

# Minimal standards for the description of new genera and species of rhizobia and agrobacteria

Philippe M. de Lajudie,<sup>1,\*</sup> Mitchell Andrews,<sup>2</sup> Julie Ardley,<sup>3</sup> Bertrand Eardly,<sup>4</sup> Estelle Jumas-Bilak,<sup>5</sup> Nemanja Kuzmanović,<sup>6</sup> Florent Lassalle,<sup>7</sup> Kristina Lindström,<sup>8</sup> Ridha Mhamdi,<sup>9</sup> Esperanza Martínez-Romero,<sup>10</sup> Lionel Moulin,<sup>11</sup> Seyed Abdollah Mousavi,<sup>8</sup> Xavier Nesme,<sup>12</sup> Alvaro Peix,<sup>13</sup> Joanna Putawska,<sup>14</sup> Emma Steenkamp,<sup>15</sup> Tomasz Stępkowski,<sup>16</sup> Chang-Fu Tian,<sup>17</sup> Pablo Vinuesa,<sup>10</sup> Gehong Wei,<sup>18</sup> Anne Willems,<sup>19</sup> Jerri Zilli<sup>20</sup> and Peter Young<sup>21</sup>

## Abstract

Herein the members of the Subcommittee on Taxonomy of Rhizobia and Agrobacteria of the International Committee on Systematics of Prokaryotes review recent developments in rhizobial and agrobacterial taxonomy and propose updated minimal standards for the description of new species (and genera) in these groups. The essential requirements (minimal standards) for description of a new species are (1) a genome sequence of at least the proposed type strain and (2) evidence for differentiation from other species based on genome sequence comparisons. It is also recommended that (3) genetic variation within the species is documented with sequence data from several clearly different strains and (4) phenotypic features are described, and their variation documented with data from a relevant set of representative strains. Furthermore, it is encouraged that information is provided on (5) nodulation or pathogenicity phenotypes, as appropriate, with relevant gene sequences. These guidelines supplement the current rules of general bacterial taxonomy, which require (6) a name that conforms to the International Code of Nomenclature of Prokaryotes, (7) validation of the name by publication either directly in the *International Journal of Systematic and Evolutionary Microbiology* or in a validation list when published elsewhere, and (8) deposition of the type strain in two international culture collections in separate countries.

## INTRODUCTION

The general term ‘rhizobia’ covers all bacteria that induce nodule formation on the roots, or occasionally stems, of a legume host (plant family *Fabaceae*), where they fix atmospheric nitrogen (N<sub>2</sub>) [1]. Nitrogen fixation by rhizobia in association with legumes can be a major input of nitrogen

(N) into natural and agricultural ecosystems and can reduce the need for chemical N fertilizer. There has been extensive research on the diversity of rhizobia over the past 25 years, with the result that many new species and genera have been formally described. Taxonomically, rhizobia represent several lineages within the Proteobacteria,

**Author affiliations:** <sup>1</sup>IRD, University of Montpellier, CIRAD, INRA, SupAgro, LSTM, Montpellier, France; <sup>2</sup>Faculty of Agriculture and Life Sciences, Lincoln University, Lincoln 7647, New Zealand; <sup>3</sup>School of Veterinary and Life Sciences, Murdoch University, Murdoch, Australia; <sup>4</sup>Penn State Berks College, Reading, PA 19610, USA; <sup>5</sup>UMR 5569, Department of Microbiology, Faculty of Pharmacy, University of Montpellier, France; <sup>6</sup>Julius Kühn-Institut, Federal Research Centre for Cultivated Plants, Institute for Epidemiology and Pathogen Diagnostics, Messeweg 11/12, 38104 Braunschweig, Germany; <sup>7</sup>Department of Infectious Disease Epidemiology - MRC Centre for Outbreak Analysis and Modelling, St Mary's Hospital, Praed Street, London W2 1NY, UK; <sup>8</sup>Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki FI-00014, Finland; <sup>9</sup>Centre of Biotechnology of Borj-Cedria, BP 901 Hammam-lif 2050, Tunisia; <sup>10</sup>Centro de Ciencias Genómicas, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, Mexico; <sup>11</sup>IRD, CIRAD, University of Montpellier, IPME, Montpellier, France; <sup>12</sup>LEM, UCBL, CNRS, INRA, Univ Lyon, Villeurbanne, France; <sup>13</sup>Instituto de Recursos Naturales y Agrobiología, IRNASA-CSIC, c/Cordel de Merinas 40-52, 37008 Salamanca, Spain; <sup>14</sup>Department of Phytopathology, Research Institute of Horticulture, ul. Konstytucji 3 Maja 1/3, 96-100 Skierniewice, Poland; <sup>15</sup>Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria 0002, South Africa; <sup>16</sup>Autonomous Department of Microbial Biology, Faculty of Agriculture and Biology, Warsaw University of Life Sciences (SGGW), Nowoursynowska 159, 02-776 Warsaw, Poland; <sup>17</sup>State Key Laboratory of Agrobiotechnology, MOA Key Laboratory of Soil Microbiology, Rhizobium Research Center, College of Biological Sciences, China Agricultural University, 100193, Beijing, PR China; <sup>18</sup>Northwest A&F University, Yangling, Shaanxi, PR China; <sup>19</sup>Department Biochemistry and Microbiology, Lab. Microbiology, Ghent University, Belgium; <sup>20</sup>Embrapa Agrobiologia, BR 465 km 07, Seropédica, Rio de Janeiro, Brazil, 23891-000, Brazil; <sup>21</sup>Department of Biology, University of York, York YO10 5DD, UK.

\*Correspondence: Philippe M. de Lajudie, philippe.lajudie@ird.fr

**Keywords:** Minimal standards for taxonomy; rhizobia; agrobacteria; genome-based taxonomy.

**Abbreviations:** ANI, average nucleotide identity; dDDH, digital DDH; DDH, DNA–DNA hybridization; GGDC, Genome-to-Genome Distance Calculator; HGT, horizontal gene transfer; ICSP, International Committee on Systematics of Prokaryotes; MLSA, multilocus sequence analysis; NCBI, National Center for Biotechnology Information; NGS, next generation sequencing; OGRI, overall genome relatedness index. One supplementary table is available with the online version of this article.

and therefore the term ‘rhizobia’ does not represent a single taxon, but refers to a polyphyletic assemblage of bacterial lineages having similar functions [1]. Most known rhizobia belong to alphaproteobacterial genera in the families *Rhizobiaceae* (*Rhizobium*, *Ensifer* (syn. *Sinorhizobium*), *Allorhizobium*, *Pararhizobium*, *Neorhizobium*, *Shinella*), *Phyllobacteriaceae* (*Mesorhizobium*, *Aminobacter*, *Phyllobacterium*), *Brucellaceae* (*Ochrobactrum*), *Methylobacteriaceae* (*Methylobacterium*, *Microvirga*), *Bradyrhizobiaceae* (*Bradyrhizobium*), *Xanthobacteraceae* (*Azorhizobium*) and *Hyphomicrobiaceae* (*Devosia*), but some belong to betaproteobacterial genera in the family *Burkholderiaceae* (*Paraburkholderia*, *Cupriavidus*, *Trinickia*) [2, 3]. For a nodule isolate to be considered a rhizobium, evidence must be provided for its ability to produce nodules on a legume, because non-rhizobial bacteria are common in nodules [4–6]. The symbiosis (sym) genes that enable rhizobial strains to induce N<sub>2</sub>-fixing nodules on legumes can be gained or lost as they are often encoded on dispensable but transferable plasmids or symbiosis islands [7, 8].

The general term ‘agrobacteria’ covers all bacteria that incite – or potentially may incite – the formation of galls or root hypertrophies (crown gall or hairy root disease) by T-DNA transfer to any of a wide range of plant species, including many important agricultural crops [1]. To date, all recognized agrobacteria belong to the family *Rhizobiaceae* (*Agrobacterium* spp., *Rhizobium rhizogenes* [9], *Allorhizobium vitis* [9] and *Rhizobium tumorigenes* [10]) in the class *Alphaproteobacteria*. Most of the genes involved in crown gall and hairy root diseases are borne on large conjugative plasmids, the Ti (tumour-inducing) and Ri (root-inducing) plasmids, respectively. These elements are transmissible and, as for rhizobial symbiosis genes, can be gained or lost by strains. The recently described species *Pararhizobium polonicum* is not an agrobacterium in this sense since it is not pathogenic, although it was isolated from crown gall tumours on various plants and harbours a plasmid encoding catabolism of opines, which are produced in tumours induced by tumorigenic strains [11]. As this example shows, not all bacteria isolated from galls are agrobacteria: it is necessary to demonstrate pathogenic capability.

