

# Nomenclature of prokaryotic ‘*Candidatus*’ taxa: establishing order in the current chaos

A. Oren

The Institute of Life Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Jerusalem, Israel

## Abstract

In the mid-1990s, the category ‘*Candidatus*’ was established for putative taxa of as yet uncultivated prokaryotes. The status of ‘*Candidatus*’ is not formally included in the rules of the International Code of Nomenclature of Prokaryotes. Thus, ‘*Candidatus*’ names do not have standing in the nomenclature. Curated annotated lists of ‘*Candidatus*’ names (not including phyla) have been published since 2020. By April 2021, about 2700 names of ‘*Candidatus*’ taxa had been published. The International Committee on Systematics of Prokaryotes recently rejected proposals to allow gene sequence data as nomenclatural types. An alternative code for naming uncultivated microorganisms (the ‘SeqCode’) is now being developed for naming the majority of prokaryotes that are as yet uncultivated. In the opinion of the author, there is no need for such a code, as the existing system, with nomenclature quality control also for ‘*Candidatus*’ names, fulfills the needs. Computer programs such as GAN which generates large numbers of correctly formed names from the short lists of Latin and Greek word elements and Protologger that produce descriptions directly from genome sequences will become important in the future for automated naming and description of large numbers of ‘*Candidatus*’ taxa from metagenomic and single cell genome data. However, the formation of interesting and meaningful names is encouraged whenever possible.

© 2021 The Author(s). Published by Elsevier Ltd.

**Keywords:** International code of nomenclature of prokaryotes, names, registration, ‘*Candidatus*’ taxa

**Original Submission:** 4 June 2021; **Revised Submission:** 5 August 2021; **Accepted:** 9 August 2021

**Article published online:** 13 September 2021

**Corresponding author:** A. Oren, Department of Plant and Environmental Sciences, The Institute of Life Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus, 9190401, Jerusalem, Israel

**E-mail:** [aharon.oren@mail.huji.ac.il](mailto:aharon.oren@mail.huji.ac.il)

## The origin of the ‘*Candidatus*’ concept

In the early 1990s, techniques such as 16S rRNA gene sequencing and fluorescent *in situ* hybridization became widely available. This enabled the molecular description of prokaryotes without the need for cultivation. Soon, the first Latin-based names were proposed for uncultivated prokaryotes whose uniqueness was defined only by limited characteristics, such as differences in molecular sequences. The International Code of Nomenclature of Bacteria (1990 revision) [1] (ICNB,

now the International Code of Nomenclature of Prokaryotes, ICNP [2]) did not have provisions for the naming of new taxa of uncultivated prokaryotes based solely or mainly on gene sequences. Already then, it was predicted that the problem would increase as the new technologies gain wider use and become easier to apply. Therefore a formal proposal for the establishment of the category ‘*Candidatus*’ to record the properties of putative taxa of prokaryotes was submitted for discussion by the International Committee on Systematics of Bacteria (ICSB, now the International Committee on Systematics of Prokaryotes, ICSP) [3].

The ICSP recommended that the category ‘*Candidatus*’ to record the properties of putative taxa of prokaryotes should be implemented for describing prokaryotic entities for which more than a mere sequence is available, but that could not be described in sufficient detail for valid publication of newly proposed names under the rules of the code of nomenclature [4]. Based on the recommendation by the ICSP, genomic information such as sequences that determine the phylogenetic

position of the organism should be supplemented by as much additional information as possible, including structural, metabolic and reproductive features for the description of a provisional taxon. Information about the natural environment in which the organism can be identified by *in situ* hybridization or other techniques for cell identification should also be provided [5].

Recent developments in metagenomics allow the recognition of discrete populations of DNA sequences in environmental samples that can be considered to belong to individual closely related populations that may be identified as members of yet to be described species. In view of the large amount of information that can be retrieved from metagenome-assembled genomes, ‘*Candidatus*’ taxa can also be proposed for environmentally occurring, as yet uncultured species not exhibiting exceptional morphologies, phenotypes or ecological relevancies [6]. This practice has already become widely used in recent years.

The ICSB/ICSP did not include the status of ‘*Candidatus*’ in the rules of the code of nomenclature so that ‘*Candidatus*’ names do not have standing in the nomenclature and are not validly published. Their status is explained in Appendix 11 of the ICNP [2], which is largely based on the original proposal approved in 1994 [4]. When an organism described as ‘*Candidatus*’ is later isolated, and the properties of the pure culture are sufficiently described, the name can be validly published, and the former ‘*Candidatus*’ organism’s name is deleted from the list of names of ‘*Candidatus*’ taxa. Appendix 11 needs to be updated in the near future in accordance with the recent developments as outlined below. An updated version is in preparation as part of the ongoing revision of the ICNP [7].

