinhos e em ambientes quentes artificiais. As estratégias de adaptação utilizadas por estes microrganismos devem-se muitas vezes à produção de biomoléculas específicas que têm vindo a ser estudadas e aplicadas em diversas áreas da biotecnologia ambiental, industrial e farmacêutica.

Em Portugal, continental e ilhas, podem ser encontradas várias fontes termais com uma vasta diversidade em microrganismos termófilos. O objetivo deste trabalho de investigação foi expandir os conhecimentos sobre a microbiologia destes ambientes termofílicos, nomeadamente na nascente termal de São Pedro do Sul, localizada no centro de Portugal. A água termal emerge à superfície terrestre a uma temperatura de cerca de 68,0ºC. Tendo em conta os avanços tecnológicos da biologia molecular, especialmente relacionados com a genómica, esta tese aborda a taxonomia de novos isolados da fonte termal de S. Pedro do Sul e a revisão da classificação do género *Meiothermus*, com a integração da sequenciação de genomas e ferramentas de genómica comparativa na abordagem polifásica para descrição e classificação de microrganismos.

Um recente organismo, foi isolado de uma amostra de biofilme da fonte termal em S. Pedro do Sul. Este microrganismo que forma colónias pigmentadas de laranja está relacionado com as espécies do género *Thermonema* da família *Thermonemataceae*, e representa o primeiro membro cultivado de uma nova linhagem distinta da ordem *Cytophagales* do filo *Bateroidetes*. O isolado SPSPC-11^T possui características genotípicas e fenotípicas que se assemelham às das espécies do género *Thermonema*, mas apresentam diferenças notáveis em diversas características. Análises da sequência do genoma corroboraram muitas das características fenotípicas da estirpe SPSPC-11^T, como a incapacidade de assimilar açúcares. Com base na caracterização fisiológica, bioquímica, quimiotaxonómica, genómica e filogenética foi possível descrever uma nova espécie de um novo género representada pela estirpe SPSPC-11^T para a qual foi proposto o nome *Raineya orbicola*. A família *Raineyaceae* também foi descrita para acomodar este novo género e espécie.

Representantes do filo *Proteobacteria* foram também isolados de uma amostra de água de S. Pedro do Sul, comprovando a diversidade microbiana destes ambientes termofílicos. A análise filogenética de duas estirpes, SPSP-6^T e SPSPC-18, que formam colónias não pigmentadas e demonstraram ter 100% de similaridade na sequência do gene 16S rRNA, posicionou estes isolados num dos dois grupos formados pelo género *Tepidimonas*. Uma vez que apenas um pequeno número de características fenotípicas e quimiotaxonómicas permitem a distinção das espécies do género *Tepidimonas*, realizou-se uma análise comparativa dos genomas das estirpes tipo do género

Tepidimonas, do novo isolado SPSP-6^T e de estirpes filogeneticamente relacionadas para avaliar a taxonomia destes organismos. Os valores do índice global de relacionamento genómico corroboraram os resultados da análise filogenética baseada na sequência do gene 16S rRNA e de 400 sequências de genes conservados, confirmando as linhagens distintas das oito espécies. Considerando a análise fenotípica, quimiotaxonómica, genómica e filogenética, todas as estirpes foram circunscritas ao género *Tepidimonas*. A estirpe SPSP-6^T foi descrita como uma nova espécie do género *Tepidimonas* e denominada *Tepidimonas charontis*.

Um grupo de bactérias frequentemente encontrado em fontes termais com características como as de S. Pedro do Sul inclui bactérias do género *Meiothermus*, que pertence à família *Thermaceae*, e representa uma linhagem distinta dentro do filo *Deinococcus-Thermus*. Estirpes do género *Meiothermus* foram isoladas de várias áreas geotérmicas por todo o mundo. A estirpe tipo da espécie *M. timidus* SPS-243^T foi isolada da fonte termal de S. Pedro do Sul. Das treze espécies do género *Meiothermus*, nove formam colónias pigmentadas de vermelho e quatro formam colónias pigmentadas de amarelo. As relações filogenéticas intragenéricas das treze espécies do género *Meiothermus* foram avaliadas por filogenómica. Os resultados indicaram que estas espécies formam pelo menos três linhagens distintas principais que podem ser consideradas como representativas de géneros distintos. A heterogeneidade do género foi sustentada pelos parâmetros do índice global de relacionamento genómico. Com base nos resultados da análise filogenética e genómica, e nas características quimiotaxonómicas e fenotípicas, o género *Meiothermus* foi circunscrito a oito espécies, o que levou a emendar a descrição do género *Meiothermus* e à reclassificação das quatro espécies que formam colónias pigmentadas de amarelo *M. chliarophilus*, *M. roseus*, *M. terrae* e *M. timidus* como membros de um novo género chamado *Calidithermus*, no entanto não foi possível clarificar a classificação de *M. silvanus*.

Vários grupos taxonómicos dos domínios *Bacteria* e *Archaea* foram sujeitos a reclassificações a partir do momento que a informação do genoma começou a ser utilizada mais frequentemente pelos taxonomistas. Estudos taxonómicos recentes comprovam que a incorporação da genómica na taxonomia dos procariontes permite alcançar conclusões mais robustas sobre a caracterização de microrganismos, constituindo um passo importante para o aperfeiçoamento da classificação dos procariontes.

Palavras-chave: Taxonomia, classificação, polifásica, procariontes, termófilo, termofílico, fonte termal, fenotípica, genómica, biotecnologia.

Abstract

Life in extreme environments has been studied intensively over the years, focusing on the diversity of organisms and the molecular mechanisms involved in their adaptation in these environments. Organisms that thrive under extreme environmental conditions are referred to as extremophiles. Organisms that grow optimally at high temperatures are designated thermophiles and are found in continental geothermal, marine hydrothermal and man-made hot environments. The search for the existence of microorganisms in these environments has brought great insight into microbial diversity and the adaptation mechanisms that allow thermophiles to live and grow in these extreme conditions. Their adaptation strategies are often due to the production of specific biomolecules that have been studied for years as sources for environmental, industrial, and pharmaceutical biotechnological applications. However, much remains to be known about the microbial diversity of these extreme environments.

In Portugal, in the mainland and islands, there are several hot springs with great diversity in thermophilic microorganisms. This research work aimed to extend the current knowledge on the microbiology of these thermophilic environments, specifically on the hot spring of São Pedro do Sul, located in central Portugal. This thermal water has a temperature of about 68.0ºC. Benefiting from the technological advances in molecular biology, especially those related to genomics, this thesis addresses the taxonomic study of new isolates from S. Pedro do Sul hot spring and the revision of the classification of the genus *Meiothermus*, integrating genome sequencing and comparative genomics tools into the polyphasic approach for microorganism description and classification.

The new strain SPSPC-11^T was isolated from a biofilm sample from the S. Pedro do Sul hot spring. This orange-pigmented colony forming microorganism was most closely related to the species of the genus *Thermonema* of the family *Thermonemataceae* and represents the first cultured member of a new distinct lineage of the order *Cytophagales* of the phylum *Bateroidetes*. The isolate SPSPC-11^T possesses genotypic and phenotypic features that resemble those of the species of the genus *Thermonema*, but there are notable differences in several characteristics. The high-quality draft genome sequence analyses corroborated many of the phenotypic characteristics of the strain SPSPC-11^T, such as the inability to assimilate sugars. Based on physiological, biochemical, chemotaxonomic, genomic and phylogenetic characterization, a new species of a novel genus was described, represented by strain SPSPC-11^T for which the name *Raineya orbicola* was proposed. The family *Raineyaceae* was also described to accommodate this new genus and species.

Representatives of the phylum *Proteobacteria* were also isolated from a water sample of S. Pedro do Sul hot spring, supporting the diversity of microbes that inhabits thermophilic environments. The phylogenetic analysis of two non-pigmented strains, SPSP-6^T and SPSPC-18, that share 100% pairwise similarity of the 16S rRNA gene sequence, located these isolates within one of the two clusters formed by the genus *Tepidimonas*. Only a small number of phenotypic and

chemotaxonomic characteristics distinguish the species of *Tepidimonas*, which led to a comparative analysis of the genome sequences of the eight type strains of the genus *Tepidimonas*, the new isolate SPSP-6^T and three closely related strains to assess the taxonomic position of the organisms. The values of the overall genome relatedness index corroborated the results of the phylogenetic analysis based on 16S rRNA gene sequence and 400 conserved genes sequences, regarding the distinct lineages of the eight species. Considering the phenotypic, chemotaxonomic, genomic, and phylogenetic analysis, all strains were circumscribed to the genus *Tepidimonas*. Strain SPSP-6^T was described as a new species of the genus *Tepidimonas* and named *Tepidimonas charontis.*

A group of bacteria most commonly found in hot springs with characteristics such as those of S. Pedro do Sul is the genus *Meiothermus*, which belong to the family *Thermaceae*, and represents a distinct lineage within the phylum *Deinococcus-Thermus*. Strains of the genus *Meiothermus* have been isolated from several geothermal areas around the world. The type strain *M. timidus* SPS-243^T was isolated from S. Pedro do Sul hot spring. Of the thirteen species of genus *Meiothermus* nine species form red-pigmented colonies and four species form yellow-pigmented colonies. The intrageneric phylogenetic relationships of the thirteen type strains of the genus *Meiothermus* were assessed by phylogenomics. Results indicated that species of the genus *Meiothermus* form at least three major distinct lineages that may be considered to represent distinct genera. The overall genome relatedness index parameters supported the heterogeneous nature of the genus. Based on the results of the phylogenetic and genomic analysis, chemotaxonomic and phenotypic characteristics, the genus *Meiothermus* was circumscribed to eight species, which led to emend the description of the genus *Meiothermus* and the reclassification of the four yellow-pigmented species *M. chliarophilus*, *M. roseus*, *M. terrae* and *M. timidus* as members of a novel genus named *Calidithermus*, however the classification of *M. silvanus* could not be clarified.

Several taxonomic taxa of *Bacteria* and *Archaea* underwent revisions and reclassifications when the genomic data start to be more commonly used by taxonomists. Recent taxonomic studies support that the approach of incorporating genomics into prokaryotic taxonomy provides more robust conclusions about the characterization of microorganisms, constituting an important step towards improving classification of prokaryotes.

Keywords: Taxonomy, classification, polyphasic, prokaryotes, thermophile, thermophilic, hot spring, phenotypic, genomics, biotechnology.

Abbreviations

AAI	average amino acid identity
ALs	aminolipids
ANI	average nucleotide identity
ANIb	average nucleotide identity based on the BLASTn algorithm
ANIm	average nucleotide identity based on the MUMmer algorithm
APLs	aminophospholipids
ATCC	The American Type Culture Collection
BLAST	basic local alignment search tool
cAAI	core-gene average amino acid identity
CAPSO	3-(cyclohexylamino)-2-hydroxy-1-propanesulfonic acid
CDS	protein coding sequences
CECT	Colección Española de Cultivos Tipo
COG(s)	cluster(s) of orthologous groups of protein
CRISPR	clustered regularly interspaced short palindromic repeats
dDDH	digital DNA-DNA hybridization
DDH	DNA-DNA hybridization
DNA	deoxyribonucleic acid
DSM	Deutsche Sammlung von Mikroorganismen (German Collection of Microorganisms and Cell Cultures)
EPS(s)	exopolysaccharide(s)
FAMEs	fatty acid methyl esters
G+C	guanine plus cytosine
GC	gas chromatography
GLC	gas liquid chromatography
GenBank/EMBL/DDBJ	International Nucleotide Sequence Database / European Molecular Biology Laboratory / DNA Data Bank of Japan
GL(s)	glycolipid(s)
GTDB	Genome Taxonomy Database
H_2S	hydrogen sulfide
HEPES	4-(2-hydroxyethyl)-1-piperazinethanesulfonic acid
HPLC	high-performance liquid chromatography
HQG	high-quality draft genome
ICNP	International Code of Nomenclature of Prokaryotes
ICSP	International Committee on Systematics of Prokaryotes
KCTC	Korean Collection for Type Cultures
KEGG	Kyoto Encyclopedia of Genes and Genomes

KO	KEGG orthology
LGT	lateral gene transfer
LMG	Laboratorium voor Microbiologie Universiteit Gent
LTP	Living Tree Project
LUCA	last universal common ancestor
MALDI-TOF MS	matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MEGA	molecular evolutionary genetics analysis
MES	2-(N-morpholino)-ethanesulfonic acid
MIGS	minimum information about a genome sequence
MIS	Microbial Identification System
MK	menaquinone
ML	maximum likelihood
MSLA	multilocus sequence analysis
MUMi	maximally unique matches index
NBRC	NITE (National Institute of Technology and Evaluation) Biological Resource Center, Japan
NCBI	National Center for Biotechnology Information
NGS	next-generation sequencing
NJ	neighbor joining
ODS2	octadecyl-silica 2
OGRI	overall genome relatedness index
OGs	orthologous groups
OH	hydroxy
OrthoANI	average nucleotide identity by orthology
PCR	polymerase chain reaction
PE	paired-end
PGP	Prokaryotic Genome Prediction
PGP2	Prokaryotic Genome Prediction 2
PHA	polyhydroxyalkanoates
PL	phospholipid
POPC	percentage of conserved proteins
RAPD	randomly amplified polymorphic DNA
RaxML	randomized axelerated maximum likelihood
RBM(s)	reciprocal best matche(s)
RBR	relative binding ratio
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RuBisCo	ribulose 1,5-biphosphate carboxylase/oxygenase

SEM	scanning electron microscopy
SSU(s)	small subunit(s)
TAPS	N-[tris(hydroxymethyl)methyl]-3-aminopropanesulfonic acid
TEM	transmission electron microscopy
TETRA	tetranucleotide signature regression
TLC	thin-layer chromatography
TSBA	tryptic soy broth agar
ULs	unidentified lipids
UV	ultraviolet
v/v	volume/volume
VP	validly published
w/v	weight/volume
WGS	whole genome sequence

Figures Index

Tables Index

Contents

Chapter 1

Introduction

1.1 Thermophilic Microorganisms

1.1.1 Life at high temperatures

Over the last decades, the boundary conditions under which life can thrive have been studied in diverse physical as well as geochemical extreme conditions, covering broad ranges of temperature, pH, salinity, pressure, desiccation, radiation, oxygen tension, chemical extremes, energy and nutrient limitation (Lever *et al*., 2015; Rothschild and Mancinelli, 2001). Microorganisms do not only thrive under such a wide range of parameters on Earth but can also survive the harsh conditions of outer space, an environment with extreme radiation, vacuum pressure, extremely variable temperature and microgravity (DasSarma and DasSarma, 2018; DasSarma *et al*., 2017; Horneck *et al*., 2010; Merino *et al*., 2019; Yamagishi *et al*., 2018; Yang *et al*., 2008). Organisms that thrive in an extreme environment, particularly hostile to humans and to the majority of the known living organisms, capable of growing optimally at or near to the extreme ranges of a particular environmental parameter, but also requiring them to proliferate are designated extremophiles (Canganella and Wiegel, 2011; Gupta *et al*., 2014; Horikoshi and Bull, 2011; Rothschild and Mancinelli, 2001). Extremophiles that thrive in environments with more than one extreme parameter are designated polyextremophiles (Capece *et al*., 2013; Gupta *et al*., 2014). Many terrestrial and extraterrestrial environments feature conditions that fall within more than one extreme, which led to the extensive study of extremophiles in different environments through the last decades to answer questions about microbial diversity, biogeography, biotechnology and astrobiology (Coker, 2019; Harrison *et al*., 2013; Martin and McMinn, 2018; Merino *et al*., 2019; Pikuta *et al*., 2007; Satyanarayana *et al*., 2005). Although extremophiles cover all the three domains of life, they are mainly prokaryotic belonging to the domain *Archaea* and *Bacteria*, even though extremophily is being increasingly reported among algae and fungi (Busk and Lange, 2013; Malavasi *et al*., 2020; Varshney *et al*., 2015; Zhang *et al*., 2018).

One of the major groups and most studied among extremophiles are the microorganisms that thrive at high temperatures. Temperature is one of the most important environmental factors affecting growth of organisms and amazingly different optimal growth temperatures have been encountered. The classification of organisms based on their growth temperature is considered a fundamental feature of microbiological taxonomy. The use of cardinal growth temperature as criteria for grouping organisms is a suitable method. The cardinal temperatures can be defined as T_{max} or T_{min}, corresponding to the highest or lowest temperature where growth and multiplication occur, respectively, and T_{opt}, the temperature at which the shortest doubling time of biomass or cell number occurs (Wiegel *et al*., 1985). Organisms that usually thrive in a temperature range of about 7–10°C to 35–42°C are designated mesophiles, with thermophiles and psychrophiles growing optimally in higher and lower temperature ranges, respectively. An organism that has T_{opt} in the mesophilic range but is able to grow and multiply in temperatures out of the range is considered thermotolerant. Thermophilic microorganisms or thermophiles are commonly classified into

moderate or slightly thermophilic (T_{opt} at 45–60ºC), extreme thermophiles (T_{opt} at 60–80ºC) and hyperthermophiles (T_{opt} at 80ºC or above) (Canganella and Wiegel, 2011; Gupta *et al*., 2014; Wiegel *et al*., 1985). The organisms with the highest growth temperatures (103–122ºC) belong to the domain *Archaea*. They are members of the genera *Pyrococcus*, *Pyrodictium*, *Hyperthermus*, *Pyrolobus* and *Methanopyrus*, while in the case of the domain *Bacteria*, the species *Thermotoga maritima*, *Thermosulfurimonas dismutans*, *Aquifex pyrophilus* and *"Aquifex aeolicus"*, and "*Geothermobacterium ferrireducens"* exhibit the highest growth temperatures of 90, 92, 95 and 100ºC respectively (Clarke, 2014; Huber *et al*., 1992; Slobodkin *et al*., 2012). For several years, the archaeon *Pyrolobus fumarii*, isolated from a hydrothermal vent at the Mid Atlantic Ridge, was the record holder of 113ºC for the maximum growth temperature (Blöchl *et al*., 1997; Stetter, 2006a). In 2003, a new record was established with the isolation of the strain 121 from a water sample from a hydrothermal vent, located in the Mothra hydrothermal vent field in the Northeast Pacific Ocean, that is able to grow at 121ºC, a strain most closely related to species *Pyrodictium occultum* and *Pyrobaculum aerophilum* of the domain *Archaea* (Kashefi and Lovley, 2003). Strain 121 is designated as "*Geogemma barossii"* but the name of the genus and species was not validly published. Subsequently, strain 116 isolated from the Kairei hydrothermal field in the Central Indian Ridge was found to grow at 122°C under 40MPa pressure, just exceeding the previous record for 121°C for strain 121 (Takai *et al*., 2008). Strain 116 belongs to the species *Methanopyrus kandleri* and the type strain of *M. kandleri* $AV19^T$ isolated from sediment samples from Guaymas Basin hot vents in the Gulf of California is able to grow at 110ºC (Huber *et al*., 1989; Kurr *et al*., 1991). All the thermophilic microorganisms described belong to the domain *Archaea* and *Bacteria*, and almost all of the hyperthermophiles belong to the domain *Archaea*. Eukaryotes do not seem to be able to live at the very highest temperatures as archaea and bacteria. The upper limit for eukaryotes appears to be around 60ºC, a temperature suitable for some algae, fungi and protozoa (Clarke, 2014; Rothschild and Mancinelli, 2001). For most known species of the domain *Eukarya*, temperatures approaching 100ºC usually denature proteins and nucleic acids, degrade many metabolites and increase the fluidity of membranes to lethality (Clarke, 2014; Dilly *et al*., 2012; Pikuta *et al*., 2007; Rothschild and Mancinelli, 2001).

One of the cornerstones of the past century in microbiology was the division of the known living world in three domains, *Archaea*, *Bacteria* and *Eukarya*, to build a phylogenetic tree of life, based on the phylogenetic analysis of the small subunits (SSUs) 16S/18S of the rRNA gene sequences present in all living organisms (Woese and Fox, 1977; Woese *et al*., 1990). One of the first attempts to determine the temperature of ancient life was based on the distribution of hyperthermophilic archaea and bacteria in the Tree of Life (Stetter, 1996, 2006b; Pace, 1997; Wiegel and Canganella, 2001). In contrast to the *Eukarya*, the bacterial and archaeal domains where the thermophiles and hyperthermophiles are included, exhibit some extremely short and deep branches near to the phylogenetic root indicating a rather slow rate of evolution and suggesting that the last universal common ancestor (LUCA) was a hyperthermophile (Stetter, 1996). Through the years, a diversity of studies has corroborated the theory that LUCA probably was a hyperthermophile or a

thermophile (Brooks *et al*., 2004; Di Giulio, 2000; Gaucher *et al*., 2010; Iwabata *et al*., 2005; Shimizu *et al*., 2007) living at an environmental temperature similar to our days hot springs (Gaucher *et al*., 2010). However, other authors sustained that a mesophilic or a moderate thermophilic organism represent the first life form (Becerra *et al*., 2007; Boussau *et al*., 2008; Forterre, 1996; Galtier *et al*., 1999; Glansdorff *et al*., 2008). New versions for the universal tree were proposed, using the phylogenomic analysis of various universal protein markers available from the increasing number of sequenced genomes as an alternative to SSU rRNA (Ciccarelli *et al*., 2006; Gribaldo *et al*., 2010; Forterre, 2015). More recently, the inclusion of genomes sequences of uncultivated organisms from metagenomics studies, intensely expanded versions of Tree of Life have been proposed, however, an outstanding feature of these trees is a large number of major lineages without isolated representatives (Castelle and Banfield, 2018; Cavalier-Smith and Chao, 2020; Hug *et al*. 2016; Parks *et al*., 2018; Zhu *et al*., 2019). In 2016 Weiss and collaborators proposed a newer two-domain tree of life, based on the phylogeny of universal genes, positioning LUCA as the common ancestor of *Bacteria* and *Archaea* and with *Eukarya* arising from prokaryotes, both *Bacteria* and *Archaea* (Weiss *et al*., 2016, 2018) (Figure 1.1). In this proposal, LUCA could have been thermophilic and lived in a hydrothermal vent setting. In 2020, Williams and collaborators also proposed a two-domain tree of life using phylogenomics, supporting a close relationship between eukaryotes and Asgard, one of the recently discovered superphylum of archaea (Imachi *et al*., 2020; Liu *et al*., 2021; Williams *et al*., 2020).

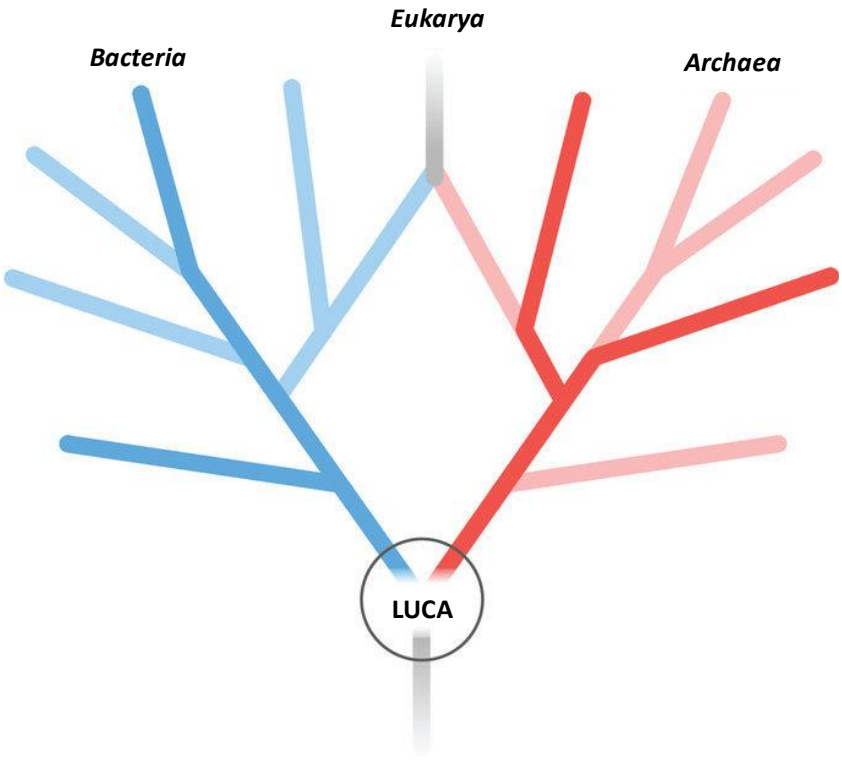

Figure 1.1 A schematic of the two-domain tree of life hypothesis (adapted from Weiss *et al*., 2016).

1.1.2 Habitats

The basis of our understanding of microbial life at high temperatures stems from the pioneering work of Thomas Brock in Yellowstone National Park, EUA (Brock, 1978). Brock first reported microorganisms (microalgae) growing at high temperatures in the 1960s (Brock and Brock, 1966; Brock, 1967b) and isolated several bacterial strains with an optimum growth temperature of 70°C which described as *Thermus aquaticus*, a new species of a novel genus of a thermophilic bacterium (Brock and Freeze, 1969). This organism has proved to be of enormous significance as the DNA polymerase source, the fundamental enzyme to the polymerase chain reaction procedures (PCR) that has revolutionized biotechnology. Based on his experiences in Yellowstone, Brock anticipated that life could be found wherever water was liquid, a prediction that was justified by the discovery of microbial life at very high temperatures and pressures associated with hydrothermal vents in the 1970s (Corliss *et al*., 1979; Rona *et al*., 1986).

The natural habitats of thermophilic microorganisms range from continental geothermal areas associated with tectonic activity with temperatures ranging from slightly above ambient to boiling point of water, geothermally heated subterranean oil reservoirs and submarine hydrothermal fields associated with the spreading ridge systems of the deep ocean, like sediments, submarine volcanoes, fumaroles and vents (black smokers or white smokers) with temperatures exceeding 300°C (Mehta and Satyanarayana, 2013). Other naturally occurring hot places usually are more transient, such as solar-heated ponds and soils with temperatures up to 65ºC. There are also human-made hot environments such as compost piles where the temperature is usually around 60–70ºC but as high as 100ºC, industrial processes and water heaters (Freitas *et al*., 2003; Oshima and Moriya, 2008; Rastogi *et al*., 2010).

The continental geothermal areas are mainly two types resulting from geological differences in the heat source: the low pH type designated solfatara fields and neutral to alkaline pH type characterized by freshwater hot springs. Solfatara fields are also called high-temperature fields, primarily located within active volcanic zones, in the form of boiling mudpot, steam holes, or fumaroles with issuing volcanic gases with relatively high concentrations of sulfuric compounds, mainly H_2S (with a characteristic odour of rotten eggs) (Figure 1.2 a, b, c). On the surface, H_2S is oxidized to sulfur and then to sulfuric acid, which lowers pH, causing corrosion of the surrounding rocks and formation of the typical acidic mud of solfatara fields (Kristjansson *et al*., 2000). The classical example is the Solfatara Crater located in the town of Pozzuoli, Naples. It represents the most prominent surface hydrothermal manifestation in the Campi Flegrei caldera (Caliro *et al*., 2007; Crognale *et al*., 2018). Microorganisms that can survive and grow under these acidic and thermophilic conditions are called thermoacidophiles and are widely distributed in the bacterial and archaeal domains (Baker-Austin and Dopson, 2007; Oren, 2018). In addition to the Solfatara of Pozzuoli in Italy, there are other places in the world with solfataric fields where thermoacidophilic microorganisms were recovered or identified, as is the case of the genera *Sulfolobus* (Brock, 1972; Colman *et al*., 2018; Huber and Stetter, 1991), *Saccharolobus* (Sakai and Kurosawa, 2018; Zillig

et al., 1980), *Thermoplasma* (Crognale *et al*., 2018; Segerer *et al*., 1988), *Ferroplasma* and *Acidithiobacillus* (Crognale *et al*., 2018). In Portugal, bacterial strains of the thermoacidophilic genus *Alicyclobacillus*, namely strains of the species *Alicyclobacillus hesperidum* and strains of a genomic species similar to species *A. acidocaldarius* have been isolated from the solfataric area at Furnas on the Island of São Miguel in the Azores (Albuquerque *et al*., 2000). Freshwater hot springs, also called low temperature fields, are mainly located outside active volcanic zones and are heated by extinct deep lava flows or dead magma chambers (Kristjansson *et al*., 2000). Groundwater percolates into these hot areas, warms up and returns to the surface to form hot springs, some of them with pleasantly warm water that can be frequented by bathers (Figure 1.2 d) and many of them explored as a Spa ("*Salus per Aquam*").

Figure 1.2 Images of different types of geothermal springs in Furnas, Island of São Miguel, Azores, Portugal. a, fumarole; b, fumarole and mudpot; c, boiling mudpot; d, the spring water pool of Terra Nostra Park (photographs by Albuquerque L. during the 8th International Conference on Extremophiles, Ponta Delgada, 2010).

Beyond Yellowstone National Park and Italy many other geothermal areas are found on Earth where geothermally heated water reaches the surface, namely Turkey (Pamukkale), India, China, Japan, New Zealand, Russia (Kamchatka) and Chile. In Portugal, the hot springs of the Island of São Miguel, Azores are the most studied, where through the years, a diversity of thermophilic bacteria from different phyla such as *Actinobacteria*, *Aquificae*, *Bacteroidetes*, *Deinococcus–Thermus*, *Firmicutes*, *Proteobacteria* have been isolated (Aguiar *et al*., 2004; Albuquerque *et al*., 2000, 2002, 2005, 2008, 2010a, 2010b, 2011b, 2012a, 2012b, 2013, 2014; França *et al*., 2006; Nunes *et al*., 1992; Pires *et al*., 2005a; Williams *et al*., 1996). Recently, studies on the microbial diversity of deep-sea hydrothermal sediments of Azores vent fields have been performed using a comparative metagenomic analysis where thermophilic microorganisms were detected, suggesting a reasonable diversity of thermophilic taxa associated with these sediments (Cerqueira *et al*., 2017, 2018) and a novel thermophilic species was isolated (Reiner *et al*., 2018). In Portugal mainland, we also find several hot springs throughout the territory (Cantista, 2008), namely Alcafache, São Gemil, Vizela, Chaves and São Pedro do Sul, from where several thermophiles have been isolated and described (Alves *et al*., 2003; Moreira *et al*., 2000; Pires *et al*., 2005a; Rainey *et al*., 2003; Santos *et al*., 1989; Tenreiro *et al*., 1995; Albuquerque L. personal unpublished results).

In this thesis, the study focused on the São Pedro do Sul hot spring, located in central Portugal, where slightly alkaline and sulfurous water emerges from the interior of the earth to the surface with a temperature of about 68.0ºC (https://termas-spsul.com/en/termas-s-pedro-do-sul-2/natural-mineral-water/). In the last years, several slightly thermophilic bacteria have been isolated from this hot spring. Sampling in the years 2013, 2015 and 2016 resulted in the isolation of 197 thermophilic organisms, preserved and maintained in the private culture collection of the Microbiology Laboratory of the Center for Neuroscience and Cell Biology, University of Coimbra (Figure 1.3). The most represented microorganisms in São Pedro do Sul hot spring belong to the genera *Tepidimonas*, *Meiothermus* and *Thermus* and are the subject of study in this thesis.

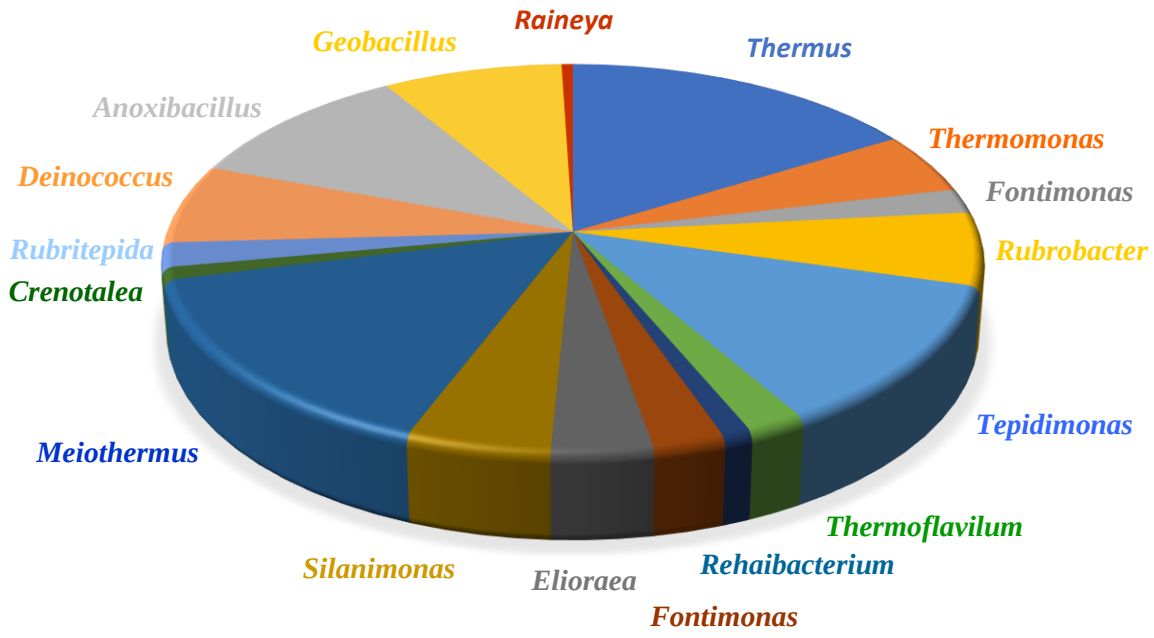

Figure 1.3 Diversity of thermophilic bacteria of the São Pedro do Sul hot spring. The isolates were recovered from sampling in the years 2013, 2015 and 2016. Taxonomic affiliations were determined through the 16S rRNA gene sequences (Albuquerque L. personal unpublished results).

1.1.3 Biochemical features of thermophiles heat stability

The thermophilic nature of an organism is primarily based on the inherent thermostability of its cellular structures and macromolecules. The maintenance of appropriate membrane fluidity in thermophilic bacteria is one of the strategies for the ability of these microorganisms to live in high-temperature environments (Brock, 1967a; Siliakus *et al*., 2017). Generally, membrane fluidity increases with the increment in temperature (Siliakus *et al*., 2017), and to maintain and to keep the optimum fluidity of the membrane to cope with life at high temperature, the cell requires a proper lipid composition (Koga, 2012; Sohlenkamp and Geiger, 2016). Changes in the fatty acid of membrane lipids can occur due to adaptation to temperature, such as the degree of fatty acid unsaturation, chain-length, branching and cyclization (Suutari and Laakso, 1994). Thermophilic bacteria mainly adjust fluidity by increasing the amount of saturated fatty acids (Oshima and

Miyagawa, 1974), branched-chain iso-fatty acids (Patel *et al*., 1991) or through the presence of long-chain diols (Pond and Langworthy, 1987; Wait *et al*., 1997). The archaea, which compose most of the hyperthermophiles, instead of fatty acids, have lipids linked with ether on the membrane and sometimes form a tetraether monolayer, a possible strategy to limit mobility and to guarantee membrane functionality (Bartucci *et al*., 2005; De Rosa *et al*., 1994; Gambacorta *et al*., 1995; Siliakus *et al*., 2017). Likewise, in the hyperthermophilic bacteria *Aquifex pyrophilus* and *Thermotoga maritima*, the core lipids are characterized by the presence of ether bonds, and in the latter the ether lipids can be arranged in tetraethers structures (Damsté *et al*., 2007; Huber *et al*., 1986, 1992). However, a significant number of hyperthermophilic archaea do not contain tetraether lipids in their membranes, suggesting that the presence of bipolar tetraether lipids is not a requirement for thermal adaptation (Koga, 2012; Ulrih *et al*., 2009).

The hyperthermophilic archaea and bacteria contain a reverse DNA gyrase introducing positive supercoils, which increases DNA stability at high temperatures (López-García, 1999; Forterre, 2002; Ogawa *et al*., 2015). The presence of histones also accounts for DNA stability in hyperthermophilic archaea (Grayling *et al*., 1996; Henneman *et al*., 2018; Stevens *et al*., 2020). Some polyamines, as triamines (spermidine, nonspermidine, homospermidine), quaternary branched penta-amines and linear penta- and hexa-amines, play important roles in the stabilization of DNA and RNA molecules in several thermophilic and hyperthermophilic organisms (Hosoya *et al*., 2004; Michael, 2016). Unusual longer polyamines (i.e., caldopentamine and caldohexamine) and branched polyamines (tetrakis(3-aminopropyl)ammonium) are produced by the species *Thermus thermophilus* and *Thermomicrobium roseum* providing thermal protection to nucleic acids (Hamana *et al*., 1990; Terui *et al*., 2005; Oshima, 2007). Additionally, the G+C (guanine plus cytosine) content of the secondary structures of ribonucleic acids increases with growth temperature (Galtier and Lobry, 1997), in RNA, the high G+C content is concentrated in the double-stranded stem region, which improves the thermostability (Hickey and Singer, 2004; Paz *et al*. 2004) as well by post-transcriptional modifications of tRNA that occur in thermophilic organisms (Kowalak *et al*., 1994; Shigi *et al*., 2002). On the other hand, there is no correlation between the G+C content of the genome and the optimal growth temperature of an organism. For instance, many thermophilic species, such as *Pyrococcus furious*, *Ignisphaera aggregans* and "*Aquifex aeolicus"*, have genomic G+C content of less than 50% while some mesophiles have much higher G+C contents in their genomes (Aptekmann and Nadra, 2018; Hickey and Singer, 2004).

Thermophilic proteins have prominent hydrophobic groups in the catalytic domains and increased number of charged groups for electrostatic (hydrogen bonds, ion-pairs) interactions to keep them functional at extreme thermophilic conditions (Kumar *et al*., 2018a; Kumar and Nussinov, 2001; Reed *et al*., 2013; Vielle and Zeikus, 2001). Furthermore, increased disulfide bridges lead to enhanced stability of the thermophilic proteins (Boutz *et al*., 2007; Cacciapuoti *et al*., 2012; Jorda and Yeates, 2011). The frequency of use of specific amino acids correlates with an optimal growth temperature of organisms and the nucleotides contents of their genomes (Klipcan *et al*., 2006; Vieira-Silva and Rocha, 2010; Zeldovich *et al*., 2007). An increase in purines (adenine

and guanine) in the genome of some thermophiles was reported as a possible primary adaptation mechanism (Lin and Forsdyke, 2006; Zeldovich *et al*., 2007). According to Sabath and collaborators (2013), the growth temperature and genome size in prokaryotes are negatively correlated, suggesting genomic streamlining during thermal adaptation. These authors propose the small genome size as an additional genomic signature of thermophilic adaptation. Examples of this phenomenon can be seen with the type strain of *Hyperthermus butylicus*, which has an T_{opt} between 95–107°C and a genome size of 1.67 Mbp (Brügger *et al*., 2007; Zillig *et al*., 1991), and the type strain of *Methanothermus fervidus* with a genome size of only 1.24 Mbp and a T_{opt} of 83ºC (Anderson *et al*., 2010). However, the strains of the species *Dehalococcoides mccartyi*, all mesophiles, have a genome size ranging from 1.34 to 1.47 Mbp (Löffler *et al*., 2013; Martínez-Cano *et al*., 2015). Therefore, there is no simple correlation between the T_{opt} and genomic features since thermophilicity results from a combination of assorted factors.

Thermophilic organisms produce specialized proteins, known as heat shock proteins, some of which function as molecular chaperones that allow the folding of enzymes into their native state, thereby helping retain their functionality in high temperatures (Conway de Macario and Macario, 2000; Fu, 2014; Zhang *et al*., 2015). The accumulation by thermophilic organisms of low molecular-mass organic compounds (compatible solutes/osmolytes) in the cytoplasm also supports a role in protecting cell components against thermal denaturation (Empadinhas and da Costa, 2011; Santos and da Costa, 2002).

As a summary, an outline of biochemical features of thermophiles heat stability is represented in Figure 1.4.

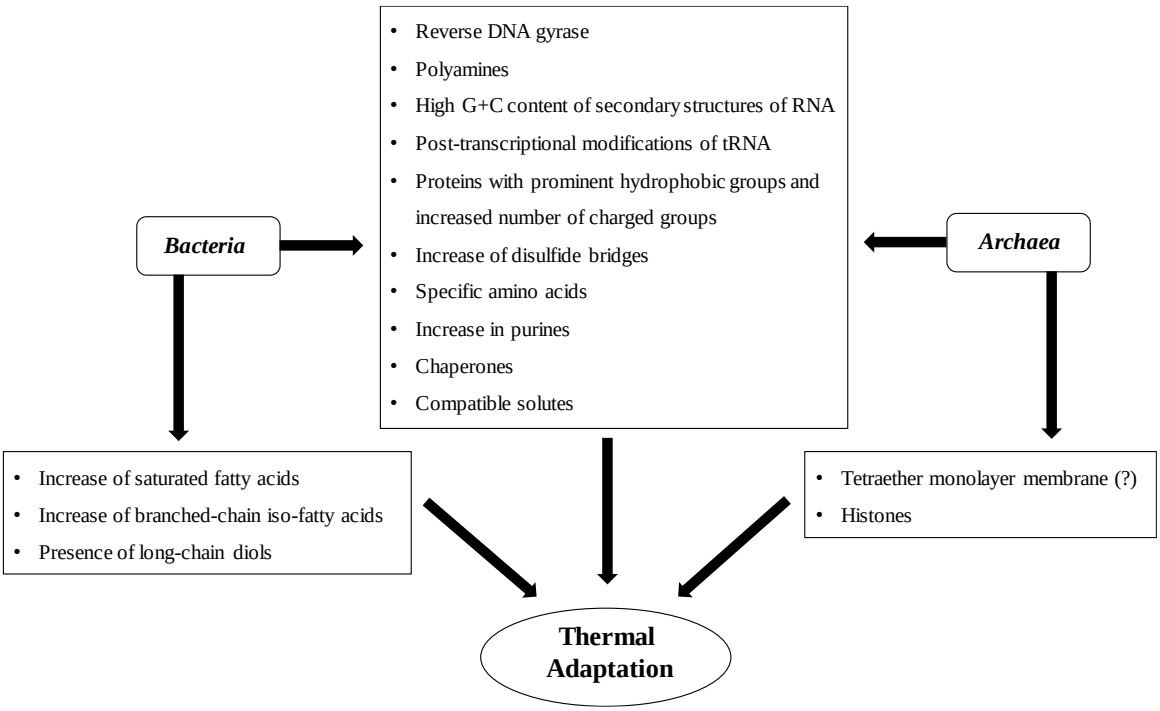

Figure 1.4 Outline of biochemical features of adaptation of *Bacteria* and *Archaea* to thermophilic conditions.

1.1.4 Biotechnological potential

There are many definitions for biotechnology, but the definition of Polish Ministry of Science and Higher Education, adapted after OECD (Organization for Economic Co-operation and Development) is enlightening: "*Biotechnology is the interdisciplinary branch of science and technology dealing with transformation of living and inanimate matter by the use of living organisms, their parts or products derived from them, as well as creation of models of biological processes in order to produce knowledge, goods and services*" (Kafarski, 2012).

Biotechnological processes frequently occur in inhospitable conditions to microorganisms. However, extremophiles are readily available to work under those conditions and are known producers of enzymes highly resistant to extreme conditions designated extremozymes (Dumorné *et al*., 2017; Elleuche *et al*., 2014). Thermophiles have shown tremendous promise in biotechnology applications with several advantages such as the high metabolic activity leading to enhanced product formation rates, reduced risk of contamination by mesophilic microorganisms, or production of heat-stable macromolecules and metabolites (Bergquits *et al*., 2014; Elleuche *et al*., 2015; Urbieta *et al*., 2015). Moreover, metabolic reactions occur at the same high temperature at which the substrates solubilize and no cooling steps are required after heating steps. High temperature also increases the bioavailability and solubility of numerous organic compounds leading to faster reaction rates and direct recovery of volatile products. The low bacterial mass formation yields higher ratios of desired product over assimilated substrate and lower waste production. Thermophiles are a relevant source of thermostable enzymes which possess the capacity to resist denaturation and proteolysis (Kumar and Nussinov, 2001).

Biotechnological applications of thermophiles can be divided between applications using whole cells (Table 1.1) and applications using their macromolecules or metabolites. Among the numerous biotechnological applications that use thermophiles, those concerning bioremediation strategies and clean production technologies are among the most remarkable. Thermophiles also have well known and potentially highly productive applications in bioenergy and biomining. The use of biomolecules produced by them as biosurfactants, osmolytes, and thermozymes has application in several biotechnology areas (Table 1.2). In recent years, the use of thermophiles in nanotechnology with nanoparticles biosynthesis has shown potential application in a wide spectrum of areas, including pharmacology, medical diagnostics, electronics and bioremediation (Beeler and Singh, 2016; Li *et al*., 2011; Moayad *et al*., 2017; Tiquia-Arashiro and Rodrigues, 2016). Exopolysaccharides (EPS) biosynthesis also has industrial and medical potential applications (Kambourova, 2018; Molina *et al*., 2013; Wang *et al*., 2021). The microbial polyhydroxyalkanoates (PHA) polyesters are considered to be a "green" alternative to fossil-based conventional plastics due to their biodegradability properties and renewable origin (Koller and Mukherjee, 2020; Obruca *et al*., 2021).

Table 1.1 Biotechnological applications of thermophiles using whole cells*.

Application	Organism	Action
Bioremediation		
	Geobacillus sp., *Anoxybacillus flavithermus*, *Thermus thermophilus*, *Thermococcus zilligii*	Biosorption of toxic metals
	Thermus scotoductus, *Pyrobaculum islandicum*, *Thermoanaerobacter* sp., *Carboxydothermus ferrireducens*	Immobilization of radionuclides
	Aeribacillus sp., *Geobacillus* sp.	Biodegradation of recalcitrant aromatic compounds and hydrocarbons
	Anoxybacillus sp.	Degradation of azo-dyes
Bioenergy		
	Caldicellulosiruptor bescii, *Caldanaerobius polysaccharolyticus*	Xylan degrading activity
	Thermoanaerobacterium thermosaccharolyticum	Biobutanol production
	Caldicellulosiruptor saccharolyticus	Hydrogen production
	Methanoculleus thermophilus	Methane production from coal mine substrates
Bioleaching		
	Sulfobacillus sp., *Ferroplasma* sp., *Acidianus infernus*	Copper extraction from chalcopyrite
	Acidianus brierleyi, *Acidianus manzaensis*, *Metallosphaera sedula*, *Sulfolobus metallicus*	Metal solubilization from nickel-copper sulphide

*Adapted from Urbieta *et al*. (2015).

1.1.4.1 Bioremediation / Bioenergy / Biomining

Bioremediation is a process to remove contaminants, pollutants, or unwanted substances from soil, water or air using living organisms, mainly microbes. Due to the increase in the costs of physical and chemical treatments, microbe-mediated eco-friendly treatments are getting more attractive, have reduced cost and eliminate or transform environmental organic or inorganic contaminants into benign products (Hazen and Tabak, 2005; Tabak *et al*., 2005). The use of thermophiles has become a promising alternative to treat metal-contaminated sites. Some thermophiles can tolerate high metal concentrations, which may increase metal solubilization through oxidation processes. Thermophilic microbial communities are also able to couple metal reduction with the oxidation of different organic and inorganic substrates (Sen *et al*., 2014). The thermophilic bacteria *Thermus*

thermophilus tolerates very high concentrations of arsenate, arsenite and cadmium (Antonucci *et al*., 2018; Del Giudice *et al*., 2013). The biodegradation of hydrocarbons by thermophiles can be used to remove organic compounds such as aliphatic and aromatic hydrocarbons (Margesin and Schinner, 2001; Mnif *et al*., 2014; Nzila, 2018). The thermophilic species *Aerobacillus pallidus* strain SL-1 efficiently degrades short-chain alkenes and aromatic hydrocarbon components of crude oil, being a good candidate for bioremediation of environments contaminated with polycyclic aromatic hydrocarbons (Tao *et al*., 2020). *Anoxybacillus* spp. produce several hydrolases and oxidoreductases, useful for the bioremediation of wastewater and phenol reduction, a constituent of many pollutants, and degrades azo dyes, an extremely hazardous compound widely used in the industry (Deive *et al*., 2010; Jardine *et al*., 2018).

The search for renewable energy sources, especially bioenergy, has become a topic of worldwide interest due to increased concerns over the decline of fossil fuel reserves and climate change. Biofuels are obtained from biomass (i.e., sugar cane, corn, beets, wheat) and from the biodegradable components of industrial, municipal and agricultural wastes (Barnard *et al*., 2010). Bioethanol, biodiesel, biobutanol and biokerosene (liquid biofuels) are obtained by fermenting materials such as starch and lignocellulosic biomass or by extraction of the lipid fraction from plants and microorganisms (Urbieta *et al*., 2015). Thermophiles produce thermozymes that efficiently degrade lignocellulosic biomass (i.e., cellulose, hemicellulose, xylan) for liquid biofuel production (Bhalla *et al*., 2013; Broeker *et al*., 2018; Han *et al*., 2012; Jiang *et al*., 2017; Patel *et al*., 2019; Peng *et al*., 2015; Su *et al*., 2013). Methane and hydrogen are examples of gas biofuels that can be obtained by the anaerobic fermentation of different feedstocks and waste materials. Several anaerobic thermophiles were described as capable of producing gas biofuels (Canganella and Wiegel, 2014).

Biomining comprises different biological processes in order to enhance the recovery of metals from ores. Bioleaching and bio-oxidation are two bio-extractive processes applied to sulfide minerals performed by microorganisms with the same mechanisms, however, during bioleaching, the metal is directly solubilized, while in bio-oxidation, microorganisms dissolve the mineral matrix that blocks the metal from being recovered, which can be later dissolved using other chemical leaching agents (Urbieta *et al*., 2015). In biomining, microorganisms create oxidizing and acidic conditions to release the metal to the acidic water solution as soluble sulfates (Donati *et al*., 2016). The main commercial application is copper bioleaching. Several studies have revealed that thermoacidophilic microorganisms can generate satisfactory copper recovery yields, much higher than those obtained with mesophilic microorganisms (Abdollahi *et al*., 2014; d'Hugues *et al*., 2002; Li *et al*., 2014; Safar *et al*., 2020; Qin *et al*., 2013). The thermoacidophilic archaeon species of the genus *Acidianus* are capable of metal extraction under highly extreme conditions (Safar *et al*., 2020; Wheaton *et al*., 2015).

1.1.4.2 Biosurfactants / Osmolytes / Thermozymes

Biosurfactants are amphiphilic compounds produced by microorganisms that help to increase the emulsification of hydrophobic compounds. Biosurfactants can be glycolipids, lipopolysaccharides, lipoproteins, fatty acids, phospholipids and neutral lipids. Biodegradability and low toxicity have led to the intensification of the use of biosurfactants in a wide range of industrial applications in the field of bioremediation as well as in petroleum, food processing, textile, detergent, pharmaceutics, cosmetics, agricultural applications and nanotechnology industries (Jimoh and Lin, 2019; Santos *et al*., 2016). Thermophiles have demonstrated potential for the production of biosurfactants with higher temperature stabilities and increased resistance to other extreme physicochemical parameters, such as pH and salinity. Thermophilic organisms with enhanced performance thus far identified have been related to the genera *Alcaligenes*, *Aneurinibacillus*, *Geobacillus*, *Brevibacillus* and *Bacillus* (Bharali *et al*., 2011; Joshi *et al*., 2008; Mehetre *et al*., 2019; Mnif *et al*., 2011; Sharafi *et al*., 2014).

Compatible solutes or osmolytes are another source of important biomolecules with various applications. One example is mannosylglycerate (MG), also called firoin, which is related to microbial adaptation to high temperature and osmoprotection. MG is accumulated in several thermophilic species of the genera *Pyrococcus, Thermococcus*, *Palaeococcus*, *Archaeoglobus*, *Aeropyrum*, *Stetteria*, *Rhodothermus*, *Thermus* and *Rubrobacter* (Borges *et al*., 2014). Preservation of protein native conformation and/or inhibition of protein aggregation seem pertinent targets for drug development for neurodegenerative disorders, such as Alzheimer's, Parkinson's, and Prion diseases. Several studies have shown that osmolytes produced by thermophiles prevent denaturation and aggregation of model proteins, among them, MG, effective in preventing amyloid formation (Faria *et al*., 2013; Jorge *et al*., 2016), act as stabilizer of enzymes against thermal stress and freeze-drying (Borges *et al*., 2002) or stabilizer of retroviral vaccines (Cruz *et al*., 2006) among other applications (Lentzen and Schwarz, 2006).

A large number of thermozymes have been characterized in the last decades, such as polysaccharide-degrading enzymes (amylases, pullulanases, xylanases, cellulases, hemicellulases, pectinases and chitinases), proteases, esterases, lipases, glucosidases, isomerases, hydrogenases, dehydrogenases, DNA polymerases among others, which have attracted great interest due to their potential for versatile applications in pharmaceutical/medical, chemical, textile, detergents, paper, leather, food and beverage industries, as well as in biofuels production (Table 1.2) (Akanbi *et al*., 2020; Allala *et al*., 2020; Antranikian *et al*., 2005; Bergquist *et al*., 2014; Egorova and Antranikian, 2005; Haki and Rakshit, 2003; Klippel and Antranikian, 2011; Kohli *et al*., 2020; Mehta *et al*., 2016; Sahoo *et al*., 2020; Urbieta *et al*., 2015). Thermostable DNA polymerases play a major role in various molecular biological applications, such as DNA amplification and sequencing. Taq polymerase from the thermophilic species *Thermus aquaticus* was the first thermostable DNA polymerase characterized (Chien *et al*., 1976) which is widely use, as well as DNA polymerases from hyperthermophilic species of the genera *Pyrococcus*, *Thermococcus* and *Thermotoga* (Ishino

and Ishino, 2014). However, the most extensively used thermozymes are the amylases in the food and beverage industry. Other thermozymes are used in food processing and preservation (Akanbi *et al*., 2020; Fernandes, 2010; Raveendran *et al*., 2018). Thermostable starch-hydrolysing enzymes have been characterized from several thermophilic species from *Bacteria* and *Archaea* (Elleuche and Antranikian, 2013). Recently an *α*-amilase of the thermophilic species *Tepidimonas fonticaldi* strain HB23 was purified and characterized with great potencial in the detergent industry (Allala *et al*., 2020). Species of the thermophilic genus *Thermus* produce several thermozymes with potencial industrial application. *Thermus thermophilus* produces enzymes that biodegrade proteins, polysaccharides or key enzymes that are involved in amino acid metabolism, protein folding or in other fundamental biological processes such as DNA replication, DNA repair, and RNA maturation, with potential use in different biotechnological processes (Pantazaki *et al*., 2002). Also, lipases and esterases produced by this species have properties that support their potential for biotechnological applications (Fuciños *et al*., 2005, 2012). However, the purification and characterization of an *α*-amilase from the species *Thermus filiformis* strain Ork A2 was the first report of this kind of enzyme purified from a strain of the genus *Thermus* (Egas *et al*., 1998). Strains of the species *Thermus thermophilus* were demonstrated to be an alternative cell factory for overproduction of thermophilic enzymes that fail to be expressed or were produced in lower amounts in the traditional mesophilic hosts (Aulitto *et al*., 2017; Cava *et al*., 2009; Hidalgo *et al*., 2004). *Meiothermus ruber* strain H328 and *Meiothermus taiwanensis* strain WR-220, moderately thermophiles, possess a strong keratinolytic activity leading to the complete degradation of feathers providing an eco-friendly way to convert keratin wastes (Kataoka *et al*., 2014; Wu *et al*., 2017). Several pullulanases were described from species of the genus *Geobacillus* as potentially valuable enzymes for starch and detergent industries (Ece *et al*., 2015; Nisha and Satyanarayana, 2015). Recently, several thermophilic isolates of the genera *Thermoanaerobacterium* exhibited significant cellulase and xylanase activity suggest promising applications of these thermoanaerobic bacteria (Harnvoravongchai *et al*., 2020).

1.1.4.3 Nanoparticles / Exopolysaccharides / Polyhydroxyalkanoates

The ability of some thermophiles to reduce heavy metal ions makes them good candidates for nanoparticle synthesis. The thermophilic species *Ureibacillus thermosphaericus* and *Geobacillus stearothermophilus* showed high potential for the biosynthesis of silver nanoparticles, and silver and gold nanoparticles, respectively (Fayaz *et al*., 2010b; Juibari *et al*. 2011). Silver nanoparticles are significant in pharmacology because of their antimicrobial properties and are thought to be one possible answer to the increasing antibiotic resistance of microbes (Beeler and Singh, 2016; Fayaz *et al*., 2010a). The thermophilic strain ID17 of the species *Geobacillus* has the ability to biosynthesize and accumulated gold nanoparticles, providing a potential applicability in immunostaining of specific molecules and also provides a potential application of this microorganism in bioremediation of gold-bearing waste (Correa-Llantén *et al*., 2013).

High molecular mass extracellular carbohydrate polymers, called EPSs, constitute part of the outer envelope of many prokaryotic microorganisms. EPSs production by *Streptococcus thermophilus*, a thermophilic organism widely used in the dairy industry, has been studied as it can improve the properties of the dairy product (Cui *et al*., 2017). Likewise, in the thermophilic species *Brevibacillus thermoruber* (Radchenkova *et al*., 2018) and in species of the genus *Geobacillus* the production and characterization of EPSs were investigated (Panosyan *et al*., 2018; Wang *et al*., 2021). EPSs from *Geobacillus* sp. strain WSUCF1 provide a valuable resource for utilization in biomedical fields such as drug delivery carriers (Wang *et al*., 2021).

Polyhydroxyalkanoates are microbial polyesters that are accumulated in the form of intracellular inclusions by several prokaryotic organisms. Apart from their primary carbon and energy storage function, PHA are also involved in the stress response of microorganisms (Obruca *et al*., 2021). The capability of PHA accumulation by species *Rubrobacter xylanophilus* and *Rubrobacter spartanus* (Kouřilová *et al*., 2021) make these thermophilic species good candidates for industrial production of PHA as "bioplastic" as an alternative to petrochemical polymers (Koller and Mukherjee, 2020; Obruca *et al*., 2021).

Table 1.2 Examples of thermozymes produced by thermophiles with known or potential biotechnological applications*.

Enzyme	Organism	Application / Industry
Amilase		
α-amilase		
	Pyrococcus woesei	Starch processing into glucose syrup
	Bacillus licheniformis	Clarification of fruit juice, bakery industry
	Geobacillus sp., *Anoxibacillus* sp.	Food industry
	Tepidimonas fonticaldi	Bio-additive in detergent formulations
	Thermococcus sp., *Anaeobranca gottschalkii*	Gelling, thickening, stabilizing agents in food industry
Glucoamylase		
	Thermoplasma acidophilum	Sugar industry and starch processing
	Thermoanaerobacter tengcongensis, *Picrophilus torridus*	Food industry
Pullulanase		
	Fervibacterium pennivorans, *Staphylothermus marinus*	Sugar industry and starch processing
	Geobacillus sp, *Thermococcus* sp., *Anaerobranca gottschalkii*, *Thermus thermophilus*, *Rhodothermus marinus*	Starch processing and detergent industry
	Thermotoga neapolitana	Biofuel production
Cellulase		
	Pyrococcus furiosus, *Acidothermus cellulolyticus*, *Rhodothermus marinus*	Clarification of fruit juice
	Thermomonospora sp.	Colour brightness improvement, superior cleaning without damaging fibers
	Clostridium thermocellum, *Acidothermus cellulolyticus*, *Thermobifida fusca*	Bioefuel industry

Table 1.2 (continued)

Xylanase		
	Pyrodictium abyssi, *Thermotoga maritima*, *Dictyoglomus thermophilum*	Bleaching of paper
	Caldicoprobacter algeriensis, *Anoxybacillus kaynarcensis*, *Roseithermus sacchariphilus*, *Bacillus licheniformis*, *Geobacillus* sp.	Food industry
	Acidothermus cellulolyticus, *Caldicellulosiruptor bescii*, *Caldanaerobius polysaccharolyticus*	Biofuel industry
Chitinase		
	Thermococcus kodakaraensis	Utilization of biomass of marine environment
	Sulfolobus tokadaii	Pharmaceutical industry
	Bacillus licheniformis, *Silanimonas lenta*, *Streptomyces roseolilacinus*	Agriculture industry and health products
Protease		
	Fervidobacterium pennivorans	Soaking in leather industry, feather degradation
	Meiothermus ruber	Feather degradation (decompose feathers of industrial waste)
	Anoxybacillus kamchatkensis, *Thermus aquaticus*, *Coprothermobacter proteolyticus*	Food industry
Esterase		
	Sulfolobus tokadaii	Biotransformation in organic solvents
	Ureibacillus thermosphaericus, *Pyrococcus furiosus*	Food industry
	Geobacillus sp., *Anoxyvbacillus* sp., *Alicyclobacillus acidocaldarius*, *Thermus thermophilus*, *Fervidobacterium nodosum*	Agriculture, food, detergent and pharmaceutical industries

Table 1.2 (continued)

Lipase		
	Bacillus pumilus	Treatment of palm oil-containing wastewater
	Geobacillus sp., *Bacillus licheniformis*	Food and pharmaceutical industries
***α*-glucosidase**		
	Thermococcus hydrothermalis	Starch processing into glucose syrup
***β*-glucosidase**		
	Alyciclobacilllus sp.	Conversion of soybean isoflavones in the feed industry
	Anoxybacillus flavithermus	Treatment of food industry wastes high in complex sugars
***β*-galactosidase**		
	Pyrococcus woesei	Production of milk with low lactose content
Glucose isomerase		
	Thermotoga maritima	Production of high-fructose corn syrup
Hydrogenase		
	Pyrococcus furiosus	Biohydrogen production
Alcohol dehydrogenase		
	Sulfolobus solfataricus	Reduction of ketones
DNA polymerase		
***Taq* polymerase**	*Thermus aquaticus*	PCR, DNA sequencing
***Pfu* polymerase**	*Pyrococcus furiosus*	PCR, DNA sequencing
***Pwo* polymerase**	*Pyrococcus woesei*	PCR, DNA sequencing

*Data from Akanbi *et al*. (2020), Allala *et al*. (2020), Antranikian *et al*. (2005), Kohli *et al*. (2020), Mehta *et al*. (2016), Sahoo *et al*. (2020) and Urbieta *et al*. (2015).

1.2 Prokaryotic Taxonomy

1.2.1 Definition and historical overview

The term taxonomy is often used synonymously with systematics; however, it is more appropriate to regard taxonomy as a part of systematics (Tindall *et al*., 2007). Systematics can be defined as the study of the diversity and relationships among organisms to characterize and arrange organisms in an orderly manner, while taxonomy is the theoretical and practical study of classification, including its bases, principles and rules (Kämpfer and Glaeser, 2013; Mayr, 1969; Rosselló-Móra and Amann, 2001; Trüper and Schleifer, 2006). Therefore, prokaryotic taxonomy is defined as the study of the classification of *Archaea* and *Bacteria* and is constituted of three independent but correlated areas: classification, that is, the organization of organisms into previously established groups (taxa) based on their phenotypic and genotypic similarities, and their phylogenetic and evolutionary relationship; nomenclature, the process of assigning a name to the taxa identified in the classification following the rules established by the International Code of Nomenclature of Prokaryotes (ICNP) (Oren *et al*., 2011b; Parker *et al*., 2019); and identification, which consists of determining whether an isolate belongs to a taxon already established in the classification and named in the nomenclature. Identification is the practical application on the foundation of classification and nomenclature; classification is often confused with identification, but classification is rather a requirement for identification. The binomial system of nomenclature, a combination of a generic and a specific name (specific epithet) in Latin, created by Carl Linnaeus, is used to this day with the recognition of species as the basic unit. Species are then organized in taxa of successively higher ranks (genus, family, order, class and phylum) (Kämpfer and Glaeser, 2013; Trüper and Schleifer, 2006). Assigning names can bring implications and assumptions to the organism, such as the pathogenic or biotechnological potential it harbors and the safety necessary for its handling; therefore, the act of assigning a taxonomic designation to an organism may have wide-reaching effects (Moore *et al*., 2010). Prokaryotic taxonomy should be predictive, the microorganism name should indicate some properties of the organism, should be universal, applicable to all kind of organisms of the discipline and finally pragmatic, with no need for users to deal with the theoretical issues (Rosselló-Móra and Amann, 2001; Rosselló-Móra, 2012; Rosselló-Móra and Whitman, 2019). The areas of taxonomy associated with the ICNP provide well-founded and stable guidelines for characterize and classify microorganisms, providing an efficient organizational system for dealing with the variety of cultured microbial diversity. The increase in the diversity of uncultivated *Archaea* and *Bacteria* and the advances in the cultivation-independent methods gave rise to the emergence of several proposals in recent years for a nomenclatural system for uncultivated taxa of *Archaea* and *Bacteria* (Chuvochina *et al*., 2019; Konstantinidis *et al*., 2017, 2020; Konstantinidis and Rosselló-Móra, 2015; Whitman, 2015, 2016). The priority of the names of uncultivated taxa to be recognized, and DNA genome sequence considered as the type material are the straightforward changes that these authors suggest, however, these changes are not consensual (Bisgaard *et al*., 2019; Overmann *et al*.,

2019). Oren and Garrity, 2018 were concerned with the proposal of some authors to create an independent nomenclature system, and stated: "*for the nomenclature of the prokaryotes, cultivated as well as uncultivated, we must choose between order to be established by the International Committee on Systematics of Prokaryotes (ICSP) (without involvement of another international microbiological society in charge of the nomenclature of the uncultivated taxa) or pluralism that will inevitably lead to chaos and to the destruction of now well ordered nomenclature system*". In march 2020, the ICSP discussed a proposal to use sequence data as type material for naming of prokaryotes, but this proposal was rejected (Sutcliffe *et al*., 2020). Soon after, a Consensus Statement provinding two alternatives plans (Plan A and Plan B) for the inclusion of uncultivated microorganisms into the classical Linnaean nomenclature system was proposed to clarify the way to effectively communicate microbial diversity (Murray *et al*., 2020). Plan A proposes the formal revision of the ICNP to include uncultivated organisms represented by DNA sequence information as the nomenclatural type (DNA sequence as type material). Plan B proposes the creation of a parallel code for uncultivated taxa, the 'Uncultivated Code'. Presently, the ICNP guidelines recommend designating these uncultivated microorganisms as *Candidatus*, a term first proposed by Murray and Schleifer in 1994 (Murray and Schleifer, 1994). Despite *Candidatus* has no standing in prokaryotic nomenclature (Hugenholtz *et al*., 2021), the information on *Candidatus* taxa is kept and updated by the Judicial Commission of the ICSP in cooperation with the Editorial Board of the International Journal of Systematic and Evolutionary Microbiology (IJSEM) and published in that journal (Oren *et al*., 2020a, 2020b; Oren and Garrity, 2021).

