



## Roadmap for naming uncultivated Archaea and Bacteria

Alison Murray, John Freudenstein, Simonetta Gribaldo, Roland Hatzenpichler, Philip Hugenholtz, Peter Kämpfer, Konstantinos Konstantinidis, Christopher Lane, R. Thane Papke, Donovan Parks, et al.

### ► To cite this version:

Alison Murray, John Freudenstein, Simonetta Gribaldo, Roland Hatzenpichler, Philip Hugenholtz, et al.. Roadmap for naming uncultivated Archaea and Bacteria. *Nature Microbiology*, 2020, 5 (8), pp.987-994. 10.1038/s41564-020-0733-x . hal-02863761v1

**HAL Id: hal-02863761**

**<https://hal.science/hal-02863761v1>**

Submitted on 18 Sep 2020 (v1), last revised 16 Feb 2024 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



OPEN

# Roadmap for naming uncultivated Archaea and Bacteria

Alison E. Murray<sup>1</sup>✉, John Freudenstein<sup>2</sup>, Simonetta Gribaldo<sup>3</sup>, Roland Hatzepichler<sup>4</sup>, Philip Hugenholtz<sup>5</sup>, Peter Kämpfer<sup>6</sup>, Konstantinos T. Konstantinidis<sup>7</sup>, Christopher E. Lane<sup>8</sup>, R. Thane Papke<sup>9</sup>, Donovan H. Parks<sup>5</sup>, Ramon Rossello-Mora<sup>10</sup>, Matthew B. Stott<sup>11</sup>, Iain C. Sutcliffe<sup>12</sup>, J. Cameron Thrash<sup>13</sup>, Stephanus N. Venter<sup>14</sup>, William B. Whitman<sup>15</sup>, Silvia G. Acinas<sup>16</sup>, Rudolf I. Amann<sup>17</sup>, Karthik Anantharaman<sup>18</sup>, Jean Armengaud<sup>19</sup>, Brett J. Baker<sup>20</sup>, Roman A. Barco<sup>21</sup>, Helge B. Bode<sup>22,23</sup>, Eric S. Boyd<sup>24</sup>, Carrie L. Brady<sup>25</sup>, Paul Carini<sup>26</sup>, Patrick S. G. Chain<sup>27</sup>, Daniel R. Colman<sup>24</sup>, Kristen M. DeAngelis<sup>28</sup>, Maria Asuncion de los Rios<sup>29</sup>, Paulina Estrada-de los Santos<sup>30</sup>, Christopher A. Dunlap<sup>31</sup>, Jonathan A. Eisen<sup>32</sup>, David Emerson<sup>33</sup>, Thijs J. G. Ettema<sup>34</sup>, Damien Eveillard<sup>35</sup>, Peter R. Girguis<sup>36</sup>, Ute Hentschel<sup>37</sup>, James T. Hollibaugh<sup>38</sup>, Laura A. Hug<sup>39</sup>, William P. Inskeep<sup>40</sup>, Elena P. Ivanova<sup>41</sup>, Hans-Peter Klenk<sup>42</sup>, Wen-Jun Li<sup>43</sup>, Karen G. Lloyd<sup>44</sup>, Frank E. Löffler<sup>45,46</sup>, Thulani P. Makhalanyane<sup>47</sup>, Duane P. Moser<sup>48</sup>, Takuro Nunoura<sup>49</sup>, Marike Palmer<sup>50</sup>, Victor Parro<sup>51</sup>, Carlos Pedrós-Alió<sup>52</sup>, Alexander J. Probst<sup>53</sup>, Theo H. M. Smits<sup>54</sup>, Andrew D. Steen<sup>55</sup>, Emma T. Steenkamp<sup>56</sup>, Anja Spang<sup>57,58</sup>, Frank J. Stewart<sup>24</sup>, James M. Tiedje<sup>59</sup>, Peter Vandamme<sup>60</sup>, Michael Wagner<sup>61</sup>, Feng-Ping Wang<sup>62</sup>, Brian P. Hedlund<sup>50</sup>✉ and Anna-Louise Reysenbach<sup>63</sup>✉

**The assembly of single-amplified genomes (SAGs) and metagenome-assembled genomes (MAGs) has led to a surge in genome-based discoveries of members affiliated with Archaea and Bacteria, bringing with it a need to develop guidelines for nomenclature of uncultivated microorganisms. The International Code of Nomenclature of Prokaryotes (ICNP) only recognizes cultures as 'type material', thereby preventing the naming of uncultivated organisms. In this Consensus Statement, we propose two potential paths to solve this nomenclatural conundrum. One option is the adoption of previously proposed modifications to the ICNP to recognize DNA sequences as acceptable type material; the other option creates a nomenclatural code for uncultivated Archaea and Bacteria that could eventually be merged with the ICNP in the future. Regardless of the path taken, we believe that action is needed now within the scientific community to develop consistent rules for nomenclature of uncultivated taxa in order to provide clarity and stability, and to effectively communicate microbial diversity.**

As new environments are explored and technological innovations improve tools for the characterization of microbial biodiversity, insights into bacterial and archaeal diversity are continually emerging<sup>1,2</sup>, including improved understanding of physiological capacity, ecology and evolution of organisms across the tree of life. These advances are based on both cultivation strategies<sup>3,4</sup> and cultivation-independent methods that directly access diversity using single-cell<sup>5,6</sup> or metagenomic sequencing<sup>7–9</sup> (Box 1). Though our ability to culture fastidious microorganisms is improving, success seems to vary depending on the environment. For example, the microbial diversity of host-associated systems such as the human microbiome<sup>11,12</sup> may be more amenable to cultivation compared to some environments such as soil. At present, it seems clear that most archaeal and bacterial diversity remains yet to be cultured<sup>10,13</sup>. The reasons are many, but as demonstrated recently by the cultivation of a member of the Asgard archaea<sup>14</sup>, syntrophic interactions, slow growth and media optimization can present formidable challenges.