### Conventions for naming ‘*Candidatus*’ taxa – then and now

According to the original guidelines proposed by Murray and Stackebrandt [5], copied in Appendix 11 of the ICNP, a name of an organism in the status of ‘*Candidatus*’ consists of the word ‘*Candidatus*’, followed by a ‘vernacular epithet’ that consists of either a genus name with a specific epithet, or only a genus name, or only a specific epithet. The following examples were given:

- ‘*Candidatus* *Liberobacter asiaticum*’ and ‘*Candidatus* *Liberobacter africanum*’ – cases of a generic name with different specific epithets. Unfortunately, these examples are problematic for two reasons: (1) the names contravene Rule 12(c)1 of the ICNB/ICNP that states that an adjective used as a specific epithet must agree in gender with the generic name, as required by the rules of Latin

grammar. Correct epithets are therefore *asiaticus* and *africanus*; (2) based on Appendix 9 – Orthography of the ICNP [2,8], the connecting vowel must be -o- when the preceding word element is of Greek origin, it is -i- when the preceding word element is of Latin origin. *Liberobacter* is therefore correct, and that name was indeed validly published in 2014 after the first representative of the group was brought into culture (*Liberobacter crescens*) [9]. ‘*Candidatus* *Liberobacter asiaticus*’ corrig. and ‘*Candidatus* *Liberobacter africanus*’ corrig. are still awaiting isolation.

- ‘*Candidatus* *magnetobacterium*’ – an example of a genus name, originally spelt with a lower case ‘m’. The practice of forming ‘*Candidatus*’ names in this way has been abandoned, and today we have records of ‘*Candidatus* *Magnetobacterium bavaricum*’ and ‘*Candidatus* *Magnetobacterium casense*’ (originally published with the epithet *casensis*, contravening Rule 12c(1) of the ICNP [10,11].
- ‘*Candidatus* *intracellularis*’ – a ‘*Candidatus*’ name, consisting of only a specific epithet. No more such cases were published. The validly published name *Lawsonia intracellularis* McOrist et al. 1995 now replaces ‘*Candidatus* *intracellularis*’ [11].

All ‘*Candidatus*’ names of the rank of genus and species proposed in the past two decades followed the conventions of Linnaean binominal nomenclature, as adopted in the rules of the ICNP. There are also many ‘*Candidatus*’ names for higher taxa that are not yet represented by cultivated representatives. Here the nomenclature rules of the ICNP were typically followed (e.g. the ending *-aceae* for the rank of family, *-ales* for the rank of order). The word ‘*Candidatus*’, but not the genus name and/or the vernacular epithet, is printed in italics. ‘*Candidatus*’ is often abbreviated ‘*Ca*’. The original proposal by Murray and Schleifer [3] suggested using quotation marks for ‘*Candidatus*’ taxa, but this proposal was not copied in the emended proposal by Murray and Stackebrandt [5] and in Appendix 11 of the ICNP [2]. The List of Prokaryotic names with Standing in Nomenclature (LPSN) website (<https://www.bacterio.net>; <https://lpsn.dsmz.de>) [12] places the names of *Candidatus* taxa in quotation marks.

### How many ‘*Candidatus*’ names have been proposed?

At its meeting in Prague in 1994, the Judicial Commission of the ICSB recommended that a ‘*Candidatus*’ list should be established in the IJSB [4]. Therefore, Murray and Stackebrandt [4] wrote, “Workers who have described a bacterial *Candidatus* taxon should submit to the International Journal of Systematic

**TABLE 1.** The number of '*Candidatus*' names in List no. 1 (covering the period 1995-2018) [11], List no. 2 (new names added in 2019 and addenda to List no. 1) [14], List no. 3 (new names added in 2020 and addenda to the previous lists; in preparation) and List no. 4 (new names added in 2021 and addenda to the previous lists; in preparation; numbers listed are updated to April 30, 2021)

Period covered	List no. 1	List no. 2	List no. 3	List no. 4
	1995–2018	2019 + addenda	2020 + addenda	January – April 2021 + addenda
Classes	7	7 + 5	9 + 2	7
Orders	12	11 + 5	17 + 5	10
Families	25	16 + 6	19 + 3	10
Genera	329	53 + 14	46 + 30	250
Species	706	86 + 24	83 + 90	796
Subspecies	10	0	1	0
Other	2	0	0	0
Total	1091	173 + 54	175 + 130	1073
' <i>Candidatus</i> ' taxa grown in pure culture, ' <i>Candidatus</i> ' status lost	54	2	2	0

Bacteriology a reprint along with an entry in the appropriate format and a cover letter. Lists will appear at appropriate intervals (annually at the outset)." A similar statement is found in Appendix 11 of the ICNP: "A list in the form of a codified record of organisms of the status *Candidatus* is kept by the Judicial Commission of the ICSP in cooperation with the Editorial Board of the IJSEM and is published in that journal in appropriate intervals" [2]. However, this task of the Judicial Commission and the Editorial Board of the IJSEM was never formalized in the statutes of the ICSB/ICSP. Until recently, no such lists were ever compiled.

