

Review Article

New Insights into the Taxonomy of Bacteria in the Genomic Era and a Case Study with Rhizobia

Luisa Caroline Ferraz Helene ^{1,2} Milena Serenato Klepa ^{1,3,4}
and Mariangela Hungria ^{1,2,3}

¹Embrapa Soja, CP 4006, 86085-981 Londrina, PR, Brazil

²Conselho Nacional de Desenvolvimento Científico e Tecnológico, SHIS QI 1 Conjunto B, Blocos A, B, C e D, Lago Sul, 71605-001 Brasília, DF, Brazil

³Department of Microbiology, Universidade Estadual de Londrina, CP 10011, 86057-970 Londrina, PR, Brazil

⁴Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, SBN, Quadra 2, Bloco L, Lote 06, Edifício Capes, 70040-020 Brasília, DF, Brazil

Correspondence should be addressed to Mariangela Hungria; mariangela.hungria@embrapa.br

Received 4 February 2022; Accepted 9 May 2022; Published 21 May 2022

Academic Editor: Zhun Li

Copyright © 2022 Luisa Caroline Ferraz Helene et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Since early studies, the history of prokaryotes taxonomy has dealt with many changes driven by the development of new and more robust technologies. As a result, the number of new taxa descriptions is exponentially increasing, while an increasing number of others has been subject of reclassification, demanding from the taxonomists more effort to maintain an organized hierarchical system. However, expectations are that the taxonomy of prokaryotes will acquire a more stable status with the genomic era. Other analyses may continue to be necessary to determine microbial features, but the use of genomic data might be sufficient to provide reliable taxa delineation, helping taxonomy to reach the goal of correct classification and identification. Here we describe the evolution of prokaryotes' taxonomy until the genomic era, emphasizing bacteria and taking as an example the history of rhizobia taxonomy. This example was chosen because of the importance of the symbiotic nitrogen fixation of legumes with rhizobia to the nitrogen input to both natural ecosystems and agricultural crops. This case study reports the technological advances and the methodologies used to classify and identify bacterial species and indicates the actual rules required for an accurate description of new taxa.

1. Introduction

The taxonomy terminology has been broadly discussed. Some researchers consider the taxonomy as systematics, while others define taxonomy as the classification of organisms and part of the systematics, which would have a broader scope, including studies with evolutionary and phylogenetic components [1]. Taxonomy is the science responsible for the orderly arrangement of living organisms respecting a hierarchical system that can presume the evolutionary relationships; it also offers relevant information about the origin of life and how it evolved on Earth [2, 3]. The objective of general taxonomy is to establish a

classification system based on genealogical relationships, aiming to reach a natural system that mirrors the “order in nature” [4–6].

Prokaryotes include living organisms belonging to both domains, Archaea and Bacteria, known as archaeobacteria or archaea and eubacteria, respectively. Those microorganisms do not have a distinct nucleus or other organelles due to the lack of internal membranes, main characteristics distinguishing them from the eukaryotes. The prokaryotic taxonomy is traditionally split into three correlated areas: classification, nomenclature, and identification [7, 8]. The orderly arrangement of organisms into taxonomic ranks designates the classification. The nomenclature objective is

to name the organisms following the International Code of Nomenclature of Prokaryotes (ICNP) rules. The identification involves allocating the strains into the described taxonomic groups [7, 9, 10]. However, if the taxonomy of a strain remains undefined, the characterization and phylogenetic analysis should be carefully carried out to describe and name the new taxon.

The basic unit of taxonomy is the species. Bergey's Manual of Systematic Bacteriology defines a bacterial species as a group of strains with certain distinctive features that generally resemble each other in essential features of an organization [11]. Estimates are of about 1012 bacterial species on Earth [12]. In November 2021, there were about 17,845 valid species names (without synonyms) [13]. Considering this number, we can conclude that bacterial diversity is not sufficiently explored and that most might still be composed of still uncultivable species.

Besides arranging the organisms, the taxonomic tools are used to study microbial diversity and establish phylogenetic relationships. Biodiversity represents the basis of the stability of the ecosystems, providing environmental resilience [14]. Assessing microbial biodiversity also provides various biotechnological resources for food production and regenerative agriculture, environmentally-friendly technologies such as bioremediation and bioprospection of new enzymes, antibiotics, and biological processes. Studies on microbial diversity and taxonomy are crucial to guarantee the environmental and socioeconomic benefits be still unknown to the microbial world. In one attempt to put a monetary value on goods and services provided by ecosystems, the worth of biodiversity was estimated at the outstanding value of US\$ 33 trillion per year, close to the GDP (growth domestic product) of the United States and China combined [15]. Other estimates point out to lower or similar values, but undoubtedly, all of them are important [16].

Although many techniques, rules, and concepts are used for prokaryotes in general, in this review we describe how bacterial taxonomy evolved until genomic era and important tools developed to assess bacterial diversity and guide to proper classification. We also present a case study with rhizobia to clarify how the evolution of the taxonomic science impacted this group of bacteria, probably amongst the most important for ecosystems and agricultural sustainability, and improved our knowledge about them.

2. The Bacteria Species Concept and Description of New Taxa

The species concept is considered a universal theory limiting the category "species" for all living organisms. Concerning the prokaryotes, several incongruences are discussed, since they do not fit into the most common eukaryotic species perceptions, such as the morphological, biological, or evolutionary concepts [17]. For years, the bacterial species were classified based on phenotypic features, followed by polyphasic analyses involving phenotypic, genotypic, and phylogenetic properties, but with the methods differing among bacterial groups. As a result, each described taxon is represented by a lower taxon, and a type species represents each

genus. Nowadays, the advancement of next-generation sequencing (NGS) has encouraged the description of novel species based mainly on genomic sequences [18], and this practice is revolutionizing the bacterial species concept and the taxonomic groups [19–23].

For the description of novel bacterial taxa, the taxonomists follow guidelines from the International Committee on Systematic of Prokaryotes (ICSP), split into several subcommittees, according to the knowledge areas. Regarding the rhizobial species, the Subcommittee on Taxonomy of Rhizobia and Agrobacteria is responsible for the guidelines. For many years, the only guideline for rhizobia taxa description available was published by Graham and collaborators in 1991 [24]. More recently, subcommittee members proposed an improvement to the minimal standards for the species and genera description [25], provided on the website <https://sites.google.com/view/taxonomyagrorhizo/home>. The genomic profile comparison of strains studied and related type strains is required among the updates, and the genome sequence of the type strain representing the new species must be deposited in databases. This requirement will increase the number of genomes available for further studies.

Another critical step in taxonomy concerns the proper nomenclature, which the ICNP regulates. The scientific name of a novel bacterial taxon needs to be in Latin, referring to the history of the taxon, and be published in the "Approved List" of prokaryotes to become a valid name. In the International Journal of Systematic and Evolutionary Microbiology (IJSEM), the official journal of the ICSP, a clear statement of the name (i.e., fam. nov., gen. nov., sp. nov., etc.) and its etymology, as well as the characterization data of the taxon and the type strain designation, must be provided [26]. The valid names outside the IJSEM must follow these same guidelines and request the official journal validation [27].

The List of Prokaryotic names with Standing in Nomenclature (LPSN) is an online tool constantly updated by the Leibniz Institute DSMZ—German Collection of Microorganisms and Cell Cultures GmbH (<https://lpsn.dsmz.de/>). It includes a broad range of taxonomic information for each described taxon, such as etymology, nomenclatural and taxonomic status, type strain designations, and the link to the description publication [13]. However, it is worth mentioning that some prokaryotes, such as those still uncultivable, or genetic material retrieved from the environment, although might have sequences indicating that could represent new species, do not have sufficient information required to be described as a novel taxon [28]. Therefore, these organisms are classified as a "candidate" of a new species or a new genus, receiving the taxonomic status *Candidatus* [29]. Nowadays, with the development of metagenomics, which involves acquiring genome sequences from sampled materials of a community of organisms inhabiting a common environment, millions of 16S rRNA sequences have been deposited in databases [30], and a large number of them share less than the suggested threshold value of similarity for species circumscription. Therefore, even though the uncultivable prokaryotes have not been isolated yet, the analyses of those sequences imply that they might represent different species. Given this evidence, many

efforts have been carried out for gene sequences of uncultivable taxa to be considered by ICSP and to change the status of *Candidatus* to valid names [31].

3. A Brief View of Bacteria Taxonomy

Ferdinand Cohn developed the first studies of bacterial classification in the 1870s. He used morphologic traits as criteria to define six different genera [1]. From this, different methods of classification based on phenotypic traits were published, making the studies of taxonomists difficult because of the lack of minimal standards and common agreement on classification [17]. From the first edition of the Bergey's Manual of Determinative Bacteriology (1984–1989), this status was changed, with the definition of a reference guideline for bacterial classification, providing conditions for the microbiologists to merge the criteria adopted among them [32].

Even though the clustering methods emerged before the advent of computers, they were applied in bacteriology just after using computers, around 1957, by Sneath and Ludwig et al. [33, 34]. This era of bacterial taxonomy was marked by numerical studies where several phenotypic properties were electronically tabulated and used as a relatedness measurement. However, although the numerical taxonomy improved the identification of bacteria, it did not consider phylogenetic analyses [35]. Together with the evolution of numerical taxonomy, new techniques considering the physiologic and biochemical traits started to be developed, such as chromatographic and electrophoretic methods to define chemotaxonomic markers, aiming to reach the goal of a universal classification system for bacteria [17].

In the early 1960s, the development of molecular techniques supported the inclusion of methodologies such as the DNA guanine and cytosine content (GC mol%) and DNA-DNA hybridization (DDH) [36, 37] for taxonomic studies. Those techniques allowed the comparison of genomes, improving the classification of bacteria. Since the first experiments based on single-stranded DNA reassociation conducted by Schildkraut et al. in 1961 [38], and for about 50 years, the DDH was used as a standard technique for bacterial species circumscription [8, 9, 17].

A remarkable breakthrough in the attempts to determine relationships between distantly related organisms came around the 1970s when the Taq polymerase enzyme was discovered and used for DNA amplification through the polymerase chain reaction (PCR) techniques [39], and the dideoxy sequencing technology was described by Sanger et al. [40]. Almost simultaneously, the small ribosomal subunits 16S rRNA for prokaryotes and 18S rRNA for eukaryotes were described and started to be broadly used as molecular markers to organize all living organisms into three superior taxa, Archaea, Bacteria, and Eukarya, named as domains [41]. This new hierarchical taxonomic system based on these molecular markers allowed the incorporation of phylogenetic information in the prokaryotic classification [2, 42–45].

The phylogeny studies the evolutionary relationships among organisms, and using conserved molecular data

became commonly accepted in taxonomy. After the ribosomal sequences, other conserved genes started to be used [46]. Although less sequences can be used, the Multilocus Sequence Typing (MLST) generally evaluates 8 to 12 concatenated housekeeping genes or other protein-coding genes to identify genotypes and differentiate closely related strains, becoming broadly applied in molecular epidemiology. In the MLST, each gene (locus) contains different alleles; the allelic differences are converted in values resulting in the “sequence type” for the bacteria, which are available in a specific database for comparisons (as <https://pubmlst.org/>) of several pathogenic bacteria. Even though the MLST is not commonly used to infer phylogeny in epidemiology studies, it was applied for this purpose in taxonomic studies, contributing to the development of the Multilocus Sequence Analysis (MLSA) [47–52].