## Rules of prokaryotic nomenclature and current challenges

Describing biodiversity and identifying organisms are the scientific goals of taxonomy. Taxonomy integrates classification and nomenclature to describe biological diversity. Classification circumscribes and ranks taxa, and nomenclature is the process of assigning names. The commonly used Linnaean nomenclatural system focuses on the recognition of species as the basic unit, which are included in taxa of successively higher ranks (genus, family, order, class and phylum). There is some flexibility on how to circumscribe microbial species using phylogenetic, genotypic and phenotypic data. Once a species is delineated, rules of nomenclature given in the International Code of Nomenclature of Prokaryotes (ICNP or 'the Code'<sup>14</sup>; see Box 2) guide the creation and assignment of names. This is true of all codes of nomenclature that currently exist—prokaryotes, viruses, animals, algae, fungi and plants—in addition to separate codes for cultivated plants and plant associations.

A full list of affiliations appears at the end of the paper.

## Box 1 | Glossary of terms

Term	Definition
Single-amplified genome (SAG)	Assembled genome in which the DNA sequence is derived from a single microbial cell.
Metagenome-assembled genome (MAG)	Assembled genome from a heterogeneous assemblage of cells; the MAG may represent multiple populations (with minor sequence variants) and thus represents a species-level group rather than a single genetic variant—or strain—that is typically represented by a cultivated isolate.
International Code of Nomenclature of Prokaryotes (ICNP)	Typically referred to as ‘the Code’. The definitive set of rules, principles and recommendations for naming Bacteria and Archaea.
International Committee on Systematics of Prokaryotes (ICSP)	The body that oversees the nomenclature of prokaryotes and supervises the publication of the ICNP. Their Judicial Commission issues opinions concerning nomenclatural matters and revisions to the Code.
International Code of Nomenclature of Uncultivated Prokaryotes (ICNUP)	The Uncultivated Code that will be developed if plan B is adopted to circumscribe nomenclature for uncultivated organisms.
International Journal of Systematic and Evolutionary Microbiology (IJSEM)	The primary journal for publication of descriptions of novel microbial taxa and lists of valid names; also where proposals to rename taxa are submitted. The official publication of the ICSP and the Bacteriology and Applied Microbiology Division of the International Union of Microbiological Societies.
International Sequence Database Collaboration (INSDC)	International coalition of databases comprised of the National Center for Biotechnology Information, the European Bioinformatics Institute and the DNA Databank of Japan.
Genome Standards Consortium (GSC)	The GSC is an open-membership working body that enables genomic data integration, discovery and comparison through international community-driven standards.
Minimal information about a single-amplified genome (MISAG)	Standard developed by the GSC for reporting single-amplified genome assembly quality, completeness and contamination information along with additional metadata.
Minimal information about a metagenome-assembled genome (MIMAG)	Standard developed by the GSC for reporting metagenome-assembled genome assembly quality, completeness and contamination information along with additional metadata.
Systematics	Systematics is the study of the units of biodiversity and their relationships. This includes the discovery of the basic units of biodiversity (species), reconstructing the patterns of relationships of species at successively higher levels and building classifications based on these patterns and taxa. This term is often used synonymously with taxonomy.
Taxonomy	The branch of science concerned with the classification, identification and nomenclature pertaining to organisms, in particular.
Nomenclature	A system for giving names to organisms.
Classification	The placement of organisms into groups on the basis of evolutionary relationships as well as similarities and shared qualities or characteristics.
Etymology	In prokaryotic nomenclature, the etymology of a name is the semantic derivation of the Latinized name describing what it was based on.
Protologue	A protologue summarizes the new name, its etymology, the taxon properties and the designated type material for which the microorganism is known.

As the development of prokaryotic taxonomy has mainly been informed by cultivation, there is currently no mechanism in the Code to either assign rank, or formally name, members of Archaea or Bacteria discovered using cultivation-independent approaches. The absence of stable names for uncultivated prokaryotes results in confusion in the literature and stands in the way of an integrated classification system. Therefore, an urgent need exists to reconsider the rules of nomenclature to include the entirety of archaeal and bacterial diversity. By recognizing this challenge, the intent is not to dissuade efforts for cultivation; quite the contrary, there remains a crucial need in the field of microbiology to bring microorganisms into pure culture or stable co-culture, and for culture collections to provide material for all matters related to the study of living microorganisms (such as physiology, growth characteristics and cell division).

## Naming conventions

Thus far, two conventions have been used to name uncultivated taxa. The first applies alphanumeric identifiers to 16S ribosomal RNA gene sequence clusters (for example, SAR11), which is now being extended to metagenome-assembled genomes (MAGs) and single-amplified genomes (SAGs). Alphanumeric identifiers are convenient as placeholders and can be used for communicating the underlying taxonomic or phylogenetic relationships and organizing diversity. However, the lack of consensus amongst the scientific community on rules for alphanumeric identifiers has resulted in frequent synonymies and confusion in the literature<sup>15</sup>. Secondly, an International Committee on Systematics of Prokaryotes (ICSP)-sanctioned approach for naming uncultivated taxa under the provisional ‘*Candidatus*’ classification has been in place for

more than two decades<sup>15</sup>. However, *Candidatus* is a category with no standing in nomenclature; thus, *Candidatus* names do not necessarily complement official nomenclature and do not have priority—that is, they do not have to be retained if representatives of the taxon are subsequently cultivated.