The first curated and annotated list of '*Candidatus*' taxa was published in the IJSEM in 2020. It contained 1091 names of '*Candidatus*' taxa published until the end of 2018 [11]. The list was accompanied by an editorial in which the purpose of the '*Candidatus*' lists was explained [13]. They are intended to comply at least in part with the guidelines of Appendix 11 of the ICNP and provide an inventory of published '*Candidatus*' names. The list editors of the IJSEM are making an effort to keep the record of published names of '*Candidatus*' taxa as complete as possible. These lists are not official documents, and names listed there are not validly published or officially approved in any other way. Many of the names are also listed in the List of Prokaryotic names with Standing in Nomenclature (LPSN) [12].

List no. 2 contains names of new '*Candidatus*' taxa published in 2019 and addenda to the first list [14]. List no. 3 (names published in 2020) and List no. 4 (names published in 2021) are in preparation. Table 1 summarizes the total number of '*Candidatus*' names that the list editors retrieved from the literature (not including names of phyla, see section 4 below). Most of the 250 new '*Candidatus*' genera and nearly 800 '*Candidatus*' species names added in 2021 were proposed in a single paper describing the chicken gut microbiome [15].

### Toward an inventory of names of '*Candidatus*' phyla

The published '*Candidatus*' lists [11,14] include names for those taxonomic ranks covered by the rules of the ICNP: from subspecies to class. As the rank of phylum was not included in the rules of the current version of the ICNP [2], names of '*Candidatus*' phyla that are not yet represented by cultivated members were not listed.

Recently, the ICSP has voted to include the rank of phylum in the rules of the ICNP [16]. The nomenclatural type of a phylum must be one of the contained genera. The upcoming revision of the code, now in preparation [7], will therefore contain emendations necessary to add the rank of phylum. A list of names of phyla with cultivated representatives, complete with the etymologies of the names, descriptions and proposed nomenclatural types, is currently in preparation to effect the valid publication of the names.

As the rank of phylum will soon be included in the rules of the ICNP, it also will be necessary to prepare lists of names of '*Candidatus*' phyla. There are over 150 such names in the literature, not including phyla currently designated by alphanumeric names only. Most of those names were proposed following massive metagenomics sequencing approaches that greatly expanded our understanding of the microbial diversity in nature, including the yet uncharted 'microbial dark matter'. These studies resulted in an expanded version of the tree of life [17,18]. Some of the proposed '*Candidatus*' phylum names may turn out to be synonyms. Many of the '*Candidatus*' phylum names were not formed in accordance with the orthography rules of Appendix 9 of the ICNP, and for most of those phyla, no subordinate taxa were ever defined. This causes problems as

**TABLE 2.** Examples of ‘*Candidatus*’ names proposed in the literature that do not comply with the rules and recommendations of the ICNP and its orthography appendix. The examples listed were included in Lists no. 1 and 2 [11,14], and some were discussed in an earlier essay [10]. Note that the suggested alternative names are proposals only and do not exclude other possibilities

Published ‘ <i>Candidatus</i> ’ name	Reasons why corrections are needed	Relevant rules of the ICNP	Proposed corrected name
Blochmannia	A generic name that exists in botany	Principle 2	Blochmanniella
Brownia	A generic name that exists in zoology	Principle 2	Spencerbrownia
Polytropus	A generic name that exists in zoology	Principle 2	Allopolytropus
Chaer	A name not treated as Latin	Principle 3, Rule 6	Chaera
Uab	A name not treated as Latin	Principle 3, Rule 6	Uabimicrobium
Rohrkolberia	Words from languages other than Latin or Greek should be avoided as long as equivalents exist in Latin or Greek or can be constructed by combining word elements from these two languages	Recommendation 6(3)	Typhincola
Bartonella woyliei	Words from languages other than Latin or Greek should be avoided as long as equivalents exist in Latin or Greek or can be constructed by combining word elements from these two languages	Recommendation 6(3)	Bartonella bettongiae
Pelagibacter ubique <sup>a</sup>	The epithet is an adverb and not an adjective or a noun	Rule 12c	Pelagibacter communis
Electronema palustris	The gender of the adjective used as a specific epithet does not agree with the gender of the genus name	Rule 12c(1)	Electronema palustre
Roseovibrio tepidum	The gender of the adjective used as specific epithet does not agree with the gender of the genus name	Rule 12c(1)	Roseovibrio tepidus
Roseilinea gracile	The gender of the adjective used as specific epithet does not agree with the gender of the genus name	Rule 12c(1)	Roseilinea gracilis
Ovatusbacter	The connecting vowel is -i- when the preceding word element is of Latin origin	Appendix 9	Ovatibacter
Liberobacter	The connecting vowel is -i- when the preceding word element is of Latin origin	Appendix 9	Liberibacter
Abysso bacteria	The connecting vowel is -o- when the preceding word element is of Greek origin	Appendix 9	Abysso bacteria
Altarchaeum	A connecting vowel is dropped when the following word element starts with a vowel	Appendix 9	Altarchaeum
Magnetoovum	A connecting vowel is dropped when the following word element starts with a vowel	Appendix 9	Magnetoovum

<sup>a</sup>Pure cultures of ‘*Candidatus* Pelagibacter ubique’ have been obtained [23], but no type strain was designated and the organism is not yet available from culture collections.

the nomenclatural type of a phylum must be a genus. Compiling a curated list of all proposed names of ‘*Candidatus*’ phyla will be a task for the future.

