



The importance of designating type material for uncultured taxa

Maria Chuvochina^{a,*}, Christian Rinke^a, Donovan H. Parks^a, Michael S. Rappé^b, Gene W. Tyson^a, Pelin Yilmaz^c, William B. Whitman^d, Philip Hugenholtz^{a,*}

^a Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, The University of Queensland, QLD 4072, Australia

^b Hawaii Institute of Marine Biology, University of Hawaii at Manoa, Kaneohe, HI, USA

^c Microbial Physiology Group, Max Planck Institute for Marine Microbiology, Bremen, Germany

^d Department of Microbiology, University of Georgia, 527 Biological Sciences Building, Athens, GA 30602-2605, USA

ARTICLE INFO

Keywords:

Prokaryotic nomenclature
Type material
Metagenome-assembled genomes
Candidatus
ICNP

ABSTRACT

Naming of uncultured Bacteria and Archaea is often inconsistent with the International Code of Nomenclature of Prokaryotes. The recent practice of proposing names for higher taxa without designation of lower ranks and nomenclature types is one of the most important inconsistencies that needs to be addressed to avoid nomenclatural instability. The Code requires names of higher taxa up to the rank of class to be derived from the type genus name, with a proposal pending to formalise this requirement for the rank of phylum. Designation of nomenclature types is crucial for providing priority to names and ensures their uniqueness and stability. However, only legitimate names proposed for axenic cultures can be used for this purpose. *Candidatus* names reserved for taxa lacking cultured representatives may be granted this right if recent proposals to use genome sequences as type material are endorsed, thereby allowing the Code to be fully applied to lineages represented by metagenome-assembled genomes (MAGs) or single amplified genomes (SAGs). Genome quality standards need to be considered to ensure unambiguous assignment of type material. Here, we illustrate the recommended practice by proposing nomenclature type material for four major uncultured prokaryotic lineages based on high-quality MAGs in accordance with the Code.

© 2018 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

Culture-independent characterization of the microbial world has prompted calls for a unified nomenclature to describe uncultured organisms based primarily on sequence data [22,24,39]. The naming of cultured Bacteria and Archaea is governed by the International Code of Nomenclature of Prokaryotes (ICNP or the Code; [32]), which currently does not officially extend to uncultured microorganisms. Many uncultured prokaryotes have been named using the *Candidatus* status to indicate putative taxa [24], but the majority of these names are inconsistent with the ICNP. Apart from linguistic errors [26], common issues include: failure to designate the nomenclatural type, naming of higher taxa without naming the type genus and species, absence of a description of the taxon, absence of an etymology for the name. In light of a pending proposal to incorporate nomenclature of uncultured prokaryotes under the Code [39], correction

of these issues is necessary for names to obtain standing in nomenclature and priority. There is some urgency to implement such nomenclature standards as the number of metagenome-assembled genomes (MAGs) is rapidly outgrowing the number of species based upon type cultures. As an example, we provide descriptions of nomenclature type material for four uncultured lineages based on high-quality MAGs that are consistent with the ICNP.

In the description of any novel taxon, the ICNP requires the designation of type material (Section 4 of the Code; [32]), which is key for providing priority to the name. The combination of priority and type material ensures the uniqueness and stability of names by permitting only the name with priority to be used to denote the taxon that includes the type material. Likewise, the name cannot be used to denote a taxon that does not contain the type material. This rule implies a simple and objective criterion for type material. The type must be sufficiently well described by the type material to distinguish it from other taxa of the same rank. Delineation of species typically includes genetic comparisons, either DNA–DNA hybridization, average nucleotide identities, average amino acid identities or similar criteria [11,22,38]. Delineation of genera and higher taxa is typically performed using phylo-

* Corresponding authors.

E-mail addresses: m.chuvochina@uq.edu.au (M. Chuvochina), p.hugenholtz@uq.edu.au (P. Hugenholtz).

genetic relationships inferred from 16S rRNA genes [15,41,42] or concatenated protein alignments [12,23,36]. The justification for a genomic sequence to serve as the type material is based upon its ability to distinguish closely related species and has been outlined elsewhere [21,38,39]. MAGs and SAGs have recently been proposed to be used for the description of uncultured taxa [22] with which we will likely be able to achieve even strain-level resolution in the near future using new techniques such as genome reconstruction based on DNA methylation profiles [6] or improving MAG assemblies with long-read sequencing platforms [10].

In the Code, lower taxonomic ranks are the types for higher ranks (Fig. 1; Section 4 of the Code). Thus, the type of a genus is a species, the type of a family and order is a genus, and the type of a class is an order. If the recommendation of Oren et al. [28] (addendum by Whitman et al. [40]) is followed, the type of a phylum will be a class. The logic of this system is simple. The types and thus the names themselves can all be traced back to a specific biological entity whose existence can be clearly demonstrated, *i.e.* a culture, a DNA sample or information such as genome sequence. Use of type material to allocate names to taxa is a fundamental process in biological nomenclature under all three codes (International Code of Nomenclature for algae, fungi, and plants; International Code of Zoological Nomenclature and International Code of Nomenclature of Prokaryotes). Even as taxonomic opinions change and new discoveries are made, the names of all taxa continue to refer to their types, thus, allowing nomenclature to be distinct from taxonomic classification. This system allows flexibility in classification while maintaining stability in nomenclature.