The MLSA accesses the evolutionary information of concatenated housekeeping genes to build phylogenetic trees with more robust data than the analysis with single sequences [53, 54]. It is an important tool for studying prokaryotic diversity and classifying taxonomic groups due to its high-resolution power on species delineation. The sequences are aligned, and informative sites are compared. As in the 16S rRNA analysis, the nucleotide identity (NI), which is a percentage of sequence similarity, must respect the threshold values suggested for each respective group of study [46, 55, 56]. The phylogenetic reconstructions can be achieved using different classes of inference methods. Each one has its particular strengths and weaknesses, requiring careful considerations to choose the best method that fits the analysis [25, 57].

Today, advanced sequencing technologies allow the taxonomy to use genomic data in silico to compare microorganisms, helping to allocate them in their respective taxa or describe new taxa to accommodate the new group. With sequenced genomes, the taxonomists can calculate the overall genome-related indices (OGRIs) and estimate the relatedness among microorganisms; however, suggested threshold values must also be considered. The OGRIs came to replace the DDH due to its low cost, reproducibility, and quality of genomic information. Furthermore, the genome sequences can be deposited in databases so that other scientists can use the data without cultivating the respective bacteria [51, 58]. The OGRIs include average nucleotide identity (ANI) and digital DNA-DNA hybridization (dDDH), the most broadly used to calculate the relatedness between orthologous sequences of two genomes. Those sequences are descended from the same ancestral sequence, which kept conserved across the evolution. Studies showing that ANI is a robust measure of evolutionary distance that correlates with DDH values have been published since 2005 [59, 60].

In conclusion, until the genomic era, the polyphasic taxonomy was used to identify, classify, and name prokaryotes according to phenotypic, genotypic, and phylogenetic characteristics. It enabled considerable progress and stability in microbial taxonomy. However, with advancements in genome sequencing, there are today better tools to delineate species, study phylogeny, and ordinate microbial

diversity. The history of microbial taxonomy incorporates the most advanced technologies and adheres to standards and rules, representing a scientific field where the progress goes alongside conservatism [61, 62].

4. Bioinformatics, Evolutionary Markers, and Threshold Values

The evolutionary molecular markers are constitutive genes that reflect the phylogeny of the organisms because the bases' substitution on DNA sequences (given by mutations and recombination) is proportional to the evolution that each species underwent from its ancestors, allowing estimates of the differentiation level of the species [63]. In bacterial taxonomy, those molecular markers started to be amplified, sequenced, and used to phylogenetically ordinate bacteria with the development of the PCR and sequencing techniques. The sequences from amplicons of molecular markers of strains under study and related strains need to be aligned using specific algorithms, such as MUSCLE [64] and ClustalW [65], to identify significant sites for the analysis.

In the next step, it is recommended to choose the best substitution model for the multiple sequence alignment, which depends on the phylogenetic method used to understand the phylogeny of the group under study. Models of substitution are algorithms responsible for evaluating the frequency of each nucleotide and its frequency of substitution, differentiating between transitions (exchange for nitrogen bases of the same biochemical class) and transversions (exchange for nitrogen bases of a different biochemical class). With this, the models can infer the evolutive history for the alignment [66, 67]. Finally, the phylogenetic methods are used to construct the phylogenies based on the alignment and best evolutive model. The result is a graphic representation of the evolution of the strains given by a phylogenetic tree.

In a phylogenetic tree, the extremities are represented by the investigated lineages (sequences, strains). The horizontal lines are called branches, and the nodes that connect the branches represent the most recent common ancestry among the strains [67]. However, a reliable phylogeny reconstruction needs to consider the number of sequences, the size of the alignment, the statistical support, and the outgroup. The bootstrap is commonly used in phylogenies to provide statistical support for the tree nodes [68]. This confidence test consists of multiple resampling of a dataset with the same size as the original alignment. For each resampling, the algorithm randomly chooses the sites, that is, the columns of the alignment, which may be repeated or excluded. Different resamplings result in phylogenetic trees that are compared with the tree obtained in the original alignment. The number of resamplings is variable and may be chosen by the researcher to ensure a reliable phylogeny [69]. The bootstrap value corresponds to the percentage of each group found among all trees obtained in the resampling. The outgroup represents a related taxonomic group, and it is used to show the ancestral strains, helping to determine the tree's root [70].

It is also possible to calculate the NI, a mathematic parameter used to measure the percentage of similarity among the nucleotide sequences from the alignment, but it does not include evolutionary analyses. Specific software is available to calculate the NI, such as BioEdit [71]. Several studies compare NI of particular sequences and genomic analysis, such as dDDH or ANI, of a large group of strains to suggest standard values of NI that reflect whole-genome features of prokaryotes, saving time and costs. These values are commonly known as threshold or cutoff values and are used for taxon circumscription [56, 72].

The 16S rRNA gene, a critical molecular marker in bacterial taxonomy, contains approximately 1,500 base pairs (bp) and plays a role in synthesizing essential proteins to the functioning of every prokaryote. It is originated from a common ancestor among all prokaryotes, being homologous and keeping conserved throughout the evolution process, but having variable sites with evolutionary information [42, 73]. The degree of divergence in the 16S rRNA sequences allows to estimate the phylogenetic distance among strains and, for this reason, is considered a molecular chronometer in bacterial taxonomy [1, 67].

Some numerical values of 16S rRNA NI have been suggested to delimit species boundaries. For example, Stackebrandt and Goebel, in 1994 [74] suggested that strains sharing less than 97% NI on the 16S rRNA sequences should not belong to the same species because the homology would not give a DNA-DNA reassociation above 60–70%. After that, a threshold ranging from 98.7 to 99% was recognized as corresponding to the DDH reassociation value for species delineation [1]. More recently, Kim et al. [72] carried out a large comparative study between 16S rRNA sequence similarities and ANI and proposed the threshold of 98.65% similarity in the 16S rRNA sequence for differentiating two species.

The increase in the number of novel taxa described using the 16S rRNA sequence data has revolutionized our knowledge of the microbial taxonomy, especially at the species level [51]. However, the resolution power of the 16S rRNA analysis is limited due to the predominance of conserved sites in these sequences, not giving enough evolutionary information for the analysis and restricting the identification at the genus level [6, 75–77], with the suggested threshold of 95% for genus delimitation [78]. In addition, studies have reported that those sequences can have multiple heterogeneous copies and, in some groups of microorganisms, the horizontal gene transference (HGT) may happen [79, 80]. These events may lead to misinterpretation of the evolutionary data [81]. Concerning these issues involving the ribosomal sequences, the information from the 16S rRNA phylogeny must corroborate with the suggested values of molecular analyses with a better resolution power and should not replace other methodologies.

Many taxonomists analyzed the internal transcribed spacer (ITS) as an alternative molecular marker to increase the knowledge about the ribosomal region [75, 82–84]. Those sequences are located between the small-subunit (16S rRNA) and the large subunit (23S rRNA) of the ribosomal RNA and contain highly variable sequence regions, allowing

the differentiation of bacteria even at the strain level, showing higher phylogenetic diversity [85–87]. For example, a study with the *Bradyrhizobium* genus demonstrated that strains sharing 95.5% NI on the ITS sequences would correspond to 60% of reassociation on the DDH, belonging to the same species [88]. However, as the 16S rRNA, these sequences do not have a high-resolution power to infer taxonomic groups at the species level. Despite being not often used, the 23S rRNA may also be applied as a molecular marker. The 23S rRNA is located in the large subunit of the prokaryotic ribosome, has a conserved function, universal distribution, and presents sequences with different levels of variation. However, it is commonly characterized by showing insertions and deletions with large sizes, resulting in more information to improve phylogenetic resolution as an additional analysis [66, 89].

Even though the 16S rRNA sequences have been broadly used as an effective tool for basic evolutionary analyses of cultivable and uncultivable bacteria, for closely related groups, they are unable to determine their nearest neighbors since different species can share identical or nearly identical 16S rRNA sequences [83, 90–93]. An approach that can be used to fill this gap is the MLSA. As mentioned before, the MLSA puts together evolutionary information from concatenated housekeeping genes to better assure the correct taxa identification of strains under study. The housekeeping genes are responsible for fundamental roles in the basal microbial metabolism, encoding for essential proteins of the cell functions. Therefore, they have high conservation levels but faster evolutionary rates when compared with the ribosomal sequences, carrying more phylogenetic information [94].

The MLSA allows the analysis of genes together as a larger phylogenetic dataset [35, 46, 52, 95]. Therefore, the methodology aggregates sufficient phylogenetic signals, offering a buffer effect to the noncongruent signals that influence smaller datasets, being more reliable [77, 94, 96]. However, it is worth mentioning that the single housekeeping genes analysis does not reflect the entire evolutionary pattern of the strain; it is recommended to proceed to the comparison between the 16S rRNA and each housekeeping gene phylogeny to detect and avoid potential HGT and recombination events, since each housekeeping gene may have a different history of evolution [75, 92].

The main requirements of MLSA involve the selection of housekeeping genes that should be present in the genome of all organisms object of study as a single copy and spread in the genome. They must also have a consistent size to allow phylogenetic reconstructions and sequencing. Consequently, different genera may vary in the set of genes used in the analysis [46, 52]. In addition, the resulting data from MLSA must corroborate with the 16S rRNA, dDDH, and ANI analyses. It is recommended to use at least five housekeeping genes, considering that the number of genes reflects directly in the discriminatory power of the technique [52]. Some of the most common housekeeping genes applied for the classification of new taxa are ATP synthase β -subunit (*atpD*), chaperone protein (*dnaK*), glutamine synthase II (*glnII*), glutamate synthase (*gltB*), DNA gyrase β -subunit

(*gyrB*), recombinase A (*recA*), RNA polymerase β -subunit (*rpoB*), and tryptophan synthase β -subunit (*trpB*) [10, 75, 97–99]. In addition, new genes can be studied, aiming to determine if they provide relevant and congruent phylogenetic information as those genes cited above.

After the housekeeping genes selection, each set of single-gene sequences should be aligned and trimmed to keep the same region of comparison and the same size for the alignment. Subsequently, the phylogeny of each single gene is individually built and compared to each other, and if they are congruent, the alignments must be concatenated to proceed with the MLSA. The concatenation process can be carried out manually using software or any text editor program. Some of the most common software for alignment of prokaryotic sequences are MEGA (Molecular Evolutionary Genetics Analysis) [100] which provides alignment tools as MUSCLE and ClustalW, and BioEdit [71] which also provides the alignment tool ClustalW. To concatenate the sequences, the SeaView [101] is commonly used. In addition, it is important to search for updated sequences in databases to construct reliable phylogenies, verifying the quality of the sequences and the availability of the sequences of the closely related strains.

In 2002, the ad hoc committee of the ICSP recommended the analysis of the concatenated housekeeping genes as a promising method to replace the DDH association in bacterial taxonomy [102]. Following this, Konstantinidis et al. [56] studied extensive whole-genome comparisons of four important bacterial groups. They concluded that the phylogeny of six to eight genes reflects the threshold of 70% DDH and 96% ANI for species circumscription. The authors did not suggest any specific set of genes and affirmed that even if the genes are randomly combined, they provide a robust phylogeny for the group studied [56, 103]. Since then, the NI of concatenated housekeeping genes has been accepted to predict the whole-genome information and differentiate bacterial groups [98, 104–106]. However, no universal threshold for MLSA has been determined for species circumscription yet [18]. Nevertheless, the MLSA has been used as a great strategy to differentiate species in bacterial taxonomic studies, and it is considered as a critical step of bacterial taxonomy since it is more phylogenetic sensible than 16S rRNA analysis and involves conserved regions able to infer phylogeny, unlike OGRIs, which are mathematic parameters [51].