The Subcommittee, initially created by the International Committee on Systematics of Prokaryotes (ICSP) to overview the taxonomy of *Rhizobium* and *Agrobacterium*, has expanded its scope to cover the functional and polyphyletic groups collectively named ‘rhizobia’ and ‘agrobacteria’ [1] in order to continue to serve the scientific communities that study these organisms. Accordingly, its name is now the ICSP Subcommittee on the Taxonomy of Rhizobia and Agrobacteria [12]. Since the defining features of rhizobia and agrobacteria are encoded by genes that can be gained or lost, and are not therefore directly useful for defining species and genera, the Subcommittee also covers species and strains that are in the relevant genera but do not (currently) have symbiotic or pathogenic phenotypes.

One of the missions of the ICSP subcommittees is to recommend minimal standards for the description of new species [13]. Such minimal standards were proposed for rhizobial species description by the ICSP Subcommittee on the Taxonomy of *Rhizobium* and *Agrobacterium* in 1991 [14]. Since then, further comments and recommendations regarding new findings, technical developments and good taxonomic practice have been published regularly as minutes of the meetings of the ICSP Subcommittee [1, 15–21] and are available on the web page of the Subcommittee ([sites.google.com/view/taxonomyagrorhizo](https://sites.google.com/view/taxonomyagrorhizo)).

In this paper we, the members of the Subcommittee, briefly review recent developments in the taxonomy of rhizobia and agrobacteria, and focus on guidelines to communicate what we currently consider to be the minimal evidence that researchers need to assemble in a taxonomic study involving the description of new rhizobial or agrobacterial genera and species. We agree with Whitman [22] that, in line with the spirit of the International Code of Nomenclature of Prokaryotes, standards must be kept minimal so that ‘the process of naming a new taxon should be straightforward and easy’, and should be ‘prepared for all users of systematics, which include biologists with a wide range of interests, and not just experts in taxonomy’. These guidelines consider current recommendations for general bacterial taxonomy and their application to rhizobia and agrobacteria, in addition to specific considerations/features that are useful but not mandatory for description of rhizobial and agrobacterial taxa.

## THE NEED FOR REVISED GUIDELINES FOR RHIZOBIAL AND AGROBACTERIAL SPECIES DESCRIPTIONS

### Species-level bacterial assignments in general taxonomy of prokaryotes

For the past 30 years, guidelines for the general description of new bacterial species have recommended DNA–DNA hybridization (DDH) [23, 24], 16S rRNA gene sequencing and the polyphasic approach [25, 26]. Guidelines for formal bacterial species description and an extensive review of established methods were published in 2010 [27]. However, in recent years the limitations of these taxonomic practices have become clear (for a review see [28]), and there is a need to revise the guidelines and to recommend that genomic information should provide the primary evidence for species affiliation [29, 30].

DDH has long been used as a proxy for whole-genome sequence information. In 1987, Wayne *et al.* [24] recommended it as ‘the best applicable procedure at the present time’. However, the use of DDH can no longer be recommended because genome sequences provide much better evidence for genomic comparison [31]. Genome-based measures of similarity, such as average amino acid identity [32], digital DDH (dDDH) [33, 34], MUM-distance index [35, 36] and average nucleotide identity (ANI) [37–39], render classical (wet-lab) DDH unnecessary and help describe

new species quickly and accurately [40]. Genome G+C content can also be determined easily and precisely from good-quality genome sequences, so sequence-derived G+C values are expected to replace laboratory determinations, which can diverge considerably from the true values [41]. Over time, it became widely accepted that, with the increased availability of genomic data, the traditional polyphasic bacterial taxonomy needed to be revised [28, 42–44], and several authors advocated using genomic-based species descriptions in the classification of prokaryotes [22, 45–49]. Indeed genome-based approaches can form the basis for a robust species description using open-access databases, as sequence data are fully portable and unambiguously comparable across locations and time. We agree that bacterial taxonomists ‘must urgently reconsider how to describe and name novel bacteria in the genomic era, using a full genome sequence and a minimal description of phenotypic characteristics as a basic, sufficient, cost-effective and appropriate biological identity card for a species description’ [44].

### Species definition in rhizobia and agrobacteria

Over the past three decades, the general approach for the determination of species boundaries within the rhizobia and agrobacteria, as in bacterial taxonomy more widely, has been polyphasic, attempting to identify natural groupings based on shared phenotypic and genotypic characters. Graham *et al.* [14] published recommendations for describing species of rhizobia using a polyphasic approach that involved the analysis of a range of phenotypic characters, along with 16S rRNA gene sequences and DDH values. At that time (1991), only four rhizobial genera existed (*Rhizobium*, *Bradyrhizobium*, *Azorhizobium* and *Sinorhizobium*), with a total of nine species. Agrobacteria were classified in five named species (*Agrobacterium tumefaciens*, *Agrobacterium radiobacter*, *Agrobacterium rhizogenes*, *Agrobacterium rubi* and *Agrobacterium vitis*). Twenty-seven years later, the number of rhizobial species is now more than 200, distributed in 18 genera. For some years, a proposal to merge the whole genus *Agrobacterium* into *Rhizobium* [50] was accepted by some taxonomists, though by few in the wider scientific community, but the separate identity of *Agrobacterium* is now firmly established [51, 52]. A review of the taxonomy of *Agrobacterium* in the genomic era was published recently [53]. Currently, seven *Agrobacterium* species are generally accepted (*Agrobacterium larrymoorei*, *Agrobacterium nepotum*, *Agrobacterium pusense*, *Agrobacterium radiobacter*, *Agrobacterium rubi*, *Agrobacterium salinitolerans*, *Agrobacterium skierniewicense*) and four more species have been described but their names are awaiting validation.

Over the period 1991–2018, strategies for the description of new rhizobial and agrobacterial species evolved and sets of chromosomal genes were sequenced for many species [54–70]. This generated an important database, which gradually served as a framework to identify novel rhizobial and agrobacterial species and to affiliate new isolates. In practice,

authors complemented the original polyphasic taxonomy with sequences of several single-copy housekeeping genes scattered in the genome (multi locus sequence analysis, MLSA). MLSA proved helpful to clarify a number of long-standing unsolved taxonomic situations in rhizobia and agrobacteria [9, 52, 71]. These genes provide higher phylogenetic resolution than the 16S rRNA gene, which is still useful for initial approximate phylogenetic placement [72]. Using several genes minimizes the impact of possible horizontal gene transfer (HGT). It has been shown that almost all genes – even the 16S rRNA gene – are to some extent horizontally transferred during evolution [71, 73, 74] – so the phylogeny of a single gene may not reflect the phylogeny of the species. In recent years many new rhizobial and agrobacterial species descriptions have included whole genome sequence information [11, 75–97], providing data that are quickly and cheaply generated, highly informative, accurate and reproducible, and conveniently storable and portable among laboratories. We emphasize that the current rhizobial and agrobacterial classification based on polyphasic studies (including DDH, 16S rRNA gene phylogeny and phenotypic features) is essentially corroborated by MLSA and genomic data. A shift to genome-based species descriptions will not therefore require major rewriting of current taxonomy, but it will make future taxonomic work simpler and more robust.

## SPECIFIC CONSIDERATIONS IN ESTABLISHING NEW GUIDELINES FOR DEFINING RHIZOBIAL AND AGROBACTERIAL SPECIES AND GENERA

### A minimal set of diverse strains for assessment of intraspecific diversity

Although single-strain species descriptions are currently published, the inclusion of multiple genetically distinct strains is considered good taxonomic practice [98]. Intraspecific diversity should be described using as many strains as possible, and it is up to the authors to provide convincing evidence that the strains are clearly different. A simple genome fingerprinting method (RAPD, rep/BOX-PCR) or MLSA (described in the following section), are possible approaches to distinguish non-clonal strains. If genomic information is provided for more than one strain, an ANI value significantly less than 100 % would demonstrate that they were different. An important benefit of studying several strains is that it allows a type strain to be chosen that shares most similarity with the other strains and is hence truly ‘typical’ of the new species that it represents.

### Description of intraspecific diversity with MLSA

The use of MLSA in bacterial species definition was strongly endorsed by an *ad hoc* committee of the ICSP in 2002 [26]. Now that whole genome sequences will be available for type strains, there is no need to select just a small number of genes for comparing type strains, but MLSA can still be valuable for exploring the genetic diversity within a

proposed species. MLSA is based on single-copy core genes, not accessory genes, located on the chromosome and not on plasmids or chromids [99]. In recent descriptions of new rhizobial and agrobacterial species, the genes most often used were *recA*, *atpD*, *glnII*, *rpoB*, *gyrB* and *dnaK* (Table S1, available in the online version of this article, which lists information provided in descriptions of rhizobial and agrobacterial species in years 2014–2018), with *recA*, *atpD* and *glnII* being most prevalent for rhizobia, and *recA*, *atpD* and *rpoB* for agrobacteria [9, 85, 100]. These genes were originally chosen because they were easily amenable to PCR amplification, but they are unlikely to be optimal. Indeed, Zhang *et al.* [101] recommended a different set of three housekeeping genes (*SMc00019*, *truA* and *thrA*) and Vinuesa *et al.* [49] have described a computational pipeline to select optimal markers for phylogenomics. It should also be noted that HGT of core genes can be expected within a species, so the phylogenies of different genes may not be congruent.