### ‘*Candidatus*’ names in need of nomenclatural quality control

The ICNP [2] and its orthography appendix (Appendix 9) [8] contain the rules and recommendations for the correct formation of names of species, genera and higher taxa based on the Linnaean binomial nomenclature system. To assist authors who are not competent in Latin, classical Greek and nomenclature Latin to propose correctly formed names for newly discovered prokaryotes, practical guides are readily available [19–22].

For naming new taxa of cultivated bacteria and archaea, there is a well-functioning and reliable quality control mechanism to ensure that validly published names meet the requirements set by the rules of the ICNP. The nomenclature reviewers of the IJSEM check every name of newly described prokaryotes in the journal, and if necessary, they suggest corrections to the authors to be implemented before papers are accepted for publication. The editors of several other journals also consult nomenclature experts to review newly proposed

names. The list editors of the IJSEM check the new prokaryotic names published in the journal once more when preparing the monthly Notification Lists, as well as names published in other journals when these are submitted for validation in the bimonthly Validation Lists.

Unfortunately, nomenclatural quality control of ‘*Candidatus*’ names proposed has been largely absent. A survey published in 2017, encompassing ~400 ‘*Candidatus*’ names found in the literature, showed that 120 contravened the rules of the ICNP or were otherwise problematic. A widely known example of an incorrectly formed ‘*Candidatus*’ name is *Pelagibacter ubique*, the famous marine ‘SAR-11’ phylotype [23]. The Latin word *ubique* is an adverb, and adverbs do not qualify as specific epithets according to Rule 12c of the ICNP. The result of the survey led to a plea for linguistic accuracy also for ‘*Candidatus*’ taxa, and well-formed alternatives were proposed for many problematic names [10]. The systematic search for ‘*Candidatus*’ names in the literature that led to the publication of the first lists of ‘*Candidatus*’ names in the IJSEM [11,14] confirmed the earlier picture: more than one-quarter of all names needed to be corrected.

Table 2 gives representative examples of the kinds of errors made in published ‘*Candidatus*’ names, together with proposals from nomenclature experts on how to improve them. Alternatives may be proposed by the authors of the original name,

and this has been done in at least one case already. In List no. I [11], change of the genus name *Rohrkolberia* (derived from the German word *Rohrkolben*, cattail, the Latin genus name being *Typha*) to *Typhincola* (a dweller of *Typha*) was suggested by nomenclature experts, but in a recent publication, the name '*Candidatus Symbiopectobacterium*' was proposed instead [24].

The publication of so many malformed '*Candidatus*' names once more shows the importance of nomenclatural quality control. The number of microbiologists who have an interest in nomenclature, as well as sufficient knowledge of Latin and Greek is very small. Still, there is no reason why some limited linguistic skills should not be part of the education of any microbial taxonomist. A few years ago, the list editors of the IJSEM called upon microbiologists with basic knowledge of Latin and Greek to join the 'nomenclature quality control' team [25]. Still, the question remains whether, in the future also there will be sufficient microbiologists with appropriate linguistic skills who are able and willing to check the proposed names in the descriptions of bacterial and archaeal taxa, cultivated, as well as uncultivated, which are being published at an ever-increasing rate.

### A separate code of nomenclature for the uncultivated majority of prokaryotes?

The ever-increasing numbers of sequenced genomes derived from uncultivated bacteria and archaea threaten to overwhelm the existing nomenclature framework that is mainly based on cultivated and well-characterized isolates. The challenge ahead is to incorporate the uncultivated microbial majority within a systematic taxonomic framework and adapt existing nomenclatural codes to scale to the task [26].

One approach proposed in the past years is to approve gene and genome sequences as type material for the valid publication of names under the rules of the ICNP. Based on this proposal, the code should be modified to allow gene sequences to serve as the type material for genera in the absence of cultivated species [27,28]. If gene sequences are accepted as suitable type material for the description of prokaryotic species, many taxa currently designated '*Candidatus*' will fulfil all the requirements in the ICNP for priority [29].

The questions of whether gene sequences may serve as type material and whether names of '*Candidatus*' taxa should have priority in the prokaryotic nomenclature have been extensively discussed in the past years [30–34]. These discussions led to a 'consensus statement', signed by a large number of colleagues, in which two potential paths were proposed to solve the nomenclatural conundrum. One option is the adoption of

previously proposed modifications to the ICNP to allow DNA sequences as acceptable type material; the other option is to create a separate nomenclatural code for uncultivated Archaea and Bacteria that could eventually be merged with the ICNP in the future [35].