Although acknowledging the disdain that some scientists have for taxonomy, the American paleontologist, evolutionary biologist and historian of science, Stephen Jay Gould, frequently highlighted in his writings how classifications arising from a good taxonomy both reflect and direct our thinking, stating, "*the way we order reflects the way we think*" (Chung e*t al*., 2018). Taxonomy has been considered one of the most progressive scientific disciplines once the way to classify microorganisms has changed over time as new technological advances were introduced. In just two centuries, we have gone from classifying microorganisms based on their basic phenotypic characteristics to classifying them based on the complete sequence of their genomes. Since the first sequencing of a bacterial genome in 1995 (Fleischmann *et al*., 1995), together with the subsequent development of sequencing techniques, a profound change occurred in taxonomic practice by allowing access to the entire genomic content of a strain (Figure 1.5).

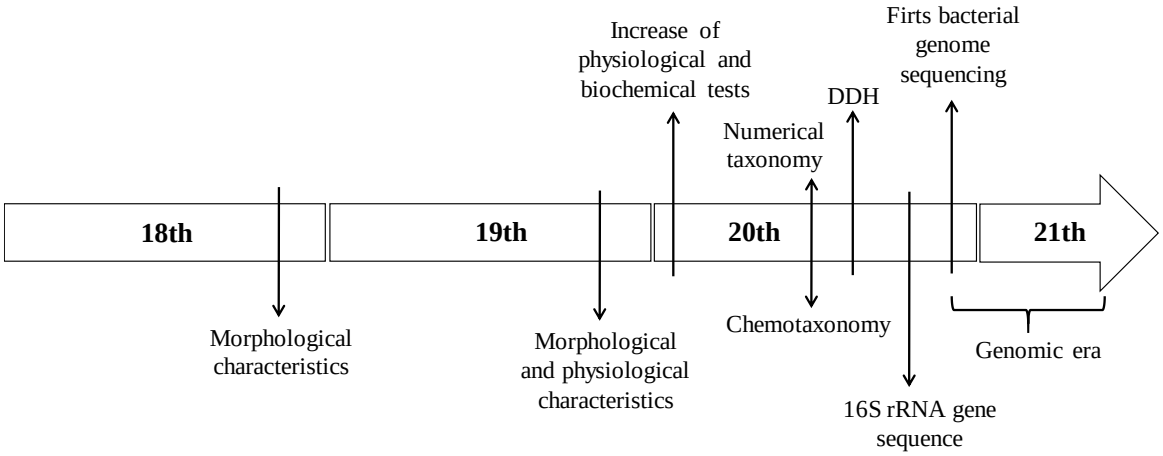

Figure 1.5 Evolution of prokaryotic taxonomy through the centuries.

The earliest effort to create microbial classifications was merely based on morphological observations made in the microscope. At the end of the 18th century, Otto Müller was the first to attempt a systematic arrangement of microorganisms; he described and named two genera, *Monas* and *Vibrio*, based on morphological characteristics. In the late 19th century, Ferdinand Cohn recognized a wide diversity of bacteria and classified them in six genera based on morphological characteristics but considered that physiologies and pathogenesis of similar-shaped organisms might differ. At the beginning of the 20th century, physiological and biochemical data were increasingly used and, in addition to morphology, became important markers for the classification and identification of microorganisms (Drews 2000; Rosselló-Móra and Amann, 2001; Schleifer, 2009; Trüper and Schleifer, 2006). In 1923, with its first edition, the *Bergey's Manual of Determinative Bacteriology*, together with the subsequent nine editions (last edition published in 1994), became the reference on bacterial taxonomy, providing essential support for microbiologists to unify the criteria used on classification and nomenclature (Guerrero, 2001). However, only in the 8th edition of the Manual, published in 1974, bacteria were no longer considered plants and were recognized members of the kingdom *Procaryotae* (Buchanan and Gibbons, 1974; Schleifer, 2009). Currently, the *Bergey's Manual of Systematics of Archaea and Bacteria* is the reference in the taxonomy of prokaryotes. In the middle of the 20th century, with the development of computing science, the introduction of numerical taxonomy enhanced phenotypic identification by increasing the number of tests used and calculating coefficients of phenetic similarities between strains and species, which allowed the comparison of large numbers of phenotypic traits for large numbers of strains (Sneath, 1957). This period coincided with the rise of chemotaxonomy that complemented

and improved the classification system (Rosselló-Móra and Amann, 2001). The increasing knowledge of the DNA properties and the development of molecular biological techniques in the early 1960s led to the introduction of the genotypic approach in the classification of microorganisms, initially including the mol% G+C compositions of DNA, and later DNA-DNA hybridization (DDH) that became the standard technique for the circumscription of bacterial species (Mandel, 1969; Wayne *et al*., 1987). In the 1980s, the emergence of amplification techniques and automatic sequencing, mainly of the 16S rRNA gene, allowed a significant advance in identifying microorganisms based on their phylogenetic relationships (Ludwig and Schleifer, 1994; Stackebrandt and Goebel, 1994). All these advances led to the use of a suitable system of classification of prokaryotes called the polyphasic approach (Vandamme *et al*., 1996).

1.2.2 Polyphasic approach

The polyphasic approach refers to a type of consensus taxonomy whose objective is to use all available phenotypic and genomic information to analyse diversity within and between taxa. The term polyphasic taxonomy was introduced by Colwell in 1970 (Colwell, 1970) and is still used today for the circumscription of taxa at all levels through a combination of as many different information as possible, including morphological, physiological, biochemical, chemotaxonomic, genomic, and phylogenetic characteristics (Adiguzel *et al*., 2020; Gevers *et al*., 2006; Gillis *et al*., 2015; Kämpfer and Glaeser, 2012; Prakash *et al*., 2007; Rainey, 2011; Vandamme *et al*., 1996), following the recommendations of the ICSP for the characterization of prokaryote strains (Chun *et al*., 2018; Stackebrandt *et al*., 2002; Tindall *et al*., 2010). Species are the taxonomic rank unit, however, there is no official definition of species in microbiology (Rosselló-Móra and Amann, 2001; Rosselló-Móra and Kämpfer, 2004). Among microbial taxonomists, species are described as a monophyletic group of organisms with high genomic and phenotypic homogeneity (Rosselló-Móra and Amann, 2001; 2015; Rosselló-Móra and Kämpfer, 2004). The species definition refers to the parameters used to embrace the unit, and that can change along time as it depends on the capability to observe characters and develops in parallel with technical advances (Rosselló-Móra and Amann, 2001, 2015; Stackebrandt *et al*., 2002). One of the most important premises to classify organisms is to have them in pure culture, and this means to isolate the organisms from their environment and grow them in the laboratory as pure culture (Tindall and Garrity, 2008; Tindall *et al*., 2010). Another essential aspect is the maintenance and long-term storage of the isolates. Generally, the organisms can be stored frozen at –80°C in glycerol without loss of viability for several years, the long-term preservation could also be done by freeze-drying or storage in liquid nitrogen (Morgan *et al*., 2006; Prakash *et al*., 2013). It is highly recommended, whenever possible, that species-level classifications should be based on the description of more than one strain to guarantee the universality of the characteristics measured within the taxon and the strain diversity within a species (Rosselló-Móra and Amann, 2015; Stackebrandt *et al*., 2002; Tindal *et al*., 2010). After the classification of a new species, it is mandatory to identify one of the strains as the type

strain, defined as living cultures of an organism that are descended from a strain designated as the nomenclatural type (Rosselló-Móra and Amann, 2001). The type strain should be the reference strain for other scientists for comparison studies, it is taxonomic common sense to include all type strains that are relevant to a study (Tindal *et al*., 2010). The type species of the genus is the most important reference organism to which a novel species has to be compared if it is considered to be a member of the same genus (Tindal *et al*., 2010).

1.2.2.1 Phenotypic methods

Phenotype alludes to how information encoded in the nucleotide sequence (genomic information) is expressed. It is the visible expression of the genotype, the observable or measurable characteristics that result from the expression of genes with the interaction of the environment. Since the beginning of prokaryotic taxonomy, phenotypic characteristics have been established to support microbial description and classification. The classical phenotypic analysis used in taxonomy comprises morphological, physiological and biochemical features of microorganisms. The morphology of a microorganism includes both cellular (Gram staining, shape, size, presence of flagella, the formation of endospores and inclusion bodies) and colonial (color, dimensions, form) characteristics. The physiological and biochemical features include temperature, pH and salt concentrations growth range, atmospheric conditions (aerobic or anaerobic requirements), growth in the presence of different substances, enzyme activities, metabolism of compounds, pathogenicity, resistance or sensitivities to antimicrobials, among others (Caumette *et al*., 2015; Oren *et al*., 2011a; Vandamme *et al*., 1996). Microbial identification systems based on classical methods are commercially available, such as API (bioMérieux) and Biolog (Biolog Inc.), developed for clinical microbiology; thus, they should be applied cautiously to samples that are not of clinical origin (Rosselló-Móra and Amann, 2001; Tindall *et al*., 2010).

The phenotypic analysis also integrates the study of the chemotaxonomic characteristics of the different chemical constituents that comprise the structural components of the prokaryotic cell, including the outer cell layers (peptidoglycan, teichoic acids, mycolic acids), the cell membrane (polar lipids, respiratory lipoquinones, fatty acids, pigments) or constituents of the cytoplasm (polyamines) (Tindal *et al*., 2010; Vandamme *et al*., 1996). The cell wall composition is generally used for the classification of Gram-positive bacteria that can have peptidoglycan with different structures and with different amino acid compositions. The peptidoglycan composition can be specific to genus or species (Chen *et al*., 2020; Schleifer and Kandler, 1972; Schumann, 2011). The structural diversity of teichoic acids can also be used as a taxonomic marker of Gram-positive bacteria (Potekhina *et al*., 2011) and mycolic acids are useful for the classification of members of the high G+C Gram-positive bacteria, specifically for the identification of *Mycobacterium* species (Yassin, 2011). The composition of polar lipids, isoprenoid quinones and fatty acids are generally analysed by chromatographic methods and are used for discriminating among taxa (Albuquerque *et al*., 2014; França *et al*., 2015; Lage *et al*., 2017). Polar lipids are analysed by thin-layer

chromatography (TLC) and specific staining (da Costa *et al*., 2011a). Isoprenoid quinones are found in most prokaryotes with an important function in electron transport. The more common respiratory lipoquinones found in prokaryotes are menaquinones (naphthoquinone) and ubiquinones (benzoquinones); the large variability of their side chains can be examined by high performance liquid chromatography (HPLC) and used to characterize organisms at different taxonomic levels (da Costa *et al*., 2011b). The determination of the fatty acid methyl esters (FAMEs) composition can be used to distinguish closely related species with similar phenotypic characteristics, such as in the case of most *Legionella* spp. (Diogo *et al*., 1999). FAMEs composition is assessed by gas chromatography (GC) and their identification made by comparison of the peak retention times of samples with those of known standards of Sherlock Microbial Identification System (MIS) (Microbial ID Inc., MIDI) database. However, bacteria have to be cultivated under standardized growth conditions since fatty acid composition may vary with the growth temperature, medium composition, and growth phase (da Costa *et al*., 2011c). Fatty acyl compounds not identified by MIS can be identified by comparison with FAMEs from other bacteria where they have been identified (Albuquerque *et al*., 2014). Unknown fatty acids that are relevant for the characterizaction of a organism need to be identified by mass spectroscopy (MS). For example, a new family of internally branched iso-fatty acids were identified by gas chromatography-mass spectroscopy (GC/MS) during the description of the species *Gaiella occulta*, the only cultured representative of the order *Gaiellales* a deep branching lineage of the phylum *Actinobacteria* (Albuquerque *et al*., 2011a; Albuquerque *et al*., 2018b). Colonies of many prokaryotes display a variety of colours that can be easily visually assessed, but for some groups of prokaryotes like phototrophic bacteria, the identification and detailed characterization of the pigments produced is necessary for the classification of the organism (Oren, 2011a). The polyamines are found in most prokaryotes in a wide concentration range and their pattern can be discriminative for taxa above the rank of genus (Busse, 2011; Busse and Auling, 1988; Hosoya and Hamana, 2004).

Techniques such as serotyping, electrophoretic profiles (whole-cell protein profiles, lipopolysaccharide profiles, multilocus enzyme electrophoresis), and spectroscopy (Fourier-Transform Infrared Spectroscopy, UV Resonance Raman Spectroscopy) provide unique patterns that can be useful for identification and discrimination purposes between strains (Caierão *et al*., 2016; Gaus *et al*., 2006; Kersters *et al*., 1994; Veríssimo *et al*., 1996; Vogt *et al*., 2019; Wattiau *et al*., 2011). The phenotype typing methods produce single-strain fingerprints useful for establishing relationships within a given taxon at the species level but lack discriminative power in higher taxa (Rosselló-Móra and Kämpfer, 2004).

Methods that rely on mass spectrometry analyse of different cellular fractions can be used as high-throughput phenotypic methods. For example, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) is capable of detecting large molecules (proteomics) (Wiser *et al*., 2012). In standard conditions used for species identification, the size range detects mainly ribosomal proteins (Karlsson *et al*., 2015; Munoz *et al*., 2011; Rosselló-Móra, 2012; Seuylemezian *et al*., 2018). High-field ion cyclotron Fourier transform mass spectrometry

(ICR-FT MS) aims at the detection and comparison of small molecules (metabolomics) and has been used to access metabolic adaptation of microorganisms to environmental variations (Rosselló-Móra *et al*., 2008; Rosselló-Móra, 2012). Databases of MALDI-TOF MS spectra of known bacterial species are available for microbial identification and diagnostic purposes (Dridi and Drancourt, 2011). However, identification of new isolates is possible only if the spectral database contains peptide mass fingerprints of the type strains of the related species (Singhal *et al*., 2015).

Standardized methods should be used and assessed in parallel with reference strains to minimize phenotypic data reproducibility and comparability problems.

1.2.2.2 Genotypic methods

Genomic information is derived from all the data retrieved from nucleic acids (DNA and RNA) present in the cell. The first nucleic acid-based technique applied to taxonomy was the determination of the base ratio of a DNA molecule, defined as the relative abundance of the pair G+C, commonly called G+C content and expressed in mol% (De Ley, 1970; Marmur *et al*., 1963). The classical methods to determine genomic G+C content are buoyant density centrifugation (B.d.), thermal denaturation (T_m), fluorometric determination of melting temperature and HPLC, the latter being the most used (Mesbah *et al*., 2011). In prokaryotes, the G+C content varies between 17 and 75 mol%, and organisms that have more than 10 mol% difference in DNA G+C content may not belong to the same genus, and a range of 3–5 mol% is found within a species (Mesbah *et al*., 2011; Rosselló-Móra and Amann, 2001). Although differences in the percentage of G+C of the DNA are taxonomically useful for separating taxa, similar DNA base compositions do not necessarily imply close relationships since it does not provide information on the linear sequences of bases in the DNA (Rosselló-Móra and Amann, 2001). Currently, the DNA G+C content can be calculated directly from the genome sequence of the organism and the result is the percentage (%) of the number of the guanine and cytosine bases over the number of total bases observed in the genome (Meier-Kolthoff *et al*., 2014). According to Meier-Kolthoff *et al*. (2014) when the G+C content was inferred from genome sequences, within species differences are almost exclusively below 1%. Despite the values of the G+C content of DNA assessed from genomic sequences show a greater precision than the values obtained by indirect methods, several studies showed small differences (< 2%) on the G+C values of several bacteria and archaea, when determined by classical methods and predicted from the genome (Albuquerque *et al*., 2016; Mesbah *et al*., 2011), corroborating the accuracy of the indirect methods used for determination of DNA content, as long as the experimental conditions are standardized and replicated. However, species descriptions have to be emended when there are discrepancies between data in the literature and the values of the G+C content based on genome sequence reported for the same type strain (Palaniappan *et al*., 2013).

In 1968, Johnson and Ordal developed a method to measure the degree of genetic relationship of two organisms based on the ability of nucleic acids to reassociate or hybridize, once denatured and under standardized conditions. This method allows the DNA of two different organisms to

hybridize based on the similarity of their nucleotide sequences, and this only occurs if the overall DNA base composition is similar and if the organisms are genetically related (McCarthy and Bolton, 1963; Schildkraut *et al*., 1961). The technique that determines the whole genome DNA-DNA similarity between two organisms is called DDH or DNA-DNA reassociation. Several methods were developed to determine DDH values. All have in common the measurement of the extent and/or stability of the hybrid double-stranded DNA resulting from a denatured mixture of DNAs incubated under stringent conditions, which allow only the renaturation of complementary sequences. Depending on the method used, there are two main parameters that can be determined: the relative binding ratio (RBR) and the increment of melting temperature (ΔTm) (Rosselló-Móra, 2006; Rosselló-Móra *et al*., 2011). In prokaryotic taxonomy, the DDH methods have concentrated mainly on the use of RBR expressed as % of similarity. With this technique, phenotypically coherent microorganisms could be regarded as a single species if they shared high DDH values, in general above 70%. An ad hoc committee recommended this cutoff value as an approximate threshold for circumscribing species, where values greater than 70% of RBR or 5ºC or less of ΔTm indicate a relationship at the species level (Wayne *et al*., 1987). However, this value must be evaluated within a diversity of parameters that need to show genomic and phenotypic consistency (Wayne *et al*., 1987). Furthermore, a strain that is indistinguishable phenotypically from an established species but with a low DDH value between them cannot be formally named as a new species and is considered a genomic species or genomovar (Richer and Rosselló-Móra, 2009; Ursing *et al*., 1995; Wayne *et al*., 1987). An ilustrative example is strain FR-6^T (DSM 11984^T) isolated from solfataric soil in the Azores that possesses a DDH of 53.3% with the type strain of *Alicyclobacillus acidocaldarius*, however, the biochemical, physiological and chemotaxonomic characteristics of the strain FR-6^T are indistinguishable from those of the type strain of *A. acidocaldarius*, being described as *Alycyclobacillus* genomic species (Albuquerque *et al*., 2000). Even though the diversity of methods developed, DDH protocols are considered laborious and time-consuming, difficult to implement and standardize between laboratories, and few laboratories are equipped to apply this methodology (Sentausa and Fournier, 2013). Despite the limitations, DDH is a methodology that can be applied to all cultivable prokaryotes regardless of their growth requirements and provides a unified measure for the circumscription of bacterial and archaeal species, and was considered during nearly 50 years the gold standard for prokaryotic species circumscriptions (Rosselló-Móra and Amann, 2001; Richer and Rosselló-Móra, 2009). Advances in sequencing technologies and the availability of a large number of genome sequences opened the door to *in silico* genome-to-genome comparison enabling the production of digital DDH (dDDH) values. Its results have proven a good correlation to the 70% threshold of DDH (Auch *et al*., 2010; Mahato *et al*., 2017; Meier-Kolthoff *et al*., 2013).

The increasing knowledge of the DNA molecule and the development of PCR techniques (Saiki *et al*., 1988) led to the emergence of new nucleic acid-based methods (van Steenbergen *et al*., 1993) that started to be intensely used in the taxonomy of microorganisms. DNA-based typing methods or DNA fingerprinting methods can reveal the diversity of close relative strains, but these methods

are only applicable to understand intraspecific diversity and not proper for the circumscription of prokaryotic species as well as for higher taxonomic taxa. Some of the techniques used are macrorestriction analysis after pulsed-field gel electrophoresis (PFGE), ribotyping, amplified fragment length polymorphism (AFLP), amplified ribosomal DNA restriction analysis (ARDRA), randomly amplified polymorphic DNA (RAPD), repetitive element sequenced-based PCR (rep-PCR), among others (Kämpfer and Glaeser, 2012; Tindall *et al*., 2010). These techniques depend on the electrophoretic separation and succeeding visualization of DNA fragments (Khosravi and Dolatabad, 2020; Lopez-Canovas *et al*., 2019; Neoh *et al*., 2019; Zare *et al*., 2019). However, it was the introduction of the analysis of the 16S rRNA gene by cataloguing (Fox *et al*., 1977) and the development of PCR-based sequencing techniques (Sanger *et al*., 1977; Böttger, 1989) that made a revolution in the history of rRNA sequence analysis. Later, the development of automated DNA sequencing technology led to a rapid increase in the number of descriptions of novel taxa.

The genes encoding rRNA proved to be very useful phylogenetic markers, since they are universal, composed of highly conserved as well as variable domains, functionally stable and contain considerable genetic information (Ludwig and Schleifer, 1994; Patwardhan *et al*., 2014). Furthermore, assuming that lateral gene transfer (LGT) does not occur or occurs poorly between rRNA genes, the variations in the rRNA primary structures among the prokaryotes will reflect evolutionary distances between organisms (Ludwig and Schleifer, 1994; Ramasamy *et al*., 2014; Rosselló-Móra and Amann, 2001; Schleifer, 2009). Moreover, rRNA genes are evolving more slowly than protein coding genes and are particularly important for the phylogenetic analysis of distantly related species (Patwardhan *et al*., 2014; Yarza *et al*., 2014). Among the three rRNA molecules present in prokaryotes, the 16S rRNA gene, due to its size of approximately 1,500 bp, was chosen as the universal marker for phylogenetic analysis, and the phylogenetic reconstructions allowed a more objective classification system among prokaryotes (Olsen *et al*., 1994; Woese, 1987). The ad hoc committee for evaluation of species definition has recommended that all species descriptions should include an almost complete 16S rRNA gene sequence with the respective accession number of the GenBank/EMBL/DDJJ database, a publicly accessible database (Stackebrandt *et al*., 2002; Tindall *et al*., 2010), and a certificate of the mandatory deposit of the type strain in two international culture collections (Tindall, 2008; Tindall and Garrity, 2008; Tindall *et al*., 2010). Unlike DDH, once a 16S rRNA gene sequence is obtained from an isolate, it can be compared, using BLAST (Altschul *et al*., 1990), against available repositories dedicated to universal 16S sequences from prokaryotes that hold carefully curated 16S rRNA gene sequences, such as the Ribosomal Database Project (RDP) (Cole *et al*., 2009), the Greengenes (GG) (DeSantis *et al*., 2006), the EzBioCloud (Chun *et al*., 2007; Yoon *et al*., 2017a) and the Living Tree Project (LTP) (Yarza *et al*., 2008; Ludwig *et al*., 2021) compatible with the rRNA databases of SILVA (Quast *et al*., 2013; Yilmaz *et al*., 2013). The taxonomic rank information of EzBioCloud and LTP are based on the List of Prokaryotic Names with Standing in Nomenclature (LPSN) (Parte, 2014, 2018) and regularly updated with 16S rRNA gene sequences of type strains of species with validly published names (Ludwig *et al*., 2021). Therefore, sequencing and phylogenetic analysis of the 16S

rRNA gene has been considered a standard method for the classification of prokaryotes at various taxonomic levels (Ludwig and Klenk, 2015; Ludwig *et al*., 2011; Tindall *et al*., 2010). The 16S rRNA gene has become the most sequenced taxonomic marker and considered the cornerstone for the prokaryotic taxonomy (Yarza *et al*., 2014) and is still a necessary marker for taxonomic purposes (Ludwig *et al*., 2021).

There are many different algorithms available for calculating similarity between two gene sequences, however for obtaining nucleotide similarity values for taxonomic purposes, it is necessary to carry out a pairwise sequence alignment using a multiple alignment program and then calculate the similarity value; pairwise similarity values obtained from local alignment programs, such as BLAST and FASTA, should not be used (Tindall *et al*., 2010). In the comparative analysis of the 16S rRNA gene sequence, it is assumed that those prokaryotes with a similarity of less than 97% should be considered as members of different species since such differences were empirically correlated with values lower than 70% of DDH (Stackebrandt and Goebel, 1994). This cutoff value at the species level was reviewed and increased first to 98.7% (Stackebrandt and Ebers, 2006) and later to 98.65% (Kim *et al*., 2014). The cutoff value at the genus level was established at 95% similarity (Tindall *et al*., 2010; Yarza *et al*., 2008). Regrettably, the conservative nature of the 16S rRNA gene did not show enough resolution to distinguish between closely related species (Beye *et al*., 2018; Schleifer, 2009). Another limiting factor is the possibility of the gene being acquired by LGT, according to Tian *et al*. (2015). Lateral gene transfer of 16S rRNA genes can occur at low rate between closely related organisms (Konstantinidis and Tiedje, 2007; Tian *et al*., 2015). Additionally, the possibility of the presence of multiple 16S rRNA genes in a bacterial genome, usually with identical copies, but occasionally with nucleotide variations, may also limit the resolving power of the 16S rRNA gene for phylogenetic reconstruction (Beye *et al*., 2018; Pei *et al*., 2010; Ramasamy *et al*., 2014; Rossi-Tamisier *et al*., 2015).

Other highly conserved single-copy genes have been proposed as genetic markers as an alternative to the 16S rRNA gene to complement the DDH data for taxonomic analysis at the species level (Tindall *et al*., 2010). The conserved protein coding genes, housekeeping genes that can be used are the genes for the GroEL chaperonin (*gro*EL), RNA polymerase beta-subunit (*rpo*B), DNA gyrase beta-subunit (*gyr*B), the heat shock protein (*dna*K), among others (Rajendhran and Gunasekaran, 2011). The phylogenetic analysis of concatenated sequences of the several housekeeping genes is defined as multilocus sequence analysis (MLSA), a term introduced by Gevers *et al*., 2005, and is a method that can elucidate phylogenetic relationships within species and above species level (Glaeser and Kämpfer, 2015). The use of multiple genes circumvents the possible effects of genetic recombination or LGT that distort phylogenies based on single genes (Gevers *et al*., 2005; Schleifer, 2009). Furthermore, the MLSA can clarify the distinction between closely related species, in which the sequence analysis of the 16S rRNA gene shows insufficient resolution, since protein coding genes, in contrast to the 16S rRNA gene, are supposed to evolve at a slow (although faster than 16S rRNA genes) but constant rate (Glaeser and Kämpfer, 2015; Rong and Huang, 2014). The different ways MLSA is performed can vary greatly for the selection of

genes that should be ubiquitous in the taxon under study, their number, and the calculation method used when comparing the sequences obtained (Glaeser and Kämpfer, 2015). For instance, a threshold of 97% similarity in the MLSA study of four housekeeping genes (16S rRNA, *gyrB*, *rpoB*, and *rpoD* genes) is correlated with ANIb values for species differentiation in the genus *Pseudomonas* (Gomila *et al*., 2015; Mulet *et al*., 2010). The increase in the availability of genome sequence data led to the development of large-scale MLSA studies based in *in silico* analyses of the sequences of a greater number of housekeeping genes (Gupta and Sharma, 2015; Lang *et al*., 2013; Rong and Huang, 2014).

Bearing in mind the above exposure, it is noticeable that the traditional polyphasic approach in which multidimensional aspects of the organisms are considered, including phenotypic and genotypic traits, is essential for prokaryotic taxonomy. In the last years, the use of genome sequence and its comparison in microbial taxonomy became feasible due to the increase in the number of sequenced bacterial and archaeal genomes, thus introducing an additional layer to the polyphasic approach. Several authors have expressed the need to address the use of genome sequences in the prokaryotic taxonomy since it provides a reproducible, consistent and highly informative method to infer phylogenetic relationships among prokaryotes (Chun and Rainey, 2014; Kim *et al*., 2014; Ramasamy *et al*., 2014; Richter and Rosselló-Móra, 2009; Rosselló-Móra and Amann, 2015; Sangal *et al*., 2016; Sutcliffe, 2015; Sentausa and Fournier, 2013; Thompson *et al*., 2015; Zhi *et al*., 2012). The main journals of taxonomic descriptions now demand the inclusion of the high-quality genome sequences of at least the type strain of the novel prokaryote taxa, which led Chun and collaborators to propose the minimal standards for the use of genome data for the taxonomy of prokaryotes (Chun *et al*., 2018), detailed in the next section. Nowadays, prokaryotic taxonomy cannot be based solely on the traditional polyphasic approach but requires approaches for the integration of genomic information for the description of new taxa and their classifications (Caputo *et al*., 2019; Goh *et al*., 2020; Lalucat *et al*., 2020; Liu *et al*., 2018b; Viver *et al*., 2018; Xu *et al*., 2019).

As a summary, an outline of the steps and processes involved in the polyphasic characterization of a novel prokaryote is represented in Figure 1.6.

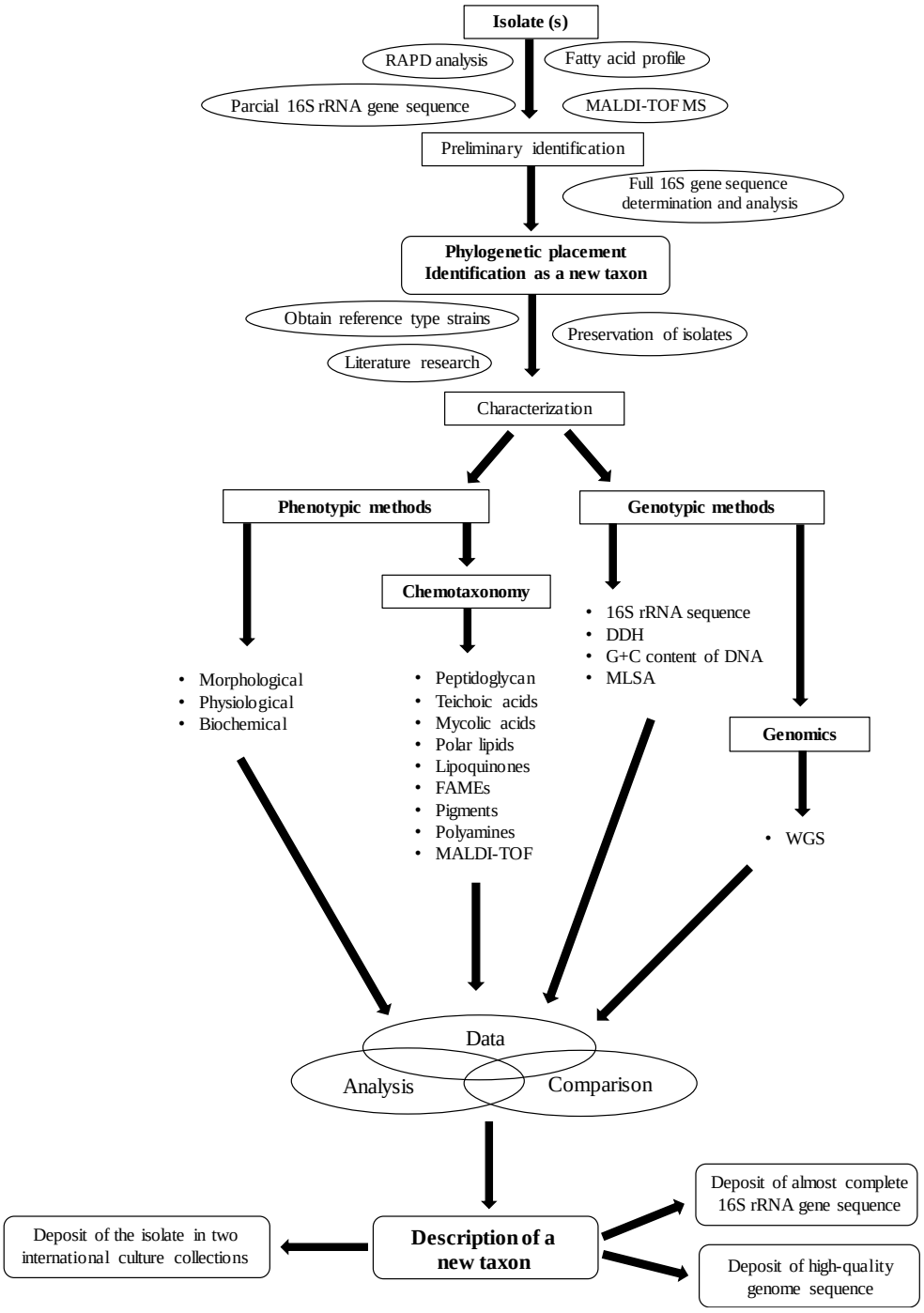

Figure 1.6 Outline of the steps and processes involved in the polyphasic characterization of a novel prokaryote (adapted from Rainey, 2011 and Ramasamy *et al*., 2014). In practice, the taxonomic characterization of a collection of isolates starts with a screening that allows the more closely related isolates to be clustered and to be distinguished from unrelated isolates; often-used screening methods are illustrated; after screening 16S rRNA gene sequence analysis is performed on representatives of the different clusters, and these sequences are compared with those of known species. The characterization of the new taxon should integrate all the different kinds of data and information as possible.

1.2.3 Genomic era

In the decade following the pioneering sequencing of the genome of *Haemophilus influenza* in 1995 (Fleischmann *et al*., 1995), sequencing of a bacterial or archaeal genome was carried out by the high-cost, laborious and time-consuming process of the conventional Sanger sequencing technique (Sanger and Coulson, 1975). Only in 2005, when high-throughput DNA sequencing technology known as next-generation sequencing (NGS) replaced Sanger sequencing, it was possible to achieve a rapid and automated sequencing method that overcame the disadvantages imposed by the traditional method. The first NGS platform widely used in microbiology was the Roche 454 sequencing system, which adopted the principle of pyrosequencing (Margulies *et al*., 2005), and was followed by other NGS platforms (Borriss *et al*., 2011). However, an important limitation continued to exist, its high cost. For this reason, the true revolution in genomic sequencing only occurred with the decrease in the costs of NGS technologies combined with advances in bioinformatics (Chun and Rainey, 2014; Sangal *et al*., 2014; Soon *et al*., 2013), which promoted an exponential increase in the number of prokaryotic genomes sequenced and deposited in public databases (Wu *et al*., 2009; Garrity, 2016; Vernikos *et al*., 2015). The increase in the whole genome sequence (WGS) data meant great progress in genomic studies. The information from WGS can be used to establish a solid base for the identification and classification of prokaryotes species, even populations, and clarify the evolutionary relationships between the different taxa and predict the metabolic, structural, functional potential of the different microorganisms (Liu *et al*., 2018b; Thompson *et al*., 2013, 2015; Viver *et al*., 2018). The WGS led prokaryotic taxonomy into the genomic era, with the possibility of establishing systematics based on information retrieved from complete genomes complemented with the phenotypic methods (Raina *et al*., 2019; Xu *et al*., 2019). This "new" polyphasic strategy that includes phenotypic as well as genomic information obtained from the WGS has been designated taxono-genomics, a term proposed by Ramasamy *et al*. (2014). Currently, this taxono-genomic approach is the one that is strongly recommended in prokaryotic taxonomy (Chun *et al*., 2018).

1.2.3.1 Genome assembly

In the last decade, NGS platforms were commercially introduced and proved to provide adequate genome data with quality requirements for taxonomic purposes (Goodwin *et al*., 2016). In 2008, Field and collaborators (Field *et al*., 2008) introduced the minimum information about a genome sequence (MIGS) that specifies a formal way to describe genomes in detail. In 2018, Chun and collaborators (Chun *et al*., 2018) established several statistical parameters to describe the quality of the final genome assembly: the genome size, defined as the sum of the length of all contigs; the number of contigs and N50, defined as the length of the shortest contig that accumulatively show 50% or more of the genome size; sequencing depth of coverage ≥50X (i.e., each base in the final assembly was read in 50 times on average); and the DNA G+C ratio. It was also established that,

for a description of a new species, a full-length 16S sequence of the type strain should be obtained by the Sanger sequencing and compared with the 16S sequence extracted from the whole genome assembly to ensure the authenticity of genome data. Housekeeping genes can also be used to support the authenticity of the final genome assembly. Contamination in the genome assembly should be checked using bioinformatic tools (Lee *et al*., 2017; Parks *et al*., 2015). The final genome assembly should be deposited in GenBank/EMBL/DDJJ database (Chun *et al*., 2018).

1.2.3.2 New genome-based parameters

With high-throughput sequencing, huge amounts of high-quality genomic sequences can easily be obtained, and together with bioinformatics tools, new methods emerged based on the comparison of genomic sequences that allow the delineation of bacterial and archaeal species (Tanaka *et al*., 2018; Teng *et al*., 2016). Different parameters emerged, reminiscent of the DDH, called overall genome relatedness index (OGRI), a term first coined by Chun and Rainey (2014) that represents any measurements indicating similarity or distance between two genomes, without gene-finding and functional annotation of predicted genes, providing a fast and reproducible way of comparing two genomes. These new parameters utilize whole genome sequences instead of individual gene sequences or a set of sequences. They refer to digital genomic relatedness or *in silico* genomic relatedness that uses the entire genome sequence to calculate the degree of relationship between two genomes.

There are several digital genomic relatedness indices to calculate OGRI values, but the most widely used for classification and identification of bacteria and archaea is the average nucleotide identity (ANI) (Arahal, 2014; Beaz-Hidalgo *et al*., 2015; Chun and Rainey, 2014; Ciufo *et al*., 2018; Rosselló-Móra and Amann, 2015). ANI is a genomic similarity index that represents the average of identity values between multiple sets of orthologous regions shared by two genomes, between the query genome and the reference genome, using BLAST alignments for genome comparisons (Altschul *et al*., 1997). This method that finds the shared orthologous protein coding genes between two genomes was proposed by Konstantinidis and Tiedje (2005a) as a robust measure of evolutionary distance, strongly correlated with DDH values and with the 16S rRNA gene sequence similarity (Kim *et al*., 2014). ANI values of 95–96% correspond to values of the threshold for differentiating two species comparable to a DDH value of 70% and a 16S rRNA gene similarity of 98.65% (Kim *et al*., 2014). A variation to the original method was introduced by Goris *et al*. (2007), which involves the *in silico* segmentation of the query genome into consecutive fragments of 1,020 nucleotides to simulate the fragmentation of genomic DNA that occurs during the DDH trials. These fragments are then used to search against the reference genome using the BLASTn algorithm (ANIb). Richter and Rosselló-Móra (2009) reported that the MUMmer algorithm (ANIm) is more efficient for comparisons of large DNA sequences, using a data structure named suffix tree to calculate alignments (Kurtz *et al*., 2004), that performs fast genome alignment without losing precision. However, ANIb is more widely used than ANIm, since there is little correlation for

distant genome comparisons (Li *et al*., 2015; Rosselló-Móra and Amann, 2015; Yoon *et al*., 2017b). Richter and Rosselló-Móra (2009) also suggested that ANI could be used as an alternative to DDH for species circumscriptions, overcoming the inconveniences of the traditional technique, and corroborate that ANI values of 95–96% equate to the value of 70% DDH. Likewise, Tindall *et al*. (2010) also proposed in a taxonomic note on the characterization of prokaryotic strains that the ANI index could substitute the DDH analyses. The average nucleotide identity by orthology algorithm (OrthoANI) (Lee *et al*., 2016) has been introduced as an alternative to the ANIb index, being increasingly used in taxonomic studies (Corral *et al*., 2018, de la Haba *et al*., 2019; Diop *et al*., 2020; Riesco *et al*., 2018). This improved algorithm solves the problem of reciprocal inconsistency of the original ANI algorithm and correlates well with ANIb (Lee *et al*., 2016). OrthoANI can be calculated with two algorithms, the BLASTn algorithm (OrthoANIb) or the USEARCH algorithm (OrthoANIu). Furthermore, this new method also reduces computational time as it does not require reciprocal calculations (Yoon *et al*., 2017b). Recently, Jain and collaborators developed FastANI, a new method to estimate ANI using alignment-free approximate sequence mapping that proved to be faster when compared with alignment-based approaches and providing identical ANI values (Jain *et al*., 2018).

Unlike ANI, which is a similarity-type index, dDDH is a distance-type index that uses the genome-to-genome distance (Auch *et al.*, 2010). The genome-to-genome distance calculator tool (GCDC) is based on the genome blast distance phylogeny (GBDP) algorithm, which calculates intergenomic distances (Henz *et al*., 2005). This algorithm locally aligns the two genomic sequences with each other, using alignment tools such as BLAST to obtain sets of high-scoring segments pairs (HSPs) that will be converted into distance values, dDDH values. Like DDH, the cutoff limit for the circumscription of prokaryotic species with the dDDH is 70% (Mahato *et al*., 2017; Chun and Rainey, 2014; Meier-Kolthoff *et al*., 2013). Consequently, the dDDH has been successfully applied in the description of novel species and the elucidation of evolutionary relationships between closely related species (Feng *et al*., 2019; Liu *et al*., 2015, 2017). Generally, dDDH is widely used to corroborate ANI results of the closely related genomes, but both fail in determining more distant relationships (Colston *et al*., 2014; Gomila *et al*., 2015; Sant'Anna *et al*., 2017).

The maximally unique matches index (MUMi) is another distance-type index based on DNA maximal unique matches (MUMs) shared by two genomes (Deloger *et al*., 2009). The method was developed to estimate the distance between closely related bacterial genomes rapidly. Like ANIm, MUMi uses the MUMmer algorithm for a faster pairwise comparison of the genomic sequences and has shown a good correlation with ANI values (Deloger *et al*., 2009). MUMi values vary from 0 for very similar genomes to 1 for distant genomes (Matsumoto *et al*., 2013). Some studies used MUMi over ANI when comparing subspecies due to its higher robustness on intraspecies differentiation (Ang *et al*., 2016; Tan *et al*., 2017).

Tetranucleotide signature regression (TETRA) is an alignment-independent parameter based on the differences in the frequency of the occurrence of the four nucleotides between two genomes (Richter and Rosselló-Móra, 2009; Teeling *et al*., 2004). This parameter identifies genomes at the

species level, but only highly similar genomes with regression values above 0.999 will correspond to ANI values of > 94% (Rosselló-Móra and Amann, 2015; Tambong, 2019). This very fast calculation method streamlines the screening of very large sets of genomes (Rosselló-Móra and Amann, 2015). Other alignment-independent parameter is the codon usage bias that refers to the difference in the frequency of occurrence of synonymous codons in coding DNA. Codon usage bias creates a pattern by selecting specific codons for an amino acid over others and this pattern can be specific for each genome (Lal *et al*., 2016; Mahato *et al*., 2017). It is possible to generate and compare the codon usage bias even in closely related organisms in the form of codon usage bias tables (Athey *et al*., 2017; Alexaki *et al*., 2019). Brbić and collaborators considered the possibility that evolution of codon usage bias within gene families may be predictive of microbial phenotypes and that the overall pattern of codon adaptation across many genes of an organism can predict its phenotype (Brbić *et al*., 2016).

Bioinformatic tools for calculating the several digital genomic relatedness indexes are available as web-services or as standalone software (Chun *et al*., 2018; Hugenholtz *et al*., 2021; Ludwig *et al*., 2021; Sant'Anna *et al*., 2019). The general procedure for genome-based species circumscription is summarized in Figure 1.7.

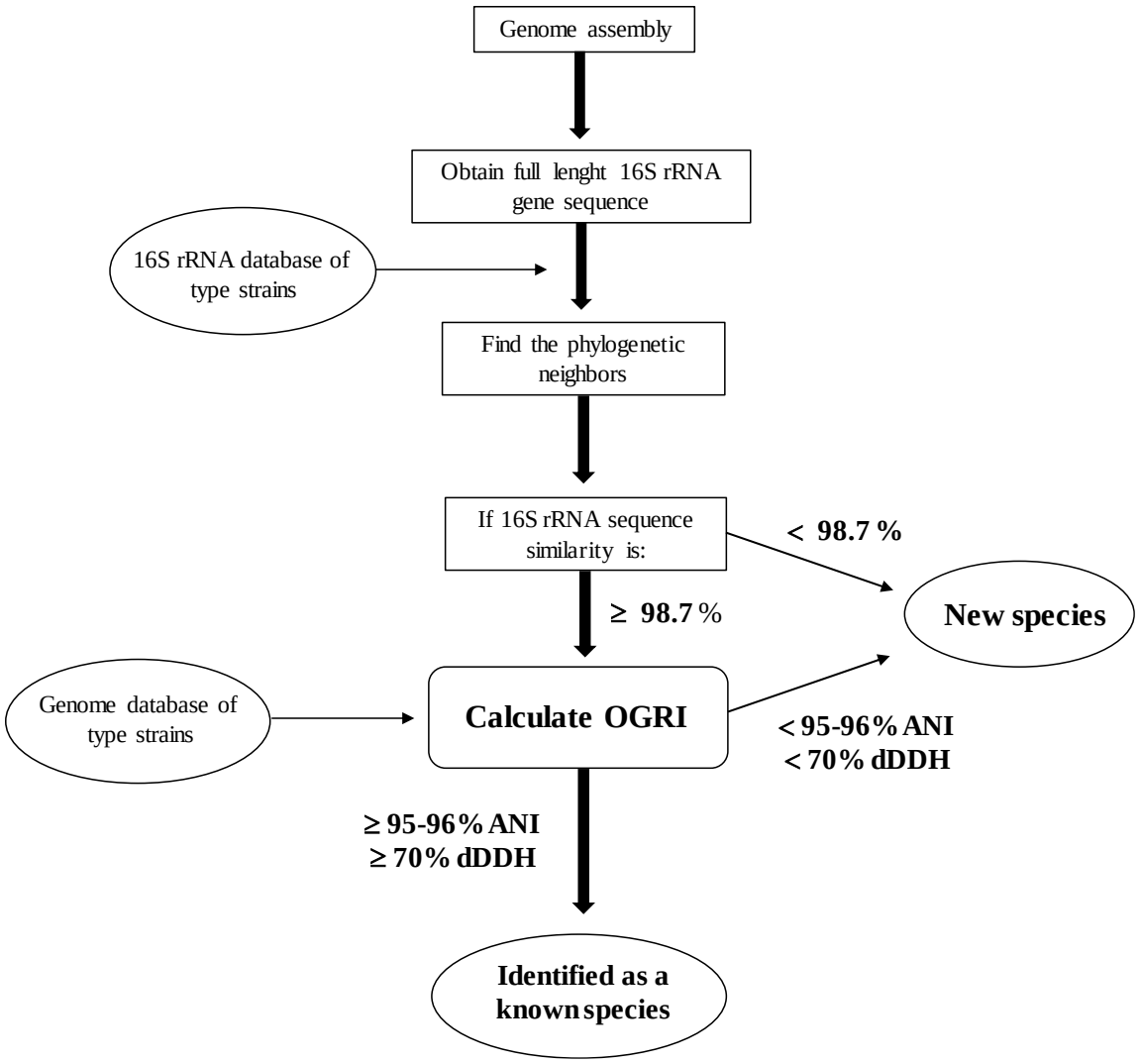

Figure 1.7 Workflow of genome-based classification at the species level (adapted from Chun *et al*., 2018).

These several OGRIs were proposed and developed for species delineation, however, they do not have a taxonomic resolution above the species level. Prokaryotic genera and higher ranks of the taxonomy remained for several years defined based on the sequence of the 16S rRNA gene. A new prokaryotic genus was proposed if it formed a monophyletic group in the phylogenetic analysis with an average divergence of less than 6% with respect to the sequence of the 16S rRNA gene between its closest neighbors and if, in addition, it had distinguished phenotypic characteristics from the closest genera (França *et al*., 2015; Sangal *et al.*, 2016; Yarza *et al*., 2014). Some attempts have been made to define generic boundaries between prokaryotes using genomic sequences. Konstantinidis and Tiedje (2005b) proposed the average amino acid identity (AAI) index for the

circumscription of higher taxa than species, showing that there is a strong correlation between 16S rRNA gene identity and AAI. This index is the equivalent of ANI but based on amino acid sequences rather than nucleotide sequences. Due to the nature of the constraints of the macromolecules, the ANI and AAI comparisons provide two levels of relatedness, ANI indicates close relationships while AAI more distant relationships, which makes the latter being more suitable and offering better resolution in the comparison of distant genomes (Cabal *et al*., 2018; Nicholson *et al*., 2020; Rosselló-Móra, 2005). Luo *et al*. (2014) demonstrated that AAI values between members of related but different genera typically vary between 60–80% and do not exceed 85%. On the other hand, Qin *et al*. (2014) suggested the percentage of conserved proteins (POCP) between genomes as a new index to estimate the evolutionary and phenotypic distance between two microorganisms, indicating POCP as a genomic index for establishing the genus boundary for prokaryotic groups. POCP values are obtained by aligning the query genome against the reference genome through the BLASTp algorithm. POCP values can vary from 0–100% depending on the similarity of the protein contents of the two genomes. Qin *et al*. (2014) also proposed that a prokaryotic genus can be defined as a group of species with POCP values higher than 50%, establishing this value as a cutoff for the delimitation of prokaryotic genera. POCP has been applied in several taxonomic studies (Maejima *et al*., 2020; Margos *et al*., 2018; Pérez-Cataluña *et al*., 2018; Ying *et al*., 2019). More recently, the core-gene average amino acid identity (cAAI) has also been suggested for the delimitation of prokaryotic genera (Wirth and Whitman, 2018). This parameter is determined by calculating the average similarity of the protein sequences of each of the orthologous genes found in the core-genome of a pair of organisms. The studies by Wirth and Whitman (2018) support that cAAI values correlated with the criteria proposed by Luo *et al*. (2014) for the conventional AAI, however in the same study, they concluded that the 50% POCP boundary could not be applied to the *Roseobacter* group suggesting that a cutoff defined by a single value is unlikely to be a universal threshold for delimiting prokaryotic genera.

As a summary, an outline of the new genome-based methods for taxonomic classification of prokaryotes is represented in Figure 1.8.

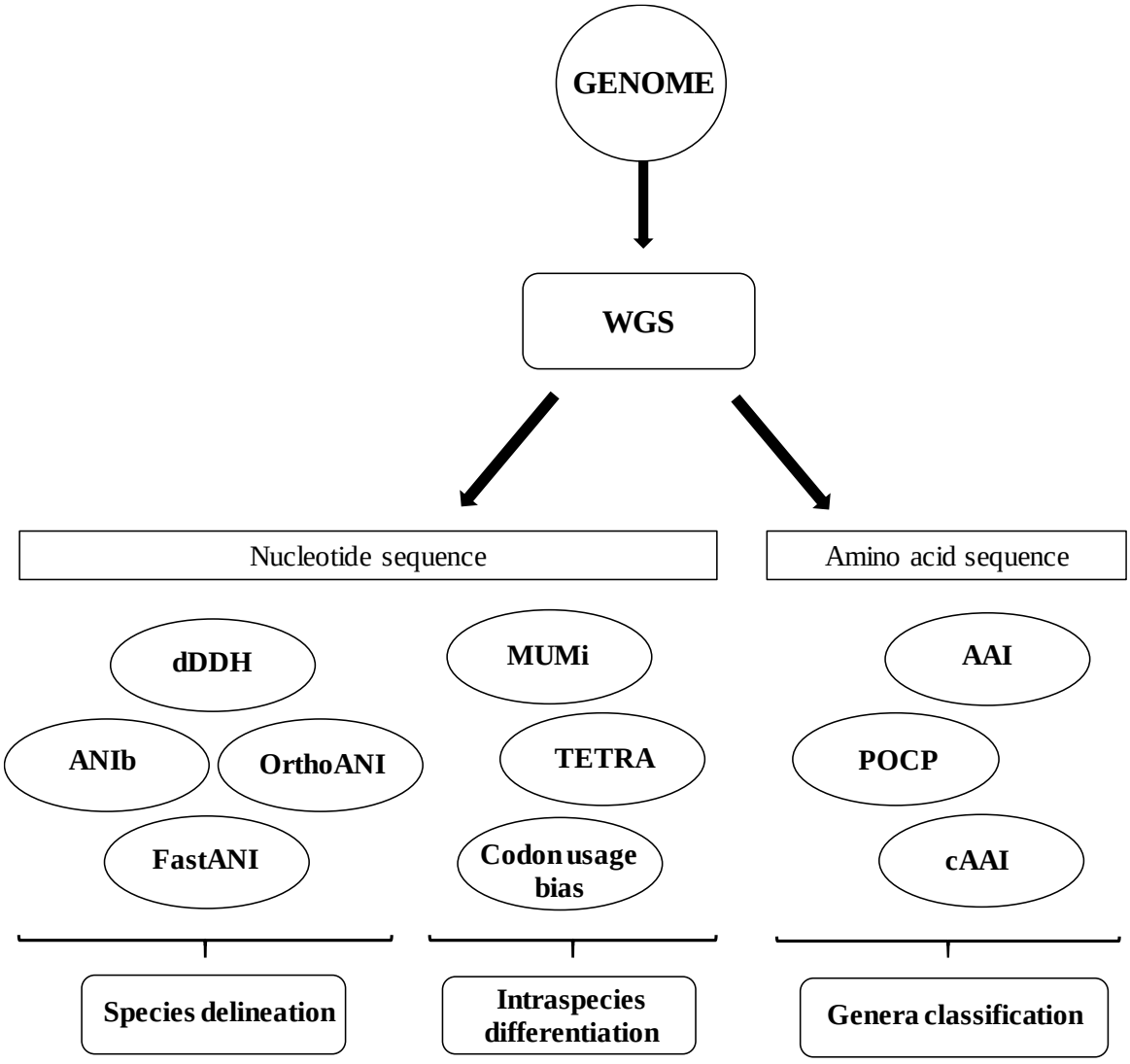

Figure 1.8 New genome-based methods for taxonomic classification of prokaryotes.

1.2.3.3 Phylogenomics

Phylogenomics aims to infer information about the evolutionary histories of organisms by using whole genomes rather than just a single gene or a few genes. Inferring whole-genome phylogeny can be accessed using multiple orthologous genes since they are unlikely to undergo lateral transfer events (Chun *et al*., 2018; Patané *et al*., 2018; Setubal and Stadler, 2018). Chun *et al*. (2018) proposed that a multigene-based phylogenomic treeing approach, which consists of phylogenetic analysis using multiple genes retrieved from the genome data, should be used to define genera or higher taxa, complementing the already established 16S rRNA gene phylogeny (Patel and Gupta, 2020; Salam *et al*., 2020). The number of chosen genes varies depending on the taxonomic scope of the study and on the algorithm used to select orthologous genes (Chun *et al*., 2018). This

approach can be distinguished from the MLSA method because of the higher number of orthologous genes selected using bioinformatics tools in the comparative genomic analysis (Hahnke *et al*., 2016; Munoz *et al*., 2016; Segata *et al*., 2013; Wu and Scott, 2012). Phylogenomic analyses can be based in the core-proteome alignments, where a concatenated alignment is constructed using the amino acid sequences encoded by shared genes between the organisms (de la Haba *et al*., 2019; Viver *et al*., 2018; Tettelin *et al*., 2005) or can be based in the core-genome alignments, where the alignment is constructed using the nucleotide sequences (Chung *et al*., 2018; de la Haba *et al*., 2019). Chung *et al*. 2018 proposed a workflow for assigning genus and species based on the length and sequence identity of the core-genome alignments: the length of the core-genome alignment at ≥ 10% is used as a cutoff at genus level, and the sequence identity of the core genome alignment at ≥ 96.8% indicates similarity at species level. With this approach, these authors reorganized the taxonomy of the order *Rickettiales* within the class *Alphaproteobacteria* where the species definitions in the families are inconsistent (Chung *et al*., 2018). The *Rickettsiaceae* taxonomy had already been subject to a reorganization in 2001 based in phylogenetic analysis using 16S rRNA and *groESL* gene sequences (Dumler *et al*., 2001).

The phylogenetic trees could be reconstructed using a distance-based algorithm, like the neighbour joining (NJ) method, or a character-based algorithm, like maximum parsimony (MP), maximum likelihood (ML) or bayesian inference (BI) methods (Patané *et al*., 2018). Combining phylogenomic treeing and the highly conserved phenotypes, including chemotaxonomic markers, has proved to be an excellent approach for the classification of genera and higher taxa, providing improved taxonomic studies (Hahnke *et al*., 2016; Infante-Domínguez *et al*., 2020; Liu *et al*., 2018b; Sangal *et al*., 2016; Viver *et al*., 2018; Xu *et al*., 2019).

The pan-genome is defined as the summation of all gene sets of the genome and is composed of both the core-genome, i.e., the genes that are present in all members of the taxon, and the variable- or accessory-genome content (Mahato *et al*., 2017; Tettelin *et al*., 2008). Therefore, according to its definition, the pan-genome analysis also includes the study of those genes that are not shared by all the members of the studied taxon and that are ignored by the traditional phylogenetic analysis since these approaches are based on the core-genome. The study of the pan-genome provides an additional resource to taxonomy since it increases the content of genetic information analysed, thus allowing further elucidation on the evolutionary relationships between different species (Ding *et al*., 2018; Marschall, 2018; Viver *et al*., 2018). Furthermore, it has been reported that the estimation of the content of the accessory genome content can be significant for delineating closely related species (Caputo *et al*., 2015, 2019; Méric *et al*., 2014). This method was applied to different species and subspecies of the genus *Klebsiella* by Caputo *et al*., 2015. These authors compared the core/pan-genome ratio of *Klebsiella* spp. and found that some subspecies exhibit as many differences between them as with other species of the genus, and should be considered distinct species of the genus *Klebsiela* (Caputo *et al*., 2015).

The analysis of the degree of synteny is another approach to examine the phylogenetic relationships between microorganisms. Synteny compares the order of arrangement of genes on a

chromosome or a plasmid among different genomes (Mahato *et al*., 2017; Snir, 2016). It is generally accepted that closely related genomes will have a similar genetic arrangement, which means that less synteny is observed as the taxonomic distance increases (Viver *et al*., 2018). Therefore, it is assumed that closely related organisms will present an arrangement and orientation of genes similar to each other, and, consequently, they will share a greater degree of synteny (Salazar and Abeel, 2018). Synteny analysis is very useful in the classification and differentiation among closely related species and can help infer the evolutionary relationships between microorganisms (de la Haba *et al*., 2019; Viver *et al*., 2018; Garcia and Gola, 2016; Stewart *et al*., 2015). As an example, the complete genome synteny analysis of two strains, JH146^{T} and strain FS406-22, of the hyperthermophilic methanogen species *Methanocaldococcus bathoardescens* showed 97% genome synteny, showing that these strains are closely related (Stewart *et al*., 2015). Likewise, synteny analysis between the two phylotypes, EHB-1 and EHB-2, of the extremely halophilic bacterium *Salinibacter ruber*, showed a conserved gene order of 84% (Antón *et al*., 2002, 2008; Peňa *et al*., 2010; Viver *et al*., 2018). Although analysis of synteny among closely related species is now widely used in new published genomes, this analysis is regularly performed on assembled sequences that are fragmented, ignoring the fact that most synteny methods were developed using complete genomes (Liu *et al*., 2018a).

1.2.3.4 Functional genomics

Functional genomics is one of the areas of study in genomics that deals with the identification of genes and their products and attempts to establish their biological functions and their interactions in different metabolic pathways. Although at first the study of functional genomics does not have any direct link with prokaryotic taxonomy, the fact is that it plays a significant role in understanding the evolutionary relationships between microorganisms since obtaining the functional profile of the genomes of the microorganisms analysed can be useful to make comparisons between different species through the study of metabolic pathways (Coenye *et al*., 2005; Mahato *et al*., 2017). This approach of inferring the phenotype based on tracing the gene content with which it is possible to predict the phenotypic traits of each of the genomes analysed is known as *in silico* phenotyping (Amaral *et al*., 2014; Weimann *et al*., 2016). The prediction of physiological traits based on screening of the presence of specific genes involved in molecular pathways can provide insight into the biology of microorganisms (Brbić *et al*., 2016; Rosselló-Móra and Amann, 2015; Thompson *et al*., 2015). As already mentioned, for taxonomic purposes, the genomes must be assembled and minimally annotated following several parameters to obtain a high-quality genome assembly (Chun *et al*., 2018). The genome annotation process consists of two steps: the gene finding process that aims to predict the sections of the genomes containing genes and the function assignment (Setubal *et al*., 2018). Several software tools can be used for genome assembly and prediction of the protein coding sequences (CDS), i.e., predicting the open reading frames (ORFs) and corresponding functional annotation. The KEGG Automated Annotation Server (KAAS) provides functional

annotation of genes using the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Moriya *et al*., 2007). KEGG is an integrated database resource that establishes links from genes in the genome to high-level functions of the cell and the organism. The genomes are annotated with the KEGG orthology (KO) database, a KO identifier (K number) is assigned as a functional ortholog defined from experimental characterized genes and proteins in specific organisms, which are then used to assign orthologous genes in other organisms based on sequence similarity. With the KO identifier, it is possible to reconstruct the KEGG pathway maps enabling interpretation of high-level functions in the BRITE hierarchies and KEGG modules (Kanehisa *et al*., 2016; Moriya *et al*., 2007). The clusters of orthologous groups of proteins (COGs) database is another functional annotation tool that uses a family-based approach which uses the functions of the characterized members of the protein family to assign a functional category to an identified ortholog (Galperin *et al*., 2015, 2019; Tatusov *et al*., 2001). Amino acid sequences of the genome can also be annotated by comparison to other databases; the SwissProt Protein Knowledgebase (Apweiler *et al*., 2004), a curated protein sequence database that provides a high level of annotation with a minimal level of redundancy; HAMAP (Pedruzzi *et al*., 2015), a collection of manually curated family profiles for protein classification; TIGRFAMs (Haft *et al*., 2003) and Pfam (Finn *et al*., 2016), databases of protein families with multiple sequence alignments and Hidden Markov Models (HMMs). Functional genomic analyses to infer phylogenetic relationships show divergent functional profiles between taxa (Chai *et al*., 2014). This correlation can be used to identify specific cellular functions to some taxa and could be helpful in the characterization of new organisms (Mahato *et al*., 2017). Despite the advances of assuming the inference of the phenotype based on the genome, it is still necessary for this inference to be validated by experimental phenotypic tests since the presence of a group of genes in a bacterial genome does not necessarily mean that the organism will present that phenotype (Amaral *et al*., 2014).

1.3 Objectives and Thesis Outline

The study of extremophilic environments and the microorganisms that inhabit them is motivated by several reasons, namely the scientific knowledge of the diversity of life in extreme environments, the molecular mechanisms of the adaptations that microorganisms use to thrive in these environments and the potential uses of these molecular mechanisms in biotechnological applications. This work was designed to address the first motivation to extend the knowledge of the microbial diversity of the São Pedro do Sul hot spring, with the isolation, characterization and description of microorganisms from this thermophilic environment. Several isolates collected in this thermal spring were grown in the lab and those deemed to represent new taxa were characterized through a polyphasic approach, comprising the morphological, biochemical, physiological, chemotaxonomic, genomic and phylogenetic characterization for taxonomic and systematics positioning. The technological advances in genomics, with an increase in the number of sequenced bacterial genomes available in public databases and the lower cost of small genome sequencing, incited the use of genome sequence data as an additional taxonomic tool. Genome sequences of the new isolates and comparative genomics with genomes of closely related strains were applied for taxonomic purposes combined with phenotypic data to improve the classification.

The main objectives of this thesis were:

- classify, identify and assign a name to new thermophilic organisms isolated from São Pedro do Sul hot spring;
- perform high-quality draft genome sequencing of fifteen species of bacteria;
- obtain insights into the metabolism of the microorganisms through the analysis of genome sequences and correlate the genotype with phenotypic characteristics;
- use comparative genome analysis through the overall genome relatedness index (OGRI) to classify new isolates;
- apply genome data for phylogenetic analysis (phylogenomics);
- revise the classification of the genus *Meiothermus* using comparative genomics and phylogenomics combined with phenotypic characteristics and chemotaxonomic markers.

Chapter 1 gives an overview of the definition of thermophiles, their habitats and their biotechnological potential. This chapter also includes an overview of the prokaryotic taxonomy, emphasizing the polyphasic approach and the new tools that the genomic era brought into taxonomy.

In **chapter 2**, a slightly thermophilic organism recovered from the São Pedro do Sul hot spring is described. Analysis of the 16S rRNA gene sequence indicated the new isolate was a novel cultured lineage within the order *Cytophagales* of the phylum *Bateroidetes*. The high-quality draft genome was sequenced and analysed, complementing the characterization of the new organism. Phenotypic characteristics, like assimilation of sugars and reduction of nitrate, were corroborated

by genome analysis. Based on physiological, biochemical, chemotaxonomic, genomic and phylogenetic characterization a new genus *Raineya* and a new species *Raineya orbicola* were described. A new family *Raineyaceae* was also described to accommodate this new genus and species. Results of this chapter were published in the International Journal of Systematic and Evolutionary Microbiology (doi.org/10.1099/ijsem.0.002556) and in the Bergey's Manual of Systematics of *Archaea* and *Bacteria* (doi.org/10.1002/9781118960608.fbm00370; doi.org/10.1002/9781118960608.gbm01902).

In **chapter 3**, high-quality draft genomes of eight type strains of the genus *Tepidimonas*, classified in the class *Betaproteobacteria*, were sequenced, and the available genomes of three closely related strains were examined. The genome sequence analyses of these organisms clarified the probable reasons why *Tepidimonas taiwanensis* is the only species of the genus *Tepidimonas* able to grow on hexoses (glucose and fructose). Comparative genome sequence analyses using average nucleotide identity (ANIb), digital DNA–DNA hybridization (dDDH), average amino acid identity (AAI) and phylogenetic analysis based on 16S rRNA gene sequence and on 400 conserved genes contributed to complement the classification of the organisms. A new slightly thermophilic organism isolated from São Pedro do Sul hot spring was additionally described as a new species of the genus *Tepidimonas*, *T. charontis*, based on phenotypic, chemotaxonomic, genomic and phylogenetic analysis. Results of this chapter were published in the International Journal of Systematic and Evolutionary Microbiology (doi.org/10.1099/ijsem.0.003942; doi.org/10.1099/ijsem.0.004563) and in the Bergey's Manual of Systematics of *Archaea* and *Bacteria* (doi.org/10.1002/9781118960608.gbm00959.pub2).

In **chapter 4**, four species of the genus *Meiothermus* of the family *Thermaceae*, a distinct lineage within the phylum *Deinococcus-Thermus*, were reclassified in a new genus, named *Calidithermus*. The high-quality draft genome of six type strains of the genus *Meiothermus* was sequenced, and genome sequences of seven type strains were retrieved from the databases. Pathways for the red-pigment or yellow-pigment synthesis were examined in all genomes to correlate to species colony colour. The comparative genome sequences analyses of the thirteen type species of the genus *Meiothermus* using the average nucleotide identity (ANIb), amino acid identity (AAI), the phylogenetic analysis of the 16S rRNA gene sequence, 90 housekeeping genes and 855 core-genes, and the phenotypic and chemotaxonomic characteristics were used to circumscribe the genus *Meiothermus* to eight species, which led to emend description of the genus *Meiothermus* and reclassification of four species of the genus *Meiothermus* as members of the novel genus *Calidithermus*. Results of this chapter were published in the International Journal of Systematic and Evolutionary Microbiology (doi.org/10.1099/ijsem.0.003270).