Largely driven by culture-independent sequence analysis, it has become common practice to designate new phyla or other higher taxa without designating lower ranks [1,3,8,14,17]. This practice is in direct conflict with the ICNP mandate that higher taxa names follow from lower taxa names that serve as their types (Rule 15 of the Code; Fig. 1), resulting in nomenclatural instability of such names. Moreover, there is an inherent ambiguity of higher taxa. For instance, while there is some consensus about what constitutes species and genera, there is less consistency in the designation of higher taxa [17], with estimates of the number of bacterial phyla ranging from ~100 to ~1350 [18,30,42]. Lastly, this practice asserts a meaning to a taxonomic rank, which contradicts one of the central tenets of the Code, *i.e.* independence of nomenclature from taxonomy (General Consideration 4). Direct consequences of ignoring the principle of hierarchy and typification would be the inability to appropriately name taxa if taxonomic opinion changes (*i.e.*, transfer, union or change in rank) leading to unconstructive debates on name authority and 'priority'. In order to have a well-defined taxonomic framework, it is crucial to ensure that the naming hierarchy is followed and that type material is provided for each taxon starting from species.

Lastly, names only gain standing in nomenclature after they have been announced in the International Journal of Systematics and Evolutionary Microbiology (IJSEM), the official journal of the ICSU (Rules 24a and 27). Although *Candidatus* taxa currently lack standing, their names should be announced, either as a publication in the IJSEM or inclusion on the "List of *Candidatus* Names", which will soon be added to the lists periodically published in the IJSEM (see Appendix A1 of the Code). Authors and other individuals wishing to have new names and/or combinations included in the Lists should send an electronic copy of the published paper to the IJSEM Editorial Office. This procedure ensures that the name becomes widely known.

Material and methods

Selection of candidate phyla for type material designations

Candidate bacterial and archaeal phyla lacking designated type material were selected based on the classification of the Genome Taxonomy Database (GTDB), release 2-RS83, which only includes genomes available prior to July 2017 in National Center for Biotechnology Information (NCBI) RefSeq/GenBank release 83, passing GTDB genome quality filters (see Ref. [31]). Genome-based phylogenetic reconstructions on which the classifications are based are available from the GTDB website under the download option (gtdb.ecogenomic.org/downloads).

Selection of MAG-based type material

Once candidate phyla were selected, metadata for all representative genomes was obtained from a metadata file available from the GTDB download page. This included assembly statistics, genome quality (completeness and contamination estimates) and properties (*e.g.*, GC content, estimated genome size, number of contigs). Number of tRNAs and the corresponding encoded amino acids were predicted using tRNA scan SE-1.23 using the domain-specific settings for Bacteria and Archaea. Presence and length of 16S, 23S and 5S rRNAs were verified using HMMER and domain-specific SSU/LSU HMM models as implemented in the 'ssu-finder' method of CheckM [29]. MAGs were considered as candidates for type material for a given lineage based on the following quality criteria: i) high-quality draft MIMAG standards for genome reporting (>90% completion; <5% contamination; multiple fragments where gaps span repetitive regions; presence of the 23S, 16S, and 5S rRNA genes and tRNAs encoding for at least 18 out of 20 standard amino acids; Bowers et al. [7]); ii) full or nearly full-length 16S rRNA gene with a threshold of 1200 bp for both Bacteria and Archaea [33]; iii) a draft assembly consisting of 100 contigs or less with less than 10 ambiguous bases. A quality score defined as completeness – 4 × contamination was used to rank the MAGs, and the one with the highest ranking score that meets the above selection criteria, was chosen to represent the given type. In the event of equal or similar quality scores for multiple MAGs, we chose the one with the most complete 16S rRNA gene sequence.

Results and discussion

Typifying the uncultured majority

The International Code of Nomenclature of Prokaryotes (ICNP, [32]), Rule 15, states: "For each named taxon of the various taxonomic categories (...), there shall be designated a nomenclature type." Rule 17 then adds: "The type determines the application of the name of a taxon if the taxon is subsequently divided or united with another taxon." Therefore, the official prokaryotic nomenclature recognises nomenclature types as a means of naming organisms. The naming hierarchy is then based on the nomenclature types for each rank ascending from the lower to higher categories (Rule 15). Thus, the names of the highest categories up to the rank of class are based upon the designation of type genus, and derived from the stem of its name by addition of the corresponding rank suffixes (Rule 8 and 9). To date, all described uncultured prokaryotes have been named without designation of formal nomenclature types. Some studies aimed to designate type species for higher taxa [2,35] but many uncultured lineages are only named at the phylum-level [1,3,8,14,17]. This situation has arisen because neither the rank of phylum nor uncultured prokaryotes are recognised in the official nomenclature [32]. A recent call