In 2020, the ICSP rejected the proposals to amend the ICNP to allow the use of genome and gene sequence data as types [36]. Thus, the priority rules of the ICNP do not currently apply to gene sequence data and '*Candidatus*' taxa. Following this decision by the ICSP, a new, alternative code for naming uncultivated microorganisms (the "International Code of Nomenclature of Prokaryotes Described from Sequence Data" - the 'SeqCode') is now being developed (see <https://www.isme-microbes.org/seqcode-initiative>; accessed 24 May 2021). The rules for naming uncultivated prokaryotes under the proposed 'SeqCode' are similar to those of the ICNP and are based on Latin and Greek word elements while adopting the orthography rules of the ICNP. Thus, implementation of the 'SeqCode' will still require the same kind of nomenclature quality control as applied today to names validly published under the provisions of the ICNP and now also implemented in the published lists of '*Candidatus*' taxa [11,14].

### Automated approaches to the creation of new names of prokaryotes and descriptions of cultivated, as well as non-cultivated taxa

With the advancement of new technologies of rapid sequencing of metagenomes and single-cell genomes, as well as cultivation-dependent 'culturomics' techniques, there is an urgent need to create new names for all newly discovered taxa based on Linnaean binomial nomenclature. Two recent papers analyzing genomes from the chicken caecum exemplify this. In the first article, 42 new '*Candidatus*' genera and 60 new '*Candidatus*' species were described [37]. The second paper, providing a comprehensive genomic blueprint of the chicken gut microbiome, contains descriptions of one new '*Candidatus*' family, 158 '*Candidatus*' genera and 657 '*Candidatus*' species, based on an analysis of over 5500 metagenome-assembled genomes. In addition, 41 novel species were described in this paper following massive cultivation attempts [15].

In the latter paper, for the first time, an automated approach for the formation of new generic names and specific epithets was used, employing combinatorial concatenation of roots from Latin and Greek to create linguistically correct names. This approach should facilitate the generation of the next million or more names for Bacteria and Archaea. For achieving this, a Python program was written, named GAN – an acronym of the

“Great Automatic Nomenclator.” This name paraphrases the “Great Automatic Grammatizator,” the title of a short story written by Roald Dahl in 1953, describing a computer that, with the proper input information, could write short stories and novels [38]. For GAN, the input information consists of lists of Latin, Neo-Latin and Greek word stems that can be combined to yield correctly formed names, generally consisting of three- or four word elements. The program automatically adjusts the proper use of connecting vowels (-i- after a Latin word element, -o- after a Greek word element, no connecting vowel if the following word element begins with a vowel, to comply with the demands of the orthography appendix of the ICNP) and the gender of adjectives used as specific epithet, to comply with Rule 12c(1). This scalable combinatorial approach provides a stable, clear and memorable nomenclature for novel species, as shown by the hundreds of well-formed Latin binomials generated by GAN and already used in the nomenclature of both cultivated prokaryotes and ‘*Candidatus*’ taxa. The software is freely available (<https://github.com/telatin/gan>; accessed 24 May 2021) [39].

Another step forward toward the automated description and naming of new prokaryotes, cultivated, as well as uncultivated, is the automated generation of ‘protologue’ descriptions using genome sequences as the input information. Protologger is a bioinformatic tool that automatically generates all the necessary readouts for writing a detailed protologue. The program produces taxonomic outputs, including functional and ecological features and analysis of the 16S rRNA gene and genome sequences. Thus, it substantially reduces the time needed to gather the information necessary for describing novel taxa. It is linked to the GAN nomenclature program described above. Protologger is open source; hence, all scripts and datasets are available, along with a webserver at [www.protologger.de](http://www.protologger.de) (accessed 24 May 2021) [40].

### ‘*Candidatus* names’ – ‘the bold and the beautiful’

Automated creation of names as pioneered by the GAN program unavoidably leads to the creation of adequate and well-formed but arguably uninteresting and dull names. The same can be said for ‘geographical’ specific epithets, formed by adding the ending *-ensis* (masculine, feminine) or *-ense* (neuter) to the name of the geographical location from which an organism or a gene sequence was retrieved. The late Hans Georg Trüper termed this ‘localimania’ [41]. This way of forming correct names is widely practised, and it is in agreement with the Recommendation 12c(1) of the ICNP (“Choose

a specific epithet that, in general, gives some indication of a property or of the source of the species”). Some colleagues encourage it [42].

A recent paper entitled “Prokaryotic names: the bold and the beautiful” [43] shows that many authors have exploited the opportunities given by the rules of the ICNP to create names in more original ways. The article lists numerous examples of names of prokaryotic genera and species for which the authors proposed intriguing names intended to arouse curiosity and thus increase the interest in biological nomenclature. One of the examples given was the name of a ‘*Candidatus*’ species, ‘*Candidatus Desulforudis audaxviator*’ [44]. The etymology of the epithet is as follows: L. masc. adj. *audax* daring, courageous; L. masc. n. *viator* traveler; N.L. masc. n. *audaxviator* a courageous traveller. This refers to a text in dog Latin, encrypted in runic script, found in Jules Verne’s “Voyage au centre de la Terre”: *In Sneffels Yoculis craterem kem delibat umbra Scartaris Julii intra calendas descende, audas viator, et terrestre centrum attinges ...* (Descend into the crater of Yocul of Sneffels, which the shade of Scartaris caresses, before the kalends of July, audacious traveler and you will reach the center of the earth ...). The organism was recently isolated in axenic culture [45], but the name has not yet been submitted for validation.