In **chapter 5**, the concluding remarks of this thesis are presented along with future perspectives supported by this work.

Chapter 2

Raineya orbicola - a New Genus and a New Thermophilic Species of the Novel Family *Raineyaceae*

Results published in:

ALBUQUERQUE, L., POLÓNIA, A.R.M., BARROSO, C., FROUFE, H.J.C., LAGE, O., LOBO-DA CUNHA, A., EGAS, C. and DA COSTA, M.S. (2018a). *Raineya orbicola* gen. nov., sp. nov., a slightly thermophilic bacterium of the phylum Bacteroidetes and the description of *Raineyaceae* fam. nov. *Int J Syst Evol Microbiol* 68: 982–989. doi.org/10.1099/ijsem.0.002556

ALBUQUERQUE, L. and EGAS, C. (2020). *Raineyaceae*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.fbm00370

ALBUQUERQUE, L. and EGAS, C. (2021a). *Raineya*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.gbm01902

2.1 Abstract

An isolate designated SPSPC-11^T, with an optimum growth temperature of about 50ºC and an optimum pH for growth between 7.5 and 8.0, was recovered from a hot spring at São Pedro do Sul in central Portugal. Based on the phylogenetic analysis of the 16S rRNA sequence, the new organism is most closely related to the species of the genus *Thermonema* but with a pairwise sequence similarity of less than 85%. The isolate formed non-motile long filaments and rod-shaped cells that stain Gram-negative; colonies were orange-pigmented. The organism was strictly aerobic, cytochrome *c* oxidase and catalase positive. The major fatty acids were iso-$C_{15:0}$, iso-$C_{15:0}$ 2-OH and iso-$C_{17:0}$ 3-OH. The major polar lipids were one unidentified aminophospholipid, two unidentified aminolipids and three unidentified lipids. Menaquinone 7 was the major respiratory quinone. The DNA guanine plus cytosine content of strain SPSPC-11^T was 37.6% (genome sequence). The high-quality draft genome sequence corroborated many of the phenotypic characteristics of strain SPSPC-11^T. Based on genomic, phylogenetic, physiological and biochemical characteristics, we describe a new species of a novel genus represented by strain SPSPC-11^T (=CECT 9012^T =LMG 29233^T) for which we propose the name *Raineya orbicola* gen. nov., sp. nov. We also describe the family *Raineyaceae* to accommodate this new genus and species.

2.2 Introduction

The vast majority of the species of the phylum *Bacteroidetes* have optimum growth temperatures that range from about 25ºC and 45ºC, while slightly thermophilic or thermophilic species are very rare. Some organisms such as *Pseudozobellia thermophila* (Nedashkovskaya *et al*., 2009) and *Lutaonella thermophila* (Arun *et al*., 2009) have slightly elevated optimum growth temperatures of around 40–45ºC, while other species like *Anaerophaga thermohalophila* are slightly thermophilic (Denger *et al*., 2002), with an optimum growth temperature around 50ºC. Two other species classified in the phylum *Bacteroidetes* are thermophilic, namely *Thermonema lapsum* (Hudson *et al*., 1989) and *Thermonema rossianum* (Tenreiro *et al*., 1997) with optimum growth temperatures of about 60ºC and a maximum growth temperature around 65ºC. The two species of the genus *Rhodothermus*, *Rhodothermus marinus* and *Rhodothermus profundi* (Alfredsson *et al*., 1988; Nunes *et al*., 1995; Silva *et al*., 2000; Marteinsson *et al*., 2010), with optimum growth temperatures of over 65ºC and maximum growth temperatures below 80ºC, were included in the phylum *Bacteroidetes* when described but are now classified in the novel phylum named *Rhodothermaeota* (Munoz *et al*., 2016).

We isolated one strain of a slightly thermophilic organism with an optimum growth temperature around 50ºC and a maximum growth temperature of 60ºC. Phylogenetic analysis of the 16S rRNA gene sequence showed that this organism represents a distinct lineage within the phylum *Bacteroidetes*. Based on genomic, phylogenetic, physiological and biochemical parameters, we are of the opinion that strain SPSPC-11^T represents a novel genus and species for which we propose

the name *Raineya orbicola* gen. nov., sp. nov. We also propose that this organism represents a new family for which we propose the name *Raineyaceae* fam. nov.

2.3 Material and Methods

2.3.1 Isolation, culture conditions and maintenance procedures

Strain SPSPC-11^T was isolated from a reddish biofilm at the hot spring at São Pedro do Sul in central Portugal (40º 46' N, 8º 4' W). The sample was collected in screw cap tubes, transported and maintained without temperature control for 1 day. Volumes of 0.001 to 0.1 ml in 10 ml of water were filtered through membrane filters (Gelman type GN-6; pore size 0.45 µm; diameter 47 mm). The filters were placed on the surface of solidified *Thermus* medium (Albuquerque and da Costa, 2014), the plates were wrapped in plastic to prevent evaporation and incubated at 45°C for up to 5 days. Cultures were purified by sub-culturing and the isolates stored at –70ºC in *Thermus* medium with 15% (w/v) glycerol. The organism is routinely grown at 45ºC rather than at the optimum growth temperature of about 50ºC because the cultures remained viable for longer periods of time. *Thermus* medium contains (per liter of water) 1 g yeast extract (Difco), 1 g tryptone (Difco), 100 ml of a macroelements solution (10x concentrated), 10 ml of a trace elements solution (100x concentrated) and 10 ml of 0.17 mM $FeCl_3.6H_2O$, pH adjusted to 7.5 before autoclaving. The 10x concentrated macroelements solution contains per liter of water: 1 g nitrilotriacetic acid, 0.6 g $CaSO_4.2H_2O$, 1 g $MgSO_4.7H_2O$, 0.08 g NaCl, 1.03 g KNO_3, 6.89 g $NaNO_3$, 1.11 g Na_2HPO_4. The 100x concentrated trace elements solution contains per liter of water: 0.22 g $MnSO_4.H_2O$, 0.05 g $ZnSO_4.7H_2O$, 0.05 g H_3BO_3, 0.0025 g $CuSO_4.5H_2O$, 0.0025 g $Na_2MoO_4.2H_2O$, 0.0046 g $CoCl_2.6H_2O$.

2.3.2 Cell morphology, motility and determination of pigments

Cell morphology and motility were examined by phase contrast microscopy (1,000 X magnification) during the exponential growth phase in liquid *Thermus* medium at 45ºC. For transmission electron microscopy (TEM), bacteria were fixed for 2 h with 2.5% glutaraldehyde in 0.1 M cacodylate buffer (pH 7.2), washed in buffer, postfixed for 4 h with buffered 2% OsO_4, washed in buffer, followed by 1 h in 1% uranyl acetate, dehydrated in ethanol and embedded in Epon. Ultrathin sections were stained with uranyl acetate and lead citrate. For scanning electron microscopy (SEM), bacteria were initially processed as for TEM, but after postfixation a drop of bacteria suspended in buffer was laid on each coverslip coated with poly-lysine. After resting for 15 min with the buffer, the bacteria on the coverslips were dehydrated in ethanol and critical-point dried. Samples were coated with Au before being observed.

The presence of flexirubin-type pigments was determined by flooding bacterial cells with 20% KOH (Bernardet *et al*., 2002). The absorption spectra of pigments extracted using acetone:methanol (7:2, v/v) was determined at 200–900 nm with a UV-Visible spectrophotometer (ThermoScientific).

2.3.3 Biochemical and physiological characterization

Unless otherwise stated, all biochemical and physiological tests were performed in liquid *Thermus* medium or *Thermus* agar plates (Albuquerque and da Costa, 2014; for details 2.3.1) at 45ºC for up to 7 days. Catalase activity was determined by the formation of bubbles with a 3% (v/v) hydrogen peroxide solution; cytochrome *c* oxidase activity was determined by the oxidation of 1% (w/v) aqueous tetramethyl-p-phenylenediamine on filter paper room temperature. DNase activity, nitrate reduction and hydrolysis of esculin, arbutin, casein, gelatin, hippurate, starch and xylan were examined as described previously (Albuquerque *et al*., 2013; Smibert and Krieg, 1981). Additional enzymatic activities were obtained using the API ZYM system (bioMérieux) at 45ºC following the manufacturer´s instructions. Anaerobic growth was assessed in culture medium containing KNO_3 (1.0 g l^{-1}) incubated in anaerobic chambers (GENbox anaer, bioMérieux). Results were recorded after 30 days of incubation at 45ºC. Single-carbon source assimilation tests were performed in a minimal medium composed of *Thermus* medium basal salts containing filter-sterilized single carbon sources (2.0 g l^{-1}), ammonium sulfate (0.5 g l^{-1}), yeast extract (0.2 g l^{-1}) and a vitamin and nucleotide solution at a final concentration of 40 μg l^{-1} (Sharp and Williams, 1988) consisting of thiamine, riboflavin, pyridoxine, biotin, folic acid, inositol, nicotinic acid, pantothenic acid, p-aminobenzoic acid, cyanocobalamin, adenine, thymine, cytosine, guanine, cytidine, uracil and inosine (10 ml l^{-1}). Growth of the strain on single carbon sources was examined by measuring the turbidity of cultures in 20 ml screw capped tubes containing 10 ml of medium.

The optimum growth conditions were examined by measuring the turbidity (610 nm) of cultures incubated in 300 ml metal-capped Erlenmeyer flasks, containing 100 ml of *Thermus* medium, in a rotary water-bath shaker at 150 rpm. The growth temperature range of the strain was examined with 5ºC increments between 30 and 65ºC. The pH range for growth was examined at 45ºC by using 50 mM MES, HEPES, TAPS and CAPSO over a pH range of 6.0 to 9.0 with intervals of 0.5 unit. Growth with added salt, 1% (w/v) NaCl, was determined at 45ºC, pH 7.5.

2.3.4 Polar lipids, lipoquinones and fatty acids analysis

The cultures for polar lipids and lipoquinones analysis were grown at 45ºC in 1 L metal-capped Erlenmeyer flasks, containing 300 ml of *Thermus* medium, in a rotatory water-bath shaker at 150 rpm, until late-exponential phase of growth. The cultures were harvested, washed and then freeze-dried (da Costa *et al*., 2011a). The polar lipids were extracted from freeze-dried cells and the individual polar lipids were separated by TLC on glass silica gel 60 plates (Merck 1.05626, 0.25 mm thickness). The solvent system used in first direction was chroroform:methanol:water (65:25:4,

by vol) and in second direction was chroroform:methanol:acetic acid:water (80:12:15:4, by vol). To visualize phospholipids, aminolipids, glycolipids and total lipids the following reagents were used respectively, molybdenum blue, ninhydrin, *α*-naphthol-sulfuric acid and molybdophosphoric acid (da Costa *et al*., 2011a).

Lipoquinones were extracted from freeze-dried cells and purified by TLC on plastic silica gel 60 F_{254} plates coated with fluorescent indicator (Merck 105735, 0.20 mm thickness). The purified lipoquinones were separated by HPLC with a Gilson HPLC system using a reverse phase column ODS2 (Hichrom 5 C18). The solvent system used for elution of lipoquinones was methanol:heptane (10:2, v/v) at a flow rate of 2.0 min ml^{-1} at 37ºC and were detected at 269 nm (da Costa *et al*., 2011b).

Cultures for fatty acid analysis were grown in *Thermus* liquid medium at 45ºC for 5, 8 and 24 h. FAMEs were obtained from fresh wet biomass by saponification, methylation and extraction as described previously (da Costa *et al*., 2011c). FAMEs were separated using a Hewlett-Packard model 6890N gas-chromatograph equipped with an automated injector with a flame ionization detector (FID) fitted with a 5% phenyl methyl silicone capillary column (0.2 mm x 25 m; Hewlett-Packard). The carrier gas was hight-purity H_2 as a flow of 30 ml min^{-1} in the detector; the column head pressure was 60 kPa; the septum purge was 5 ml min^{-1}; the column split ratio was 55:1; and the injection port temperature was 300ºC. The temperature of the oven was programmed from 170ºC to 270ºC at a rate of 5ºC min^{-1}. As auxiliary gas, nitrogen was used at a flow of 30 ml min^{-1} and as a flame support in the detector, synthetic air (20% O_2 and 80% N_2) was used with a flow of 400 ml min^{-1}. Identification and quantification of the FAMEs, as well as the numerical analysis of the fatty acids profiles, were performed by using the standard MIS Library Generation Software (Sherlock Microbial Identification System), version 6.0, aerobe TSBA method (Microbial ID Inc., MIDI) (da Costa *et al*., 2011c).

2.3.5 Extraction of DNA and determination of the G+C content

Total genomic DNA was extracted following the method of Nielsen *et al*. (1995). Cells were lysed with a solution of lysozyme, guanidium thiocyanate and sodium *n*-lauryl sarcosine. DNA was extracted with chloroform:isoamyl alcohol (24:1, v/v), precipitated with isopropanol and washed with 70% ethanol, dried and resuspended in water. RNase was included in the extraction process. The purity of DNA was verified by 1% agarose gel electrophoresis. DNA quantity was measured by fluorescence in an Invitrogen Qubit 2.0 fluorometer (Thermo Fisher Scientific). This DNA was used for the different analyses performed.

The G+C content of DNA was determined using the HPLC method as described by Mesbah *et al*. (1989) and by genome sequencing (for details 2.3.7). For the HPLC method the DNA was degraded with P1 nuclease (Sigma, N8630) and alkaline phosphatase. The resulting deoxyribonucleosides were analysed using HPLC system using a reverse phase column ODS2 (Hichrom 5 C18). The solvent system used for elution of the nucleosides contained 12% methanol

and 20 mM trimethylamine phosphate (pH 5.1) at a flow rate of 1.0 min ml^{-1} at 37ºC and were detected at 254 nm. Calibration procedures were performed with non-methylated lambda phage DNA (Sigma, D3654). The G+C value was calculated from the ratio of deoxyguanosine (dG) and thymidine (dT) according to the method of Mesbah *et al*. (1989).

2.3.6 Phylogenetic analysis of the 16S rRNA gene sequences

The 16S rRNA gene was amplified by PCR using the forward primer 27F (5'-GAGTTTGATCCTGGCTCAG-3') and the reverse primer 1525R (5'-AGAAAGGAGGTGATCCAGCC-3') as described by Rainey *et al*. (1996). The 16S rRNA gene sequence was determined by Sanger sequencing (Macrogen, Netherlands) and by genome sequencing (for details 2.3.7). The taxonomic affiliation of strain SPSPC-11^T including the designation of its closest relatives based on the 16S rRNA gene sequence, was determined using online EzBioCloud database version 2017.5 (Yoon *et al*., 2017a). The phylogenetic dendrograms showing the position of strain SPSPC-11^T within the related taxa and with the environmental clone sequences were generated by the neighbor joining (NJ) method (Saitou and Nei, 1987) using the MEGA 6.0 software package (Tamura *et al*., 2013).

2.3.7 Genome sequencing, assembly, annotation and analysis

The genomic DNA was prepared with the Nextera XT DNA Library Preparation Kit and sequenced using paired-end (PE) 2x300 bp on the MiSeq (Illumina). Sequenced reads were quality filtered with Trimmomatic (Bolger *et al.*, 2014) and assembled with SPAdes version 3.7.1 (Bankevich *et al*., 2012) and the resulting contigs annotated with Prokaryotic Genome Prediction (PGP) (Egas *et al*., 2014). Genome estimated completeness and contamination were verified with CheckM version 1.0.7 (Parks *et al*., 2015). RNAmmer version 1.2 (Lagesen *et al*., 2007) and Usearch61 (Edgar, 2010) (against Greengenes database, version 13.8) were used for complete or partial 16S rRNA genes analysis. The two 16S rRNA genes identified were scattered in three contigs, but the complete ribosomal genes were manually reconstructed based on the mapping of PE reads against the assembled contigs with Bowtie 2 (Langmead and Salzberg, 2012). The genome of strain SPSPC-11^T was compared to the genomes of several organisms of the order *Cytophagales*, namely *Bacteroides fragilis* YCH46 (NC_006347.1), *Hymenobacter roseosalivarius* DSM 11622^T (GCA_900176135.1), *Cyclobacterium marinum* DSM 745^T (NC_015914), *Cytophaga hutchinsonii* ATCC 33406^T (NC_008255.1) and *Thermonema rossianum* DSM 10300^T (NZ_AUGC00000000) with GET_HOMOLOGUES using BLASTP and OrthoMCL (Contreras-Moreira and Vinuesa, 2013). Orthologous genes were annotated against KEGG and assigned to metabolic pathways (sequence similarity cutoff e-values of $1e^{-5}$) using KOBAS 2.0 (Xie *et al*., 2011).

2.3.8 16S rRNA gene sequences and draft genome accession numbers

The 16S rRNA gene sequences of strain SPSPC-11^T (=CECT 9012^T =LMG 29233^T) determined in this study are deposited in the GenBank/EMBL/DDBJ under the following accession numbers: KY990922, MF125287 and MF125288. The draft genome accession number of strain SPSPC-11^T is also deposited in the GenBank/EMBL/DDBJ under the accession number NKXO00000000.

2.4 Results and Discussion

2.4.1 Cell morphology, motility and colony characteristics

Isolate SPSPC-11^T formed rod-shaped cells and long filaments during the exponential phase of growth, 0.5–0.8 µm in width and 5.0–15.0 µm in length (Figure 2.1a and 2.1b). Cell wall septa were rarely seen to divide into smaller cells (Figure 2.1c). The bacterium had a Gram-negative type of cell wall (Figure 2.1d) and a few small electron-dense inclusions could be seen in the cytoplasm. Flagella or motility have not been observed. Colonies were orange-pigmented on *Thermus* medium due to carotenoids; flexirubin-type pigments were not detected.

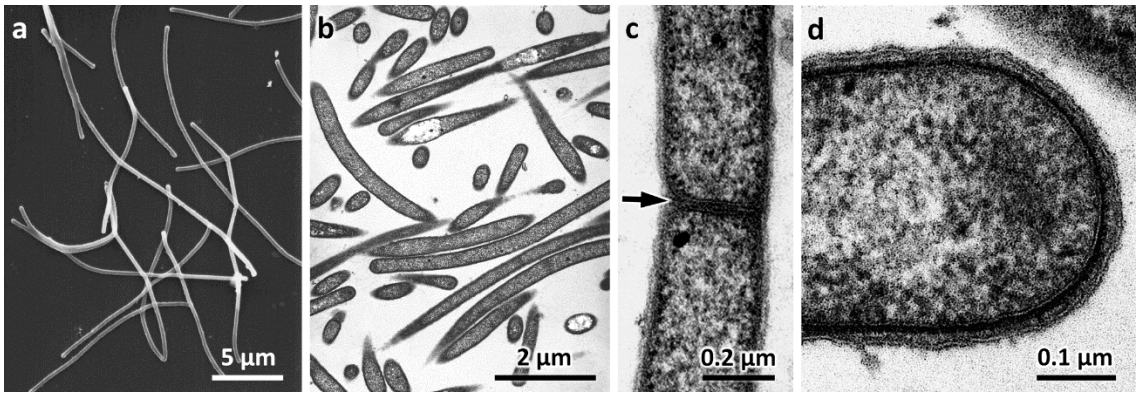

Figure 2.1 Electron microscopy by SEM and TEM of exponential phase cells of strain SPSPC-11^T. (a) Filamentous cells from a young culture (2-5 h) observed by SEM. (b) Filamentous cells from a young culture (2-5 h) observed by TEM. (c) A septum is indicated by an arrow. (d) High magnification showing the Gram-negative type of cell wall.

2.4.2 Growth conditions, biochemical and physiological characteristics

This organism was slightly thermophilic with an optimum growth temperature of about 50ºC and a temperature range for growth between 35ºC and 60ºC. The optimum pH for growth was about 7.5–8.0 with a range of growth between pH 6.5 and 8.5. Optimum growth of this organism occurred without added NaCl; no growth occurred in media with NaCl 1% (w/v). Cytochrome *c* oxidase and catalase are positive. Nitrate was not reduced to nitrite and anaerobic growth with nitrate as an

electron acceptor was not observed. The organism was strictly aerobic and chemoorganotrophic. The isolate did not utilize any of the carbohydrates or polyols examined, but grew well in a minimal medium supplemented with growth factors (yeast extract and/or a vitamin and nucleotide supplements) containing single carbon sources, namely organic acids, amino acids, casamino acids, tryptone, peptone or yeast extract. Other biochemical and physiological characteristics are listed in Table 2.1.

2.4.3 Chemotaxonomic characteristics

The polar lipid pattern on thin-layer chromatography of the new organism revealed the presence of unidentified aminophospholipid (APLs), unidentified aminolipids (ALs) and additional unidentified lipids (ULs) (Figure 2.2). The major respiratory lipoquinone was menaquinone (MK-7). The fatty acid composition of the new isolate were dominated by iso-branched and hydroxy fatty acids, namely iso-$C_{15:0}$, iso-$C_{15:0}$ 2-OH and iso-$C_{17:0}$ 3-OH and were similar during several phases of growth despite the notable changes in morphology (Table 2.2).

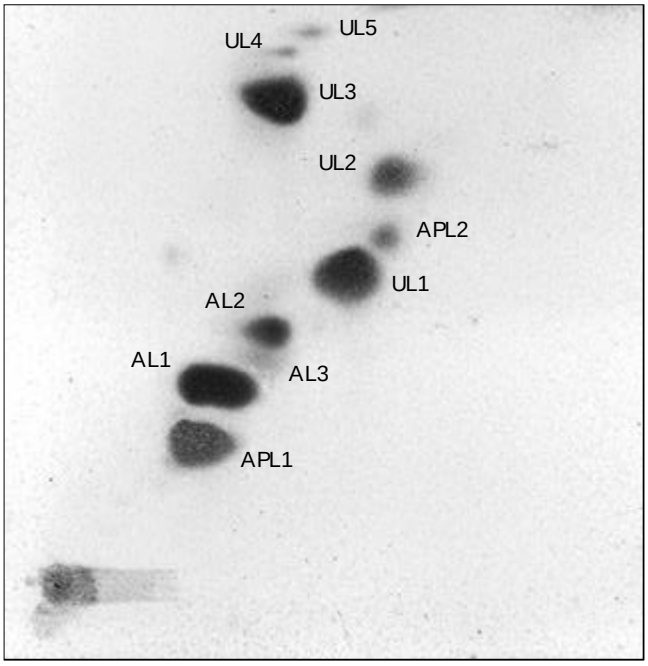

Figure 2.2 Two-dimensional thin layer chromatography of polar lipids of strain SPSPC-11^T grown at 45ºC. The lipids were stained by spraying with 5% molybdophosphoric acid in ethanol followed by heating at 160ºC. AL1,2,3, unidentified aminolipids 1,2,3; APL1,2, unidentified aminophospholipids 1,2; UL1,2,3,4,5, unidentified lipids 1,2,3,4,5.

Table 2.1 Distinguishing characteristics between strain SPSPC-11^T, *Thermonema lapsum* DSM 5718^T and *Thermonema rossianum* DSM 10300^T.

Characteristics	**SPSPC-11^T**	***Thermonema lapsum*[a,b]**	***Thermonema rossianum*[b]**
Cell size (µm)	0.5–0.8 x 5.0–15.0	0.25–0.3 x 60	0.7 in width
Temperature for growth (ºC)			
Optimum	50	60	60
Range	35–60	35–65	35–65
pH for growth			
Optimum	7.5–8.0	6.5	7.0–7.5
Range	6.5–8.5	nd	5.5–9.5
NaCl for growth (%)			
NaCl optimum	0	0	1–3
NaCl range	0	0–3	0.5–5
Assimilation of			
Acetate	+	–	–
Pyruvate	+	–	–
Aspartate	+	–	–
L-glutamate	+	–	–
L-alanine	+	–	–
L-proline	+	–	–
L-glutamine	+	–	–
L-serine	+	–	–
Tryptone	+	–	–
Peptone	+	+	nd
G+C content (mol%) (HPLC method)	39.2	47.0	50.9
G+C content (%) (genome sequencing)	37.6	nd	48.6

+, positive; –, negative; nd, not determined.
All strains were catalase and cytochrome *c* oxidase positive. Strain SPSPC-11^T and *Thermonema rossianum* DSM 10300^T do not reduce nitrate. In the API ZYM test strips strain SPSPC-11^T is positive for alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase (C14), leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin, *α*-chymotrypsin, acid phosphatase and naphthol-AS-BI-phosphohydrolase, but negative for *α*-galactosidase, *β*-galactosidase, *β*-glucuronidase, *α*-glucosidase, *β*-glucosidase, *N*-acetyl-*β*-glucosaminidase, *α*-mannosidase and *α*-fucosidase. Strain SPSPC-11^T does not hydrolyse DNA, esculin and arbutin. All strains hydrolyse casein, gelatin and hippurate but none of the strains hydrolyse starch and xylan. All strains assimilated casamino acids and yeast extract but do not assimilate D-glucose, D-fructose, D-galactose, D-mannose, L-rhamnose, L-fucose, L-sorbose, D-ribose, D-xylose, D-arabinose, L-arabinose, sucrose, maltose, D-cellobiose, lactose, D-trehalose, D-raffinose, D-melibiose, methyl-*α*-D-glucopyranoside, glycerol, ribitol, xylitol, sorbitol, D-mannitol, *myo*-inositol, erythritol, D-arabitol, *α*-ketoglutarate, DL-lactate, succinate, malate, citrate, benzoate, fumarate, formate, D-gluconate, D-glucuronate, L-asparagine, glycine, L-histidine, L-lysine, L-arginine, L-valine, L-phenylalanine, L-leucine, L-isoleucine, L-ornithine, L-methionine, L-threonine, L-glucosamine, *N*-acetylglucosamine, cysteine, cystine, tyrosine, tryptophan, glycine-betaine and dextrin.

[a]Data from Hudson *et al*. (1989).

[b]Data from Tenreiro *et al*. (1997).

Table 2.2 Fatty acid composition of strain SPSPC-11^T grown in *Thermus* liquid medium at 45ºC, and *Thermonema lapsum* DSM 5718^T and *Thermonema rossianum* DSM 10300^T grown on Degryse medium 162 agar plates at 60ºC.

		SPSPC-11^T			***T. lapsum*[a]**	***T. rossianum*[a]**
Fatty acids	**ECL**	**5 h**	**8 h**	**24 h**	**24 h**	**24 h**
Unknown 11.543	11.543	0.7 ± 0.1	0.6 ± 0.1	0.6 ± 0.1	–	–
iso-$C_{14:0}$	13.619	–	–	–	1.0	1.5
Unknown 13.565	13.565	5.6 ± 0.4	4.9 ± 0.5	4.5 ± 0.7	–	–
iso-$C_{15:0}$	14.623	59.4 ± 1.5	61.8 ± 1.6	57.5 ± 1.7	37.4	39.9
anteiso-$C_{15:0}$	14.713	1.2 ± 0.2	1.1 ± 0.2	1.4 ± 0.2	8.4	8.6
$C_{15:0}$	15.000	1.0 ± 0.2	1.0 ± 0.2	1.4 ± 0.2	2.7	1.8
iso-$C_{16:0}$	15.627	tr	tr	tr	1.0	1.5
iso-$C_{15:0}$ 2-OH	15.852	8.7 ± 0.4	8.1 ± 0.3	10.8 ± 0.6	7.4	5.4
$C_{16:0}$	16.000	3.7 ± 0.3	3.6 ± 0.2	2.9 ± 0.3	1.2	1.2
iso-$C_{15:0}$ 3-OH	16.134	2.8 ± 0.2	2.9 ± 0.2	3.0 ± 0.2	8.9	8.3
$C_{15:0}$ 2-OH	16.219	tr	tr	tr	1.2	0.6
iso-$C_{17:0}$	16.630	tr	tr	tr	1.1	1.2
Unknown 16.582	16.582	0.7 ± 0.1	0.6 ± 0.1	0.6 ± 0.1	2.0	2.1
$C_{17:1}$ ω6*c*	16.860	–	–	–	1.3	0.8
iso-$C_{16:0}$ 3-OH	17.150	tr	tr	tr	1.1	1.4
$C_{16:0}$ 2-OH	17.233	0.5 ± 0.1	tr	0.5 ± 0.1	–	–
$C_{16:0}$ 3-OH	17.519	1.9 ± 0.3	1.9 ± 0.3	2.4 ± 0.4	1.5	1.3
iso-$C_{17:0}$ 2-OH	17.880	–	–	–	1.1	0.5
iso-$C_{17:0}$ 3-OH	18.161	10.1 ± 0.7	10.0 ± 0.6	10.1 ± 0.9	17.4	18.7
anteiso-$C_{17:0}$ 3-OH	18.254	–	–	–	0.7	1.0

Results are percentage of the total fatty acids. ±, results are the mean plus the standard deviation of two to four analyses; values for fatty acids present at levels of less than 0.5% in all strains are not shown; tr, trace (< 0.5%); –, not detected; ECL, equivalent chain length.

[a]Data from Tenreiro *et al*. (1997).

2.4.4 Phylogenetic analysis of the 16S rRNA gene sequences

The analysis of the 16S rRNA gene sequence of strain SPSPC-11^T (KY990922) demonstrated that strain SPSPC-11^T belonged to the phylum *Bacteroidetes* and represented a novel cultured lineage that shared less than 85% similarity with previously described taxa. The SPSPC-11^T lineage clustered with the lineage of the family *Thermonemataceae* within the order *Cytophagales* (Figure 2.3). Comparison of the two 16S rRNA gene sequences (MF125287; M125288) determined from the draft genome sequence with environmental sequences showed it to share 90–99% similarity with sequences recovered from a range of aquatic environments (Figure 2.4 and Table 2.3).

Published studies on the phylogeny of the phylum *Bacteroidetes*, based on whole genome comparisons, have demonstrated the existence of a number of lineages representing new taxa at the phylum, class, order and family levels (Hahnke *et al*., 2016; Munoz *et al*., 2016), although Munoz *et al*. (2016) designated 16S rRNA gene sequence similarity ranges outside the taxonomic levels proposed by Hahnke *et al*. (2016). Phylogenetic analysis of the 16S rRNA gene sequence of strain SPSPC-11^T showed its position within this classification of the phylum *Bacteroidetes* and related taxa (Figure 2.3). Based on the 16S rRNA gene sequence similarity values to related taxa, less than 85%, and the position within the phylogenetic tree it is demonstrated that strain SPSPC-11^T represents a novel lineage at the family level within the order *Cytophagales*.

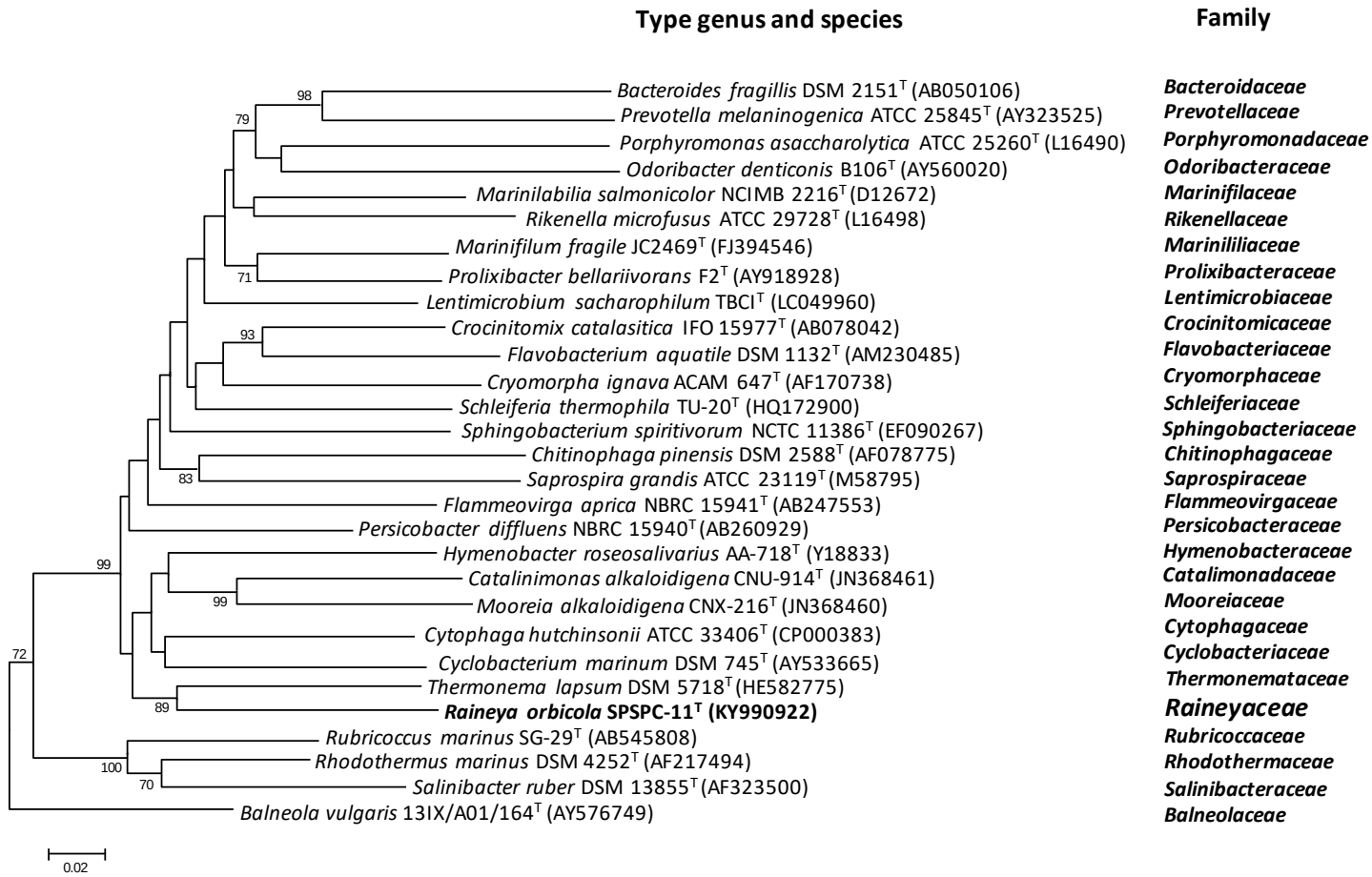

Figure 2.3 Phylogenetic position of strain SPSPC-11^T within the radiation of representatives of the families of the phyla *Bacteroidetes* and *Rhodothermaeota*. The phylogenetic dendrogram was generated by the NJ method using the MEGA 6.0 software package. Bootstrap values, expressed as percentages of 1000 replications, are given at branching points. The scale bar represents 2 inferred nucleotide substitutions per 100 nucleotides.

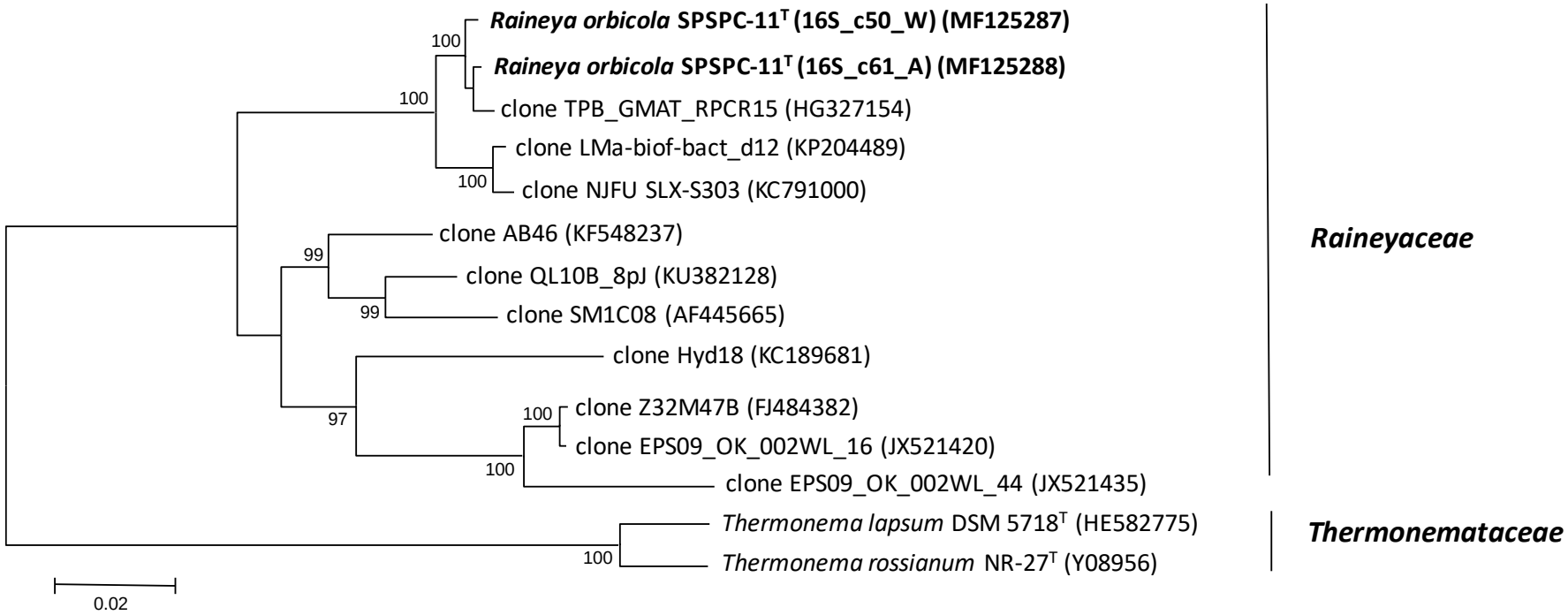

Figure 2.4 Phylogenetic position of strain SPSPC-11^T within the radiation of representatives of environmental clone sequences to belong to the *Raineya* lineage. The source of the environmental clone sequences is shown in Table 2.3. The phylogenetic dendrogram was generated by the NJ method using the MEGA 6.0 software package. Bootstrap values, expressed as percentages of 1000 replications, are given at branching points. The scale bar represents 2 inferred nucleotide substitutions per 100 nucleotides.

Table 2.3 Information on environmental clone sequences that belong to the *Raineyaceae* lineage within the *Bacteroidetes*.

Accession number	Clone designation	Pairwise sequence similarity (%) of 16S rRNA gene with SPSPC-11^T	Query cover (%)	Source and location
HG327154	TPB_GMAT_RPCR15	99	93	Microbial mat, Tapoban hot springs, Garhwal, India
KP204489	LMa-biof-bact_d12	98	99	Microbial biofilm, acidic geothermal area of Copahue, Neuquen, Argentina
KC791000	NJFU SLX-S303	98	95	Whitewater of paper making machine, ChangShu, China
KF548237	AB46	93	98	Anerobic tank of wastewater treatment plant, China
KU382128	QL10B_8pJ	93	98	Microbial mat, Queen's Laundry hot spring, Yellowstone National Park, USA
AF445665	SM1C08	92	99	Travertine deposition, Angel Terrace, Mammoth Hot Springs, USA
FJ484382	Z32M47B	92	91	Wall biomat, phreatic sinkhole, El Zacaton, Mexico
JX521420	EPS09_OK_002WL_16	92	99	Terrestrial sulfidic spring, Sulphur Springs, Oklahoma, USA
KC189681	Hyd18	90	100	Freshwater spring, Wakulla Spring, Florida, USA
JX521435	EPS09_OK_002WL_44	89	99	Terrestrial sulfidic spring, Sulphur Springs, Oklahoma, USA
*HG327161	TPB_GMAT_RPCR32	100	55	Microbial mat, Tapoban hot springs, Garhwal, India
*EU815166	yang-W129	98	58	Thermal vent boiling pool, Tibet, China

*Not included in phylogenetic analysis and Figure 2.4 due to short sequences and low query coverage.

2.4.5 High-quality draft genome sequence and analysis

The observation that strain SPSPC-11^T was unable to grow on any of the sugars examined prompted us to produce a high-quality draft genome sequence to assess the possibility that some genes involved in sugar catabolism would not be present. Additionally, the genome was searched for other metabolic processes and compared with the genomes sequences of members of the order *Cytophagales* that assimilate carbohydrates, namely *Bacteroides fragilis* YCH46 (NC_006347.1), *Hymenobacter roseosalivarius* DSM 11622^T (GCA_900176135.1), *Cyclobacterium marinum* DSM 745^T (NC_015914) and *Cytophaga hutchinsonii* ATCC 33406^T (NC_008255.1), as well as the genome sequence of *Thermonema rossianum* DSM 10300^T (NZ_AUGC00000000) that does not utilize any sugars tested (Tenreiro *et al*., 1997).

The SPSPC-11^T DNA sequence run generated 2,112,714 PE reads of which 1,796,859 high-quality reads remained after quality filtering. The *de novo* read assembly produced 104 contigs with an N50 size of 67,061 bp (Table 2.4). The high-quality draft assembled genome sequence consisted of 3,070,213 bp with a DNA G+C content of 37.6%. CheckM estimated the genome to be near-completion (98.2%) and the level of contamination to be extremely low (0.3%). No contamination was detected for 16S rRNA genes as tested by RNAmmer and Usearch61. The genome had a total of 2,730 genes, including 2,685 protein-coding genes, 39 tRNA genes and 6 rRNA genes (two 5S, two 16S and two 23S) (Table 2.4). Analysis of the whole genome sequence demonstrated the presence of two 16S RNA gene coding sequences. The two 16S rRNA gene sequences differed at 8 positions over 1501 compared nucleotides representing 99.47% identity. The presence of multiple 16S rRNA gene copies with such levels of similarity between the gene copies of the same organism have been reported across many bacterial taxa and in representatives of the phylum *Bacteroidetes* (Pei *et al*., 2010; Sun *et al*., 2013).

The draft genome comprised 2,115 genes with putative functions (~79% of total protein-coding genes) and 1,320 allocated to COG functional categories (~49% of total protein-coding genes). The most abundant COG category was “Translation, ribosomal structure and biogenesis” followed by “Cell wall / membrane biogenesis” and “Amino acid transport and metabolism” (Table 2.5).

Table 2.4 Genome sequencing project information and statistics of strain SPSPC-11^T and *Thermonema rossianum* DSM 10300^T.

		SPSPC-11^T	***Thermonema rossianum*[a]**
MIGS ID*	**Attribute**	**Value/comment**	
MIGS 28	Libraries used	Illumina PE library (2 × 300 bp insert size)	Illumina PE library
MIGS 29	Sequencing platforms	Illumina MiSeq	Illumina HiSeq 2000 and HiSeq2500
	Size of raw data included in the assembly process (Mbp)	820	176.2
MIGS 30	Assembler	Spades v. 3.7.1	–
MIGS 31	Finishing quality	High-quality draft	High-quality draft
MIGS 31.2	Sequencing depth of coverage	250x	–
MIGS 31.3	Number of contigs	104	26
MIGS 32	Gene calling method	PGP	Prodigal 2.5
	N50 (bp)	67,061	202,966
	Estimated genome completeness (%)	98.2	–
	Assembled genome size (bp)	3,070,213	2,956,866
	DNA coding (bp)	2,806,590	2,723,503
	DNA G+C (bp)	1,151,283	1,441,896
	DNA G+C (%)	37.6	48.6
	Total genes	2,730	2,654
	Protein-coding genes	2,685	2,599

Table 2.4 (continued)

RNA genes	45	55
tRNA genes	39	44
rRNA genes	6	9
5S	2	3
16S	2	3
23S	2	3
Genes with function prediction	2,115	1,935
Genes assigned to COGs	1,320	1,511
Genes with Pfam domains	2,048	1,998
Genes with Tfam domains	749	–
CRISPR repeats	4	2
Estimated contamination (%)	0.3	–
Authenticity of strain checked by	16S (rRNA gene from Sanger and genome sequencing)	–
Accession number of the assembly	NKXO00000000	ASM42682v1
Accession number of raw data the assembly	SRR5815076	SRP054817

[*]Based on MIGS recommendations (Field *et al*., 2008).
[a]Data from NCBI Bioproject PRJNA195851 and JGI Project 1015836.

Table 2.5 Number of genes associated with general COG functional categories.

Code	Value	%*	Description
R	139	5.18	General function prediction only
J	135	5.03	Translation, ribosomal structure and biogenesis
M	116	4.32	Cell wall / membrane biogenesis
E	112	4.17	Amino acid transport and metabolism
L	95	3.54	Replication, recombination and repair
C	82	3.05	Energy production and conversion
H	77	2.87	Coenzyme transport and metabolism
I	71	2.64	Lipid transport and metabolism
O	61	2.27	Posttranslational modification, protein turnover, chaperones
K	56	2.09	Transcription
P	54	2.01	Inorganic ion transport and metabolism
F	53	1.97	Nucleotide transport and metabolism
T	45	1.68	Signal transduction mechanisms
G	41	1.53	Carbohydrate transport and metabolism
Q	30	1.12	Secondary metabolites biosynthesis, transport and catabolism
D	19	0.71	Cell cycle control, cell division, chromosome partitioning
V	25	0.93	Defence mechanisms
U	21	0.78	Intracellular trafficking and secretion
N	3	0.11	Cell motility
B	1	0.04	Chromatin structure and dynamics
S	84	3.13	Function unknown
–	1365	50.84	Not in COGs

*The percentage is based on the total number of protein-coding genes in the annotated genome.

2.4.6 Insights from the genome sequences

Several genes coding for enzymes involved in the initial catabolism of carbohydrates to glucose were not identified in the new strain, thus preventing the utilization of hexoses or pentoses through the Embden-Meyerhof-Parnas or the Entner-Doudoroff pathways. It is noteworthy that *T. rossianum*, also lacks the same genes for the initial catabolism of sugars and is, like strain SPSPC-11^T, unable to grow on any of the sugars examined (Tenreiro *et al*., 1997). In contrast, the genome sequence of *B. fragilis*, *H. roseosalivarius*, *Cyclobacterium marinum* and *Cytophaga hutchinsonii* predict the assimilation of hexoses and pentoses through these pathways, as also confirmed by assimilation tests (Hirsch *et al*., 1998; Larkin, 1989; Varel and Bryant, 1974).

It is possible that strain SPSPC-11^T lacks the genetic ability to metabolize carbohydrates confirming the results of the phenotypic tests that show that sugars do not serve as carbon and energy sources for growth. Similar to other members of the order *Cytophagales* (*B. fragilis*, *H. roseosalivarius*, *Cyclobacterium marinum* and *Cytophaga hutchinsonii*), the putative gene for fructose-1,6-bisphosphatase (EC:3.1.3.11) was identified, suggesting that strain SPSPC-11^T can perform gluconeogenesis. The genome sequence of strain SPSPC-11^T predicts that the tricarboxylic acid cycle is complete.

The draft genome of strain SPSPC-11^T indicated that oxidative phosphorylation occurs via NADH dehydrogenase, succinate dehydrogenase, cytochrome *c*, cytochrome *c* oxidase and an F-type ATPase. The *T. rossianum* genome sequence appears to possess several genes coding for the same oxidative phosphorylation functions that were identified in the strain SPSPC-11^T with exception of the NuoEG subunits of the NADH dehydrogenase complex. In contrast to strain SPSPC-11^T, genes coding for cytochrome bd complex were identified in *Cyclobacterium marinum* and *B. fragilis*. The genome of *B. fragilis* lacks not only cytochrome *c* oxidase like-genes but also the NuoEFG subunits of the NADH dehydrogenase complex. The latter organisms also possess some V/A Type ATPase-associated genes in addition to F-type ATPase.

The absence of assimilatory nitrate or dissimilatory nitrite reduction genes by strain SPSPC-11^T, *H. roseosalivarius* and *T. rossianum* confirms the absence of phenotypic nitrate reduction. The genes involved in nitrate/nitrite transport and nitrate reduction, namely the assimilatory nitrate reductase and the enzymes for denitrification were not encountered. The other members of the order *Cytophagales*, namely *B. fragilis*, *Cyclobacterium marinum* and *Cytophaga hutchinsonii* possess putative genes involved in nitrite reduction, while *Cyclobacterium marinum* and *Cytophaga hutchinsonii* also had genes involved in the assimilatory nitrate reduction to nitrite.

2.4.7 Description of a novel family, genus and species

From the comparison of environmental sequences from uncultured organisms it was demonstrated that strain SPSPC-11^T is a cultured representative of a family level phylogenetic lineage within the phylum *Bacteroidetes* that has been already detected and is represented by 16S rRNA gene

sequences recovered from geographically distant aquatic environments, many of them geothermal (Figure 2.4 and Table 2.3). Based on the 16S rRNA gene sequence similarities within the lineage represented by environmental sequences and now strain SPSPC-11^T it is clear that this lineage contains a number of novel genera and species yet to be cultured. Phylogenetic analysis demonstrated that strain SPSPC-11^T represents the first cultured member of a novel family level lineage within the order *Cytophagales* of the phylum *Bacteroidetes* (Figure 2.3 and Figure 2.4).

The new lineage represented by strain SPSPC-11^T possesses genotypic and phenotypic features that resembled those of the species of the genus *Thermonema*. However, the strain SPSPC-11^T can be distinguished from the genus *Thermonema* in several characteristics: strain SPSPC-11^T assimilates some single amino acids and organic acids while the species of the genus *Thermonema* assimilate only complex mixtures of amino acids; the optimum growth temperatures of the organisms differ by about 10ºC; the inability of the new species to grow in medium with added NaCl; the large difference between the DNA G+C % of strain SPSPC-11^T and the species of the genus *Thermonema* (Table 2.1).

On the basis of these results, we propose that strain SPSPC-11^T represents a novel species of a new genus for which we recommend the name *Raineya orbicola* gen. nov., sp. nov. Moreover, we are of the opinion that the genomic, phylogenetic, chemotaxonomic and phenotypic characteristics warrant a new family within the phylum *Bacteroidetes* for which we propose the name *Raineyaceae* fam. nov.

2.4.7.1 Description of *Raineyaceae* fam. nov.

Raineyaceae (Albuquerque *et al*., 2018a^{VP}; Albuquerque and Egas, 2020).

Rai.ney.a.ce'ae. N.L. fem. dim. n. *Raineya*, type genus of the family; suff. *-aceae*, ending denoting a family; N.L. fem. pl. *Raineyaceae*, the *Raineya* family.

Cells stain Gram-negative and form rod-shaped cells. Endospores are not formed. Chemorganotrophic and strictly aerobic. Slightly thermophilic. MK-7 is the major respiratory lipoquinone. Represents a distinct phylogenetic lineage within the order *Cytophagales*. The type genus of this family is *Raineya*.

2.4.7.2 Description of *Raineya* gen. nov.

Raineya (Albuquerque *et al*., 2018a^{VP}; Albuquerque and Egas, 2021a).

Rai.ney.a. N.L. fem. n. *Raineya* referring to Fred A. Rainey, for his contributions to the taxonomy and phylogeny of archaea and bacteria.

Cytochrome *c* oxidase and catalase positive. Flexirubin-type pigments are not present. Carbohydrates are not utilized for growth. The polar lipid profile is composed of ALs, APLs and

ULs. The fatty acid composition is dominated by iso-branched fatty acids and hydroxyl fatty acids. The type species of the genus is *Raineya orbicola*.

2.4.7.3 Description of *Raineya orbicola* sp. nov.

Raineya orbicola (Albuquerque *et al*., 2018a[VP]; Albuquerque and Egas, 2021a).

or.bi'co.la. L. masc. n. *orbis*, the whole world; L. suff. *cola*, inhabitant, dweller; N.L. n. *orbicola*, inhabitant of the whole world.

Forms long filaments and rod-shaped cells with 0.5–0.8 µm in width and 5.0–15.0 µm in length; colonies on *Thermus* medium are orange-pigmented due to carotenoids. Growth occurs between 35 and 60ºC; the optimum growth temperature is about 50ºC. The optimum pH for growth is about 7.5–8.0; growth occurs between pH 6.5 and 8.5. Optimum growth occurs without added NaCl; no growth occurs with 1% NaCl. Yeast extract and/or a vitamin and nucleotide solution is required for growth. Nitrate is not reduced to nitrite. Gelatine, casein and hippurate are degraded; starch, esculin, arbutin and xylan are not degraded. DNase negative. In the API ZYM alkaline phosphatase, esterase (C4), esterase lipase (C8), lipase (C4), leucine arylamidase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, acid phosphatase and naphthol-AS-BI-phosphohydrolase are positive; other enzyme activities are negative. Acetate, pyruvate, aspartate, L-glutamate, L-alanine, L-proline, L-glutamine, L-serine, yeast extract, tryptone, peptone and casamino acids are assimilated. Other single carbon sources tested are not assimilated (Table 2.1). The major fatty acids are iso-$C_{15:0}$, iso-$C_{15:0}$ 2-OH and iso-$C_{17:0}$ 3-OH. The DNA of strain SPSPC-11^T has a G+C content of 39.2 mol% (HPLC method) and 37.6% (genome sequencing). The type strain SPSPC-11^T (=CECT 9012^T =LMG 29233^T) was isolated from a hot spring at São Pedro do Sul in central Portugal. GenBank/EMBL/DDBJ accession numbers for the 16S rRNA gene sequences of strain SPSPC-11^T are KY990922, MF125287 and MF125288. The draft genome sequence of SPSPC-11^T (NKXO00000000) has been deposited in GenBank/EMBL/DDBJ.

Acknowledgments

We thank Bernhard Schink (University of Constance, Germany) for the etymology of the name of the new organism.

Chapter 3

Comparative Genome Sequence Analysis of Species of the Genus *Tepidimonas* and the Description of *Tepidimonas charontis*, a New Thermophilic Species

Results published in:

ALBUQUERQUE, L., CASTELHANO, N., RAPOSO, P., FROUFE, H.J.C., TIAGO, I., SEVERINO, R., ROXO, I., GREGÓRIO, I., BARROSO, C., EGAS, C. and DA COSTA, M.S. (2020a). Comparative genome sequence analysis of several species in the genus *Tepidimonas* and description of a novel species *Tepidimonas charontis* sp. nov. *Int J Syst Evol Microbiol* 70: 1596–1604. doi.org/10.1099/ijsem.0.003942

ALBUQUERQUE, L., CASTELHANO, N., RAPOSO, P., FROUFE, H.J.C., TIAGO, I., SEVERINO, R., ROXO, I., GREGÓRIO, I., BARROSO, C., EGAS and DA COSTA, M.S. (2020b). Corrigendum: Comparative genome sequence analysis of several species in the genus *Tepidimonas* and description of a novel species *Tepidimonas charontis* sp. nov. *Int J Syst Evol Microbiol* 70: 6539. doi.org/10.1099/ijsem.0.004563

ALBUQUERQUE, L. and EGAS, C. (2021b). *Tepidimonas*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.gbm00959.pub2

3.1 Abstract

We performed high-quality genome sequencing of eight strains of the species of *Tepidimonas* and examined the genomes of closely related strains from the databases to understand why *Tepidimonas taiwanensis* is the only strain of this genus that utilizes glucose and fructose for growth. We found that the assimilation of these hexoses by *T. taiwanensis* was due to the presence of two transporters that are absent in all other genomes of *Tepidimonas* strains examined. Some strains lack genes coding for glucokinase, but the Embden-Meyerhof-Parnas pathway appears to be otherwise complete. The pentose-phosphate pathway has a complete set of genes, but genes of the Entner-Doudoroff were not identified in the genomes of any of the strains examined. Genome analysis using average nucleotide identity, digital DNA-DNA hybridization, average amino acid identity and phylogenetic analysis of 400 conserved genes were performed to assess the taxonomic classification of the organisms. Two isolates of the genus *Tepidimonas* from the hot spring at São Pedro do Sul, Portugal, designated SPSP-6^T and SPSPC-18 were also examined. These organisms are mixotrophic, have an optimum growth temperature of about 50ºC, utilize several organic acids and amino acids for growth but do not grow on sugars. Distinctive phenotypic, 16S rRNA gene sequence similarities and genomic characteristics of strains SPSP-6^T and SPSPC-18 lead us to propose a novel species based on strain SPSP-6^T for which we recommend the name *Tepidimonas charontis* sp. nov. (=CECT 9683^T =LMG 30884^T).

3.2 Introduction

The slightly thermophilic species of the genus *Tepidimonas* are classified in the class *Betaproteobacteria* of the order *Burkholderiales*. The genus *tepidimonas* was recently accomodated in the family *Comamonadaceae*, and for a long time was considered a genus *incertae sedis VII* (Albuquerque and Egas, 2021b). The genus comprises seven validly named species, *Tepidimonas ignava* (Moreira *et al*., 2000), *Tepidimonas aquatica* (Freitas *et al*., 2003), *Tepidimonas taiwanensis* (Chen *et al*., 2006), *Tepidimonas thermarum* (Albuquerque *et al*., 2006), *Tepidimonas fonticaldi* (Chen *et al*., 2013), *Tepidimonas alkaliphilus* and *Tepidimonas sediminis* (Habib *et al*., 2018), while the name of the species "*Tepidimonas arfidensis*" has not been validated (Ko *et al*., 2005).

With the exception of *T. taiwanensis*, none of the strains of this genus grow in a minimal medium with glucose and fructose as sole carbon and energy source. The type strains of all species of this genus assimilate individual organic acids and amino acids for growth. Moreover, the strains examined oxidize thiosulfate in the presence of an organic carbon source indicating that the strains are mixotrophic. Chemoorganotrophic and mixotrophic bacteria that do not utilize sugars for growth are not rare (Albuquerque *et al*., 2018d; da Costa *et al*., 2019); these organisms may lack specific sugar transporters or enzymes involved in the Embden-Meyerhof-Parnas, the Entner-Doudoroff or the pentose-phosphate pathways. With the objective of understanding the conundrum that only one type strain of this genus is capable of growing on hexoses, we performed high-quality

draft genome sequences of eight type strains of *T. ignava*, *T. aquatica*, *T. fonticaldi*, *T. taiwanensis*, *T. thermarum*, *T. sediminis*, *T. alkaliphilus* and the type strain of one isolate of the genus *Tepidimonas* from the hot spring at São Pedro do Sul, Portugal, designated SPSP-6^T. The type strain of the genus *Tepidimonas*, *T. ignava* SPS-1037^T, was also isolated from the same hot spring at São Pedro do Sul (Moreira *et al*., 2000). We also scrutinized two genome sequences of strains closely related to the type strain of *T. taiwanensis* I1-1^T, namely strains VT154-175 and MB2 as well as a strain closely related to *T. fonticaldi* designated PL17 (Dhakan *et al*., 2016; Valeriani *et al*., 2016). We also propose that strain SPSP-6^T represents a novel species for which we recommend the name *Tepidimonas charontis* sp. nov.

3.3 Material and Methods

3.3.1 Isolation, culture conditions, maintenance procedures and bacterial strains

Strains SPSP-6^T and SPSPC-18 were isolated from water samples at the hot spring at São Pedro do Sul in Central Portugal (40º 46' N, 8º 4' W) with temperatures of 65ºC and 50ºC, respectively. The samples were maintained without temperature control for 1 day. Samples or dilutions of the water samples were filtered through membrane filters (Gelman type GN-6; pore size 0.45 µm; diameter 47 mm). The filters were placed on the surface of solidified *Thermus* medium (section 2.3.1 for details) (Albuquerque *et al*., 2018c). The plates were wrapped in plastic to prevent evaporation and incubated at 50ºC for up to 5 days. Cultures were purified by sub-culturing and the isolates stored at –70ºC in *Thermus* medium with 15% (w/v) glycerol. Cultivation in Degryse medium 162 was later adopted because this medium resulted in higher growth yields (Albuquerque *et al*., 2018d). Degryse medium 162 contains (per liter of water) 2.5 g yeast extract (Difco), 2.5 g tryptone (Difco), 100 ml of a macroelements solution (10x concentrated), 5 ml of a trace elements solution (identical to the trace elements of *Thermus* medium, section 2.3.1 for details), 15 ml of 0.2 M $Na_2HPO_4.12H_2O$, 10 ml of 0.2 M KH_2PO_4 and 0.5 ml of 0.01 M ferric citrate, pH adjusted to 7.5 before autoclaving. The 10x concentrated macroelements solution contained per liter of water: 1 g nitrilotriacetic acid, 0.4 g $CaSO_4.2H_2O$ and 2 g $MgCl_2.6H_2O$.

The type strains of *Tepidimonas alkaliphilus* YIM 72238^T (KCTC 52717^T), *T. aquatica* CLN-1^T (DSM 14833^T), *T. fonticaldi* AT-A2^T (KCTC 23862^T), *T. ignava* SPS-1037^T (DSM 12034^T), *T. sediminis* YIM 72259^T (NBRC 112410^T), *T. taiwanensis* I1-1^T (LMG 22826^T) and *T. thermarum* AA-1^T (LMG 23094^T) were used for comparative purposes.

3.3.2 Phenotypic and chemotaxonomic characterization

Cell morphology and motility of strains SPSP-6^T and SPSPC-18 were examined by phase contrast microscopy (1,000X magnification) during the exponential growth phase in liquid Degryse medium

162 at 50ºC. Cell dimensions were determined with an ocular micrometer with a stage micrometer; motility with the Ryu stain (Heimbrook *et al*., 1989).

The optimum growth conditions were examined by measuring the turbidity (610 nm) of cultures incubated in 300 ml metal-capped Erlenmeyer flasks, containing 100 ml of Degryse medium 162, in a rotary water-bath shaker at 150 rpm. The growth temperature ranges of strains SPSP-6^T and SPSPC-18 were examined with 5ºC increments between 20 and 65ºC. The pH range for growth was examined at 50ºC by using 50 mM MES, HEPES, TAPS and CAPSO over a pH range of 6.0 to 10.0 with 0.5 unit increments. Growth with added salt, 0.5 and 1% (w/v) NaCl, was determined at 50ºC and pH 7.5.

Catalase and cytochrome *c* oxidase activities, nitrate reduction and anaerobic growth were examined as described previously by Smibert and Krieg (1981) (for details 2.3.3). Single-carbon source assimilation tests were performed in a minimal medium composed of Degryse medium 162 basal salts containing filter-sterilized single carbon sources (2.0 g l^{-1}), ammonium sulfate (0.5 g l^{-1}) and a vitamin and nucleotide solution at a final concentration of 40 µg l^{-1} described previously by Sharp and Williams (1988) (for details 2.3.3). Growth on thiosulfate was assessed on modified medium DSM 69 containing the following components per liter: 10.6 g $Na_2HPO_4.12H_2O$, 1.5 g KH_2PO_4, 0.3 g NH_4Cl, 1.0 g yeast extract, 1 ml trace elements solution SL-6 of medium DSM 27 without the addition of sulfate (0.03 g $MnCl_2.4H_2O$, 0.3 g H_3BO_3, 0.2 g $CoCl_2.6H_2O$, 0.01 g $CuCl_2.2H_2O$, 0.02 g $NiCl_2.6H_2O$, 0.03 g $Na_2MoO_4.2H_2O$, per liter of water). The modified medium DSM 69 was supplemented with the same amount of the vitamin and nucleotide solution used for the single-carbon source assimilation tests. Concentrations of 0.5 g l^{-1} and 1 g l^{-1} of thiosulfate was added to this media. At appropriate intervals, the turbidity of the cultures was measured and the levels of thiosulfate and sulfate in the supernatants were measured using the methods described by Westley (1987) and Sörbo (1987). Thiosulfate was quantified by measurement of the ferric thiocyanate complex ion [$Fe(SCN)_3^6$] by the colorimetric method (Westley, 1987). An aliquot of sample (10 nmol to 1 µmol) was prepared in 1.8 ml of water, and 100 µl of KCN 0.25 M and 100 µl $CuSO_4$ 0.2 M were added successively. Ferric thiocyanate was measurement by absorbance at 460 nm, after the addition of 1 ml of Sörbo reagent (100 g l^{-1} $Fe(NO_3)_3.9H_2O$ and 200 ml l^{-1} HNO_3 65%). Sulfate was quantified according to the turbidimetric method by measurement of the absorbance at 600 nm (Sörbo, 1987). A sample aliquot (up to 2.5 µmol) was prepared in 3 ml of water and 1 ml of HCl 0.5 M was added, followed by 1 ml of Ba-PEG reagent (2.44 g $BaCl_2.H_2O$, 37.5 g polyethylene glycol 6000, 500 µl Na_2SO_4 50 mM).

Cultures for fatty acid analysis were grown on R2A and Degryse medium 162 at 50ºC for 24 h. FAMEs were obtained from fresh wet biomass, separated, identified and quantified with the standard Sherlock MIS Library Generation Software, version 6.0, aerobe TSBA method (Microbial ID Inc., MIDI) as described previously by da Costa *et al*., 2011c (for details 2.3.4).

3.3.3 Extraction of DNA

Total genomic DNA of *T. thermarum* AA-1^T (LMG 23094^T), *T. ignava* SPS-1037^T (DSM 12034^T), *T. aquatica* CLN-1^T (DSM 14833^T), *T. fonticaldi* AT-A2^T (KCTC 23862^T), *T. taiwanensis* I1-1^T (LMG 22826^T), *T. sediminis* YIM 72259^T (NBRC 112410^T), *T. alkaliphilus* YIM 72238^T (KCTC 52717^T) and SPSP-6^T (LMG 30884^T) was extracted following the method of Nielsen *et al*. (1995) (for details 2.3.5). The purity and quantity of the DNA were verified as described in 2.3.5.

3.3.4 Phylogenetic analysis of the 16S rRNA gene sequences

PCR-amplification of 16S rRNA genes of strains SPSP-6^T and SPSPC-18 was carried out as described in section 2.3.6. The 16S rRNA gene sequence was determined by Sanger sequencing (Macrogen). Multiple sequence alignments were performed using MUSCLE (Edgar, 2004). Phylogenetic trees were reconstructed with the NJ and maximum likelihood (ML) algorithms using MEGA (version X) (Kumar *et al*., 2018b). For the NJ and ML algorithms, genetic distances were calculated with the Jukes-Cantor model (Jukes and Cantor, 1969). Bootstrap analysis based on 1000 replicates evaluated resulting tree topologies.

3.3.5 Genome sequencing, assembly, annotation and analysis

The DNA was prepared for genome sequencing using the Nextera XT DNA Library Preparation Kit (Illumina). Bacterial genomes were sequenced on the MiSeq (Illumina) with PE 2x300 bp reads. The draft genomes of strains MB2 (GCF_001481285.1), VT154-175 (GCF_000807215.1) and PL17 (GCF_001675355.1) members of the genus *Tepidimonas* were obtained from public databases.