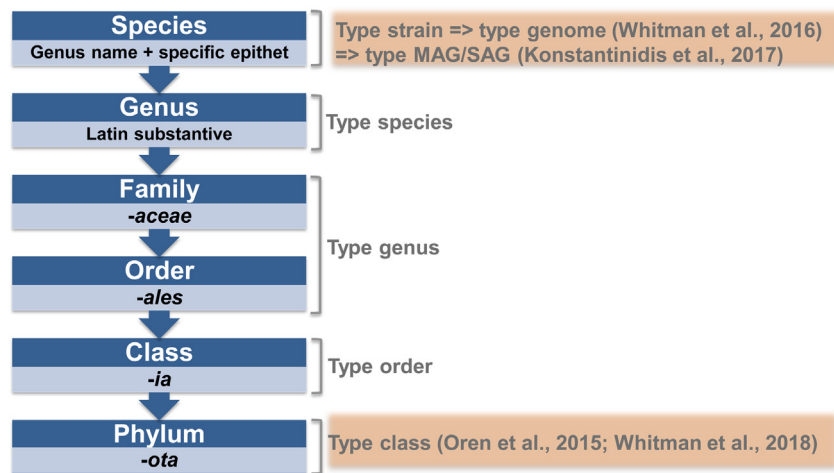


Fig. 1. Naming formation and hierarchy of standard taxonomic categories based on their nomenclature types (modified after Parker et al. [32]). Recent proposals to alter the Code are highlighted in orange.

to formalise the rank of phylum under the Code provision ([28]; addendum by Whitman et al. [40]) will mean that phylum names will be formed by the addition of the suffix *-ota* to the name of one of the subordinate classes. As the name of the latter is based on the type genus of one of the contained orders, one will only need to name a genus to provide the basis for all higher taxa names. Currently, this is not possible for uncultured prokaryotes because only validly published names can serve as the nomenclature types as stated in Rule 20a. However, the recent ability to both genotypically and phenotypically characterize uncultured prokaryotes at the species level has led to the proposal to expand type material for naming prokaryotes to genomic sequences, including those derived from metagenomic or single cell datasets [22,39]. In this instance, official nomenclature practice could be applied to name uncultured microorganisms that could form the basis of a parallel nomenclature system or become part of a unified Code [22,39]. We support the latter proposal as a unified nomenclature will ultimately be the most beneficial outcome for taxonomic classification and communication of results (see also [27]). An important consideration for proposing genome-based type material is the quality of the assembled genome sequence, which must reliably discriminate a new taxon from others, as is the case for cultured isolates. Below we discuss quality standards of genomes to be used as type material and propose type species for four candidate phyla based on high quality MAGs.

Genome type material quality considerations

Quality criteria for genomes that will serve as type material are important for taxonomy and nomenclature as the designated type must unambiguously identify the taxon (Section 4 of the Code; [11,39]). Low quality type genomes, e.g. highly incomplete, chimeric assemblies, or missing marker genes, could lead to misclassifications and unnecessary creation of names, thus contradicting Principle 1 of the Code. A number of publications have proposed genome quality standards for isolate genomes [9,11,25] and genomes from uncultured microorganisms [7], including one of the two recent proposals specific to their use as type material for description of uncultured taxa [22]. The 'high-quality draft' MIMAG/SAG (Minimum Information about a metagenome-assembled genome/single amplified genome) standard uses a combination of estimated completeness (>90%) and contamination (<5%) with nearly complete complements of rRNAs (23S, 16S, 5S) and tRNAs (≥ 18) as defining criteria [7], which we suggest should serve as a benchmark for MAG type material with the following

caveats. Firstly, completeness estimates should be considered in the context of genome sequence evolution. In particular the completeness of small streamlined genomes such as those belonging to parasitic or symbiotic microorganisms can be underestimated due to *bona fide* absence of marker genes [2,5,29]. In these cases, the number of contigs may be a better indication of whether a given genome is suitable as type material, with priority given to genomes with fewer contigs. Secondly, given the importance of a complete or nearly complete 16S rRNA gene for classifying microorganisms [11,22], care should be taken to confirm that this gene belongs to the MAG, especially if it occurs in isolation on a small contig. Thirdly, although the presence of tRNAs encoding at least 18 of the 20 standard amino acids is desirable, the correlation between the number of tRNAs and genome completeness is weak, thus, reflecting its poor suitability for estimating genome completeness [30]. Furthermore, domain-specific features such as presence of fragmented or split tRNAs in Archaea [16] can result in underestimation of the true number of encoded tRNAs. Therefore, we suggest that, while this criterion should still be reported, it should not be mandatory when considering a MAG for use as type material.

Case studies of genome-based type material for uncultured prokaryotes

Major bacterial and archaeal lineages currently lacking cultured representatives were reviewed to select examples suitable for proposal of type material based on MAGs. For delineation and annotation of higher and lower ranks including species and genera, we used a recently proposed standardised genome-based taxonomy (GTDB; [31]). Four MAGs representing two archaeal and two bacterial phyla were identified as suitable candidates for type material as they conform to the MIMAG criteria for high-quality draft as outlined in Bowers et al. [7] with the caveats discussed above. We propose four species based on MAG type material (Table 1). The corresponding quality metrics of each MAG are summarised in Supplementary Table 1. Below we describe how the essential rules used in official nomenclature practice can be applied to uncultured prokaryotes through the proposed use of MAG type material from simple to more complex case studies.