Here are a few other interesting ‘*Candidatus*’ names found in the literature:

- *Cellulosimonas argentiregionis* (L. neut. n. *argentum* silver; L. fem. n. *regio* territory; N.L. gen. n. *argentiregionis* of Nevada, the silver state) [46].
- *Umbricyclops* (L. fem. n. *umbra* shadow; L. masc. n. *cyclops* (from Gr. masc. n. *kyklops* round eye; Cyclops) a cyclops; N.L. masc. n. *Umbricyclops* a round-eye (based on the ovoid shape of the cells) living in the shade) [47].
- *Vampirococcus lugosii* (N.L. gen. n. *lugosii* named after Bela Lugosi (1882-1956), who played the role of the vampire in the iconic 1931 film *Dracula*) [48]. Such a name may contravene the recommendation to refrain from naming organisms after persons, quite unconnected with bacteriology or at least with natural science [2], but is in my opinion appropriate, based on the properties of the bacterium.

### Conclusions

The current system of naming uncultivated prokaryotes using the ‘*Candidatus*’ concept as proposed in the middle of the 1990s and outlined in Appendix 11 of the ICNP is still working well. The ‘*Candidatus*’ status is not formally included in the rules of the ICNP, and ‘*Candidatus*’ names are therefore not validly

published and have no priority in the nomenclature (see section 6: “A separate code of nomenclature for the uncultivated majority of prokaryotes?”).

Appendix 11 needs updating in the upcoming revision of the ICNP [7] to ensure that the same rules for the naming of cultivated prokaryotes should also apply to the naming of uncultivated taxa, based on Latin, Greek and the orthography rules given in Appendix 9. Correctly formed names of ‘*Candidatus*’ taxa can then be used unchanged when the organisms are later cultivated, and the provisional names can be validly published under the rules of the ICNP.

As of April 2021, names of 57 taxa originally described as ‘*Candidatus*’ were later validly or effectively published after the organisms were cultivated. In some cases, the names had to be corrected to comply with the rules of the ICNP and its orthography appendix. This is in part due to the fact that the rules of the Code do not formally apply to ‘*Candidatus*’ names. However, with the publication of annotated lists of ‘*Candidatus*’ names in the IJSEM, including alternatives proposed for problematic names, the names listed are ready to be validated when the taxa can be cultivated.

The original ‘*Candidatus*’ proposal [5] and Appendix 11 of the ICNP call for a list in the form of a codified record of organisms of the status ‘*Candidatus*’, to be kept by the Judicial Commission of the ICSP in cooperation with the Editorial Board of the IJSEM, to be published in that journal at appropriate intervals. However, registration of ‘*Candidatus*’ names started only recently [11,13,14].

Now we have well-ordered and curated lists of names of the ‘*Candidatus*’ taxa, with proper quality control of the names proposed, it can be argued that there is no need for a separate code of nomenclature for the uncultivated majority of prokaryotes. In particular, the establishment of the ‘SeqCode’ to be independent of the ICNP will cause unnecessary confusion. A single system with a single set of rules is better suited to consistent nomenclature quality control. This also agrees with the spirit of General Consideration 5 of the ICNP that begins with the words: “This Code of Nomenclature of Prokaryotes applies to all Prokaryotes.”.

The use of new sequencing technologies and bioinformatics tools to interpret the wealth of sequence material has already led to proposals of very large numbers of new taxa to be named as ‘*Candidatus*’ taxa. Thus far, the list editors of the IJSEM have managed to cope with the workload involved in the preparation of curated ‘*Candidatus*’ lists [11,14]. However, this may no longer be the case in the near future. The publication of a paper on the chicken intestine metagenome that included proposals for 816 new ‘*Candidatus*’ names [15] is an example of what we may expect more extensively in the coming years.

The development of computer programs such as GAN [39] that can automatically generate large numbers of correctly formed names that meet the requirements of the rules of the ICNP and its orthography appendix will be essential to reduce the need for manual quality control of every single name proposed. Combined with programs such as Protologger [40] that produce descriptions of ‘*Candidatus*’ taxa directly from the genome sequence, at least part of the quality control of newly proposed names and taxa descriptions can be taken over by machines.

Computer-generated names, formed by combining Latin, Greek and Neo-Latin word elements in the correct way, satisfy the demands of the ICNP. When there is a need to name dozens or hundreds of new ‘*Candidatus*’ genera and species, searching for exciting names is no longer feasible. As stated before, “it takes some effort and it requires proper use of the lexicon of Classical Greek and Latin, as well as an understanding of the Code and the guidelines of its orthography appendix. Creation of attractive names will boost the general interest in prokaryotic nomenclature.” [43]

---

## Transparency declaration

The author declares that the content of this article expresses his personal views, unconnected to his functions of Executive Secretary of the International Committee on Systematics of Prokaryotes, Editor-in-Chief of the International Code of Nomenclature of Prokaryotes, List Editor of the International Journal of Systematic and Evolutionary Microbiology, or any other function.