This Consensus Statement proposes two potential paths to develop a system of nomenclature for uncultivated microorganisms that allows them to be classified and named with a high degree of fidelity using MAG and/or SAG sequences. This would allow these microorganisms to be described according to predicted characteristics, and to be linked to environmental and ecological contextual data, resulting in similar integrity and reproducibility to the current system used to name and classify cultivated microorganisms<sup>17</sup>. One path requires modifications to the Code to allow the use of DNA sequence data as ‘type material’ as proposed by Whitman<sup>18</sup>, whereas the other creates a parallel code for uncultivated microorganisms, as previously proposed<sup>16,19</sup>. The concept of type material in these cases reflects that DNA sequence deposited in an International Sequence Database Collaboration repository is the international entity, supplanting the current ICNP requirement to deposit viable cultures in at least two culture collections. The focus of this Consensus Statement pertains to cases in which Archaea or Bacteria that are represented by DNA sequence information are to be named formally (at all levels of the taxonomic rank appropriate for the microorganism to be described), and where descriptive protologues are generated based on the DNA sequence information and supporting cultivation-independent and environmental data<sup>20</sup>. This Consensus Statement does not specifically address the overwhelming abundance of MAG and SAG data that will not be formally named, that is, those that are not studied in sufficient detail to provide meaningful insight into their structural, physiological, ecological or evolutionary properties. However, we advocate the adoption of quality standard frameworks<sup>21</sup> for both formally named and alphanumerically identified MAGs and SAGs.

### Lessons from the history of prokaryotic taxonomy

Prokaryotic taxonomy underwent two revolutions in the latter half of the twentieth century that, we posit, are analogous to the present situation. Initially, methodological limitations in archaeal and bacterial classification, particularly the reliance on staining, morphology and physiological properties, led to a confusing proliferation of names and a poorly ordered taxonomy rife with synonyms<sup>22</sup>. Consequently, an ICSP ad hoc committee was appointed in 1973 to review the legitimate names of bacteria and compile the Approved Lists of Bacterial Names, which designated type strains, and, in some cases, type material (the description, illustration or preserved specimen), with valid names<sup>23</sup>. Names not included in the list lost their standing in nomenclature (Box 1).

Several years later, in 1987 and 1990, two ad hoc committees discussed the incorporation of DNA sequence information into the bacterial species definition which resulted in the integrated use of phylogenetic and phenotypic characteristics, or polyphasic taxonomy<sup>24</sup>. In 2002, a fourth ad hoc committee revisited the species definition in light of new molecular-sequence-based methods, encouraged the use of the *Candidatus* provisional category and recommended data standards utilizing sequence databases<sup>25</sup>. Since then, no additional committees have been appointed to address the opportunities and complexities of massive increases in genomic data provided by advancements in DNA sequencing technology.

### Stabilizing *Candidatus* names

At present, there is a need to formally account for all *Candidatus* taxa that have been described according to the original proposal<sup>15</sup>; such an effort is already underway<sup>26</sup>. Since 1995, more than 700 *Candidatus* names have been proposed but, due to the lack of official rules and oversight, a significant proportion do not comply with the

### Box 2 | Brief introduction to the ICNP

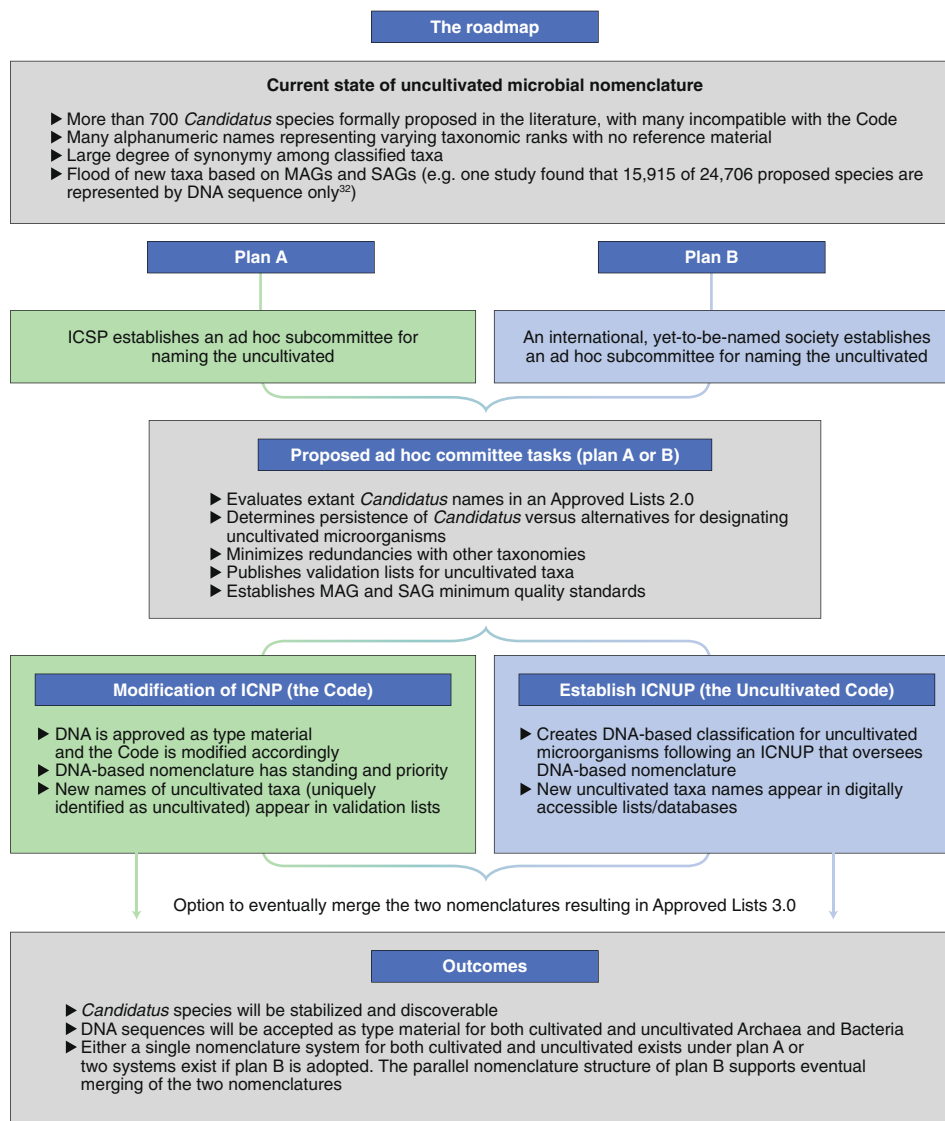
“The primary purpose of giving a name to a taxon is to supply a means of referring to it rather than to indicate the characters or the history of the taxon.” is stated in principle 4 of the Code (ref.<sup>15</sup>). The Code requires that every name is associated to a type material, that all names are announced and that names that are validated first have priority. Names are validated either by direct publication of papers in IJSEM (summarized in notification lists) or, if published in another journal (referred to as ‘effective publication’), by submission of any appropriately documented request to IJSEM for inclusion of the name in the validation list. This process ensures that nomenclature rules are followed, that names are communicated to the scientific community and that redundancies are prevented for different taxa. The taxon description (or protologue) includes, at a minimum, the Latin etymology of the name, a brief description of the taxon, designation of type material and, currently, details of pure culture submission to two culture collections. Priority simply means that the first validly published legitimate name is the ‘correct’ name of a taxon.