Case 1: proposal of *Candidatus Binatus soli*

Representatives of several novel bacterial and archaeal phyla have been recovered recently from metagenomic data deposited in public repositories [30]. Among those, the first genomic representatives of candidate bacterial phylum UBP10 were reconstructed from

Table 1
List of proposed *Candidatus* names and their corresponding types based on metagenome-assembled genomes (MAGs).

Proposed type ^a : MAG accession number (NCBI)	Proposed name	Suggested type class and phylum name ^b	Current name of the lineage	Original study
Bacteria				
GCA.002479255.1	<i>Binatus soli</i>	<i>Binatia</i> ; <i>Binatota</i>	UBP10	Parks et al. [30]
GCA.002011615.1	<i>Hydrothermus pacificus</i>	<i>Hydrothermia</i> ; <i>Hydrothermota</i>	<i>Hydrothermae</i>	Jungbluth et al. [19]
Archaea				
GCA.002011125.1	<i>Hydrothermarchaeum profundum</i>	<i>Hydrothermarchaeia</i> ; <i>Hydrothermarchaeota</i>	<i>Hydrothermarchaeota</i>	Jungbluth et al. [19]
GCA.001515205.2	<i>Hadarchaeum yellowstonense</i>	<i>Hadarchaeia</i> ; <i>Hadarchaeota</i>	<i>Hadesarchaea</i>	Baker et al. [4]

^a Based on the pending proposal by Whitman [40].

^b Based on the pending proposal by Oren et al. [28], addendum Whitman et al. [40].

a soil metagenome. As of today, UB10 appears to be represented by a single class as annotated in GTDB, including MAGs identified in other studies. Here we propose *Candidatus Binatus soli* based on the high-quality MAG, UBA7539 (GCA_002479255.1) to represent the first named species of the UB10 phylum. This species can serve as the type for higher-ranking taxa including phylum if the Oren et al. [28] proposal is accepted, which would be named: *Binataceae* (family), *Binatales* (order), *Binatia* (class), and *Binatota* (phylum).

Case 2: proposal of *Candidatus Hydrothermarchaeum profundum*

Candidate archaeal phylum *Hydrothermarchaeota* was recently proposed to accommodate the first genomic representatives of Marine Benthic Group E (MBG-E) [19], which was previously identified via 16S rRNA gene sequences [37]. Although three MAGs representing this phylum were recovered, none were proposed as *Candidatus* species or genera. Of these, JdFR-18 (GCA_002011125.1) is the highest quality MAG meeting the requirements described above. We therefore propose JdFR-18 as the type material of a new species *Candidatus Hydrothermarchaeum profundum*. The generic name was formed from the stem of the originally proposed higher taxon name as the latter is not validly published and could only be valid under the Code with the designation of its nomenclature type. However, with the proposal of *Ca. H. profundum*, stability of the phylum name can be achieved as follows: *Hydrothermarchaeaceae* (family), *Hydrothermarchaeales* (order), *Hydrothermarchaeia* (class), and *Hydrothermarchaeota* (phylum). Note that the originally proposed phylum name 'Hydrothermarchaeota', will need to be re-proposed as the phylum containing the genus *Ca. Hydrothermarchaeum* to be consistent with the Code.

Case 3: proposal of *Candidatus Hadarchaeum yellowstonense*

Originally proposed as a class-level lineage within the archaeal phylum Euryarchaeota [4] and later as a member of the superclass Stygia [1], the class Hadesarchaea (formerly known as SAGMEG [South-African Gold Mine Miscellaneous Euryarchaeal Group]) is currently considered a separate phylum based on taxonomic rank normalisation [31]. No named species or genera have been proposed for the Hadesarchaea, therefore, this class cannot be validated according to the Code (assuming unified rules or recent proposals). Here we propose *Candidatus Hadarchaeum yellowstonense* based on the highest quality MAG available, YNP_45 (GCA_001515205.2) as the first named species within the lineage that can serve as the type for naming higher taxa containing this species: *Hadarchaeaceae* (family), *Hadarchaeales* (order), *Hadarchaeia* (class), and *Hadarchaeota* (phylum).

Case 4: proposal of *Candidatus Hydrothermus pacificus*

Candidate bacterial phylum *Hydrothermae* was recently proposed to accommodate the first genomic representatives of the EM3 lineage [19], which was previously known only from 16S rRNA sequences [34]. According to taxonomic rank normalization [31], the *Hydrothermae* are a class within a larger monophyletic unit comprising candidate phylum WOR-3 [4], which itself has been

proposed to be named *Candidatus Stahlbacteria* after Dr. David Stahl, an environmental microbiologist and early proponent of 16S rRNA-based microbial ecology [13]. However, both *Hydrothermae* and *Candidatus Stahlbacteria* are examples of high level lineages that lack designation of lower ranks and nomenclature types. Here we propose *Candidatus Hydrothermus pacificus* as the type and first named species of the group based on the highest quality MAG currently available, JdFR-72 (GCA_002011615.1) to conserve the link to the name *Hydrothermae* [19]. The *Hydrothermae* (EM3) lineage comprising this species could then use the root of the genus name for higher ranks: *Hydrothermaceae* (family), *Hydrothermales* (order), *Hydrothermia* (class) and *Hydrothermota* (phylum). Similarly, if a type species is designated based on a high quality MAG in the *Candidatus Stahlbacteria* (WOR-3) lineage using the original etymology (e.g. named after David Stahl as genus *Stahlia* or *Stahlbacter*), the name and associated recognition can be salvaged without leading to nomenclatural discord as ensured by the use of type material. If both lineages have valid classes, either can then be used to form the phylum name.