5. Evaluating Genomic Traits

As the DNA molecules represent the identity of the species, studying the genomic profiles allows obtaining relevant information for taxonomic purposes. Prokaryotic genomes contain repetitive sequences distributed throughout the chromosome; however, the sites, length, and the number of times they are repeated are characteristic of each strain, representing a fingerprint of each genome. To evaluate those genomic profiles, the taxonomists use DNA amplification by PCR with specific primers for those regions or restriction enzymes to cut the chromosome in the restriction sites. As a result, in both procedures, there is a mixture of different

fragments of DNA that can be separated by electrophoresis revealing the respective genetic profile [9, 107, 108]. For taxonomic studies, there are three main sets of repetitive elements: repetitive extragenic palindromic (REP) [109], enterobacterial repetitive intergenic consensus (ERIC) [110], and BOX elements [111–113]. Those analyses are limited for species circumscription, but if properly used, they can be applied to identify variability from strains of the same species, to find clones, and to authenticate strains in a study or culture collection [8, 108, 114].

The DDH evaluates the extension and stability of DNA hybrid strands after the dissociation and consecutive reassociation of a two-genomes mixture incubated under controlled conditions [115]. The cutoff is based on the percentage of reassociation using the difference in melting temperature of the hybrid DNA strand (heterologous) compared to 100% of the original strand (homologous). If this association is 70% or higher, or the melting temperature difference (ΔT_m) is below 5°C, the genomes compared belong to the same species [7, 18, 35]. The DDH started to be applied when a putative novel group shared more than 97% NI in the 16S rRNA sequence with the related species, aiming to assure enough difference among the genomes of the strains to support the description of a new group. Since the introduction of the DDH in microbiology [116], it became the “gold standard” for species delineation of Bacteria and Archaea and a required technique in every prokaryote species description [8]. However, despite the cited advantages, the DDH has limitations, especially for being time-consuming, requiring intensive labor, and not allowing comparison of results obtained in different experiments, not enabling establishing a global database [44, 51, 117, 118].

The first prokaryote genome sequenced was of the bacteria *Haemophilus influenzae* in 1995, using the conventional Sanger sequencing technique [119]. Although a new genomic era was starting, the high cost and time-consuming process of genome sequencing did not allow significant progress for one decade. In 2005, with the arrival of the NGS technologies, also known as high-throughput sequencing methods, the easy access and low cost of the genome sequencing launched the report of sequenced genomes of prokaryotes [51, 120]. Today, many sequencing platforms can be used, and they are split into two main groups, depending on the type of template used for the sequencing reactions, the high-end instruments, and the bench-top instruments. The high-end platforms demand more technology infrastructure for data tracking, analysis, and storage, while the bench-top ones have more modest requirements. Nowadays, the most popular platforms are the MiSeq and HiSeq (Illumina, USA), Ion Torrent (Thermo Fisher Scientific, USA), and Pacific Biosciences (USA) [60].

The statistical parameters used to report the quality of the genome assembly recommended for taxonomic purposes are (i) the genome size, defined as the length of all contigs sequenced; (ii) the N50, defined as the length of the shortest contig that accumulatively shows 50% or more of the genome size when the lengths of the contigs are summed from the largest to shortest; and (iii) the depth of coverage from the sequencing, indicating how many sequencing reads

are generated; this value is usually given as folds and is recommended a minimum of 50X (50-fold) for the platforms cited before [60, 121]. Another critical issue to consider in taxonomy is the genome authenticity, which can be achieved by comparing complete conserved sequences, as the 16S rRNA or housekeeping genes extracted from the genome, with the same sequence obtained through the conventional Sanger sequencing. When describing a new species, this comparison should be performed at least with the type strain [51].

Presently, using genome sequencing, the taxonomists can use other analyses to study the relatedness among the DNAs of bacteria. The threshold values suggested for these analyses relate to the 70% from the DDH technique. As mentioned above, these values are known as OGRIs and effectively calculate genomes' similarities *in silico* [51, 58]. The ANI and the dDDH are now extensively used to replace the original DDH, adopting the threshold values of 95–96% and 70%, respectively, for species circumscription. The OGRIs can be calculated using software tools for taxonomic purposes. Today there are many readily available web services, such as the Genome-to-Genome Distance Calculator and the ANI calculator, and standalone tools, such as JSpecies and OrthoANI with USEARCH. Besides being fast and of low cost, the high quality of the sequences allows the improvement of genomic databases that researchers can use worldwide [60, 117, 122, 123].

With the availability of genome sequences, another parameter that can be calculated *in silico* is the DNA guanine and cytosine (GC) content, which is also used as a genotypic marker in taxonomy [124]. In the beginning, the bases' composition variation from strains of the same species should not exceed 3 mol%, and for strains from the same genus, 10 mol% [7]. Later, a survey comparing the GC content with the genomic data and dDDH indicated that the variation should not exceed 1 mol% among strains of the same species [58]. The GC content can be calculated using tools to assemble the genomes, such as the SEED viewer provided by the RAST server [125], QUAST [121], and BioEdit [71]. Although the GC content is important to distinguish nonrelated bacteria, similar DNA base composition does not necessarily imply that the two strains are closely related [35].

6. Phenotypic Traits

In contrast to the taxonomy of eukaryotes, where the phenotypic characteristics can be used to differentiate some organisms, these traits are questionable in prokaryotes since different bacterial species can present identical phenotypes [17]. Nevertheless, starting the prokaryotic classification using the phenotypic evaluation may help delimit further analyses; for example, most pathogenic bacterial groups present well-established phenotypes in the literature and are critical to quick diagnostics.

The classical phenotypic tests in microbiology include morphological, physiologic, and biochemical analysis. Morphological characteristics describe the cellular and colony features, such as the cell shape, endospore formation,

presence of flagella, Gram stain, color colony, diameter, opacity, mucus production, and their consistency. On the other hand, the physiological and biochemical characteristics include data about the culture under different growth conditions, such as range of temperature, pH (4–12), salinity tolerance (1–10%), O₂ or CO₂ requirements, tolerance to different antibiotics, enzymatic activity (i.e., urease), and metabolism of compounds (i.e., carbon source and nitrogen source) [7, 9, 28, 126]. Therefore, it is recommended to compare the data obtained with reference or type strains and to include negative and positive controls [8]. In addition, the use of commercial tests with minimal standards such as API (bioMérieux) or Biolog (Biolog, Inc.) that evaluate the carbon source utilization by bacteria is suggested to avoid incongruences among laboratories.

Another common phenotypic test is the chemical characterization of cells, which evaluates extracellular elements (peptidoglycan, teichoic acids, and mycolic acids), cell membrane composition (fatty acids, polar lipids, respiratory lipoquinones, and pigments), or cytoplasm compounds (polyamines) [8]. These features are also known as chemotaxonomic markers because they are usually stable within a bacterial group [127]. The resulting data are available in specific databases. One of the most common is the database of fatty acids profiles implemented by Sherlock Microbial Identification System (MIDI, Inc.) [8, 105, 128]. Even though the chemical characterization provides important information about the cells, it has become unusual in taxonomic studies because the tests are laborious, require specific equipment and methodologies, and usually allow the classification only at the genus level identification. Lately, the use of Matrix-Assisted Laser Desorption/Ionization Time-of-Flight (MALDI-TOF) for bacterial identification has become a rapid, precise, and cost-effective method, especially compared to traditional phenotypic and molecular techniques. It is based on the analyses of bacteria components profiles, as proteins directly extracted from intact bacteria, and compared with a reference database for identification [129, 130]. This methodology is commonly used on clinical isolates' identification, but as the reference databases are being updated, it is becoming practicable for other groups of bacteria, such as the rhizobia [131–133].

It is worth mentioning that there are many genes coding proteins without known function. Therefore, the phenotypic tests could help the search for proteins with biotechnological interest, improving the knowledge about the interactions of microorganisms with the environment [18]. However, the phenotype results from the combination of genotype and environmental conditions. Consequently, it is impossible to know the whole phenotype of a prokaryote based only on observable characteristics or on the information obtained from the genome. In taxonomy, it is common to find some incongruent data in phenotypic evaluations, probably because the accessory genome codes many phenotypic traits; for example, the plasmids easily exchanged among the organisms. In addition, as there are no laboratory conditions able to mimic the environment entirely, the phenotypic

characterization can exclude some microorganisms, such as the uncultivable. Therefore, these data need to be validated by genomic analysis [134].

7. Rhizobia Study Case

7.1. History of Rhizobia Taxonomy. More than 2,000 years ago, ancient Chinese literature reported that crop rotation of legumes with cereals was traditionally used to enhance grain production. Improvement in soil fertility by cultivating legumes was thus already noticed at that time, although the mechanisms involved were not known yet [28]. However, it was just by the end of the nineteenth century that the assimilation of atmospheric nitrogen was related to the root nodules in legumes, and in 1888 the first root-nodule bacterium was isolated from nodules of *Pisum sativum* plants (pea) and reported by the Dutch microbiologist Martinus Willem Beijerinck to be responsible for the nitrogen fixation process. The isolated bacterium was first named *Bacillus radicola* but later reclassified as *Rhizobium*, comprising the *Rhizobium leguminosarum* species [28, 135, 136]. Since then, the nodulating nitrogen-fixing bacteria have been generally called rhizobia [137].

In the early twentieth century, nodulation tests using a broad range of host plants and different bacteria were conducted, and the specificity between the host plants and the symbiotic bacteria was reported. Based on this, Baldwin and Fred [138] proposed the cross-nodulation concept, indicating specificity between rhizobia and host plants. For about 80 years, taxonomists used this concept for species definition, and six main species were described: *R. leguminosarum* [136], *R. phaseoli*, *R. trifolii*, and *R. meliloti* [139], which produced acid reaction on yeast-extract mannitol agar (YMA) medium, and *R. lupini* [140] and *R. japonicum* [141], which produced alkaline reaction on the same medium. Besides those six species, some other strains isolated from cowpea were defined as *Rhizobium* spp. [135, 142].

The taxonomy dropped the cross-nodulation concept after several studies reporting both exceptions and strains sharing high similarity and belonging to different groups. Additionally, the rhizobia classification needed to include more information to adjust to the general bacterial taxonomy [143–145]. However, even after the cross-inoculation concept was dropped in rhizobia taxonomy, it continued to be studied, representing an important feature regarding the inoculation and efficiency of the symbiosis [28].

The next step of rhizobia taxonomy was based on the numerical taxonomy using computers to compare bacteria properties. Around the 1960s, many analyses were included in taxonomic studies involving phenotypic traits, growth conditions, nutrient resources, metabolic features, and resistance to antibiotics and other chemicals, among others. Also, the DNA molecule started to be investigated, and the base composition (GC mol%) was added to bacterial classification [145–148]. Using the numerical taxonomy combined with DNA analyses, three other rhizobial genera were described. In 1982, Jordan [149] proposed the description of the genus *Bradyrhizobium* to allocate the slow-growing

species *B. japonicum* and *B. lupini*. Six years later, the genus *Azorhizobium* [150] was described, encompassing strains that can effectively nodulate roots and stems of *Sesbania rostrata* and fix nitrogen under free-living aerobic conditions. In the same year, the *Sinorhizobium* genus was proposed for the fast-growing soybean species *Rhizobium fredii* [151].