Sequenced reads were filtered for quality with Trimmomatic version 0.30 (Bolger *et al*., 2014) and assembled with SPAdes version 3.9.1 (Bankevich *et al*., 2012). The resulting contigs were annotated with PGP2. PGP2 used Prodigal version 2.6 (Hyatt *et al*., 2010) for gene prediction, Barrnap version 0.8 (https://github.com/tseemann/barrnap) for rRNA and tRNA genes detection, and Prokka version 1.12 (Seemann, 2014) for the annotation of protein-coding genes. Gene annotation with Prokka used the SwissProt (Apweiler *et al*., 2004), HAMAP (Pedruzzi *et al*., 2015), TIGRFAMs (Haft *et al*., 2003) and Pfam (Finn *et al*., 2016) repositories. Genes observed to be missing in the pathways were searched manually at the ends of the contigs and were annotated. These genes were the sulfur-oxidizing protein *soxZ* of *T. aquatica*, strain SPSP-6^T, *T. taiwanensis* I1-1^T and strain MB2, the phosphoglycerate kinase gene (*pgk*) of strain SPSP-6^T and the alpha chain of the nitrate reductase *narG* of *T. thermarum*. Genome estimated completeness and contamination were verified with CheckM version 1.0.7 (Parks *et al*. 2015).

Pairwise ANIb was analysed with JSpecies (Richter and Rosselló-Móra, 2009). dDDH was determined with the Genome-to-Genome distance Calculator (Meier-Kolthoff *et al*., 2013). AAI

and phylogenetic tree based on 400 universally conserved protein sequences were produced with PhyloPhlAn (Segata *et al*., 2013) to provide additional information on the relationships between members of the genus *Tepidimonas.*

3.3.6 16S rRNA gene sequences and draft genome accession numbers

The GenBank/EMBL/DDBJ accession numbers for the 16S rRNA gene sequence of strains SPSP-6^T (=CECT 9683^T =LMG 30884^T) and SPSPC-18 (=CECT 9684 =LMG 30885) are MH590702 and MH590703, respectively. Draft genomes accession numbers of strains *T. alkaliphilus* YIM 72238^T (KCTC 52717^T) (VJNB00000000), *T. aquatica* CLN-1^T (DSM 14833^T) (VJNA00000000), *T. fonticaldi* AT-A2^T (LMG 26746^T) (VJOO00000000), *T. ignava* SPS-1037^T (DSM 123034^T) (VJNC00000000), *T. sediminis* YIM 72259^T (NBRC 112410^T) (VJND00000000), *T. taiwanensis* I1-1^T (LMG 22826^T) (VJOM00000000), *T. thermarum* AA-1^T (LMG 23094^T) (VJOL00000000) and strain SPSP-6^T (LMG 30884^T) (VJON00000000) were deposited in GenBank/EMBL/DDBJ.

3.4 Results and Discussion

3.4.1 Hight-quality draft genome sequences

The assembled genomes of the strains of the genus *Tepidimonas* ranged from 2465 kbp for *T. alkaliphilus* strain YIM 72238^T to 3009 kbp for strain *T. fonticaldi* AT-A2^T. The DNA G+C content of genomes ranged from 66.63% for strain SPSP-6^T to 71.83% for *T. sediminis* YIM 72259^T. The completeness of the genomes examined ranged from 98.91% for the draft genomes of strains VT154-175 and AA-1^T and 100% for the draft genome of *T. fonticaldi* AT-A2^T. The genomes of strains of members of the genus *Tepidimonas* had a variable number of rRNA genes ranging from three in *T. alkaliphilus* strain YIM 72238^T to twelve in the genome of strain VT154-175 (Table 3.1).

Table 3.1 Summary of genome sequencing and annotation metrics of members of the genus *Tepidimonas*: strain SPSP-6^T (VJON00000000), *T. alkaliphilus* YIM 72238^T (VJNB00000000), *T. aquatica* CLN-1^T (VJNA00000000), *T. fonticaldi* AT-A2^T (VJOO00000000), strain PL17 (GCF_001675355.1), *T. ignava* SPS-1037^T (VJNC00000000), *T. sediminis* YIM 72259^T (VJND00000000), *T. taiwanensis* I1-1^T (VJOM00000000), strain MB2 (GCF_001481285.1), strain VT154-175 (GCF_000807215.1) and *T. thermarum* AA-1^T (VJOL00000000).

	SPSP-6^T	***Tepidimonas alkaliphilus***	***Tepidimonas aquatica***	***Tepidimonas fonticaldi***	**PL17**	***Tepidimonas ignava***	***Tepidimonas sediminis***	***Tepidimonas taiwanensis***	**MB2**	**VT154-175**	***Tepidimonas thermarum***
Assembled genome size (bp)	2,808,982	2,465,445	2,672,904	3,009,257	2,740,548	2,715,700	2,533,936	2,859,782	2,813,615	2,924,885	2,703,753
DNA G+C content (%)	66.63	69.01	68.55	69.00	69.53	68.79	71.83	68.80	68.80	68.66	68.70
Protein-coding genes	2,634	2,280	2,507	2,758	2,519	2,563	2,337	2,622	2,591	2,658	2,552
Genes with function prediction	2,208	2,049	2,256	2,407	2,290	2,217	2,141	2,310	2,291	2,362	2,260
Ribosomal genes (5S, 16S, 23S)	2, 2, 2	1, 1, 1	2, 2, 2	2, 2, 2	2, 1, 1	2, 2, 2	2, 1, 1	2, 2, 2	2, 2, 2	4, 4, 4	2, 2, 2
Estimated genome completeness (%)	99.14	99.51	99.53	100	99.53	99.53	99.07	99.42	99.42	98.91	98.91
Estimated contamination (%)	1.05	0.00	0.96	0.00	1.17	0.03	0.15	0.03	0.18	0.05	0.47

3.4.2 Insights from the genomes of members of the genus *Tepidimonas*

Genes coding for enzymes involved in the hydrolysis of starch, cyclodextrin and pullulan namely alpha-amylase (EC 3.2.1.1), beta-amylase (EC 3.2.1.2), pullulanase (EC 3.2.1.41) and cyclomaltodextrinase (EC 2.4.1.19), were not identified in the genomes of any strains of members of the genus *Tepidimonas*. Therefore, it should not be possible for these organisms to obtain glucose or maltose from starch that could be taken up in pure culture. The type strain of *T. taiwanensis* has been reported to hydrolyse starch (Chen *et al*., 2006), while another study has reported that starch was not hydrolysed by the same organism, corroborating the absence of starch hydrolysing-enzymes from the genome analyses (Albuquerque *et al*., 2006).

The ability of the type strain of *T. taiwanensis* to grow on glucose and fructose has been reproduced in laboratories that examined these phenotypic characteristics (Albuquerque *et al*., 2006; Chen *et al*., 2013). Likewise, the inability of the other species of this genus to grow on hexoses is also well attested. The genome sequences of the organisms used in this study clarified the likely reasons why the type strain of *T. taiwanensis* is able to use glucose and fructose while the other strains are not (Table 3.2). The genome analysis indicated that glucose and fructose transporters only occur in the type strain of *T. taiwanensis*, as well as strains MB2 and VT154-175, where putative ABC glucose/mannose (*gtsABCD*) and fructose (*frcABC*) transporters were the only two sugar transporters identified. Moreover, we did not identify other transport systems for hexoses, disaccharides or pentoses in the genomes of any of the strains of members of the genus *Tepidimonas*. The gene coding for glucokinase (EC 2.7.1.2) was only identified in the genomes of *T. aquatica* CLN-1^{T}, *T. taiwanensis* I1-1^{T}, strains MB2 and VT154-175. Otherwise, all other genes of the Emden-Meyerhof-Parnas pathway were identified in the genomes of strains of members of the genus *Tepidimonas*.

The genes coding for the enzymes of the pentose-phosphate pathway, specifically glucose-6-phosphate 1-dehydrogenase (EC 1.1.1.49), 6-phosphogluconolactonase (EC 3.1.1.31), 2-dehydro-3-deoxyphosphogluconate aldolase (EC 4.1.2.14) and phosphogluconate dehydratase (EC 4.2.1.12) were also identified in the genomes of *T. taiwanensis* I1-1^{T}, strains MB2 and VT154-175 but were not identified in any of the other genomes. The pentose-phosphate pathway can be predicted to channel intermediates to glyceraldehyde-3-phosphate. Additionally, the gene coding for the enzyme 6-phosphogluconate dehydrogenase (EC 1.1.1.44, EC 1.1.1.343) was not identified in any of the genomes analysed, thus precluding the utilization of the Entner-Doudoroff pathway by all strains. Gluconeogenesis, as expected, was predicted in all strains of members of the genus *Tepidimonas* examined because the key enzyme fructose-1,6-bisphosphatase (EC 3.1.3.11) was identified in all genomes.

Table 3.2 Differential characteristics of members of the genus *Tepidimonas*: strain SPSP-6^{T}, strain SPSPC-18, *T. alkaliphilus* YIM 72238^{T}, *T. aquatica* CLN-1^{T}, *T. fonticaldi* AT-A2^{T}, *T. ignava* SPS-1037^{T}, *T. sediminis* YIM 72259^{T}, *T. taiwanensis* I1-1^{T} and *T. thermarum* AA-1^{T}.

Characteristics	**SPSP-6^{T*}**	**SPSPC-18***	***Tepidimonas alkaliphilus*^{a}**	***Tepidimonas aquatica*^{b,c}**	***Tepidimonas fonticaldi*^{a,d}**	***Tepidimonas ignava*^{b,e}**	***Tepidimonas sediminis*^{a}**	***Tepidimonas taiwanensis*^{b,d,f}**	***Tepidimonas thermarum*^{b,d}**
Temperature for growth (ºC)									
Optimum	50	50	45	50	55	50–55	45–50	55	50–55
Range	25–60	30–57.5	37–55	35–62	37–60	35–65	45–60	35–60	30–57.5
pH for growth									
Optimum	7.5–9.0	7.5–9.0	7.0–9.0	7.5–8.0	7.0	7.5–8.5	6.0–7.0	7.0	7.5–8.5
Range	6.5–9.5	6.5–9.5	6.0–11.0	6.5–9.5	7.0–9.0	6.5–9.5	6.0–9.0	6.0–8.0	6.0–9.5
NaCl for growth (%)									
Optimum	0	0	0.5	0	0.2	0	0.5	0.2	0
Range	0–0.5	0–0.5	0–1	0–2	0–1	0–1	0–1	0–1	0–1
Reduction of NO_3^- to NO_2^-	–	–	–	+	+	–	–	+	$+^{d}/-^{b}$
Assimilation of									
D-glucose	–	–	–	–	–	–	–	+	–
D-fructose	–	–	–	–	–	–	–	+	–
α-ketoglutarate	–	–	nd	+	nd	+	nd	+	–
Malate	–	–	nd	$+^{*}/-^{b}$	–	+	nd	+	–
Citrate	–	–	nd	–	–	–	nd	+	–
Fumarate	–	–	nd	+	nd	+	nd	+	–

Table 3.2 (continued)

Aspartate	+	+	nd	+	–	+	nd	+	–
L-alanine	+	+	–	+	–	+	–	+	+
L-asparagine	+	+	+	+	–	+	–	+	+
L-histidine	–	–	–	–	+	–	+	+	–
L-lysine	+	–	+	–	+	–	+	+	+
L-proline	–	–	+	+	+	+	+	+	+
L-arginine	–	–	–	–	+	–	+	+	–
L-isoleucine	+	+	nd	+	–	+	nd	–	+
L-ornithine	+	–	nd	+	+	+	nd	+	+
L-threonine	–	–	+	–	–	–	–	–	–
DNA G+C content (mol%) (HPLC method)	nd	nd	68.9	68.6	70.1	69.7	71.6	68.1	67.9
DNA G+C content (%) (genome sequencing)	66.6	nd	69.0	68.6	69.0	68.8	71.8	68.8	68.7

+, positive; –, negative; nd, not determined. Strains SPSP-6^T, SPSPC-18, *T. thermarum* AA-1^T, *T. ignava* SPS-1037^T, *T. aquatica* CLN-1^T, *T. taiwanensis* I1-1^T and *T. fonticaldi* AT-A2^T assimilate succinate, L-glutamate and L-glutamine. Strains SPSP-6^T, SPSPC-18, *T. thermarum* AA-1^T, *T. ignava* SPS-1037^T, *T. aquatica* CLN-1^T and *T. taiwanensis* I1-1^T assimilate lactate, pyruvate, acetate, but do not assimilate D-galactose, D-mannose, D-trehalose, D-cellobiose, D-melibiose, D-raffinose, D-ribose, D-xylose, D-arabinose, L-arabinose, L-rhamnose, L-fucose, L-sorbose, sucrose, lactose, maltose, ribitol, xylitol, sorbitol, erythritol, D-mannitol, *myo*-inositol, glycerol, benzoate, formate, glycine, L-methionine, L-serine and valine.

[*]Data from this study.

[a]Data from Habib *et al*. (2018).

[b]Data from Albuquerque *et al*. (2006).

[c]Data from Freitas *et al*. (2003).

[d]Data from Chen *et al*. (2013).

[e]Data from Moreira *et al*. (2000).

[f]Data from Chen *et al*. (2006).

Although the type strain of *T. taiwanensis* grows on glucose and fructose, this strain, like all strains of this genus, does not grow on any other carbohydrates examined, such as mannose, galactose, trehalose, maltose, sucrose, ribose, L-arabinose, xylose or polyols (Table 3.2). We were unable to identify genes in any strains of members of the genus *Tepidimonas* that could channel these carbohydrates to the Emden-Meyerhof-Parnas or the pentose phosphate pathways.

Enzymes of the TCA cycle were identified in all genomes of members of the genus *Tepidimonas*. Genes coding for the enzymes of oxidative phosphorylation were NADH dehydrogenase (EC 1.6.5.11, complex I), succinate dehydrogenase/fumarate reductase (EC 1.3.5.1, complex II), cytochrome bc1 (EC 1.10.2.2, complex III), cytochrome *c* oxidase cbb3-type (EC 1.9.3.1, complex IV) and an F-type ATPase (EC 3.6.3.14, complex V) were identified in all genomes.

Experimental nitrate reduction to nitrite has been observed by the type strains of *T. aquatica*, *T. fonticaldi*, *T. thermarum* and *T. taiwanensis* but not by the type strain of *T. ignava* and strain SPSP-6^T (Chen *et al*., 2013). The strains of the members of the genus *Tepidimonas* have variable genes involved in nitrogen metabolism (Table 3.3). For example, the type strain of *T. fonticaldi* possesses the most complete set of genes of the species of this genus being predicted to be capable of reducing nitrate to nitrous oxide via the products of *norB* (nitric oxide reductase, large subunit) and *norC* (nitric oxide reductase, small subunit). The other strains, including the closely related strain PL17, appear to lack genes *norB* and *norC*. Genes coding for nitrate/nitrite transporters *nasA/narK* and the nitrate reductase complex *narGHIJ* were identified in the genomes of *T. aquatica* CLN-1^T, *T. fonticaldi* AT-A2^T, strain PL17, *T. taiwanensis* I1-1^T, strain VT154-175 and *T. thermarum* AA-1^T, but not in *T. alkaliphilus* YIM 72238^T, strain SPSP-6^T and strain MB2. *Tepidimonas ignava* SPS-1037^T and *T. sediminis* YIM 72259^T have genes coding for the nitrate/nitrite transporters *nasA/narK* but the nitrate reductase complex *narGHIJ* was not identified in the genome sequences. Nitrate did not appear to be reduced by *T. thermarum* AA-1^T experimentally in one study but has been reported to reduce nitrate in another study (Albuquerque *et al*., 2006; Chen *et al*., 2013). However, the genome predicts that nitrate should be reduced to nitrite because this organism possesses *narGHIJ*. The only gene involved in the reduction of nitrate identified in strain SPSP-6^T and *T. alkaliphilus* YIM 72238^T was *NirB* (nitrite reductase, NADH-dependent large subunit) (Table 3.2 and 3.3).

With the exception of strains PL17, MB2 and VT154-175 whose phenotypic characteristics are not available, and the type strains of *T. sediminis* and *T. alkaliphilus* where thiosulfate oxidation was not examined, all other strains of *Tepidimonas* oxidize thiosulfate to sulfate experimentally. However, all genomes predict that thiosulfate is oxidized to sulfate via the sox pathway, namely *soxXABCDYZ* genes.

The three subunit orthologs of the Tripartite ATP-Independent Periplasmic transporter (TRAP) that transport the C_4-dicarboxylates malate/fumarate (DctM, DctP, DctQ) were identified in *T. aquatica* CLN-1^T, *T. ignava* SPS-1037^T, *T. alkaliphilus* YIM 72238^T and *T. taiwanensis* I1-1^T (Valentini *et al*., 2011). These organisms use malate and fumarate as single carbon sources (Table 3.2). The other type strains appear to have only the genes for the DctM and DctP components

or the DctM component alone. The type strain of *T. thermarum* and strain SPSP-6^T do not grow on malate or fumarate. The type strain of *T. fonticaldi* does not grow on malate but growth on fumarate was not tested, while the type strain of *T. sediminis* and *T. alkaliphilus* were not tested for the utilization of malate or fumarate (Table 3.2).

We only identified the genes for the Tripartite Tricarboxylate ABC System transporter for citrate composed of three subunits (TctA, TctB and TctC) in the genome of the type strain of *T. taiwanensis* and strain MB2. Strain VT154-175 has two components (TctB and TctC), while the other *Tepidimonas* strains appear to have only one component (Brocker *et al*., 2009). The type strain of *T. taiwanensis* is the only organism, among those examined that uses citrate for growth (Table 3.2).

Table 3.3 Genes involved in nitrate/nitrite metabolism in genomes of members of the genus *Tepidimonas*: strain SPSP-6^T (VJON00000000), *T. alkaliphilus* YIM 72238^T (VJNB00000000), *T. aquatica* CLN-1^T (VJNA00000000), *T. fonticaldi* AT-A2^T (VJOO00000000), strain PL17 (GCF_001675355.1), *T. ignava* SPS-1037^T (VJNC00000000), *T. sediminis* YIM 72259^T (VJND00000000), *T. taiwanensis* I1-1^T (VJOM00000000), strain MB2 (GCF_001481285.1), strain VT154-175 (GCF_000807215.1) and *T. thermarum* AA-1^T (VJOL00000000).

	SPSP-6^T	***Tepidimonas alkaliphilus***	***Tepidimonas aquatica***	***Tepidimonas fonticaldi***	**PL17**	***Tepidimonas ignava***	***Tepidimonas sediminis***	***Tepidimonas taiwanensis***	**MB2**	**VT154-175**	***Tepidimonas thermarum***
narGHIJ	nd	nd	+	+	+	nd	nd	+	nd	+	+
nirB	+	+	+	+	+	+	+	+	+	+	+
nirD	nd	nd	+	+	+	nd	+	nd	nd	nd	nd
nasA	nd	nd	+	+	+	+	+	+	nd	+	+
nirK	nd	nd	+	+	+	nd	nd	+	+	+	+
nirS	nd	nd	nd	+	nd	nd	nd	nd	nd	nd	nd
norB	nd	nd	nd	+	nd	nd	nd	nd	nd	nd	nd
norC	nd	nd	nd	+	nd	nd	nd	nd	nd	nd	nd

nd, not detected; *narGHIJ*, nitrate reductase complex; *nirB*, nitrite reductase (NADH dependent subunit); *nirD*, nitrite reductase subunit; *nasA*, nitrite reductase, large subunit; *nirK*, nitrite reductase, small subunit; *nirS*, nitrite reductase, monomer; *norB*, nitric oxide reductase, large subunit; *norC*, nitric oxide reductase, small subunit.

3.4.3 Phylogenetic analysis of 16S rRNA sequences and 400 conserved genes

The pairwise 16S rRNA gene sequence similarity determined between strains SPSP-6^T and SPSC-18 was 100%. Strain SPSP-6^T shared a pairwise 16S rRNA gene sequence similarity of 98.07%, 98.38% and 98.44% with the type strains of *Tepidimonas ignava*, *T. aquatica* and *T. taiwanensis* respectively (Figure 3.1 and Table 3.4). The sequence similarity between *T. fonticaldi* AT-A2^T and strain PL17 was 99.86%, indicating an extremely close relationship between the two organisms. A close relationship of *T. fonticaldi* AT-A2^T with "*T. arfidensis*" SMC-6271 of 99.58% was also noted. The 16S rRNA gene sequence analysis also indicated *T. taiwanensis* I1-1^T to be closely related to strains MB2 and VT154-175 with sequence similarities of 99.93 and 99.58%, respectively (Table 3.4). The phylogenetic results of the 16S rRNA gene analysis indicate that SPSP-6^T is located within a cluster comprising the type strains of *T. ignava*, *T. aquatica*, *T. taiwanensis*, *T. alkaliphilus* and *T. sediminis*. However, SPSP-6^T is most closely related to the type strains of *T. ignava*, *T. aquatica* and *T. taiwanensis* (Figure 3.1 and 3.2). The phylogenetic results based on 400 conserved genes was consistent with the 16S rRNA gene sequence results (Figure 3.3), corroborating the phylogenetic relations observed for strain SPSP-6^T within the genus *Tepidimonas*. The phylogenetic analysis of 400 conserved genes also showed *T. taiwanensis* I1-1^T, strains MB2 and VT154-175 to be very closely related to each other indicating that the three strains belong to one species. Strain PL17 and *T. fonticaldi* AT-A2^T should also probably be regarded as representing one species.

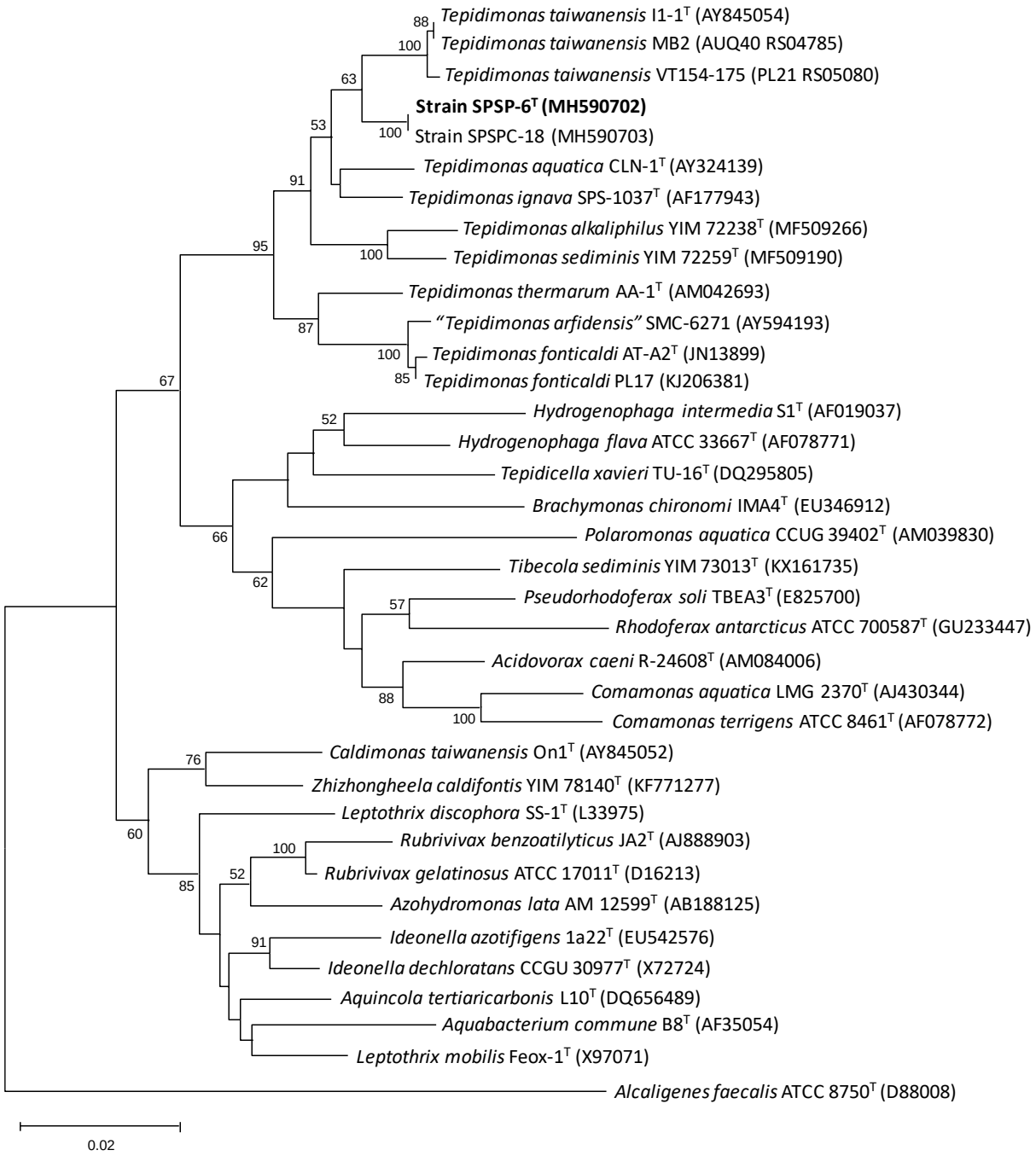

Figure 3.1 Phylogenetic reconstruction based on 16S rRNA genes of strains of members of the genus *Tepidimonas* and the type strains of genera belonging to family *Comamonadaceae* using the NJ algorithm. The numbers at branching points represent bootstrap values from 1000 replications. Bar, 0.02 substitutions per nucleotide position. The tree was rooted using the sequence of *Alcaligenes faecalis* ATCC 8750^T (D88008).

Table 3.4 Pairwise similarity values determined between the 16S rRNA sequence gene of members of the genus *Tepidimonas* and type strains of *Tepidicella xavieri* and *Acidovorax caeni* of the family *Comamonadaceae*: strain SPSP-6^T (MH590702), strain SPSPC-18 (MH590703), *T. aquatica* CLN-1^T (AY324139), *T. ignava* SPS-1037^T (AF177943), *T. taiwanensis* I1-1^T (AY845054), strain MB2 (NZ_LOQE01000009) (AUQ40_RS04785), strain VT154-175 (NZ_JTKY01000044) (PL21 RS05080), *T. thermarum* AA1^T (AM042693), *T. alkaliphilus* YIM 72238^T (MF509266), *T. sediminis* YIM 72259^T (MF509190), "*T. arfidensis"* SMC-6271 (AY594193), strain PL17 (KF206381), *T. fonticaldi* AT-A2^T (JN713899), *Tepidicella xavieri* TU-16^T (DQ295805), *Acidovorax caeni* R-24608^T (AM084006).

	SPSP-6^T	**SPSPC-18**	***Tepidimonas aquatica***	***Tepidimonas ignava***	***Tepidimonas taiwanensis***	**MB2**	**VT154-175**	***Tepidimonas thermarum***	***Tepidimonas alkaliphilus***	***Tepidimonas sediminis***	***"Tepidimonas arfidensis"***	**PL17**	***Tepidimonas fonticaldi***	***Tepidicella xavieri***	***Acidovorax caeni***
SPSP-6^T	100.00														
SPSPC-18	100.00	100.00													
T. aquatica	98.38	98.37	100.00												
T. ignava	98.07	97.94	97.98	100.00											
T. taiwanensis	98.44	98.22	97.20	97.76	100.00										
MB2	98.44	98.29	97.31	97.85	99.93	100.00									
VT154-175	98.59	98.29	96.97	97.58	99.58	99.67	100.00								
T. thermarum	96.88	96.73	95.76	97.04	96.71	96.96	97.29	100.00							
T. alkaliphilus	97.18	97.00	95.64	97.64	97.39	97.38	97.38	96.75	100.00						
T. sediminis	97.25	97.14	96.25	97.63	97.10	97.09	97.02	96.60	98.62	100.00					
"T. arfidensis"	96.97	96.84	96.46	95.73	96.31	96.37	96.02	97.42	95.12	95.03	100.00				
PL17	97.20	97.03	96.78	95.86	96.43	96.53	96.19	97.55	95.25	95.17	99.65	100.00			
T. fonticaldi	97.12	97.03	96.77	95.85	96.36	96.48	96.13	97.51	95.12	95.03	99.58	99.86	100.00		
Tepidicella xavieri	94.42	94.38	94.66	95.47	94.61	94.73	94.59	95.61	94.09	94.14	94.82	9497	94.96	100.00	
Acidovorax caeni	93.23	93.24	93.27	93.28	93.35	92.82	92.49	94.57	93.37	93.35	95.25	95.44	95.37	94.12	100.00

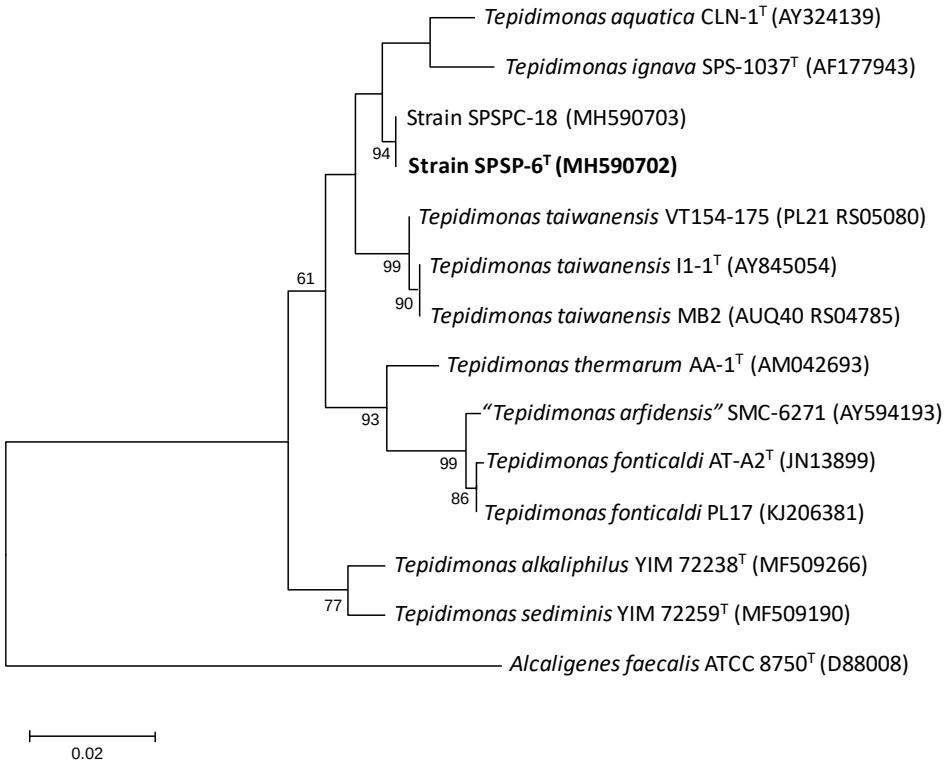

Figure 3.2 Phylogenetic reconstruction based on 16S rRNA genes of strains of members of the genus *Tepidimonas* using the ML algorithm. The numbers at branching points represent bootstrap values from 1000 replications. Bar, 0.02 substitutions per nucleotide position. The tree was rooted using the sequence of *Alcaligenes faecalis* ATCC 8750^T (D88008).

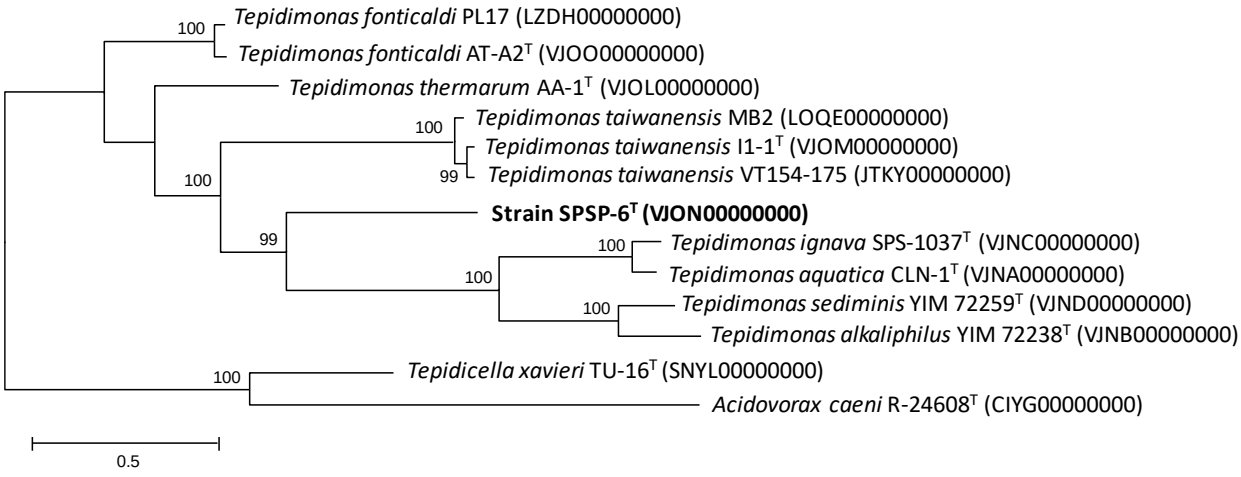

Figure 3.3 Phylogenetic tree reconstruction of members of the genus *Tepidimonas* based on a set of 400 conserved bacterial genes. GenBank accession numbers of the genomes are given in parentheses. Bootstrap values were calculated based on 1000 replicates. The scale bar indicates the number of amino acid substitutions per site.

3.4.4 Comparative genome analysis

Additionally, the results of the genome comparisons, namely ANIb, AAI and dDDH corroborated the results of the 16S rRNA sequence analysis regarding the distinct species nature of the lineages (Table 3.5, 3.6 and 3.7). Based on a threshold value of 95–96% for species delineation (Stackebrandt *et al*., 2002; Richer and Rosselló-Morá 2009), the ANIb values indicate that SPSP-6^T, with ANIb values of about 80% with other species, constitutes a separate species of the genus *Tepidimonas*. The ANIb value for *T. fonticaldi* AT-A2^T and strain PL17 of 97.46% indicates a very close relationship between these strains. The same is true for *T. taiwanensis* I1-1^T, strains MB2 and VT154-175 that share ANIb values of 97.31 to 98.48%. In addition to the genomic values, these are the only currently known strains that have genes for the transport of glucose and fructose, and utilize these hexoses for growth. On the basis of the ANIb values, the type strains of *T. aquatica* and *T. ignava* are closely related (93.22%). Nevertheless, this value is below the threshold value for delineation of species (Table 3.5).

The AAI values, generally taken to have a cutoff value of around 70% to delineate genera (Konstantinidis and Tiedje, 2005b; Luo *et al*., 2014), indicate that the strains of *Tepidimonas*, including strain SPSP-6^T, belong to one genus because of higher AAI values (Table 3.6). The genome of *Tepidicella xavieri* has high AAI values with the species of *Tepidimonas* (67.59 to 70.85%) indicating that the organisms of the two genera are related, but within a transitional zone of AAI values, making it difficult to have an opinion on the classification of the sole strain of *Tepidicella xavieri* from the genomic analysis. However, the phylogeny obtained for the 16S rRNA gene and for 400 conserved genes sequence analysis indicates that *Tepidicella xavieri* is not closely related to the species of *Tepidimonas* (Chun *et al*., 2018; França *et al*., 2006).

The genomic-based dDDH estimates have values of 79.8% to 87.20% between *T. taiwanensis* strains I1-1^T, MB2 and VT154-175 (Table 3.7). These results are above the reference dDDH value of about 70% to delineate separate species by Stackebrandt *et al*. (2002) leading us to the opinion that the three strains represent members of the species *T. taiwanensis*. Moreover, these three strains possess an ABC glucose/mannose transporter that all other strains seem to lack. The high dDDH value of 80.1% between *T. fonticaldi* AT-A2^T and strain PL17 also supports the view that these two strains represent the same species. The low dDDH values between other organisms of the genus *Tepidimonas*, notably strain SPSP-6^T sharing no more than 25.1%, support the view that the organisms examined represent distinct species of the genus *Tepidimonas* (Table 3.7).

The 16S rRNA gene sequence analysis, as well as the genomic data, circumscribes all type strains of the genus *Tepidimonas*, as well as strains PL17, MB2 and VT154-175 to the genus *Tepidimonas*. Moreover, strains MB2 and VT154-175 appear, based on the presence of glucose/ mannose and fructose transporters and the genomic results, to represent *Tepidimonas taiwanensis*, while the close phylogenetic and genomic results ascribe strain PL17 to the species *Tepidimonas fonticaldi*.

Table 3.5 ANIb (%) values between genomes of members of the genus *Tepidimonas* and type strains of *Tepidicella xavieri* and *Acidovorax caeni* (the percentage of aligned nucleotides is given between brackets): strain SPSP-6^T (VJON00000000), *T. alkaliphilus* YIM 72238^T (VJNB00000000), *T. aquatica* CLN-1^T (VJNA00000000), *T. fonticaldi* AT-A2^T (VJOO00000000), strain PL17 (GCF_001675355.1), *T. ignava* SPS-1037^T (VJNC00000000), *T. sediminis* YIM 72259^T (VJND00000000), *T. taiwanensis* I1-1^T (VJOM00000000), strain MB2 (GCF_001481285.1), strain VT154-175 (GCF_000807215.1), *T. thermarum* AA-1^T (VJOL00000000), *Tepidicella xavieri* TU-16^T (GCF_004363315.1), *Acidovorax caeni* R-24608^T (GCF_001298675.1).

	SPSP-6^T	***Tepidimonas alkaliphilus***	***Tepidimonas aquatica***	***Tepidimonas fonticaldi***	**PL17**	***Tepidimonas ignava***	***Tepidimonas sediminis***	***Tepidimonas taiwanensis***	**MB2**	**VT154-175**	***Tepidimonas thermarum***	***Tepidicella xavieri***	***Acidovorax caeni***
SPSP-6^T	–	77.99 [47.98]	78.82 [50.92]	80.65 [54.62]	80.15 [54.11]	79.84 [54.32]	79.17 [51.41]	79.18 [53.49]	79.43 [52.57]	79.07 [52.92]	79.42 [55.36]	75.11 [40.96]	72.94 [28.66]
T. alkaliphilus	77.99 [55.16]	–	81.13 [68.68]	81.09 [63.44]	80.88 [65.69]	80.90 [68.25]	87.96 [75.53]	78.87 [60.48]	79.12 [59.76]	78.86 [60.54]	79.59 [63.23]	74.86 [48.8]	72.89 [34.46]
T. aquatica	78.80 [54.28]	81.01 [63.78]	–	81.57 [67.81]	81.36 [66.74]	93.27 [79.68]	82.37 [69.13]	79.45 [60.25]	80.06 [60.00]	79.50 [60.25]	79.65 [66.93]	75.80 [51.47]	72.85 [34.36]
T. fonticaldi	80.57 [51.89]	80.88 [52.74]	81.46 [60.00]	–	97.46 [79.29]	80.16 [58.86]	82.52 [58.84]	81.55 [58.56]	81.95 [59.31]	81.53 [59.08]	82.67 [64.58]	77.07 [49.37]	73.81 [34.22]
PL17	80.05 [56.05]	80.68 [59.59]	81.29 [64.90]	97.61 [86.05]	–	80.08 [63.40]	82.51 [66.27]	81.45 [64.28]	81.9 [63.72]	81.48 [64.56]	82.91 [71.07]	76.56 [54.71]	73.72 [35.81]
T. ignava	79.58 [57.82]	80.89 [62.39]	93.22 [78.48]	80.00 [65.17]	80.06 [64.31]	–	81.97 [71.32]	79.74 [58.63]	79.88 [58.09]	79.53 [58.38]	79.84 [66.53]	75.36 [50.14]	72.95 [32.68]
T. sediminis	79.12 [57.58]	87.83 [74.00]	82.26 [73.04]	82.51 [70.02]	82.63 [71.86]	81.92 [76.02]	–	80.42 [63.22]	80.63 [63.02]	80.43 [63.52]	80.75 [68.69]	75.30 [53.97]	73.33 [36.04]

Table 3.5 (continued)

T. taiwanensis	79.31 [52.32]	78.97 [52.41]	79.48 [57.09]	81.55 [61.86]	81.45 [62.56]	79.86 [55.83]	80.56 [56.23]	–	97.53 [85.43]	98.48 [94.19]	80.87 [59.48]	75.56 [45.37]	73.27 [31.42]
MB2	79.24 [53.55]	79.10 [52.97]	79.97 [57.61]	81.86 [63.76]	81.69 [62.91]	79.94 [56.56]	80.73 [54.44]	97.31 [86.32]	–	96.87 [85.92]	80.93 [60.04]	75.87 [47.01]	73.52 [30.95]
VT154-175	79.09 [52.26]	78.87 [52.21]	79.64 [56.28]	81.73 [61.70]	81.46 [61.98]	79.73 [54.03]	80.68 [54.95]	98.48 [93.28]	97.10 [84.11]	–	80.89 [58.67]	75.69 [45.59]	73.52 [30.95]
T. thermarum	79.36 [58.66]	79.50 [58.72]	79.83 [65.54]	82.77 [72.32]	82.98 [72.59]	79.96 [66.69]	80.86 [65.24]	81.02 [62.51]	81.27 [61.46]	81.00 [62.64]	–	75.81 [51.2]	73.36 [34.2]
Tepidicella xavieri	75.04 [41.69]	74.80 [43.63]	75.72 [48.51]	77.11 [52.64]	76.74 [53.21]	75.31 [48.80]	75.53 [48.80]	75.56 [47.02]	76.16 [47.39]	75.57 [46.85]	75.79 [49.44]	–	73.44 [34.43]
Acidovorax caeni	72.60 [20.31]	72.57 [21.26]	72.70 [22.40]	73.63 [24.63]	73.63 [23.89]	72.68 [22.35]	73.23 [22.48]	73.13 [22.14]	73.24 [21.93]	73.22 [22.35]	73.28 [22.49]	73.63 [23.21]	–

Table 3.6 AAI (%) values between genomes of members of the genus *Tepidimonas* and type strains of *Tepidicella xavieri* and *Acidovorax caeni*: strain SPSP-6^T (VJON00000000), *T. alkaliphilus* YIM 72238^T (VJNB00000000), *T. aquatica* CLN-1^T (VJNA00000000), *T. fonticaldi* AT-A2^T (VJOO00000000), strain PL17 (GCF_001675355.1), *T. ignava* SPS-1037^T (VJNC00000000), *T. sediminis* YIM 72259^T (VJND00000000), *T. taiwanensis* I1-1^T (VJOM00000000), strain MB2 (GCF_001481285.1), strain VT154-175 (GCF_000807215.1), *T. thermarum* AA-1^T (VJOL00000000), *Tepidicella xavieri* TU-16^T (GCF_004363315.1), *Acidovorax caeni* R-24608^T (GCF_001298675.1).

	SPSP-6^T	***Tepidimonas alkaliphilus***	***Tepidimonas aquatica***	***Tepidimonas fonticaldi***	**PL17**	***Tepidimonas ignava***	***Tepidimonas sediminis***	***Tepidimonas taiwanensis***	**MB2**	**VT154-175**	***Tepidimonas thermarum***	***Tepidicella xavieri***	***Acidovorax caeni***
SPSP-6^T	–	73.38	74.25	76.62	76.04	74.99	74.42	75.94	75.76	75.82	75.86	68.58	63.34
T. alkaliphilus	73.38	–	76.51	74.84	74.72	76.27	82.07	72.82	73.10	72.79	74.30	67.59	62.31
T. aquatica	74.25	76.51	–	75.69	75.71	85.68	77.76	74.40	74.61	74.46	75.12	69.00	62.66
T. fonticaldi	76.62	74.84	75.69	–	88.33	75.41	76.97	77.34	77.84	77.56	79.52	70.85	63.47
PL17	76.04	74.72	75.71	88.33	–	75.39	76.89	77.59	77.95	77.70	79.43	70.49	63.24
T. ignava	74.99	76.27	85.68	75.41	75.39	–	77.35	74.39	74.33	74.42	75.07	68.09	62.67
T. sediminis	74.42	82.07	77.76	76.97	76.89	77.35	–	74.16	74.24	74.18	75.47	68.12	62.84
T. taiwanensis	75.94	72.82	74.40	77.34	77.59	74.39	74.16	–	88.66	89.96	77.52	69.21	63.16
MB2	75.76	73.10	74.61	77.84	77.95	74.33	74.24	88.66	–	88.31	77.48	69.89	63.15
VT154-175	75.82	72.79	74.46	77.56	77.70	74.42	74.18	89.96	88.31	–	77.55	69.38	63.16
T. thermarum	75.86	74.30	75.12	79.52	79.43	75.07	75.47	77.52	77.48	77.55	–	69.72	63.26
Tepidicella xavieri	68.58	67.59	69.00	70.85	70.49	68.09	68.12	69.21	69.89	69.38	69.72	–	64.35
Acidovorax caeni	63.34	62.31	62.66	63.47	63.24	62.67	62.84	63.16	63.15	63.16	63.26	64.35	–

Table 3.7 dDDH (%) values between genomes of members of the genus *Tepidimonas* and type strains of *Tepidicella xavieri* and *Acidovorax caeni*: strain SPSP-6^T (VJON00000000), *T. alkaliphilus* YIM 72238^T (VJNB00000000), *T. aquatica* CLN-1^T (VJNA00000000), *T. fonticaldi* AT-A2^T (VJOO00000000), strain PL17 (GCF_001675355.1), *T. ignava* SPS-1037^T (VJNC00000000), *T. sediminis* YIM 72259^T (VJND00000000), *T. taiwanensis* I1-1^T (VJOM00000000), strain MB2 (GCF_001481285.1), strain VT154-175 (GCF_000807215.1), *T. thermarum* AA-1^T (VJOL00000000), *Tepidicella xavieri* TU-16^T (GCF_004363315.1), *Acidovorax caeni* R-24608^T (GCF_001298675.1).

	SPSP-6^T	***Tepidimonas alkaliphilus***	***Tepidimonas aquatica***	***Tepidimonas fonticaldi***	**PL17**	***Tepidimonas ignava***	***Tepidimonas sediminis***	***Tepidimonas taiwanensis***	**MB2**	**VT154-175**	***Tepidimonas thermarum***	***Tepidicella xavieri***	***Acidovorax caeni***
SPSP-6^T	–	22.4	22.9	25.1	23.9	24.2	22.9	23.8	23.8	23.7	23.2	21.3	19.2
T. alkaliphilus	22.4	–	24.2	25.1	24.9	23.9	34.8	23.1	23.3	23.2	22.9	20.1	19.2
T. aquatica	22.9	24.2	–	25.1	24.5	53.0	25.4	23.1	23.8	23.2	22.9	21.1	19.3
T. fonticaldi	25.1	25.1	25.1	–	80.1	23.2	26.2	25.5	26.0	25.5	26.1	21.7	20.4
PL7	23.9	24.9	24.5	80.1	–	22.9	26.0	25.2	25.6	25.2	26.2	21.2	19.8
T. ignava	24.2	23.9	53.0	23.2	22.9	–	25.1	23.4	23.6	23.2	22.8	20.6	19.3
T. sediminis	22.9	34.8	25.4	26.2	26.0	25.1	–	24.0	24.1	24.0	23.6	20.2	19.4
T. taiwanensis	23.8	23.1	23.1	25.5	25.2	23.4	24.0	–	79.8	87.2	24.8	21.0	19.8
MB2	23.8	23.3	23.8	26.0	25.6	23.6	24.1	79.8	–	76.0	24.8	21.8	19.8
VT154-175	23.7	23.2	23.2	25.5	25.2	23.2	24.0	87.2	76.0	–	24.8	21.2	19.8
T. thermarum	23.2	22.9	22.9	26.1	26.2	22.8	23.6	24.8	24.8	24.8	–	20.4	19.7
Tepidicella xavieri	21.3	20.1	21.1	21.7	21.2	20.6	20.2	21.0	21.8	21.2	20.4	–	20.4
Acidovorax caeni	19.2	19.2	19.3	20.4	19.8	19.3	19.4	19.8	19.8	19.8	19.7	20.4	–

3.4.5 Phenotypic and chemotaxonomic characteristics

A small number of phenotypic characteristics of the type strains of the species of the genus *Tepidimonas* distinguish the strains from each other (Table 3.2). The fatty acid composition of the strains was obtained after the organisms were grown in Degryse medium 162 agar plates and R2A for 24 h at 50ºC. These results indicated that the medium influenced the fatty acid composition to a large extent (Table 3.8 and 3.9). For example, $C_{17:0}$ cyclo was not detected in *T. ignava* SPS-1037^T and *T. sediminis* YIM 72259^T grown on Degryse medium 162 but reached levels of 6.7 and 7.3%, respectively, when they were grown on R2A agar. Nevertheless, the major fatty acids of all strains were $C_{16:0}$, summed feature 3 ($C_{16:1}$ ω6*c* and/or $C_{16:1}$ ω7*c*) and in some cases, $C_{17:0}$ cyclo and summed feature 8 ($C_{18:1}$ ω6*c* and/or $C_{18:1}$ ω6*c*). However, there were differences in the concentration of these fatty acids among the type strains. For example, the combination of $C_{17:0}$ cyclo and $C_{17:0}$, after growth of the organisms on Degryse medium 162 and R2A, can distinguish strains SPSP-6^T and SPSPC-18 from the other strains of species of this genus.

Many of the validly described prokaryotic species are only based on a few distinctive phenotypic characteristics that could represent interspecific diversity, since these novel organisms are proposed on the basis of the description of one strain alone. The species of the genus *Tepidimonas* are an example of these considerations since strains SPSP-6^T and SPSPC-18 have identical 16S rRNA sequences but have slightly different phenotypic and fatty acid characteristics.

Table 3.8 Fatty acid composition of the species of the genus *Tepidimonas* grown on Degryse medium 162 agar plates at 50ºC for 24h: strain SPSP-6^T, strain SPSPC-18, *T. alkaliphilus* YIM 72238^T, *T. aquatica* CLN-1^T, *T. fonticaldi* AT-A2^T, *T. ignava* SPS-1037^T, *T. sediminis* YIM 72259^T, *T. taiwanensis* I1-1^T and *T. thermarum* AA-1^T.

Fatty acids	**ECL**	**SPSP-6^T**	**SPSPC-18**	***Tepidimonas alkaliphilus***	***Tepidimonas aquatica***	***Tepidimonas fonticaldi***	***Tepidimonas ignava***	***Tepidimonas sediminis***	***Tepidimonas taiwanensis***	***Tepidimonas thermarum***
$C_{8:0}$ 3-OH	9.392	3.2 ± 0.7	2.1 ± 0.1	3.3 ± 0.4	3.2 ± 0.3	3.8 ± 0.3	2.5 ± 0.6	3.1 ± 0.1	3.7 ± 0.5	2.4 ± 0.5
$C_{15:1}$ ω6*c*	14.856	0.7 ± 0.1	tr	0.6 ± 0.1	tr	tr	7.0 ± 0.6	2.6 ± 0.3	0.7 ± 0.1	2.4 ± 0.5
$C_{15:0}$	15.000	1.5 ± 0.2	0.5 ± 0.2	1.6 ± 0.2	0.7 ± 0.1	0.7 ± 0.1	9.0 ± 0.7	8.3 ± 0.5	1.0 ± 0.3	2.9 ± 0.3
Summed feature 3	15.822	30.8 ± 2.2	27.7 ± 4.1	21.2 ± 1.2	26.7 ± 1.9	17.5 ± 1.4	29.6 ± 1.2	15.5 ± 0.4	31.8 ± 1.5	37.1 ± 0.9
$C_{16:0}$	16.000	39.1 ± 2.1	44.5 ± 5.1	39.9 ± 2.0	37.5 ± 1.2	38.6 ± 1.3	19.2 ± 0.6	25.4 ± 0.9	38.7 ± 1.2	35.3 ± 1.1
$C_{17:1}$ ω8*c*	16.792	–	–	–	–	–	1.1 ± 0.1	1.1 ± 0.1	–	–
$C_{17:1}$ ω6*c*	16.860	–	–	–	–	–	3.8 ± 0.2	4.8 ± 0.2	–	–
$C_{17:0}$ cyclo	16.888	2.7 ± 0.2	4.5 ± 2.0	2.2 ± 0.1	22.0 ± 3.6	13.8 ± 0.5	–	–	5.8 ± 0.5	6.1 ± 0.7
$C_{17:0}$	17.000	2.9 ± 0.2	1.2 ± 0.1	6.8 ± 0.4	2.4 ± 0.2	3.9 ± 0.3	16.5 ± 1.2	17.7 ± 0.5	2.2 ± 0.2	3.5 ± 0.4
Summed feature 8	17.823	8.7 ± 0.6	12.2 ± 0.9	15.0 ± 0.7	–	14.0 ± 0.7	6.2 ± 0.5	11.0 ± 0.6	13.0 ± 0.7	6.8 ± 0.5
$C_{18:0}$	18.000	2.3 ± 0.3	1.7 ± 0.9	3.2 ± 0.3	2.0 ± 0.2	2.8 ± 0.2	0.7 ± 0.1	3.1 ± 0.1	1.2 ± 0.2	tr
$C_{18:1}$ ω7*c* 11-methyl	18.081	tr	1.0 ± 0.3	4.0 ± 0.3	–	–	2.0 ± 0.2	2.2 ± 0.1	0.7 ± 0.1	1.0 ± 0.2
$C_{18:0}$ 12-methyl	18.430	tr	–	–	1.4 ± 0.3	0.9 ± 0.1	–	–	0.5 ± 0.1	tr
Summed feature 7	18.846	1.4 ± 0.2	1.5 ± 0.2	1.0 ± 0.1	–	–	tr	3.4 ± 0.2	–	–
$C_{19:0}$ cyclo ω8*c*	18.902	–	tr	tr	–	1.7 ± 0.2	–	tr	tr	–

Results are the percentage of the total fatty acids. ±, results are the mean plus the standard deviation of two to four analyses; values for fatty acids present at less than 0.5% in all strains are not shown; tr, trace (< 0.5%); –, not detected; ECL, equivalent chain length. A summed feature represents groups of two or three fatty acids that could not be separated by GLC with the MIDI System: summed feature 3 comprises $C_{16:1}$ ω7*c* and/or $C_{16:1}$ ω6*c* and/or iso-$C_{15:0}$ 2-OH; summed feature 8 comprises $C_{18:1}$ ω7*c* and/or $C_{18:1}$ ω6*c*; summed feature 7 comprises unknown 18.846 and/or $C_{19:1}$ ω6*c*.

Table 3.9 Fatty acid composition of the species of the genus *Tepidimonas* grown on R2A medium at 50ºC for 24h: strain SPSP-6^T, strain SPSPC-18, *T. alkaliphilus* YIM 72238^T, *T. aquatica* CLN-1^T, *T. fonticaldi* AT-A2^T, *T. ignava* SPS-1037^T, *T. sediminis* YIM 72259^T, *T. taiwanensis* I1-1^T and *T. thermarum* AA-1^T.

Fatty acids	**ECL**	**SPSP-6^T**	**SPSPC-18**	***Tepidimonas alkaliphilus***	***Tepidimonas aquatica***	***Tepidimonas fonticaldi***	***Tepidimonas ignava***	***Tepidimonas sediminis***	***Tepidimonas taiwanensis***	***Tepidimonas thermarum***
$C_{8:0}$ 3-OH	9.392	3.5 ± 0.3	2.6 ± 0.4	2.6 ± 0.2	2.5 ± 0.6	4.2 ± 0.6	3.2 ± 0.6	2.8 ± 0.1	3.5 ± 0.5	3.5 ± 0.6
$C_{15:1}$ ω6*c*	14.856	tr	–	tr	tr	tr	2.5 ± 0.5	0.8 ± 0.1	–	1.0 ± 0.1
$C_{15:0}$	15.000	0.8 ± 0.1	tr	1.8 ± 0.1	0.5 ± 0.1	1.4 ± 0.1	6.6 ± 0.7	6.3 ± 0.3	0.8 ± 0.1	3.6 ± 0.1
$C_{16:0}$ iso	15.627	–	–	–	–	0.8 ± 0.1	–	–	–	–
Summed feature 3	15.822	24.8 ± 0.4	15.7 ± 0.7	13.5 ± 0.3	12.3 ± 0.8	15.2 ± 1.0	20.4 ± 1.9	9.0 ± 0.4	14.5 ± 0.6	16.7 ± 0.8
$C_{16:0}$	16.000	45.4 ± 1.0	50.6 ± 0.8	39.9 ± 0.9	45.9 ± 1.5	38.2 ± 1.0	29.3 ± 1.0	30.1 ± 0.9	45.8 ± 1.2	40.9 ± 1.1
Unknown 16.090	16.090	–	–	–	1.4 ± 0.3	0.6 ± 0.3	–	–	tr	tr
$C_{17:1}$ ω8*c*	16.792	–	–	0.6 ± 0.1	–	–	0.9 ± 0.1	0.8 ± 0.1	–	–
$C_{17:0}$ cyclo	16.888	6.1 ± 0.4	12.4 ± 0.4	3.5 ± 0.1	25.3 ± 2.0	17.5 ± 0.8	6.7 ± 0.4	7.3 ± 0.4	11.5 ± 0.4	12.6 ± 0.5
$C_{17:0}$	17.000	1.9 ± 0.2	0.6 ± 0.1	8.5 ± 0.4	4.6 ± 0.6	4.9 ± 0.6	14.9 ± 2.8	15.9 ± 0.5	2.8 ± 0.2	5.8 ± 0.4
Unknown 17.747	17.747	–	–	–	1.0 ± 0.3	0.6 ± 0.3	–	–	–	–
Summed feature 8	17.823	11.9 ± 0.5	11.2 ± 0.7	17.5 ± 0.5	–	10.7 ± 0.3	7.6 ± 1.0	11.9 ± 0.4	12.8 ± 0.3	9.3 ± 0.5
$C_{18:0}$	18.000	2.3 ± 0.3	1.9 ± 0.2	4.9 ± 0.3	4.4 ± 0.8	2.1 ± 0.4	1.6 ± 0.5	4.6 ± 0.3	3.0 ± 0.3	1.8 ± 0.4
$C_{18:1}$ ω7*c* 11-methyl	18.081	–	0.7 ± 0.1	4.1 ± 0.2	–	–	1.2 ± 0.4	2.4 ± 0.2	1.9 ± 0.3	2.0 ± 0.3
$C_{18:0}$ 12-methyl	18.430	tr	0.7 ± 0.1	–	1.6 ± 0.2	1.0 ± 0.1	–	–	0.7 ± 0.1	0.7 ± 0.1
$C_{19:0}$ iso	18.639	–	–	–	0.8 ± 0.2	tr	–	–	–	–
Unknown 18.814	18.814	–	2.1 ± 0.1	–	–	–	–	–	–	–
Summed feature 7	18.846	2.2 ± 0.2	–	1.9 ± 0.1	–	–	3.6 ± 0.6	4.7 ± 0.2	0.7 ± 0.1	tr
$C_{19:0}$ cyclo ω8*c*	18.902	tr	1.0 ± 0.1	1.0 ± 0.1	–	1.4 ± 0.2	tr	1.9 ± 0.1	1.0 ± 0.1	–

Results are the percentage of the total fatty acids. ±, results are the mean plus the standard deviation of two to four analyses; values for fatty acids present at less than 0.5% in all strains are not shown; tr, trace (< 0.5%); –, not detected; ECL, equivalent chain length. A summed feature represents groups of two or three fatty acids that could not be separated by GLC with the MIDI System: summed feature 3 comprises $C_{16:1}$ ω7*c* and/or $C_{16:1}$ ω6*c* and/or iso-$C_{15:0}$ 2-OH; summed feature 8 comprises $C_{18:1}$ ω7*c* and/or $C_{18:1}$ ω6*c*; summed feature 7 comprises unknown 18.846 and/or $C_{19:1}$ ω6*c*.

3.4.6 Description of a novel species

The novel species of the genus *Tepidimonas* represented by strains SPSP-6^T has very few phenotypic and chemotaxonomic characteristics that distinguish this strain from the type strains of the other species. The single carbon source assimilations, with the exception of *T. taiwanensis*, sulfur oxidation and the fatty acid composition are similar in all type strains of species of the genus *Tepidimonas* (Table 3.2, 3.8 and 3.9). However, some phenotypic and chemotaxonomic characteristics indicate that the organism represents a novel species; the fatty acid composition indicated that SPSP-6^T and SPSPC-18 can be distinguished from other strains of members of the genus *Tepidimonas* by combining the relative proportions of $C_{17:0}$ cyclo and $C_{17:0}$. Except for strains SPSP-6^T and *T. alkaliphilus* YIM 72238^T, which possess only homologues for *nirB*, all other strains possess genes coding for proteins involved in the reduction of nitrate to nitrite. Considering the phenotypic, genomic and the phylogenetic analysis based on 16S rRNA gene sequence and on 400 conserved genes sequences clearly confirms that strain SPSP-6^T represents a species level taxon, leading us to propose the name *Tepidimonas charontis* sp. nov.

3.4.6.1 Description of *Tepidimonas charontis* sp. nov.

Tepidimonas charontis (Albuquerque *et al*., 2020a^{VP}; Albuquerque and Egas, 2021b).

cha.ron'tis. L. gen. n. *charontis* of Charon, the boatman who required payment to ferry the ancient dead Greeks across the Rivers Styx and Acheron to Hades.

Forms short rod-shaped cells 0.5–0.8 µm in width and 1.0–2.0 µm in length. Endospores are not formed. The cells stain Gram-negative and are motile by one polar flagellum. Colonies on Degryse medium 162 are not pigmented and are 1 to 2 mm in diameter after 48 h of growth. The optimum growth temperature is about 50ºC; growth occurs in the range of 25–60ºC. The optimum pH is between 7.5 and 9.0; the pH range for growth is pH 6.5–9.5. Mixotrophic. Aerobic. Nitrate is not reduced to nitrite. Cytochrome *c* oxidase and catalase positive. The major fatty acids are $C_{16:0}$ and $C_{16:1}$ ω6*c* and/or $C_{16:1}$ ω7*c*. Yeast extract or growth factors are required for growth. Thiosulfate is oxidized to sulfate. Several organic acids and amino acids are utilized for growth, namely succinate, lactate, pyruvate, acetate, glutamate, aspartate, L-alanine, L-asparagine, L-lysine, L-glutamine, L-isoleucine and L-ornithine, but the strains do not utilize carbohydrates or polyols. The type strain SPSP-6^T (=CECT 9683^T =LMG 30884^T) was isolated from a hot spring at São Pedro do Sul in central Portugal. The genomic DNA G+C content is 66.63% (determined by genome sequencing). Strain SPSPC-18 (=CECT 9684 =LMG 30885) is an additional strain of this species. GenBank/EMBL/DDBJ accession numbers for the 16S rRNA gene sequence of strains SPSP-6^T and SPSPC-18 are MH590702 and MH590703, respectively. The draft genome sequence of SPSP-6^T (VJON00000000) has been deposited in GenBank/EMBL/DDBJ.

Acknowledgments

We thank Aharon Oren (The Hebrew University of Jerusalem, Israel) for the etymology of the name of the novel organism. We also thank Ramon Rosselló-Móra (IMEDEA, Illes Baleares, Spain).

Chapter 4

Reclassification of Four Yellow-Pigmented Species of the Genus *Meiothermus* to the Novel Genus *Calidithermus* and Emended Description of the Genus *Meiothermus*

Results published in:

[*]RAPOSO, P., [*]VIVER, T., [*]ALBUQUERQUE, L., FROUFE, H., BARROSO, C., EGAS, C., ROSSELLÓ-MÓRA, R. and DA COSTA, M.S. (2019). Transfer of *Meiothermus chliarophilus* (Tenreiro *et al*. 1995) Nobre *et al*. 1996, *Meiothermus roseus* Ming *et al*. 2016, *Meiothermus terrae* Yu *et al*. 2014 and *Meiothermus timidus* Pires *et al*. 2005, to *Calidithermus* gen. nov., as *Calidithermus chliarophilus* comb. nov., *Calidithermus roseus* comb. nov., *Calidithermus terrae* comb. nov. and *Calidithermus timidus* comb. nov., respectively, and emended description of the genus *Meiothermus*. *Int J Syst Evol Microbiol* 69: 1060–1069. doi.org/10.1099/ijsem.0.003270

[*]These authors contributed equally to this work.

4.1 Abstract

Chemotaxonomic parameters, phylogenetic analysis of the 16S rRNA gene, phylogenetic analysis of 90 housekeeping genes and 855 core-genes, average amino acid identity, average nucleotide identity and genomic characteristics were used to examine the thirteen species of the genus *Meiothermus* with validly published names to reclassify this genus. The results indicate that the species of the genus *Meiothermus* can be divided into three lineages on the basis of the results of the phylogenetic analysis, average amino acid identity, the G+C ratio, the ability to synthesize the red-pigmented carotenoid canthaxanthin and the colony colour, as well as other genomic characteristics. The results presented in this study circumscribe the genus *Meiothermus* to the species *Meiothermus ruber*, *Meiothermus cateniformans*, *Meiothermus taiwanensis*, *Meiothermus cerbereus*, *Meiothermus hypogeus*, *Meiothermus luteus*, *Meiothermus rufus* and *Meiothermus granaticius*, for which it is necessary to emend the genus *Meiothermus*. The species *Meiothermus silvanus*, which clearly represents a separate genus level lineage was not reclassified in this study for lack of any distinctive phenotypic or genotypic characteristics. The results of this study led us to reclassify the species *M. chliarophilus*, *M. timidus*, *M. roseus* and *M. terrae* as species of a novel genus for which we propose the epithet *Calidithermus* gen. nov.

4.2 Introduction

The genus *Meiothermus* was proposed by (Nobre *et al*., 1996) to reclassify three species included in the genus *Thermus* but grew at lower temperatures, formed red- or yellow-pigmented colonies and possessed two variants of glycolipid 1 (GL-1). The species of the genus *Meiothermus* comprise thirteen species with validly published names, namely *M. ruber* (Loginova *et al*., 1984), *M. chliarophilus* (Nobre *et al*., 1996; Tenreiro *et al*., 1995), *M. silvanus* (Nobre *et al*., 1996; Tenreiro *et al*., 1995), *M. cerbereus* (Chung *et al*., 1997), *M. taiwanensis* (Chen *et al*., 2002), *M. timidus* (Pires *et al*., 2005a), *M. rufus* (Albuquerque *et al*., 2009), *M. cateniformans* (Zhang *et al*., 2010), *M. granaticius* (Albuquerque *et al*., 2010b), *M. hypogaeus* (Mori *et al*., 2012), *M. terrae* (Yu *et al*., 2014), *M. roseus* (Ming *et al*., 2015) and *M. luteus* (Habib *et al*., 2017). Most of the type strains of the genus *Meiothermus* are red-pigmented but the type strains of *M. chliarophilus*, *M. timidus*, *M. terrae* and *M. roseus* form yellow-pigmented colonies. The type strains of *M. timidus* and *M. chliarophilus* were isolated from the hot spring of S. Pedro do Sul in central Portugal and from the hot spring of Alcafache in the same region of Portugal, respectively (Pires *et al*., 2005a; Tenreiro *et al*., 1995). The name *M. roseus* refers to the colour of a diffusible pink pigment on solid R2A medium, although the colonies are yellow-pigmented.