Some authors have calculated the ANI of the genes they have used for MLSA [102–104], and have used the abbreviation ‘ANI’ to describe this value. It should be emphasized that this is not the ANI metric, based on whole genome comparisons, that has been proposed as a universal tool for defining species boundaries [33, 34, 38]. The ANI of small sets of conserved genes will depend on the specific genes chosen, but may differ from the standard whole-genome ANI.

PCR-based methods that generate fragment size patterns, such as restriction fragment length polymorphism (RFLP) [105] and repetitive sequence-based PCR (repPCR) [106] can still be useful for exploring the diversity of large strain collections. However, the data they generate are not readily compared between studies, in contrast to the sequence data obtained by MLSA, which provide a common scale to compare intraspecific and interspecific distances.

### Definition of new species with genomic data

There is now general agreement in the bacterial systematics community that genome sequence data should be provided when describing a new species [35], and routine description of prokaryotic species based on their genomic sequence would fulfil all the requirements of the International Code of Nomenclature of Prokaryotes [13, 22].

Ormeño-Orrillo *et al.* [86] reconstructed a genome-based phylogeny with a sample of 113 genome sequences representing various genospecies in the *Rhizobiaceae*. This confirmed the overall MLSA-based phylogeny [9, 52], and proved highly decisive to remove some uncertainties between closely related species in *Rhizobium*, *Neorhizobium* and *Ensifer* (formerly *Sinorhizobium*).

We consider that it is mandatory, when describing a new agrobacterial or rhizobial species, to deposit the genome sequence of at least the type strain in a public database with open access within the International Nucleotide Sequence Database Collaboration (i.e. GenBank, ENA or DDBJ).

Authors should consider the minimal standards for the quality of genome data used for the taxonomy of prokaryotes established by Chun *et al.* [31]. Next-generation sequencing (NGS) is currently a cost-effective high-throughput DNA sequencing approach widely applicable to microbial taxonomy and within the reach of most laboratories [72]. The assemblies generated from the short reads tend to be fragmented, but the most important information for taxonomy will still be present provided the coverage is sufficient to ensure a reliable sequence for each gene. Chun *et al.* [31] detailed useful indices to check the quality of the genome sequence. Genome sequencing with short-read techniques does not allow the genome replicon structure to be resolved, so it may be useful to perform complementary pulse-field gel electrophoresis (PFGE) or Eckhart gel electrophoresis for this (see the subsection: Specific considerations in establishing new guidelines for defining rhizobial and agrobacterial species and genera, Genus-level assignments). As newer technologies emerge and become affordable, complete genome assemblies will doubtless become the expected standard.

Appropriate quantitative comparisons must be made between the genome(s) of the new species and genomes of the type strains of all closely related species. Authors should carefully check reference sequences in databases since they may harbour misleading uncertainties and errors, and several examples of divergent sequences for different subcultures of the same type strain have been pointed out [107, 108]. Issues that we are aware of are reported in the minutes of the Subcommittee [21]; recent minutes and the web page of the subcommittee ([sites.google.com/view/taxonomyagrorhizo](https://sites.google.com/view/taxonomyagrorhizo)) should be consulted.

Currently, the most widely used measures of between-strain genomic similarity are whole-genome ANI and digital DDH. There are several variants of the ANI approach [38], an online tool is available for dDDH [35], software is available for statistical analysis of the results [49], and faster comparison methods are being developed [109, 110]. A weakness of all these measures is that they are based on an ill-defined combination of core and accessory genes. Since accessory genomes vary among isolates within a species, it is the core genes that provide the clearest taxonomic information. New measures of genome similarity are continually being developed, including some that differentiate the core and accessory component of genome variation [111], and authors should be aware of ongoing advances in bioinformatics and use the most robust tools available.

Several studies have attempted to define thresholds in the use of overall genome relatedness indexes (OGRIs) in taxonomy [31]. Whole-genome ANI values of 95–96% have been proposed to be equivalent to 70% DDH and hence to delineate species [33, 38, 112]. There is no strict correlation between ANI and 16S rRNA gene similarity; in general, strains with less than 98.7% 16S rRNA sequence identity do not belong to the same species [107]. However the converse is not always true because different species may have

identical 16S rRNA gene sequences [113]. Thompson *et al.* [30] calibrated different metrics and concluded that strains from the same microbial species share >95 % average amino acid identity (AAI) and ANI, >95 % identity based on multiple genes, a Karlin genomic signature <10, and >70 % *in silico* genome-to-genome hybridization (GGDH) similarity. We endorse the recommendation of Chun *et al.* [31] that OGRI values between closely related species (those showing 98.7 % or higher 16S rRNA gene sequence similarity) should be provided for the proposal of new prokaryote species [31, 34, 114, 115]. In addition, in case of large divergences between strains within a species, and in the absence of experimental DDH values, we strongly recommend that authors sequence the genomes of divergent strains to provide a measure of ANI to ascertain the affiliation of the strains to the same species.

It should be noted that thresholds should be used with caution: they are specific to a particular analytical method, used with specific parameters, and there is ongoing development of improved methods (requiring different thresholds). In addition, the whole idea that species boundaries can be defined by a universal cut-off in a distance measure reflects a particular approach to taxonomy that is largely confined to microbiologists. An alternative view is that species are natural units maintained by patterns of gene exchange, and the role of taxonomy is to identify the boundaries of these units [116, 117].

### Choice of phylogenetic reconstruction methods and substitution models

Phylogenetic reconstruction is a problem of statistical inference, and as such can be achieved using different classes of inference methods, including maximum-parsimony, distance-based methods (such as neighbour-joining), maximum-likelihood procedures and Bayesian inference. Each of these analytical methods has its particular strengths and weaknesses, which require careful consideration [118]. However, maximum-likelihood (ML) and Bayesian methods, which employ explicit models of molecular evolution, typically provide more realistic and robust phylogenies. Such model-based methods are easily within the capability of any modern computer, and fast software is available for large-scale studies [119]. Before applying them, it is important to identify appropriate substitution models and estimate best-fit model parameters [120–123]. For this purpose, mathematical methods such as likelihood ratio tests and Akaike and Bayesian information criteria may be used [124], and can easily be implemented using software such as jModelTest [125] or the model-finding module in MEGA7 [126]. Due to diverse evolutionary forces and processes (recombination, HGT, differential loss of paralogues, retention of ancestral polymorphisms and excessively high or low mutation rates, among others), many loci in genomes have undesirable properties for phylogenetic reconstruction. If undetected, these can lead to erroneous or biased estimates [127, 128] although, ironically, with strong branch support [129]. Their impact is particularly severe in concatenated

datasets [130, 131], which are standard in microbial phylogenomics [132]. Hence, robust phylogenomic inference requires careful selection of well-suited markers for the job; software is available for this, e.g. GET\_PHYLOMARKERS [49]. This package can also estimate ML and parsimony phylogenies from pan-genome matrices, and provide convenient statistical analyses of ANI (and other OGRI) matrices to decide the (mathematically) optimal number of clusters using diverse goodness-of-clustering statistics.

### Inclusion of phenotypic information and metadata

The International Code of Nomenclature of Prokaryotes [13] specifies that:

‘Descriptions of taxa should include the following information: (a) those characteristics which are essential for membership in the taxon, i.e., those characteristics which constitute the basic concept of the taxon; (b) those characteristics which qualify the taxon for membership in the next higher taxon; (c) the diagnostic characteristics, i.e., those characteristics which distinguish the taxon from closely related taxa; and (d) in the case of species, the total number of strains studied, the strain designations, and the number of strains which are either positive or negative for each characteristic.’

In modern bacterial systematics, the characters used for species definition and diagnosis are derived from DNA sequence, so the need for an extensive description of phenotypic properties is greatly reduced. Nevertheless, it is still useful to provide a basic phenotypic characterisation, including information relevant for culturing the bacteria. It is important, as the Code emphasizes, to make it clear whether the properties apply to all known strains or only to some.

Phenotypic features such as those derived from carbon and nitrogen source utilization, Biolog tests and fatty acid methyl esters (FAMES) are variable at the intraspecies level and so have limited interspecies discriminative and informative powers [11, 42, 117]. Phenotypic and chemotypic tests may possibly be used as additional characteristics of strains but should not be required as species markers [32, 86, 112, 133]. It is important to establish whether they are consistent across multiple strains of the proposed species [27]; there is little value in reporting results just for the type strain. MALDI-TOF mass spectrometry of cellular proteins is a rapid strain characterization method that is proving increasingly useful to survey the variation in large collections of isolates. It is routinely used for the identification of clinical isolates, and is becoming practicable for rhizobia as reference databases are extended to include more relevant species [134, 135]. Although the ability to deduce the chemotaxonomic properties from genomic data is still in its infancy [28, 47], a ‘genomotaxonomy’ approach has been proposed to predict stable phenotypes from genomic data [42]. Lassalle *et al.* [136, 137] used comparative

genomics to identify discriminant biochemical traits in agrobacteria, at the species and clade levels.