Around the 1990s, many other analyses were included in taxonomic studies. The polyphasic taxonomy confirmed some of the taxa proposed with the numerical taxonomy but also pointed out that the numerical taxonomy lacked information about the evolutionary relationships among rhizobia. Consistent DNA studies allowed the taxonomists to assess the diversity and phylogenetic relationship among bacteria at a molecular level [28]. The 16S rRNA sequence analysis was included in rhizobia species descriptions and reclassifications in 1991 when Graham and collaborators [24] published the minimal standards for species description of new rhizobia and *Agrobacterium* [152]. Considering the phylogeny and similarity of 16S rRNA, rhizobia were classified in the phylum Proteobacteria, subdivision α -Proteobacteria [153]. In the same decade, two more genera were described. First, Jarvis et al. [154] proposed the genus *Mesorhizobium* to allocate five *Rhizobium* species with an intermediate growth rate than fast-growing *Rhizobium* and *Sinorhizobium* and slow-growing *Bradyrhizobium*. Following, de Lajudie et al. [155] proposed the description of the genus *Allorhizobium* of symbiotic bacteria associated with the aquatic legume *Neptunia natans*.

The six rhizobia genera were allocated in four distinct families: *Bradyrhizobium* [149], *Mesorhizobium* [154], and *Azorhizobium* [150] belonging to families Bradyrhizobiaceae (today Nitrobacteraceae), Phyllobacteriaceae, and Xanthobacteraceae, respectively, and *Rhizobium* [136], *Sinorhizobium* (reclassified as *Ensifer*) [151, 156], and *Allorhizobium* [155] belonging to the family Rhizobiaceae [157]. Additionally, a study with members of the family Rhizobiaceae reported similarities of the 16S rRNA gene higher than 92% among the genera, suggesting that this value could be a threshold for family delineation [158].

The new century started a revolution on the rhizobia taxonomy. A first milestone occurred in January of 2001, with the first report of a non-rhizobia nitrogen-fixing legume-symbiotic bacterium isolated from the nodules of *Crotalaria*, classified as *Methylobacterium* [159], with the species *M. nodulans* [160]. In the same month, based on the 16S rRNA analysis [161], Young et al. [162] suggested the inclusion of all species of *Agrobacterium* [152] and *Allorhizobium undicola* [155] in the genus *Rhizobium*. In 2002, a new species of the genus *Devosia* was reported to induce nitrogen-fixing root-nodules in *Neptunia natans* [163]. In 2005, the first rhizobia in the genus *Phyllobacterium* were described as *P. trifolii*, isolated from the nodules of *Trifolium pratense* [164]. From 2005 to 2007, the genus *Ochrobactrum* (today, *Brucella*) [165, 166] allocated two nodulating species, *O. lupini* and *O. cytisi* [167, 168], isolated from nodules of *Lupinus albus* and *Cytisus scoparius*, respectively. In 2008, Lin and collaborators [169] described in the *Shinella* genus [170] the symbiotic bacterium *S. kummerowiae* isolated from

root nodules of *Kummerowia stipulacea*. In 2009, the first rhizobial isolate (BA135) belonging to the species *Aminobacter aminovorans* was reported, isolated from nodules of *Lotus tenuis* [171]. In the following decade, from 2012 to 2014, the genus *Microvirga* [172] allocated four nodulating and nitrogen-fixing new species, *M. lupini* isolated from *Lupinus texensis*, *M. lotononidis*, and *M. zambiensis* isolated from *Listia angolensis*, and *M. vignae* isolated from *Vigna unguiculata* [173, 174].

Besides all those changes in the α -Proteobacteria class, 2001 was outstanding by the report of a nodulating β -Proteobacteria [175], belonging to the genus *Burkholderia*, described by Yabuuchi and collaborators in 1992 [176]. It was the first time that the essential nodulation genes (*nod*) and the nodulation capacity were reported in symbiotic bacteria not belonging to α -Proteobacteria [175]. After that, several nodulating bacteria of β -Proteobacteria genera were described, including the *Ralstonia* genus [177], reclassified in 2004 as *Cupriavidus* [178–180], in addition to several species of the *Burkholderia* genus, later reclassified as the new genus *Paraburkholderia* [181, 182]. In the following years, the introduction of single and concatenated housekeeping genes in phylogenetic studies culminated in the reclassification of many species and the proposal of the new genera *Neorhizobium* and *Pararhizobium*, and also the revision of the genera *Agrobacterium* and *Allorhizobium* [183, 184]. More recently, using whole-genome analyses, Santos and collaborators [21] suggested the description of two new genera, *Mycetohabitans* and *Trinickia*, this last one containing the nodulating nitrogen-fixing species *T. symbiotica*.

With the evolution of taxonomic analyses, we may conclude that many descriptions of nodulating bacteria, isolated from nodules of different hosts and belonging to nonrhizobial genera have been published, and many taxonomic groups were reclassified. Most of those bacteria have a diverse set of nodulation (*nod*) and nitrogen fixation (*nif*) genes, some of which are related to genes of different members of classical rhizobial genera. All those findings show that the ability to establish symbiosis with legumes is more widespread in bacteria than anticipated before [137, 175].

Today, rhizobia are distributed in eight families, seven belonging to the α -Proteobacteria class and one in the β -Proteobacteria. From both subclasses, the α -rhizobia are reported as broadly distributed, both in geography and host-plant range, and although the β -rhizobia are well established, their distribution seems more restricted [28]. The α -Proteobacteria families are the Rhizobiaceae that allocate seven genera, the Phyllobacteriaceae with three genera, the Methylobacteriaceae with two genera, and the other four families allocating only one genus, Nitrobacteriaceae (old Bradyrhizobiaceae), Brucellaceae, Hyphomicrobiaceae, and Xanthobacteriaceae. The β -Proteobacteria family is Burkholderiaceae with three genera. The list of genera and the number of species with valid names standing in nomenclature (without synonyms) according to the LPSN in October of 2021 are listed in Table 1.

In 2020, a study performed phenotypic, genomic, and phylogenetic analyses of the genus *Ensifer* and suggested that

TABLE 1: List of genera comprising rhizobial species with valid names standing in nomenclature (LPSN) in October 2021. The number of valid names does not include synonyms.

α -Proteobacteria		
<i>Hyphomicrobiales</i> (=Rhizobiales)		
Rhizobiaceae		
<i>Agrobacterium</i>	11 valid names	2 rhizobia
<i>Allorhizobium</i>	8 valid names	2 rhizobia
<i>Ensifer</i> (=Sinorhizobium)	20 valid names	20 rhizobia
<i>Neorhizobium</i>	5 valid names	4 rhizobia
<i>Pararhizobium</i>	6 valid names	2 rhizobia
<i>Rhizobium</i>	92 valid names	49 rhizobia
<i>Shinella</i>	8 valid names	1 rhizobia
Phyllobacteriaceae		
<i>Aminobacter</i>	7 valid names	1 rhizobia
<i>Phyllobacterium</i>	12 valid names	4 rhizobia
<i>Mesorhizobium</i>	59 valid names	48 rhizobia
Nitrobacteraceae (=Bradyrhizobiaceae)		
<i>Bradyrhizobium</i>	63 valid names	58 rhizobia
Methylobacteriaceae		
<i>Microvirga</i>	18 valid names	4 rhizobia
<i>Methylobacterium</i>	47 valid names	1 rhizobia
Brucellaceae		
<i>Brucella</i> (=Ochrobactrum)	25 valid names	2 rhizobia
Hyphomicrobiaceae		
<i>Devosia</i>	26 valid names	1 rhizobia
Xanthobacteraceae		
<i>Azorhizobium</i>	3 valid names	2 rhizobia
β -Proteobacteria		
Burkholderiales		
Burkholderiaceae		
<i>Cupriavidus</i>	18 valid names	2 rhizobia
<i>Paraburkholderia</i>	79 valid names	22 rhizobia
<i>Trinickia</i>	7 valid names	1 rhizobia

the genus should be separated into two genera, one for the symbiotic clade and the other for the nonsymbiotic clade [185]. More recently, a publication on the IJSEM suggested the revision of family Rhizobiaceae [186]. The extensive study suggests a threshold for core-proteome average amino acid identity (cpAAI) of approximately 86% as a new framework for genus delimitation. It is noteworthy to mention that cpAAI must not be used as sole information for genus delimitation, and the authors specify “approximately 86%” to provide some flexibility regarding the evolution of each genus. Although not yet updated in the LPSN website as validated names, among the reclassifications based on the study of Kuzmanović and collaborators [186], there are arguments that the unification of the genera *Ensifer* and *Sinorhizobium* is no longer justified, and eight new combinations were suggested, but not all involving rhizobial strains [186].

From Table 1 based on the LPSN, we may conclude that today over 200 known rhizobial species are split into 19 genera of the α - and β -Proteobacteria subclasses, and the number increases every year. However, less than half of these valid names from the 19 genera have species comprising strains already reported for their symbiotic properties,

including nodulation and nitrogen fixation abilities. Furthermore, many species are reported as endophytes, or were isolated from environmental samples, or from nodules but unable to reestablish symbiotic associations. Therefore, the symbiotic capacity remains largely unknown for many species.

Additionally, in 2004, Benhizia and collaborators [187] published for the first time the isolation of γ -Proteobacteria species from legume nodules. In this study, 52 isolates belonging to the *Pseudomonas*, *Escherichia*, *Leclercia*, *Pantoea*, and *Enterobacter* genera were isolated from three *Hedysarum* species, and rhizobia-like bacteria were found occupying the nodules. However, Koch’s postulates and the symbiotic parameters from the isolates were not investigated. Shiraishi and collaborators [188] also reported in 2010 the existence of rhizobial strains in the γ -Proteobacteria subdivision. The authors described *nod* and *nif* genes in the *Pseudomonas* sp. strain Ch10048, sharing high similarity with the symbiotic sequences of *Agrobacterium* sp., suggesting the acquisition of these genes through HGT from rhizobial species in the soil. Despite the report of *nod* and *nif* genes in strain Ch10048, and the confirmation of the ability to nodulate the host legume *Robinia pseudoacacia*, the existence of γ -rhizobia remains controversial until additional evidence confirms that the genes were not provided by other bacteria coexisting in the nodules and that the nitrogen fixation ability of the strain is tested [28].

Undoubtedly, rhizobia taxonomy advanced together with prokaryotes’ taxonomy, and improvements regarding the origin and evolution of these bacteria were obtained. However, there is a need to increase the studies relating taxonomy and phylogeny with the phylogeny of nitrogen fixation and biotechnological properties of rhizobia.

8. Rhizobial Symbiotic Parameters and Genome Architecture

As commented before, the members of the Subcommittee on Taxonomy of Rhizobia and Agrobacteria of the ICSP reviewed the taxonomic developments for this group of bacteria and updated the minimal standards for taxonomic studies, including additional considerations specific to rhizobia and agrobacteria. According to them, taxonomic definitions should not include symbiotic or pathogenic characters because the interactions with plants are determined by accessory genes that may be present in several bacterial species, and be gained or lost, imposing taxonomic limits [25]. Instead, if a strain has a phenotype regarding plant-interaction, this should be described in the taxonomic proposal but considered as a property of the strain, not of the whole taxon. It is worth mentioning that symbiovars are also studied in rhizobial surveys, although not accepted as a formal taxonomic category. The term symbiovar was proposed in 2011 to name a group of strains able to nodulate and fix nitrogen with a range of legumes or a specific legume, and today this definition is based mainly on symbiotic genes’ phylogeny [25, 189].