Naming under the Code is hierarchical, that is, the names of the higher ranks are derived from the names of the lower ranks. For instance, the names for families and orders are based on a type genus, the names for classes are based upon a type order, the names for phyla are based upon a type class and, reversely, the type of a phylum is a class, the type of a class is an order, the type of an order is a genus and the type of a genus is a species. If a new MAG represents a new class then the genus and species would be named first, for example, ‘*Ca. Caldatribacterium californiense*’, in the family Caldatribacteriaceae, order Caldatribacteriales and class Caldatribacteria<sup>4</sup>. A consequence of this system is that even the highest ranks are based upon some type material, which is the evidence for the existence of the taxon. Thus, this hierarchical system communicates phylogenetic information and differing levels of relatedness. If a future integrated Code for cultured and uncultivated organisms can be adopted, synonymies will be avoided. The requirement to deposit viable pure cultures into two international culture collections as type material is a relatively recent development in taxonomy. Prior to changes to the Code initiated in 1996, detailed descriptions or nonviable specimens were admissible as type material, as is currently the standard in other codes of nomenclature (such as the International Code of Nomenclature for algae, fungi, and plants; International Code of Zoological Nomenclature; and The International Code of Virus Classification and Nomenclature).

Code<sup>27</sup>. Many names have not been captured in a unified list; some names lack key aspects of the description such as the designation of a type or an etymology, and the quality of available data to serve as type material for *Candidatus* taxa varies greatly. For instance, some *Candidatus* species are only linked to 16S rRNA gene sequences, or to no genetic data at all, which complicates linking legacy with modern datasets. In addition, there are numerous higher taxa (such as candidate phyla) for which no lower rank or type has been designated. This also contradicts the principles of all codes of taxonomic nomenclature. Naming higher taxa has become common practice (compounded by the problem that the rank of phylum currently lacks official status in the Code<sup>28</sup>), particularly for newly discovered uncultivated lineages<sup>17</sup>.

### The path forward

The path described is the result of engagement with a large consortium of scientists who provided input (both co-authors and endorsees; see Supplementary Table 1). Plan A proposes the formal



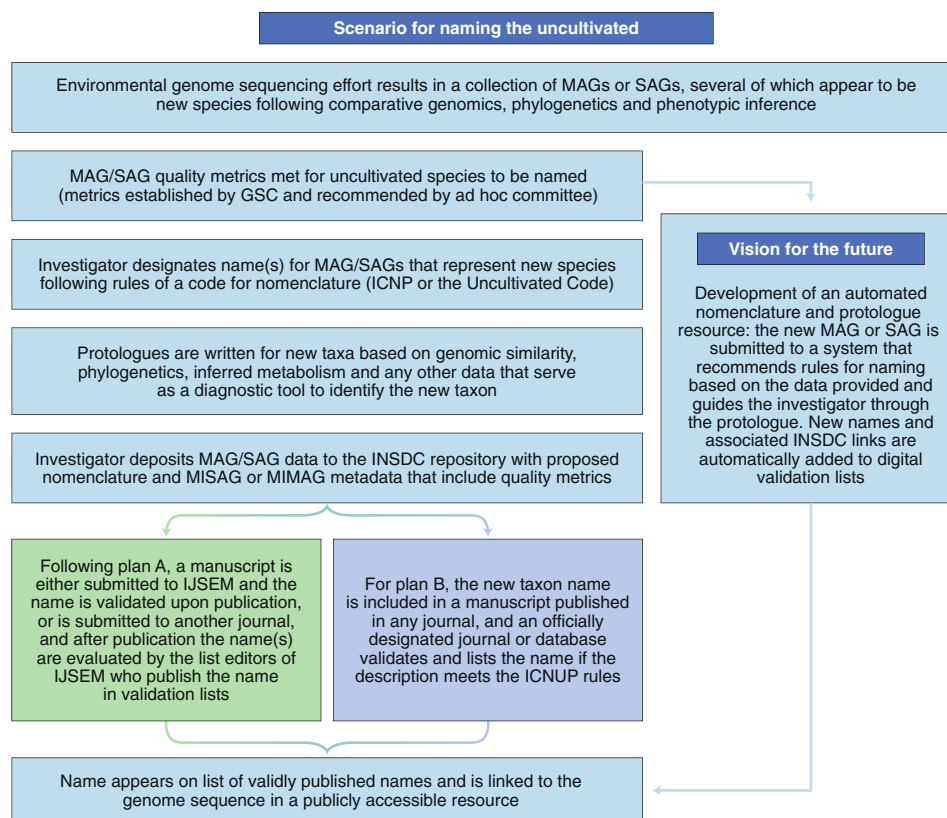
**Fig. 1 | Proposed roadmaps for nomenclature of uncultivated Archaea and Bacteria.** Plans A and B provide two alternatives for inclusion of uncultivated Archaea and Bacteria into the classical Linnaean nomenclature system.