### Valid publication

The rules for publication are defined in the International Code of Nomenclature of Prokaryotes [13]. The names of new species or genera may be ‘effectively’ published in any journal, in print or online, but are only ‘validly’ published once they appear in the *International Journal of Systematic and Evolutionary Microbiology* (IJSEM). Names published in other journals appear in ‘validation lists’ in IJSEM. It is the date and order of the valid publication that determines the priority of the name, so authors who publish elsewhere are urged to submit their paper to IJSEM immediately on publication, in order that it can appear in the next ‘validation list’. There are a number of requirements for valid publication. The proposed name must conform to the nomenclatural rules, there must be a description that is recommended (Rule 30 of the Code) to meet at least the minimal standards (this document), and a type strain must be designated (and a type species, in the case of a new genus). A type strain ‘must be deposited in at least two publicly accessible service collections in different countries from which subcultures must be available’.

### Genus-level assignments

There is no recognized ‘golden standard’ for genus delineation and only limited attempts have been made to define generic boundaries between prokaryotes [138]. Genus affiliation has, so far, primarily relied on the single estimate of evolutionary relationships inferred from sequence variation of the 16S rRNA gene. It is widely recognized that species of the same genus form monophyletic groups on the basis of 16S rRNA gene sequences, MLSA and supertree analysis [30]. Prokaryotic genera remain loosely defined, typically based on monophyly of strains; there is no clear rule on sequence divergence between genera but, in the literature, the boundary between genera lies between 94–95 % 16S rRNA gene sequence similarity [27, 107]. For the classification of prokaryotic genera or higher taxa, Chun *et al.* [139] proposed that phylogenomic analysis should be used based on at least 30 single-copy orthologous genes (unlikely to have undergone lateral transfer events), complementing 16S rRNA-derived phylogeny and highly conserved phenotypes, including chemotaxonomic markers.

Other proposals for the delimitation of genera have been based on the proportion of shared proteins [139] or the level of amino acid identity of shared proteins [140]. In the absence of agreed standards, authors are free to choose appropriate evidence to support the creation of new genera, but it is important that new and amended genera should be monophyletic.

Recently, Parks *et al.* [141] used genome-based phylogeny to explore taxon boundaries across the whole of the Bacteria, based on relative evolutionary distance from ancestors, and proposed many shifts in taxonomic level to remove obvious anomalies. Their analysis shows that genus-level

assignments are generally satisfactory for rhizobia and agrobacteria, but many GenBank entries described as *Rhizobium* or *Agrobacterium* have not been updated, and actually belong to more-recently defined genera.

Many rhizobia and agrobacteria harbour chromids, which have been proposed as potential genus-specific indicators because they are similar within a genus, but those in different genera tend to use different replication genes and to carry different sets of conserved genes [142]. A very relevant example is the genus *Agrobacterium*, in which the chromid not only carries a distinctive replication system and conserved genes, but is linear rather than circular [51]. Therefore, it is suggested, but not mandatory, that evidence is sought for the presence of protelomerase, which is necessary for the maintenance of the linear chromid (see section Additional considerations specific to rhizobia and agrobacteria, Agrobacteria: pathogenic features, below). Alternatively, the replicon structure could be investigated directly by pulse-field or Eckhardt gel electrophoresis [51].

The Subcommittee encourages authors to have comprehensive evidence before describing a new agrobacterial or rhizobial genus or transferring species to a different genus. Phylogenetic reconstructions should include sequences from all type strains of species in a given genus including, most importantly, the type species [27]. It is recommended that authors provide information on percentage gene sequence similarity with species accommodated in the same genus (over full pairwise comparisons).

## ADDITIONAL CONSIDERATIONS SPECIFIC TO RHIZOBIA AND AGROBACTERIA

Group-specific guidelines help species descriptions to be useful to the research communities that are interested in those organisms and, specifically here, the communities that study nodulating and plant tumour-inducing bacteria. In rhizobia and agrobacteria these plant interactions are encoded by accessory genes that can function in a limited range of bacterial species, which imposes some taxonomic limits. Non-symbiotic and non-pathogenic members must be considered, and taxonomic definitions should not include symbiotic or pathogenic characters because a bacterium cannot change species merely by gaining or losing some accessory genes. Nevertheless, if the type strain or other strains have plant-interaction phenotypes, these should be described in the taxonomic proposal, because the ability to host these functions is an important aspect of the biology of the species. The genomes of type strains of new species should be analysed for their completeness and the presence of relevant functions. The gene content and the phenotypic properties should be treated as characteristics of the type strain but not of the entire taxon.

### Rhizobia: symbiotic capacities and symbiovars

Rhizobia should be tested for their infectivity (nodulation capacity) and effectiveness (symbiotic/free-living nitrogen fixation). Experience shows that the nodulation test itself

may not be sufficient to demonstrate that a particular strain is able to nodulate a specific plant. The strain must be isolated from the nodule again and shown to be identical to the inoculated strain (Koch's postulate). Regarding the host plant range, it may be considered sufficient to demonstrate the nodulation of a particular host since it is difficult to define which plant species should be tested. Testing nodulation ability on a broad symbiont-range legume may be an alternative for rhizobia originating from legume species difficult to cultivate, such as trees for which seeds are difficult to obtain and/or germinate, or if there is a long delay before nodulation. For instance, *Phaseolus vulgaris*, *Macroptilium atropurpureum* or *Vigna unguiculata* may be used for alpha-rhizobia and some beta-rhizobia (from South Africa), and *Mimosa pudica* for many beta-rhizobia isolated from mimosoids. Nitrogen fixation can be confirmed by comparing the appearance, biomass and/or N content of inoculated plants with that of uninoculated controls, or by the acetylene reduction assay [143]. Several bacterial genes are involved in the plant nodulation process and some are used for symbiotic characterization and are predictive for host range. Among them *nodC* and *nodA* are the most studied, and many sequences are available in the GenBank database for these genes. Sequences of nitrogen fixation genes (particularly *nifH*) should also be studied; we know that a few rhizobia have *nif* genes but lack *nod* genes [144]. We encourage a broader consideration of the biological insights that can be gained from a genome sequence as done, for example, by Aserse et al. [145].

Symbiovars [146] are defined to describe the nodulation properties of a rhizobium, reflecting similarity in host range and *nod* genes. It should be noted that the term 'symbiovar' is not a formal taxonomic category, but the concept has led to a valuable classification of the accessory genes involved in symbiosis.

### Agrobacteria: pathogenic features

Just as with symbiosis, pathogenicity encoded by Ti and Ri plasmids is not relevant for the taxonomy of agrobacteria. However, for agrobacterial strains isolated from tumours or hairy roots it is useful (not mandatory) to test pathogenicity. This test may also be indicated for strains isolated from other environments, if genome sequences reveal the occurrence of genes known to be related to pathogenicity.

In the case of *Agrobacterium* spp., it is necessary to demonstrate that strains possess a linear chromid, which seems to be a characteristic feature of this genus [51, 52]. In this regard, it would be sufficient to detect a gene for protelomerase by PCR, and/or to detect it in the whole genome sequence of the type strain [51]. No 'universal' *telA* primers exist so far but *telA* sequences can easily be searched in sequenced genomes, even though it requires a good level of sequence completion not to miss the gene, reinforcing the need for high-quality sequenced genomes. Alternatively, it is possible to demonstrate the presence of a linear chromid by PFGE [147].

## CONCLUSIONS AND CONSENSUS RECOMMENDATIONS

The use of genome sequence data in bacterial taxonomy has long been recommended in the literature [32, 33, 38, 112] and is now becoming standard [31, 40, 44, 72, 148–151]. This is increasingly the case in rhizobial and agrobacterial taxonomy (Table S1) [86, 88, 152], demonstrating the pertinence of the comment by Graham et al. [14]: 'Clearly, the greater the percentage of the bacterial genome considered in the classification of strains, the greater the precision obtained'.

For a proper description of a new genus or species of rhizobia or agrobacteria, this Subcommittee believes that the following minimal standards are necessary.

- (1) A genome sequence must be available for the type strain. Genome sequences may be incompletely assembled, but should follow the proposed minimal standards for the use of genome data for the taxonomy of Prokaryotes [31].
- (2) Evidence for differentiation from other species must be provided and be based on genome sequences of the type strains.

It is recommended that, when relevant, the following additional useful information be included (but this is not mandatory).

- (3) Intra-specific molecular variation should be described based on several clearly different strains.
- (4) Phenotypic data, particularly information that is useful for culturing the organisms and understanding their ecology, should be provided for a set of strains representing the variation in the species.
- (5) It is of interest to provide information on plant-interaction capacity and related gene sequences, although these are not taxonomic features: nodulation phenotype (infectivity, effectiveness, host range) and *nod/nif* gene sequences for rhizobia, pathogenicity assay and Ti/Ri-related sequences for tumorigenic or rhizogenic strains.