Concerning the description of new rhizobial species, it is especially recommended to evaluate the symbiotic ability of

the strains based on Koch's postulates using the original host and/or other legume species. This last alternative may be used to expand the information about the host range of the strains and to define symbiovar groups or when the seeds of the original host plant are not available. The species *Phaseolus vulgaris*, *Macroptilium atropurpureum*, *Vigna unguiculata*, and *Mimosa pudica* are promiscuous legumes commonly used in taxonomic studies of rhizobia. The symbiotic ability may be evaluated compared to negative controls by the presence/absence of root nodules, plant biomass, N content, or the acetylene reduction assay. In addition, the strains must be reisolated from the nodules, keeping the original phenotypic, phylogenetic, or genotypic features, obeying Koch's postulates [25, 190, 191].

Based on the meta-analysis of 1,708 completed bacterial genomes performed in 2017 by diCenzo and Finan [192], the average and median of bacterial genomes found were 3.65 Mb and 3.46 Mb, respectively. In a review study of 2020, Geddes and collaborators [193] compared the genomes of representative strains of α -rhizobia and β -rhizobia and showed that rhizobial genomes range from 3.42 Mb in *Cupriavidus taiwanensis* LMG 19424 to 9.36 Mb in *Microvirga lupini* Lut6. However, the authors highlighted that some strains of *Bradyrhizobium*, *Mesorhizobium*, and *Azorhizobium* genera might have higher genomes, which means that the rhizobial genomes can be twice or more times higher than the average size of bacterial genomes reported in the two studies [192, 193]. In contrast, another study reported that *Ensifer* strains from the symbiotic clade carried an average of 325 fewer genes and appeared to have fewer rRNA operons when compared to strains belonging to the nonsymbiotic clade [185]. Large genomes may be related to adaptations to the soil and rhizosphere conditions [194].

In general, the rhizobial genes responsible for plant infection, nodulation, and nitrogen fixation are clustered together in symbiotic plasmids or symbiotic islands in the chromosome, or even in both genomic regions [195–200]. Those clusters are frequently part of mobile genetic elements (MGE) that have independent evolutionary pathways [201]. An exception was first reported in 2007 revealing a group of *Bradyrhizobium* strains with photosynthetic ability that does not possess nodulation genes and can induce nodulation without nodulation (Nod) factors [202]. It is known that Nod factors play a crucial role in host specificity in the rhizobia-legume interactions; those molecules differ on the symbiosis specific backbone length and other structures, determining the set of plants that the rhizobia can nodulate [203, 204]. In a review recently published, Patra and Mandal [205] pointed out other studies reporting that even in absence of Nod factors, other bradyrhizobia strains, not belonging to the photosynthetic group, are also able to establish successful nodulation. Among the hypotheses discussed, there is the possibility that Nod factors independent nodulation start with the host infection through crack invasion process, instead of the formation of the common infection thread. After the infection, the nodulation might take place using similar signals and mechanisms present in Nod-dependent nodulation [205, 206].

Besides the main chromosome, some bacteria have a "second chromosome" or "megaplasmid," for which the term "chromid" was proposed. These elements have some core genes and nucleotide composition similar to the associated chromosomes, but most of their genes are accessory. Some rhizobia and agrobacteria also have genus-specific chromids, similar within a genus but with different sets of conserved genes among genera. An example is some *Agrobacterium* species with linear chromids carrying a unique replication system and conserved genes [207].

9. Rhizobial Origin Hypothesis and Evolution

As biological nitrogen fixation is considered one of the most important biological processes for life on Earth, there is great biotechnological interest in diazotrophic bacteria [208]. Studying the nodulation ability and nitrogen fixation efficiency, together with the phylogenetic comparison of core and symbiotic genes, gives insights about the origin of the diazotrophic bacteria and the evolution of the biological fixation ability on prokaryotes [106, 209, 210].

Remigi and collaborators [201] reviewed that the nitrogen fixation ability is older than the nodulation process and broadly spread among Bacteria and Archaea. The nodulation genes might have clustered early in the symbiosis evolution path by duplication and specialization of other functional genes. Nowadays, the high diversity of nodulation genes makes it difficult to suggest which bacterial lineage was ancestral. The *nif* and *nod* genes have different phylogenies, implying that rhizobia inherited the nitrogen fixation ability of their free-living relatives [211]. Besides, as the proximity of those genes is not essential for function, it suggests a relatively recent HGT event as a symbiosis set.

Evidence indicates that the *Bradyrhizobium* genus might be the rhizobia's ancestor [212]. Using the genes from the glutamine synthetase enzymes (GSI and GSII), essential for nitrogen assimilation, the estimates are that *Bradyrhizobium* originated 553 million years ago (MYA). Other rhizobia evolved 400–324 MYA, originating the *Mesorhizobium*, *Rhizobium*, and *Sinorhizobium* (= *Ensifer*) genera. Interestingly, the first legumes ascended on Earth long after, around 70 MYA [63, 213]. Another piece of evidence of *Bradyrhizobium* ancestry is that some strains were detected with nitrogen fixation ability as free-living bacteria, as observed in some *Azorhizobium*, and both lineages are very distant from the other rhizobial genera [28].

It is well known that bacteria have different mechanisms to exchange genetic material. This event is more recurrent among organisms sharing the same ecological environment, reinforcing that some rhizobia evolved by acquiring symbiosis genes from other species by HGT [197, 214, 215]. Furthermore, a study with *A. caulinodans* reported an increase in horizontal transference frequency of its symbiosis island in the legume rhizosphere or in the presence of plant flavonoids, suggesting a host-dependent evolution [216]. Over evolutionary time, the horizontal transference of symbiotic functional genes among symbiotic and non-symbiotic bacteria is hypothetically responsible for the growing number of studies reporting great rhizobia diversity

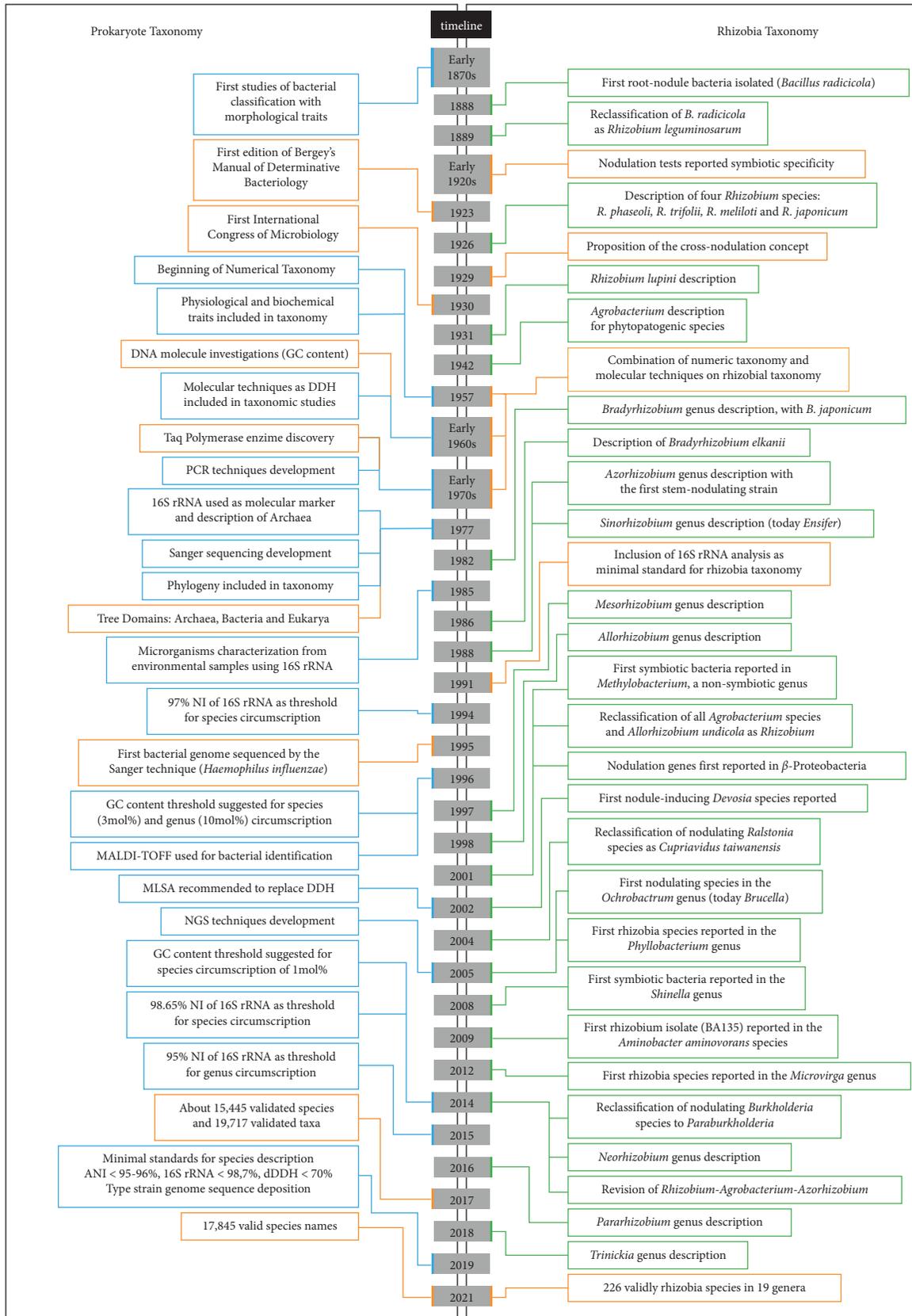


FIGURE 1: Timeline showing the evolution of methods and achievements in prokaryote and rhizobia taxonomy.

[201, 217]. Given the vast population of rhizospheric bacteria, it might seem paradoxical that the symbiosis is restricted to only nineteen genera. However, an increasing number of studies report the coexistence of nonrhizobial bacteria inside the nodules, which deserve further studies and might help to explain that additional partners help the symbiosis [153, 218–222].

10. Conclusions

As presented in this review, the main goal of taxonomy is to ordinate living organisms in a stable and hierarchical system. As shown in Figure 1, remarkable progress has been achieved in both prokaryote and rhizobia taxonomy and phylogeny. However, profound changes may arise with the genomic era. Nevertheless, robust taxonomic methodologies are becoming gradually available in an increasing number of laboratories, allowing researchers to conduct surveys of great interest. These studies have contributed to new insights about the origin, evolution, and diversity of bacteria on Earth and the description of almost 18,000 valid species, of which more than 220 are rhizobia. However, more studies are needed to correlate taxonomy with biotechnological properties of nitrogen-fixing rhizobia to improve their contribution to agricultural and environmental sustainability.