The result of phylogenetic analyses based on 16S rRNA gene sequence indicate that the species of the genus *Meiothermus* form a separate line of descent from the species of the genus *Thermus* with which they share about 83.0–87.0% 16S rRNA gene sequence similarity. Moreover, the species of the genus *Meiothermus* form, at least, three well defined lineages as determined by 16S

rRNA gene sequence analysis (Albuquerque *et al*., 2018e). These lineages included the clusters comprising *M. ruber, M. cateniformans, M. taiwanensis, M. cerbereus, M. hypogaeus* and *M. rufus*. Another lineage includes the four yellow-pigmented species *M. chliarophilus, M. timidus, M. roseus* and *M. terrae*. A third lineage comprises *M. silvanus*. The species *M. granaticius* could also be assumed to represent a separate lineage. At the time *M. luteus* had not been described and was not included in the 16S rRNA analysis (Albuquerque *et al*., 2018e).

Despite the low 16S rRNA gene sequence similarities of some lineages, there are few phenotypic characteristics that distinguish each of the deep-rooted lineages classified as members of the genus *Meiothermus*. The fatty acids, for example, of all strains are predominantly composed of iso- and anteiso-branched C_{15} and C_{17} fatty acids that display only minor differences among type strains of the genus *Meiothermus* (Albuquerque *et al*., 2018e). Iso- and anteiso-branched fatty acids are also the predominant acyl chains of the strains of the related genera *Thermus*, *Oceanithermus*, *Vulcanithermus*, *Rhabdothermus* and *Marinithermus* (Albuquerque and da Costa 2014; Miroshnichenko *et al*., 2003a, 2003b; Mori *et al*., 2004; Sako *et al*., 2003; Steinsbu *et al*., 2011).

With only a paucity of phenotypic characteristics to reclassify the genus *Meiothermus* we resorted to an extensive characterization of genomic data allied with a few phenotypic characteristics. Our analysis led us to define one new genus to comprise the four yellow-pigmented species *M. chliarophilus, M. timidus, M. terrae* and *M. roseus* for which we propose the name *Calidithermus* gen. nov.

4.3 Material and Methods

4.3.1 Chemotaxonomic characterization

Cultures for polar lipids and fatty acid analysis were grown in *Thermus* liquid medium at 50ºC until the late-exponential growth phase (for details 2.3.1 and 2.3.4).

4.3.2 Extraction of DNA, genome sequencing, assembly and annotation

Total genomic DNA was extracted following the method of Nielsen *et al*. (1995) (for details 2.3.5). The purity and quantity of the DNA were verified as described in 2.3.5. The genomic DNA was prepared with the Nextera XT DNA Library Preparation Kit and sequenced using PE 2x300 bp on the MiSeq (Illumina, San Diego, USA). Sequenced reads were quality filtered with Trimmomatic (Bolger *et al*., 2014) and assembled with SPAdes version 3.9.1 (Bankevich *et al*., 2012). Resulting contigs were annotated with PGP2. PGP2 used Prodigal version 2.6 (Hyatt *et al*., 2010) for gene prediction, Barrnap version 0.8 (https://github.com/tseemann/barrnap) for rRNA and tRNA genes detection, and Prokka version 1.12 (Seemann, 2014) for the annotation of protein-coding genes. Gene annotation with Prokka used the SwissProt (Apweiler *et al*., 2004), HAMAP (Pedruzzi *et al*.,

2015), TIGRFAMs (Haft *et al*., 2003) and Pfam (Finn *et al*., 2016) repositories. Genome estimated completeness and contamination were verified with CheckM version 1.0.7 (Parks *et al*., 2015).

High-quality draft genome sequences were performed with the type strains of *M. cateniformans* JCM 15151^T (QWKX00000000), *M. granaticius* AF-68^T (=DSM 23260^T) (QWLB00000000), *M. hypogaeus* DSM 23238^T (QWKY00000000), *M. luteus* KCTC 52599^T (QWKZ00000000), *M. roseus* NBRC 110900^T (QWLA00000000) and *M. terrae* DSM 26712^T (QXDL00000000). The genome sequences of *M. ruber* DSM 1279^T (CP001743.1), *M. taiwanensis* DSM 14542^T (AXWR00000000.1), *M. silvanus* DSM 9946^T (CP002042.1, CP002043.1 and CP002044.1), *M. cerbereus* DSM 11376^T (JHVI00000000.1), *M. rufus* DSM 22234^T (AUHY00000000.1), *M. chliarophilus* DSM 9957^T (AUQW00000000.1), *M. timidus* DSM 17022^T (ARDL00000000.1), *Oceanithermus profundus* DSM 14977^T (CP002561.1) and *Thermus aquaticus* Y51MC23^T (CP010822.1) were obtained from the databases.

4.3.3 Tree reconstructions based on 16S rRNA genes

The complete 16S rRNA genes sequences recovered from genomes were extracted using the RNAmmer 1.2 Server (Lagesen *et al*., 2007). The genes were aligned using the SINA v1.2.12 tool (SILVA Incremental Aligner [Pruesse *et al*., 2007]) implemented within the ARB software package version 5.5 (Ludwig *et al*., 2004) and added by parsimony to the LTPs128_SSU database (Yarza *et al*., 2014). Final alignments were manually improved following the reference alignment in ARB-editor. The aligned sequences were used to reconstruct *de novo* trees using the NJ (Saitou and Nei, 1987) algorithm with the Jukes-Cantor correction and randomized axelerated maximum likelihood (RaxML) v8.2.0 (Stamatakis, 2006) algorithm with the GTRGAMMA correction.

4.3.4 Core-pan-genome analysis, phylogenetic reconstruction and ANI/AAI calculation

Genomic analyses were performed as detailed by Viver *et al*. (2018). CDS from assembled genomes were conducted by using GeneMark.hmm with default parameters (Besemer *et al*., 2001). The CDS were compared using an all-versus-all BLAST v2.2.28 (Altschul *et al*., 1990) with available reference sequences in order to identify the reciprocal best matches (RBM) in all pairwise genome comparisons using a 50% sequence similarity cutoff and over 50% of the query sequence length. The orthologous groups (OGs) in the RBMs analysis were identified using the Markov cluster algorithm implemented in *ogs.mcl.rb* script from Enveomics collection (Rodriguez-R and Konstantinidis, 2016). Proteins shared between all genomes (core-genome) were aligned individually using MUSCLE v3.8.31 (Edgar, 2004). The concatenated and aligned OGs were used to reconstruct NJ phylogenetic trees using the NJ (Saitou and Nei, 1987) algorithm implemented in ARB software (Ludwig *et al*., 2004). The CDS present in two or more genomes were defined as the variable genes (pan-genome). The presence or absence of variable genes was used to cluster the genomes with the Euclidian distance using the ggplot2 package from Wickham (Wickham, 2016).

The estimates of the core-pan-genome sizes were predicted using the script *ogs.core-pan.rb* implemented in Enveomics collection (Rodriguez-R and Konstantinidis, 2016).

The housekeeping CDS from all genomes were extracted using the script *HMM.essential.rb* from Enveomics collection (Rodriguez-R and Konstantinidis, 2016). The genes were concatenated, aligned and phylogenetic tree was constructed as detailed for the core-genome phylogeny.

The ANIb and the AAI between all genomes were determined according to Konstantinidis and Tiedje (2005a, 2005b) using the webserver available at http://enve-omics.gatech.edu/ (Rodriguez-R and Konstantinidis, 2016).

4.3.5 Draft genome accession numbers

Draft genomes accession numbers of strains *M. cateniformans* JCM 15151^T (QWKX00000000), *M. granaticius* AF-68^T (=DSM 23260^T) (QWLB00000000) *M. hypogaeus* DSM 23238^T (QWKY00000000), *M. luteus* KCTC 52599^T (QWKZ00000000), *M. roseus* NBRC 110900^T (QWLA00000000) and *M. terrae* DSM 26712^T (QXDL00000000) were deposited in GenBank/EMBL/DDBJ.

4.4 Results and Discussion

4.4.1 Phylogenomic and comparative genomic analyses

The results of phylogenetic analysis based on 16S rRNA gene sequence of all type strains indicated that the species of the genus *Meiothermus* form a separate monophyletic line of descent from the species of *Thermus* with which they share about 83.0–87.0% 16S rRNA gene sequence similarity (Albuquerque *et al*., 2018e). The tree topology (Figure 4.1, 4.2 and 4.3) indicates that there are three major lineages forming the *Meiothermus* branch. One comprises the type strain of the genus *M. ruber*, monophyletic with *M. taiwanensis*, *M. cateniformis* and *M. cerbereus* sharing 16S rRNA sequence similarities ranging from 98.5 to 98.7%, and more loosely affiliated with *M. granaticius*, *M. rufus*, *M. luteus* and *M. hypogaeus* sharing 93.2 to 95.2% with *M. ruber*. On the other hand, a second lineage, loosely affiliated with the former, comprised the type strains of *M. chliarophilus*, *M. roseus, M. timidus* and *M. terrae* with sequence similarities of less than 91% with *M. ruber*, but high 16S rRNA with each other of 93.4 to 98.7%. Finally, a third lineage was formed by the single sequence of the species *M. silvanus*, which was very loosely related to any the other members of the genus *Meiothermus*, with less than 88% 16S rRNA sequence identity. All such results indicated that at least the three lineages, showing less than 94% identity among themselves may be considered to represent distinct genera according to previously published thresholds (Yarza *et al*., 2014). The results of the genome comparisons were consistent with our 16S rRNA observations regarding the distinct putative genus nature of the three lineages. The whole-genome analyses further reinforced our observations based on the ribosomal small subunit gene sequences. The phylogenetic

reconstructions based on the concatenates of the complete core-genome of 855 shared genes among all genomes (Figure 4.4), and also that of a subset of the 90 housekeeping genes present in all genomes (Figure 4.5), both were consistent with the 16S rRNA gene lineage distinctness. In both cases, the three major lineages, despite some small branching order differences, remained stable and corroborated the three-lineage nature of the genus *Meiothermus*. The heterogeneous nature of the genus was finally supported by the OGRI parameters useful for taxonomic purposes (Chun and Rainey, 2014). The ANI values are valuable for delineating species at the threshold levels ranging from 95 to 96% identity (Richter and Rosselló-Móra, 2009). In this respect, the two branches containing more than one species showed interspecific ANIb values greater than 77% (data not shown) within each branch also confirming that these single lineages formed coherent genera (Kim *et al*., 2014). However, as the 16S rRNA gene sequences diverged, the ANI signal decreased and the AAI became the parameter to consider for genome comparisons; this is proposed to discern genera at thresholds around 70% (Konstantinidis and Tiedje, 2005b; Luo *et al*., 2014). In this regard, we encountered that the three lineages could be considered separate genera within these bacteria. The true *Meiothermus* genus, formed by the lineage comprising the type species of the genus *M. ruber*, and the additional type strains of *M. cateniformans*, *M. taiwanensis*, *M. cerbereus*, *M. luteus*, *M. rufus*, *M. hypogaeus* and *M. granaticius*, with AAI values in the neighbourhood of 70% or higher. The second lineage comprising *M. chliarophilus*, *M. roseus*, *M. terrae* and *M. timidus* showed internal AAI values between the species of 85%, whereas with the neighbour, the true *Meiothermus* lineage, AAI values were always less than 70%. Finally, the most distant lineage formed by the single species *M. silvanus* showed values always less than 70.7% with any of the currently classified species of the genus (Table 4.1). Altogether, the genome to genome comparison parameters reinforce the idea of *Meiothermus* comprising at least three different genera, each of them formed by one of the distinct lineages.

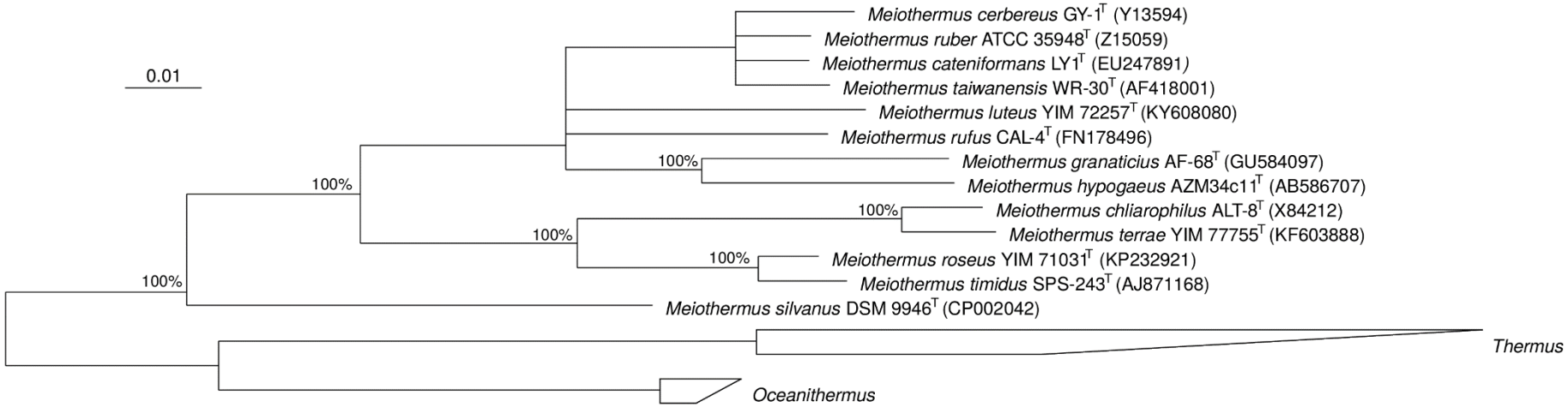

Figure 4.1 Phylogenetic reconstruction based on 16S rRNA genes of type strains of members of the genus *Meiothermus*. This is a consensus tree between the NJ and RaxML reconstructed trees using the 30% conservative filter. Multifurcations show branching order that could not be resolved. Bootstrap values were obtained using the RAxML algorithm, and are only shown for branches having 50% or more stability. Bar, 0.01 substitutions per nucleotide position.

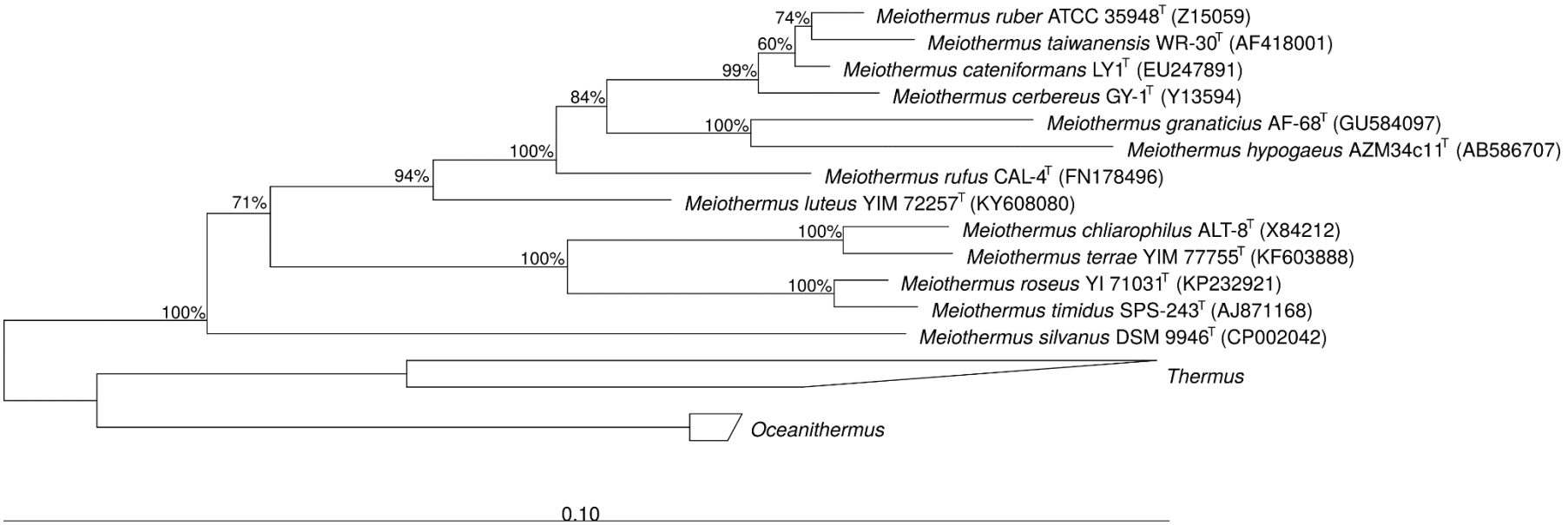

Figure 4.2 Phylogenetic reconstruction based on 16S rRNA genes of type strains of members of the genus *Meiothermus* using the NJ algorithm. Bar, 0.1 substitutions per nucleotide position.

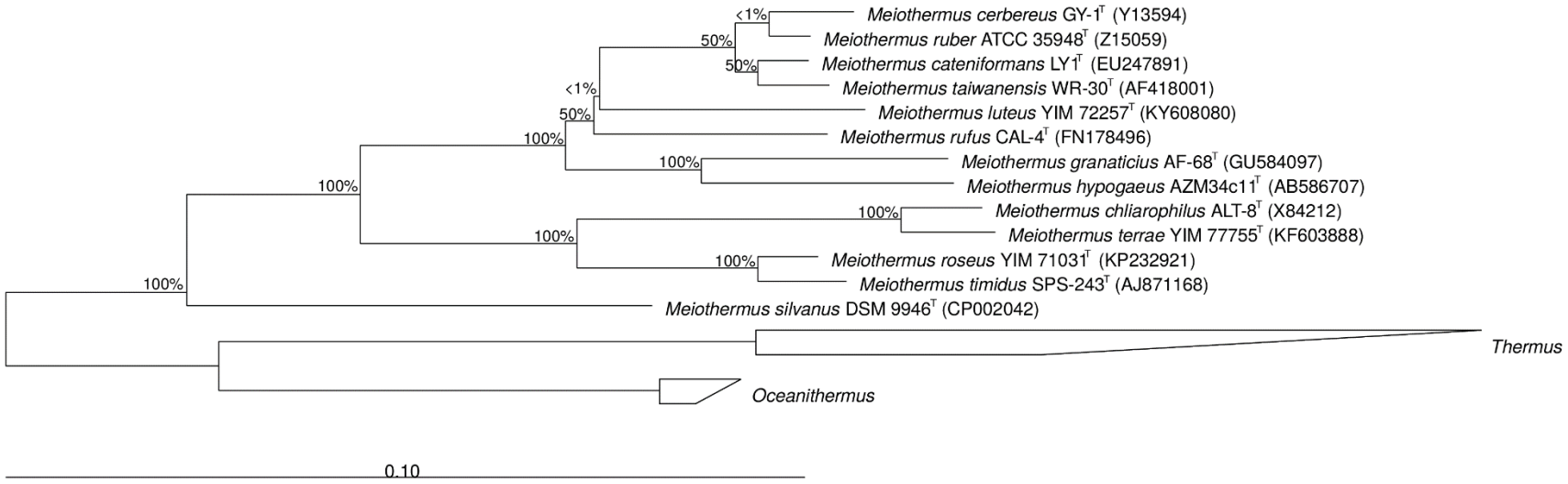

Figure 4.3 Phylogenetic reconstruction based on 16S rRNA genes of type strains of members of the genus *Meiothermus* using RAxML method. Bar, 0.1 substitutions per nucleotide position.

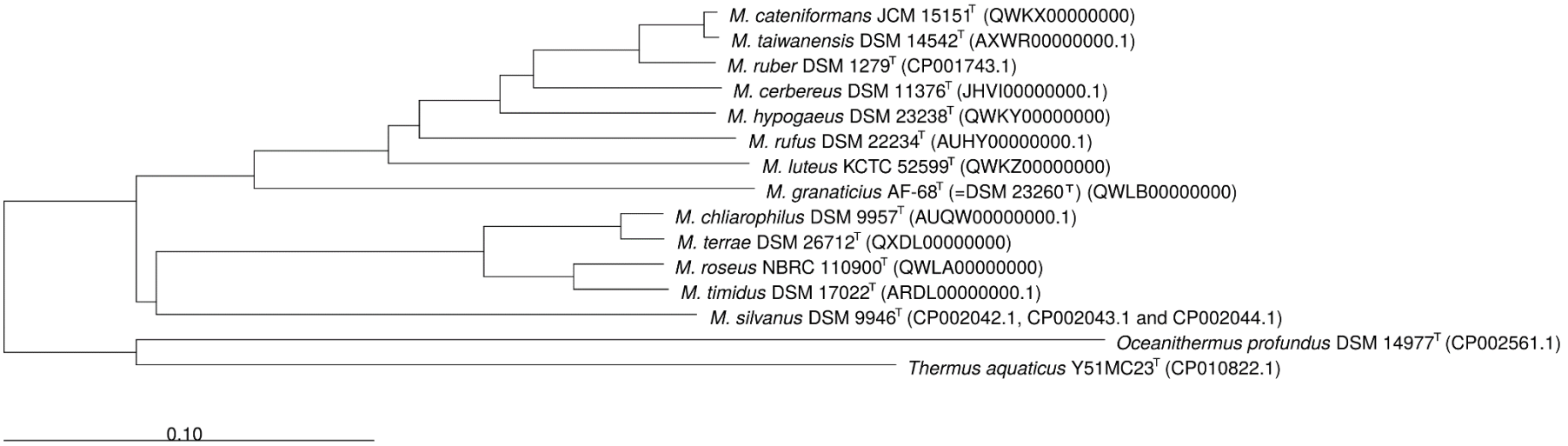

Figure 4.4 Phylogenetic tree reconstruction based on an NJ calculation for the 855 core-genes of genomes of type strains of members of the genus *Meiothermus*, *Oceanithermus profundus* DSM 14977^T and *Thermus aquaticus* Y51MC23 (genome accession numbers in parentheses). Bar, 0.1 substitutions per amino acid position.

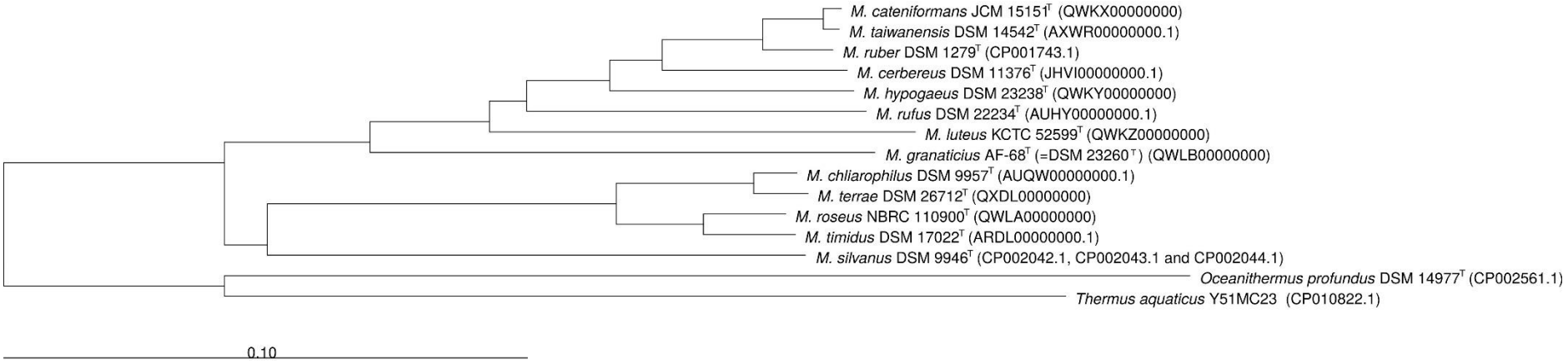

Figure 4.5 Phylogenetic tree reconstruction based on 90 housekeeping genes (essential genes present in almost all Bacterial genomes) of type strains of members of the genus *Meiothermus*, *Oceanithermus profundus* DSM 14977^T and *Thermus aquaticus* Y51MC23. Bar, 0.1 substitutions per amino acid position.

Table 4.1 AAI values between genomes of members of the genus *Meiothermus*. AAI values between all genes encoded in the genomes are given in the upper triangle and the numbers of proteins shared are given in brackets. The AAI values between the 873 core-genes are given in the lower triangle. *M. cateniformans* JCM 15151^T (QWKX00000000), *M. taiwanensis* DSM 14542^T (AXWR00000000.1), *M. ruber* DSM 1279^T (CP001743.1), *M. cerbereus* DSM 11376^T (JHVI00000000.1), *M. hypogaeus* DSM 23238^T (QWKY00000000), *M. rufus* DSM 22234^T (AUHY00000000.1), *M. granaticius* AF-68^T (=DSM 23260^T) (QWLB00000000), *M. luteus* KCTC 52599^T (QWKZ00000000), *M. chliarophilus* DSM 9957^T (AUQW00000000.1), *M. terrae* DSM 26712^T (QXDL00000000), *M. roseus* NBRC 110900^T (QWLA00000000), *M. timidus* DSM 17022^T (ARDL00000000.1), *M. silvanus* DSM 9946^T (CP002042.1, CP002043.1 and CP002044.1). *Oceanithermus profundus* DSM 14977^T (CP002561.1) and *Thermus aquaticus* Y51MC23 (CP010822.1) were also included in this table.

	M. cateniformans	***M. taiwanensis***	***M. ruber***	***M. cerbereus***	***M. hypogaeus***	***M. rufus***	***M. granaticius***	***M. luteus***	***M. chliarophilus***	***M. terrae***	***M. roseus***	***M. timidus***	***M. silvanus***	***Oceanithermus profundus***	***Thermus aquaticus***
M. cateniformans	–	98.84% [2600]	93.38% [2488]	85.78% [2249]	82.81% [2402]	79.95% [2178]	70.27% [2115]	78.14% [2245]	68.06% [2153]	67.47% [2116]	68.59% [2178]	68.90% [2136]	68.00% [2112]	55.89% [1500]	59.78% [1651]
M. taiwanensis	99.12%	–	92.98% [2555]	85.14% [2308]	83.20% [2426]	80.19% [2217]	70.21% [2110]	78.07% [2229]	68.21% [2177]	67.44% [2102]	68.51% [2180]	68.86% [2173]	67.84% [2160]	55.76% [1529]	59.58% [1687]
M. ruber	95.70%	95.62%	–	85.60% [2403]	82.96% [2479]	79.85% [2247]	70.14% [2156]	78.03% [2342]	67.32% [2242]	66.73% [2171]	68.37% [2255]	69.42% [2272]	68.04% [2270]	55.62% [1551]	59.36% [1708]
M. cerbereus	89.92%	89.85%	90.33%	–	84.80% [2489]	79.84% [2211]	70.10% [2081]	77.84% [2141]	66.84% [2168]	66.60% [2109]	68.24% [2122]	68.51% [2119]	67.06% [2188]	55.48% [1547]	59.53% [1695]
M. hypogaeus	88.11%	88.05%	88.35%	88.94%	–	80.24% [2274]	69.97% [2298]	77.7% [2218]	67.95% [2378]	66.57% [2283]	67.60% [2278]	68.15% [2223]	67.21% [2266]	55.12% [1597]	58.85% [1691]
M. rufus	84.54%	84.56%	84.56%	84.29%	84.62%	–	69.86% [2010]	77.51% [2132]	67.68% [2051]	67.34% [1983]	68.43% [2025]	68.50% [2056]	68.25% [2034]	56.38% [1489]	60.36% [1639]

Table 4.1 (continued)

M. granaticius	76.93%	76.92%	76.88%	76.63%	77.17%	76.36%	–	68.52% [2070]	65.37% [2276]	65.10% [2242]	67.40% [2257]	67.70% [2170]	67.32% [2238]	55.27% [1547]	58.61% [1616]
M. luteus	83.88%	83.84%	83.79%	83.38%	92.49%	83.14%	76.84%	–	66.80% [2097]	66.40% [2076]	67.90% [2137]	69.98% [2166]	67.38% [2111]	56.33% [1492]	61.09% [1672]
M. chliarophilus	74.12%	74.09%	74.13%	73.90%	74.07%	73.62%	73.43%	74.58%	–	94.93% [3384]	85.99% [2792]	85.61% [2475]	69.19% [2494]	55.57% [1678]	59.12% [1749]
M. terrae	74.09%	74.09%	74.13%	73.87%	74.06%	73.60%	73.40%	74.64%	97.64%	–	85.54% [2858]	84.57% [2405]	68.51% [2390]	55.47% [1621]	58.96% [1692]
M. roseus	74.14%	74.08%	74.06%	73.79%	73.98%	73.76%	73.38%	74.75%	90.69%	90.35%	–	90.16% [2574]	69.77% [2409]	56.08% [1620]	59.99% [1710]
M. timidus	74.06%	74.07%	74.07%	73.73%	73.86%	73.76%	73.30%	74.80%	90.43%	90.01%	94.95%	–	70.68% [2323]	56.96% [1570]	60.61% [1687]
M. silvanus	73.12%	73.08%	73.10%	72.99%	73.14%	72.70%	72.82%	73.92%	74.94%	74.89%	74.86%	74.71%	–	56.47% [1621]	60.20% [1756]
O. profundus	62.25%	62.29%	62.19%	62.06%	61.88%	62.53%	61.81%	64.16%	63.03%	63.08%	62.95%	62.95%	63.21%	–	58.54% [1510]
T. aquaticus	65.34%	65.36%	65.27%	65.00%	64.75%	65.87%	64.67%	67.81%	65.73%	65.77%	65.76%	65.80%	66.36%	63.85%	–

4.4.2 Phenotypic and chemotaxonomic characteristics

In contrast to the genome data, the phenotypic and chemotaxonomic characteristics of the members of the genus *Meiothermus* cannot lead to conclusions on the classification of the species as separate genera. The single carbon source assimilations, many of which were performed under different growth conditions and by different methods, are similar. The studies show a fairly homogenous set of results of growth with single carbon sources that could even be the result of interspecies diversity. There are no noticeable differences in growth temperature or pH range. Fatty acid composition of all the type strains examined under identical growth conditions shows minor variations in fatty acid levels and polar lipid patterns cannot be correlated with the phylogenetic results that would divide the genus *Meiothermus* into three putative genera (Table 4.2). Most of the strains of the species with validly published names of the genus *Meiothermus* produce 3- and 2-OH fatty acids (Ferreira *et al*., 1999); 3-OH fatty acids are very rare in species of the genus *Thermus* while 2-OH fatty acids have never been identified in that genus. The major fatty acids in both genera are always iso-$C_{15:0}$, anteiso-$C_{15:0}$, iso-$C_{17:0}$ and anteiso-$C_{17:0}$. The strains of species of the genera *Thermus* and *Meiothermus* have one identical major phospholipid (PL) on TLC, designated PL-2 and one major glycolipid, designated GL-1, whose structures are known (Figure 4.6) (Carreto *et al*., 1996; Ferreira *et al*., 1999; Lagutin *et al*., 2014). The species of the genus *Meiothermus* generally have two glycolipid variants of GL-1 as determined by TLC, instead of only one as in members of the genus *Thermus*. The faster-migrating glycolipid on TLC is designated GL-1b, while the slower-migrating component is designated GL-1a (Ferreira *et al*., 1999; Yang *et al*., 2006). The slower-migrating GL-1a variant has the hexosamine of the polar head group exclusively N-acylated with 2-OH iso fatty acids while GL-lb may comprise glycolipids that are N-acylated with 3-OH iso fatty acids, non-hydroxylated iso fatty acids or a mixture of both. The glycolipid variant GL-1a is not detected by TLC in *M. rufus* and *M. granaticius* which, unlike the other type strains, do not possess 2-OH fatty acids. It is noteworthy that the growth temperature affects the levels of GL-1a (Albuquerque *et al*., 2009), thus affecting our insight into the value of the glycolipids in the classification of the genus. Two minor glycolipids, designated GL-2a and GL-2b are always detected in the two closely related species *M. chliarophilus* and *M. terrae* (Albuquerque *et al*., 2009; Yu *et al*., 2014). These glycolipids are intermediates in the synthesis of GL-1a and GL-1b and are also sometimes visible in other species of this genus (Figure 4.7) (Wait *et al*., 1997).

Conspicuously, the colony colour seems to be the only phenotypic characteristic that could be used to distinguish the species of the genus *Meiothermus*.

Table 4.2 Fatty acid composition of the species of the genus *Meiothermus* grown in *Thermus* liquid medium at 50ºC until the late-exponential phase of growth. Strains: 1, *M. ruber* ATCC25948^T; 2, *M. cateniformans* JCM 15151^T; 3, *M. taiwanensis* DSM 14542^T; 4, *M. cerbereus* DSM 11376^T; 5, *M. luteus* KCTC 52599^T; 6, *M. rufus* CAL-4^T (=DSM 22234^T); 7, *M. hypogaeus* DSM 23238^T; 8, *M. granaticius* AF-68^T (=DSM 23260^T); 9, *M. roseus* NBRC 110900^T; 10, *M. timidus* SPS-243^T (=DSM 17022^T); 11, *M. terrae* DSM 26712^T; 12, *M. chliarophilus* ATCC 700543^T; 13, *M. silvanus* ATCC 700542^T.

Fatty acids	ECL	1	2	3	4	5	6	7	8	9	10	11	12	13
Straight-chain														
$C_{14:0}$	14.000	tr	tr	–	–	tr	tr	1.4 ± 0.1	0.8 ± 0.2	tr	tr	tr	tr	–
$C_{15:0}$	15.000	1.3 ± 0.2	1.4 ± 0.2	1.1 ± 0.2	0.9 ± 0.2	2.1 ± 0.3	2.0 ± 0.1	2.5 ± 0.2	0.5 ± 0.1	0.9 ± 0.1	0.9 ± 0.1	2.6 ± 0.1	1.6 ± 0.3	tr
$C_{16:0}$	16.000	0.9 ± 0.2	2.3 ± 0.3	1.1 ± 0.3	1.0 ± 0.2	3.2 ± 0.2	2.3 ± 0.2	5.5 ± 0.3	4.8 ± 0.4	2.4 ± 0.2	1.4 ± 0.2	4.5 ± 0.4	3.9 ± 0.5	tr
$C_{17:0}$	17.000	0.7 ± 0.2	1.5 ± 0.3	1.9 ± 0.3	–	1.0 ± 0.2	1.1 ± 0.1	0.6 ± 0.1	–	1.1 ± 0.1	1.1 ± 0.1	3.6 ± 0.3	1.3 ± 0.1	tr
$C_{18:0}$	18.000	–	–	tr	–	–	tr	–	–	tr	tr	0.6 ± 0.1	tr	–
$C_{15:1}$ ω8*c*	14.793	tr	tr	tr	1.1 ± 0.3	–	tr	–	–	–	–	–	–	–
$C_{17:1}$ ω6*c*	16.860	0.8 ± 0.2	tr	0.6 ± 0.1	1.6 ± 0.3	tr	–	–	–	–	–	0.6 ± 0.1	–	tr
$C_{16:1}$ ω7*c* alcohol	15.387	0.7 ± 0.2	0.5 ± 0.1	tr	3.4 ± 0.2	–	–	–	–	–	–	–	–	–
Branched-chain														
iso-$C_{11:0}$	10.606	tr	tr	0.7 ± 0.1	1.3 ± 0.3	tr	tr	tr	0.6 ± 0.3	–	tr	tr	tr	–
iso-$C_{13:0}$	12.614	0.5 ± 0.1	tr	0.8 ± 0.1	2.8 ± 0.4	0.9 ± 0.1	0.5 ± 0.1	0.9 ± 0.1	tr	1.0 ± 0.2	1.2 ± 0.1	1.2 ± 0.1	1.8 ± 0.2	tr
iso-$C_{14:0}$	13.619	0.8 ± 0.1	1.2 ± 0.2	1.2 ± 0.3	3.4 ± 0.3	1.1 ± 0.2	0.7 ± 0.1	2.5 ± 0.2	0.8 ± 0.2	tr	0.5 ± 0.1	2.9 ± 0.2	1.9 ± 0.3	0.7 ± 0.1
iso-$C_{15:0}$	14.623	35.2 ± 1.5	28.7 ± 0.6	35.3 ± 1.0	18.0 ± 1.1	46.9 ± 1.5	38.0 ± 1.3	32.8 ± 1.2	11.2 ± 1.3	46.1 ± 1.3	41.6 ± 1.4	35.3 ± 1.5	41.7 ± 1.5	22.5 ± 0.9
anteiso-$C_{15:0}$	14.713	6.4 ± 0.4	9.5 ± 0.4	3.4 ± 0.2	4.7 ± 0.2	3.5 ± 0.4	11.5 ± 0.7	23.6 ± 1.0	55.0 ± 1.5	4.5 ± 0.3	6.3 ± 0.5	7.6 ± 0.5	12.4 ± 0.8	29.9 ± 1.2
iso-$C_{16:0}$	15.627	3.5 ± 0.3	5.8 ± 0.3	6.0 ± 0.3	2.5 ± 0.6	3.8 ± 0.4	3.4 ± 0.2	5.3 ± 0.3	0.6 ± 0.1	1.1 ± 0.1	1.1 ± 0.1	4.0 ± 0.3	2.7 ± 0.2	2.3 ± 0.2
iso-$C_{17:0}$	16.630	20.3 ± 0.9	21.5 ± 0.8	26.4 ± 1.2	3.7 ± 0.5	17.9 ± 2.0	22.2 ± 0.8	9.2 ± 0.3	8.6 ± 0.9	32.5 ± 1.4	33.6 ± 1.2	22.5 ± 0.8	22.1 ± 1.0	12.0 ± 0.5
anteiso-$C_{17:0}$	16.723	5.6 ± 0.4	12.0 ± 0.5	4.4 ± 0.3	1.9 ± 0.4	2.2 ± 0.3	10.2 ± 0.6	9.2 ± 0.3	8.4 ± 0.8	3.9 ± 0.3	5.5 ± 0.3	3.4 ± 0.2	4.6 ± 0.2	10.7 ± 0.4
iso-$C_{18:0}$	17.632	0.5 ± 0.1	0.5 ± 0.1	0.8 ± 0.1	–	tr	tr	–	–	tr	tr	1.2 ± 0.1	tr	0.5 ± 0.1
iso-$C_{19:0}$	18.634	–	tr	0.7 ± 0.1	–	–	tr	–	–	1.2 ± 0.1	0.8 ± 0.1	0.8 ± 0.1	0.7 ± 0.2	2.6 ± 0.2
anteiso-$C_{19:0}$	18.731	–	tr	–	–	–	tr	–	–	tr	tr	0.6 ± 0.1	tr	2.6 ± 0.2

Table 4.2 (continued)

iso-$C_{15:1}$ F*	14.415	2.0 ± 0.2	0.6 ± 0.1	0.5 ± 0.1	8.9 ± 1.0	–	1.0 ± 0.1	tr	–	–	–	–	–	–
iso-$C_{16:1}$ H	15.461	tr	tr	tr	1.5 ± 0.3	–	tr	–	–	–	–	–	–	–
anteiso-$C_{17:1}$ ω9*c*	16.524	0.8 ± 0.2	1.0 ± 0.2	tr	1.6 ± 0.2	–	0.5 ± 0.1	–	–	–	–	–	–	–
2-Hydroxy														
$C_{15:0}$ 2-OH	16.219	tr	–	tr	0.8 ± 0.1	0.5 ± 0.1	–	–	–	–	–	tr	–	–
$C_{16:0}$ 2-OH	17.233	tr	tr	tr	–	tr	–	0.6 ± 0.1	–	–	–	–	–	–
$C_{17:0}$ 2-OH	18.254	tr	tr	tr	1.9 ± 0.3	tr	–	–	–	–	0.6 ± 0.1	tr	–	–
iso-$C_{17:0}$ 2-OH	17.880	8.4 ± 0.4	4.0 ± 0.3	8.6 ± 0.5	6.3± 1.0	5.2 ± 0.3	–	1.1 ± 0.2	–	0.6 ± 0.1	2.3 ± 0.2	4.8± 0.4	2.1 ± 0.2	8.7 ± 0.3
anteiso-$C_{17:0}$ 2-OH	17.970	tr	tr	tr	–	–	–	–	–	–	–	–	–	3.1 ± 0.2
3-Hydroxy														
$C_{12:0}$ 3-OH	13.454	–	–	–	0.8 ± 0.1	–	–	–	–	–	–	–	–	–
$C_{17:0}$ 3-OH	18.536	–	–	–	–	0.5 ± 0.1	–	–	–	–	–	–	–	tr
iso-$C_{11:0}$ 3-OH	12.089	–	–	–	0.5 ± 0.1	tr	tr	–	–	–	–	–	–	–
iso-$C_{13:0}$ 3-OH	14.109	tr	tr	tr	–	1.1 ± 0.2	tr	–	1.2 ± 0.4	–	–	–	–	tr
iso-$C_{15:0}$ 3-OH	16.134	tr	–	–	2.2 ± 0.2	1.1 ± 0.2	–	–	–	1.3 ± 0.3	0.8 ± 0.1	1.0 ± 0.1	0.7 ± 0.2	–
iso-$C_{16:0}$ 3-OH	17.150	–	–	–	1.3 ± 0.2	tr	–	–	–	–	–	–	–	–
iso-$C_{17:0}$ 3-OH	18.161	0.9 ± 0.2	0.6 ± 0.1	tr	8.4 ± 0.6	5.8 ± 0.4	tr	–	–	tr	tr	–	–	–
Diol														
iso-$C_{15:0}$ 1,2-diol	16.090	–	0.6 ± 0.1	–	–	–	tr	1.0 ± 0.2	2.0 ± 0.5	–	–	–	tr	tr
iso-$C_{18:0}$ 1,2-diol	19.060	–	0.7 ± 0.2	1.2 ± 0.3	–	tr	tr	–	–	–	tr	–	–	tr
Summed feature 9	16.416	5.0 ± 0.3	2.7 ± 0.2	2.1 ± 0.1	7.0 ± 1.0	–	2.5 ± 0.1	0.5 ± 0.1	–	–	–	–	–	–
Summed feature 4	16.486	0.7 ± 0.1	tr	–	1.6 ± 0.3	–	tr	–	–	–	–	–	–	–
Unknown 14.502	14.502	tr	tr	–	3.0 ± 0.3	–	tr	–	–	–	–	–	–	–

Results are the percentage of the total fatty acids. ±, results are the mean plus the standard deviation of two to four analyses; values for fatty acids present at less than 0.5% in all strains are not shown; tr, trace (< 0.5%); –, not detected; ECL, equivalent chain length. A summed feature represents groups of two or three fatty acids that could not be separated by GLC with the MIDI System: summed feature 9 comprises iso-$C_{17:1}$ ω9*c* and/or $C_{16:0}$ 10-methyl; summed feature 4 comprises iso-$C_{17:1}$ I and/or anteiso-$C_{17:1}$ B.

*The double-bond position of this fatty acid is not known.

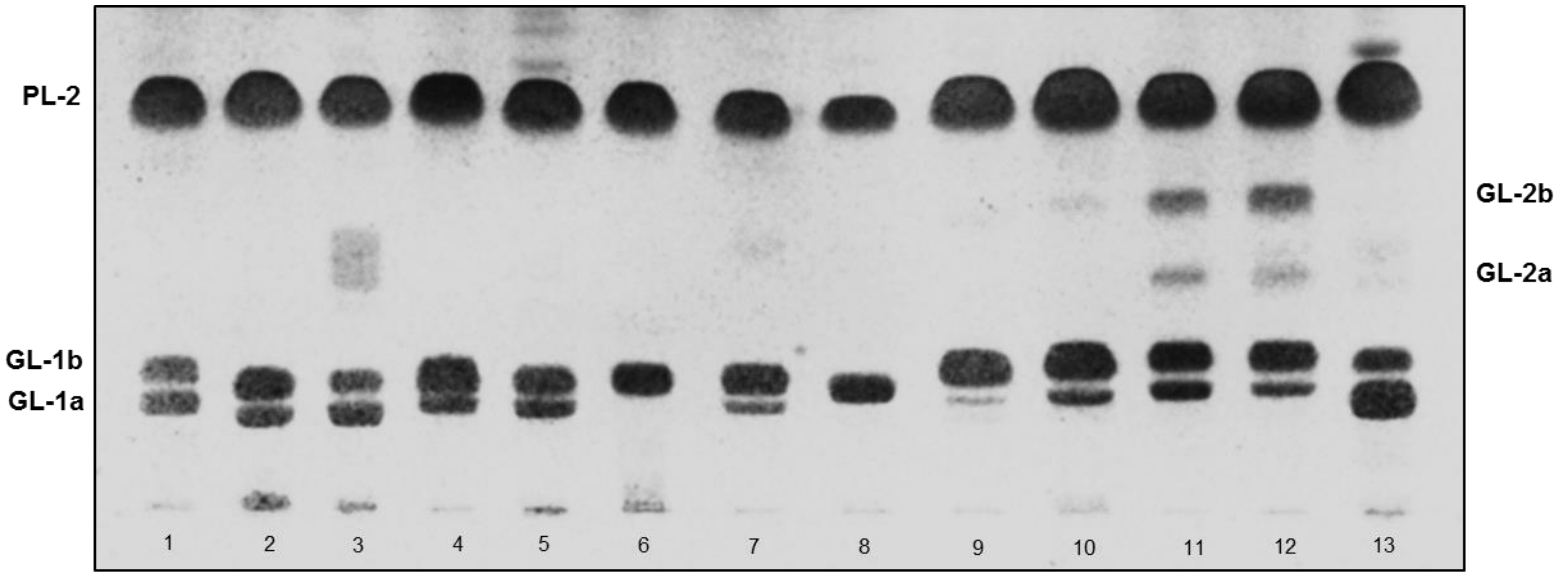

Figure 4.6 One-dimensional TLC of polar lipids of the species of the genus *Meiothermus* grown in *Thermus* liquid medium at 50ºC until late-exponential phase of growth. The lipids were stained by spraying with 5% molybdophosphoric acid in ethanol followed by heating at 160ºC. GL-1a, glycolipid 1a; GL-1b, glycolipid 1b; GL-2a, glycolipid 2a; GL-2b, glycolipid 2b; PL-2, phospholipid 2. Strains: 1, *M. ruber* ATCC25948^T; 2, *M. cateniformans* JCM 15151^T; 3, *M. taiwanensis* DSM 14542^T; 4, *M. cerbereus* DSM 11376^T; 5, *M. luteus* KCTC 52599^T; 6, *M. rufus* CAL-4^T (=DSM 22234^T); 7, *M. hypogaeus* DSM 23238^T; 8, *M. granaticius* AF-68^T (=DSM 23260^T); 9, *M. roseus* NBRC 110900^T; 10, *M. timidus* SPS-243^T (=DSM 17022^T); 11, *M. terrae* DSM 26712^T; 12, *M. chliarophilus* ATCC 700543^T; 13, *M. silvanus* ATCC 700542^T.

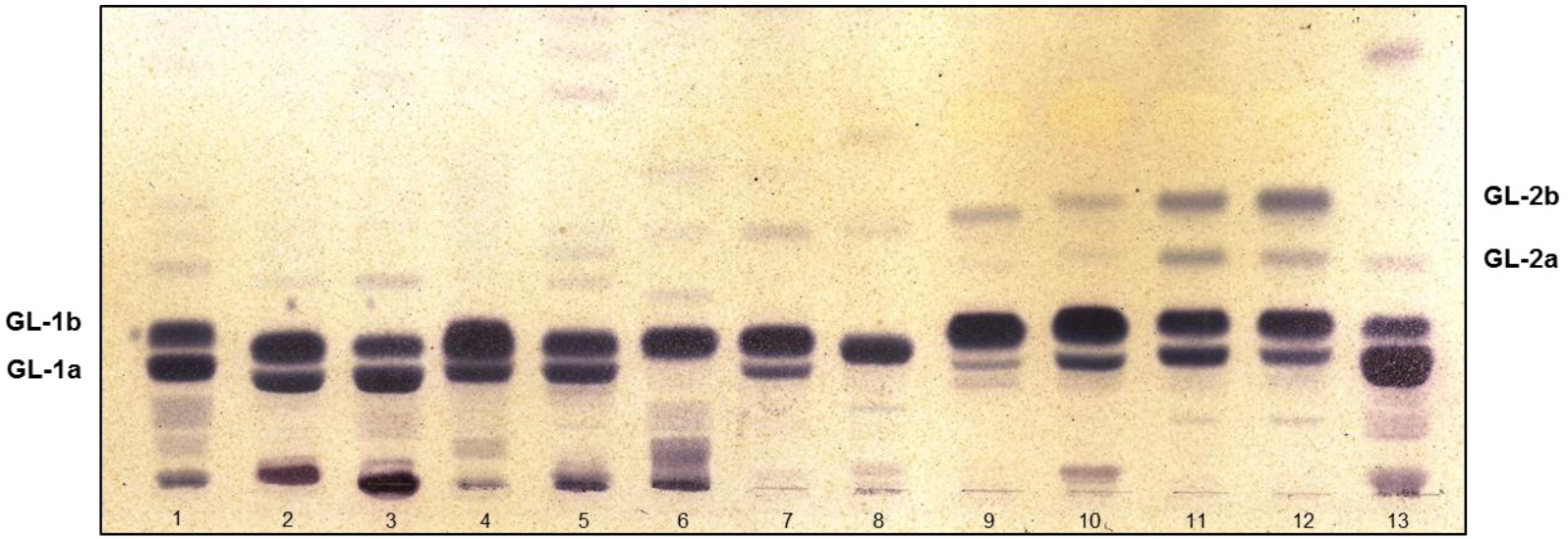

Figure 4.7 One-dimensional TLC of polar lipids of the species of the genus *Meiothermus* grown in *Thermus* liquid medium at 50ºC until late-exponential phase of growth. The lipids were stained by spraying with α-naphthol-sulfuric acid followed by heating at 120ºC. GL-1a, glycolipid 1a; GL-1b, glycolipid 1b; GL-2a, glycolipid 2a; GL-2b, glycolipid 2b. Strains: 1, *M. ruber* ATCC25948^T; 2, *M. cateniformans* JCM 15151^T; 3, *M. taiwanensis* DSM 14542^T; 4, *M. cerbereus* DSM 11376^T; 5, *M. luteus* KCTC 52599^T; 6, *M. rufus* CAL-4^T (=DSM 22234^T); 7, *M. hypogaeus* DSM 23238^T; 8, *M. granaticius* AF-68^T (=DSM 23260^T); 9, *M. roseus* NBRC 110900^T; 10, *M. timidus* SPS-243^T (=DSM 17022^T); 11, *M. terrae* DSM 26712^T; 12, *M. chliarophilus* ATCC 700543^T; 13, *M. silvanus* ATCC 700542^T.

4.4.3 Insights from the genomes of members of the genus *Meiothermus*

To support our phylogenetic and genomic differences observed, we searched the genomes of all type strains of species of the genus *Meiothermus* for several pathways for insights that would lead to the reclassification of some species as member of different genera. Most genomes had very similar or identical genes and pathways. The G+C content of the DNA of the type strains of species of the genus *Meiothermus*, calculated from the draft genome sequences, ranged between 61.0% in *M. hypogaeus* to 69.5% in *M. terrae*. Actually, this content was consistent with the distinct lineages that we observed. On the one hand, the lineage of *M. chliarophilus*, *M. roseus*, *M. terrae* and *M. timidus* always showed the highest G+C % values of greater than 65.3%, whereas the other lineages had values of less than 65.1% from *M. luteus*, but in general the G+C ratio of the other members of the genus *Meiothermus* was less than 63.5%, indicating that the former lineage had a high G+C content (Table 4.3). The size of the draft genomes of the members of the genus *Meiothermus* ranged from 2.75 Mbp in *M. rufus* to 4.69 Mbp in *M. chliarophilus* (Table 4.3).

The Embden-Meyerhof-Parnas pathway was deemed complete and able to catabolize the sugars to acetate in all strains. However, the gene coding for lactate dehydrogenase (EC1.1.1.27) was only identified in the yellow-pigmented type strains. Unfortunately, growth on lactate has only been examined in a few type strains and its significance for the taxonomy of these organisms is unknown. None of the type strains of the species of the genus *Meiothermus* have putative genes that code for enzymes of the Entner-Doudoroff pathway, but all strains have genes indicating that the pentose phosphate cycle is active and can channel glyceraldehyde-3-phosphate into the Embden-Meyerhof-Parnas pathway. The genes coding for the citric acid cycle enzymes are present in all type strains, although the yellow-pigmented species as well as *M. silvanus* lack the genes for the ancillary enzymes 2-oxacid oxidoreductase (EC 1.2.7.11) and 2-ketoglutarate ferredoxin oxidoreductase (1.2.7.3). Oxidative phosphorylation occurs via NADH dehydrogenase, succinate dehydrogenase and cytochrome *c* oxidase. The ATPase of all strains are of the V/A type.

The type strains of most species of the genus *Meiothermus* can reduce nitrate to nitrite via NarGHI or NasAB (in *M. rufus* and *M. luteus*), however, *M. ruber*, *M. cerbereus*, *M. cateniformans* and *M. taiwanensis* do not have these genes for nitrate reduction. These results confirm the phenotypic observation that the latter group of type strains do not reduce nitrate to nitrite. The type strains of *M. chliarophilus*, *M. roseus* and *M. terrae* also possess a *nirK* gene (nitrite reductase, NO-forming, EC 1.7.2.1) that could lead nitrite reduction to nitric oxide, but *M. timidus* does not.

With the exception of the genome of *M. granaticius*, all other strains predict a Calvin-Benson-Basham (CBB) cycle with a type I RuBisCO (EC 4.1.1.39) lacking a sedoheptulose-bisphosphatase (EC 3.1.3.37) gene that leads to the synthesis of sedoheptulose-phosphate. The strains of species of the genus *Meiothermus* are generally red-pigmented due to the synthesis of canthaxanthin and 1′-*β*-glucopyranosyl-3,4,3′,4′-tetradehydro-1′,2′-dihydro-*β*,*ψ*-caroten-2-one (Burgess *et al*., 1999). The pathway for the synthesis of canthaxanthin has been examined by genome analysis in an *M. taiwanensis* strain, designated RP (Mukherjee *et al*., 2016). We found that this pathway is present

in all red-pigmented strains where the gene coding for *β*-carotene ketose leading to the synthesis of canthaxanthin from *β*-carotene is present, but is interrupted in the yellow-pigmented strains at the level of *β*-carotene (Kim *et al*., 2010). All strains of species of the genus *Meiothermus* possess the gene that codes for lycopene *β*-cyclase (EC 5.5.1.19) that leads to the synthesis of the yellow-pigmented 7,8-dihydro-*β*-carotene from neurosporene. The genes leading to synthesis of the thermozeaxanthins through zeaxanthin in strains of species of the genus *Thermus* were not identified in the strains of the genus *Meiothermus* (Yokoyama *et al*., 1995).

The genes of several ABC transport systems for sorbitol/mannitol, maltose/maltodextrin, raffinose/stachyose/melibiose, trehalose/maltose, glucose/mannose, ribose, fructose, D-xylose and lactose/L-arabinose were identified in the type strains of species of the genus *Meiothermus*. Most ABC transporters discussed here have four subunits composed of a substrate binding protein, two permease proteins and an ATP binding protein, while the ABC transport systems for fructose and xylose have three subunits, having only one permease (Boos and Shuman, 1998). By and large the presence or absence of some components of the transport systems are consistent with the ability of type strains to grow on single hexoses, pentoses and polyols (76%), where growth was examined using the identical methods (Albuquerque *et al*., 2018e). However, the lactose/L-arabinose transport system lacks the ATP binding protein gene *lacK* in all the genomes, but most strains lacking this component grow on lactose (Albuquerque *et al*., 2018e) indicating that this component may be shared among other ABC transport systems or that there may be another unidentified transport system.

Sox genes for the oxidation of reduced sulfur compounds were not found in the genomes of any of the type strains. Some type strains of species of the genus *Meiothermus* are catalase-positive, while others are not. We identified manganese catalase (EC 1.11.1.6) genes in the genome of *M. ruber*, *M. taiwanensis*, *M. cateniformans*, *M. luteus*, *M. rufus*, *M. roseus* and *M. timidus* all of which are catalase-positive (Albuquerque *et al*., 2018e; Habib *et al*., 2017). The type strain of *M. granaticius* is also catalase-positive but has a gene for a bifunctional catalase/peroxidase (EC 1.11.1.21) not identified in any of the other strains. The other type strains, namely *M. terrae*, *M. chliarophilus*, *M. cerbereus*, *M. hypogaeus* and *M. silvanus* are catalase-negative and do not possess these two genes.

Table 4.3 Genome sequence information of members of the genus *Meiothermus*. *M. cateniformans* JCM 15151^T (QWKX00000000), *M. taiwanensis* DSM 14542^T (AXWR00000000.1), *M. ruber* DSM 1279^T (CP001743.1), *M. cerbereus* DSM 11376^T (JHVI00000000.1), *M. hypogaeus* DSM 23238^T (QWKY00000000), *M. rufus* DSM 22234^T (AUHY00000000.1), *M. granaticius* AF-68^T (=DSM 23260^T) (QWLB00000000), *M. luteus* KCTC 52599^T (QWKZ00000000), *M. chliarophilus* DSM 9957^T (AUQW00000000.1), *M. terrae* DSM 26712^T (QXDL00000000), *M. roseus* NBRC 110900^T (QWLA00000000), *M. timidus* DSM 17022^T (ARDL00000000.1), *M. silvanus* DSM 9946^T (CP002042.1, CP002043.1 and CP002044.1).

	M. cateniformans	***M. taiwanensis***	***M. ruber***	***M. cerbereus***	***M. hypogaeus***	***M. rufus***	***M. granaticius***	***M. luteus***	***M. chliarophilus***	***M. terrae***	***M. roseus***	***M. timidus***	***M. silvanus***
Assembled genome size (Mbp)	2.92	3.02	3.10	3.03	3.68	2.75	3.27	2.88	4.69	4.43	3.68	3.19	3.72
Protein-coding genes	2798	2824	3015	2827	3476	2582	3220	2864	4161	4065	3556	3021	3476
Finishing quality	HQD*	HQD*	Finished	HQD*	HQD*	HQD*	HQD*	HQD*	HQD*	HQD*	HQD*	HQD*	HQD*
Completeness	99.57	99.57	99.57	100.00	97.01	100.00	96.58	97.44	97.91	99.79	100.00	99.57	99.79
Contamination	0.00	0.00	0.00	0.00	0.00	0.43	0.00	1.71	1.29	1.50	0.43	0.00	0.14
DNA G+C content (%)	63.4	63.5	63.4	61.4	61.0	63.1	63.0	65.1	68.9	69.5	65.8	65.3	62.7
rRNA genes													
5S	1	1	2	3	1	1	1	1	2	1	1	1	2
16S	1	4	2	4	1	2	1	1	1	1	1	3	2
23S	1	2	2	2	1	2	1	1	1	1	1	1	2

*HQD – High-quality draft genome, corresponding to an overall coverage representing at least 90% of the genome.

4.4.4 Emended description of a genus, description of a novel genus and reclassification of four species

The fatty acid composition and the polar lipids do not clarify the classification of the species of the genus *Meiothermus*. The yellow pigmentation of the colonies of the four species on all media tested seems to be the only phenotypic characteristic that is stable among the species of *M. chliarophilus*, *M. roseus*, *M. terrae* and *M. timidus*. It is noteworthy that *M. timidus* strains were isolated from two different sites, one at São Pedro do Sul in mainland Portugal and the other on the Island of São Miguel, that are all yellow-pigmented (Pires *et al*., 2005a). Moreover, the genomic results indicate that the yellow-pigmented type strains have a higher G+C content than the other type strains of species of the genus *Meiothermus*, possess genes coding for lactate dehydrogenase, generally possess the gene encoding the enzyme that reduces nitrite to nitric oxide and do not have the gene coding for *β*-carotene ketose leading to the synthesis of canthaxanthin.

We are of the opinion that the genus *Meiothermus* should be circumscribed to the species *M. ruber*, *M. cateniformans*, *M. taiwanensis*, *M. cerbereus*, *M. hypogeus*, *M. luteus* and *M. rufus* on the basis of the genomic results. One species with low AAI values, namely *M. granaticius* can be maintained in the genus *Meiothermus* because no distinctive phenotypic characteristics are known to classify this species as a member of a separate genus. The species *M. silvanus* cannot, at present, be reclassified as a member of a separate genus, because no distinctive phenotypic characteristics are available, although the phylogenetic analysis strongly indicates that this organism represents a member of a separate genus. On the other hand, the species, *M. chliarophilus*, *M. roseus*, *M. terrae* and *M. timidus*, on the basis of the distinctive yellow pigmentation of all strains examined and the genomic results should be classified in a separate genus for which we offer the name *Calidithermus* gen. nov. for which, the type species is *Calidithermus chliarophilus*. The alteration in the classification of the species of the genus *Meiothermus* also requires that the description of this genus should be emended.

4.4.4.1 Emended description of the genus *Meiothermus*

Meiothermus (Nobre *et al*., 1996[VP]; Emend Raposo *et al*., 2019).

Characteristics are as given in the description of the genus by Nobre *et al*. (1996). The colonies are red-, pink- or orange-red-pigmented and possess genes for the synthesis of canthaxanthin. Strains do not possess the gene coding for lactate dehydrogenase. Many strains do not reduce nitrate via NarGHI (dissimilatory nitrate reductase) or NasAB (assimilatory nitrate reductase). Most strains produce GL-1 variants (GL-1a and GL-1b), but some only produce GL-1b variant. The G+C content of the DNA ranges from 61.0 to 65.1% (genome sequence). The type species of the genus is *Meiothermus ruber*.

4.4.4.2 Description of *Calidithermus* gen. nov.

Calidithermus (Raposo *et al*., 2019[VP]).

Ca.li.di.ther' mus. L. adj. *calidus* warm; N.L. masc. n. *Thermus* a bacterial genus; N.L. masc. n. *Calidithermus,* a lukewarm *Thermus*).

Non-motile rod-shaped cells that stain Gram-negative with variable length forming filaments. Form yellow-pigmented colonies due to the lack of the gene coding for *β*-carotene ketose leading to the synthesis of canthaxanthin. Moderately thermophilic, with optimum growth temperatures of about 50 to 60ºC. Growth is chemoorganotrophic on sugars, polyols, organic acids and amino acids. Aerobic with a strictly respiratory type of metabolism; but all strains reduce nitrate to nitrite by nitrate reductase. Cytochrome *c* oxidase-positive and catalase-variable. Most strains also possess a *nirK* gene that codes for a nitrite reductase (NO-forming). All strains possess the gene coding for lactate dehydrogenase. Fatty acids are primarily iso- and anteiso-branched; 2-OH and 3-OH fatty acids are also present. MK-8 is the only respiratory lipoquinone identified. One major PL (PL-2) and two GLs variants GL-1a and GL-1b, are present.

The type species of the genus is *Calidithermus chliarophilus.* The genus *Calidithermus* belongs to the family *Thermaceae* (Nobre *et al*., 1996). The G+C content of the DNA range is about 65.3–69.5% (determined from genome sequences).

4.4.4.3 Description of *Calidithermus chliarophilus* comb. nov.

Calidithermus chliarophilus (Raposo *et al*., 2019[VP]). Basonym: *Thermus chliarophilus* Tenreiro *et al*. (1995); *Meiothermus chliarophilus* Nobre *et al*. (1996).

The description of *Calidithermus chliarophilus* is based on data from Tenreiro *et al.* (1995), Nobre *et al*. (1996) and Albuquerque *et al*. (2018e). The optimum growth temperature is about 50°C. Possess the *nirK* gene. Catalase-negative, lacks the manganese catalase gene.

The type strain is *Calidithermus chliarophilus* ALT-8[T] (=DSM 9957[T] =ATCC 700543[T] =BCRC 17113[T] =NCIMB 13439[T] =NBRC 106474[T]). Accession number of 16S rRNA gene for *M. chliarophilus* DSM 9957[T] is X84212. Accession number for genome is AUQW00000000.1. The G+C content of the DNA is 69.5%, determined from the genome sequence.

4.4.4.4 Description of *Calidithermus roseus* comb. nov.

Calidithermus roseus (Raposo *et al*., 2019[VP]). Basonym: *Meiothermus roseus* Ming *et al*. (2016).

The description of *Calidithermus roseus* is based on data from Ming *et al*. (2015) and Albuquerque *et al*. (2018e). The optimum growth temperature is about 50°C. Possess the *nirK* gene. Catalase-positive, the manganese catalase gene is present.

The type strain of *Calidithermus roseus* is YIM 71031^{T} (=NBRC 110900^{T}= KCTC 42495^{T}). Accession number of 16S rRNA gene for *M. roseus* NBRC 110900^{T} is KP232921. Accession number for genome is QWLA00000000. The G+C content of the DNA is 65.8%, determined from the genome sequence.

4.4.4.5 Description of *Calidithermus terrae* comb. nov.

Calidithermus terrae (Raposo *et al*., 2019^{VP}). Basonym: *Meiothermus terrae* Yu *et al*. (2014).

The description of *Calidithermus terrae* is based on data from Yu *et al*. (2014) and Albuquerque *et al*. (2018e). The optimum growth temperature is about 50°C. Possess the *nirK* gene. Catalase-negative, lacks the manganese catalase gene.

The type strain of *Calidithermus terrae* is YIM 77755^{T} (=DSM 26712^{T} =CCTCC AB 2012942^{T}). Accession number of 16S rRNA gene for *M. terrae* DSM 26712^{T} is X84212. Accession number for genome is QXDL00000000. The G+C content of the DNA is 69.5%, determined from the genome sequence.

4.4.4.6 Description of *Calidithermus timidus* comb. nov.

Calidithermus timidus (Raposo *et al*., 2019^{VP}). Basonym: *Meiothermus timidus* Pires *et al*. (2005b).

The description of *Calidithermus timidus* is based on data from Pires *et al*. (2005a) and Albuquerque *et al*. (2018e). The optimum growth temperature is about 55–60ºC. Catalase-positive, the manganese catalase gene is present. Does not possess the *nirK* gene.

The type strain of *Calidithermus timidus* is SPS-243^{T} (=DSM 17022^{T} =CIP 108604^{T} =LMG 22897^{T} =NBRC 103207^{T}). Accession number of 16S rRNA gene for *M. timidus* DSM 17022^{T} is AJ871168. Accession number for genome is ARDL00000000.1. The G+C content of the DNA is 65.3%, determined from the genome sequence.

Chapter 5

Conclusions and Future Perspectives

5.1 Conclusions

A deeper understanding of the microbiology of the São Pedro do Sul hot spring was achieved in this study. The use of genomic information of strains confirmed the vital contribution of genomics in taxonomic studies to improve prokaryotes classification.