---

## Credit author statement

Not applicable.

---

## Acknowledgements

I thank Mark J. Pallen for reading and commenting on drafts of the manuscript.

---

## References

- [1] Lapage SP, Sneath PHA, Lessel EF, Skerman VBD, Seeliger HPR, Clark WA. International code of nomenclature of bacteria. Bacteriological code, 1990 revision. Washington (DC): ASM Press; 1992.

- [2] Parker CT, Tindall BJ, Garrity GM. International code of nomenclature of prokaryotes. Prokaryotic code (2008 revision). *Int J Syst Evol Microbiol* 2019;69:S1–111.
- [3] Murray RG, Schleifer KH. Taxonomic notes: a proposal for recording the properties of putative taxa of prokaryotes. *Int J Syst Bacteriol* 1994;44:174–6.
- [4] Frederiksen W. Judicial commission of the international committee on systematic bacteriology. Minutes of the meetings, 2 and 6 July 1994, Prague, Czech republic. *Int J Syst Bacteriol* 1995;45:195–6.
- [5] Murray RG, Stackebrandt E. Taxonomic note: implementation of the provisional status *Candidatus* for incompletely described prokaryotes. *Int J Syst Bacteriol* 1995;45:186–7.
- [6] Konstantinidis KT, Rosselló-Móra R. Classifying the uncultivated microbial majority: a place for metagenomic data in the *Candidatus* proposal. *Syst Appl Microbiol* 2015;38:223–30.
- [7] Oren A, Sutcliffe I, Moore ERB, Arahal DA, Rosselló-Móra R. Preparing a revision of the international code of nomenclature of prokaryotes. *Int J Syst Evol Microbiol* 2021;71:004598.
- [8] Trüper HG, Euzéby JP. International code of nomenclature of prokaryotes. Appendix 9: orthography. *Int J Syst Bacteriol* 2009;59:2107–13.
- [9] Fagen JR, Leonard MT, Coyle JF, McCullough CM, Davis-Richardson AG, Davis MJ, et al. *Liberibacter crescens* gen. nov., sp. nov., the first cultured member of the genus *Liberibacter*. *Int J Syst Evol Microbiol* 2014;64:2461–6.
- [10] Oren A. A plea for linguistic accuracy – also for *Candidatus* taxa. *Int J Syst Evol Microbiol* 2017;67:1085–94.
- [11] Oren A, Garrity GM, Parker CT, Chuvochina M, Trujillo ME. Lists of names of prokaryotic *Candidatus* taxa. *Candidatus* List no. 1. *Int J Syst Evol Microbiol* 2020;70:3956–4042.
- [12] Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. *Int J Syst Evol Microbiol* 2020;70:5607–12.
- [13] Oren A, Garrity GM, Trujillo ME. Registration of names of prokaryotic *Candidatus* taxa in the IJSEM. *Int J Syst Evol Microbiol* 2020;70:3955.
- [14] Oren A, Garrity GM. Lists of names of prokaryotic *Candidatus* taxa. *Candidatus* List no. 2. *Int J Syst Evol Microbiol* 2021;71:004671.
- [15] Gilroy R, Ravi A, Getino M, Pursley I, Horton DL, Alikhan N-F, et al. Extensive microbial diversity within the chicken gut microbiome revealed by metagenomics and culture. *PeerJ* 2021. <https://doi.org/10.7717/peerj.10941>.
- [16] Oren A, Arahal DR, Rosselló-Móra R, Sutcliffe IC, Moore EJB. Emendation of rules 5b, 8, 15 and 22 of the international code of nomenclature of prokaryotes to include the rank of phylum. *Int J Syst Evol Microbiol* 2021;71:004851.
- [17] Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, et al. A new view of the tree of life. *Nat Microbiol* 2016;1:16048.
- [18] Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 2013;499:431–7.
- [19] MacAdoo TO. Nomenclatural literacy. In: Goodfellow M, O'Donnell AG, editors. *Handbook of new bacterial systematics*. London: Academic Press; 1993. p. 339–58.
- [20] Oren A. How to name new taxa of prokaryotes?. In: Rainey FA, Oren A, editors. *Taxonomy of prokaryotes - methods in Microbiology*, vol. 38. Amsterdam: Elsevier/Academic Press; 2011. p. 438–63.
- [21] Oren A. How to name new taxa of archaea and bacteria. In: *Bergey's manual of systematics of archaea and bacteria*. John Wiley & Sons, in association with Bergey's Manual Trust; 2019.
- [22] Trüper HG. How to name a prokaryote? Etymological considerations, proposals and practical advice in prokaryote nomenclature. *FEMS Microbiol Rev* 1999;23:231–49.
- [23] Rappé MS, Connon SA, Vergin KL, Giovannoni SJ. Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* 2002;418:30–3.