Abbreviations

ANI:	Average nucleotide identity
DDH:	DNA-DNA hybridization
dDDH:	Digital DNA-DNA hybridization
ERIC:	Enterobacterial repetitive intergenic consensus
GDP:	Growth domestic product
HGT:	Horizontal gene transference
ICNP:	International Code of Nomenclature of Prokaryotes
ICSP:	International Committee on Systematic of Prokaryotes
IJSEM:	International Journal of Systematic and Evolutionary Microbiology
ITS:	Internal transcribed spacer
LPSN:	List of prokaryotic names with standing in nomenclature
MEGA:	Molecular evolutionary genetics analysis
MGE:	Mobile genetic elements
MLSA:	Multilocus sequence analysis
MLST:	Multilocus sequence typing
MYA:	Million years ago
NGS:	Next-generation sequencing
NI:	Nucleotide identity
OGRI:	Overall genome-related indices
PCR:	Polymerase chain reaction
REP:	Repetitive extragenic palindromic.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publishing of this paper.

Authors' Contributions

Luisa Caroline Ferraz Helene and Milena Serenato Klepa contributed equally to this work. The authors declare that they have consented to participate in the manuscript and publish it.

Acknowledgments

L.C.F. Helene acknowledges a postdoctoral fellowship from CAPES (INCT). M.S. Klepa acknowledges a PhD fellowship from CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Finance Code 001) and M. Hungria acknowledges a research fellow from CNPq (Brazilian Council for Scientific and Technological Development, 303026/2020-0). This study was partially financed by INCT—Plant-Growth Promoting Microorganisms for Agricultural Sustainability and Environmental Responsibility (CNPq 465133/2014-4, Fundação Araucária-STI 043/2019, CAPES).

References

- [1] E. Stackebrandt and J. Ebers, "Taxonomic parameters revisited: tarnished gold standards," *Microbiology Today*, vol. 33, pp. 152–155, 2006.
- [2] P. Kämpfer, "Systematics of prokaryotes: the state of the art," *Antonie Van Leeuwenhoek*, vol. 101, 2012.
- [3] I. C. Sutcliffe, M. E. Trujillo, and M. Goodfellow, "A call to arms for systematists: revitalising the purpose and practises underpinning the description of novel microbial taxa," *Antonie Van Leeuwenhoek*, vol. 101, no. 1, pp. 13–20, 2012.
- [4] H. G. Schlegel and W. Köhler, "Bacteriology paved the way to cell biology: a historical account," *Biology of the Prokaryotes*, Wiley, Stuttgart, Germany, 1999.
- [5] G. Drews, "The roots of microbiology and the influence of Ferdinand Cohn on microbiology of the 19th century," *FEMS Microbiology Reviews*, vol. 24, no. 3, pp. 225–249, 2000.
- [6] P. Kämpfer and S. P. Glaeser, "Prokaryotic taxonomy in the sequencing era and the role of MLSA in classification," *Microbiology Australia*, vol. 32, no. 2, p. 66, 2011.
- [7] P. Vandamme, B. Pot, M. Gillis, P. De Vos, K. Kersters, and J. Swings, "Polyphasic taxonomy, a consensus approach to bacterial systematics," *Microbiological Reviews*, vol. 60, no. 2, pp. 407–438, 1996.
- [8] B. J. Tindall, R. Rosselló-Móra, H. J. Busse, W. Ludwig, and P. Kämpfer, "Notes on the characterization of prokaryote strains for taxonomic purposes," *International Journal of Systematic and Evolutionary Microbiology*, vol. 60, no. 1, pp. 249–266, 2010.
- [9] E. R. B. Moore, S. A. Mihaylova, P. Vandamme, M. I. Krichevsky, and L. Dijkshoorn, "Microbial systematics and taxonomy: relevance for a microbial commons," *Research in Microbiology*, vol. 161, no. 6, pp. 430–438, 2010.
- [10] P. Kämpfer and S. P. Glaeser, "Prokaryotic taxonomy in the sequencing era—the polyphasic approach revisited," *Environmental Microbiology*, vol. 14, no. 2, pp. 291–317, 2012.
- [11] D. J. Brenner, J. T. Staley, and N. R. Krieg, "Classification of prokaryotic organisms and the concept of bacterial speciation," *Bergey's Manual of Systematic Bacteriology*, Wiley, Boston, MA, USA, 2005.
- [12] K. J. Locey and J. T. Lennon, "Scaling laws predict global microbial diversity," *Proceedings of the National Academy of Sciences*, vol. 114, no. 12, pp. 3177–3182, 2017.

- Sciences of the United States of America*, vol. 113, no. 21, pp. 5970–5975, 2016.
- [13] A. C. Parte, J. Sardà Carbasse, J. P. Meier-Kolthoff, L. C. Reimer, and M. Göker, “List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ,” *International Journal of Systematic and Evolutionary Microbiology*, vol. 70, no. 11, pp. 5607–5612, 2020.
- [14] M. Loreau and C. de Mazancourt, “Biodiversity and ecosystem stability: a synthesis of underlying mechanisms,” *Ecology Letters*, vol. 16, pp. 106–115, 2013.
- [15] G. Bos, “Tomorrow’s production systems will be closer to nature,” 2017, <https://www.iucn.org/news/business-and-biodiversity/201701/tomorrows-production-systems-will-be-closer-nature>.
- [16] World Economic Forum, “The global risks report 2020, insight report,” 2020, https://www3.weforum.org/docs/WEF_Global_Risk_Report_2020.pdf.
- [17] R. Rosselló-Móra and R. Amann, “The species concept for prokaryotes,” *FEMS Microbiology Reviews*, vol. 25, no. 1, pp. 39–67, 2001.
- [18] R. Rosselló-Móra and R. Amann, “Past and future species definitions for Bacteria and Archaea,” *Systematic & Applied Microbiology*, vol. 38, no. 4, pp. 209–216, 2015.
- [19] L. C. Ferraz Helene, R. F. Dall’Agnol, J. R. M. Delamuta, and M. Hungria, “*Mesorhizobium atlanticum* sp. nov., a new nitrogen-fixing species from soils of the Brazilian Atlantic Forest biome,” *International Journal of Systematic and Evolutionary Microbiology*, vol. 69, no. 6, pp. 1800–1806, 2019.
- [20] E. S. P. Bromfield, S. Cloutier, and H. D. T. Nguyen, “Description and complete genome sequence of *Bradyrhizobium amphicarphae* sp. nov., harbouring photosystem and nitrogen-fixation genes,” *International Journal of Systematic and Evolutionary Microbiology*, vol. 69, no. 9, pp. 2841–2848, 2019.
- [21] E. L. P. Santos, M. Palmer, B. Chávez-Ramírez 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*, vol. 9, 2018.
- [22] J. D. Flores-Félix, E. Menéndez, A. Peix, P. García-Fraile, and E. Velázquez, “History and current taxonomic status of genus *Agrobacterium*,” *Systematic & Applied Microbiology*, vol. 43, no. 1, Article ID 126046, 2020.
- [23] J. R. Avontuur, M. Palmer, C. W. Beukes et al., “Genome-informed *Bradyrhizobium* taxonomy: where to from here?” *Systematic & Applied Microbiology*, vol. 42, no. 4, pp. 427–439, 2019.
- [24] P. H. Graham, M. J. Sadowsky, H. H. Keyser et al., “Proposed minimal standards for the description of new genera and species of root- and stem-nodulating bacteria,” *International Journal of Systematic Bacteriology*, vol. 41, no. 4, pp. 582–587, 1991.
- [25] P. M. de Lajudie, M. Andrews, J. Ardley et al., “Minimal standards for the description of new genera and species of rhizobia and agrobacteria,” *International Journal of Systematic and Evolutionary Microbiology*, vol. 69, no. 7, pp. 1852–1863, 2019.
- [26] C. T. Parker, B. J. Tindall, and G. M. Garrity, “Prokaryotic code (2008 revision) International code of nomenclature of prokaryotes,” *International Journal of Systematic and Evolutionary Microbiology*, vol. 69, 2019.
- [27] M. E. Trujillo, “Publication of new taxa,” 2021, <https://www.microbiologyresearch.org/journal/ijsem/scope>.
- [28] E. T. Wang, C. F. Tian, W. F. Chen, J. P. W. Young, and W. X. Chen, *Ecology and Evolution of Rhizobia: Principles and Applications*, Springer, Berlin, Germany, 2019.
- [29] R. G. Murray and E. Stackenbrandt, “Taxonomic Note: Implementation of the provisional status Candidatus for incompletely described prokaryotes,” *International Journal of Systematic Bacteriology*, vol. 45, no. 1, pp. 186–187, 1995.
- [30] C. Quast, E. Pruesse, P. Yilmaz et al., “The SILVA ribosomal RNA gene database project: improved data processing and web-based tools,” *Nucleic Acids Research*, vol. 41, no. D1, pp. D590–D596, 2012.
- [31] W. B. Whitman, I. C. Sutcliffe, and R. Rossello-Mora, “Proposal for changes in the International code of nomenclature of prokaryotes: granting priority to *Candidatus* names,” *International Journal of Systematic and Evolutionary Microbiology*, vol. 69, no. 7, pp. 2174–2175, 2019.
- [32] D. J. Brenner, J. T. Staley, and N. R. Krieg, *Classification of Prokaryotic Organisms: an Overview*, Wiley, New York, NY, USA, 1989.
- [33] P. H. A. Sneath, “The application of computers to taxonomy,” *Microbiology*, vol. 17, no. 1, pp. 201–226, 1957.
- [34] W. Ludwig, K. Schleifer, and E. Stackenbrandt, “Databases,” *The Prokaryotes*, Springer, Berlin, Germany, 2006.
- [35] K. H. Schleifer, “Classification of Bacteria and Archaea: past, present and future,” *Systematic & Applied Microbiology*, vol. 32, no. 8, pp. 533–542, 2009.
- [36] L. G. Wayne, W. E. C. Moore, E. Stackenbrandt et al., “Report of the Ad Hoc Committee on reconciliation of approaches to bacterial systematics,” *International Journal of Systematic and Evolutionary Microbiology*, vol. 37, no. 4, pp. 463–464, 1987.
- [37] D. J. Brenner, G. R. Fanning, A. V. Rake, and K. E. Johnson, “Batch procedure for thermal elution of DNA from hydroxyapatite,” *Analytical Biochemistry*, vol. 28, pp. 447–459, 1969.
- [38] C. L. Schildkraut, J. Marmur, and P. Doty, “The formation of hybrid DNA molecules and their use in studies of DNA homologies,” *Journal of Molecular Biology*, vol. 3, no. 5, pp. 595–IN16, 1961.
- [39] R. K. Saiki, D. H. Gelfand, S. Stoffel et al., “Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase,” *Science*, vol. 239, no. 4839, pp. 487–491, 1988.
- [40] F. Sanger, S. Nicklen, and A. R. Coulson, “DNA sequencing with chain-terminating inhibitors,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 74, no. 12, pp. 5463–5467, 1977.
- [41] C. R. Woese and G. E. Fox, “Phylogenetic structure of the prokaryotic domain: the primary kingdoms,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 74, no. 11, pp. 5088–5090, 1977.
- [42] C. R. Woese, O. Kandler, and M. L. Wheelis, “Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 87, no. 12, pp. 4576–4579, 1990.
- [43] J. T. Staley, “The bacterial species dilemma and the genomic-phylogenetic species concept,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 361, 2006.
- [44] A. Oren and G. M. Garrity, “Then and now: a systematic review of the systematics of prokaryotes in the last 80 years,” *Antonie Van Leeuwenhoek*, vol. 106, no. 1, pp. 43–56, 2014.