Initial studies focused on the characterization of new isolates retrieved from biofilm and water samples. One orange-pigmented colony forming isolate, strain SPSPC-11^T, was designated *Raineya orbicola* gen. nov., sp. nov. and forms rod-shaped cells and long filaments during the exponential phase of growth. This species is slightly thermophilic, with an optimum growth temperature of about 50ºC and a temperature range for growth between 35ºC and 60ºC. The fatty acids of this species are primarily iso-branched and hydroxy fatty acids, a characteristic of the new genus *Raineya*, and the major respiratory quinone is MK-7, a characteristic of the novel family *Raineyaceae*. Two other new isolates, strains SPSP-6^T and SPSPC-18, that form non-pigmented colonies, were designated *Tepidimonas charontis* sp. nov. These organisms also form rod-shaped cells but shorter than the species *Raineya orbicola* and are also slightly thermophilic, with the same optimum growth temperature of about 50ºC but with a wider temperature range for growth between 25ºC and 60ºC. The fatty acids of species *T. charontis* are primarily straight-chain saturated and monounsaturated, characteristic common to all species of the genus *Tepidimonas*.

The genome of *Raineya orbicola* was sequenced and compared with the genome sequences of closely related microorganisms to gain insights into their metabolism. The analysis of the draft genome of *Raineya orbicola* SPSPC-11^T indicated that the genes involved in nitrate/nitrite transport and nitrate reduction, namely the assimilatory nitrate reductase and the enzymes for denitrification were not encountered, confirming the absence of phenotypic nitrate reduction in strain SPSPC-11^T. The lack of several genes encoding for enzymes involved in the initial catabolism of carbohydrates to glucose also confirmed that strain SPSPC-11^T was unable to grow on any of the carbohydrates examined indorsing that sugars do not serve as carbon and energy sources for growth.

The analysis of the assembled genomes of all the type strains of the validly named species of the genus *Tepidimonas*, of strain PL17 of *T. fonticaldi*, strains MB2 and VT154-175 of *T. taiwanensis* and strain SPSP-6^T of *T. charontis* elucidated the inability of the species of this genus to grow on hexoses with the exception of the type strain I1-1^T of the species *T. taiwanensis*. This strain is the only strain of the genus able to use glucose and fructose as carbon and energy sources for growth, as experimentally verified. Glucose and fructose transporters were only identified in the genome of the type strain of *T. taiwanensis.* In strains MB2 and VT154-175, only putative glucose and fructose transporters were identified, however, in these latter strains the growth on hexoses was not examined. Also, genes encoding for the pentose-phosphate pathway enzymes were identified in the genomes of *T. taiwanensis* strains I1-1^T, MB2 and VT154-175 but were not identified in any of the other genomes. However, strain I1-1^T, like all tested strains of the genus *Tepidimonas*, does not grow on any other carbohydrates examined, such as mannose, galactose, trehalose, maltose, sucrose, ribose, L-arabinose, xylose, or polyols. The genes that could channel these carbohydrates

to the Emden-Meyerhof-Parnas or the pentose phosphate pathways were not identified in any strains of the genus *Tepidimonas*, confirming the phenotype. The strains of the members of the genus *Tepidimonas* have variable genes involved in nitrogen metabolism. The species *T. fonticaldi* possesses the most complete set of genes of the species of this genus for nitrogen metabolism, corroborating the experimental reduction of nitrate to nitrate. On the other hand, only one gene involved in the reduction of nitrate was identified in the type strain of *T. charontis*, confirming the phenotypic absence of nitrate reduction. In all species where thiosulfate oxidation has been tested, thiosulfate was oxidized to sulfate in the presence of an organic carbon source, with an increase in the biomass detected in some species, indicating that these organisms are mixotrophic. All the genome sequences analysed predict that thiosulfate is oxidized to sulfate via the sox pathway. The lack of agreement regarding starch hydrolysis by strain I1-1^T of *T. taiwanensis* was elucidated. The absence of starch hydrolyzing-enzymes in the genomes of any strain of members of the genus *Tepidimonas* predicts strain I1-1^T is not capable of degrading starch.

The comparative genomic analysis based on the OGRI and the phylogenomic approach allowed the classification of the new species *T. charontis* as a member of the genus *Tepidimonas*. This was only possible through the use of genome analysis tools since a small number of phenotypic and chemotaxonomic characteristics distinguish this new isolate from the other type strains of the species of the genus *Tepidimonas*. The ANIb and dDDH values corroborated the 16S rRNA sequence analysis results indicating strain SPSP-6^T as a new species, and the AAI values indicated that this species belongs to the genus *Tepidimonas*. The ANIb and dDDH values for *T. fonticaldi* AT-A2^T and strain PL17 indicated a very close relationship between these strains, the same for *T. taiwanensis* I1-1^T and strains MB2 and VT154-175. The phylogenetic study with the 16S rRNA gene and the 400 conserved genes sequence analysis circumscribed all type strains of the genus *Tepidimonas*, the new isolate strain SPSP-6^T, as well as strains PL17, MB2 and VT154-175 to the genus *Tepidimonas*.

The search in all the assembled genomes of all the type strains of the validly named species of the genus *Meiothermus* for the genes encoding for the synthesis of pigments revealed that strains that produced the red, pink, or orange pigments characteristic of the colonies possess the genes for the synthesis of canthaxanthin. Strains that form yellow-pigmented colonies lack the gene for *β*-carotene ketose, which leads to the synthesis of canthaxanthin. This distinct phenotypic feature correlated with the genome sequences of the strains. The presence or absence of other genes like the gene encoding for lactate dehydrogenase or the genes involved in the nitrogen metabolism interrelated with the comparative genomic analysis and with the phylogenomic analysis leads to the emended description of the genus *Meiothermus*, with the reclassification of the species *M. chliarophilus*, *M. roseus*, *M. terrae* and *M. timidus* and to the description of the new genus *Calidithermus* to harbour these four yellow-pigmented species.

Complementing the distinct phenotypic characteristic, the colony colour of the strains of the species of the genus *Meiothermus*, with the G+C content of the DNA, with the results of the OGRI and with the phylogenetic analysis of the 16S rRNA gene, phylogenetic analysis of 90 housekeeping

genes and 855 core-genes, supported the reclassification of this genus. The phylogenetic analysis based on 16S rRNA gene sequence of all type strains indicated the species of the genus *Meiothermus* form a separate monophyletic line of descent from the species of the genus *Thermus*, but split it into three lineages representing distinct genera. The phylogenetic reconstructions based on the concatenates of the complete core-genome of 855 shared genes of all genomes, and also that of a subset of the 90 housekeeping genes present in all genomes, were both consistent with the 16S rRNA gene lineage distinctness. The heterogeneous nature of the genus was also supported by the values of ANIb, dDDH and AAI, delineating the type species of the genus *M. ruber*, and the type strains of *M. cateniformans*, *M. taiwanensis*, *M. cerbereus*, *M. luteus*, *M. rufus*, *M. hypogaeus* and *M. granaticius* in a group, a second lineage comprising the yellow-pigmented strains, and one most distant lineage formed by the single species *M. silvanus*. The fatty acid and the polar lipid compositions did not clarify the classification of the species of this genus, not allowing the species *M. silvanus* to be reclassified as a member of a separate genus, because no distinctive phenotypic characteristics are available, although the phylogenetic analysis strongly indicates that this organism represents a member of a separate genus.

Therefore, it is clear that in this study the integration of genomics into prokaryotic taxonomy provides more robust and stable conclusions about the characterization and phylogenetic assignment of new taxa and existing taxa. For example, only with the use of genomic information and comparative genomic analysis was it possible to describe the new species *Tepidimonas charontis* and reclassify some species of the genus *Meiothermus* as members of the new genus *Calidithermus*.

Similar approaches were recently used in the classification of bacteria assigned to the phylum *Actinobacteria*, one of the largest lineages in the domain *Bacteria* (Nouioui *et al*., 2018; Salam *et al*., 2020). The inconsistencies that occurred in the 16S rRNA gene phylogeny were resolved using phylogeny based on concatenated sequences of several universal protein marker genes retrieved from genome sequences (Salam *et al*., 2020). The phylogenomics and the comparative genomic analysis based on AAI allowed several new descriptions and emended descriptions of the higher taxa (class, order, family) of the phylum *Actinobacteria* (Salam *et al*., 2020). This taxonomic approach was also applied to the genus *Bacillus*, an extremely diverse group of bacteria within the phylum *Firmicutes*, supporting the existence of six distinct clades representing six novel genera (Patel and Gupta, 2020) and of seventeen distinct clades representing seventeen novel genera (Gupta *et al*., 2020). Gupta *et al*., 2020 proposed an emended description of the genus *Bacillus* to restrict it to only members of the "Subtilis" clade that contains the type species of the genus *B. subtilis*, and "Cereus" clade that contains many important human pathogens, like *B. anthracis* and *B. cereus* (Bhandari *et al*., 2013; Liu *et al*., 2015). Also, the study of de la Haba *et al*., 2019 clarified the relationships of several strains, isolated from hypersaline environments with available genomic information and review their taxonomic affiliation to the genus *Salinivibrio* of the family *Vibrionaceae* of the phylum *Proteobacteria*. Phylogenomics based on core- and pan-genome, ANI, AAI and synteny analyses showed an elevated level of genetic relatedness of *Salinivibrio* strains

within members of the genus *Salinivibrio*, and the phenotypic characteristics described for the member of this group were in consensus with the information retrieved from the annotated genomes, guiding the classification of all the *Salinivibrio* strains with available genomes in seven separated species (de la Haba *et al*., 2019). Likewise, in the domain *Archaea*, a comparative taxonomic study of four species of the genus *Halorubrum* based on several approaches like 16S rRNA gene sequence analysis, MLSA, phylogenomic analysis based on core-genome, ANI, dDDH, synteny analysis and polar lipid profile revealed that *H. distributum*, *H. terrestre*, *H. arcis* and *H. litoreum* constitute a single species. The latter three were considered synonyms of *H. distributum* based on the rules for priority of names, which led to the emended description of the species *H. distributum* (Infante-Domíngez *et al*., 2020). Already in 2018, Corral and collaborators started to reclassify the genus *Halorobrum* based on phenotypic, chemotaxonomic and comparative genomic studies (Corral *et al*., 2018).

Ongoing genomics studies in our research group may lead to taxa reclassification. The genus *Thermus* which, like the genera *Meiothermus* and *Calidithermus* belong to the family *Thermaceae* of the phylum *Deinococcus-Thermus*, is characterized by its thermophilic species (Albuquerque *et al*., 2018c). The distinctness of the species in some phenotypic characteristics complementing with phylogenetic position and comparative genomic studies can led to the reclassification of the genus *Thermus* (personal Albuquerque L. unpublished results). Also the genus *Rubrobacter* of the phylum *Actinobacteria* that comprise thermophilic and mesophilic species and some species resistant to ionizing radiation and desiccation (Albuquerque *et al*., 2014), may be subject of reclassification (personal Albuquerque L. unpublished results).

All these recent studies in the *Bacteria* and *Archaea* domains support that the integration of genomics into prokaryotic taxonomy established an important step towards the improvement of the taxonomy of prokaryotes.

5.2 Future Perspectives

This work contributed to increasing the knowledge on the microbiology of the extreme environments, specifically of the thermophilic environment of the São Pedro do Sul hot spring. However, many of the microorganisms that live in this kind of environment cannot be cultivated using established laboratory methods, thus requiring alternative approaches for their isolation and characterization. The main drawback of culture-dependent methods is that it is not possible to reproduce all the culture conditions of the environment to isolate and grow all the microorganisms in the laboratory successfully. Moreover, microorganisms occurring in low numbers are often outcompeted *in vitro* by numerically more abundant organisms and some of them may be unable to grow in the laboratory at all. Additionally, culture-dependent methods can be time-consuming due to long culture periods and to elaborate culture techniques. New procedures have been developed to allow the culture-independent analysis of the totality of microbial genomes in a particular environment, called the metagenome, to overcome culture-dependent limitations. Metagenomics is a culture-independent genomic analysis of microbial communities, divided into sequence-based and function-driven analyses, and these two branches of metagenomics address the challenge of studying microbial communities and functions in several environments (Ramganesh *et al*., 2014; Mirete *et al*., 2016). Despite the high-throughput sequencing approaches of metagenomics enabling advances in exploring microbial diversity, their inability to isolate the unknown content of microbial communities, and the need for living cultures have allowed the rebirth of culture approaches (Diakite *et al*., 2020). Culturomics, a large-scale culturing methods based on the use of a huge variety of culture conditions and media, was improve along with metagenomics to keep pace with the advances in exploring microbial diversity (Browne *et al*., 2016; Diakite *et al*., 2019). Joining the genomic analysis of microbial communities with an optimized culture-dependent approach can enable the isolation of a maximal diversity of microorganisms. Inferring the hidden traits of uncultured populations from metagenomes can provide clues for enrichment and isolation (Karthikeyan *et al*., 2019; Zhang *et al*., 2019). Identifying novel genes retrieved from the environment through functional metagenomics, combined with further biochemical studies, may provide deeper insights into the molecular elucidation of diverse microbial processes under extreme conditions that can be useful in designing biotechnological processes (Pascoal *et al*., 2020). The development of metagenomics and the single-cell amplified genome, the assembly of metagenome-assembled genomes (MAGs) and single-amplified genomes (SAGs), respectively, has led to an increase in genome-based discoveries of new organisms of the domains *Archaea* and *Bacteria* and to expand the knowledge on the microbial diversity of many environments. However, the description of new microorganisms based solely on genomic information raises nomenclature problems since the ICNP, that follows the recommendations of the ICSP, only recognizes pure and living cultures as type material (Parker *et al*., 2019), by this means preventing the naming of uncultivated organisms. Recently, a Consensus Statement was published with a proposal of two potential plans to solve this nomenclature challenge and develop a nomenclature system for

uncultivated microorganisms based on DNA genome sequences as the type material (Murray *et al*., 2020). The advantage of adopting MAGs and SAGs as alternative type material for uncultivated taxa is to expand the taxonomic framework to the vast uncultivated majority of prokaryotes (Konstantinidis *et al*., 2020). This issue is under discussion within the community of prokaryotic taxonomists to develop consistent rules for the nomenclature of uncultivated taxa. Presently, the challenge is to reach a consensus on the taxonomic framework and adapt the existing nomenclatural code, or create a new code, to systematically incorporate uncultured taxa into the chosen plan (Hugenholtz *et al*., 2021). The Genome Taxonomy Database (GTDB), a phylogenetically consistent genome-based taxonomy, provides classifications for bacterial and archaeal genomes from domain to species, however almost 40% of the genomes lack a species name (Parks *et al*., 2020). The GTDB identified genomes assembled from the type strain of the species (type strain genomes) and used these as the representative of species clusters circumscribed using ANI. The genomes that were not assigned to a named species cluster were organized into *de novo* species clusters with representative genomes selected based on genome quality and acting as effective nomenclatural type material following the proposal that gene sequences are suitable type material for *Bacteria* and *Archaea* (Parks *et al*., 2020).

With the technological innovations, improving tools for the characterization of bacterial and archaeal diversity, the insights into microbial diversity are continually developing with better understanding of the physiology, ecology and evolution of microorganisms, however, at this time, most prokaryotic diversity remains yet to be cultured. The global diversity and distribution of prokaryotic organisms in the biosphere remains a subject of intense controversy (Curtis *et al*., 2002; Kallmeyer *et al*., 2012; Locey and Lennon, 2016; Louca *et al*., 2019; McMahon and Parnell, 2014; Schloss *et al*., 2016; Straub *et al*., 2017). Several years ago, Whitman and collaborators estimated $4\text{-}6x10^{30}$ prokaryotic organisms present in the biosphere, and the majority occur in subseaflor sediment (Whitman *et al*., 1998). Twenty years later, Bar-On and collaborators assembled the overall biomass composition of the biosphere estimating that the second major biomass component are the Domain *Bacteria* constituting about 15% of the global biomass, where plants represent 80% of the biomass; the majority biomass of bacteria is concentrated in terrestrial deep subsurface environments, and the global biomass of archaea (± 1.3%) is distributed by terrestrial and marine deep-subsurface (Bar-On *et al*., 2018). There are some different calculations to how many prokaryotic species exist, varying widely from a more optimistic view of $1x10^{12}$ species (Dykhuizen, 1998; Locey and Lennon, 2016) to a more pessimistic view of $3x10^{4}$ species (Mora *et al*., 2011). Considering the number of $1x10^{7}$ a feasible number of species that we might have in the biosphere (Yarza *et al*., 2014), this number contrast sharply with the current number of classified species of 17,000, which represents only a minor fraction of the species catalogued (Ludwig *et al*., 2021).

To date, only four new species of *Bacteria* were isolated and characterized from S. Pedro do Sul hot spring. The new culture-independent genomic analyses of microbial communities can be used in São Pedro do Sul to increase the knowledge on the thermal spring complete microbial diversity.

This new technique can help elucidate the microbial communities of this specific environment, and with this information, new culture approaches can be developed to isolate and describe many more microorganisms that thrive in that community. São Pedro do Sul hot spring remains a great source of new microorganisms with valuable biotechnological potential.

> "*The adequacy of characterization of a bacterium is a reflexion of time; it should be as full as modern techniques make possible. Unfortunately, one now regarded as adequate is likely, in ten years time, to be hopelessly inadequate*!"
>
> Statement from the taxonomist Cowan S.T. (Cowan, 1965).

References

ABDOLLAHI, H., SHAFAEI, S.Z., NOAPARAST, M., MANIFI, Z., NIEMELÄ, S.I. and TUOVINEN, O.H. (2014). Mesophilic and thermophilic bioleaching of copper from a chalcopyrite-containing molybdenite concentrate. *Int J Miner Process* 128: 25–32. doi.org/10.1016/j.minpro.2014.02.003

ADIGUZEL, A., AY, H., BALTACI, M.O., AKBULUT, S., ALBAYRAK, S. and OMEROGLU, M.A. (2020). Genome-based classification of *Calidifontibacillus erzurumensis* gen. nov., sp. nov., isolated from a hot spring in Turkey, with reclassification of *Bacillus azotoformans* as *Calidifontibacillus azotoformans* comb. nov. and *Bacillus oryziterrae* as *Calidifontibacillus oryziterrae* comb. nov. *Int J Syst Evol Microbiol* 70: 6418–6427. doi.org/10.1099/ijsem.0.004549

AGUIAR, P., BEVERIDGE, T.J. and REYSENBACH, A.L. (2004). *Sulfurihydrogenibium azorense*, sp. nov., a thermophilic hydrogen-oxidizing microaerophile from terrestrial hot springs in the Azores. *Int J Syst Evol Microbiol* 54: 33–39. doi.org/10.1099/ijs.0.02790-0

AKANBI, T.O., JI, D. and AGYEI, D. (2020). Revisiting the scope and applications of food enzymes from extremophiles. *J Food Biochem* 44: e13475. doi.org/10.1111/jfbc.13475

ALBUQUERQUE, L. and DA COSTA, M.S. (2014). Family *Thermaceae*. In *The Prokaryotes-Other Major Lineages of Bacteria and The Archaea*. 4th Ed, Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E. and Thompson, F. (editors). Berlin Heidelberg, Springer-Verlag; pp 955–987. doi.org/10.1007/978-3-642-38954-2_128

ALBUQUERQUE, L. and EGAS, C. (2020). *Raineyaceae*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.fbm00370

ALBUQUERQUE, L. and EGAS, C. (2021a). *Raineya*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.gbm01902

ALBUQUERQUE, L. and EGAS, C. (2021b). *Tepidimonas*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.gbm00959.pub2

ALBUQUERQUE, L., CASTELHANO, N., RAPOSO, P., FROUFE, H.J.C., TIAGO, I., SEVERINO, R., ROXO, I., GREGÓRIO, I., BARROSO, C., EGAS, C., *et al*. (2020a). Comparative genome sequence analysis of several species in the genus *Tepidimonas* and description of a novel species *Tepidimonas charontis* sp. nov. *Int J Syst Evol Microbiol* 70: 1596–1604. doi.org/10.1099/ijsem.0.003942

ALBUQUERQUE, L., CASTELHANO, N., RAPOSO, P., FROUFE, H.J.C., TIAGO, I., SEVERINO, R., ROXO, I., GREGÓRIO, I., BARROSO, C., EGAS, C., *et al*. (2020b). Corrigendum: Comparative genome sequence analysis of several species in the genus *Tepidimonas* and description of a novel species *Tepidimonas charontis* sp. nov. *Int J Syst Evol Microbiol* 70: 6539. doi.org/10.1099/ijsem.0.004563

ALBUQUERQUE, L., FERREIRA, C., TOMAZ, D., TIAGO, I., VERÍSSIMO, A., DA COSTA, M.S. and NOBRE, F. (2009). *Meiothermus rufus* sp. nov., a new slightly thermophilic red-pigmented species and emended description of the genus *Meiothermus*. *Syst Appl Microbiol* 32: 306–313. doi.org/10.1016/j.syapm.2009.05.002

ALBUQUERQUE, L., FRANÇA, L., RAINEY, F.A., SCHUMANN, P., NOBRE, M.F. and DA COSTA, M.S. (2011a). *Gaiella occulta* gen. nov., sp. nov., a novel representative of a deep branching phylogenetic lineage within the class *Actinobacteria* and proposal of *Gaiellaceae* fam. nov. and *Gaiellales* ord. nov. *Syst Appl Microbiol* 34: 595–599. doi.org/10.1016/j.syapm.2011.07.001

ALBUQUERQUE, L., JOHNSON, M.M., SCHUMANN, P., RAINEY, F.A. and DA COSTA, M.S. (2014). Description of two new thermophilic species of the genus *Rubrobacter*, *Rubrobacter calidifluminis* sp. nov. and *Rubrobacter naiadicus* sp. nov., and emended description of the genus *Rubrobacter* and the species *Rubrobacter bracarensis*. *Syst Appl Microbiol* 37: 235–243. doi.org/10.1016/j.syapm.2014.03.001

ALBUQUERQUE, L., KOWALEWICZ-KULBAT, M., DRZEWIECKA, D., STACZEK, P., D'AURIA, G., ROSSELLÓ-MÓRA, R. and DA COSTA, M.S. (2016). *Halorhabdus rudnickae* sp. nov., a halophilic archaeon from a salt mine borehole in Poland. *Syst Appl Microbiol* 39: 100–105. doi.org/10.1016/j.syapm.2015.12.004

ALBUQUERQUE, L., POLÓNIA, A.R.M., BARROSO, C., FROUFE, H.J.C., LAGE, O., LOBO-DA CUNHA, A., EGAS, C. and DA COSTA, M.S. (2018a). *Raineya orbicola* gen. nov., sp. nov., a slightly thermophilic bacterium of the phylum *Bacteroidetes* and the description of *Raineyaceae* fam. nov. *Int J Syst Evol Microbiol* 68: 982–989. doi.org/10.1099/ijsem.0.002556

ALBUQUERQUE, L., RAINEY, F.A. and DA COSTA, M.S. (2018b). *Gaiella*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.gbm01469

ALBUQUERQUE, L., RAINEY, F.A. and DA COSTA, M.S. (2018c). *Thermus*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.gbm00477.pub2

ALBUQUERQUE, L., RAINEY, F.A. and DA COSTA, M.S. (2018d). *Tepidicella*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley

& Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.gbm01464

ALBUQUERQUE, L., RAINEY, F.A. and DA COSTA, M.S. (2018e). *Meiothermus*. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.gbm00476

ALBUQUERQUE, L., RAINEY, F.A., ALDINA, P., TIAGO, I., VERÍSSIMO, A., NOBRE, M.F. and DA COSTA, M.S. (2010a). *Tepidamorphus gemmatus* gen. nov., sp. nov., a slightly thermophilic member of the *Alphaproteobacteria*. *Syst Appl Microbiol* 33: 60–65. doi.org/10.1016/j.syapm.2010.01.002

ALBUQUERQUE, L., RAINEY, F.A., CHUNG, A.P., SUNNA, A., NOBRE, M.F., GROTE, R., ANTRANIKIAN, G. and DA COSTA, M.S. (2000). *Alicyclobacillus hesperidum* sp. nov. and a related genomic species from solfataric soils of São Miguel in the Azores. *Int J Syst Evol Microbiol* 50: 451–457. doi.org/10.1099/00207713-50-2-451

ALBUQUERQUE, L., RAINEY, F.A., NOBRE, M.F. and DA COSTA, M.S. (2010b). *Meiothermus granaticius* sp. nov., a new slightly thermophilic red-pigmented species from the Azores. *Syst Appl Microbiol* 33: 243–246. doi.org/10.1016/j.syapm.2010.04.001

ALBUQUERQUE, L., RAINEY, F.A., NOBRE, M.F. and DA COSTA, M.S. (2008). *Elioraea tepidiphila* gen. nov., sp. nov., a slightly thermophilic member of the *Alphaproteobacteria*. *Int J Syst Evol Microbiol* 58: 773–778. doi.org/10.1099/ijs.0.65294-0

ALBUQUERQUE, L., RAINEY, F.A., NOBRE, M.F. and DA COSTA, M.S. (2011b). *Schleiferia thermophila* gen. nov., sp. nov., a slightly thermophilic bacterium of the phylum '*Bacteroidetes'* and the proposal of *Schleiferiaceae* fam. nov. *Int J Syst Evol Microbiol* 61: 2450–2455. doi.org/10.1099/ijs.0.028852-0

ALBUQUERQUE, L., RAINEY, F.A., NOBRE, M.F. and DA COSTA, M.S. (2012a). *Hydrotalea sandarakina* sp. nov., isolated from a hot spring runoff, and emended descriptions of the genus *Hydrotalea* and the species *Hydrotalea flava*. *Int J Syst Evol Microbiol* 62: 1603–1608. doi.org/10.1099/ijs.0.034496-0

ALBUQUERQUE, L., RAINEY, F.A., NOBRE, M.F. and DA COSTA, M.S. (2012b). *Oceanicella actignis* gen. nov., sp. nov., a halophilic slightly thermophilic member of the *Alphaproteobacteria*. *Syst Appl Microbiol* 35: 385–389. doi.org/10.1016/j.syapm.2012.07.001

ALBUQUERQUE, L., SANTOS, J., TRAVASSOS, P., NOBRE, M.F., RAINEY, F.A., WAIT, R., EMPADINHAS, N., SILVA, M.T. and DA COSTA, M.S. (2002). *Albidovulum inexpectatum* gen. nov., sp. nov., a nonphotosynthetic and slightly thermophilic bacterium from a marine hot spring that is very closely related to members of the photosynthetic genus *Rhodovulum*. *Appl Environ Microbiol* 68: 4266–4273. doi.org/10.1128/AEM.68.9.4266-4273.2002

ALBUQUERQUE, L., SIMÕES, C., NOBRE, M.F., PINO, N.M., BATTISTA, J.R., SILVA, M.T., RAINEY, F.A. and DA COSTA, M.S. (2005). *Truepera radiovictrix* gen. nov., sp. nov., a new

radiation resistant species and the proposal of *Trueperaceae* fam. nov. *FEMS Microbiol Lett* 247: 161–169. doi.org/10.1016/j.femsle.2005.05.002

ALBUQUERQUE, L., TIAGO, I., NOBRE, M.F., VERÍSSIMO, A. and DA COSTA, M.S. (2013). *Cecembia calidifontis* sp. nov., isolated from a hot spring runoff, and emended description of the genus *Cecembia*. *Int J Syst Evol Microbiol* 63: 1431–1436. doi.org/10.1099/ijs.0.044537-0

ALBUQUERQUE, L., TIAGO, I., VERÍSSIMO, A. and DA COSTA, M.S. (2006). *Tepidimonas thermarum* nov., a new slightly thermophilic betaproteobacterium isolated from the Elisenquelle in Aachen and emended description of the genus *Tepidimonas*. *Syst Appl Microbiol* 29: 450–456. doi.org/10.1016/j.syapm.2005.12.004

ALEXAKI, A., KAMES, J., HOLCOMB, D.D., ATHEY, J., SANTANA-QUINTERO, L.V., LAM, P.V.N., HAMASAKI-KATAGIRI, N., OSIPOVA, E., SIMONYAN, V., BAR, H., *et al*. (2019). Codon and codon-pair usage tables (CoCoPUTs): facilitating genetic variation analyses and recombinant gene design. *J Mol Biol* 431: 2434–2441. doi.org/10.1016/j.jmb.2019.04.021

ALFREDSSON, G.A., KRISTJANSSON, J.K., HJÖRLEIFSDOTTIR, S. and STETTER, K.O. (1988). *Rhodothermus marinus* gen. nov., sp. nov., a thermophilic, halophilic bacterium from submarine hot springs in Iceland. *J Gen Microbiol* 134: 299–306.

ALLALA, F., BOUACEM, K., BOUCHERBA, N., MECHRI, S., KRIAA, M., ARKOUB-DJOUDI, W., AZZOUZ, Z., BENALLAOUA, S., HACENE, H., JAOUADI, B., *et al*. (2020). α-amylase production by *Tepidimonas fonticaldi* strain HB23: statistical optimization and compatibility study for use in detergent formulations. *Environ Sci Pollut Res* 27: 37164–37172. doi.org/10.1007/s11356-020-10142-2

ALTSCHUL, S.F., GISH, W., MILLER, W., MYERS, E.W. and LIPMAN, D.J. (1990). Basic local alignment search tool. *J Mol Biol* 215: 403–410. doi.org/10.1016/S0022-2836(05)80360-2

ALTSCHUL, S.F., MADDEN, T.L., SCHÄFFER, A.A., ZHANG, J., ZHANG, Z., MILLER, W. and LIPMAN, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402. doi.org/10.1093/nar/25.17.3389

ALVES, M.P., RAINEY, F.A., NOBRE and DA COSTA, M.S. (2003). *Thermomonas hydrothermalis* sp. nov., a new slightly thermophilic γ-proteobacterium isolated from a hot spring in central Portugal. *Syst Appl Microbiol* 26: 70–75. doi.org/10.1078/072320203322337335

AMARAL, G.R.S., DIAS, G.M., WELLINGTON-GURI, M., CHIMETTO, L., CAMPEÃO, M.E., THOMPSON, F.L. and THOMPSON, C.C. (2014). Genotype to phenotype: identification of diagnostic vibrio phenotypes using whole genome sequences. *Int J Syst Evol Microbiol* 64: 357–365. doi.org/10.1099/ijs.0.057927-0

ANDERSON, I., DUPLEX, O., DJAO, N., MISRA, M., CHERTKOV, O., NOLAN, M., LUCAS, S., LAPIDUS, A., GLAVINA DEL RIO, T., TICE, H., *et al*. (2010). Complete genome sequence of *Methanothermus fervidus* type strain ($V24S^T$). *Stand Genomic Sci* 3: 127–135. doi.org/10.4056/sigs.1283367

ANG, M.Y., DUTTA, A., WEE, W.Y., DYMOCK, D., PATERSON, I.C. and CHOO, S.W. (2016). Comparative genome analysis of *Fusobacterium nucleatum*. *Genome Biol Evol* 8: 2928–2938. doi.org/10.1093/gbe/evw199

ANTÓN, J., OREN, A., BENLLOCH, S., RODRÍGUEZ-VALERA, F., AMANN, R. and ROSSELLÓ-MÓRA, R., (2002). *Salinibacter ruber* gen. nov., sp. nov., a novel, extremely halophilic member of the *Bacteria* from saltern crystallizer ponds. *Int J Syst Evol Microbiol* 52: 485–491. doi.org/10.1099/00207713-52-2-485

ANTÓN, J., PENÄ, A., SANTOS, F., MARTÍNEZ-GARCÍA, M, SCHMITT-KOPPLIN, P. and ROSSELLÓ-MÓRA, R. (2008). Distribution, abundance and diversity of the extremely halophilic bacterium *Salinibacter ruber*. *Saline Syst* 4: 15. doi.org/10.1186/1746-1448-4-15

ANTONUCCI, I., GALLO, G., LIMAURO, D., CONTURSI, P., BLES, A., BERENGUER, J., BARTOLUCCI, S. and FIORENTINO, G. (2018). Characterization of a promiscuous cadmium and arsenic resistance mechanism in *Thermus thermophilus* HB27 and potential application of a novel bioreporter system. *Microb Cell Fact* 17: 78. doi.org/10.1186/s12934-018-0918-7

ANTRANIKIAN, A., VORGIAS, C.E. and BERTOLDO, C. (2005). Extreme environments as a resource for microorganisms and novel biocatalysts. *Adv Biochem Eng Biotechnol* 96: 219–262. doi.org/10.1007/b135786

APTEKMANN, A.A. and NADRA, A.D. (2018). Core promoter information content correlates with optimal growth temperature. *Sci Rep* 8: 1313. doi.org/10.1038/s41598-018-19495-8

APWEILER, R., BAIROCH, A., WU, C.H., BARKER, W.C., BOECKMANN, B., FERRO, S., GASTEIGER, E., HUANG, H., LOPEZ, R., MAGRANE, M., *et al*. (2004). UniProt: the universal protein knowledgebase. *Nucleic Acids Res* 45: D115–D119. doi.org/10.1093/nar/gkh131

ARAHAL, D.R. (2014). Whole-genome analyses: average nucleotide identity. In *Methods in Microbiology (New Approaches to Prokaryotic Systematics)*. Goodfellow, M., Sutcliffe, I. and Chun, J. (editors). Elsevier Ltd. Vol. 41, pp. 103–122. doi.org/10.1016/bs.mim.2014.07.002

ARUN, A.B., CHEN, W.M., LAI, W.A., CHOU, J.H., SHEN, F.T., REKHA, P.D. and YOUNG, C.C. (2009). *Lutaonella thermophila* gen., nov., sp. nov., a moderately thermophilic member of the family *Flavobacteriaceae* isolated from a coastal hot spring. *Int J Syst Evol Microbiol* 59: 2069–2073. doi.org/10.1099/ijs.0.005256-0

ATHEY, J., ALEXAKI, A., OSIPOVA, E., ROSTOVTSEV, A., SANTANA-QUINTERO, L.V., KATNENI, U., SIMONYAN, V. and KIMCHI-SARFATY, C. (2017). A new and updated resource for codon usage tables. *BMC Bioinformatics* 18: 391. doi.org/10.1186/s12859-017-1793-7

AUCH, A.F., VON JAN, M., KLENK, H.P. and GÖKER, M. (2010). Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2: 117–134. doi.org/10.4056/sigs.531120

AULITTO, M., FUSCO, S., FIORENTINO, G., LIMAURO, D., PEDONE, E., BARTOLUCCI, S. and CONTURSI, P. (2017). *Thermus thermophilus* as source of thermozymes for

biotechnological applications: homologous expression and biochemical characterization of an α-galactosidase. *Microb Cell Fact* 16: 28. doi.org/10.1186/s12934-017-0638-4

BAKER-AUSTIN, C. and DOPSON, M. (2007). Life in acid: pH homeostasis in acidophiles. *Trends Microbiol* 15: fny239. doi.org/10.1016/j.tim.2007.02.005

BANKEVICH, A., NURK, S., ANTIPOV, D., GUREVICH, A.A., DVORKIN, M., KULIKOV, A.S., LESIN, V.M., NIKOLENKO, S.I., PHAM, S., PRJIBELSKI, A.D., *et al*. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19: 455–477. doi.org/10.1089/cmb.2012.0021

BARNARD, D., CASANUEVA, A., TUFFIN, M. and COWAN, D. (2010). Extremophiles in biofuel synthesis. *Environ Technol* 31: 871–888. doi.org/10.1080/09593331003710236

BAR-ON, Y.M., PHILLIPS, R. and MILO, R. (2018). The biomass distribution on earth. *PNAS* 115: 6506–6511. doi.org/10.1073/pnas.1711842115

BARTUCCI, R., GAMBACORTA, A., GLIOZZI, A., MARSH, D. and SPORTELLI, L. (2005). Bipolar tetraether lipids: chain flexibility and membrane polarity gradients from spin-label electron spin resonance. *Biochemistry* 44: 15017–15023. doi.org/10.1021/bi051101i

BEAZ-HIDALGO, R., HOSSAIN, M., LILES, M.R. and FIGUERAS, M.J. (2015). Strategies to avoid wrongly labelled genomes using as example the detected wrong taxonomic affiliation for *Aeromonas* genomes in the GenBank database. *PLoS One* 10: e0115813. doi.org/10.1371/journal.pone.0115813

BECERRA, A., DELAYE, L., LAZCANO, A. and ORGEL, L.E. (2007). Protein disulfide oxidoreductases and the evolution of thermophily: was the last common ancestor a heat-loving microbe? *J Mol Evol* 65: 296–303. doi.org/10.1007/s00239-007-9005-0

BEELER, E. and SINGH, O.V. (2016). Extremophiles as sources of inorganic bio-nanoparticles. *World J Microbiol Biotechnol* 32: 156. doi.org/10.1007/s11274-016-2111-7

BERGQUIST, P.L., MORGAN, H.W. and SAUL, D. (2014). Selected enzymes from extreme thermophiles with applications in biotechnology. *Curr Biotechnol* 3: 45–59. doi.org/10.1016/j.biotechadv.2015.04.007

BERNARDET, J.F., NAKAGAWA, Y. and HOLMES, B. (2002). Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended description of the family. *Int J Syst Evol Bacteriol* 52: 1049–1070. doi.org/10.1099/ijs.0.02136-0

BESEMER, J., LOMSADZE, A. and BORODOVSKY, M. (2001). GeneMarkS: a self-training method for prediction of gene starts in microbial genomes: implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res* 20: 2607–2618. doi.org/10.1093/nar/29.12.2607

BEYE, M., FAHSI, N., RAOULT, D. and FOURNIER, P.E. (2018). Careful use of 16S rRNA gene sequence similarity values for the identification of *Mycobacterium* species. *New Microbes New Infect* 22: 24–29. doi.org/10.1016/j.nmni.2017.12.009

BHALLA, A., BANSAL, N., KUMAR, S., BISCHOFF, K.M. and SANI, R.K. (2013). Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. *Bioresour Technol* 128: 751–759. doi.org/10.1016/j.biortech.2012.10.145

BHANDARI, V., AHMOD, N.Z., SHAH, H.N. and GUPTA, R.S. (2013). Molecular signatures for *Bacillus* species: demarcation of the *Bacillus subtilis* and *Bacillus cereus* clades in molecular terms and proposal to limit the placement of new species into the genus *Bacillus*. *Int J Syst Evol Microbiol* 63: 2712–2726. doi.org/10.1099/ijs.0.048488-0

BHARALI, P., DAS, S., KONWAR, B.K. and THAKUR, A.J. (2011). Crude biosurfactant from thermophilic *Alcaligenes faecalis*: feasibility in petro-spill bioremediation. *Int Biodeter Biodegr* 65: 682–690. doi.org/10.1016/j.ibiod.2011.04.001

BISGAARD, M., CHRISTENSEN, H., CLERMONT, D., DIJKSHOORN, L., JANDA, J.M., MOORE, E.R.B., NEMEC, A., NØRSKOV-LAURITSEN, N., OVERMANN, J. and REUBSAET, F.A.G. (2019). The use of genomic DNA sequences as type material for valid publication of bacterial species names will have severe implications for clinical microbiology and related disciplines. *Diagn Microbiol Infect Dis* 95: 102–103. doi.org/10.1016/j.diagmicrobio.2019.03.007

BLÖCHL, E., RACHEL, R., BURGGRAF, S., HAFENBRADL, D., JANNASCH, H.W. and STETTER, K.O. (1997). *Pyrolobus fumarii*, gen. and sp. nov., represents a novel group of archaea, extending the upper temperature limit for life to 113°C. *Extremophiles* 1: 14–21. doi.org/10.1007/s007920050010

BOLGER, A.M., LOHSE, M. and USADEL, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30: 2114–2120. doi.org/10.1093/bioinformatics/btu170

BOOS, W. and SHUMAN, H. (1998). Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism and regulation. *Microbiol Mol Biol Rev* 62: 204–229. doi.org/10.1128/MMBR.62.1.204-229.1998

BORGES, N., JORGE, C.D., GONÇALVES, L.G., GONÇALVES, S., MATIAS, P.M. and SANTOS, H. (2014). Mannosylglycerate: structural analysis of biosynthesis and evolutionary history. *Extremophiles* 18: 835–852. doi.org/10.1007/s00792-014-0661-x

BORGES, N., RAMOS, A., RAVEN, N.D.H., SHARP, R.J. and SANTOS, H. (2002). Comparative study of the thermostabilizing properties of mannosylglycerate and other compatible solutes on model enzymes. *Extremophiles* 6: 209–216. doi.org/10.1007/s007920100236

BORRISS, R., RUECKER, C., BLOM, J., BEZUIDT, O., REVA, O. and KLENK, H.P. (2011). Whole genome sequence comparisons in taxonomy. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 409–436. doi.org/10.1016/B978-0-12-387730-7.00018-8

BOUSSAU, B., BLANQUART, S., NECSULEA, A., LARTILLOT, N. and GOUY, M. (2008). Parallel adaptations to high temperatures in the archaean eon. *Nature* 456: 942–945. doi.org/10.1038/nature07393

BOUTZ, D.R., CASCIO, D., WHITELEGGE, J., PERRY, L.J. and YEATES, T.O. (2007). Discovery of a thermophilic protein complex stabilized by topologically interlinked chains. *J Mol Biol* 368: 1332–1344. doi.org/10.1016/j.jmb.2007.02.078

BRBIĆ, M., PIŠKOREC, M., VIDULIN, V., KRIŠKO, A., ŠMUC, T. and SUPEK, F. (2016). The landscape of microbial phenotypic traits and associated genes. *Nucleic Acids Res* 44: 10074–10090. doi.org/10.1093/nar/gkw964

BROCK, T.D. (1967a). Life at high temperatures. *Science* 158: 1012–1019. doi.org/10.1126/science.158.3804.1012

BROCK, T.D. (1967b). Micro-organisms adapted to high temperatures. *Nature* 214: 882–885. doi.org/10.1038/214882a0

BROCK, T.D. (1978). *Thermophilic Microorganisms and Life at High Temperatures*. Springer-Verlag, New York. pp. 465.

BROCK, T.D. and BROCK, M.L. (1966). Temperature optima for alga development in Yellowstone and Iceland hot springs. *Nature* 209: 733–734. doi.org/10.1038/209733a0

BROCK, T.D. and FREEZE, H. (1969). *Thermus aquaticus* gen. n. and sp. n., a non-sporulating extreme thermophile. *J Bacteriol* 98: 289–297. doi.org/10.1128/JB.98.1.289-297.1969

BROCK, T.D., BROCK, K.M., BELLY, R.T. and WEISS, R.L. (1972). *Sulfolobus*: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature. *Arch Mikrobiol* 84: 54–68. doi.org/10.1007/BF00408082

BROCKER, M., SCHAFFER, S., MACK, C. and BOTT, M. (2009). Citrate utilization by *Corynebacterium glutamicum* is controlled by the CitAB two-component system through positive regulation of the citrate transport genes *citH* and *tctCBA*. *J Bacteriol* 191: 3869–3880. doi.org/10.1128/JB.00113-09

BROEKER, J., MECHELKE, M., BAUDREXL, M., MENNERICH, D., HORNBURG, D., MANN, M., SCHWARZ, W., LIEBL, W. and ZVERLOV, V.V. (2018). The hemicellulose-degrading enzyme system of the thermophilic bacterium *Clostridium stercorarium*: comparative characterisation and addition of new hemicellulolytic glycoside hydrolases. *Biotechnol Biofuels* 11: 229. doi.org/10.1186/s13068-018-1228-3

BROOKS, D.J., FRESCO, J.R. and SINGH, M. (2004). A novel method for estimating ancestral amino acid composition and its application to proteins of the Last Universal Ancestor. *Bioinformatics* 20: 2251–2257. doi.org/10.1093/bioinformatics/bth235

BROWNE, H.P., FORSTER, S.C., ANONYE, B.O., KUMAR, N., NEVILLE, B.A., STARES, M.D., GOULDING, D. and LAWLEY, T.D. (2016). Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* 533: 543–546. doi.org/10.1038/nature17645

BRÜGGER, K., CHEN, L., STARK, M., ZIBAT, A., REDDER, P., RUEPP, A., AWAYEZ, M., SHE, Q., GARRET, A. and KLENK, H.P. (2007). The genome of *Hyperthermus butylicus*: a sulfur-reducing, peptide fermenting, neutrophilic crenarchaeote growing up to 108°C. *Archaea* 2: 127–135. doi.org/10.1155/2007/745987

BUCHANAN, R.E. and GIBBONS, N.E. (1974). *Bergey's Manual of Determinative Bacteriology*. 8th ed, The Williams and Wilkins Co., Baltimore, USA. doi.org/10.1111/j.1550-7408.1975.tb00935.x

BURGESS, M.L., BARROW, K.D., GAO, C., HEARD, G.M. and GLENN, D. (1999). Carotenoid glycoside esters from the thermophilic bacterium *Meiothermus ruber*. *J Nat Prod* 62: 859–863. doi.org/10.1021/np980573d

BUSK, P.K. and LANGE, L. (2013). Cellulolytic potential of thermophilic species from four fungal orders. *AMB Expr* 3: 47. doi.org/10.1186/2191-0855-3-47

BUSSE, F.A. (2011). Polyamines. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 239–259. doi.org/10.1016/B978-0-12-387730-7.00011-5

BUSSE, J. and AULING, G. (1988). Polyamine pattern as a chemotaxonomic marker within the *Proteobacteria*. *Syst Appl Microbiol* 11: 1–8. doi.org/10.1016/S0723-2020(88)80040-7

BÖTTGER, E.C. (1989). Rapid determination of bacterial ribosomal RNA sequences by direct sequencing of enzymatically amplified DNA. *FEMS Microbiol Lett* 65: 171–176. doi.org/10.1016/0378-1097(89)90386-8

CABAL, A., JUN, S.R., JENJAROENPUN, P., WANCHAI, V., NOOKAEW, I., WONGSURAWAT, T., BURGESS, M.J., KOTHARI, A., WASSENAAR, T.M. and USSERY, D.W. (2018). Genome-based comparison of *Clostridioides difficile*: average amino acid identity analysis of core genomes. *Microb Ecol* 76: 801–813. doi.org/10.1007/s00248-018-1155-7

CACCIAPUOTI, G., FUCCIO, F., PETRACCONE, L., DEL VECCHIO, P. and PORCELLI, M. (2012). Role of disulfide bonds in conformational stability and folding of 5′-deoxy-5′-methylthioadenosine phosphorylase II from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biochim Biophys Acta* 1824: 1136–1143. doi.org/10.1016/j.bbapap.2012.06.014

CAIERÃO, J., PAIVA, J.A.C.D., SAMPAIO, J.L.M., DA SILVA, M.G., SANTOS, D.R.S., COELHO, F.S., FONSECA, L.S., DUARTE, R.S., ARMSTRONG, D.T. and REGUA-MANGUIA, A.H. (2016). Multilocus enzyme electrophoresis analysis of rapidly-growing mycobacteria: an alternative tool for identification and typing. *Int J Infect Dis* 42: 11–16. doi.org/10.1016/j.ijid.2015.11.010

CALIRO, S., CHIODINI, G., MORETTI, R., AVINO, R., GRANIERI, D., RUSSO, M. and FIEBIG, F. (2007). The origin of the fumaroles of La Solfatara (Campi Flegrei, South Italy). *Geo Cosmoch Acta* 71: 3040–3055. doi.org/10.1016/j.gca.2007.04.007

CANGANELLA, F. and WIEGEL, J. (2011). Extremophiles: from abyssal to terrestrial ecosystems and possibly beyond. *Naturwissenschaften* 98: 253–279. doi.org/10.1007/s00114-011-0775-2

CANGANELLA, F. and WIEGEL, J. (2014). Anaerobic thermophiles. *Life 4*: 77–104. doi.org/10.3390/life4010077

CANTISTA, A.P.P. (2008). O termalismo em Portugal. *An Hidrol Medica 3*: 79–107. ISSN:1887-0813

CAPECE, M.C., CLARK, E., SALEH, J.K., HALFORD, D., HEINL, N., HOSKINS, S. and ROTHSCHILD, L.J. (2013). Polyextremophiles and the constraints for terrestrial habitability. In *Polyextremophiles. Cellular Origin, Life in Extreme Habitats and Astrobiology*. Seckbach,

J., Oren, A. and Stan-Lotter, H. (editors). Springer. Vol. 27, pp.3–59. doi.org/10.1007/978-94-007-6488-0_1

CAPUTO, A., FOURNIER, P.E. and RAOULT, D. (2019). Genome and pan-genome analysis to classify emerging bacteria. *Biol Direct* 14: 5. doi.org/10.1186/s13062-019-0234-0

CAPUTO, A., MERHEJ, V., GEORGIADES, K., FOURNIER, P.E., CROCE, O., ROBERT, C. and RAOULT, D. (2015). Pan-genomic analysis to redefine species and subspecies based on quantum discontinuous variation: the *Klebsiella* paradigm. *Biol Direct* 10: 55. doi.org/10.1186/s13062-015-0085-2

CARRETO, L., WAIT, R., NOBRE, M.F. and DA COSTA, M.S. (1996). Determination of the structure of a novel glycolipid from *Thermus aquaticus* 15004 and demonstration that hydroxy fatty acids are amide linked to glycolipids in *Thermus* spp. *J Bacteriol* 178: 6479–6486. doi.org/10.1128/jb.178.22.6479-6486.1996

CASTELLE, C.J. and BANFIELD, J.F. (2018). Major new microbial groups expand diversity and alter our understanding of the Tree of Life. *Cell* 172: 1181–1197. doi.org/10.1016/j.cell.2018.02.016

CAUMETTE, P., BROCHIER-ARMANET, C. and NORMAND, P. (2015). Taxonomy and phylogeny of prokaryotes. In *Environmental Microbiology: Fundamentals and Applications*. Bertrandt, J.C., Caumette, P., Lebaron, P., Matheron, R., Normand, P. and Sime-Ngando, T. (editors). Springer, Dordrech. pp. 145–190. doi.org/10.1007/978-94-017-9118-2_6

CAVA, F., HIDALGO, A., and BERENGUER, J. (2009). *Thermus thermophilus* as biological model. *Extremophiles* 13: 213–231. doi.org/10.1007/s00792-009-0226-6

CAVALIER-SMITH, T. and CHAO, E.E.Y. (2020). Multidomain ribosomal protein trees and the planctobacterial origin of neomura (eukaryotes, archaebacteria). *Propoplasma* 257: 621–753. doi.org/10.1007/s00709-019-01442-7

CERQUEIRA, T., BARROSO, C., FROUFE, H., EGAS, C. and BETTENCOURT, R. (2018). Metagenomic signatures of microbial communities in deep-sea hydrothermal sediments of Azores vent fields. *Microb Ecol* 76: 387–403. doi.org/10.1007/s00248-018-1144-x

CERQUEIRA, T., PINHO, D., FROUFE, H., SANTOS, R.S., BETTENCOURT, R. and EGAS, C. (2017). Sediment microbial diversity of three deep-sea hydrothermal vents southwest of the Azores. *Microb Ecol* 74: 332–349. doi.org/10.1007/s00248-017-0943-9

CHAI, J., KORA, G., AHN, T.H., HYATT, D. and PAN, C. (2014). Functional phylogenomics analysis of bacteria and archaea using consistent genome annotation with UniFam. *BMC Evol Biol* 14: 207. doi.org/10.1186/s12862-014-0207-y

CHEN, M.Y., LIN, G.H., LIN, Y.T. and TSAY, S.S. (2002). *Meiothermus taiwanensis* sp. nov., a novel filamentous, thermophilic species isolated in Taiwan. *Int J Syst Evol Microbiol* 52: 1647–1654. doi.org/10.1099/00207713-52-5-1647

CHEN, R.W., LI, C., HE, Y.Q., CUI, L.Q., LONG, L.J. and TIAN, X.P. (2020). *Rubrobacter tropicus* sp. nov. and *Rubrobacter marinus* sp. nov., isolated from deep-sea sediment of the South China Sea. *Int J Syst Evol Microbiol* 70: 5576–5585. doi.org/10.1099/ijsem.0.004449

CHEN, T.L., CHOU, Y.J., CHEN, W.M., ARUN, B. and YOUNG, C.C. (2006). *Tepidimonas taiwanensis* sp. nov., a novel alkaline-protease-producing bacterium isolated from a hot spring. *Extremophiles* 10: 35–40. doi.org/10.1007/s00792-005-0469-9

CHEN, W.M., HUANG, H.W., CHANG J.S., HAN, Y.L., GUO, T.R. and SHEU, S.Y. (2013). *Tepidimonas fonticaldi* sp. nov., a slightly thermophilic betaproteobacterium isolated from a hot spring. *Int J Sys Evol Microbiol* 63: 1810–1816. doi.org/10.1099/ijs.0.043729-0

CHIEN, A., EDGAR, D.B. and TRELA, J.M. (1976). Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J Bacteriol* 127: 1550–1557.

CHUN, J. and RAINEY, F.A. (2014). Integrating genomics into the taxonomy and systematics of the *Bacteria* and *Archaea*. *Int J Syst Evol Microbiol* 64: 316–324. doi.org/10.1099/ijs.0.054171-0

CHUN, J., LEE, J.H., JUNG, Y., KIM, M., KIM, S., KIM, B.K. and LIM, Y.W. (2007). EzTaxon: a web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences. *Int J Syst Evol Microbiol* 57: 2259–2261. doi.org/10.1099/ijs.0.64915-0

CHUN, J., OREN, A., VENTOSA, A., CHRISTENSEN, H., ARAHAL, D.R., DA COSTA, M.S., ROONEY, A.P., YI, H., XU, X.W., DE MEYER, S., *et al*. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 68: 461–466. doi.org/10.1099/ijsem.0.002516

CHUNG, A.P., RAINEY, F., NOBRE, M.F., BURGHARDT, J. and DA COSTA, M.S. (1997). *Meiothermus cerbereus* sp. nov., a new slightly thermophilic species with high levels of 3-hydroxy fatty acids. *Int J Syst Bacteriol* 47: 1225–1230. doi.org/doi: 10.1099/00207713-47-4-1225

CHUNG, M., MUNRO, J.B., TETTELIN, H. and HOTOPP, J.C.D. (2018). Using core genome alignments to assign bacterial species. *mSystems* 3: e00236-18. doi.org/10.1128/mSystems.00236-18

CHUVOCHINA, M., RINKE, C., PARKS, D.H., RAPPÉ, M.S., TYSON, G.W., YILMAZ, P., WHITMAN, W.B. and HUGENHOLTZ, P. (2019). The importance of designating type material for uncultured taxa. *Syst Appl Microbiol* 42: 15–21. doi.org/10.1016/j.syapm.2018.07.003

CICCARELLI, F.D., DOERKD, T., MERING, C.V., CREEVEY, C.J., SNEL, B. and BORK, P. (2006). Toward automatic reconstruction of a highly resolved Tree of Life. *Science* 311: 1283–1287. doi.org/10.1126/science.1123061

CIUFO, S., KANNAN, S., SHARMA, S., BADRETDIN, A., CLARK, K., TURNER, S., BROVER, S., SCHOCH, C.L., KIMCHI, A. and DICUCCIO, M. (2018). Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int J Syst Evol Microbiol* 68: 2386–2392. doi.org/10.1099/ijsem.0.002809

CLARKE, A. (2014). The thermal limits to life on Earth. *Int J Astrobiol* 13: 141–154. doi.org/10.1017/S1473550413000438

COENYE, T., GEVERS, D., VAN DE PEER, Y., VANDAMME, P. and SWINGS, J. (2005). Towards a prokaryotic genomic taxonomy. *FEMS Microbiol Rev* 29: 147–167. doi.org/10.1016/j.femsre.2004.11.004

COKER, J.A. (2019). Recent advances in understanding extremophiles. *F1000Research* 8. doi.org/10.12688/f1000research.20765.1

COLE, J.R., WANG, Q., CARDENAS, E., FISH, J., CHAI, B., FARRIS, R.J., KULAM-SYED-MOHIDEEN, A.S., MCGARRELL, D.M., MARSH, T., GARRITY, G.M., *et al*. (2009). The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. *Nucleic Acids Res* 37: D141–D145. doi.org/10.1093/nar/gkn879

COLMAN, D.R., POUDEL, S., HAMILTON, T.L., HAVING, J.R., SELENSKY, M.J., SHOCK, E.L. and BOYD, E.S. (2018). Geobiological feedbacks and the evolution of thermoacidophiles. *ISME J* 12: 225–236. doi.org/10.1038/ismej.2017.162

COLSTON, S.M., FULLMER, M.S., BEKA, L, LAMY, B., GOGARTEN, J.P. and GRAF, J. (2014). Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case. *mBio* 18: e02136. doi.org/10.1128/mBio.02136-14

COLWELL, R.R. (1970). Polyphasic taxonomy of the genus *Vibrio*: numerical taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and related *Vibrio* species. *J Bacteriol* 104: 410–433. doi.org/10.1128/JB.104.1.410-433.1970

CONTRERAS-MOREIRA, B. and VINUESA, P. (2013). GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Appl Environ Microbiol* 79: 7696–7701. doi.org/10.1128/AEM.02411-13

CONWAY DE MACARIO, E. and MACARIO, A.J.L. (2000). Stressors, stress and survival: overview. *Front Biosci* 5: D780–786. doi.org/10.2741/a550

CORLISS, J.B., DYMOND, J., LOUIS, I.G., EDMOND, J.M., HERZEN, R.P.V., BALLARD, R.D., GREEN, K., WILLIAMS, D., BAINBRIDGE, A., CRANE, K., *et al.* (1979). Submarine thermal springs on the Galápagos Rift. *Science* 203: 1073–1083. doi.org/10.1126/science.203.4385.1073

CORRAL, P., DE LA HABA, R.R., INFANTE-DOMÍNGUEZ, C., SÁNCHEZ-PORRO, C., AMOOZEGAR, M.A., PAPKE, R.T. and VENTOSA, A. (2018). *Halorubrum chaoviator* Mancinelli *et al*. 2009 is a later, heterotypic synonym of *Halorubrum ezzemoulense* Kharroub *et al*. 2006. Emended description of *Halorubrum ezzemoulense* Kharroub *et al*. 2006. *Int J Syst Evol Microbiol* 68: 3657–3665. doi.org/10.1099/ijsem.0.003005

CORREA-LLANTÉN, D.N., MUÑOZ, S.A., CASTRO, M.E., MUÑOZ, P.A. and BLAMEY, J.M. (2013). Gold nanoparticles synthesized by *Geobacillus* sp. strain ID17 a thermophilic bacterium isolated from Deception Island, Antarctica. *Microbial Cell Fact* 12: 75. doi.org/10.1186/1475-2859-12-75

COWAN, S.T. (1965). Principles and practice of bacterial taxonomy – a forward look. *J Gen Microbiol* 39: 143–153. doi.org/10.1099/00221287-39-1-143

CROGNALE, S., VENTURI, S., TASSI, F., ROSSETTI, S., RASHED, H., CABASSI, J., CAPECCHIACCI, F., NISI, B., VASELLI, O., MORRISON, H.G., *et al.* (2018). Microbiome profiling in extremely acidic soils affected by hydrothermal fluids: the case of the Solfatara Crater (Campi Flegrei, southern Italy). *FEMS Microbiol Eco* 94: fiy190. doi.org/10.1093/femsec/fiy190

CRUZ, P.E., SILVA, A.C., ROLDÃO, A., CARMO, M., CARRONDO, M.J.T. and ALVES, P.M. (2006). Screening of novel excipients for improving the stability of retroviral and adenoviral vectors. *Biotechnol Prog* 22: 568–576. doi.org/10.1021/bp050294y

CUI, Y., JIANG, X., HAO, M., QU, X. and HU, T. (2017). New advances in exopolysaccharides production of *Streptococcus thermophilus*. *Arch Microbiol* 199: 799–809. doi.org/10.1007/s00203-017-1366-1

CURTIS, T.P., SLOAN, W.T. and SCANNELL, J.W. (2002). Estimating prokaryotic diversity and its limits. *PNAS 99*: 10494–10499. doi.org/10.1073/pnas.142680199

D´HUGUES, P., FOUCHER, S., GALLÉ-CAVALLONI, P. and MORIN, D. (2002). Continuous bioleaching of chalcopyrite using a novel extremely thermophilic mixed culture. *Int J Miner Process* 66: 107–119. doi.org/10.1016/S0301-7516(02)00004-2

DA COSTA, M.S., ALBUQUERQUE, L., NOBRE, M.F. and WAIT, R. (2011a). The identification of polar lipids in prokaryotes. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey F.A. and Oren A. (editors). London, Elsevier Ltd. Vol. 38, pp. 165–181. doi.org/10.1016/B978-0-12-387730-7.00007-3

DA COSTA, M.S., ALBUQUERQUE, L., NOBRE, M.F. and WAIT, R. (2011b). The extraction and identification of respiratory lipoquinones of prokaryotes and their use in taxonomy. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey F.A. and Oren A. (editors). London, Elsevier Ltd. Vol. 38, pp. 197–206. doi.org/10.1016/B978-0-12-387730-7.00009-7

DA COSTA, M.S., ALBUQUERQUE, L., NOBRE, M.F. and WAIT, R. (2011c). The identification of fatty acids in bacteria. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey F.A. and Oren A. (editors). London, Elsevier Ltd. Vol. 38, pp. 183–196. doi.org/10.1016/B978-0-12-387730-7.00008-5

DA COSTA, M.S., ALBUQUERQUE, L., RAINEY, F.A., FROUFE, H.J.C., ROXO, I., RAPOSO, P. and EGAS, C. (2019). *Elioraea*. In *Bergey's Manual of Systematics of Archaea and Bacteria,* Trujillo, M.E., Dedysh, S., De Vos, P., Hedlund, B., Kämpfer, P., Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.gbm01458

DAMSTÉ, J.S.S., RIJPSTRA, W.I.C., HOPMANS, E.C., SCHOUTEN, S., BALK, M. and STAMS, A.J.M. (2007). Structural characterization of diabolic acid-based tetraester, tetraether and mixed ether/ester, membrane-spanning lipids of bacteria from the order *Thermotogales*. *Arch Microbiol 188*: 629–641. doi.org/10.1007/s00203-007-0284-z

DASSARMA, P. and DASSARMA, S. (2018). Survival of microbes in Earth's stratosphere. *Curr Opin Microbiol* 43: 24–30. doi.org/10.1016/j.mib.2017.11.002

DASSARMA, P., LAYE, V.J., HARVEY, J., REID, C., SHULTZ, J., YARBOROUGH, A., LAMB, A., KOSKE-PHILLIPS, A., HERBST, A., MOLINA, F., *et al*. (2017). Survival of halophilic *Archaea* in Earth's cold stratosphere. *Int J Astrobiol* 16: 321–327. doi.org/10.1017/S1473550416000410

DE LA HABA, R.R., LÓPEZ-HERMOSO, C., SÁNCHEZ-PORRO, C., KONSTANTINIDIS, K.T. and VENTOSA, A. (2019). Comparative genomics and phylogenomic analysis of the genus *Salinivibrio*. *Front Microbiol* 10: 2104. doi.org/10.3389/fmicb.2019.02104

DE LEY, J. (1970). Reexamination of the association between melting point, buoyant density, and chemical base composition of deoxyribonucleic acid. *J Bacteriol* 101: 738–754. doi.org/10.1128/JB.101.3.738-754.1970

DE ROSA, M., MORANA, A., RICCIO, A., GAMBACORTA, A., TRINCONE, A. and INCANI, O. (1994). Lipids of the *Archaea*: a new tool for bioelectronics. *Biosens Bioelectron* 9: 669–675. doi.org/10.1016/0956-5663(94)80064-2

DEIVE, F.J., DOMÍNGUEZ, A., BARRIO, T., MOSCOVO, F., MORÁN, P., LONGO, M.A. and SANROMÁN, M.A. (2010). Decolorization of dye reactive black 5 by newly isolated thermophilic microorganisms from geothermal sites in Galicia (Spain). *J Hazard Mater* 182: 735–742. doi.org/10.1016/j.jhazmat.2010.06.096

DEL GIUDICE, I., LIMAURO, D., PEDONE, E., BARTOLUCCI, S. and FIORENTINO, G. (2013). A novel arsenate reductase from the bacterium *Thermus thermophiles* HB27: its role in arsenic detoxification. *Biochim Biophys Acta* 1834: 2071–2079. doi.org/10.1016/j.bbapap.2013.06.007

DELOGER, M., KAROUI, M.E. and PETIT, M.A. (2009). A genomic distance based on MUM indicates discontinuity between most bacterial species and genera. *J Bacteriol* 191: 91–99. doi.org/10.1128/JB.01202-08

DENGER, K., WARTHMANN, R., LUDWIG, W. and SCHINK, B. (2002). *Anaerophaga thermohalophila* gen., nov., sp. nov., a moderately thermohalophilic, strictly anaerobic fermentative bacterium. *Int J Syst Evol Microbiol* 52:173–178. doi.org/10.1099/00207713-52-1-173

DESANTIS, T.Z., KUGENHOLTZ, P., LARSEN, N., ROJAS, M., BRODIE, E.L., KELLER, K., HUBER, T., DALEVI, D., HU, P. and ANDERSEN, G.L. (2006). Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72: 5069–5072. doi.org/10.1128/AEM.03006-05

DHAKAN, D.B., SAXENA, R., CHAUDHARY, N. and SHARMA, V.K. (2016). Draft genome sequence of *Tepidimonas taiwanensis* strain MB2, a chemolithotrophic thermophile isolated from a hot spring in central India. *Genome Announc* 4: 1–2. doi.org/10.1128/genomeA.01723-15

DI GIULIO, M.D. (2000). The universal ancestor lived in a thermophilic or hyperthermophilic environment. *J Theor Biol* 203: 203–2013. doi.org/10.1006/jtbi.2000.1086

DIAKITE, A., DUBOURG, G., DIONE, N., AFOUDA, P., BELLALI, S., NGOM, I.I., VALLES, C., MILLION, M., LEVASSEUR, A., CADORET, F., *et al*. (2019). Extensive culturomics of 8 healthy samples enhances metagenomics efficiency. *PLoS One* 14: e0223543. doi.org/10.1371/journal.pone.0223543

DIAKITE, A., DUBOURG, G., DIONE, N., AFOUDA, P., BELLALI, S., NGOM, I.I., VALLES, C., TALL, M.I., LAGIER, J.C. and RAOLT, D. (2020). Optimization and standardization of the culturomics technique for human microbiome exploration *Sci Rep* 10: 9674. doi.org/10.1038/s41598-020-66738-8

DILLY, G.F., YOUNG, C.R., LANE, W.S., PANGILINAN, J. and GIRGUIS, P.R. (2012). Exploring the limit of metazoan thermal tolerance via comparative proteomics: thermally induced changes in protein abundance by two hydrothermal vent polychaetes. *Proc R Soc B* 279: 3347–3356. doi.org/10.1098/rspb.2012.0098

DING, W., BAUMDICKER, F. and NEHER, R.A. (2018). panX: pan-genome analysis and exploration. *Nucleic Acids Res* 46: e5. doi.org/10.1093/nar/gkx977