These guidelines are in addition to current rules for general bacterial taxonomy set out in the International Code of Nomenclature of Prokaryotes [13], which include the following mandatory requirements.

- (6) New names must conform to the International Code of Nomenclature of Prokaryotes [13].
- (7) Descriptions must be validly published, either directly in *Int J Syst Evol Microbiol* (detailed instructions for submitting a proposal for naming new species can be found at <http://ijs.microbiologyresearch.org/content/journal/ijsem/about>), or in a validation list in this journal when published elsewhere. Authors wishing to have new names and/or combinations included in future lists should send an electronic copy of the published paper to the IJSEM Editorial Office with

evidence that all of the other requirements for valid publication have been met.

- (8) Type strains must be deposited in two recognized culture collections in two different countries, from which subcultures must be available.

#### Funding information

The authors received no specific grant from any funding agency.

#### Conflicts of interest

The authors declare that there are no conflicts of interest.

#### References

- de Lajudie PM, Young JPW. International Committee on Systematics of Prokaryotes, Subcommittee on the taxonomy of *Rhizobium* and *Agrobacterium*. Minutes of the meeting, Budapest, 25 August 2016. *Int J Syst Evol Microbiol* 2017;67:2485–2494.
- Estrada-de Los Santos P, Palmer M, Chávez-Ramírez B, Beukes C, Steenkamp ET et al. Whole genome analyses suggests that *Burkholderia sensu lato* contains two additional novel genera (*Mycetohabitans* gen. nov., and *Trinickia* gen. nov.): implications for the evolution of diazotrophy and nodulation in the *Burkholderiaceae*. *Genes* 2018;9:389–412.
- Andrews M, Andrews ME. Specificity in legume-rhizobia symbioses. *Int J Mol Sci* 2017;18:705–744.
- De Meyer SE, Willems A. Multilocus sequence analysis of *Bosea* species and description of *Bosea lupini* sp. nov., *Bosea lathyri* sp. nov. and *Bosea robiniae* sp. nov., isolated from legumes. *Int J Syst Evol Microbiol* 2012;62:2505–2510.
- Chou YJ, Elliott GN, James EK, Lin KY, Chou JH et al. *Labrys neptuniae* sp. nov., isolated from root nodules of the aquatic legume *Neptunia oleracea*. *Int J Syst Evol Microbiol* 2007;57:577–581.
- Martínez-Hidalgo P, Hirsch AM. The nodule microbiome: N<sub>2</sub>-fixing rhizobia do not live alone. *Phytobiomes J* 2017;1:70–82.
- Andrews M, De Meyer S, James EK, Stępkowski T, Hodge S et al. Horizontal transfer of symbiosis genes within and between rhizobial genera: occurrence and importance. *Genes* 2018;9:321–345.
- Remigi P, Zhu J, Young JPW, Masson-Boivin C. Symbiosis within symbiosis: evolving nitrogen-fixing legume symbionts. *Trends Microbiol* 2016;24:63–75.
- Mousavi SA, Willems A, Nesme X, de Lajudie P, Lindström K. Revised phylogeny of *Rhizobiaceae*: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations. *Syst Appl Microbiol* 2015;38:84–90.
- Kuzmanović N, Smalla K, Gronow S, Putawska J. *Rhizobium tumorigenes* sp. nov., a novel plant tumorigenic bacterium isolated from cane gall tumors on thornless blackberry. *Sci Rep* 2018;8:1–8.
- Putawska J, Kuzmanović N, Willems A, Pothier JF. *Pararhizobium polonicum* sp. nov. isolated from tumors on stone fruit rootstocks. *Syst Appl Microbiol* 2016;39:164–169.
- Dijkshoorn L. International Committee on Systematics of Prokaryotes. Minutes of the meetings, 7, 8 and 9 July 2017, Valencia, Spain. *Int J Syst Evol Microbiol* 2018;68:2104–2110.
- Parker CT, Tindall BJ, Garrity GM. International Code of Nomenclature of Prokaryotes. Prokaryotic code (2008 revision). *Int J Syst Evol Microbiol* 2019;69:S1–S111.
- Graham PH, Sadowsky MJ, Keyser HH, Barnet YM, Bradley RS et al. Proposed minimal standards for the description of new genera and species of root- and stem-nodulating bacteria. *Int J Syst Bacteriol* 1991;41:582–587.
- Lindstrom K. International Committee on Systematics of Prokaryotes: Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*. Minutes of the meeting, 4 July 2001, Hamilton, Canada. *Int J Syst Evol Microbiol* 2002;52:2337.
- Lindström K, Martínez-Romero ME. International Committee on Systematics of Prokaryotes; Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*. Minutes of the meeting, 26 July 2004, Toulouse, France. *Int J Syst Evol Microbiol* 2005;55:1383.
- Lindström K, Martínez-Romero E. International Committee on Systematics of Prokaryotes; Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*. Minutes of the meeting, 23–24 July 2006, Århus, Denmark. *Int J Syst Evol Microbiol* 2007;57:1365–1366.
- Lindström K, Young JPW. International Committee on Systematics of Prokaryotes; Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*. Minutes of the meeting, 31 August 2008, Gent, Belgium. *Int J Syst Evol Microbiol* 2009;59:921–922.
- Lindström K, Young JPW. International Committee on Systematics of Prokaryotes; Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*. Minutes of the meeting, 7 September 2010, Geneva, Switzerland. *Int J Syst Evol Microbiol* 2011;61:3089–3093.
- de Lajudie P, Martínez-Romero E. International Committee on Systematics of Prokaryotes; Subcommittee on the Taxonomy of *Agrobacterium* and *Rhizobium*. Minutes of the meeting, 7 September 2014, Tenerife, Spain. *Int J Syst Evol Microbiol* 2017;67:516–520.
- de Lajudie PM, Young JPW. International Committee on Systematics of Prokaryotes; Subcommittee on the taxonomy of rhizobia and agrobacteria. Minutes of the closed meeting, granada, 4 September 2017. *Int J Syst Evol Microbiol* 2018;68:3363–3368.
- Whitman WB. Intent of the nomenclatural code and recommendations about naming new species based on genomic sequences. *Bull BISMIS* 2011;2:135–139.
- Stackebrandt E, Goebel BM. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Evol Microbiol* 1994;44:846–849.
- Wayne LG, Moore WEC, Stackebrandt E, Kandler O, Colwell RR et al. Report of the *ad hoc* committee on reconciliation of approaches to bacterial systematics. *Int J Syst Evol Microbiol* 1987;37:463–464.
- Vandamme P, Pot B, Gillis M, De Vos P, Kersters K. Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* 1996;60:407–438.
- Stackebrandt E, Frederiksen W, Garrity GM, Grimont PA, Kämpfer P et al. Report of the *ad hoc* committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* 2002;52:1043–1047.
- Tindall BJ, Rosselló-Móra R, Busse HJ, Ludwig W, Kämpfer P. Notes on the characterization of prokaryote strains for taxonomic purposes. *Int J Syst Evol Microbiol* 2010;60:249–266.
- Mahato NK, Gupta V, Singh P, Kumari R, Verma H et al. Microbial taxonomy in the era of OMICS: application of DNA sequences, computational tools and techniques. *Antonie van Leeuwenhoek* 2017;110:1357–1371.
- Thompson CC, Vieira NM, Vicente AC, Thompson FL. Towards a genome-based taxonomy of Mycoplasmas. *Infect Genet Evol* 2011;11:1798–1804.
- Thompson CC, Chimento L, Edwards RA, Swings J, Stackebrandt E et al. Microbial genomic taxonomy. *BMC Genomics* 2013;14:913–921.
- Chun J, Oren A, Ventosa A, Christensen H, Arahal DR et al. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 2018;68:461–466.
- Konstantinidis KT, Tiedje JM. Towards a genome-based taxonomy for prokaryotes. *J Bacteriol* 2005;187:6258–6264.
- Auch AF, von Jan M, Klenk H-P, Göker M. Digital DNA-DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand Genomic Sci* 2010;2:117–134.



34. Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60–74.
35. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M et al. Versatile and open software for comparing large genomes. *Genome Biol* 2004;5:R12.
36. Deloger M, El Karoui M, Petit MA. A genomic distance based on MUM indicates discontinuity between most bacterial species and genera. *J Bacteriol* 2009;191:91–99.
37. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P et al. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 2007;57:81–91.
38. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 2009;106:19126–19131.
39. Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K et al. Microbial species delineation using whole genome sequences. *Nucleic Acids Res* 2015;43:6761–6771.
40. Arahall DR. Whole-genome analyses: average nucleotide identity. In: Goodfellow M, Sutcliffe I and Chun J (editors). *Methods in Microbiology* Academic Press; 2014. pp. 103–122.
41. Kim M, Park S-C, Baek I, Chun J. Large-scale evaluation of experimentally determined DNA G+C contents with whole genome sequences of prokaryotes. *Syst Appl Microbiol* 2015;38:79–83.
42. Ormeño-Orrillo E, Martínez-Romero E. Phenotypic tests in *Rhizobium* species description: an opinion and (a sympatric speciation) hypothesis. *Syst Appl Microbiol* 2013;36:145–147.
43. Sangal V, Goodfellow M, Jones AL, Schwalbe EC, Blom J et al. Next-generation systematics: An innovative approach to resolve the structure of complex prokaryotic taxa. *Sci Rep* 2016;6:38392.
44. Vandamme P, Peeters C. Time to revisit polyphasic taxonomy. *Antonie van Leeuwenhoek* 2014;106:57–65.
45. Coenye T, Gevers D, van De Peer Y, Vandamme P, Swings J. Towards a prokaryotic genomic taxonomy. *FEMS Microbiol Rev* 2005;29:147–167.
46. Gevers D, Dawyndt P, Vandamme P, Willems A, Vancanneyt M et al. Stepping stones towards a new prokaryotic taxonomy. *Philos Trans R Soc Lond B Biol Sci* 2006;361:1911–1916.
47. Sutcliffe IC, Trujillo ME, Goodfellow M. A call to arms for systematists: revitalising the purpose and practises underpinning the description of novel microbial taxa. *Antonie van Leeuwenhoek* 2012;101:13–20.
48. Sutcliffe IC, Trujillo ME, Whitman WB, Goodfellow M. A call to action for the International Committee on Systematics of Prokaryotes. *Trends Microbiol* 2013;21:51–52.
49. Vinuesa P, Ochoa-Sánchez LE, Contreras-Moreira B. GET\_PHYLOMARKERS, a software package to select optimal orthologous clusters for phylogenomics and inferring pan-genome phylogenies, used for a critical geno-taxonomic revision of the genus *Stenotrophomonas*. *Front Microbiol* 2018;9:771.
50. Young JM, Kuykendall LD, Martínez-Romero E, Kerr A, Sawada H. A revision of *Rhizobium* Frank 1889, with an emended description of the genus, and the inclusion of all species of *Agrobacterium* Conn 1942 and *Allorhizobium undicola* de Lajudie et al. 1998 as new combinations: *Rhizobium radiobacter*, *R. rhizogenes*, *R. rubi*, *R. undicola* and *R. vitis*. *Int J Syst Evol Microbiol* 2001;51:89–103.
51. Ramírez-Bahena MH, Vial L, Lassalle F, Diel B, Chapulliot D et al. Single acquisition of protelomerase gave rise to speciation of a large and diverse clade within the *Agrobacterium/Rhizobium* supercluster characterized by the presence of a linear chromid. *Mol Phylogenet Evol* 2014;73:202–207.
52. Mousavi SA, Österman J, Wahlberg N, Nesme X, Lavire C et al. Phylogeny of the *Rhizobium-Allorhizobium-Agrobacterium* clade supports the delineation of *Neorhizobium* gen. nov. *Syst Appl Microbiol* 2014;37:208–215.
53. Gan HM, Savka MA. One more decade of *Agrobacterium* taxonomy. *Current Topics in Microbiology and Immunology*. Berlin, Heidelberg: Springer; 2018. pp. 1–14.
54. Gaunt MW, Turner SL, Rigottier-Gois L, Lloyd-Macgilp SA, Young JPW. Phylogenies of *atpD* and *recA* support the small subunit rRNA-based classification of rhizobia. *Int J Syst Evol Microbiol* 2001;51:2037–2048.
55. Turner SL, Young JPW. The glutamine synthetases of rhizobia: phylogenetics and evolutionary implications. *Mol Biol Evol* 2000;17:309–319.
56. Wernegreen JJ, Riley MA. Comparison of the evolutionary dynamics of symbiotic and housekeeping loci: a case for the genetic coherence of rhizobial lineages. *Mol Biol Evol* 1999;16:98–113.
57. Martens M, Delaere M, Coopman R, De Vos P, Gillis M et al. Multilocus sequence analysis of *Ensifer* and related taxa. *Int J Syst Evol Microbiol* 2007;57:489–503.
58. Martens M, Dawyndt P, Coopman R, Gillis M, De Vos P et al. Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *Int J Syst Evol Microbiol* 2008;58:200–214.
59. Parker MA. rRNA and *dnaK* relationships of *Bradyrhizobium* sp. nodule bacteria from four papilionoid legume trees in Costa Rica. *Syst Appl Microbiol* 2004;27:334–342.
60. Moulin L, Béna G, Boivin-Masson C, Stepkowski T. Phylogenetic analyses of symbiotic nodulation genes support vertical and lateral gene co-transfer within the *Bradyrhizobium* genus. *Mol Phylogenet Evol* 2004;30:720–732.
61. van Berkum P, Terefework Z, Paulin L, Suomalainen S, Lindström K et al. Discordant phylogenies within the *rrn* loci of rhizobia. *J Bacteriol* 2003;185:2988–2998.
62. Silva C, Vinuesa P, Eguiarte LE, Souza V, Martínez-Romero E. Evolutionary genetics and biogeographic structure of *Rhizobium gallicum sensu lato*, a widely distributed bacterial symbiont of diverse legumes. *Mol Ecol* 2005;14:4033–4050.
63. Vinuesa P, Silva C, Lorite MJ, Izaguirre-Mayoral ML, Bedmar EJ et al. Molecular systematics of rhizobia based on maximum likelihood and Bayesian phylogenies inferred from *rrs*, *atpD*, *recA* and *nifH* sequences, and their use in the classification of *Sesbania* microsymbionts from Venezuelan wetlands. *Syst Appl Microbiol* 2005;28:702–716.
64. Vinuesa P, Silva C, Werner D, Martínez-Romero E. Population genetics and phylogenetic inference in bacterial molecular systematics: the roles of migration and recombination in *Bradyrhizobium* species cohesion and delineation. *Mol Phylogenet Evol* 2005;34:29–54.
65. Vinuesa P, León-Barríos M, Silva C, Willems A, Jarabo-Lorenzo A et al. *Bradyrhizobium canariense* sp. nov., an acid-tolerant endosymbiont that nodulates endemic genistoid legumes (Papilionoideae: Genisteeae) from the Canary Islands, along with *Bradyrhizobium japonicum* bv. *genistearum*, *Bradyrhizobium* genospecies alpha and *Bradyrhizobium* genospecies beta. *Int J Syst Evol Microbiol* 2005;55:569–575.
66. Vinuesa P, Rojas-Jiménez K, Contreras-Moreira B, Mahna SK, Prasad BN et al. Multilocus sequence analysis for assessment of the biogeography and evolutionary genetics of four *Bradyrhizobium* species that nodulate soybeans on the asiatic continent. *Appl Environ Microbiol* 2008;74:6987–6996.
67. Rivas R, Martens M, de Lajudie P, Willems A. Multilocus sequence analysis of the genus *Bradyrhizobium*. *Syst Appl Microbiol* 2009;32:101–110.
68. Nzoué A, Miché L, Klonowska A, Laguerre G, de Lajudie P et al. Multilocus sequence analysis of bradyrhizobia isolated from *Aeschynomene* species in Senegal. *Syst Appl Microbiol* 2009;32:400–412.