- [45] E. Stackebrandt, W. Ludwig, and G. E. Fox, "3 16 S ribosomal RNA oligonucleotide cataloguing," *Methods in Microbiology*, vol. 18, pp. 75–107, 1985.
- [46] D. Gevers, F. M. Cohan, J. G. Lawrence et al., "Re-evaluating prokaryotic species," *Nature Reviews Microbiology*, vol. 3, no. 9, pp. 733–739, 2005.
- [47] J. Feng, E. L. Schuenzel, J. Li, and N. W. Schaad, "Multilocus sequence typing reveals two evolutionary lineages of acidovorax avenae subsp. citrulli," *Phytopathology*, vol. 99, 2009.
- [48] S. Yan, H. Liu, T. J. Mohr et al., "Role of recombination in the evolution of the model plant pathogen *Pseudomonas syringae* pv. tomato DC3000, a very atypical tomato strain," *Applied and Environmental Microbiology*, vol. 74, no. 10, pp. 3171–3181, 2008.
- [49] M. C. J. Maiden, J. A. Bygraves, E. Feil et al., "Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 6, pp. 3140–3145, 1998.
- [50] C. B. Sullivan, M. A. Diggle, and S. C. Clarke, "Multilocus sequence typing: data analysis in clinical microbiology and public health," *Molecular Biotechnology*, vol. 29, pp. 245–254, 2005.
- [51] J. Chun and F. A. Rainey, "Integrating genomics into the taxonomy and systematics of the Bacteria and Archaea," *International Journal of Systematic and Evolutionary Microbiology*, vol. 64, no. Pt_2, pp. 316–324, 2014.
- [52] S. P. Glaeser and P. Kämpfer, "Multilocus sequence analysis (MLSA) in prokaryotic taxonomy," *Systematic & Applied Microbiology*, vol. 38, no. 4, pp. 237–245, 2015.
- [53] P. De Vos, "Multilocus sequence determination and analysis," *Methods in Microbiology*, vol. 38, pp. 385–407, 2012.
- [54] X. Rong and Y. Huang, "Taxonomic evaluation of the *Streptomyces hygroscopicus* clade using multilocus sequence analysis and DNA-DNA hybridization, validating the MLSA scheme for systematics of the whole genus," *Systematic & Applied Microbiology*, vol. 35, no. 1, pp. 7–18, 2012.
- [55] D. R. Zeigler, "Gene sequences useful for predicting relatedness of whole genomes in bacteria," *International Journal of Systematic and Evolutionary Microbiology*, vol. 53, no. 6, pp. 1893–1900, 2003.
- [56] K. T. Konstantinidis, A. Ramette, and J. M. Tiedje, "Toward a more robust assessment of intraspecies diversity, using fewer genetic markers," *Applied and Environmental Microbiology*, vol. 72, no. 11, pp. 7286–7293, 2006.
- [57] Z. Yang and B. Rannala, "Molecular phylogenetics: principles and practice," *Nature Reviews Genetics*, vol. 13, no. 5, pp. 303–314, 2012.
- [58] J. P. Meier-Kolthoff, H. P. Klenk, and M. Göker, "Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age," *International Journal of Systematic and Evolutionary Microbiology*, vol. 64, no. Pt_2, pp. 352–356, 2014.
- [59] K. T. Konstantinidis and J. M. Tiedje, "Genomic insights that advance the species definition for prokaryotes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 7, pp. 2567–2572, 2005.
- [60] J. Chun, A. Oren, A. Ventosa et al., "Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes," *International Journal of Systematic and Evolutionary Microbiology*, vol. 68, no. 1, pp. 461–466, 2018.
- [61] P. Vandamme and C. Peeters, "Time to revisit polyphasic taxonomy," *Antonie Van Leeuwenhoek*, vol. 106, no. 1, pp. 57–65, 2014.
- [62] C. C. Thompson, G. R. Amaral, M. Campeão et al., "Microbial taxonomy in the post-genomic era: rebuilding from scratch?" *Archives of Microbiology*, vol. 197, no. 3, pp. 359–370, 2015.
- [63] L. Loret and E. Martínez-Romero, "Evolución y filogenia de *Rhizobium*," *Revista Latinoamericana de Microbiología*, vol. 47, no. 1-2, pp. 43–60, 2005.
- [64] R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, vol. 32, no. 5, pp. 1792–1797, 2004.
- [65] J. D. Thompson, D. G. Higgins, and T. J. Gibson, "Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice," *Nucleic Acids Research*, vol. 22, no. 22, pp. 4673–4680, 1994.
- [66] W. Ludwig and K. H. Schleifer, "Bacterial phylogeny based on 16S and 23S rRNA sequence analysis," *FEMS Microbiology Reviews*, vol. 15, no. 2-3, pp. 155–173, 1994.
- [67] W. Ludwig and H. P. Klenk, "Overview: a phylogenetic backbone and taxonomic framework for prokaryotic systematics," *Bergey's Manual of Systematic Bacteriology*, Springer, Berlin, Germany, 2001.
- [68] J. Felsenstein, "Confidence limits on phylogenies: an approach using the bootstrap," *Evolution*, vol. 39, 1985.
- [69] S. B. Hedges, "The number of replications needed for accurate estimation of the bootstrap P value in phylogenetic studies," *Molecular Biology and Evolution*, vol. 9, pp. 366–369, 1992.
- [70] J. D. Retief, "Phylogenetic analysis using PHYLIP," *Bioinformatics Methods and Protocols*, Human Press, Totowa, New Jersey, USA, 2000.
- [71] T. Hall, "BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT," *Nucleic Acids Symposium Series*, vol. 41, pp. 95–98, 1999.
- [72] M. Kim, H. S. Oh, S. C. Park, and J. Chun, "Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes," *International Journal of Systematic and Evolutionary Microbiology*, vol. 64, no. Pt_2, pp. 346–351, 2014.
- [73] W. Ludwig, R. Amann, E. Martínez-Romero et al., "rRNA based identification and detection systems for rhizobia and other bacteria," *Plant and Soil*, vol. 204, pp. 1–19, 1998.
- [74] E. Stackebrandt and B. M. Goebel, "Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology," *International Journal of Systematic and Evolutionary Microbiology*, vol. 44, no. 4, pp. 846–849, 1994.
- [75] P. Menna, F. G. Barcellos, and M. Hungria, "Phylogeny and taxonomy of a diverse collection of *Bradyrhizobium* strains based on multilocus sequence analysis of the 16S rRNA gene, ITS region and *glnII*, *recA*, *atpD* and *dnaK* genes," *International Journal of Systematic and Evolutionary Microbiology*, vol. 59, no. 12, pp. 2934–2950, 2009.
- [76] D. Durán, L. Rey, J. Mayo et al., "*Bradyrhizobium paxllaeri* sp. nov. and *Bradyrhizobium icense* sp. nov., nitrogen-fixing rhizobial symbionts of Lima bean (*Phaseolus lunatus* L.) in Peru," *International Journal of Systematic and Evolutionary Microbiology*, vol. 64, no. Pt_6, pp. 2072–2078, 2014.
- [77] L. C. Ferraz Helene, G. O'Hara, and M. Hungria, "Characterization of *Bradyrhizobium* strains indigenous to Western Australia and South Africa indicates remarkable genetic diversity and reveals putative new species,"

- Systematic & Applied Microbiology*, vol. 43, no. 2, Article ID 126053, 2020.
- [78] M. Rossi-Tamisier, S. Benamar, D. Raoult, and P. E. Fournier, "Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species," *International Journal of Systematic and Evolutionary Microbiology*, vol. 65, no. Pt_6, pp. 1929–1934, 2015.
- [79] L. M. Schouls, C. S. Schot, and J. A. Jacobs, "Horizontal transfer of segments of the 16S rRNA genes between species of the *Streptococcus anginosus* group," *Journal of Bacteriology*, vol. 185, no. 24, pp. 7241–7246, 2003.
- [80] A. Y. Pei, W. E. Oberdorf, C. W. Nossa et al., "Diversity of 16S rRNA genes within individual prokaryotic genomes," *Applied and Environmental Microbiology*, vol. 76, no. 12, pp. 3886–3897, 2010.
- [81] X. Y. Zhi, W. Zhao, W. J. Li, and G. P. Zhao, "Prokaryotic systematics in the genomics era," *Antonie Van Leeuwenhoek*, vol. 101, no. 1, pp. 21–34, 2011.
- [82] J. García-Martínez, S. G. Acinas, A. I. Antón, and F. Rodríguez-Valera, "Use of the 16S–23S ribosomal genes spacer region in studies of prokaryotic diversity," *Journal of Microbiological Methods*, vol. 36, no. 1–2, pp. 55–64, 1999.
- [83] M. S. Klepa, M. C. d O. Urquiaga, P. Somasegaran, J. R. M. Delamuta, R. A. Ribeiro, and M. Hungria, "*Bradyrhizobium niftali* sp. nov., an effective nitrogen-fixing symbiont of partridge pea [*Chamaecrista fasciculata* (Michx.) Greene], a native caesalpinoid legume broadly distributed in the USA," *International Journal of Systematic and Evolutionary Microbiology*, vol. 69, no. 11, pp. 3448–3459, 2019.
- [84] M. C. d O. Urquiaga, M. S. Klepa, P. Somasegaran, R. A. Ribeiro, J. R. M. Delamuta, and M. Hungria, "*Bradyrhizobium frederickii* sp. nov., a nitrogen-fixing lineage isolated from nodules of the caesalpinoid species *Chamaecrista fasciculata* and characterized by tolerance to high temperature in vitro," *International Journal of Systematic and Evolutionary Microbiology*, vol. 69, no. 12, pp. 3863–3877, 2019.
- [85] P. van Berkum and J. J. Fuhrmann, "Evolutionary relationships among the soybean bradyrhizobia reconstructed from 16S rRNA gene and internally transcribed spacer region sequence divergence," *International Journal of Systematic and Evolutionary Microbiology*, vol. 50, no. 6, pp. 2165–2172, 2000.
- [86] A. Willems, R. Coopman, and M. Gillis, "Comparison of sequence analysis of 16S–23S rDNA spacer regions, AFLP analysis and DNA–DNA hybridizations in *Bradyrhizobium*," *International Journal of Systematic and Evolutionary Microbiology*, vol. 51, no. 2, pp. 623–632, 2001.
- [87] L. Carro, Á. Peix, and E. Velázquez, "The taxonomy of bacteria in the genomic era," *Developmental Biology in Prokaryotes and Lower Eukaryotes*, Springer, Berlin, Germany, 2021.
- [88] A. Willems, A. Munive, P. de Lajudie, and M. Gillis, "In most *Bradyrhizobium* groups sequence comparison of 16S–23S rDNA internal transcribed spacer regions corroborates DNA–DNA hybridizations," *Systematic & Applied Microbiology*, vol. 26, no. 2, pp. 203–210, 2003.
- [89] A. Pei, Cw Nossa, P. Chokshi et al., "Diversity of 23S rRNA genes within individual prokaryotic genomes," *PLoS One*, vol. 4, no. 5, Article ID e5437, 2009.
- [90] A. J. Probst, C. Hertel, L. Richter, L. Wassil, W. Ludwig, and W. P. Hammes, "*Staphylococcus condimenti* sp. nov., from soy sauce mash, and *Staphylococcus carnosus* (Schleifer and Fischer 1982) subsp. *utilis* subsp. nov.," *International Journal of Systematic Bacteriology*, vol. 48, no. 3, pp. 651–658, 1998.
- [91] J. Yan, H. Yan, L. X. Liu et al., "*Rhizobium hidalgonense* sp. nov., a nodule endophytic bacterium of *Phaseolus vulgaris* in acid soil," *Archives of Microbiology*, vol. 199, no. 1, pp. 97–104, 2017.
- [92] M. S. Klepa, L. C. Ferraz Helene, G. O'Hara, and M. Hungria, "*Bradyrhizobium agreste* sp. nov., *Bradyrhizobium glycinis* sp. nov. and *Bradyrhizobium diversitatis* sp. nov., isolated from a biodiversity hotspot of the genus *Glycine* in Western Australia," *International Journal of Systematic and Evolutionary Microbiology*, vol. 71, no. 3, 2019.
- [93] L. C. F. Helene, M. S. Klepa, G. O'Hara, and M. Hungria, "*Bradyrhizobium archetypum* sp. nov., *Bradyrhizobium australiense* sp. nov. and *Bradyrhizobium murdochi* sp. nov., isolated from nodules of legumes indigenous to Western Australia," *International Journal of Systematic and Evolutionary Microbiology*, vol. 70, no. 8, pp. 4623–4636, 2020.
- [94] J. M. Lang, A. E. Darling, and J. A. Eisen, "Phylogeny of Bacterial and Archaeal genomes using conserved genes: supertrees and supermatrices," *PLoS One*, vol. 8, no. 4, Article ID e62510, 2013.
- [95] H. Berrada and K. Fikri-Benbrahim, "Taxonomy of the rhizobia: current perspectives," *British Microbiology Research Journal*, vol. 4, no. 6, pp. 616–639, 2014.
- [96] J. Z. M. Chan, M. R. Halachev, N. J. Loman, C. Constantinidou, and M. J. Pallen, "Defining bacterial species in the genomic era: insights from the genus *Acinetobacter*," *BMC Microbiology*, vol. 12, no. 1, p. 302, 2012.
- [97] F. L. Thompson, D. Gevers, C. C. Thompson et al., "Phylogeny and molecular identification of vibrios on the basis of Multilocus Sequence Analysis," *Applied and Environmental Microbiology*, vol. 71, pp. 5107–5115, 2005.
- [98] M. Martens, P. Dawyndt, R. Coopman, M. Gillis, P. De Vos, and A. Willems, "Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 house-keeping genes in the genus *Ensifer* (including former *Sinorhizobium*)," *International Journal of Systematic and Evolutionary Microbiology*, vol. 58, no. 1, pp. 200–214, 2008.
- [99] L. Mavima, C. W. Beukes, M. Palmer et al., "*Paraburkholderia youngii* sp. nov. and "*Paraburkholderia atlantica*" Brazilian and Mexican *Mimosa*-associated rhizobia that were previously known as *Paraburkholderia tuberum* sv. *mimosae*," *Systematic & Applied Microbiology*, vol. 44, no. 1, Article ID 126152, 2021.
- [100] K. Tamura, G. Stecher, D. Peterson, A. Filipski, and S. Kumar, "MEGA6: molecular evolutionary genetics analysis version 6.0," *Molecular Biology and Evolution*, vol. 30, no. 12, pp. 2725–2729, 2013.
- [101] M. Gouy, S. Guindon, and O. Gascuel, "SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building," *Molecular Biology and Evolution*, vol. 27, no. 2, pp. 221–224, 2010.
- [102] E. Stackebrandt, W. Frederiksen, G. M. Garrity et al., "Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology," *International Journal of Systematic and Evolutionary Microbiology*, vol. 52, pp. 1043–1047, 2002.
- [103] K. T. Konstantinidis and J. M. Tiedje, "Prokaryotic taxonomy and phylogeny in the genomic era: advancements and challenges ahead," *Current Opinion in Microbiology*, vol. 10, no. 5, pp. 504–509, 2007.