Concluding remarks

A substantial proportion of bacterial and archaeal diversity has been described solely using culture-independent techniques, including recovery of near complete genomes from metagenomic data [1,3,8,14,43]. However communication of this diversity is often complicated by uncultured prokaryotes not being recognised in the official nomenclature. As a result, a number of inconsistencies exist between the way cultured and uncultured taxa are named. Here we demonstrate how some of these common inconsistencies, most notably the failure to designate nomenclature types for higher taxa, can be addressed and provide examples for two archaeal and two bacterial phyla. This is a necessary first step for candidate taxa names to obtain standing in nomenclature and priority if the nomenclatures of cultured and uncultured prokaryotes are unified following pending proposals [28,39]. Names proposed in this manner can be considered for inclusion in future validation lists. We strongly encourage authors to name at least one species and genus within any newly delineated uncultured lineage based on MAGs, with due consideration of genome quality criteria. Since the phylogenetic position of taxa including newly discovered lineages may be uncertain or disputed [31], designating type material ensures name stability with changing taxonomic opinion.

One important further consideration is that improved genomic representatives may become available for already defined type material. Recovery of improved MAG or SAG representatives of type species is highly likely given that rapid advances in sequencing and bioinformatic techniques continue to be made [6,10,20]. Improved sequences may be helpful for making taxonomic decisions within the group. However, they should have no influence on designation of the type material. To do so only leads to nomenclatural instability by making priority conditional rather than absolute.

Table 2Protologues for *Candidatus Hydrothermarchaeum profundum*, *Ca. Hadarchaeum yellowstonense*, *Ca. Hydrothermus pacificus* and *Ca. Binatus soli*.