69. Merabet C, Martens M, Mahdhi M, Zakhia F, Sy A et al. Multilocus sequence analysis of root nodule isolates from *Lotus arabicus* (Senegal), *Lotus creticus*, *Argyrobium uniflorum* and *Medicago sativa* (Tunisia) and description of *Ensifer numidicus* sp. nov. and *Ensifer garamanticus* sp. nov. *Int J Syst Evol Microbiol* 2010;60:664–674.
70. Hernández-Lucas I, Rogel-Hernández MA, Segovia L, Rojas-Jiménez K, Martínez-Romero E. Phylogenetic relationships of rhizobia based on citrate synthase gene sequences. *Syst Appl Microbiol* 2004;27:703–706.
71. Eardly BD, Nour SM, van Berkum P, Selander RK. Rhizobial 16S rRNA and *dnaK* genes: mosaicism and the uncertain phylogenetic placement of *Rhizobium galegae*. *Appl Environ Microbiol* 2005;71:1328–1335.
72. Chun J, Rainey FA. Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea. *Int J Syst Evol Microbiol* 2014;64:316–324.
73. Young JPW, Haukka KE. Diversity and phylogeny of rhizobia. *New Phytologist* 1996;133:87–94.
74. Tian CF, Zhou YJ, Zhang YM, Li QQ, Zhang YZ et al. Comparative genomics of rhizobia nodulating soybean suggests extensive recruitment of lineage-specific genes in adaptations. *Proc Natl Acad Sci USA* 2012;109:8629–8634.
75. López-López A, Rogel-Hernández MA, Barois I, Ortiz Ceballos AI, Martínez J et al. *Rhizobium grahamii* sp. nov., from nodules of *Dalea leporina*, *Leucaena leucocephala* and *Clitoria ternatea*, and *Rhizobium mesoamericanum* sp. nov., from nodules of *Phaseolus vulgaris*, *siratro*, cowpea and *Mimosa pudica*. *Int J Syst Evol Microbiol* 2012;62:2264–2271.
76. López-Guerrero MG, Ormeño-Orrillo E, Velázquez E, Rogel MA, Acosta JL et al. *Rhizobium etli* taxonomy revised with novel genomic data and analyses. *Syst Appl Microbiol* 2012;35:353–358.
77. Ribeiro RA, Rogel MA, López-López A, Ormeño-Orrillo E, Barcellos FG et al. Reclassification of *Rhizobium tropici* type A strains as *Rhizobium leucaenae* sp. nov. *Int J Syst Evol Microbiol* 2012;62:1179–1184.
78. Delamuta JR, Ribeiro RA, Ormeño-Orrillo E, Melo IS, Martínez-Romero E et al. Polyphasic evidence supporting the reclassification of *Bradyrhizobium japonicum* group Ia strains as *Bradyrhizobium diazoefficiens* sp. nov. *Int J Syst Evol Microbiol* 2013;63:3342–3351.
79. Delamuta JR, Ribeiro RA, Ormeño-Orrillo E, Parma MM, Melo IS et al. *Bradyrhizobium tropiciagri* sp. nov. and *Bradyrhizobium embrapense* sp. nov., nitrogen-fixing symbionts of tropical forage legumes. *Int J Syst Evol Microbiol* 2015;65:4424–4433.
80. Dall'Agnol RF, Ribeiro RA, Ormeño-Orrillo E, Rogel MA, Delamuta JR et al. *Rhizobium freirei* sp. nov., a symbiont of *Phaseolus vulgaris* that is very effective at fixing nitrogen. *Int J Syst Evol Microbiol* 2013;63:4167–4173.
81. Dall'Agnol RF, Ribeiro RA, Delamuta JRM, Ormeño-Orrillo E, Rogel MA et al. *Rhizobium paranaense* sp. nov., an effective N<sub>2</sub>-fixing symbiont of common bean (*Phaseolus vulgaris* L.) with broad geographical distribution in Brazil. *Int J Syst Evol Microbiol* 2014;64:3222–3229.
82. Durán D, Rey L, Mayo J, Zúñiga-Dávila D, Imperial J et al. *Bradyrhizobium paxllaeri* sp. nov. and *Bradyrhizobium icense* sp. nov., nitrogen-fixing rhizobial symbionts of Lima bean (*Phaseolus lunatus* L.) in Peru. *Int J Syst Evol Microbiol* 2014;64:2072–2078.
83. Durán D, Rey L, Navarro A, Busquets A, Imperial J et al. *Bradyrhizobium valentinum* sp. nov., isolated from effective nodules of *Lupinus mariae-josephae*, a lupine endemic of basic-lime soils in Eastern Spain. *Syst Appl Microbiol* 2014;37:336–341.
84. Sawana A, Adeolu M, Gupta RS. Molecular signatures and phylogenomic analysis of the genus *Burkholderia*: proposal for division of this genus into the emended genus *Burkholderia* containing pathogenic organisms and a new genus *Paraburkholderia* gen. nov. harboring environmental species. *Front Genet* 2014;5:429–451.
85. Kuzmanović N, Putawska J, Prokić A, Ivanović M, Zlatković N et al. *Agrobacterium arsenijevicei* sp. nov., isolated from crown gall tumors on raspberry and cherry plum. *Syst Appl Microbiol* 2015;38:373–378.
86. Ormeño-Orrillo E, Servín-Garcidueñas LE, Rogel MA, González V, Peralta H et al. Taxonomy of rhizobia and agrobacteria from the *Rhizobiaceae* family in light of genomics. *Syst Appl Microbiol* 2015;38:287–291.
87. Steenkamp ET, van Zyl E, Beukes CW, Avontuur JR, Chan WY et al. *Burkholderia kirstenboschensis* sp. nov. nodulates papilionoid legumes indigenous to South Africa. *Syst Appl Microbiol* 2015;38:545–554.
88. Rashid MH, Young JPW, Everall I, Clercx P, Willems A et al. Average nucleotide identity of genome sequences supports the description of *Rhizobium lentis* sp. nov., *Rhizobium bangladeshense* sp. nov. and *Rhizobium binae* sp. nov. from lentil (*Lens culinaris*) nodules. *Int J Syst Evol Microbiol* 2015;65:3037–3045.
89. Diouf F, Diouf D, Klonowska A, Le Queré A, Bakhoum N et al. Genetic and genomic diversity studies of *Acacia* symbionts in Senegal reveal new species of *Mesorhizobium* with a putative geographical pattern. *PLoS One* 2015;10:e0117667.
90. Ribeiro RA, Martins TB, Ormeño-Orrillo E, Marçon Delamuta JR, Rogel MA et al. *Rhizobium ecuadorensis* sp. nov., an indigenous N<sub>2</sub>-fixing symbiont of the Ecuadorian common bean (*Phaseolus vulgaris* L.) genetic pool. *Int J Syst Evol Microbiol* 2015;65:3162–3169.
91. Torres Tejerizo G, Rogel MA, Ormeño-Orrillo E, Althabegoiti MJ, Nilsson JF et al. *Rhizobium favelukesii* sp. nov., isolated from the root nodules of alfalfa (*Medicago sativa* L.). *Int J Syst Evol Microbiol* 2016;66:4451–4457.
92. Le Queré A, Tak N, Gehlot HS, Lavire C, Meyer T et al. Genomic characterization of *Ensifer aridi*, a proposed new species of nitrogen-fixing rhizobium recovered from Asian, African and American deserts. *BMC Genomics* 2017;18:85–109.
93. Aserse AA, Woyke T, Kyrpides NC, Whitman WB, Lindström K. Draft genome sequence of type strain HBR26<sup>T</sup> and description of *Rhizobium aethiopicum* sp. nov. *Stand Genomic Sci* 2017;12:14–30.
94. Yan J, Yan H, Liu LX, Chen WF, Zhang XX et al. *Rhizobium hidalgonense* sp. nov., a nodule endophytic bacterium of *Phaseolus vulgaris* in acid soil. *Arch Microbiol* 2017;199:97–104.
95. Baraúna AC, Rouws LF, Simões-Araujo JL, dos Reis Junior FB, Iannetta PP et al. *Rhizobium altiplani* sp. nov., isolated from effective nodules on *Mimosa pudica* growing in untypically alkaline soil in central Brazil. *Int J Syst Evol Microbiol* 2016;66:4118–4124.
96. Michel DC, Passos SR, Simões-Araujo JL, Baraúna AC, da Silva K et al. *Bradyrhizobium centrolobii* and *Bradyrhizobium macuxiense* sp. nov. isolated from *Centrolobium paraense* grown in soil of Amazonia, Brazil. *Arch Microbiol* 2017;199:657–664.
97. Bournaud C, Moulin L, Cnockaert M, De Faria S, Prin Y et al. *Paraburkholderia piptadeniae* sp. nov. and *Paraburkholderia ribeironis* sp. nov., two root-nodulating symbiotic species of *Piptadenia gonoacantha* in Brazil. *Int J Syst Evol Microbiol* 2017;67:432–440.
98. Oren A, Garrity GM. Then and now: a systematic review of the systematics of prokaryotes in the last 80 years. *Antonie van Leeuwenhoek* 2014;106:43–56.
99. Young JPW. Bacteria are smartphones and mobile genes are apps. *Trends Microbiol* 2016;24:931–932.
100. Kuzmanović N, Putawska J, Smalla K, Nesme X. *Agrobacterium rosae* sp. nov., isolated from galls on different agricultural crops. *Syst Appl Microbiol* 2018;41:191–197.
101. Zhang YM, Tian CF, Sui XH, Chen WF, Chen WX. Robust markers reflecting phylogeny and taxonomy of rhizobia. *PLoS One* 2012;7:e44936.