DIOGO, A., VERÍSSIMO, A., NOBRE, M.F. and DA COSTA, M.S. (1999). Usefulness of fatty acid composition for differentiation of *Legionella* species. *J Clin Microbiol* 37: 2248–2254. doi.org/10.1128/JCM.37.7.2248-2254.1999

DIOP, A., KARKOURI, K.E., RAOULT, D. and FOURNIER, P.E. (2020). Genome sequence-based criteria for demarcation and definition of species in the genus *Rickettsia*. *Int J Syst Evol Microbiol* 70: 1738–1750. doi.org/10.1099/ijsem.0.003963

DONATI, E.R., CASTRO, C. and URBIETA, M.S. (2016). Thermophilic microorganisms in biomining. *World J Microbiol Biotechnol* 32: 179–1750. doi.org/10.1007/s11274-016-2140-2

DREWS, G. (2000). The roots of microbiology and the influence of Ferdinand Cohn on microbiology of the 19th century. *FEMS Microbiol Rev* 24: 225–249. doi.org/10.1111/j.1574-6976.2000.tb00540.x

DRIDI, B.B. and DRANCOURT, M. (2011). Characterization of prokaryotes using MALDI-TOF mass spectrometry. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey F.A. and Oren A. (editors). London, Elsevier Ltd. Vol. 38, pp. 283–297. doi.org/10.1016/B978-0-12-387730-7.00013-9

DUMLER, J.S., BARBET, A.F., BEKKER, C.P.J., DASCH, G.A., PALMER, G.H., RAY, S.C., RIKIHISA, Y. and RURANGIRWA, F.R. (2001). Reorganization of genera in the families *Rickettsiaceae* and *Anaplasmataceae* in the order *Rickettsiales*: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia* equi and 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. *Int J Syst Evol Microbiol* 51: 2145–2165. doi.org/10.1099/00207713-51-6-2145

DUMORNÉ, K., CÓRDOVA, D.C., ASTORGA-ELÓ, M. and RENGANATHAN, P. (2017). Extremozymes: a potential source for industrial applications. *J Microbiol Biotechnol* 27: 649–659. doi.org/10.4014/jmb.1611.11006

DYKHUIZEN, D.E. (1998). Santa Rosalia revisited: why are there so many species of bacteria? *Antonie van Leeuwenhoek* 73: 25–33. doi.org/10.1023/a:1000665216662

ECE, S., EVRAN, S., JANDA, J.O., MERKL, R. and STERNER, R. (2015). Improving thermal and detergent stability of *Bacillus stearothermophilus* neopullulanase by rational enzyme design. *Protein Eng Des Sel* 28: 147–151. doi.org/10.1093/protein/gzv001

EDGAR, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32: 1792–1797. doi.org/10.1093/nar/gkh340

EDGAR, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460–2461. doi.org/10.1093/bioinformatics/btq461

EGAS, C., BARROSO, C., FROUFE, H.J., PACHECO, J., ALBUQUERQUE, L. and DA COSTA M.S. (2014). Complete genome sequence of the radiation-resistant bacterium *Rubrobacter radiotolerans* RSPS-4. *Stand Genomic Sci* 9: 1062–1075. doi.org/10.4056/sigs.5661021

EGAS, M.C.V., DA COSTA, M.S., COWAN, D.A. and PIRES, E.M.V. (1998). Extracellular α-amylase from *Thermus filiformis* Ork A2: purification and biochemical characterization. *Extremophiles* 2: 23–32. doi.org/10.1007/s007920050039

EGOROVA, K. and ANTRANIKIAN, A. (2005). Industrial relevance of thermophilic Archaea. *Curr Opin Microbiol* 8: 649–655. doi.org/10.1016/j.mib.2005.10.015

ELLEUCHE, S. and ANTRANIKIAN, G. (2013). Starch-hydrolyzing enzymes from thermophiles. In *Thermophilic Microbes in Environmental and Industrial Biotechnology*. Satyanarayana, T., Littlechild, J. and Kawarabayasi, Y. (editors). Springer, Dordrecht. pp. 509–533. doi.org/10.1007/978-94-007-5899-5_20

ELLEUCHE, S., SCHÄFERS, C., BLANK, S., SCHRÖDER, C. and ANTRANIKIAN, G. (2015). Exploration of extremophiles for high temperature biotechnological processes. *Curr Opin Microbiol* 23: 113–119. doi.org/10.1016/j.mib.2015.05.011

ELLEUCHE, S., SCHRÖDER, C., SALM, K. and ANTRANIKIAN, G. (2014). Extremozymes – biocatalysts with unique properties from extremophilic microorganisms. *Curr Opin Biotechnol* 29: 116–123. doi.org/10.1016/j.copbio.2014.04.003

EMPADINHAS, N. and DA COSTA, M.S. (2011). Diversity, biological roles and biosynthetic pathways for sugar-glycerate containing compatible solutes in bacteria and archaea. *Environ Microbiol* 13: 2056–2077. doi.org/10.1111/j.1462-2920.2010.02390.x

FARIA, C., JORGE, C.D., BORGES, N., TENREIRO, S., OUTEIRO, T.F. and SANTOS, H. (2013). Inhibition of formation of α-synuclein inclusions by mannosylglycerate in a yeast model of Parkinson's disease. *Biochim Biophys Acta* 1830: 4065–4072. doi.org/10.1016/j.bbagen.2013.04.015

FAYAZ, A.M., BALAJI, K., GIRILAL, M., MTECH, R.Y., KALAICHELVAN, P.T. and VENKETESAN, R. (2010a). Biogenic synthesis of silver nanoparticles and their synergistic effect with antibiotics: a study against gram-positive and gram-negative bacteria. *Nanomedicine*: *NBM* 6: 103–109. doi.org/10.1016/j.nano.2009.04.006

FAYAZ, A.M., GIRILAL, M., RAHMAN, M., VENKETESAN, R. and KALAICHELVAN, P.T. (2010b). Biosynthesis of silver and gold nanoparticles using thermophilic bacterium *Geobacillus stearothermophilus*. *Process Biochem* 46: 1958–1962. doi.org/10.1016/j.procbio.2011.07.003

FENG, G.D., ZHANG, J., ZHANG, X.J., WANG, S.N., XIONG, X., ZHANG, Y.L., HUANG, H.R. and ZHU, H.H. (2019). *Hymenobacter metallilatus* sp. nov., isolated from abandoned lead–zinc ore. *Int J Syst Evol Microbiol* 69: 2142–2146. doi.org/10.1099/ijsem.0.003450

FERNANDES, P. (2010). Enzymes in food processing: a condensed overview on strategies for better biocatalysts. *Enzyme Res* 2010: 862537. doi.org/10.4061/2010/862537

FERREIRA, A.M., WAIT, R., NOBRE, M.F. and DA COSTA, M.S. (1999). Characterization of glycolipids from *Meiothermus* spp. *Microbiology* 145: 1191–1199. doi.org/10.1099/13500872-145-5-1191

FIELD, D., GARRITY, G., GRAY, T., MORRISON, N., SELENGUT, J., STERK, P., TATUSOVA, T., THOMSON, N., ALLEN, M.J., ANGIUOLI, S.V., *et al*. (2008). The minimum information about a genome sequence (MIGS) specification. *Nat Biotechnol* 26: 541–547. doi.org/10.1038/nbt1360

FINN, R.D., COGGILL, P., EBERHARDT, R.Y., EDDY, S.R., MISTRY, J., MITCHELL, A.L., POTTER, S.C., PUNTA, M., QURESHI, M., SANGRADOR-VEGAS, A., *et al*. (2016). The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Research* 44: D279–D285. doi.org/10.1093/nar/gkv1344

FLEISCHMANN, R.D., ADAMS, M.D., WHITE, O., CLAYTON, R.A., KIRKNESS, E.F., KERLAVAGE, A.R., BULT, C.J., TOMB, J.F., DOUGHERTY, B.A., MERRICK, J.M., *et al*. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenza*. *Science* 269: 496–512. doi.org/10.1126/science.7542800

FORTERRE, P. (1996). A hot topic: the origin of hyperthermophiles. *Cell* 85: 789–792. doi.org/10.1016/s0092-8674(00)81262-3

FORTERRE, P. (2002). A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein. *Trends Genet* 18: 236–237. doi.org/10.1016/s0168-9525(02)02650-1

FORTERRE, P. (2015). The universal tree of life: an update. *Front Microbiol* 6: 717. doi.org/10.3389/fmicb.2015.00717

FOX, G.E., PECHMAN, K.R., and WOESE, C.R. (1977). Comparative cataloging of 16S ribosomal ribonucleic acid: molecular approach to procaryotic systematics. *Int J Syst Bacteriol* 27: 44–57. doi.org/10.1099/00207713-27-1-44

FRANÇA, L., ALBUQUERQUE, L. and DA COSTA, M.S. (2015). *Cavicella subterranea* gen. nov., sp. nov., isolated from a deep mineral-water aquifer, and emended description of the species *Perlucidibaca piscinae*. *Int J Syst Evol Microbiol* 65: 3812–3817. doi.org/10.1099/ijsem.0.000493

FRANÇA, L., RAINEY, F.A., NOBRE, M.F. and DA COSTA, M.S. (2006). *Tepidicella xavieri* gen. nov., sp. nov., a betaproteobacterium isolated from a hot spring runoff. *Int J Syst Evol Microbiol* 56: 907–912. doi.org/10.1099/ijs.0.64193-0

FREITAS, M., RAINEY, F.A., NOBRE, M.F., SILVESTRE, A.J. and DA COSTA, M.S. (2003). *Tepidimonas aquatica* sp. nov., a new slightly thermophilic *β*-proteobacterium isolated from a hot water tank. *Syst Appl Microbiol* 26: 376–381. doi.org/10.1078/072320203322497400

FUCIÑOS, P., ABADÍN, C.M., SANROMÁN, A., LONGO, M.A., PASTRANA, L. and RÚA, M.L. (2005). Identification of extracellular lipases/esterases produced by *Thermus thermophilus* HB27: partial purification and preliminary biochemical characterisation. *J Biotechnol* 117: 233–241. doi.org/10.1016/j.jbiotec.2005.01.019

FUCIÑOS, P., GONZÁLEZ, R., ATANES, E., SESTELO, A.B.F., PÉREZ-GUERRA, N., PASTRANA and RÚA, M.R. (2012). Lipases and esterases from extremophiles: overview and case example of the production and purification of an esterase from *Thermus thermophilus* HB27. In *Lipases and Phospholipases: Methods in Molecular Biology (Methods and Protocols)*. Sandoval, G. (editor). Springer. Vol. 861, pp. 239–266. doi.org/10.1007/978-1-61779-600-5_15

FU, X. (2014). Chaperone function and mechanism of small heat-shock proteins. *Acta Biochim Biophys Sin* 46: 347–356. doi.org/10.1093/abbs/gmt152

GALPERIN, M.Y., KRISTENSEN, D.M., MAKAROVA, K.S., WOLF, Y.I. and KOONIN, E.V. (2019). Microbial genome analysis: the COG approach. *Brief Bioinform* 20: 1063–1070. doi.org/10.1093/bib/bbx117

GALPERIN, M.Y., MAKAROVA, K.S., WOLF, Y.I. and KOONIN, E.V. (2015). Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43: D261–D269. doi.org/10.1093/nar/gku1223

GALTIER, N. and LOBRY, J.R. (1997). Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. *J Mol Evol* 44: 632–636. doi.org/10.1007/pl00006186

GALTIER, N., TOURASSE, N. and GOUY, M. (1999). A nonhyperthermophilic common ancestor to extant life forms. *Science* 283: 220–221. doi.org/10.1126/science.283.5399.220

GAMBACORTA, A., GLIOZZI, A. and DE ROSA, M. (1995). Archael lipids and their biotechnological applications. *World J Microbiol Biotechnol* 11: 115–131. doi.org/10.1007/BF00339140

GARCIA, M.J. and GOLA, S. (2016). Gene and whole genome analyses reveal that the mycobacterial strain JS623 is not a member of the species *Mycobacterium smegmatis*. *Microbiol Biotechnol* 9: 269–274. doi.org/10.1111/1751-7915.12336

GARRITY, G.M. (2016). A new genomics-driven taxonomy of *Bacteria* and *Archaea*: are we there yet? *J Clin Microbiol* 54: 1956–1963. doi.org/10.1128/JCM.00200-16

GAUCHER, E.A., KRATZER, J.T. and RANDALL, R.N. (2010). Deep phylogeny–how a tree can help characterize early life on Earth. *Cold Spring Harb Perspect Biol* 2: a002238. doi.org/10.1101/cshperspect.a002238

GAUS, K., RÖSCH, P., PETRY, R., PESCHKE, K.D., RONNEBERGER, O., BURKHARDT, H., BAUMANN, K. and POPP, J. (2006). Classification of lactic acid bacteria with UV-resonance Raman spectroscopy. *Biopolymers* 82: 286–290. doi.org/10.1002/bip.20448

GEVERS, D., COHAN, F.M., LAWRENCE, J.G., SPRATT, B.G., COENYE, T., FEIL, E.J., STACKEBRANDT, E., PEER, Y.V., VANDAMME, P., THOMPSON, F.L., *et al*. (2005). Re-evaluating prokaryotic species. *Nat Rev Microbiol* 3: 733–739. doi.org/10.1038/nrmicro1236

GEVERS, D., DAWYNDT, P., VANDAMME, P., WILLEMS, A., VANCANNEYT, M., SWINGS, J. and DE VOS, P. (2006). Stepping stones towards a new prokaryotic taxonomy. *Phil Trans R Soc Lond B Biol Sci* 361: 1911–1916. doi.org/10.1098/rstb.2006.1915

GILLIS, M., VANDAMME, P., DE VOS, P., SWINGS, J. and KERSTERS, K. (2015). Polyphasic taxonomy. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Bonne, D.R., Castenholz, R.W. and Garrity G.M. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1007/978-0-387-21609-6_7

GLAESER, S.P. and KÄMPFER, P. (2015). Multilocus sequence analysis (MLSA) in prokaryotic taxonomy. *Syst Appl Microbiol* 38: 237–245. doi.org/10.1016/j.syapm.2015.03.007

GLANSDORFF, N., XU, Y. and LABEDAN, B. (2008). The Last Universal Common Ancestor: emergence, constitution and genetic legacy of an elusive forerunner. *Biology Direct* 3: 29. doi.org/10.1186/1745-6150-3-29

GOH, C.B.S., WONG, L.W., PARIMANNAN, S., RAJANDAS, H., LOKE, S., CROFT, L., YULE, C.M., PASBAKHSH, P., LEE, S.M and TAN, J.B.L. (2020). *Chitinophaga extrema* sp. nov., isolated from subsurface soil and leaf litter in a tropical peat swamp forest. *Int J Syst Evol Microbiol* 70: 6355–6363. doi.org/10.1099/ijsem.0.004539

GOMILA, M., PENÄ, A., MULET, M., LALUCAT, J. and GARCIA-VALDÉS, E. (2015). Phylogenomics and systematics in *Pseudomonas*. *Front Microbiol* 6: 214. doi.org/10.3389/fmicb.2015.00214

GOODWIN, S., MCPHERSON, J.D. and MCCOMBIE, W.R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 17: 333–351. doi.org/10.1038/nrg.2016.49

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.org/10.1099/ijs.0.64483-0

GRAYLING, R.A, SANDMAN, K. and REEVE, J.N. (1996). Histones and chromatin structure in hyperthermophilic *Archaea*. *FEMS Microbiol Rev* 18: 203–213. doi.org/10.1111/j.1574-6976.1996.tb00237.x

GRIBALDO, S., POOLE, A.M., DAUBIN, V., FORTERRE, P. and BROCHIER-ARMANET, C. (2010). The origin of eukaryotes and their relationship with the *Archaea*: are we at a phylogenomic impasse? *Nature Rev Microbiol* 8: 743–752. doi.org/10.1038/nrmicro2426

GUERRERO, R. (2001). Bergey's Manual and the classification of prokaryotes. *Int Microbiol* 4: 103–109. doi.org/10.1007/s101230100021

GUPTA, A. and SHARMA, V.K. (2015). Using the taxon-specific genes for the taxonomic classification of bacterial genomes. *BMC Genomics* 16: 396–245. doi.org/10.1186/s12864-015-1542-0

GUPTA, G.N., SRIVASTAVA, S., KHARE, S.K. and PRAKASH, V. (2014). Extremophiles: an overview of microorganism from extreme environment. *Int J Agric Environ Biotechnol* 7: 371–380. doi.org/10.5958/2230-732X.2014.00258.7

GUPTA, R.S., PATEL, S., SAINI, N. and CHEN, S. (2020). Robust demarcation of 17 distinct *Bacillus* species clades, proposed as novel *Bacillaceae* genera, by phylogenomics and comparative genomic analyses: description of *Robertmurraya kyonggiensis* sp. nov. and proposal for an emended genus *Bacillus* limiting it only to the members of the Subtilis and Cereus clades of species. *Int J Syst Evol Microbiol* 70: 5753–5798. doi.org/10.1099/ijsem.0.004475

HABIB, N., KHAN, I.U., HUSSAIN, F., ZHOU, E.M., XIAO, M., DONG, L., ZHI, X.Y. and LI, W.J. (2017). *Meiothermus luteus* sp. nov., a thermophilic bacterium isolated from a hot spring. *Int J Syst Evol Microbiol* 67: 2910–2914. doi.org/10.1099/ijsem.0.002040

HABIB, N., KHAN, I.U., SALAM, N., XIAO, M., AHMED, I., ZHI, X.Y. and LI, W.J. (2018). *Tepidimonas sediminis* sp. nov. and *Tepidimonas alkaliphilus* sp. nov., two novel moderately thermophilic species isolated from a hot spring. *Antonie van Leeuwenhoek* 111: 1023–1031. doi.org/10.1007/s10482-017-1002-8

HAFT, D.H., SELENGUT, J.D. and WHITE, O. (2003). The TIGRFAMs database of protein families. *Nucleic Acids Research* 31:371–373. doi.org/10.1093/nar/gkg128

HAHNKE, R.L., MEIER-KOLTHOFF, J.P., GARCÍA-LÓPEZ, M., MUKHERJEE, S., HUNTEMANN, M., IVANOVA, N.N., WOYKE, T., KYRPIDES, N.C., KLENK, H.P. and GÖKER, M. (2016). Genome-based taxonomic classification of *Bacteroidetes*. *Front Microbiol*: 1–37. doi.org/10.3389/fmicb.2016.02003

HAKI, G.D. and RAKSHIT, S.K. (2003). Developments in industrially important thermostable enzymes: a review. *Bioresour Technol* 89: 17–34. doi.org/10.1016/S0960-8524(03)00033-6

HAMANA, K., MATSUZAKI, S., NIITSU, M. and SAMEJIMA, K. (1990). Pentaamines and hexaamine are present in a thermophilic eubacterium, *Thermomicrobium roseum*. *FEMS Microbiol Lett* 68: 31–34. doi.org/10.1111/j.1574-6968.1990.tb04117.x

HAN, Y., AGARWAL, V., DODD, D., KIM, J., BAE, B., MACKIE, R.I., NAIR, S.K. and CANN, I.K.O. (2012). Biochemical and structural insights into xylan utilization by the thermophilic bacterium *Caldanaerobius polysaccharolyticus. J Biol Chem* 287: 34946–34960. doi.org/10.1074/jbc.M112.391532

HARNVORAVONGCHAI, P., SINGWISUT, R., OUNJAI, P., AROONNUAL, A., KOSIYACHINDA, P., JANVILISRI, T. and CHANKHAMHAENGDECHA, S. (2020). Isolation and characterization of thermophilic cellulose and hemicellulose degrading bacterium,

Thermoanaerobacterium sp. R63 from tropical dry deciduous forest soil. *PLoS One* 15: e0236518. doi.org/10.1371/journal.pone.0236518

HARRISON, J.P., GHEERAERT, N., TSIGELNITSKIY, D. and COCKELL, C.S. (2013). The limits for life under multiple extremes. *Trends Microbiol* 21: 204–212. doi.org/10.1016/j.tim.2013.01.006

HAZEN, T.C. and TABAK, H.H. (2005). Developments in bioremediation of soils and sediments polluted with metals and radionuclides: 2. Field research on bioremediation of metals and radionuclides. *Rev Environ Sci Bio/Technol* 4: 157–183. doi.org/10.1007/s11157-005-2170-y

HEIMBROOK, M.E., WANG, W.L.L. and CAMPBELL, G. (1989). Staining bacterial flagella easily. *J Clin Microbiol* 27: 2612–2615. doi.org/10.1128/jcm.27.11.2612-2615.1989

HENNEMAN, B., EMMERIK, C.V., INGEN, H.V. and DAME, R.T. (2018). Structure and function of archaeal histones. *PLoS Genet* 14: e1007582. doi.org/10.1371/journal.pgen.1007582

HENZ, S.R., HUSON, D.H., AUCH, A.F., NIESELT-STRUWE, K. and SCHUSTER, S.C. (2005). Whole-genome prokaryotic phylogeny. *Bioformatics* 21: 2329–2335. doi.org/10.1093/bioinformatics/bth324

HICKEY, D.A. and SINGER, G.A.C. (2004). Genomic and proteomic adaptations to growth at high temperature. *Gen Biol* 5: 117. doi.org/10.1186/gb-2004-5-10-117

HIDALGO, A., BETANCOR, L., MORENO, R., ZAFRA, O., CAVA, F., FERNÁNDEZ-LAFUENTE, R., GUISÁN, J.M. and BERENGUER, J. (2004). *Thermus thermophilus* as a cell factory for the production of a thermophilic Mn-dependent catalase which fails to be synthesized in an active form in *Escherichia coli*. *Appl Environ Microbiol* 70: 3839–3844. doi.org/10.1128/AEM.70.7.3839-3844.2004

HIRSCH, P., LUDWIG, W., HETHKE, C., SITTIG, M., HOFFMANN, B. and GALLIKOWSKI, C.A. (1998). *Hymenobacter roseosalivarius* gen. nov., sp. nov. from continental antarctic soils and sandstone: bacteria of the *Cytophaga*/*Flavobacterium*/*Bacteroides* line of phylogenetic descent. *Syst Appl Microbiol* 21: 374–383. doi.org/10.1016/S0723-2020(98)80047-7

HORIKOSHI, K. and BULL, A.T. (2011). Prologue: definition, categories, distribution, origin and evolution, pioneering studies, and emerging fields of extremophiles. In *Extremophiles Handbook*. Horikoshi, K., Antranikian, G., Bull, A.T., Robb, F.T., Stetter, K.O. (editors). Springer. pp. 3–15. doi.org/10.1007/978-4-431-53898-1_1

HORNECK, G., KLAUS, D.M. and MANCINELLI, R.L. (2010). Space microbiology. *Microbiol Mol Biol Rev* 74: 121–156. doi.org/10.1128/MMBR.00016-09

HOSOYA, R. and HAMANA, K. (2004). Distribution of two triamines, spermidine and homospermidine, and an aromatic amine, 2-phenylethylamine, within the phylum *Bacteroidetes*. *J Gen Appl Microbiol* 50: 255–260. doi.org/10.2323/jgam.50.255

HOSOYA, R., HAMANA, K., NIITSU, M. and ITOH, T. (2004). Polyamine analysis for chemotaxonomy of thermophilic eubacteria: polyamine distribution profiles within the orders *Aquificales*, *Thermotogales*, *Thermodesulfobacteriales*, *Thermales*, *Thermoanaerobacteriales*, *Clostridiales* and *Bacillales*. *J Gen Appl Microbiol* 50: 271–287. doi.org/10.2323/jgam.50.271

HUBER, G. and STETTER, K.O. (1991). *Sulfolobus metallicus*, sp. nov., a novel strictly chemolithoautotrophic thermophilic archaeal species of metal-mobilizers. *Syst Appl Microbiol* 14: 372–378. doi.org/10.1016/S0723-2020(11)80312-7

HUBER, R., KURR, M., JANNASCH, H.W. and STETTER, K.O. (1989). A novel group of abyssal methanogenic archaebacterial (*Methanopyrus*) growing at 110ºC. *Nature* 342: 833–834. doi.org/10.1038/342833a0

HUBER, R., LANGWORTHY, T.A., KÖNIG, H., THOMM, M., WOESE, C.R., SLEYTR, U.B. and STETTER, K.O. (1986). *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90ºC. *Arch Micobiol* 144: 324–333. doi.org/10.1007/BF00409880

HUBER, R., WILHARM, T., HUBER, D., TRINCONE, A., BURGGRAF, S., KÖNIG, H., RACHEL, R., ROCKNINGER, I., FRICKE, H. and STETTER, K.O. (1992). *Aquifex pyrophilus* gen. nov. sp. nov., represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria. *Syst Appl Microbiol* 15: 340–351. doi.org/10.1016/S0723-2020(11)80206-7

HUDSON, J.A., SCHOFIELD, K.M., MORGAN, H.W. and DANIEL, R.M. (1989). *Thermonema lapsum* gen. nov., sp. nov., a thermophilic gliding bacterium. *Int J Syst Bacteriol* 39: 485–487. doi.org/10.1099/00207713-39-4-485

HUG, L.A., BAKER, B.J., ANANTHARAMAN, K., BROWN, C.T., PROBST, A.J., CASTELLE, C.J., BUTTERFIELD, C.N., HERNSDORF, A.W., AMANO, Y., ISE, K., *et al*. (2016). The physiology and habitat of the last universal common ancestor. *Nature Microbiol* 1: 16048. doi.org/10.1038/NMICROBIOL.2016.48

HUGENHOLTZ, P., CHUVOCHINA, M., OREN, A., and PARKS, D.H. (2021). Prokaryotic taxonomy and nomenclature in the age of sequence data. *ISME J*. https://doi.org/10.1038/s41396-021-00941-x

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 Bioinformatics* 11: 119. doi.org/10.1186/1471-2105-11-119

IMACHI, H., NOBU, M.K., NAKAHARA, N., MORONO, Y., OGAWARA, M., TAKAKI, Y., TAKANO, Y., UEMATSU, K., IKUTA, T., ITO, M. *et al*. (2020). Isolation of na archaeon at the prokaryote-eukaryote interface. *Nature* 577: 519–525. doi.org/10.1038/s41586-019-1916-6

INFANTE-DOMÍNGUEZ, C., DE LA HABA, R.R., CORRAL, P., SANCHEZ-PORRO, C., ARAHAL, D.R. and VENTOSA, A. (2020). Genome-based analyses reveal a synonymy among *Halorubrum distributum* Zvyagintseva and Tarasov 1989; Oren and Ventosa 1996, *Halorubrum terrestre* Ventosa *et al*. 2004, *Halorubrum arcis* Xu *et al*. 2007 and *Halorubrum litoreum* Cui *et al*. 2007. Emended description of *Halorubrum distributum* Zvyagintseva and Tarasov 1989; Oren and Ventosa 1996. *Int J Syst Evol Microbiol* 70: 1698–1705. doi.org/10.1099/ijsem.0.003956

ISHINO, S. and ISHINO, Y. (2014). DNA polymerases as useful reagents for biotechnology – the history of developmental research in the field. *Front Microbiol* 5: 465. doi.org/10.3389/fmicb.2014.00465

IWABATA, H., WATANABE, K., OHKURI, T., TOKOBORI, S.I. and YAMAGISHI, A. (2005). Thermostability of ancestral mutants of *Caldococcus noboribetus* isocitrate dehydrogenase. *FEMS Microbiol Letters* 243: 393–398. doi.org/10.1016/j.femsle.2004.12.030

JAIN, C., RODRIGUEZ-R, L.M, PHILLIPPY, A.M., KONSTANTINIDIS, K.T. and ALURU, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9: 5114. doi.org/10.1038/s41467-018-07641-9

JARDINE, J.L., STOYCHEV, S., MAVUMENGWANA, V. and UBOMBA-JASWA, E. (2018). Screening of potential bioremediation enzymes from hot spring bacteria using conventional plate assays and liquid chromatography - Tandem mass spectrometry (Lc-Ms/Ms). *J Environ Manage* 223: 787–796. doi.org/10.1016/j.jenvman.2018.06.089

JIANG, Y., XIN, F., LU, J., DONG, W., ZHANG, W., ZHANG, M., WU, H., MA, J. and JIANG, M. (2017). State of the art review of biofuels production from lignocellulose by thermophilic bacteria. *Bioresour Technol* 245: 1498–1506. doi.org/10.1016/j.biortech.2017.05.142

JIMOH, A.A. and LIN, J. (2019). Biosurfactant: a new frontier for greener technology and environmental sustainability. *Ecotoxicol Environ Saf* 184: 109607. doi.org/10.1016/j.ecoenv.2019.109607

JORDA, J. and YEATES, T.O. (2011). Widespread disulphide bonding in proteins from thermophilic archaea. *Archaea* 2011: 409156. doi.org/10.1155/2011/409156

JORGE, C.D., BORGES, N., BAGYAN, I., BILSTEIN, A. and SANTOS, H. (2016). Potential applications of stress solutes from extremophiles in protein folding diseases and healthcare. *Extremophiles* 20: 251–259. doi.org/10.1007/s00792-016-0828-8

JOSHI, S., BHARUCHA, C., JHA, S., YADAV, S., NERURKAR, A. and DESAI, A.J. (2008). Biosurfactant production using molasses and whey under thermophilic conditions. *Bioresour Technol* 99: 195–199. doi.org/10.1016/j.biortech.2006.12.010

JUIBARI, M.M., ABBASALIZADEH, S., JOUZANI, G.S. and NORUZI, M. (2011). Intensified biosynthesis of silver nanoparticles using a native extremophilic *Ureibacillus thermosphaericus* strain. *Mater Lett* 65: 1014–1017. doi.org/10.1016/j.matlet.2010.12.056

JUKES, T.H. and CANTOR, C.R. (1969). Evolution of protein molecules. In *Mammalian Protein Metabolism*. Munro, H.N. (editor). Academic Press, New York. pp. 21–132. doi.org/10.1016/B978-1-4832-3211-9.50009-7

KAFARSKI, P. (2012). Rainbow code of biotechnology. *Chemik* 66: 811–816.

KALLMEYER, J., POCKALNY, R., ADHIKARI, R.R., SMITH, D.C. and D'HONDT, S. (2012). Global distribution of microbial abundance and biomass in subseafloor sediment. *PNAS* 109: 16213–16216. doi.org/10.1073/pnas.1203849109

KAMBOUROVA, M. (2018). Thermostable enzymes and polysaccharides produced by thermophilic bacteria isolated from Bulgarian hot springs. *Eng Life Sci* 18: 758–767. doi.org/10.1002/elsc.201800022

KÄMPFER, P. and GLAESER, S.P. (2012). Prokaryotic taxonomy in the sequencing era – the polyphasic approach revisited. *Environ Microbiol* 14: 291–317. doi.org/10.1111/j.1462-2920.2011.02615.x

KÄMPFER, P. and GLAESER, S.P. (2013). Prokaryote characterization and identification. In *The Prokaryotes*. Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E. and Thompson, F. (editors). Springer, Berlin, Heidelberg. pp. 123–147. doi.org/10.1007/978-3-642-30194-0_6

KANEHISA, M., SATO, Y., KAWASHIMA, M., FURUMICHI, M. and TANABE, M. (2016). KEEG as a reference resource for gene and protein annotation. *Nucleic Acids Res* 44: D457–D462. doi.org/10.1093/nar/gkv1070

KARLSSON, R., GONZALES-SILES, L., BOULUND, F., SVENSSON-STADLER, L., SKROVBJERG, S., KARLSSON, A., DAVIDSON, M., HULTH, S., KRISTIANSSON, E and MOORE, E.R.B. (2015). Proteotyping: proteomic characterization, classification and identification of microorganisms – a prospectus. *Syst Appl Microbiol* 38: 246–257. doi.org/10.1016/j.syapm.2015.03.006

KARTHIKEYAN, S., RODRIGUEZ-R, L.M., HERITIER-ROBBINS, P., KIM, M., OVERHOLT, W.A., GABY, J.C., HATT, J.K., SPAIN, J.C., ROSSELLÓ-MÓRA, R, HUETTEL, M, *et al*. (2019). "*Candidatus* Macondimonas diazotrophica", a novel gammaproteobacterial genus dominating crude-oil-contaminated coastal sediments. *ISME J* 13: 2129–2134. doi.org/10.1038/s41396-019-0400-5

KASHEFI, K. and LOVLEY, D.R. (2003). Extending the upper temperature limit for life. *Science* 301: 934. doi.org/10.1126/science.1086823

KATAOKA, M., YAMAOKA, A., KAWASAKI, K., SHIGERI, Y. and WATANABE, K. (2014). Extraordinary denaturant tolerance of keratinolytic protease complex assemblies produced by *Meiothermus ruber* H328. *Appl Microbiol Biotechnol* 98: 2973–2980. doi.org/10.1007/s00253-013-5155-8

KERSTERS, K., POT, B., DEWETTINCK, D., TORCK, U., VANCANNEYT, M., VAUTERIN, L. and VANDAMME, P. (1994). Identification and typing of bacteria by protein electrophoresis. In *Bacterial Diversity and Systematics*. Priest, F.G., Ramos-Cormenzana, A., and Tindall, B.J. (editors). Springer, Boston, MA. Vol 75, pp. 51–66. doi.org/10.1007/978-1-4615-1869-3_3

KHOSRAVI, A. and DOLATABAD, H.K. (2020). Identification and molecular characterization of *Azotobacter chroococcum* and *Azotobacter salinestris* using ARDRA, REP, ERIC, and BOX. *Mol Biol Rep* 47: 307–316. doi.org/10.1007/s11033-019-05133-7

KIM, M., OH, H.S., PARK, S.C. and 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: 346–651. doi.org/10.1099/ijs.0.059774-0

KIM, S.H., PARK, Y.H., SCHMIDT-DANNERT, C. and LEE, P.C. (2010). Redesign, reconstruction, and directed extension of the *Brevibacterium linens* C_{40} carotenoid pathway in *Escherichia coli*. *App Environ Microbiol* 76: 5199–5206. doi.org/10.1128/AEM.00263-10

KLIPCAN, L., SAFRO, I., TEMKIN, B. and SAFRO, M. (2006). Optimal growth temperature of prokaryotes correlates with class II amino acid composition. *FEBS Letters* 580: 1672–1676. doi.org/10.1016/j.febslet.2006.02.013

KLIPPEL, B. and ANTRANIKIAN, G. (2011). Lignocellulose converting enzymes from thermophiles. In *Extremophiles Handbook*. Horikoshi, K., Antranikian, G., Bull, A.T., Robb, F.T., Stetter, K.O. (editors). Springer. pp. 443–474. doi.org/10.1007/978-4-431-53898-1_21

KO, K.S., LEE, N.Y., OH, W.S., LEE, J.H., KI, H.K., PECK, K.R. and SONG, J.H. (2005). *Tepidimonas arfidensis* sp. nov., a novel gram-negative and thermophilic bacterium isolated from the bone marrow of a patient with leukemia in Korea. *Microbiol Immunol* 49: 785–788. doi.org/10.1111/j.1348-0421.2005.tb03669.x

KOGA, Y. (2012). Thermal adaptation of the archaeal and bacterial lipid membranes. *Archaea* 2012: 789652. doi.org/10.1155/2012/789652

KOHLI, I., JOSHI, N.C., MOHAPATRA, S. and VARMA, A. (2020). Extremophile – an adaptive strategy for extreme conditions and applications. *Curr Genomics* 21: 96–110. doi.org/10.2174/1389202921666200401105908

KOLLER, M. and MUKHERJEE, A. (2020). Polyhydroxyalkanoates – linking properties, applications, and end-of-life options. *Chem Biochem Eng Q* 34: 115–129. doi.org/10.15255/CABEQ.2020.1819

KONSTANTINIDIS, K.T. and ROSSELLÓ-MÓRA, R. (2015). Classifying the uncultivated microbial majority: a place for metagenomic data in the *Candidatus* proposal. *Syst Appl Microbiol* 38: 223–230. doi.org/10.1016/j.syapm.2015.01.001

KONSTANTINIDIS, K.T. and TIEDGE, J.M. (2007). Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead. *Curr Opin Microbiol* 10: 504–509. doi.org/10.1016/j.mib.2007.08.006

KONSTANTINIDIS, K.T. and TIEDJE, J.M. (2005a). Genomic insights that advance the species definition for prokaryotes. *PNAS* 102: 2567–2572. doi.org/10.1073/pnas.0409727102

KONSTANTINIDIS, K.T. and TIEDJE, J.M. (2005b) Towards a genome-based taxonomy for prokaryotes. *J Bacteriol* 187: 6258–6264. doi.org/10.1128/JB.187.18.6258-6264.2005

KONSTANTINIDIS, K.T., ROSSELLÓ-MÓRA, R. and AMANN, R. (2017). Uncultivated microbes in need of their own taxonomy. *ISME J* 11: 2399–2406. doi.org/10.1038/ismej.2017.113e

KONSTANTINIDIS, K.T., ROSSELLÓ-MÓRA, R. and AMANN, R. (2020). Advantages outweigh concerns about using genome sequence as type material for prokaryotic taxonomy. *Environ Microbiol* 22: 819–822. doi.org/10.1111/1462-2920.14934

KOUŘILOVÁ, X., SCHWARZEROVÁ, J., PERNICOVÁ, I., SEDLÁŘ, K., MRÁZOVÁ, K., KRZYŽÁNEK, V., NEBESÁŘOVÁ, J. and OBRUČA, S. (2021). The first insight into

polyhydroxyalkanoates accumulation in multi-extremophilic *Rubrobacter xylanophilus* and *Rubrobacter spartanus*. *Microorganisms* 9: 909. doi.org/10.3390/microorganisms9050909

KOWALAK, J.A., DALLUGE, J.J., MCCLOSKEY, J.A. and STETTER, K.O. (1994). The role of posttranscriptional modification in stabilization of transfer RNA from hyperthermophiles. *Biochemistry* 33: 7869–7876. doi.org/10.1021/bi00191a014

KRISTJANSSON, J.K., HERGGVIDSSON, G.O. and GRANT, W.D. (2000). Taxonomy of extremophiles. In *Applied Microbial Systematics*. Priest, F.G. and Goodfellow, M. (editors). Springer, Dordrecht. pp. 231–291. doi.org/10.1007/978-94-011-4020-1_9

KUMAR, A., ALAM, A., TRIPATHI, D., RANI, M., KHATOON, H., PANDEY, S., EHTESHAM, N. and HASNAIN, S.E. (2018a). Protein adaptations in extremophiles: An insight into extremophilic connection of mycobacterial proteome. *Semin Cell Dev Biol* 84: 147–157. doi.org/10.1016/j.semcdb.2018.01.003

KUMAR, S. and NUSSINOV, R. (2001). How do thermophilic proteins deal with heat? *Cell Mol Life Sci* 58: 1216–1233. doi.org/10.1007/PL00000935

KUMAR, S., STECHER, G., LI, M., KNYAZ, C. and TAMURA, K. (2018b). MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution.* 35: 1547–1549. doi.org/10.1093/molbev/msy096

KURR, M., HUBER, R., KÖNIG, H., JANNASCH, H.W., FRIKE, H., TRINCONE, A., KRISTJANSSON, J.K. and STETTER, K.O. (1991). *Methanopyrus kandleri*, gen. and sp. nov. represents a novel group of hyperthermophilic methanogens, growing at 110ºC. *Arch Microbiol* 156: 239–247. doi.org/10.1007/BF00262992

KURTZ, S., PHILLIPPY, A., DELCHER, A.L., SMOOT, M., SHUMWAY, M., ANTONESCU, C. and SALZBERG, S.L. (2004). Versatile and open software for comparing large genomes. *Genome Biol* 5: R12. doi.org/10.1186/gb-2004-5-2-r12

LAGE, O.M., ALBUQUERQUE, L., LOBO-DA-CUNHA, A. and DA COSTA, M.S. (2017). *Mariniblastus fucicola* gen. nov., sp. nov. a novel planctomycete associated with macroalgae. *Int J Syst Evol Microbiol* 67: 1571–1576. doi.org/10.1099/ijsem.0.001760

LAGESEN, K., HALLIN, P., RØDLAND, E.A., STÆRFELDT, H.H., ROGNES, T. and USSERY, D.W. (2007). RNAmmer: consistent annotation of rRNA genes in genomic sequences. *Nucleic Acids Res* 35: 3100–3108. doi.org/10.1093/nar/gkm160

LAGUTIN, K., MACKENZIE, A., HOUGHTON, K.M., STOTT, M.B. and VYSSOTSKI, M. (2014). The identification and quantification of phospholipids from *Thermus* and *Meiothermus* bacteria. *Lipids* 49: 1133–1141. doi.org/10.1007/s11745-014-3946-z

LAL, D., VERMA, M., BEHURA, S.K. and LAL, R. (2016). Codon usage bias in phylum *Actinobacteria*: relevance to environmental adaptation and host pathogenicity. *Res Microbiol* 167: 669–677. doi.org/10.1016/j.resmic.2016.06.003

LALUCAT, J., MULET, M., GOMILA, M. and GARCÍA-VALDÉS, E. (2020). Genomics in bacterial taxonomy: impact on the genus *Pseudomonas*. *Genes* 11: 139. doi.org/10.3390/genes11020139

LANG, J.M., DARLING, A.E. and EISEN, J.A. (2013). Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices. *PLoS One* 8: e62510. doi.org/10.1371/journal.pone.0062510

LANGMEAD, B. and SALZBERG, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357–359. doi.org/10.1038/nmeth.1923

LARKIN, J.M. (1989). Nonphotosynthetic, nonfruiting gliding bacteria. In *Bergey's manual of systematic bacteriology*. Staley, J.T., Bryant, M.P., Pfennig, N. and Holt, J.G. (editors). Williams and Wilkins, Baltimore, MD. Vol. 3, pp. 2010–2138.

LEE, I., CHALITA, M., HA, S.M., NA, S.I., YOON, S.H. and CHUN, J. (2017). ContEst16S: an algorithm that identifies contaminated prokaryotic genomes using 16S RNA gene sequences. *Int J Syst Evol Microbiol* 67: 2053–2057. doi.org/10.1099/ijsem.0.001872

LEE, I., KIM, Y.O., PARK, S.C. and CHUN, J. (2016). OrthoANI: an improved algorithm and software for calculating average nucleotide identity. *Int J Syst Evol Microbiol* 66: 1100–1103. doi.org/10.1099/ijsem.0.000760

LENTZEN, G. and SCHWARZ, T. (2006). Extremolytes: natural compounds from extremophiles for versatile applications. *Appl Microbiol Biotechnol* 72: 623–634. doi.org/10.1007/s00253-006-0553-9

LEVER, M.A, ROGERS, K.L., LLOYD, K.G., OVERMANN, J., SCHINK, B., THAUER, R.K., HOEHLER, T.M. and JØRGENSEN, B.B. (2015). Life under extreme energy limitation: a synthesis of laboratory- and field-based investigations. *FEMS Microbiol Rev* 39: 688–728. doi.org/10.1093/femsre/fuv020

LI, S., ZHONG, H., HU, Y., ZHOU, J., HE, Z. and GU, G. (2014). Bioleaching of a low-grade nickel–copper sulfide by mixture of four thermophiles. *Bioresour Technol* 153: 300–306. doi.org/10.1016/j.biortech.2013.12.018

LI, X., HUANG, Y. and WHITMAN, W.B. (2015). The relationship of the whole genome sequence identity to DNA hybridization varies between genera of prokaryotes. *Antonie van Leeuwenhoek* 107: 241–249. doi.org/10.1007/s10482-014-0322-1

LI, X., XU, H., CHEN, Z.S. and CHEN, G. (2011). Biosynthesis of nanoparticles by microorganisms and their applications. *J Nanomaterials* 2011: 270974. doi.org/10.1155/2011/270974

LIN, F.H. and FORSDYKE, D.R (2006). Prokaryotes that grow optimally in acid have purine-poor codons in long open reading frames. *Extremophiles* 11: 9–18. doi.org/10.1007/s00792-006-0005-6

LIU, D., HUNT, M. and TSAI, I.J. (2018a). Inferring synteny between genome assemblies: a systematic evaluation. *BMC Bioinformatics* 19: 26. doi.org/10.1186/s12859-018-2026-4

LIU, Y., DU, J., LAI, Q., ZENG, R., YE, D., XU, J. and SHAO, Z. (2017). Proposal of nine novel species of the *Bacillus cereus* group. *Int J Syst Evol Microbiol* 67: 2499–2508. doi.org/10.1099/ijsem.0.001821

LIU, Y., LAI, Q. and SHAO, Z. (2018b). Genome-based analysis reveals the taxonomy and diversity of the family *Idiomarinaceae*. *Front Microbiol* 9: 2453. doi.org/10.3389/fmicb.2018.02453

LIU, Y., LAI, Q., GÖKER, M., MEIER-KOLTHOFF, J.P., WANG, M., SUN, Y., WANG, L. and SHAO, Z. (2015). Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci Rep* 5: 14082. doi.org/10.1038/srep14082

LIU, Y., MAKAROVA, K.S., HUANG, W.C.., WOLF, Y.I., NIKOLSKAYA, A.N., ZHANG, X., CAI, M., ZHANG, C.J., XU, W., LUO, Z., *et al*. (2021). Expanded diversity of Asgard archaea and their relationships with eukaryotes. *Nature* 593: 553–557. doi.org/10.1038/s41586-021-03494-3

LOCEY, K.J. and LENNON, J.T. (2016). Scalling laws predict global microbial diversity. *PNAS* 113: 5970–5975. doi.org/10.1073/pnas.1521291113

LOGINOVA, L.G., EGOROVA, L.A., GOLOVACHEVA, R.S. and SEREGINA, L.M. (1984). *Thermus ruber* sp. nov., nom. rev. *Int J Syst Bacteriol* 34: 498–499. doi.org/10.1099/00207713-34-4-498

LOPEZ-CANOVAS, L., BENITEZ, M.B.M., ISIDRON, J.A.H. and SOTO, E.F. (2019). Pulsed field gel electrophoresis: past, present, and future. *Anal Biochem* 573: 17–29. doi.org/10.1016/j.ab.2019.02.020

LÓPEZ-GARCÍA, P. (1999). DNA supercoiling and temperature adaptation: a clue to early diversification of life? *J Mol Evol* 49: 439–452. doi.org/10.1007/pl00006567

LOUCA, S., MAZEL, F., DOEBELI, M. and PARFREY, W. (2019). A census-based estimate of Earth's bacterial and archaeal diversity. *PLoS Biol* 17: e300106. doi.org/10.1371/journal.pbio.3000106

LUDWIG, W. and KLENK, H.P. (2015). Overview: A phylogenetic backbone and taxonomic framework for procaryotic systematics. In *Bergey's Manual of Systematics of Archaea and Bacteria*. Trujillo, M.E., Dedyh, S., De Vos, P., Hedlund, B., Kämpfer, P, Rainey, F.A. and Whitman, W.B. (editors). Published by John Wiley & Sons, Inc., in association with Bergey's Manual Trust. doi.org/10.1002/9781118960608.bm00020

LUDWIG, W. and SCHLEIFER, K.H. (1994). Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. *FEMS Microbiol Rev* 15: 155–173. doi.org/10.1111/j.1574-6976.1994.tb00132.x

LUDWIG, W., GLÖCKER, F.O. and YILMAZ, P. (2011). The use of rRNA gene sequence data in the classification and identification of prokaryotes. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 349–384. doi.org/10.1016/B978-0-12-387730-7.00016-4

LUDWIG, W., STRUNK, O., WESTRAM, R., RICHTER. L., MEIER, H., YADHUKUMAR, BUCHNER, A., LAI, T., STEPPI, S., JOBB, G., *et al*. (2004). ARB: a software environment for sequence data. *Nucleic Acids Res* 32: 1363–1371. doi.org/10.1093/nar/gkh293

LUDWIG, W., VIVER, T., WESTRAM, R., WEATRAM, R., GAGO, J.F., BUSTOS-CAPARROS, E., KNITTEL, K., AMANN, R. and ROSSELLÓ-MÓRA, R (2021). Release LTP_12_2020, featuring a new ARB alignment and improved 16S rRNA tree for prokaryotic type strains. *Syst Appl Microbiol* 126218. doi.org/10.1016/j.syapm.2021.126218

LUO, C., RODRIGUEZ-R, L.M. and KONSTANTINIDIS, K.T. (2014). MyTaxa: an advanced taxonomic classifier for genomic and metagenomic sequences. *Nucleic Acids Res* 42: e73. doi.org/10.1093/nar/gku169

LÖFFLER, F.E., YAN, J., RITALAHTI, K.M., ADRIAN, L., EDWARDS, E.A., KONSTANTINIDIS, K.T., MÜLLER, J.A., FULLERTON, H., ZINDER, S.H. and SPORMANN, A.M. (2013). *Dehalococcoides mccartyi* gen. nov., sp. nov., obligately organohalide-respiring anaerobic bacteria relevant to halogen cycling and bioremediation, belong to a novel bacterial class, *Dehalococcoidia* classis nov., order *Dehalococcoidales* ord. nov. and family *Dehalococcoidaceae* fam. nov., within the phylum *Chloroflexi*. *Int J Syst Evol Microbiol* 63: 625–635. doi.org/10.1099/ijs.0.034926-0

MAEJIMA, Y., LINO, T., MURAGUCHI, Y., FUKUDA, K., OHKUMA, M., SUZUKI, T., MORIUCHI, R., DOHRA, H., KIMBARA, K. and SHINTANI, M. (2020). *Chryseotalea sanaruensis* gen. nov., sp., nov., a member of the family *Cytophagaceae*, isolated from a brackish lake in Hamamatsu Japan. *Curr Microbiol* 77: 306–312. doi.org/10.1007/s00284-019-01823-4

MAHATO, N.K., GUPTA, V., SINGH, P., KUMARI, R., VERMA, H., TRIPATHI, C., RANI, P., SHARMA, A., SINGHVI, N. and SOOD, U. (2017). Microbial taxonomy in the era of OMICS: application of DNA sequences, computational tools and techniques. *Antonie van Leeuwenhoek* 110: 1357–1371. doi.org/10.1007/s10482-017-0928-1

MALAVASI, V., SORU, S. and CAO, G. (2020). Extremophile microalgae: the potential for biotechnological application. *J. Phycol* 56: 559–573. doi.org/10.1111/jpy.12965

MANDEL, M. (1969). New approaches to bacterial taxonomy: perspectives and prospects. *Annu Rev Microbiol* 23: 239–274. doi.org/10.1146/annurev.mi.23.100169.001323

MARGESIN, R. and SCHINNER, F. (2001). Biodegradation and bioremediation of hydrocarbons in extreme environments. *Appl Microbiol Biotechnol* 56: 650–663. doi.org/10.1007/s002530100701

MARGOS, G., GOFTON, A., WIBBERG, D., DANGEL, A., MAROSEVIC, D., LOH, S.M., OSKAM, C. and FINGERLE, V. (2018). The genus *Borrelia* reloaded. *PLoS One* 13: e02084332. doi.org/10.1371/journal.pone.0208432

MARGULIES, M., EGHOLM, M., ALTMAN, W.E., ATTIYA, S., BADER, J.S., BEMBEN, L.A., BERKA, J., BRAVERMAN, M.S., CHEN, Y.J., CHEN, Z., *et al*. (2005). Genome sequencing in open microfabricated high density picoliter reactors. *Nature* 437: 376–380. doi.org/10.1038/nature03959

MARMUR, J., FALKOW, S. and MANDEL, M. (1963). New approaches to bacterial taxonomy. *Ann Rev Microbiol* 17: 329–372. doi.org/10.1146/annurev.mi.17.100163.001553

MARSCHALL, T. (2018). The computational pan-genomics consortium. *Brief Bioinform* 19: 118–135. doi.org/10.1093/bib/bbw089m

MARTEINSSON, V.T., BJORNSDOTTIR, S.H., BIENVENU, N., KRISTJANSSON, J.K. and BIRRIEN, J.L. (2010). *Rhodothermus profundi* sp. nov., a thermophilic bacterium isolated from a deep-sea hydrothermal vent in the Pacific Ocean. *Int J Syst Evol Microbiol* 60: 2729–2734. doi.org/10.1099/ijs.0.012724-0

MARTIN, A. and MCMINN, A. (2018). Sea ice, extremophiles and life on extra-terrestrial ocean worlds. *Int J Astrobiol* 17: 1–16. doi.org/10.1017/S1473550416000483

MARTÍNEZ-CANO, D.J., REYES-PRIETO, M., MARTÍNEZ-ROMERO, E., PARTIDA-MARTÍNEZ, L.P., LATORRE, A., MOYA, A. and DELAYE, L. (2015). Evolution of small prokaryotic genomes. *Front Microbiol* 5: 742. doi.org/10.3389/fmicb.2014.00742

MATSUMOTO, A., KASAI, H., MATSUO, Y., SHIZURI, Y., ICHIKAWA, N., FUJITA, N., ÖMURA, S. and TAKAHASHI, Y. (2013). *Ilumatobacter nonamiense* sp. nov. and *Ilumatobacter coccineum* sp. nov., isolated from seashore sand. *Int J Syst Evol Microbiol* 63: 3404–3408. doi.org/10.1099/ijs.0.047316-0

MAYR, E. (1969). Principles of systematic zoology. McGraw-Hill, *New York.*

MCCARTHY, B.J. and BOLTON, E.T. (1963). An approach to the measurement of genetic relatedness among organisms. *PNAS* 50: 156–164. doi.org/10.1073/pnas.50.1.156

MCMAHON, S. and PARNELL, J. (2014). Weighing the deep continental biosphere. *FEMS Microbiol Ecol* 87: 113–120. doi.org/10.1111/1574-6941.12196

MEHETRE, G.T., DASTAGER, S.G. and DHARNE, M.S. (2019). Biodegradation of mixed polycyclic aromatic hydrocarbons by pure and mixed cultures of biosurfactant producing thermophilic and thermo-tolerant bacteria. *Sci Total Environ* 679: 52–60. doi.org/10.1016/j.scitotenv.2019.04.376

MEHTA, D. and SATYANARAYANA, T. (2013). Diversity of hot environments and thermophilic microbes. In *Environmental and Industrial Biotechnology*. Satyanarayana, T., Littlechild, J. and Kawarabayasi, Y. (editors). Springer, Dordrecht. doi.org/10.1007/978-94-007-5899-5_1

MEHTA, R., SINGHAL, P., SINGH, H., DAMLE, D. and SHARMA, A.K. (2016). Insight into thermophiles and their wide-spectrum applications. *3 Biotech* 6: 81. doi.org/10.1007/s13205-016-0368-z

MEIER-KOLTHOFF, J.P., AUCH, A.F., KLENK, H.P. and GÖKER, M. (2013). Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 14: 60. doi.org/10.1186/1471-2105-14-60

MEIER-KOLTHOFF, J.P., KLENK, H.P. and GÖKER, M. (2014). Taxonomic use of DNA G+C content and DNA–DNA hybridization in the genomic age. *Int J Syst Evol Microbiol* 64: 352–356. doi.org/10.1099/ijs.0.056994-0

MÉRIC, G., YAHARA, K., MAGEIROS, L., PASCOE, B., MAIDEN, M.C.J., JOLLEY, K.A. and SHEPPARD, S.K. (2014). A reference pan-genome approach to comparative bacterial

genomics: identification of novel epidemiological markers in pathogenic *Campylobacter*. *PLoS One* 9: e92798. doi.org/10.1371/journal.pone.0092798

MERINO, N., ARONSON, H.S., BOJANOVA, D.P., FEYHI-BUSKA, J., WONG, M.L., JHANG, S. and GIOVANELLI, D. (2019). Living at the extremes: extremophiles and the limits of life in a planetary context. *Front Microbiol* 10: 780. doi.org/10.3389/fmicb.2019.00780

MESBAH, M., PREMACHANDRAN, U. and WHITMAN, W.B. (1989). Precise measurement of the G+C content of deoxyribonucleic acid by high-performance liquid chromatography. *Int J Syst Bacteriol* 39: 159–167. doi.org/10.1099/00207713-39-2-159

MESBAH, N.M., WHITMAN, W.B. and MESBAH, M. (2011). Determination of the G+C content of prokaryotes. In *Taxonomy of Prokaryotes*. Rainey, R. and Oren, A. (editors). Waltham, MA, Academic Press. pp. 299–324. doi.org/10.1016/b978-0-12-387730-7.00014-0

MICHAEL, A.J. (2016). Polyamines in eukaryotes, Bacteria, and Archaea. *J Biol Chem* 291: 14896–14903. doi.org/10.1074/jbc.R116.734780

MING, H., DUAN, Y.Y., GUO, Q.Q., YIN, Y.R., ZHOU, E.M., LIU, L., LI, S., NIE, G.X. and LI, W.J. (2015). *Meiothermus roseus* sp. nov. a thermophilic bacterium isolated from a geothermal area. *Antonie van Leeuwenhoek* 108: 897–905. doi.org/10.1007/s10482-015-0544-x

MING, H., DUAN, Y.Y., GUO, Q.Q., YIN, Y.R., ZHOU, E.M., LIU, L., LI, S., NIE, G.X. and LI, W.J. (2016). Validation List no. 167. *Int J Syst Evol Microbiol* 66: 1–3. doi.org/10.1099/ijsem.0.000737

MIRETE, S., MORGANTE, V. and GONZÁLEZ-PASTOR, J.E. (2016). Functional metagenomics of extreme environments. *Curr Opin Biotechnol* 38: 143–149. doi.org/10.1016/j.copbio.2016.01.017

MIROSHNICHENKO, M.L, L´HARIDON, S., NERCESSIAN, O., ANTIPOV, A.N., KOSTRIKINA, N.A.., TINDALL, B.J., SCHUMANN, P., SPRING, S., STACKEBRANDT, E., BONCH-OSMOLOVSKAYA, E.A., *et al*. (2003a). *Vulcanithermus mediatlanticus* gen. nov., sp. nov., a novel member of the family *Thermaceae* from a deep-sea hot vent. *Int J Syst Evol Microbiol* 53: 1143–1148. doi.org/10.1099/ijs.0.02579-0

MIROSHNICHENKO, M.L., L´HARIDON, S., JEANTHON, C., ANTIPOV, A.N., KOSTRIKINA, N.A., TINDALL, B.J., SCHUMANN, P., SPRING, S., STACKEBRANDT, E. and BONCH-OSMOLOVSKAYA, E.A. (2003b). *Oceanithermus profundus* gen. nov., sp. nov., a thermophilic, microaerophilic, facultatively chemolithoheterotrophic bacterium from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* 53: 747–752. doi.org/10.1099/ijs.0.02367-0

MNIF, S., CHAMKHA, M., LABAT, M. and SAYADI, S. (2011). Simultaneous hydrocarbon biodegradation and biosurfactant production by oilfield-selected bacteria. *J Appl Microbiol* 111: 525–536. doi.org/10.1111/j.1365-2672.2011.05071.x

MNIF, S., SAYADI, S. and CHAMKHA, M. (2014). Biodegradative potential and characterization of a novel aromatic-degrading bacterium isolated from a geothermal oil field under saline and thermophilic conditions. *Int Biodeter Biodegr* 86: 258–264. doi.org/10.1016/j.ibiod.2013.09.015

MOAYAD, W., ZHA, G. and YAN, Y. (2017). Extremophile current challenges and new gate of knowledge by nanoparticles pathways. *J Pharm Biol Sci* 12: 10–17. doi.org/10.9790/3008-1201021017

MOLINA, I.J., RUIZ-RUIZ, C., QUESADA, E. and BEJAR, V. (2013). Biomedical application of exopolysaccharides produced by microorganisms isolated from extreme environments. In *Extremophiles: sustainable resources and biotechnological implications*. Singh, O.V. (editor). Wiley, Hoboken. pp. 335–355. doi.org/10.1002/9781118394144.ch14

MOORE, E.R.B., MIHAYLOVA, S.A., VANDAMME, P., KRICHEVSKY, M.I. and DIJKSHOORN, L. (2010). Microbial systematics and taxonomy: relevance for a microbial commons. *Res Microbiol* 161: 430–438. doi.org/10.1016/j.resmic.2010.05.007

MORA, C., TITTENSOR, D.P., ADL, S., SIMPSON, A.G.B. and WORM, B. (2011). How many species are there on Earth and in the Ocean? *PLoS Biol* 9: e1001127. doi.org/10.1371/journal.pbio.1001127

MOREIRA, C., RAINEY, F.A., NOBRE, M.F., DA SILVA, M.T. and DA COSTA, M.S. (2000). *Tepidimonas ignava* gen. nov., sp. nov., a new chemolithoheterotrophic and slightly thermophilic member of the *β-Proteobacteria*. *Int J Sys Evol Microbiol* 50: 735–742. doi.org/10.1099/00207713-50-2-735

MORGAN, C.A., HERMAN, N., WHITE, P.A. and VESEY, G. (2006). Preservation of micro-organisms by drying; a review. *J Microbiol Methods* 66: 183–193. doi.org/10.1016/j.mimet.2006.02.017

MORI, K., KAKEGAWA, T., HIGASHI, Y., NAKAMURA, K., MARUYAMA, A. and HANADA, S. (2004). *Oceanithermus desulfurans* sp. nov., a novel thermophilic, sulfur-reducing bacterium isolated from a sulfide chimney in Suiyo Seamount. *Int J Syst Evol Microbiol* 54: 1561–1566. doi.org/10.1099/ijs.0.02962-0

MORI, K., LINO, T., ISHIBASHI, J., KIMURA, H., HAMADA, M. and SUZUKI, K. (2012). *Meiothermus hypogaeus* sp. nov., a moderately thermophilic bacterium isolated from a hot spring. *Int J Syst Evol Microbiol* 62: 112–117. doi.org/10.1099/ijs.0.028654-0

MORIYA, Y., ITOH, M., OKUDA, S., YOSHIZAWA, A.C. and KANEHISA, M. (2007). KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* 35: W182–W185. doi.org/10.1093/nar/gkm321

MUKHERJEE, T., BOSE, S., SEN, U., ROY, C., RAMEEZ, M.J., GHOSH, W. and MUKHOPADHYAY, K. (2016). Genome sequence of the red pigment-forming *Meiothermus taiwanensis* strain RP isolated from Paniphala hot spring, India. G*enome Announc* 4: e00629-16. doi.org/10.1128/genomeA.00629-16

MULET, M., LALUCAT, J. and GARCÍA-VALDÉS, E. (2010). DNA sequence-based analysis of the *Pseudomonas* species. *Environm Microbiol* 12: 1513–1530. doi.org/10.1111/j.1462-2920.2010.02181.x

MUNOZ, R., LÓPEZ-LÓPEZ, A., URDIAIN, M., MOORE, E.R.B. and ROSSELLÓ-MÓRA, R. (2011). Evaluation of matrix-assisted laser desorption ionization-time of flight whole cell

profiles for assessing the cultivable diversity of aerobic and moderately halophilic prokaryotes thriving in solar saltern sediments. *Syst Appl Microbiol* 34: 69–75. doi.org/10.1016/j.syapm.2010.11.012

MUNOZ, R., ROSSELLÓ-MÓRA, R. and AMANN, R. (2016). Revised phylogeny of *Bacteroidetes* and proposal of sixteen new taxa and two new combinations including *Rhodothermaeota* phyl. nov. *Syst Appl Microbiol* 39: 281–296. doi.org/10.1016/j.syapm.2016.04.004

MURRAY, A.E., FREUDENSTEIN, J., GRIBALDO, S., HATZZENPICHLES, R., HUGENHOLTZ, P., KÄMPFER, P., KONSYANTINIDIS, K.T., LANE, C.E., PAPKE, R.T., PARKS, D.H., *et al*. (2020). Roadmap for naming uncultivated Archaea and Bacteria. *Nature Microbiol* 5: 987–994. doi.org/10.1038/s41564-020-0733-x

MURRAY, R.G.E. and SCHLEIFER, K.H. (1994). Taxonomic notes: a proposal for recording the properties of putative taxa of prokaryotes. *Int J Syst Bacteriol* 44: 174–176. doi.org/10.1099/00207713-44-1-174

NEDASHKOVSKAYA, O.I., SUZUKI, M., LEE, J.S., LEE, K.C., SHEVCHENCKO, L.S. and MIKHAILOV, V.V. (2009). *Pseudozobellia thermophila* gen., nov., sp. nov., a bacterium of the family *Flavobacteriaceae*, isolated from the green alga *Ulva fenestrata*. *Int J Syst Evol Microbiol* 59: 806–810. doi.org/10.1099/ijs.0.004143-0

NEOH, H.M., TAN, X.E., SAPRI, H.F. and TAN, T.L. (2019). Pulsed-field gel electrophoresis (PFGE): A review of the "gold standard" for bacteria typing and current alternatives. *Infect Genet Evol* 74: 103935. doi.org/10.1016/j.meegid.2019.103935

NICHOLSON, A.C., GULVIK, C.A., WHITNEY, A.M., HUMRIGHOUSE, B.W., BELL, M.E., HOLMES, B., STEIGERWALT, A.G., VILLARMA, A., SHETH, M., BATRA, D., *et al*. (2020). Division of the genus *Chryseobacterium*: observation of discontinuities in amino acid identity values, a possible consequence of major extinction events, guides transfer of nine species to the genus *Epilithonimonas*, eleven species to the genus *Kaistella*, and three species to the genus *Halpernia* gen. nov., with description of *Kaistella daneshvariae* sp. nov. and *Epilithonimonas vandammei* sp. nov. derived from clinical specimens. *Int J Syst Evol Microbiol* 70: 4432–4450. doi.org/10.1099/ijsem.0.003935

NIELSEN, P., FRITZE, D. and PRIEST, F.G. (1995). Phenetic diversity of alkaliphilic *Bacillus* strains: proposal for nine new species. *Microbiology* 141: 1745–1761. doi.org/10.1099/13500872-141-7-1745

NISHA, M. and SATYANARAYANA, T. (2015). The role of N1 domain on the activity, stability, substrate specificity and raw starch binding of amylopullulanase of the extreme thermophile *Geobacillus thermoleovorans*. *Appl Microbiol Biotechnol* 99: 5461–5474. doi.org/10.1007/s00253-014-6345-8

NOBRE, M.F., TRÜPER, H.G. and DA COSTA, M.S. (1996). Transfer of *Thermus ruber* (Loginova *et al*. 1984), *Thermus silvanus* (Tenreiro *et al*. 1995), and *Thermus chliarophilus* (Tenreiro *et al*. 1995) to *Meiothermus* gen. nov. as *Meiothermus ruber* comb. nov., *Meiothermus*

silvanus comb. nov., and *Meiothermus chliarophilus* comb. nov., respectively, and emendation of the genus *Thermus*. *Int J Syst Bacteriol* 46: 604-606. doi.org/10.1099/00207713-46-2-604