Taxonumber	CA00024	CA00025	CA00027	CA00028
Species name	<i>Candidatus Hydrothermarchaeum profundum</i>	<i>Candidatus Hadarchaeum yellowstonense</i>	<i>Candidatus Hydrothermus pacificus</i>	<i>Candidatus Binatus soli</i>
Genus name	<i>Candidatus Hydrothermarchaeum profundum</i>	<i>Candidatus Hadarchaeum yellowstonense</i>	<i>Candidatus Hydrothermus pacificus</i>	<i>Candidatus Binatus soli</i>
Specific epithet				
Genus etymology	Hy.dro.therm.ar.chae'um. Gr. n. <i>hydōr</i> water; Gr. adj. <i>thermos</i> hot; N.L. neut. n. <i>archaeum</i> archaeon from Gr. adj. <i>archaios</i> –ê –on ancient; N.L. neut. n. <i>Hydrothermarchaeum</i> an archaeon from a hydrothermal environment.	Had.ar.chae'um. N.L. neut. n. <i>archaeum</i> archaeon from Gr. adj. <i>archaios</i> –ê –on ancient; N.L. neut. n. <i>Hadarchaeum</i> the archaeon named after Hades (gen. <i>Hadou</i>) the Greek god of the underworld, referring to the origin of the organism from the subsurface biosphere.	Hy.dro.ther'mus. Gr. n. <i>hydōr</i> water; Gr. adj. <i>thermos</i> hot; N.L. masc. n. <i>Hydrothermus</i> an organism living in hot water.	Bi.na'tus. L. adv. num. <i>bis</i> twice; L. part. adj. <i>natus</i> born, made; N.L. masc. n. <i>Binatus</i> born-twice, referring to the discovery of the organism from the reprocessing (second study) of the originally studied metagenomic data.
Type species of the genus	<i>Candidatus Hydrothermarchaeum profundum</i>	<i>Candidatus Hadarchaeum yellowstonense</i>	<i>Candidatus Hydrothermus pacificus</i>	<i>Candidatus Binatus soli</i>
Taxonumber of the type species	CA00024	CA00025	CA00027	CA00028
Genus status	gen. nov.	gen. nov.	gen. nov.	gen. nov.
Species etymology	pro.fun'di. L. gen. n. <i>profundi</i> of/from the depth of the sea.	ye.low.ston.en'se. N.L. neut. adj. <i>yellowstonense</i> pertaining to the Yellowstone National Park, the place of sampling from where organism was found.	pa.ci'fi.cus. L. masc. adj. <i>pacificus</i> pertaining to the Pacific Ocean.	so'li. L. gen. n. <i>soli</i> of the soil.
Species status	sp. nov.	sp. nov.	sp. nov.	sp. nov.
Designation of the type MAG	GCA.002011125.1	GCA.001515205.2	GCA.002011615.1	GCA.002479255.1
MAG/SAG accession number	GCA.002011125.1	GCA.001515205.2	GCA.002011615.1	GCA.002479255.1
Genome status	Draft	Draft	Draft	Draft
Genome size	2,062.13	1,231.13	2,059.80	3,671.36
GC mol %	39.14	54.84	34.75	61.34
Country of origin	USA	USA	USA	USA
Region of origin	Pacific Ocean	Wyoming	Pacific Ocean	Minnesota
Source of sample	Hydrothermal fluid	Hydrothermal sediment	Hydrothermal fluid	Soil
Sampling date	12/07/11	1/01/14	12/07/11	1/01/12
Geographic location	Juan de Fuca Ridge	Yellowstone National Park	Juan de Fuca Ridge	Marcell Experimental Forest
Latitude	47°76'N	44°N	47°76'N	N/A
Longitude	127°76'W	110°W	127°76'W	N/A
Depth	N/A	0–0.1	N/A	N/A
Altitude	N/A	2232 m	N/A	N/A
Temperature of the sample [in celsius degrees]	N/A	68.8	N/A	N/A
pH of the sample	N/A	8.6	N/A	N/A
Relationship to O ₂	Anaerobe	Anaerobe	Anaerobe	N/A
Energy metabolism	Chemolithotroph	Mixotroph	Chemoorganotroph	N/A
Assembly	1 sample	1 sample	1 sample	1 sample
Sequencing technology	Illumina HiSeq 2000	Illumina HiSeq	Illumina HiSeq 2000	Illumina HiSeq 2000
Binning software used	CONCOCT	ESOM	CONCOCT	MetaBAT v.0.26.3
Assembly software used	SOAPdenovo v. 1.05; Newbler v. 2.7; minimus2 (AMOS) v. 3.1.0	IDBA-UD v. 1.0.9	SOAPdenovo v. 1.05; Newbler v. 2.7; minimus2 (AMOS) v. 3.1.0	CLC de novo assembler v4.4.1
Habitat	Hydrothermal fluid	Hot spring	Hydrothermal fluid	Peatland
Miscellaneous, extraordinary features relevant for the description	The organism was formerly identified as JdFR-18 by Jungbluth et al. (2017) Sci. Data 4, 170037. doi: 10.1038/sdata.2017.37. An annotation of the genome is provided in the same original study. Metabolic capabilities inferred from the genome reveal a thermophilic carboxydrotroph capable of chemotaxis and motility. The organism appears to be capable of anaerobic respiration with sulfate serving as a terminal electron acceptor.	A detailed metabolic description of the proposed <i>Ca. H. yellowstonense</i> , formerly identified as YNP-45, is given in the original work by Baker et al. (2016) Nat. Microbiol. 1, 16002. doi: 10.1038/NMICROBIOL.2016.2. The reduced genome size and previously inferred gene content (821) suggests that the genome has undergone streamlining. The inferred metabolic capabilities indicates oxidation of carbon monoxide, which may be coupled to H ₂ O or nitrite reduction to ammonia. Also inferred to contain a variety of central carbon metabolic (C ₁ pathway) genes found in methanogens, which may be used for carbon fixation. The organism is inferred to be thermophilic.	An annotation of the genome for <i>Ca. H. pacificus</i> , formerly identified as JdFR-72, is given in the original work by Jungbluth et al. (2017) Sci. Data 4, 170037. doi: 10.1038/sdata.2017.37. Despite a low GC mol%, genomic and environmental evidence reveals organism to be a thermophilic, anaerobic chemoorganotroph capable of chemotaxis in the deep seafloor.	The genome sequence was obtained by assembly and binning of metagenomic sequences as described previously Parks et al. (2017) Nature Microbiol. 2, 1533–1542. Doi: 10.1038/s41564-017-0012-7.

N/A not available.

Improvements are a taxonomic and not a nomenclatural issue. Moreover, if sequences are shown to possess errors that preclude their ability to truthfully represent a biological entity, procedures within the Code already allow designation of a neotype. Likewise, if a representative of a taxon is subsequently cultured, Rule 18f allows the culture to be designated as the type. However, this is quite different from making incremental changes in the type sequence, which ultimately may change the original intention and is, in fact, unnecessary. We also suggest that genomes designated as types should be recognised by public repositories such as those within The International Nucleotide Sequence Database Collaboration (INSDC) and include a designation such as ‘type genome’ to facilitate comparison to other taxa.

The formal proposals of the new *Candidatus* genera “*Candidatus* Hydrothermarchaeum”, “*Candidatus* Hadarchaeum”, “*Candidatus* Hydrothermus”, “*Candidatus* Binatus” and new species “*Candidatus* Hydrothermarchaeum profundum”, “*Candidatus* Hadarchaeum yellowstoneense”, “*Candidatus* Hydrothermus pacificus” and “*Candidatus* Binatus soli” are given in Table 2 with the corresponding Taxonumbers; CA00024, CA00025, CA00027 and CA00028.