revision of the Code to include uncultivated organisms represented by DNA sequence information as the nomenclatural type<sup>18</sup>, albeit with an allowance to distinguish cultivated and uncultivated taxa<sup>17</sup> (Fig. 1). Here, we refer to microorganisms available in pure culture (or stable co-culture) that can be named according to the rules of the Code as ‘cultivated’, and those that are recognized through their DNA sequence information as ‘uncultivated’ (including mixed cultures in which the members are recognized through DNA sequence information). Plan A could be initiated by establishing a subcommittee of the ICSP to review and stabilize the current *Candidatus* nomenclature, develop standards for DNA sequence data to serve as type material, address the use of *Candidatus* or other alternatives (such as superscripts ‘u’, ‘c’ or ‘e’ to represent uncultivated, *Candidatus* or environmental, respectively) to distinguish between names of uncultivated organisms versus those derived from cultures<sup>15,17</sup> and to establish an updated list (an ‘Approved Lists 2.0’) of approved nomenclature to include previously named taxa with DNA sequence as type material. These new names and descriptions (protologues including etymology, taxon properties, inferred phenotype and sequence deposit accession numbers), whether for single taxa or large-scale MAG and SAG studies, would be communicated

through the literature, reviewed by the International Journal of Systematic and Evolutionary Microbiology (IJSEM) list editors and then included in the revised list, granting them priority over subsequent names. Plan A creates a framework within the Code that will lead to a smooth integration of a harmonized nomenclature and will facilitate future unified nomenclature.

An alternate, near-term solution (plan B) would be the creation of a parallel code for uncultivated taxa (the ‘Uncultivated Code’; Fig. 1) under the auspices of an international entity with enough authority to provide a unified framework. This entity could take on the responsibility for supporting the development of an International Code of Nomenclature of Uncultivated Prokaryotes (ICNUP; that is, the Uncultivated Code) ruling on its actions and publishing a list (a digital record) of valid names for uncultivated taxa. The ICNUP would appoint an ad hoc committee to address the current *Candidatus* names and develop an ‘Approved Uncultivated Lists 2.0’ inclusive of *Candidatus* species names to stabilize the nomenclature and ascertain priority. The *Candidatus* designation could be preserved, or some other notation recommended to identify uncultivated status. Likewise, the ad hoc committee could provide guidelines regarding quality standards and full taxonomic





**Fig. 2 | Scenario for naming uncultivated Archaea and Bacteria.** In cases where naming a new species is warranted, the steps outlined here are a likely process for nomenclature regardless of whether plan A or plan B (Fig. 1) is adopted.

classification for MAGs and SAGs to be named going forward, possibly with input from the Genome Standards Consortium (GSC). As recommended<sup>15</sup>, the rules of the Uncultivated Code would be analogous to the Code, and *Candidatus* names already published and supported by DNA sequence information would be granted priority as in plan A. This parallel structure allows the two nomenclature systems to be merged to yield a single, unified collection of validly published names (for example, an ‘Approved Lists 3.0’) if and when supported by the scientific community. Alternatively, the two systems could exist in parallel and never be unified (Fig. 1). We also recommend that names established under the Uncultivated Code be conserved in cases where uncultivated taxa are brought into pure culture—the ultimate path for microbiological characterization.

Plan A works within the Code to avoid decentralizing the process of nomenclature—thus mitigating disputes over priority in the future—and could be implemented rapidly to effectively meet the immediate demands of the scientific community. However, a practical, expedient solution is required. If ratification of the revised Code via the ICSP is prolonged (as it has been recently<sup>16</sup>), adoption of the scenario described in plan B could provide a timely solution to avoid conflict in the nomenclatural system and promote communication across stakeholders in the prokaryotic sciences. In a practical sense, both plans result in a similar process for naming uncultivated microorganisms in which the uncultivated representatives have a unique identifier (Fig. 2).

### Quality standards and digital protologues

Regardless of the path forward (plan A or B), we propose the development of genomic standards to guide the naming of uncultivated taxa to the extent possible, across all taxonomic ranks. Relevant

standards for MAGs and SAGs have recently been published, including recommendations on contextual information or meta-data (for example, geographic location, biome and sampled material characteristics)<sup>16</sup>, minimal standards based on MAG and SAG completeness and contamination<sup>21</sup>, and type material<sup>17</sup> (Table 1). In addition, the overseeing body (such as the ICSP) could also provide direction to the scientific community on how and when to name (and not name) a MAG or SAG. Likewise, the overseeing body could also recommend standardized naming practices that could be applied to high quality MAGs and SAGs currently deposited in public repositories.

With the impending adoption of minimal information about a single-amplified genome (MISAG) and metagenome-assembled genome (MIMAG) checklists in GenBank and the European Nucleotide Archive<sup>29</sup>, it is now up to the scientific community, through peer review and journal policies, to ensure reporting of SAG and MAG data quality. While MAGs may not always represent single genomes, if they are of high quality and have minimal contamination, they likely represent the consensus genome of a natural microbial population. Thus, while the designation of a type strain is unlikely (albeit advances in long-read sequencing technology may aid in this respect), MAGs can act as the nomenclatural type for a species despite their mosaic nature. This distinction should be carried forward regardless of whether plan A or B are adopted. If standards are not enforced by the scientific community, the risk is that poor-quality genomes with contaminating sequences could exacerbate transitive errors in annotation (such as cases in which a contaminating sequence could be misidentified as being associated with the particular MAG) and species assignments in downstream phylogenomic studies—a clearly undesirable situation that is not limited to MAGs and SAGs.