- [24] Martinson VG, Gawryluk RMR, Gowen BE, Curtis CI, Jaenike J, Perlman SJ. Multiple origins of obligate nematode and insect symbionts by a clade of bacteria closely related to plant pathogens. *Proc Natl Acad Sci USA* 2020;117:31979–86.
- [25] Oren A, Schink B, Garrity GM. Wanted: microbiologists with basic knowledge of Latin and Greek to join our 'nomenclature quality control' team. *Int J Syst Evol Microbiol* 2015;65:3761–2.
- [26] Hugenholtz P, Chuvochina M, Oren A, Parks DH, Soo RM. Microbial taxonomy and nomenclature in the age of big sequence data. *ISME J* 2021. <https://doi.org/10.1038/s41396-021-00941-x>.
- [27] Whitman WB. Genome sequences as the type material for taxonomic descriptions of prokaryotes. *Syst Appl Microbiol* 2015;38:217–22.
- [28] Whitman WB. Modest proposals to expand the type material for naming of prokaryotes. *Int J Syst Evol Microbiol* 2016;66:2108–12.
- [29] Whitman WB, Sutcliffe IC, Rossello-Mora R. Proposal for changes in the international code of nomenclature of prokaryotes: granting priority to *Candidatus* names. *Int J Syst Evol Microbiol* 2019;69:2174–5.
- [30] Konstantinidis KT, Rosselló-Móra R, Amann R. Uncultivated microbes in need of their own taxonomy. *ISME J* 2017;11:2399–406.
- [31] Konstantinidis KT, Rosselló-Móra R, Amann R. Reply to the commentary "Uncultivated microbes – in need of their own nomenclature?" *ISME J* 2018;12:653–4.
- [32] Oren A, Garrity GM. Uncultivated microbes – in need of their own nomenclature? *ISME J* 2018;12:309–11.
- [33] Rosselló-Móra R, Whitman WB. Dialogue on the nomenclature and classification of prokaryotes. *Syst Appl Microbiol* 2019;42:5–14.
- [34] Rossello-Mora R, Konstantinidis KT, Sutcliffe I, Whitman W. Opinion: response to 4 prokaryotes. *Syst Appl Microbiol* 2020;43:126070.
- [35] Murray AE, Freudenstein J, Gribaldo S, Hatzepichler R, Hugenholtz P, Kämpfer P, et al. Roadmap for naming uncultivated archaea and bacteria. *Nat Microbiol* 2020;5:987–94.
- [36] Sutcliffe IC, Dijkshoorn L, Whitman WB, on behalf of the ICSP Executive Board. 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* 2020;70:4416–7.
- [37] Glendinning L, Stewart RD, Pallen MJ, Watson KA, Watson M. Author Correction: assembly of hundreds of novel bacterial genomes from the chicken caecum Genome. *Biol* 2021;22:60.
- [38] Dahl R. The great automatic grammatizator. In: *Someone like you*. Alfred A. Knopf; 1953.
- [39] Pallen MJ, Telatin A, Oren A. The next million bacterial names. *Trends Microbiol* 2021;29:289–98.
- [40] Hitch TCA, Riedel T, Oren A, Overmann J, Lawley T, Clavel T. Automated analysis of genomic sequences facilitates high-throughput and comprehensive description of prokaryotes. *ISME Commun* 2021;1:16.
- [41] Trüper HG. Is 'localimania' becoming a fashion for prokaryote taxonomists? *Int J Syst Evol Microbiol* 2005;55:753.
- [42] Lagier JC, Bilen M, Cadoret F, Drancourt M, Fournier E, La Scola B, et al. Naming microorganisms: the contribution of the IHU Méditerranée Infection, Marseille, France. *New Microbes. New Infect* 2018;26:S89–95.
- [43] Oren A. Prokaryotic names: the bold and the beautiful. *FEMS Microbiol Lett* 2020;367:fnaa096.
- [44] Chivian D, Brodie EL, Alm EJ, Culley DE, Dehal PS, DeSantis TZ, et al. Environmental genomics reveals a single-species ecosystem deep within Earth. *Science* 2008;322:275–8.
- [45] Karnachuk OV, Frank YA, Lukina AP, Kadnikov VV, Beletsky AV, Mardanov AV, et al. Domestication of previously uncultivated



- Candidatus* Desulforudis audaxviator from a deep aquifer in Siberia sheds light on its physiology and evolution. ISME J 2019;13:1947–59.
- [46] Doud DFR, Bowers RM, Schulz F, De Raad M, Deng K, Tarver A, et al. Function-driven single-cell genomics uncovers cellulose-degrading bacteria from the rare biosphere. ISME J 2020;14:659–75.
- [47] Mehrshad M, Salcher MM, Okazaki Y, Nakano S-i, Šimek K, Andrei A-S, et al. Hidden in plain sight – highly abundant and diverse planktonic freshwater *Chloroflexi*. Microbiome 2018;6:176.
- [48] Moreira D, Zivanovic Y, López-Archilla AI, Iniesto M, López-García P. Reductive evolution and unique predatory mode in the CPR bacterium *Vampirococcus lugosii*. Nat Commun 2021;12:2454.