Acknowledgments

We thank Pierre Chaumeil for his help with metadata preparation and general maintenance of the GTDB website and associated data preparation. We also thank Aharon Oren, Bernard Schink and Stefano Ventura for assistance with etymology of the proposed names, and Aharon Oren for initiating regular announcements of *Candidatus* taxa in IJSEM with the provisional name “List of *Candidatus* names”. This study was supported by an Australian Research Council Laureate Fellowship (FL150100038) awarded to PH.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.syapm.2018.07.003>.

References

- [1] Adam, P.S., Borrel, G., Brochier-Armanet, C., Gribaldo, S. (2017) The growing tree of Archaea: new perspectives on their diversity, evolution and ecology. *ISME J.* 11, 2407–2425.
- [2] Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K.L., Tyson, G.W., Nielsen, P.H. (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat. Biotechnol.* 31, 533–538.
- [3] Anantharaman, K., Brown, C.T., Hug, L.A., Sharon, I., Castelle, C.J., Probst, A.J., Thomas, B.C., Singh, A., Wilkins, M.J., Karaoz, U., Brodie, E.L. (2016) Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat. Commun.* 7, 13219.
- [4] Baker, B.J., Saw, J.H., Lind, A.E., Lazar, C.S., Hinrichs, K.U., Teske, A.P., Ettema, T.J. (2016) Genomic inference of the metabolism of cosmopolitan subsurface Archaea, Hadesarchaea. *Nat. Microbiol.* 1, 16002.
- [5] Baker, B.J., Comolli, L.R., Dick, G.J., Hauser, L.J., Hyatt, D., Dill, B.D., Land, M.L., VerBerkmoes, N.C., Hettich, R.L., Banfield, J.F. (2010) Enigmatic, ultrasmall, uncultivated Archaea. *Proc. Natl. Acad. Sci. U. S. A.* 107, 8806–8811.
- [6] Beaulaurier, J., Zhu, S., Deikus, G., Mogno, I., Zhang, X.S., Davis-Richardson, A., Canepa, R., Triplett, E.W., Faith, J.J., Sebra, R., Schadt, E.E. (2018) Metagenomic binning and association of plasmids with bacterial host genomes using DNA methylation. *Nat. Biotechnol.* 36, 61–69.
- [7] Bowers, R.M., Kyrpides, N.C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T.B.K., Schulz, F., Jarett, J., Rivers, A.R., Eloe-Fadrosh, E.A., Tringe, S.G. (2017) Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* 35, 725–731.
- [8] Brown, C.T., Hug, L.A., Thomas, B.C., Sharon, I., Castelle, C.J., Singh, A., Wilkins, M.J., Wrighton, K.C., Williams, K.H., Banfield, J.F. (2015) Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* 523, 208–211.
- [9] Chain, P.S.G., Grafham, D.V., Fulton, R.S., Fitzgerald, M.G., Hostetler, J., Muzny, D., Ali, J., Birren, B., Bruce, D.C., Buhay, C., Cole, J.R. (2009) Genome project standards in a new era of sequencing. *Science* 326, 236–237.
- [10] Chin, C.S., Alexander, D.H., Marks, P., Klammer, A.A., Drake, J., Heiner, C., Clum, A., Copeland, A., Huddleston, J., Eichler, E.E., Turner, S.W. (2013) Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* 10, 563–569.
- [11] Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahall, D.R., da Costa, M.S., Rooney, A.P., Yi, H., Xu, X.W., De Meyer, S., Trujillo, M.E. (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 68, 461–466.
- [12] Ciccarelli, F.D., Doerks, T., Von Mering, C., Creevey, C.J., Snel, B., Bork, P. (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science* 311, 1283–1287.
- [13] Dombrowski, N., Seitz, K.W., Teske, A.P., Baker, B.J. (2017) Genomic insights into potential interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal sediments. *Microbiome* 5, 106.
- [14] Eme, L., Spang, A., Lombard, J., Stairs, C.W., Ettema, T.J. (2017) Archaea and the origin of eukaryotes. *Nat. Rev. Microbiol.* 15, 711–723.
- [15] Fox, G.E., Pechman, K.R., Woese, C.R. (1977) Comparative cataloging of 16S ribosomal ribonucleic acid: molecular approach to prokaryotic systematics. *Int. J. Syst. Evol. Microbiol.* 27, 44–57.
- [16] Fujishima, K., Sugahara, J., Kikuta, K., Hirano, R., Sato, A., Tomita, M., Kanai, A. (2009) Tri-split tRNA is a transfer RNA made from 3 transcripts that provides insight into the evolution of fragmented tRNAs in archaea. *Proc. Natl. Acad. Sci. U. S. A.* 106, 2683–2687.
- [17] Gribaldo, S., Brochier-Armanet, C. (2012) Time for order in microbial systematics. *Trends Microbiol.* 20, 209–210.
- [18] 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., Suzuki, Y. (2016) A new view of the tree of life. *Nat. Microbiol.* 1, 16048.
- [19] Jungbluth, S.P., Amend, J.P., Rappé, M.S. (2017) Metagenome sequencing and 98 microbial genomes from Juan de Fuca Ridge flank subsurface fluids. *Sci. Data* 4, 170037.