**Table 1 | Data inputs for an ad hoc committee naming uncultivated SAGs and MAGs**

Category	Parameters for data quality recommendations
Genome quality	Percent completion
	Percent contamination
	Presence of 5S, 16S 23S rRNA gene (and level of completeness)
	Number of tRNA genes
Assembly quality	N50 (defined as the length of the shortest contig in the set of largest contigs that together constitute at least half of the total assembly size)
	Number of contigs
Naming conventions	Status of <i>Candidatus</i> and future nomenclature (for example, superscript <sup>u</sup> , <sup>c</sup> or <sup>e</sup> designating uncultivated, <i>Candidatus</i> or environmental microorganisms, respectively)
Description requirements	Phenotype or metabolic prediction based on DNA sequence
	Ecological and biogeographic consideration
	Additional metadata (including guanine and cytosine content, genome size and number of protein coding genes)
Validation of MAG and SAG nomenclature	Potential for informational system of classification, validation and DOI assignment

Current publishing capabilities will continue to struggle to keep pace with the anticipated number of taxonomic descriptions, especially if MAGs and SAGs were allowed as type material. Therefore, the future of this field requires breakthroughs in information access and advances in database interoperability. Examples of these breakthroughs include the creation of standardized, machine-readable formats for nomenclature that can capture name changes, automated taxonomic assignment based on big-data analysis (with best criteria discussed and widely adopted in the community) and nomenclature pipelines guiding the user through rules for naming by following guidelines of the Code or the Uncultivated Code. Automated mechanisms to create properly formatted protologues (Fig. 2) are also urgently needed<sup>9,30,31</sup>.

### Concluding remarks

This Consensus Statement addresses the need to provide a stable nomenclature and taxonomy for uncultivated Archaea and Bacteria that will enable scientific discourse among the many fields that communicate microbial diversity information. The proposed plans (A or B) enable a roadmap for communicating the enormous diversity of the prokaryotic world. This includes a standardized framework for naming uncultivated Archaea and Bacteria that will provide a needed structure to the classification system and allow for scientific communication regarding diversity across the microbial sciences. The proposed roadmap is not meant to suggest that all MAGs and SAGs will be named according to the Linnaean nomenclature—many will remain with alphanumeric identifiers. Instead, the roadmap provides a path for naming MAGs and SAGs that meet high quality standards. There are additional needs for discovering and classifying both named and unnamed MAGs and SAGs based on phylogenomics, and for identifying high-quality, well-curated representative (or ‘type’) genomes<sup>32</sup> that are not addressed here.

Regardless, implementation of either of our proposed plans will require engagement from the scientific community (including the ICSP) to address the finer details, some of which were not captured herein. As evidenced by our effort here, there is substantial interest from the scientific community to participate in the decision-making process for determining standards in nomenclature that affect the entire microbiology field. We can look to the virus community for guidance in their adoption of nomenclature rules based on viral genome sequences<sup>33</sup> in which the International Committee on Taxonomy of Viruses endorsed the proposal to include (meta) genome sequence data. The utility of DNA sequences as type material is relevant to organisms across the tree of life and biologists in other fields, including fungi and protists, face similar challenges<sup>34</sup>. We hope that the solutions identified in this roadmap might also apply to the naming of other organisms in these diverse fields.

**Note added in proof:** Whilst this manuscript was in revision, the ICSP held an e-mail discussion forum, followed by voting on the Whitman<sup>18</sup> proposals to modify the ICNP to allow sequence data as type material (plan A). In the subsequent ICSP vote, these proposals were rejected. Minutes of the e-mail discussion will be made available on the ICSP website. Although further proposals to modify the ICNP may be forthcoming, this result makes the imminent adoption of plan A unlikely and therefore increases the likelihood of plan B being enacted.

Received: 14 June 2019; Accepted: 1 May 2020;  
Published online: 8 June 2020

### References

- Hug, L. A. et al. A new view of the tree of life. *Nat. Microbiol.* **1**, 16048 (2016).
- Almeida, A. et al. A new genomic blueprint of the human gut microbiota. *Nature* **568**, 499–504 (2019).
- Berdy, B., Sporing, A. L., Ling, L. L. & Epstein, S. S. In situ cultivation of previously uncultivable microorganisms using the ichip. *Nat. Protoc.* **12**, 2232–2242 (2017).
- Cross, K. L. et al. Targeted isolation and cultivation of uncultivated bacteria by reverse genomics. *Nat. Biotechnol.* **37**, 1314–1321 (2019).
- Rinke, C. et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**, 431–437 (2013).
- Pachiadaki, M. et al. Charting the complexity of the marine microbiome through single-cell genomics. *Cell* **179**, 1623–1635 (2019).
- Brown, C. T. et al. Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523**, 208–211 (2015).
- Spang, A. et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173–179 (2015).
- Parks, D. H. et al. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat. Microbiol.* **2**, 1533–1542 (2017).
- Steen, A. D. et al. High proportions of bacteria and archaea across most biomes remain uncultured. *ISME J.* **13**, 3126–3130 (2019).
- Forster, S. C. et al. A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat. Biotechnol.* **37**, 186–192 (2019).
- Lagier, J. C. et al. Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol.* **1**, 16203 (2016).
- Lloyd, K. G., Steen, A. D., Ladau, J., Yin, J. Q. & Crosby, L. Phylogenetically novel uncultured microbial cells dominate Earth microbiomes. *mSystems* **3**, e00055–18 (2018).
- Imachi, H. et al. Isolation of an archaeon at the prokaryote–eukaryote interface. *Nature* **577**, 519–525 (2020).
- Parker, C. T., Tindall, B. J. & Garrity, G. M. International Code of Nomenclature of Prokaryotes. Prokaryotic Code (2008 revision). *Int. J. Syst. Evol. Microbiol.* **69**, S1–S111 (2019).
- Konstantinidis, K. T., Rossello-Mora, R. & Amann, R. Reply to the commentary “Uncultivated microbes—in need of their own nomenclature?”. *ISME J.* **12**, 653–654 (2018).
- Chuvpochina, M. et al. The importance of designating type material for uncultured taxa. *Syst. Appl. Microbiol.* **42**, 15–21 (2019).
- Whitman, W. B. Modest proposals to expand the type material for naming of prokaryotes. *Int. J. Syst. Evol. Microbiol.* **66**, 2018–2112 (2016).
- Konstantinidis, K. T., Rossello-Mora, R. & Amann, R. Uncultivated microbes in need of their own taxonomy. *ISME J.* **11**, 2399–2406 (2017).
- Hatzenpichler, R., Krukenberg, V., Spietz, R. L. & Jay, Z. J. Next-generation physiology approaches to studying microbiome function at the single cell level. *Nat. Microbiol. Rev.* **18**, 241–256 (2020).