NOUIOUI, I., CARRO, L., GARCIA-LÓPEZ, M., MEIER-KOLTHOFF, J.P., WOYKE, T., KYRPIDES, N.C., PUKALL, R., KLENK, H.P., GOODFELLLOW, M. and GÖKER, M. (2018). Genome-based taxonomic classification of the phylum *Actinobacteria*. *Fron Microbiol* 9: 2007. doi.org/10.3389/fmicb.2018.02007

NUNES, O.C., MANAIA, C.M., DA COSTA, M.S. and SANTOS, H. (1995). Compatible solutes in the thermophilic bacteria *Rhodothermus marinus* and "*Thermus thermophiles*". *Appl Environ Microbiol* 61: 2351–2357. doi.org/10.1128/AEM.61.6.2351-2357.1995

NUNES, O.G., DONATO, M.M. and DA COSTA, M.S. (1992). Isolation and characterization of *Rhodothermus* strains from S. Miguel, Azores. *Syst Appl Microbiol* 15: 92–97. doi.org/10.1016/S0723-2020(11)80144-X

NZILA, A. (2018). Current status of the degradation of aliphatic and aromatic petroleum hydrocarbons by thermophilic microbes and future perspectives. *Int J Environ Res Public Health* 15: 2782. doi.org/doi:10.3390/ijerph15122782

OBRUCA, S., SEDLACEK, P. and KOLLER, M. (2021). The underexplored role of diverse stress factors in microbial biopolymer synthesis. *Bioresour Tecnhol* 326: 124767. doi.org/10.1016/j.biortech.2021.124767

OGAWA, T., YOGO, K., FURUIKE, S., SUTOH, K., KIKUCHI, A. and KINOSITA JR, K. (2015). Direct observation of DNA overwinding by reverse gyrase. *PNAS* 112: 7495–7500. doi.org/10.1073/pnas.1422203112

OLSEN, G.J., WOESE, C.R. and OVERBEEK, R. (1994). The winds of (evolutionary) change: breathing new life into microbiology. *J Bacteriol* 176: 1–6. doi.org/10.1128/jb.176.1.1-6.1994

OREN, A. (2011a). Characterization of pigments of prokaryotes and their use in taxonomy and classification. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 261–282. doi.org/10.1016/B978-0-12-387730-7.00012-7

OREN, A. (2011b). How to name new genera and species of prokaryotes. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 437–463. doi.org/10.1016/B978-0-12-387730-7.00019-X

OREN, A. (2018). Acidophiles. In *eLS*. John Wiley & Sons, Ltd (Ed). doi.org/10.1002/9780470015902.a0000336.pub3

OREN, A. and GARRITY, G.M. (2018). Uncultivated microbes – in need of their own nomenclature? *ISME J* 12: 309–311. doi.org/10.1038/ismej.2017.188

OREN, A. and GARRITY, G.M. (2021). Candidatus List No. 2. Lists of names of prokaryotic *Candidatus* taxa. *Int J Syst Evol Microbiol* 71: 004671. doi.org/10.1099/ijsem.0.004671

OREN, A., GARRITY, G.M. and TRUJILLO, M.E. (2020a). Registration of names of prokaryotic Candidatus taxa in IJSEM. *Int J Syst Evol Microbiol* 70: 3955. doi.org/10.1099/ijsem.0.003791

OREN, A., GARRITY, G.M., PARKER, C.T., CHUVOCHINA, M. and TRUJILLO, M.E. (2020b). Lists of names of prokaryotic *Candidatus* taxa. *Int J Syst Evol Microbiol* 70: 3956–4042. doi.org/10.1099/ijsem.0.003789

OSHIMA, M. and MIYAGAWA, T. (1974). Comparative studies on the fatty acid composition of moderately and extremely thermophilic bacteria. *Lipids* 9: 476–480. doi.org/10.1007/BF02534274

OSHIMA, T. (2007). Unique polyamines produced by an extreme thermophile, *Thermus thermophilus*. *Amino Acids* 33: 367–372. doi.org/10.1007/s00726-007-0526-z

OSHIMA, T. and MORIYA, T. (2008). A preliminary analysis of microbial and biochemical properties of high-temperature compost. *Ann N Y Acad Sci* 1125: 338–344. doi.org/10.1196/annals.1419.012

OVERMANN, J., HUANG, S., NÜBEL, U., HAHNKE, R.L. and TINDALL, B.J. (2019). Relevance of phenotypic information for the taxonomy of not-yet-cultured microorganisms. *Syst Appl Bacteriol* 42: 22–29. doi.org/10.1016/j.syapm.2018.08.009

PACE, N.R. (1997). A molecular view of microbial diversity and the biosphere. *Science* 276: 734–740. doi.org/10.1126/science.276.5313.734

PALANIAPPAN, K., MEIER-KOLTHOFF, J.P., TESHIMA, H., NOLAN, M., LAPIDUS, A., TICE, H., DEL RIO, T.G., CHENG, J.F., HAN, C., TAPIA, R., *et al*. (2013). Genome sequence of the moderately thermophilic sulfur-reducing bacterium *Thermanaerovibrio velox* type strain (Z-9701^{T}) and emended description of the genus *Thermanaerovibrio*. *Stand Genomic Sci* 9: 57–70. doi.org/10.4056/sigs.4237901

PANOSYAN, H., DONATO, P.D., POLI, A. and NICOLAUS, B. (2018). Production and characterization of exopolysaccharides by *Geobacillus thermodenitrificans* ArzA-6 and *Geobacillus toebii* ArzA-8 strains isolated from an Armenian geothermal spring. *Extremophiles* 22: 725–737. doi.org/10.1007/s00792-018-1032-9

PANTAZAKI, A.A., PRITSA, A.A. and KYRIAKIDIS, D.A. (2002). Biotechnologically relevant enzymes from *Thermus thermophilus*. *Appl Microbiol Biotechnol* 58: 1–12. doi.org/10.1007/s00253-001-0843-1

PARKER, C.T., TINDALL, B.J. and GARRITY, G.M. (2019). International code of nomenclature of prokaryotes. Prokaryotic code (revision 2008). *Int J Syst Evol Microbiol* 69: S1–S111. doi.org/10.1099/ijsem.0.000778

PARKS, D.H., CHUVOCHINA, M., CHAUMEIL, P.A., RINKE, C., MUSSIG, A.J. and HUGENHOLTZ, P. (2020). A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat Biotechnol* 38: 1079–1086. doi.org/10.1038/s41587-020-0501-8

PARKS, D.H., CHUVOCHINA, M., WAITE, D.W., RINKE, C., SKARSHEWSKI, A, CHAUMEIL, P.A and HUGENHOLTZ, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 36: 996–1004. doi.org/10.1038/nbt.4229

PARKS, D.H., IMELFORT, M., SKENNERTON, C.T., HUGENHOLTZ, P. and TYSON, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25: 1043–1055. doi.org/10.1101/gr.186072.114

PARTE, A.C. (2014). LPSN – list of prokaryotic names with standing in nomenclature. *Nucleic Acids Res* 42: D613–D616. doi.org/10.1093/nar/gkt1111

PARTE, A.C. (2018). LPSN – list of prokaryotic names with standing in nomenclature. *Int J Syst Evol Microbiol* 68: 1825–1829. doi.org/10.1099/ijsem.0.002786

PASCOAL, F., MAGALHÃES, C. and COSTA, R. (2020). The link between the ecology of the prokaryotic rare biosphere and its biotechnological potential. *Front Microbiol* 11: 231. doi.org/10.3389/fmicb.2020.00231

PATANÉ, J.S.L., MARTINS, J. and SETUBAL, J.C. (2018). Phylogenomics. In *Comparative Genomics: Methods and Protocols, Methods in Molecular Biology*. Setubal, J., Stoye, J. and Stadler, P. (editors). Humana Press, New York. Vol. 1704, pp. 103–187. doi.org/10.1007/978-1-4939-7463-4_5

PATEL, A.K., SINGHANIA, R.R., SIM, S.J. and PANDEY, A. (2019). Thermostable cellulases: current status and perspectives. *Bioresour Technol* 279: 385–392. doi.org/10.1016/j.biortech.2019.01.049

PATEL, B.K.C., SKERRATT, J.H. and NICHOLS, P.D. (1991). The phospholipid ester-linked fatty acid composition of thermophilic bacteria. *Syst Appl Microbiol* 14: 311–316. doi.org/10.1016/S0723-2020(11)80304-8

PATEL, S. and GUPTA, R.S. (2020). A phylogenomic and comparative genomic framework for resolving the polyphyly of the genus *Bacillus*: Proposal for six new genera of *Bacillus* species, *Peribacillus* gen. nov., *Cytobacillus* gen. nov., *Mesobacillus* gen. nov., *Neobacillus* gen. nov., *Metabacillus* gen. nov. and *Alkalihalobacillus* gen. nov. *Int J Syst Evol Microbiol* 70: 406–438. doi.org/10.1099/ijsem.0.003775

PATWARDHAN, A., RAY, S. and ROY, A. (2014). Molecular markers in phylogenetic studies-a review. *J Phylogen Evolution Biol* 2: 131. doi.org/10.4172/2329-9002.1000131

PAZ, A., MESTER, D., BACA, I., NEVO, E. and KOROL, A. (2004). Adaptive role of increased frequency of polypurine tracts in mRNA sequences of thermophilic prokaryotes. *PNAS* 111: 2951–2956. doi.org/10.1073/pnas.0308594100

PEDRUZZI, I., RIVOIRE, C., AUCHINCLOSS, A.H., COUDERT, E., KELLER, G., DE CASTRO, E., BARATIN, D., CUCHE, B.A., BOUGUELERET, L., POUX, S., *et al*. (2015). HAMAP in 2015: updates to the protein family classification and annotation system. *Nucleic Acids Res* 43: D1064–D1070. doi.org/10.1093/nar/gku1002

PEI, A.Y., OBERDORF, W.E., NOSSA, C.W., AGARWAL, A., CHOKSHI, P., GERZ, E.A., JIN, Z., LEE, P., YANG, L., POLES, M., *et al*. (2010). Diversity of 16S rRNA genes within individual prokaryotic genomes. *Appl Environ Microbiol* 76: 3886–3897. doi.org/10.1128/AEM.02953-09

PEÑA, A., TEELING, H., HUERTA-CEPAS, J., SANTOS, F., YARZA, P., BRITO-ECHEVERRIÁ, J., LUCIO, M., SCHMITT-KOPPLIN, P., MESEGUER, I., SCHENOWITZ, C. *et al*. (2010). Fine-scale evolution: genomic, phenotypic and ecological differentiation in two coexisting *Salinibacter ruber* strains. *ISME J* 4: 882–895. doi.org/10.1038/ismej.2010.6

PENG, X., QIAO, W., MI, S., JIA, X., SU, H. and HAN, Y. (2015). Characterization of hemicellulase and cellulase from the extremely thermophilic bacterium *Caldicellulosiruptor owensensis* and their potential application for bioconversion of lignocellulosic biomass without pretreatment. *Biotechnol Biofuels* 8: 131. doi.org/10.1186/s13068-015-0313-0

PÉREZ-CATALUÑA, A., SALAS-MASSÓ, N., DIÉGUEZ, A.L., BALBOA, S., LEMA, A., ROMALDE, J.L. and FIGUERAS, M.J. (2018). Revisiting the taxonomy of the genus *Arcobacter*: getting order from the chaos. *Front Microbiol* 9: 2077. doi.org/10.3389/fmicb.2018.02077

PIKUTA, E.V., HOOVER, R.B. and TANG, J. (2007). Microbial extremophiles at the limits of life. *Crit Rev Microbiol* 33: 183–209. doi.org/10.1080/10408410701451948

PIRES, A.L., ALBUQUERQUE, L., TIAGO, I., NOBRE, M.F., EMPADINHAS, N., VERÍSSIMO, A. and DA COSTA, M.S. (2005a). *Meiothermus timidus* sp. nov., a new slightly thermophilic yellow-pigmented species. *FEMS Microbiol Lett* 245: 39–45. doi.org/doi: 10.1016/j.femsle.2005.02.011

PIRES, A.L., ALBUQUERQUE, L., TIAGO, I., NOBRE, M.F., EMPADINHAS, N., VERÍSSIMO, A. and DA COSTA, M.S. (2005b). Validation List no.104. *Int J Syst Evol Microbiol* 55: 1395–1394. doi.org/10.1099/ijs.0.63883-0

POND, J.L. and LANGWORTHY, T.A. (1987). Effect of growth temperature on the long-chain diols and fatty acids of *Thermomicrobium roseum. J Bacteriol* 169: 1328–1330. doi.org/10.1128/jb.169.3.1328-1330.1987

POTEKHINA, N.V., STRESHINSKAYA, G.M., TUL'SKAYA, E.M. and SHASHKOV, A.S. (2011). Peptidoglycan structure. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 131–164. doi.org/10.1016/B978-0-12-387730-7.00006-1

PRAKASH, O., NIMONKAR, Y. and SHOUCHE, Y.S. (2013). Practice and prospects of microbial preservation. *FEMS Microbiol Lett* 339: 1–9. doi.org/10.1111/1574-6968.12034

PRAKASH, O., VERMA, M., SHARMA, P., KUMAR, M., KUMARI, K., SINGH, A., KUMARI, H., JIT, S., GUPTA, S.K., KHANNA, M., *et al*. (2007). Polyphasic approach of bacterial classification – an overview of recent advances. *Indian J Microbiol* 47: 98–108. doi.org/10.1007/s12088-007-0022-x

PRUESSE, E., QUAST, C., KNITTEL, K., FUCHS, B.M., LUDWIG, W., PEPLIES, J. and GLÖCKNER, F.O. (2007). Silva: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* 35: 7188–7196. doi.org/10.1093/nar/gkm864

QIN, Q.L., XIE, B.B., ZHANG, X.Z., CHEN, X.L., ZHOU, J., OREN, A. and ZHANG, Y.Z. (2014). A proposed genus boundary for the prokaryotes based on genomic insights. *J Bacteriol* 196: 2210–2215. doi.org/10.1128/JB.01688-14

QIN, W., YANG, C., LAI, S., WANG, J., LIU, K. and ZHANG, B. (2013). Bioleaching of chalcopyrite by moderately thermophilic microorganisms. *Bioresour Technol* 129: 200–208. doi.org/10.1016/j.biortech.2012.11.050

QUAST, C., PRUESSE, E., YILMAZ, P., GERKEN, J., SCHWEER, T., YARZA, P., PEPLIES, J. and GLÖCKNER, F.O. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41: D590–D596. doi.org/10.1093/nar/gks1219

RADCHENKOVA, N., TOMOVA, A. and KAMBOUROVA, M. (2018). Biosynthesis of an exopolysaccharide produced by *Brevibacillus Thermoruber* 438. *Biotechnol Biotechnol Equip* 25: 77–79. doi.org/10.5504/bbeq.2011.0115

RAINA, V., NAYAK, T., RAY, L., KUMARI, K. and SUAR, M. (2019). A polyphasic taxonomic approach for designation and description of novel microbial species. In *Microbial Diversity in the Genomic Era*. Das, S. and Dash, H.R. (editors). Academic Press, Elsevier Inc. pp. 137–152. doi.org/10.1016/B978-0-12-814849-5.00009-5

RAINEY, F.A. (2011). How to describe new species of prokaryotes. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 7–14. doi.org/10.1016/B978-0-12-387730-7.00002-4

RAINEY, F.A., SILVA, J., NOBRE, M.F., SILVA, M.T. and DA COSTA, M.S. (2003). *Porphyrobacter cryptus* sp. nov., a novel slightly thermophilic, aerobic, bacteriochlorophyll *a*-containing species. *Int J Syst Evol Microbiol* 53: 35–41. doi.org/10.1099/ijs.0.02308-0

RAINEY, F.A., WARD-RAINEY, N., KROPPENSTEDT, R.M. and STACKEBRANDT, E. (1996). The genus *Nocardiopsis* represents a phylogenetically coherent taxon and a distinct Actinomycete lineage: proposal of *Nocardiopsaceae* fam. nov. *Int J Syst Bacteriol* 46: 1088–1092. doi.org/10.1099/00207713-46-4-1088

RAJENDHRAN, J. and GUNASEKARAN, P. (2011). Microbial phylogeny and diversity: small subunit ribosomal RNA sequence analysis and beyond. *Microbiol Res* 38: 237–245. doi.org/10.1016/j.micres.2010.02.003

RAMASAMY, D., MISHRA, A.K, LAGIER, J.C., PADHMANABHAN, R., ROSSI, M., SENTAUSA, E., RAOULT, D. and FOURNIER, P.E. (2014). A polyphasic strategy incorporating genomic data for the taxonomic description of novel bacterial species. *Int J Syst Evol Microbiol* 64: 384–391. doi.org/10.1099/ijs.0.057091-0

RAMGANESH, S., MAREDZA, A.T. and TEKERE, M. (2014). Microbial exploration in extreme conditions: metagenomic analysis and future perspectives. In *Metagenomics: Methods, Applications and Perspectives*. Camilla, B. (editor). Nova Science Publishers, Inc. pp. 157–181. ISBN: 978-1-61122-358-3

RAPOSO, P., VIVER, T., ALBUQUERQUE, L., FROUFE, H., BARROSO, C., EGAS, C., ROSSELLÓ-MÓRA, R. and DA COSTA, M.S. (2019). Transfer of *Meiothermus chliarophilus* (Tenreiro *et al.* 1995) Nobre *et al*. 1996, *Meiothermus roseus* Ming *et al*. 2016, *Meiothermus terrae* Yu *et al*. 2014 and *Meiothermus timidus* Pires *et al*. 2005, to *Calidithermus* gen. nov., as *Calidithermus chliarophilus* comb. nov., *Calidithermus roseus* comb. nov., *Calidithermus terrae* comb. nov. and *Calidithermus timidus* comb. nov., respectively, and emended description of the genus *Meiothermus*. *Int J Syst Evol Microbiol* 69: 1060–1069. doi.org/10.1099/ijsem.0.003270

RASTOGI, G., BHALA, A., ADHIKARI, A., BISCHOFF, K.M., HUGHES, S.R., CHRISTOPHER, L.P and SANI, R.K. (2010). Characterization of thermostable cellulases produced by *Bacillus* and *Geobacillus* strains. *Bioresour Technol* 101: 8798–8806. doi.org/10.1016/j.biortech.2010.06.001

RAVEENDRAN, S., PARAMESWARAN, B., UMMALYMA, S.B., ABRAHAM, A., KURUVILLA, A., MADHAVAN, A., REBELLO, S. and PANDEY, A. (2018). Applications of microbial enzymes in food industry. *Food Technol Biotechnol* 56: 16–30. doi.org/10.17113/ftb.56.01.18.5491

REED, C.J., LEWIS, H., TREJO, E., WINSTON, V. and EVILIA, C. (2013). Protein adaptations in archaeal extremophiles. *Archaea* 2013: 373275. doi.org/10.1155/2013/373275

REINER, J.B., JUNG, T., LAPP, C.J., SIEDLER, M., BUNK, B., OVERMANN, J. and GESCHER, J. (2018). *Kyrpidia spormannii* sp. nov., a thermophilic, hydrogen-oxidizing, facultative autotroph, isolated from hydrothermal systems at São Miguel Island, and emended description of the genus *Kyrpidia. Int J Syst Evol Microbiol* 68: 3735–3740. doi.org/10.1099/ijsem.0.003037

RICHER, M. and ROSSELLÓ-MÓRA, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *PNAS* 106: 19126–19131. doi.org/10.1073/pnas.0906412106

RIESCO, R., CARRO, L., ROMÁN-PONCE, B., PRIETO, C., BLOM, J., KLENK, H.P., NORMAND, P. and TRUJILLO, M.E. (2018). Defining the species *Micromonospora saelicesensis* and *Micromonospora noduli* under the framework of genomics. *Front Microbiol* 9: 1360. doi.org/10.3389/fmicb.2018.01360

RODRIGUEZ-R, L.M. and KONSTANTINIDIS, K.T. (2016). The enveomics collection: a toolbox for specialized analyses of microbial genomes and metagenomes. *Peer J. Preprints* 4: e1900v1. doi.org/10.7287/peerj.preprints.1900v1

RONA, P.A., KLINKHAMMER, G., NELSEN, T.A., TREFRY, J.H. and ELDERFIELD, H. (1986). Black smokers, massive sulphides and vent biota at the Mid-Atlantic Ridge. *Nature* 321: 33–37. doi.org/10.1038/321033a0

RONG, X. and HUANG, Y. (2014). Multi-locus sequence analysis: taking prokaryotic systematics to the next level. In *Methods in Microbiology (New Approaches to Prokaryotic Systematics)*. Goodfellow, M., Sutcliffe, I. and Chun, J. (editors). Academic Press. Vol. 41, pp. 221–251. doi.org/10.1016/bs.mim.2014.10.001

ROSSELLÓ-MÓRA, R. (2005). Updating prokaryotic taxonomy. *J Bacteriol* 187: 6255–6257. doi.org/10.1128/JB.187.18.6255-6257.2005

ROSSELLÓ-MÓRA, R. (2006). DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation. In *Molecular Identification, Systematics, and Population Structure of Prokaryotes*. Stackebrandt, E. (editor). Springer, Berlin, Heidelberg. pp. 23–50. doi.org/10.1007/978-3-540-31292-5_2

ROSSELLÓ-MÓRA, R. (2012). Towards a taxonomy of *Bacteria* and *Archaea* based on interactive and cumulative data repositories. *Environ Microbiol* 14: 318–334. doi.org/10.1111/j.1462-2920.2011.02599.x

ROSSELLÓ-MÓRA, R. and AMANN, R. (2001). The species concept for prokaryotes. *FEMS Microbiol Rev* 25: 39–67. doi.org/10.1111/j.1574-6976.2001.tb00571.x

ROSSELLÓ-MÓRA, R. and AMANN, R. (2015). Past and future species definitions for *Bacteria* and *Archaea*. *Syst Appl Microbiol* 38: 209–216. doi.org/10.1016/j.syapm.2015.02.001

ROSSELLÓ-MÓRA, R. and KÄMPFER, P. (2004). Defining microbial diversity – the species concept for prokaryotic and eukaryotic microorganisms. In *Microbial Diversity and Bioprospecting*. Bull, A.T. (editor). ASM Press, Washington, DC. pp. 29–39. doi.org/10.1128/9781555817770.ch3

ROSSELLÓ-MÓRA, R. and WHITMAN, W.B. (2019). Dialogue on the nomenclature and classification of prokaryotes. *Syst Appl Microbiol* 42: 5–14. doi.org/10.1016/j.syapm.2018.07.002

ROSSELLÓ-MÓRA, R., LUCIO, M., PEÑA, A., BRITO-ECHEVERRÍA, J., LÓPEZ-LÓPEZ, A., VALENS-VADELL, M., FROMMBERGER, M., ANTÓN, J. and SCHMITT-KOPPLIN, P. (2008). Metabolic evidence for biogeographic isolation of the extremophilic bacterium *Salinibacter ruber*. *ISME J* 2: 242–253. doi.org/10.1038/ismej.2007.93

ROSSELLÓ-MÓRA, R., URDIAIN, M. and LÓPEZ-LÓPEZ, A. (2011). DNA-DNA hybridization. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 325–347. doi.org/10.1016/B978-0-12-387730-7.00015-2

ROSSI-TAMISIER, M., BENAMAR, S., RAOULT, D. and FOURNIER, P.E. (2015). Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. *Int J Syst Evol Microbiol* 65: 1929–1934. doi.org/10.1099/ijs.0.000161

ROTHSCHILD, L.J. and MANCINELLI, R.L. (2001). Life in extreme environments. *Nature* 409: 1092–1101. doi.org/10.1038/35059215

SABATH, N., FERRADA, E., BARVE, A. and WAGNER, A. (2013). Growth temperature and genome size in bacteria are negatively correlated, suggesting genomic streamlining during thermal adaptation. *Genome Biol Evol* 5: 966–977. doi.org/10.1093/gbe/evt050

SAFAR, C., CASTRO, C. and DONATI, E. (2020). Importance of initial interfacial steps during chalcopyrite bioleaching by a thermoacidophilic archaeon. *Minerals* 8: 1009. doi.org/10.3390/microorganisms8071009

SAHOO, K., SAHOO, R.K., GAUR, M. and SUBUDHI, E. (2020). Cellulolytic thermophilic microorganisms in white biotechnology: a review. *Folia Microbiol* 65: 25–43. doi.org/10.1007/s12223-019-00710-6

SAIKI, R.K., GELFAND, D.H., STOFFEL, S., SCHARF, S.J., HIGUCHI, R., HORN, G.T., MULLIS, K.B. and ERLICH, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487–491. doi.org/10.1126/science.239.4839.487

SAITOU, N. and NEI, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4: 406–425. doi.org/10.1093/oxfordjournals.molbev.a040454

SAKAI, H.D. and KUROSAWA, N. (2018). *Saccharolobus caldissimus* gen. nov., sp. nov., a facultatively anaerobic iron-reducing hyperthermophilic archaeon isolated from an acidic terrestrial hot spring, and reclassification of *Sulfolobus solfataricus* as *Saccharolobus solfataricus* comb. nov. and *Sulfolobus shibatae* as *Saccharolobus shibatae* comb. nov. *Int J Syst Evol Microbiol* 68: 1271–1278. doi.org/10.1099/ijsem.0.002665

SAKO, Y., NAKAGAWA, S., TAKAI, K. and HORIKOSHI, K. (2003). *Marinithermus hydrothermalis* gen. nov., sp. nov., a strictly aerobic, thermophilic bacterium from a deep-sea hydrothermal vent chimney. *Int J Syst Evol Microbiol* 53: 59–65. doi.org/10.1099/ijs.0.02364-0

SALAM, N., JIAO, J.Y., ZHANG, X.T. and LI, W.J. (2020). Update on the classification of higher ranks in the phylum *Actinobacteria*. *Int J Syst Evol Microbiol* 70: 1331–1355. doi.org/10.1099/ijsem.0.003920

SALAZAR, A.N. and ABEEL, T. (2018). Approximate, simultaneous comparison of microbial genome architectures via syntenic anchoring of quiver representations. *Bioinformatics* 34: i732–i742. org/10.1093/bioinformatics/bty614

SANGAL, V., GOODFELLOW, M., JONES, A.L., SCHWALBE, E.C., BLOM, J., HOSKISSON, P.A. and SUTCLIFFE, I.C. (2016). Next-generation systematics: an innovative approach to resolve the structure of complex prokaryotic taxa. *Sci Rep* 6: 38392. doi.org/10.1038/srep38392

SANGAL, V., NIEMINEM, L., TUCKER, N.P. and HOSKISSON, P.A. (2014). Revolutionizing prokaryotic systematics through next generation sequencing. In *Methods in Microbiology (New Approaches to Prokaryotic Systematics)*. Goodfellow, M., Sutcliffe, I. and Chun, J. (editors). Academic Press. Vol. 41, pp. 75–101. doi.org/10.1016/bs.mim.2014.07.001

SANGER, F. and COULSON, A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 94: 441–448. doi.org/10.1016/0022-2836(75)90213-2

SANGER, F., NICKLEN, S. and COULSON, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *PNAS* 74: 5463–5467. doi.org/10.1073/pnas.74.12.5463

SANT'ANNA, F.H., AMBROSINI, A., DE SOUZA, R., FERNANDES, G.C., BACH, E., BALSANELLI, E., BAURA, V., BRITO, L.F., WENDISCH, V.F., PEDROSA, F.O., *et al*. (2017). Reclassification of *Paenibacillus riograndensis* as a genomovar of *Paenibacillus sonchi*:

genome-based metrics improve bacterial taxonomic classification. *Front Microbiol* 8: 1849. doi.org/10.3389/fmicb.2017.01849

SANT'ANNA, F.H., BACH, E., PORTO, R.Z., GUELLA, F., SANT'ANNA, E.H. and PASSAGLIA, L.M.P. (2019). Genomic metrics made easy: what to do and where to go in the new era of bacterial taxonomy. *Cri Rev Microbiol* 45: 182–200. doi.org/10.1080/1040841X.2019.1569587

SANTOS, D.K.F., RUFINO, R.D., LUNA, J.M., SANTOS, V.A. and SARUBBO, L.A. (2016). Biosurfactants: multifunctional biomolecules of the 21st century. *Int J Mol Sci* 17: 401. doi.org/10.3390/ijms17030401

SANTOS, H. and DA COSTA, M.S. (2002). Compatible solutes of organisms that live in hot saline environments. *Environ Microbiol* 4: 501–509. doi.org/10.1046/j.1462-2920.2002.00335.x

SANTOS, M.A., WILLIAMS, R.A. and DA COSTA, M.S. (1989). Numerical taxonomy of *Thermus* isolates from hot springs in Portugal. *Syst Appl Microbiol* 12: 310–315. doi.org/10.1016/S0723-2020(89)80079-7

SATYANARAYANA, T., RAGHUKUMAR, C. and SHIVAJI, S. (2005). Extremophilic microbes: diversity and perspectives. *Curr Sci* 89: 78–90. https://www.jstor.org/stable/24110434 (Corpus ID: 83251782)

SCHILDKRAUT, C.L., MARMUR, J. and DOTY, P. (1961). The formation of hybrid DNA molecules and their use in studies of DNA homologies. *J Mol Biol* 3: 595–617. doi.org/10.1016/s0022-2836(61)80024-7

SCHLEIFER, K.H. (2009). Classification of *Bacteria* and *Archaea*: past, present and future. *Syst Appl Microbiol* 32: 533–542. doi.org/10.1016/j.syapm.2009.09.002

SCHLEIFER, K.H. and KANDLER, O. (1972). Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol Rev* 36: 407–477. doi.org/10.1128/br.36.4.407-477.1972

SCHLOSS, P.D., GIRARD, R.A., MARTIN, T., EDWARDS, J. and THRASH, J.C. (2016). Status of the archaeal and bacterial census: an update. *mBio* 7: e00201-16. doi.org/10.1128/mBio.00201-16

SCHUMANN, P. (2011). Peptidoglycan structure. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 101–129. doi.org/10.1016/B978-0-12-387730-7.00005-X

SEEMANN, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* 30: 2068–2069. doi.org/10.1093/bioinformatics/btu153

SEGATA, N., BÖRNIGEN, D., MORGAN, X.C. and HUTTENHOWER, C. (2013). PhyloPhlAn is a new method for improved phylogenetic and taxonomic placement of microbes. *Nat Commun* 4: 2304. doi.org/10.1038/ncomms3304

SEGERER, A., LANGWORTHY, T.A. and STETTER, K.O. (1988). *Thermoplasma acidophilum* and *Thermoplasma volcanium* sp. nov. from solfatara fields. *Syst Appl Microbiol* 10: 161–171. doi.org/10.1016/S0723-2020(88)80031-6

SEN, S.K., RAULT, S., DORA, T.K. and DAS MOHAPATRA, P.K. (2014). Contribution of hot spring bacterial consortium in cadmium and lead bioremediation through quadratic programming model. *J Hazard Mater* 265: 47–60. doi.org/10.1016/j.jhazmat.2013.11.036

SENTAUSA, E. and FOURNIER, P.E. (2013). Advantages and limitations of genomics in prokaryotic taxonomy. *Clin Microbiol Infect* 19: 790–795. doi.org/10.1111/1469-0691.12181

SETUBAL, J.C. and STADLER, P.F. (2018). Gene phylogenies and orthologous groups. In *Comparative Genomics: Methods and Protocols, Methods in Molecular Biology*. Setubal, J., Stoye, J. and Stadler, P. (editors). Humana Press, New York. Vol. 1704, pp. 1–28. doi.org/10.1007/978-1-4939-7463-4_1

SETUBAL, J.C., ALMEIDA, N.F. and WATTAM, A.R. (2018). Comparative genomics for prokaryotes. In *Comparative Genomics: Methods and Protocols, Methods in Molecular Biology*. Setubal, J., Stoye, J. and Stadler, P. (editors). Humana Press, New York. Vol. 1704, pp. 55–78. doi.org/10.1007/978-1-4939-7463-4_3

SEUYLEMEZIAN, A., ARONSON, H.S., TAN, J., LIN, M., SCHUBERT, W. and VAISHAMPAYAN, P. (2018). Development of a custom MALDI-TOF MS database for species-level identification of bacterial isolates collected from spacecraft and associated surfaces. *Front Microbiol* 9: 780. doi.org/10.3389/fmicb.2018.00780

SHARAFI, H., ABDOLI, M., HAJFARAJOLLAH, H., SAMIE, N., ALIDOUST, L., ABBASI, H., FOOLADI, J., ZAHIRI, H.S. and NOGHABI, K.A. (2014). First report of a lipopeptide biosurfactant from thermophilic bacterium *Aneurinibacillus thermoaerophilus* MK01 newly isolated from municipal landfill site. *Appl Biochem Biotechnol* 173: 1236–1249. doi.org/10.1007/s12010-014-0928-9

SHARP, R.J. and WILLIAMS, R.A.D. (1988). Properties of *Thermus ruber* strains isolated from Icelandic hot springs and DNA:DNA homology of *Thermus ruber* and *Thermus aquaticus*. *Appl Environ Microbiol* 54: 2049–2053. doi.org/10.1128/AEM.54.8.2049-2053.1988

SHIGI, N., SUZUKI, T., TAMAKOSHI, M., OSHIMA, T. and WATANABE, K. (2002). Conserved bases in the TΨC loop of tRNA are determinants for thermophile-specific 2-thiouridylation at position 54. *J Biol Chem* 42: 39128–39135. doi.org/10.1074/jbc.M207323200

SHIMIZU, H., YOKOBORI, S.I., OHKURI, T., YOKOGAWA, T., NISHIKAWA, K. and YAMAGISHI, A. (2007). Extremely thermophilic translation system in the common ancestor commonote: ancestral mutants of glycyl-tRNA synthetase from the extreme thermophile *Thermus thermophilus. JMB* 369: 1060–1069. doi.org/10.1016/j.femsle.2004.12.030

SILIAKUS, M.F., VAN DER OOST, J. and KENGEN, W.M. (2017). Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. *Extremophiles* 21: 651–670. doi.org/10.1007/s00792-017-0939-x

SILVA, Z., HORTA, C., DA COSTA, M.S., CHUNG, A.P. and RAINEY, F.A. (2000). Polyphasic evidence for the reclassification of *Rhodothermus obamensis* Sako *et al*., 1996 as a member of the species *Rhodotermus marinus* Alfredsson *et al*., 1998. *Int J Syst Evol Bacteriol* 50: 1457–1461. doi.org/10.1099/00207713-50-4-1457

SINGHAL, N., KUMAR, M., KANAUJIA, P.K. and VIRDI, J.S. (2015). MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Front Microbiol* 6: 791. doi.org/10.3389/fmicb.2015.00791

SLOBODKIN, A.I., REYSENBACH, A.L., SLOBODKINA, G.B., BASLEROV, R.V., KOSTRIKINA, N.A., WAGNER, I.D. and BONCH-OSMOLOVSKAYA, E.A. (2012). *Thermosulfurimonas dismutans*, gen. nov., sp. nov., an extremely thermophilic sulfur-disproportionating bacterium from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* 62: 2565–2571. doi.org/10.1099/ijs.0.034397-0

SMIBERT, R.M. and KRIEG, N.R. (1981). General Characterization. In *Manual of methods for general bacteriology*. Gerhardt, P., Murray, R.G.E., Costilow, R.N., Nester, E.W., Wood, W.A., Krieg, N.R. and Phillips, G.B. (editors). American Society for Microbiology, Washington DC. pp. 409–442.

SNEATH, H.A. (1957). The application of computers to taxonomy. *J Gen Microbiol* 17: 201–226. doi.org/10.1099/00221287-17-1-201

SNIR, S. (2016). Ordered orthology as a tool in prokaryotic evolutionary inference. *Mob Genet Elements* 6: e1120576. doi.org/10.1080/2159256X.2015.1120576

SOHLENKAMP, C. and GEIGER, O. (2016). Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol Rev* 40: 133–159. doi.org/10.1093/femsre/fuv008

SOON, W.W., HARIHARAN, M. and SNYDER, M.P. (2013). High-throughput sequencing for biology and medicine. *Mol Syst Biol* 9: 640. doi.org/10.1038/msb.2012.61

SÖRBO, B. (1987). Sulfate: Turbidimetric and nephelometric methods. In *Methods Enzymology. Sulfur and Sulfur Amino Acids*. Jakoby, W.B. and Griffith, O.W. (editors). Academic Press. Vol. 143, pp. 3–6. doi.org/10.1016/0076-6879(87)43003-6

STACKEBRANDT, E. and EBERS, J. (2006). Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 33: 152–155.

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 Bacteriol* 44: 846–849. doi.org/10.1099/00207713-44-4-846

STACKEBRANDT, E., FREDERIKSEN, W., GARRITY, G.M., GRIMONT, P.A.D., KÄMPFER, P., MAIDEN, M.C.J., NESME, X., ROSSELLÓ-MÓRA, R., SWINGS, J., TRÜPER, H.G., *et al*. (2002). Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* 52: 1043–1047. doi.org/10.1099/00207713-52-3-1043

STAMATAKIS, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22: 2688–2690. doi.org/10.1093/bioinformatics/btl446

STEINSBU, B.O., TINDALL, B.J., TORSVIK, V.L., THORSETH, I.H., DAAE, F.L. and PEDERSEN, R.B. (2011). *Rhabdothermus arcticus* gen. nov., sp. nov., a member of the family *Thermaceae* isolated from a hydrothermal vent chimney in the Soria Moria vent field on the

Arctic Mid-Ocean Ridge. *Int J Syst Evol Microbiol* 61: 2197–2204. doi.org/10.1099/ijs.0.027839-0

STETTER, K.O. (1996). Hyperthermophilic procaryotes. *FEMS Microbiol Rev* 18: 149–158. doi.org/10.1111/j.1574-6976.1996.tb00233.x

STETTER, K.O. (2006a). History of discovery of the first hyperthermophiles. *Extremophiles* 10: 357–362. doi.org/10.1007/s00792-006-0012-7

STETTER, K.O. (2006b). Hyperthermophiles in the history of life. *Phil Trans R Soc B* 361: 1837–1843. doi.org/10.1098/rstb.2006.1907

STEVENS, K.M, SWADLING, J.B., HOCHER, A., BANG, C., GRIBALDO, S., SCHMITZ, R.A. and WARNECKE, T. (2020). Histone variants in archaea and the evolution of combinatorial chromatin complexity. *PNAS* 117: 33384–33395. doi.org/10.1073/pnas.2007056117

STEWART, L.C., JUNG, J.H., KIM, Y.T., KWON, S.W., PARK, C.S. and HOLDEN, J.F. (2015). *Methanocaldococcus bathoardescens* sp. nov., a hyperthermophilic methanogen isolated from a volcanically active deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* 65: 1280–1283. doi.org/10.1099/ijs.0.000097

STRAUB, T.J. and ZHAXYBAYEVA, O. (2017). A null model for microbial diversification. *PNAS* 114: E5414–5423. doi.org/10.1073/pnas.1619993114

SU, X., HAN, Y., DODD, D., MOON, Y.H., YOSHIDA, S., MACKIE, R.I. and CANN, I.K.O. (2013). Reconstitution of a thermostable xylan-degrading enzyme mixture from the bacterium *Caldicellulosiruptor bescii. Appl Microbiol Biotechnol 79*: 1481–1490. doi.org/10.1128/AEM.03265-12

SUN, D.L., JIANG, X., WU, Q.L. and ZHOU, N.Y. (2013). Intragenomic heterogeneity of 16S rRNA genes causes overestimation of prokaryotic diversity. *Appl Environ Microbiol* 79: 5962–5969. doi.org/10.1128/AEM.01282-13

SUTCLIFFE, I.C. (2015). Challenging the anthropocentric emphasis on phenotypic testing in prokaryotic species descriptions: rip it up and start again. *Front Genet* 6: 218. doi.org/10.3389/fgene.2015.00218

SUTCLIFFE, I.C., DIJKSHOORN, L., WHITMAN, W.B. and ON BEHALF OF THE ICSP EXECUTIVE BOARD (2020). Minutes of the International Committee on Systematics of Prokaryotes online discussion on the proposed use of gene sequences as type for naming of prokaryotes, and outcome of vote. *Int J Syst Evol Microbiol* 70: 4416–4417. doi.org/10.1099/ijsem.0.004303

SUUTARI, M. and LAAKSO, S. (1994). Microbial fatty acids and thermal adaptation. *Crit Rev Microbiol* 20: 285–328. doi.org/10.3109/10408419409113560

TABAK, H.H., LENS, P., VAN HULLEBUSCH, E.D. and DEJONGHE, W. (2005). Developments in bioremediation of soils and sediments polluted with metals and radionuclides – 1. Microbial processes and mechanisms affecting bioremediation of metal contamination and influencing metal toxicity and transport. *Rev Environ Sci Bio/Technol* 4: 115–156. doi.org/10.1007/s11157-005-2169-4

TAKAI, K., NAKAMURA, K., TOMOHIRO, T., TSUNOGAI, U., MIYAZAKI, M., MIYAZAKI, J., HIRAYAMA, H., NAKAGAWA, S., NUNOURA, T. and HORIKOSHI, K. (2008). Cell proliferation at 122°C and isotopically heavy CH_4 production by a hyperthermophilic methanogen under high-pressure cultivation. *PNAS* 105: 10949–10954. doi.org/10.1073/pnas.0712334105

TAMBONG, J.T. (2019). Taxogenomics and systematics of the genus *Pantoea*. *Front Microbiol* 10: 2463. doi.org/10.3389/fmicb.2019.02463

TAMURA, K., STECHER, G., PETERSON, D., FILIPSKI, A. and KUMAR, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30: 2725–2729. doi.org/10.1093/molbev/mst197

TAN, J.L., NG, K.P., ONG, C.S. and NGEOW, Y.F. (2017). Genomic comparisons reveal microevolutionary differences in *Mycobacterium abscessus* subspecies. *Front Microbiol* 8: 2042. doi.org/10.3389/fmicb.2017.02042

TANAKA, M., MINO, S., OGURA, Y., HAYASHI, T. and SAWABE, T. (2018). Availability of nanopore sequences in the genome taxonomy for *Vibrionaceae* systematics: *Rumoiensis* clade species as a test case. *PeerJ* 6: e5018. doi.org/10.7717/peerj.5018

TAO, W., LIN, J., WANG, W., HUANG, H. and LI, S. (2020). Biodegradation of aliphatic and polycyclic aromatic hydrocarbons by the thermophilic bioemulsifier-producing *Aeribacillus pallidus* strain SL-1. *Ecotoxicol Enviro Saf* 189: 109994. doi.org/10.1016/j.ecoenv.2019.109994

TATUSOV, R.L., NATALE, D.A., GARKAVTSEV, I.V., TATUSOVA, T.A., SHANKAVARAM, U.T., RAO, B.S., KIRYUTIN, B., GALPERIN, M.Y., FEDOROVA, N.D. and KOONIN, E.V. (2001). The COG database: new developments in phylogenetic classification of proteins from complete genomes. *Nucleic Acids Res* 29: 22–28. doi.org/10.1093/nar/29.1.22

TEELING, H., MEYERDIERKS, A., BAUER, M., AMANN, R. and GLÕCKNER, F.O. (2004). Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environm Microbiol* 6: 938–947. doi.org/10.1111/j.1462-2920.2004.00624.x

TENG, J.L.L., TANG, Y., HUAG, Y., GUO, F.B., WEI, W., CHEN, J.H.K., WONG, S.S.Y., LAU, S.K.P. and WOO, P.C.Y. (2016). Phylogenomic analyses and reclassification of species within the genus *Tsukamurella*: insights to species definition in the post-genomic era. *Front Microbiol* 7: 1137. doi.org/10.3389/fmicb.2016.01137

TENREIRO, S., NOBRE, M.F. and DA COSTA, M.S. (1995). *Thermus silvanus* sp. nov. and *Thermus chliarophilus* sp. nov., two new species related to *Thermus ruber* but with lower growth temperatures. *Int J Syst Bacteriol* 45: 633–639. doi.org/10.1099/00207713-45-4-633

TENREIRO, S., NOBRE, M.F., RAINEY, F.A., MIGUEL, C. and DA COSTA, M.S. (1997). *Thermonema rossianum* sp. nov., a new thermophilic and slightly halophilic species from saline hot springs in Naples, Italy. *Int J Syst Bacteriol* 47: 122–126. doi.org/10.1099/00207713-47-1-122

TERUI, Y., OHNUMA, M., HIRAGA, K., KAWASHIMA, E. and OSHIMA, T. (2005). Stabilization of nucleic acids by unusual polyamines produced by an extreme thermophile, *Thermus thermophilus*. *Biochem J* 388: 427–433. doi.org/10.1042/BJ20041778

TETTELIN, H., MASIGNANI, V., CIESLEWICZ, M.J., DONATI, C., MEDINI, D., WARD, N.L., ANGIUOLI, S.V., CRABTREE, J., JONES, A.L., DURKIN, A.S. *et al*. (2005). Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial ‘‘pan-genome’’. *PNAS* 102: 13950–13955. doi.org/10.1073/pnas.0506758102

TETTELIN, H., RILEY, D., CATTUTO, C. and MEDINI, D. (2008). Comparative genomics: the bacterial pan-genome. *Curr Opin Microbiol* 12: 472–477. doi.org/10.1016/j.mib.2008.09.006

THOMPSON, C.C., AMARAL, G.R., CAMPEÃO, M., EDWARDS, R.A., POLZ, M.F., DUTILH, B.E., USSERY, D.W., SAWABE, T., SWINGS, J. and THOMPSON, F.L. (2015). Microbial taxonomy in the post-genomic era: rebuilding from scratch? *Arch Microbiol* 197: 359–370. doi.org/10.1007/s00203-014-1071-2

THOMPSON, C.C., CHIMETTO, L., EDWARDS, R.A., SWINGS, J., STACKEBRANDT, E. and THOMPSON, F.L. (2013). Microbial genomic taxonomy. *BMC Genomics* 14: 913. doi.org/10.1186/1471-2164-14-913

TIAN, R.M., CAI, L., ZHANG, W.P., CAO, H.L. and QUIAN, P.Y. (2015). Rare events of intragenus and intraspecies horizontal transfer of the 16S rRNA gene. *Gen Biol Evol* 7: 2310–2320. doi.org/10.1093/gbe/evv143

TINDALL, B.J. (2008). Confirmation of deposit, but confirmation of what? *Int J Syst Evol Microbiol* 58: 1785–1787. doi.org/10.1099/ijs.0.2008/006023-0

TINDALL, B.J. and GARRITY, P. (2008). Proposals to clarify how type strains are deposited and made available to the scientific community for the purpose of systematic research. *Int J Syst Evol Microbiol* 58: 1987–1990. doi.org/10.1099/ijs.0.2008/006155-0

TINDALL, B.J., ROSSELLÓ-MÓRA, R., BUSSE, H., 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.org/10.1099/ijs.0.016949-0

TINDALL, B.J., SIRORSKI, J., SMIBERT, R.M. and KRIEG, N.R. (2007). Phenotypic characterization and the principles of comparative systematics. In *Methods for General and Molecular Microbiology*. Reddy, C.A., Beveridge, T.J., Breznak, J.A., Marzluf, G., Schmidt, T.M., Snyder, L.R. (editors). ASM Press, Washington, DC. pp. 330–393. doi.org/10.1128/9781555817497.ch15

TIQUIA-ARASHIRO, S. and RODRIGUES, D. (2016). Thermophiles and psychrophiles in nanotechnology. In *Extremophiles: Applications in Nanotechnology*. Springer, Cham. pp. 89–127. doi.org/10.1007/978-3-319-45215-9_3

TRÜPER, H.G. and SCHLEIFER, K.H. (2006). Prokaryote characterization and identification. In *The Prokaryotes*. Dworkin, M., Falkow, S., Rosenberg, E., Schleifer, K.H., Stackebrandt, E. and Thompson, F. (editors). Springer, New York. pp. 58–79. doi.org/10.1007/0-387-30741-9_4

ULRIH, N.P., GMAJNER, D. and RASPOR, P. (2009). Structural and physicochemical properties of polar lipids from thermophilic archaea. *Appl Microbiol Biotechnol* 84: 249–260. doi.org/10.1007/s00253-009-2102-9

URBIETA, M.S., DONATI, E.R., CHAN, K.G., SHAHAR, S., SIN, L.L. and GOH, K.M. (2015). Thermophiles in the genomic era: biodiversity, science, and applications. *Biotechnol Adv* 33: 633–647. doi.org/10.1016/j.biotechadv.2015.04.007

URSING, J.B., ROSSELLÓ-MÓRA, R.A., GARCÍA-VALDÉS, E. and LALUCAT, J. (1995). Taxonomic note: a pragmatic approach to the nomenclature of phenotypically similar genomic groups. *Int J Syst Bacteriol* 45: 604. doi.org/10.1099/00207713-45-3-604

VALENTINI, M., STORELLI, N. and LAPOUGE, K. (2011). Identification of C4-dicarboxylate transport systems in *Pseudomonas aeruginosa* PAO1. *J Bacteriol* 193: 4307–4316. doi.org/10.1128/JB.05074-11

VALERIANI, F., BIAGINI, T., GIAMPAOLI, S., CROGNALE, S., SANTONI, D. and SPICA, V.R. (2016). Draft genome sequence of *Tepidimonas taiwanensis* strain VT154-175. *Genome Announc* 4: 1–2. doi.org/10.1128/genomeA.00942-16

VAN STEENBERGEN, T.J.M., MENARD, C., TIJHOF, C.J., MOUTON, C. and DE GRAAFF, J. (1993). Comparison of three molecular typing methods in studies of transmission of *Porphyromonas gingivalis*. *J Med Microbiol* 39: 416–421. doi.org/10.1099/00222615-39-6-416

VANDAMME, P., POT, B., GILLIS, M., DE VOS, P. and SWINGS, J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* 60: 407–438. doi.org/10.1128/mr.60.2.407-438.1996

VAREL, V.H. and BRYANT, M.P. (1974). Nutritional features of *Bacteroides fragilis* subsp. *fragilis*. *Appl Microbiol* 28: 251–257.

VARSHNEY, P., MIKULIC, P., VONSHAK, A., BEARDALL, J. and WANGIKAR, P.P. (2015). Extremophilic micro-algae and their potential contribution in biotechnology. *Bioresource Technology* 184: 363–372. doi.org/10.1016/j.biortech.2014.11.040

VERÍSSIMO, A., MORAIS, P.V., DIOGO, A., GOMES, C. and DA COSTA, M.S. (1996). Characterization of *Legionella* species by numerical analysis of whole-cell protein electrophoresis. *Int J Syst Bacteriol* 46: 41–49. doi.org/10.1099/00207713-46-1-41

VERNIKOS, G., MEDINI, D., RILEY, D.R. and TETTELIN, H. (2015). Ten years of pan-genome analyses. *Curr Opin Microbiol* 23: 148–154. doi.org/10.1016/j.mib.2014.11.016

VIEIRA-SILVA, S. and ROCHA, E.P.C. (2010). The systemic imprint of growth and its uses in ecological (meta)genomics. *PLoS Genet* 6: e1000808. doi.org/10.1371/journal.pgen.1000808

VIELLE, C. and ZEIKUS, G.J. (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* 65: 1–43. doi.org/10.1128/MMBR.65.1.1-43.2001

VIVER, T., ORELLANA, L., GONZÁLEZ-TORRES, P., DÍAZ, S., URDIAIN, M., FARÍAS, M.E., BENES, V., KÄMPFER, P., SHAHINPEI, A., AMOOZEGAR, M.A., *et al*. (2018). Genomic comparison between members of the *Salinibacteraceae* family, and description of a

new species of *Salinibacter* (*Salinibacter altiplanensis* sp. nov.) isolated from high altitude hypersaline environments of the Argentinian Altiplano. *Syst Appl Microbiol* 41: 198–212. doi.org/10.1016/j.syapm.2017.12.004

VOGT, S., LÖFFLER, K., DINKELACKER, A.G., BADER, B., AUTENRIETH, I.B., PETER, S. and LIESE, J. (2019). Fourier-Transform Infrared (FTIR) spectroscopy for typing of clinical *Enterobacter cloacae* complex isolates. *Front Microbiol* 10: 2582. doi.org/10.3389/fmicb.2019.02582

WAIT, R., CARRETO, L., NOBRE, M.F., FERREIRA, A.M. and DA COSTA, M.S. (1997). Characterization of novel long-chain 1,2-diols in *Thermus* species and demonstration that *Thermus* strains contain both glycerol-linked and diol-linked glycolipids. *J Bacteriol* 179: 6154–6162. doi.org/10.1128/jb.179.19.6154-6162.1997

WANG, J., SALEM, D.R. and SANI, R.K. (2021). Two new exopolysaccharides from a thermophilic bacterium *Geobacillus* sp. WSUCF1: Characterization and bioactivities. *New Biotechnol* 61: 29–39. doi.org/10.1016/j.nbt.2020.11.004

WATTIAU, P., BOLAND, C., and BERTRAND, S. (2011). Methodologies for *Salmonella enterica* subsp. *enterica* subtyping: gold standards and alternatives. *Appl Environ Microbiol* 77: 7877–7885. doi.org/10.1128/AEM.05527-11

WAYNE, L.G., BRENNER, D.J., COLWELL, R.R., GRIMONT, P.A.D., KANDLER, O., KRICHEVSKY, M.I., MOORE, L.H., MOORE, W.E.C., MURRAY, R.G.E, STACKEBRANDT, E., *et al*. (1987). Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int J Syst Bacteriol* 37: 463–464. doi.org/10.1099/00207713-37-4-463

WEIMANN, A., MOOREN, K., FRANK, J., POPE, P.B., BREMGES, A. and MCHARDY, A.C. (2016). From genomes to phenotypes: Traitar, the microbial trait analyzer. *mSystems* 27: e00101. doi.org/10.1128/mSystems.00101-16

WEISS, M.C., PREINER, M., XAVIER, J.C., ZIMORSKI, V. and MARTIN, W.F. (2018). The last universal common ancestor between ancient Earth chemistry and the onset of genetics. *PLoS Genet* 14: e1007518. doi.org/10.1371/journal.pgen.1007518

WEISS, M.C., SOUSA, F.L., MRNJAVAC, N., NEUKIRCHEN, S., ROETTGER, M., NELSON-SATHI, S. and MARTIN, W.F. (2016). A new view of the Tree of Life. *Nature Microbiol* 1: 16116. doi.org/10.1038/nmicrobiol.2016.116

WESTLEY, R.A.D. (1987). Thiocyanate and thiosulfate. In *Methods Enzymology. Sulfur and Sulfur Amino Acids*. Jakoby, W.B. and Griffith, O.W. (editors). Academic Press.Vol. 143, pp. 22–25. doi.org/10.1016/0076-6879(87)43008-5

WHEATON, G., COUNTS, J., MUKHERJEE, A., KRUH, J. and KELLY, R. (2015). The confluence of heavy metal biooxidation and heavy metal resistance: implications for bioleaching by extreme thermoacidophiles. *Minerals* 5: 397–451. doi.org/10.3390/min5030397

WHITMAN, W.B. (2015). Genome sequences as the type material for taxonomic descriptions of prokaryotes. *Syst Appl Microbiol* 38: 217–222. doi.org/10.1016/j.syapm.2015.02.003

WHITMAN, W.B. (2016). Modest proposals to expand the type material for naming of prokaryotes. *Int J Syst Evol Microbiol* 66: 2108–2112. doi.org/10.1099/ijsem.0.000980

WHITMAN, W.B., COLEMAN, D.C. and WIEBE, W.J. (1998). Prokaryotes: the unseen majority. *PNAS* 95: 6578–6583. doi.org/10.1073/pnas.95.12.6578

WICKHAM, H. (2016). ggplot2. In *Elegant Graphics for Data Analysis*. 2th Ed. Springer-Verlag New York. doi.org/10.1007/978-3-319-24277-4

WIEGEL, J. and CANGANELLA, F. (2001). Extreme thermophiles. *Enc Life Sci* 1–12. doi.org/10.1038/npg.els.0000392

WIEGEL, J., LJUNGDAHL, L.G. and DEMAIN, A.L. (1985). The importance of thermophilic bacteria in biotechnology. *Crit Rev Microbiol* 3: 39–108. doi.org/10.3109/07388558509150780

WILLIAMS, R.A.D., SMITH, K.E., WELCH, S.G. and MICALLEF, J. (1996). *Themus oshimai* sp. nov., isolated from hot springs in Portugal, Iceland, and the Azores, and comment on the concept of a limited geographical distribution of *Themus* species. *Int J Syst Bacteriol* 46: 403–408. doi.org/10.1099/00207713-46-2-403

WILLIAMS, T.A., COX, C.J., FOSTER, P.G., SZÖLLŐSI, G.J. and EMBLEY, T.M. (2020). Phylogenomics provides robust support for a two-domains tree of life. *Nat Ecol Evol* 4: 138–147. doi.org/10.1038/s41559-019-1040-x

WIRTH, J.S. and WHITMAN, W.B. (2018). Phylogenomic analyses of a clade within the roseobacter group suggest taxonomic reassignments of species of the genera *Aestuariivita*, *Citreicella*, *Loktanella*, *Nautella*, *Pelagibaca*, *Ruegeria*, *Thalassobius*, *Thiobacimonas* and *Tropicibacter*, and the proposal of six novel genera. *Int J Syst Evol Microbiol* 68: 2393–2411. doi.org/10.1099/ijsem.0.002833

WISER, A., SCHNEIDER, L., JUNG, J. and SHUBERT, S. (2012). MALDI-TOF MS in microbiological diagnostics – identification of microorganisms and beyond (mini review). *Appl Microbiol Biotechnol* 93: 965–974. doi.org/10.1007/s00253-011-3783-4

WOESE, C.R. (1987). Bacterial evolution. *Microbiol Rev* 51: 221–271. doi.org/10.1128/mr.51.2.221-271.1987

WOESE, C.R. and FOX, G.E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *PNAS* 74: 5088–5090. doi.org/10.1073/pnas.74.11.5088

WOESE, C.R., KANDLER, O. and WHEELIS, M.L. (1990). Towards a natural system of organisms: proposal for the domains *Archaea*, *Bacteria*, and *Eucarya*. *PNAS* 87: 4576–4579. doi.org/10.1073/pnas.87.12.4576

WU, D., HUGENHOLTZ, P., MAVROMATIS, K.M., PUKALL, R., DALIN, E., IVANOVA, N.N., KUNIN, V., GOODWIN, L., WU, M., TINDALL, B.J., *et al* (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* 462: 24–31. doi.org/10.1038/nature08656

WU, M. and SCOTT, A.J. (2012). Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* 28: 1033–1034. doi.org/10.1093/bioinformatics/bts079

WU, W.L., CHEN, M.Y., TU, I.F., LIN, Y.C., ESWARKUMAR, N., YI, M., HO, M.C. and WU, S.H. (2017). The discovery of novel heat-stable keratinases from *Meiothermus taiwanensis* WR-220 and other extremophiles. *Sci Rep* 7: 4658. doi.org/10.1038/s41598-017-04723-4

XIE, C., MAO, X., HUANG, J., DING, Y., WU, J., DONG, S., KONG, L., GAO, G., LI, C.Y. and WEI, L. (2011). KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* 39: W316–W322. doi.org/10.1093/nar/gkr483

XU, Z., MASUDA, Y., ITOH, H., USHIJIMA, N., SHIRATORI, Y. and SENOO, K. (2019). *Geomonas oryzae* gen. nov., sp. nov., *Geomonas edaphica* sp. nov., *Geomonas ferrireducens* sp. nov., *Geomonas terrae* sp. nov., four ferric-reducing bacteria isolated from paddy soil, and reclassification of three species of the genus *Geobacter* as members of the genus *Geomonas* gen. nov. *Front Microbiol* 10: 2201. doi.org/10.3389/fmicb.2019.02201

YAMAGISHI, A., KAWAGUCHI, Y., HASHIMOTO, H., YANO, H., IMAI, E., KODAIRA, S., UCHIHORI, Y. and NAKAGAWA, K. (2018). Environmental data and survival data of *Deinococcus aetherius* from the exposure facility of the Japan experimental module of the international space station obtained by the Tanpopo mission. *Astrobiology* 18: 11. doi.org/10.1089/ast.2017.1751

YANG, Y., ITAHASHI, S., YOKOBORI, S. and YAMAGISHI, A. (2008). UV-resistant bacteria isolated from upper troposphere and lower stratosphere. *Biol Sci Space* 22: 18–125. doi.org/10.2187/bss.22.18

YANG, Y.L., YANG, F.L., JAO, S.C., CHEN, M.Y., TSAY, S.S., OU, W. and WU, S.H. (2006). Structural elucidation of phosphoglycolipids from strains of the bacterial thermophiles *Thermus* and *Meiothermus*. *J Lipid Res* 47: 1823–1832. doi.org/10.1194/jlr.M600034-JLR200

YARZA, P., RICHTER, M, PEPLIES, J., EUZEBY, J., AMANN, R., SCHLEIFER, K.H., LUDWIG, W., GLÖCKNER, F.O. and ROSSELLÓ-MÓRA, R. (2008). The All-Species Living Tree project: a 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst Appl Microbiol* 31: 241–250. doi.org/10.1016/j.syapm.2008.07.001

YARZA, P., YILMAZ, P., PRUESSE, E., GLÖCKNER, F.O., LUDWIG, W., SCHLEIFER, K.H., WHITMAN, W.B., EUZÉBY, J., AMANN, R and ROSSELLÓ-MÓRA, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 12: 635-645. doi.org/10.1038/nrmicro3330

YASSIN, A.A.F. (2011). Detection and characterization of mycolic acids and their use in taxonomy and classification. In *Methods in Microbiology (Taxonomy of Prokaryotes)*. Rainey, F.A. and Oren, A. (editors). Elsevier Ltd. Vol. 38, pp. 207–237. doi.org/10.1016/B978-0-12-387730-7.00010-3

YILMAZ, P., PARFREY, L.W., YARZA, P., GERKEN, J., PRUESSE, E., QUAST, C., SCHWEER, T., PEPLIES, J., LUDWIG, W. and GLÖCKNER, F.O. (2013). The SILVA and ‘‘All-species Living Tree Project (LTP)’’ taxonomic frameworks. *Nucleic Acids Res* 42: D643–D648. doi.org/10.1093/nar/gkt1209

YING, J.J., ZHANG, S.L., HUANG, C.Y., XU, L., ZHAO, Z., WU, M. and SUN, C. (2019). *Algicoccus marinus* gen. nov. sp. nov., a marine bacterium isolated from the surface of brown seaweed *Laminaria japonica*. *Arch Microbiol* 201: 943–950. doi.org/10.1007/s00203-019-01664-6

YOKOYAMA, A., SANDMANN, G., HOSHINO, T., ADACHI, K., SAKAI, M. and SHIZURI, Y. (1995). Thermozeaxanthins, new carotenoid-glycoside-esters from thermophilic eubacterium *Thermus thermophilus*. *Tetrahedron Lett* 36: 4901–4904. doi.org/10.1016/0040-4039(95)00881-C

YOON, S.H., HA, S.M., KWON, S., LIM, J., KIM, Y., SEO, H. and CHUN, J. (2017a). Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole genome assemblies. *Int J Syst Evol Microbiol* 67: 1613–1617. doi.org/10.1099/ijsem.0.001755

YOON, S.H., HA, S.M., LIM, J., KWON, S and CHUN, J. (2017b). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* 110: 1281–1286. doi.org/10.1007/s10482-017-0844-4

YU, T.T., YIN, Y.R., ZHANG, Y.G., YAO, J.C., KLENK, H.P., WANG, H.F., MING, H., ZHOU, E.M. and LI, W.J. (2014). *Meiothermus terrae* sp. nov., isolated from a geothermally heated soil sample. *Int J Syst Evol Microbiol* 64: 794–798. doi.org/10.1099/ijs.0.055392-0

ZARE, S., DERAKHSHANDEH, A., HAGHKHAH, M., NAZIRI, Z. and BROUJENI, A.M. (2019). Molecular typing of *Staphylococcus aureus* from different sources by RAPD-PCR analysis. *Heliyon* 5: e02231. doi.org/10.1016/j.heliyon.2019.e02231

ZELDOVICH, K.B., BEREZOVSKY, I.N. and SHAKHNOVICH, E.I. (2007). Protein and DNA sequence determinants of thermophilic adaptation. *PLoS Comp Biol* 3: e5. doi.org/10.1371/journal.pcbi.0030005

ZHANG, A.N., MAO, Y., WANG, Y. and ZHANG, T. (2019). Mining traits for the enrichment and isolation of not-yet-cultured populations. *Microbiome* 7: 96. doi.org/10.1186/s40168-019-0708-4

ZHANG, K., EZEMADUKA, A.N., WANG, Z., HU, H., SHI, X., LIU, C., LU, X., FU, X., CHANG, Z. and YIN, C.C. (2015). A novel mechanism for small heat shock proteins to function as molecular chaperones. *Sci Rep* 5: 8811. doi.org/10.1038/srep08811

ZHANG, X., LI, S.J., LI, J.J., LIANG, Z.Z. and ZHAO, C.Q. (2018). Novel natural products from extremophilic fungi. *Marine Drugs* 16: 194. doi.org/10.3390/md16060194

ZHANG, X.Q., ZHANG, W.J., WEI, B.P., XU, X.W., ZHU, X.F. and WU, M. (2010). *Meiothermus cateniformans* sp. nov., a slightly thermophilic species from north-eastern China. *Int J Syst Evol Microbiol* 60: 840–844. doi.org/10.1099/ijs.0.007914-0

ZHI, X.Y., ZHAO, W., LI, W.J. and ZHAO, G.P. (2012). Prokaryotic systematics in the genomics era. *Antonie van Leeuwenhoek* 101: 21–34. doi.org/10.1007/s10482-011-9667-x

ZHU, Q., MAI, U., PFEIFFER, W., JANSSEN, S., ASNICAR, F., SANDERS, J.G., BELDA-FERRE, P., AL-GHALITH, G.A., KOPYLOVA, E., MCDONALD, D., *et al*. (2019).

Phylogenomics of 10,575 genomes reveals evolutionary proximity between domains Bacteria and Archaea. *Nat Commun* 10: 5477. doi.org/10.1038/s41467-019-13443-4

ZILLIG, W., HOLZ, I. and WUNDERL, S. (1991). *Hyperthermus butylicus* gen. nov. sp. nov. a hyperthermophilic, anaerobic, peptide-fermenting, facultatively H_2S-generating archaebacterium. *Int J Syst Bacteriol* 41: 169–170. doi.org/10.1099/00207713-41-1-169

ZILLIG, W., STETTER, K.O., WUNDER, S., SCHULZ, W., PRIESS, H. and SCHOLZ, I. (1980). The *Sulfolobus*-"*Caldariella*" group: taxonomy on the basis of the structure of DNA-dependent RNA polymerases. *Arch Microbiol* 125: 259–269. doi.org/10.1007/BF00446886