- [20] Kogawa, M., Hosokawa, M., Nishikawa, Y., Mori, K., Takeyama, H. (2018) Obtaining high-quality draft genomes from uncultured microbes by cleaning and co-assembly of single-cell amplified genomes. *Sci. Rep.* 8, 2059.
- [21] Konstantinidis, K.T., Rosselló-Móra, R. (2015) Classifying the uncultivated microbial majority: a place for metagenomic data in the *Candidatus* proposal. *Syst. Appl. Microbiol.* 38, 223–230.
- [22] Konstantinidis, K.T., Rosselló-Móra, R., Amann, R. (2017) Uncultivated microbes in need of their own taxonomy. *ISME J.* 11, 2399–2406.
- [23] Lang, J.M., Darling, A.E., Eisen, J.A. (2013) Phylogeny of bacterial and archaeal genomes using conserved genes: supertrees and supermatrices. *PLoS One* 8, e62510.
- [24] Murray, R.G.E., Stackebrandt, E. (1995) Taxonomic note: implementation of the provisional status *Candidatus* for incompletely described prokaryotes. *Int. J. Syst. Evol. Microbiol.* 45, 186–187.
- [25] Nelson, K.E., Weinstock, G.M., Highlander, S.K., Worley, K.C., Creasy, H.H., Wortman, J.R., Rusch, D.B., Mitreva, M., Sodergren, E., Chinwalla, A.T., et al. (2010) A catalog of reference genomes from the human microbiome. *Science* 328, 994–999.
- [26] Oren, A. (2017) A plea for linguistic accuracy— also for *Candidatus* taxa. *Int. J. Syst. Evol. Microbiol.* 67, 1085–1094.
- [27] Oren, A., Garrity, G.M. (2017) Uncultivated microbes – in need of their own nomenclature? *ISME J.* 12, 309–311.
- [28] Oren, A., da Costa, M.S., Garrity, G.M., Rainey, F.A., Rosselló-Móra, R., Schink, B., Sutcliffe, I., Trujillo, M.E., Whitman, W.B. (2015) Proposal to include the rank of phylum in the international code of nomenclature of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 65, 4284–4287.
- [29] Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055.
- [30] Parks, D.H., Rinke, C., Chuvochina, M., Chaumeil, P.A., Woodcroft, B.J., Evans, P.N., Hugenholtz, P., Tyson, G.W. (2017) Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* 2, 1533–1542.
- [31] Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.A., Hugenholtz, P. (2018) A proposal for a standardized bacterial taxonomy based on genome phylogeny. *bioRxiv*, 256800 <https://doi.org/10.1101/256800>.
- [32] Parker, C.T., Tindall, B.J., Garrity, G.M. (2016) International Code of Nomenclature of Prokaryotes and Prokaryotic Code. *Int. J. Syst. Evol. Microbiol.*, <http://dx.doi.org/10.1099/ijsem.0.000778>, In press.
- [33] Quast, C., Priebe, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O. (2012) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, D590–D596.
- [34] Reysenbach, A.L., Wickham, G.S., Pace, N.R. (1994) Phylogenetic analysis of the hyperthermophilic pink filament community in Octopus Spring, Yellowstone National Park. *Appl. Environ. Microbiol.* 60, 2113–2119.
- [35] Rinke, C., Schwiertek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., Darling, A., Malfatti, S., Swan, B.K., Gies, E.A., Dodsworth, J.A. (2013) Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499, 431–437.
- [36] Thiery, T., Landan, G., Martin, W.F. (2014) Concatenated alignments and the case of the disappearing tree. *BMC Evol. Biol.* 14, 266.
- [37] Verrini, C., Jannasch, H.W., MacGregor, B.J., Stahl, D.A., Reysenbach, A.L. (1999) Population structure and phylogenetic characterization of marine benthic archaea in deep-sea sediments. *Appl. Environ. Microbiol.* 65, 4375–4384.
- [38] Whitman, W.B. (2015) Genome sequences as the type material for taxonomic descriptions of prokaryotes. *Syst. Appl. Microbiol.* 38, 217–222.

- [39] Whitman, W.B. (2016) Modest proposals to expand the type material for naming of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 66, 2108–2112.
- [40] Whitman, W.B., Oren, A., Chuvochina, M., da Costa, M.S., Garrity, G.M., Rainey, F.A., Rossello-Mora, R., Schink, B., Sutcliffe, I., Trujillo, M.E., Ventura, S. (2018) Proposal of the suffix *-ota* to denote phyla: Addendum to 'Proposal to include the rank of phylum in the International Code of Nomenclature of Prokaryotes'. *Int. J. Syst. Evol. Microbiol.* 68, 967–969.
- [41] Woese, C.R. (1987) Bacterial evolution. *Microbiol. Rev.* 51, 221.
- [42] Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., Euzéby, J., Amann, R., 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.
- [43] Yeoh, Y.K., Sekiguchi, Y., Parks, D.H., Hugenholtz, P. (2015) Comparative genomics of candidate phylum TM6 suggests that parasitism is widespread and ancestral in this lineage. *Mol. Biol. Evol.* 33, 915–927.