21. Bowers, R. M. et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* **35**, 725–731 (2017).
22. Sneath, P. H. A. & Brenner, D. J. “Official” nomenclature lists. *ASM News* **58**, 175 (1992).
23. Sneath, P. H. A. The preparation of the approved lists of bacterial names. *Int. J. Syst. Evol. Microbiol.* **55**, 2247–2449 (2005).
24. Murray, R. G. E. et al. Report of the ad hoc committee on approaches to taxonomy within the Proteobacteria. *Int. J. Syst. Bact.* **40**, 213–215 (1990).
25. Stackebrandt, E. et al. Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int. J. Syst. Evol. Microbiol.* **52**, 1043–1047 (2002).
26. Oren, A., Garrity, G. M. & Trujillo, M. E. Registration of names of prokaryotic Candidatus taxa in the IJSEM. *Int. J. Syst. Evol. Microbiol.* <https://doi.org/10.1099/ijsem.0.003791> (2019).
27. Oren, A. A plea for linguistic accuracy – also for Candidatus taxa. *Int. J. Syst. Evol. Microbiol.* **67**, 1085–1094 (2017).
28. Rossello-Mora, R. & Whitman, W. B. Dialogue on the nomenclature and classification of prokaryotes. *Syst. Appl. Microbiol.* **42**, 5–14 (2019).
29. Amid, C. et al. The European Nucleotide Archive in 2019. *Nucleic Acids Res.* **48**, D70–D76 (2019).
30. Rodriguez-R, L. M. et al. The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. *Nucleic Acids Res.* **46**, W282–W288 (2018).
31. Rossello-Mora, R. & Sutcliffe, I. C. Reflections on the introduction of the Digital Protologue Database – a partial success? *Antonie van Leeuwenhoek* **112**, 141–143 (2019).
32. Parks, D. H. et al. A complete domain-to-species taxonomy for Bacteria and Archaea. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-020-0501-8> (2020).
33. Simmonds, P. et al. Virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol.* **15**, 161–168 (2017).
34. Lewin, H. A. et al. Earth BioGenome Project: sequencing life for the future of life. *Proc. Natl Acad. Sci. USA* **115**, 4325–4333 (2018).

## Acknowledgements

Support for this effort was provided by the National Science Foundation (grant no. DEB-1841658). We extend deep appreciation to all who attended the Microbial Taxonomy 2018 seminars, the Microbial Systematics for the Next Decade workshop held in Hood River, Oregon, USA, in October 2018, and those that participated in subsequent discussions and provided comments to mature the concept communicated here. Please

see Supplementary Table 1 for a list of representatives from the international microbial sciences community that have endorsed this Consensus Statement.

## Author contributions

A.E.M., J.F., S.G., R.H., P.H., P.K., K.T.K., C.E.L., R.T.P., D.H.P., R.R.-M., M.B.S., I.C.S., J.C.T., S.N.V., W.B.W., B.P.H. and A.-L.R. wrote the article. All authors (A.E.M., J.F., S.G., R.H., P.H., P.K., K.T.K., C.E.L., R.T.P., D.H.P., R.R.-M., M.B.S., I.C.S., J.C.T., S.N.V., W.B.W., S.G.A., R.I.A., K.A., J.A., B.J.B., R.A.B., H.B.B., E.S.B., C.L.B., P.C., P.S.G.C., D.R.C., K.M.D., M.A.R., P.E.S., C.A.D., J.A.E., D.E., T.J.G.E., D.E., P.R.G., U.H., J.T.H., L.A.H., W.P.I., E.P.I., H.-P.K., W.-J.L., K.G.L., F.E.L., T.M., D.P.M., T.N., M.P., V.P., C.P.-A., A.J.P., T.H.M.S., A.D.S., E.T.S., A.S., F.J.S., J.M.T., P.V., M.W., F.-P.W., B.P.H. and A.-L.R.) contributed to the discussion of the manuscript content and reviewed and/or edited the manuscript before submission of the final version.

## Competing interests

The authors have no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41564-020-0733-x>.