102. Aserse AA, Räsänen LA, Assefa F, Hailemariam A, Lindström K. Phylogeny and genetic diversity of native rhizobia nodulating common bean (*Phaseolus vulgaris* L.) in Ethiopia. *Syst Appl Microbiol* 2012;35:120–131.
103. Degefu T, Wolde-meskel E, Frostegård Å. Phylogenetic multilocus sequence analysis identifies seven novel *Ensifer* genospecies isolated from a less-well-explored biogeographical region in East Africa. *Int J Syst Evol Microbiol* 2012;62:2286–2295.
104. Rincón-Rosales R, Villalobos-Escobedo JM, Rogel MA, Martínez J, Ormeño-Orrillo E et al. *Rhizobium calliandrae* sp. nov., *Rhizobium mayense* sp. nov. and *Rhizobium jaguaris* sp. nov., rhizobial species nodulating the medicinal legume *Calliandra grandiflora*. *Int J Syst Evol Microbiol* 2013;63:3423–3429.
105. Laguerre G, Mavingui P, Allard MR, Charnay MP, Louvrier P et al. Typing of rhizobia by PCR DNA fingerprinting and PCR-restriction fragment length polymorphism analysis of chromosomal and symbiotic gene regions: application to *Rhizobium leguminosarum* and its different biovars. *Appl Environ Microbiol* 1996;62:2029–2036.
106. Versalovic J, Schneider M, De Bruijn FJ, Lupski JR. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol Cell Biol* 1994;5:25–40.
107. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 2014;12:635–645.
108. Ramírez-Bahena MH, García-Fraile P, Peix A, Valverde A, Rivas R et al. Revision of the taxonomic status of the species *Rhizobium leguminosarum* (Frank 1879) Frank 1889<sup>AL</sup>, *Rhizobium phaseoli* Dangeard 1926<sup>AL</sup> and *Rhizobium trifolii* Dangeard 1926<sup>AL</sup>. *R. trifolii* is a later synonym of *R. leguminosarum*. Reclassification of the strain *R. leguminosarum* DSM 30132 (=NCIMB 11478) as *Rhizobium pisi* sp. nov. *Int J Syst Evol Microbiol* 2008;58:2484–2490.
109. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* 2016;17:132–146.
110. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;9:5114–5122.
111. Lees JA, Harris SR, Tonkin-Hill G, Gladstone RA, Lo SW et al. Fast and flexible bacterial genomic epidemiology with PopPUNK. *Genome Res* 2019;29:304–316.
112. Konstantinidis KT, Ramette A, Tiedje JM. The bacterial species definition in the genomic era. *Philos Trans R Soc Lond B Biol Sci* 2006;361:1929–1940.
113. Fox GE, Wisotzkey JD, Jurtschuk P. How close is close: 16S rRNA sequence identity may not be sufficient to guarantee species identity. *Int J Syst Bacteriol* 1992;42:166–170.
114. Kim M, Oh H-S, Park S-C, Chun J. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int J Syst Evol Microbiol* 2014;64:346–351.
115. Stackebrandt E. Taxonomic parameters revisited: tarnished gold standards. *Microbiol Today* 2006;6:152–155.
116. Costechareyre D, Bertolla F, Nesme X. Homologous recombination in *Agrobacterium*: potential implications for the genomic species concept in bacteria. *Mol Biol Evol* 2009;26:167–176.
117. Kumar N, Lad G, Giuntini E, Kaye ME, Udomwong P et al. Bacterial genospecies that are not ecologically coherent: population genomics of *Rhizobium leguminosarum*. *Open Biol* 2015;5:140133.
118. Yang Z, Rannala B. Molecular phylogenetics: principles and practice. *Nat Rev Genet* 2012;13:303–314.
119. Zhou X, Shen XX, Hittinger CT, Rokas A. Evaluating fast maximum likelihood-based phylogenetic programs using empirical phylogenomic data sets. *Mol Biol Evol* 2018;35:486–503.
120. Felsenstein J. *Inferring Phylogenies*, vol. 2. Sunderland, MA: Sinauer Associates; 2004.
121. Bruno WJ, Halpern AL. Topological bias and inconsistency of maximum likelihood using wrong models. *Mol Biol Evol* 1999;16:564–566.
122. Buckley TR, Cunningham CW. The effects of nucleotide substitution model assumptions on estimates of nonparametric bootstrap support. *Mol Biol Evol* 2002;19:394–405.
123. Kelsey CR, Crandall KA, Voevodin AF. Different models different trees: the geographic origin of PTLV-I. *Mol Phylogenet Evol* 1999;13:336–347.
124. Posada D, Buckley TR. Model selection and model averaging in phylogenetics: advantages of akaike information criterion and bayesian approaches over likelihood ratio tests. *Syst Biol* 2004;53:793–808.
125. Posada D. jModelTest: phylogenetic model averaging. *Mol Biol Evol* 2008;25:1253–1256.
126. Kumar S, Stecher G, Tamura K. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33:1870–1874.
127. Parks MB, Wickett NJ, Alverson AJ. Signal, uncertainty, and conflict in phylogenomic data for a diverse lineage of microbial eukaryotes (diatoms, bacillariophyta). *Mol Biol Evol* 2018;35:80–93.
128. Shen XX, Hittinger CT, Rokas A. Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nat Ecol Evol* 2017;1:0126.
129. Kumar S, Filipski AJ, Battistuzzi FU, Kosakovsky Pond SL, Tamura K. Statistics and truth in phylogenomics. *Mol Biol Evol* 2012;29:457–472.
130. Degnan JH, Rosenberg NA. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol Evol* 2009;24:332–340.
131. Kubatko LS, Degnan JH. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst Biol* 2007;56:17–24.
132. Wu M, Eisen JA. A simple, fast, and accurate method of phylogenomic inference. *Genome Biol* 2008;9:R151.
133. Sutcliffe IC. Challenging the anthropocentric emphasis on phenotypic testing in prokaryotic species descriptions: rip it up and start again. *Front Genet* 2015;6:218.
134. Ziegler D, Pothier JF, Ardley J, Fossou RK, Pflüger V et al. Ribosomal protein biomarkers provide root nodule bacterial identification by MALDI-TOF MS. *Appl Microbiol Biotechnol* 2015;99:5547–5562.
135. Rahi P, Prakash O, Shouche YS. Matrix-assisted laser desorption/ionization time-of-flight mass-spectrometry (MALDI-TOF MS) based microbial identifications: challenges and scopes for microbial ecologists. *Front Microbiol* 2016;7:1359.
136. Lassalle F, Planel R, Penel S, Chapulliot D, Barbe V et al. Ancestral genome estimation reveals the history of ecological diversification in *Agrobacterium*. *Genome Biol Evol* 2017;9:3413–3431.
137. Lassalle F, Campillo T, Vial L, Baude J, Costechareyre D et al. Genomic species are ecological species as revealed by comparative genomics in *Agrobacterium tumefaciens*. *Genome Biol Evol* 2011;3:762–781.
138. Qin QL, Xie BB, Zhang XY, Chen XL, Zhou BC et al. A proposed genus boundary for the prokaryotes based on genomic insights. *J Bacteriol* 2014;196:2210–2215.
139. Chun J, Grim CJ, Hasan NA, Lee JH, Choi SY et al. Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *Proc Natl Acad Sci USA* 2009;106:15442–15447.
140. Konstantinidis KT, Rosselló-Móra R, Amann R. Uncultivated microbes in need of their own taxonomy. *ISME J* 2017;11:2399–2406.

141. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 2018; 36:996–1004.
142. Harrison PW, Lower RPJ, Kim NKD, Young JPW. Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol* 2010;18:141–148.
143. Howieson JG, Dilworth MJ. *Working with Rhizobia*. Canberra: Australian Centre for International Agricultural Research; 2016. <https://www.aciar.gov.au/node/12651>.
144. Giraud E, Moulin L, Vallenet D, Barbe V, Cytryn E et al. Legumes symbioses: absence of *nod* genes in photosynthetic bradyrhizobia. *Science* 2007;316:1307–1312.
145. Aserse AA, Woyke T, Kyrpides NC, Whitman WB, Lindström K. Draft genome sequences of *Bradyrhizobium shewense* sp. nov. ERR11<sup>T</sup> and *Bradyrhizobium yuanmingense* CCB AU 10071<sup>T</sup>. *Stand Genomic Sci* 2017;12:74.
146. Rogel MA, Ormeño-Orrillo E, Martínez Romero E. Symbiovars in rhizobia reflect bacterial adaptation to legumes. *Syst Appl Microbiol* 2011;34:96–104.
147. Ramírez-Bahena MH, Nesme X, Muller D. Rapid and simultaneous detection of linear chromosome and large plasmids in Proteobacteria. *J Basic Microbiol* 2012;52:736–739.
148. Whitman WB. Genome sequences as the type material for taxonomic descriptions of prokaryotes. *Syst Appl Microbiol* 2015;38: 217–222.
149. Thompson CC, Amaral GR, Campeão M, Edwards RA, Polz MF et al. Microbial taxonomy in the post-genomic era: rebuilding from scratch? *Arch Microbiol* 2015;197:359–370.
150. Konstantinidis KT, Stackebrandt E. Defining taxonomic ranks. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F et al. (editors). *The Prokaryotes*, Springer Berlin, Heidelberg; 2013. pp. 229–254.
151. Whitman WB. The need for change: embracing the genome. *Methods in Microbiology*, Elsevier; 2014. pp. 1–12.
152. Zahradník J, Nunvar J, Pařízková H, Kolářová L, Palyzová A et al. *Agrobacterium bohemicum* sp. nov. isolated from poppy seed wastes in central Bohemia. *Syst Appl Microbiol* 2018;41: 184–190.

#### Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as 'excellent' or 'very good'.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at [microbiologyresearch.org](http://microbiologyresearch.org).