**Correspondence and requests for materials** should be addressed to A.E.M., B.P.H. or A.-L.R.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

<sup>1</sup>Division of Earth and Ecosystem Sciences, Desert Research Institute, Reno, NV, USA. <sup>2</sup>Department of Evolution, Ecology and Organismal Biology, The Ohio State University, Columbus, OH, USA. <sup>3</sup>Evolutionary Biology of the Microbial Cell, Department of Microbiology, Institut Pasteur, Paris, France. <sup>4</sup>Department of Chemistry and Biochemistry, Center for Biofilm Engineering, and Thermal Biology Institute, Montana State University, Bozeman, MT, USA. <sup>5</sup>Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, The University of Queensland, Brisbane, Australia. <sup>6</sup>Department of Applied Microbiology, Justus-Liebig-Universität, Giessen, Germany. <sup>7</sup>School of Civil and Environmental Engineering, Georgia Tech, Atlanta, GA, USA. <sup>8</sup>Department of Biological Sciences, University of Rhode Island, Kingston, RI, USA. <sup>9</sup>Department of Molecular and Cellular Biology, University of Connecticut, Storrs, CT, USA. <sup>10</sup>Mediterranean Institute for Advanced Studies, CSIC-UIB, Illes Balears, Spain. <sup>11</sup>School of Biological Sciences, University of Canterbury, Christchurch, New Zealand. <sup>12</sup>Department of Applied Sciences, Northumbria University, Newcastle upon Tyne, UK. <sup>13</sup>Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA. <sup>14</sup>Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa. <sup>15</sup>Department of Microbiology, University of Georgia, Athens, GA, USA. <sup>16</sup>Department of Marine Biology and Oceanography, Institut de Ciències del Mar, CSIC, Barcelona, Spain. <sup>17</sup>Max Planck Institute for Marine Microbiology, Bremen, Germany. <sup>18</sup>Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA. <sup>19</sup>CEA Technological Innovations for Detection and Diagnosis Laboratory, CEA Pharmacology and Immunoanalysis Unit (SPI), Bagnols-sur-Cèze, France. <sup>20</sup>Department of Marine Science, Marine Science Institute, University of Texas at Austin, Port Aransas, TX, USA. <sup>21</sup>Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA. <sup>22</sup>Molecular Biotechnology, Department of Biosciences and Buchmann Institute for Molecular Life Sciences (BMLS), Goethe University Frankfurt, Frankfurt am Main, Germany. <sup>23</sup>Senckenberg Society for Nature Research, Frankfurt am Main, Germany. <sup>24</sup>Department of Microbiology and Immunology, Montana State University, Bozeman, MT, USA. <sup>25</sup>University of the West of England, Bristol, England. <sup>26</sup>Department of Environmental Science, University of Arizona, Tucson, AZ, USA. <sup>27</sup>Bioscience Division, Los Alamos National Laboratory, Los Alamos, NM, USA. <sup>28</sup>Microbiology Department, University of Massachusetts, Amherst, MA, USA. <sup>29</sup>National Museum of Natural Sciences, CSIC, Madrid, Spain. <sup>30</sup>Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Mexico City, Mexico. <sup>31</sup>National Center for Agricultural Utilization Research, Crop Bioprotection Research Unit, Peoria, IL, USA. <sup>32</sup>Department of Evolution and Ecology, Department of Medical Microbiology and Immunology, University of California, Davis, CA, USA. <sup>33</sup>Bigelow Laboratory for Ocean Sciences, East Boothbay, ME, USA. <sup>34</sup>Laboratory of Microbiology, Wageningen University & Research, Wageningen, the Netherlands. <sup>35</sup>LS2N, Université de Nantes, CNRS, Nantes, France. <sup>36</sup>Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. <sup>37</sup>GEOMAR-Helmholtz Centre for Ocean Research, RD3-Marine Ecology, RU-Marine Microbiology, Kiel, Germany. <sup>38</sup>Department of Marine Sciences, University of Georgia, Athens, GA, USA. <sup>39</sup>Department of Biology, University of Waterloo, Waterloo, Canada. <sup>40</sup>Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA. <sup>41</sup>School of Science, RMIT University, Melbourne, Victoria, Australia. <sup>42</sup>School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, UK. <sup>43</sup>School of Life Sciences, Sun Yat-Sen University, Guangzhou, China. <sup>44</sup>Department of Microbiology, University of Tennessee, Knoxville, TN, USA. <sup>45</sup>Departments of Microbiology and Civil & Environmental Engineering, Center for Environmental Biotechnology, University of Tennessee, Knoxville, TN, USA. <sup>46</sup>Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA. <sup>47</sup>Centre for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa. <sup>48</sup>Division of Hydrologic Sciences, Desert Research Institute, Las Vegas, NV, USA. <sup>49</sup>Research Center for Bioscience and Nanoscience (CeBN), Japan Agency for



Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan. <sup>50</sup>School of Life Sciences, University of Nevada, Las Vegas, NV, USA. <sup>51</sup>Astrobiology Center (CSIC-INTA), Madrid, Spain. <sup>52</sup>National Biotechnology Center, Cantoblanco, Madrid, Spain. <sup>53</sup>Department of Chemistry, Environmental Microbiology and Biotechnology, University of Duisburg-Essen, Essen, Germany. <sup>54</sup>Environmental Genomics and Systems Biology Research Group, Institute for Environment and Natural Resources, Zürich University for Applied Sciences (ZHAW), Wädenswil, Switzerland. <sup>55</sup>Departments of Microbiology and Earth and Planetary Sciences, University of Tennessee, Knoxville, TN, USA. <sup>56</sup>Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa. <sup>57</sup>Department for Marine Microbiology and Biogeochemistry, Royal Netherlands Institute for Sea Research, Den Burg, the Netherlands. <sup>58</sup>Department of Cell and Molecular Biology, Uppsala University, Uppsala, Sweden. <sup>59</sup>Center for Microbial Ecology, Department of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI, USA. <sup>60</sup>Department of Biochemistry and Microbiology, Ghent University, Ghent, Belgium. <sup>61</sup>Centre for Microbiology and Environmental Systems Science, University of Vienna, Vienna, Austria. <sup>62</sup>International Center for Deep Life Investigation, School of Oceanography and School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China. <sup>63</sup>Biology Department, Portland State University, Portland, OR, USA. <sup>64</sup>e-mail: [alison.murray@dri.edu](mailto:alison.murray@dri.edu); [brian.hedlund@unlv.edu](mailto:brian.hedlund@unlv.edu); [bwar@pdx.edu](mailto:bwar@pdx.edu)