- [104] M. Mulet, J. Lalucat, and E. Garcia-Valdes, "DNA sequence-based analysis of the *Pseudomonas* species," *Environmental Microbiology*, vol. 12, 2010.
- [105] J. R. M. Delamuta, R. A. Ribeiro, E. Ormeño-Orrillo, I. S. Melo, E. Martínez-Romero, and M. Hungria, "Polyphasic evidence supporting the reclassification of *Bradyrhizobium japonicum* group Ia strains as *Bradyrhizobium diazoefficiens* sp. nov.," *International Journal of Systematic and Evolutionary Microbiology*, vol. 63, no. Pt_9, pp. 3342–3351, 2013.
- [106] F. Paulitsch, R. F. Dall'Agnol, J. R. M. Delamuta, R. A. Ribeiro, J. S. da Silva Batista, and M. Hungria, "*Paraburkholderia atlantica* sp. nov. and *Paraburkholderia franconis* sp. nov., two new nitrogen-fixing nodulating species isolated from Atlantic forest soils in Brazil," *Archives of Microbiology*, vol. 202, no. 6, pp. 1369–1380, 2020.
- [107] J. Versalovic, M. Schneider, F. J. Bruijn, and J. R. Lupski, "Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction," *Methods in Molecular and Cellular Biology*, vol. 5, pp. 25–40, 1994.
- [108] K. S. B. Mannan, R. M. Mazumdar, and A. Chowdhury, "Identification and characterization of microorganisms: DNA fingerprinting methods," *Songklanakarin Journal of Science and Technology*, vol. 35, no. 4, pp. 397–404, 2013.
- [109] M. J. Stern, G. F. L. Ames, N. H. Smith, E. Clare Robinson, and C. F. Higgins, "Repetitive extragenic palindromic sequences: a major component of the bacterial genome," *Cell*, vol. 37, no. 3, pp. 1015–1026, 1984.
- [110] C. S. J. Hulton, C. F. Higgins, and P. M. Sharp, "ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria," *Molecular Microbiology*, vol. 5, no. 4, pp. 825–834, 1991.
- [111] B. Martin, O. Humbert, M. Camara et al., "A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*," *Nucleic Acids Research*, vol. 20, no. 13, pp. 3479–3483, 1992.
- [112] F. J. de Bruijn, "Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria," *Applied and Environmental Microbiology*, vol. 58, no. 7, pp. 2180–2187, 1992.
- [113] G. Laguerre, P. Mavingui, M. R. Allard 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," *Applied and Environmental Microbiology*, vol. 62, no. 6, pp. 2029–2036, 1996.
- [114] P. Menna, A. A. Pereira, E. V. Bangel, and M. Hungria, "rep-PCR of tropical rhizobia for strain fingerprinting, biodiversity appraisal and as a taxonomic and phylogenetic tool," *Symbiosis*, vol. 48, pp. 120–130, 2009.
- [115] R. Rosselló-Móra, M. Urdiain, and A. López-López, "DNA-DNA hybridization," *Taxonomy of Prokaryotes*, Academic Press, Cambridge, MA, USA, 2012.
- [116] J. L. Johnson and E. J. Ordal, "Deoxyribonucleic acid homology in bacterial taxonomy: effect of incubation temperature on reaction specificity," *Journal of Bacteriology*, vol. 95, no. 3, pp. 893–900, 1968.
- [117] J. Goris, K. T. Konstantinidis, J. A. Klappenbach, T. Coenye, P. Vandamme, and J. M. Tiedje, "DNA-DNA hybridization values and their relationship to whole-genome sequence similarities," *International Journal of Systematic and Evolutionary Microbiology*, vol. 57, no. 1, pp. 81–91, 2007.
- [118] N. K. Mahato, V. Gupta, P. Singh et al., "Microbial taxonomy in the era of OMICS: application of DNA sequences, computational tools and techniques," *Antonie van Leeuwenhoek*, vol. 110, 2017.
- [119] R. D. Fleischmann, M. D. Adams, O. White et al., "Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd," *Science*, vol. 269, no. 5223, pp. 496–512, 1995.
- [120] E. Sentausa and P. E. Fournier, "Advantages and limitations of genomics in prokaryotic taxonomy," *Clinical Microbiology and Infection*, vol. 19, no. 9, pp. 790–795, 2013.
- [121] A. Gurevich, V. Saveliev, N. Vyahhi, and G. Tesler, "QUAST: quality assessment tool for genome assemblies," *Bioinformatics*, vol. 29, no. 8, pp. 1072–1075, 2013.
- [122] J. P. Meier-Kolthoff, A. F. Auch, H. P. Klenk, and M. Göker, "Genome sequence based species delimitation with confidence intervals and improved distance functions," *BMC Bioinformatics*, vol. 14, no. 1, p. 60, 2013.
- [123] M. Richter and R. Rosselló-Mora, "Shifting the genomic gold standard for the prokaryotic species definition," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 45, pp. 19126–19131, 2009.
- [124] N. M. Mesbah, W. B. Whitman, and M. Mesbah, "Determination of the G+C content of prokaryotes," *Taxonomy of Prokaryotes*, Elsevier, Amsterdam, Netherlands, 1st edition, 2012.
- [125] R. Overbeek, R. Olson, G. D. Pusch et al., "The seed and the rapid annotation of microbial genomes using subsystems technology (RAST)," *Nucleic Acids Research*, vol. 42, no. D1, pp. D206–D214, 2014.
- [126] M. Hungria, L. M. D O. Chueire, R. G. Coca, and M. Megias, "Preliminary characterization of fast growing rhizobial strains isolated from soyabean nodules in Brazil," *Soil Biology and Biochemistry*, vol. 33, no. 10, pp. 1349–1361, 2001.
- [127] P. Kämpfer, "Continuing importance of the "phenotype" in the genomic era," *Methods in Microbiology*, vol. 41, pp. 307–320, 2014.
- [128] M. Sasser, "Identification of bacteria by gas chromatography of cellular fatty acids," 1990, https://www.researchgate.net/publication/284025789_Identification_of_bacteria_by_gas_chromatography_of_cellular_fatty_acids.
- [129] E. Carbonnelle, C. Mesquita, E. Bille et al., "MALDI-TOF mass spectrometry tools for bacterial identification in clinical microbiology laboratory," *Clinical Biochemistry*, vol. 44, no. 1, pp. 104–109, 2011.
- [130] R. D. Holland, J. G. Wilkes, F. Rafi et al., "Rapid identification of intact whole bacteria based on spectral patterns using matrix-assisted laser desorption/ionization with time-of-flight mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 10, pp. 1227–1232, 1996.
- [131] D. Ziegler, J. F. Pothier, J. Ardley et al., "Ribosomal protein biomarkers provide root nodule bacterial identification by MALDI-TOF MS," *Applied Microbiology and Biotechnology*, vol. 99, no. 13, pp. 5547–5562, 2015.
- [132] L. Rolim, T. R. Santiago, F. B. dos Reis Junior, I. de Carvalho Mendes, M. Hungria, and L. P. Silva, "Identification of soybean *Bradyrhizobium* strains used in commercial inoculants in Brazil by MALDI-TOF mass spectrometry," *Brazilian Journal of Microbiology*, vol. 50, no. 4, pp. 1147–1148, 2019.
- [133] P. Rahi, O. Prakash, and Y. S. Shouche, "Matrix-assisted laser desorption/ionization time-of-flight mass-spectrometry (MALDI-TOF MS) based microbial identifications:

- challenges and scopes for microbial ecologists," *Frontiers in Microbiology*, vol. 7, 2016.
- [134] E. Ormeño-Orrillo and E. Martínez-Romero, "Phenotypic tests in *Rhizobium* species description: an opinion and (a sympatric speciation) hypothesis," *Systematic & Applied Microbiology*, vol. 36, no. 3, pp. 145–147, 2013.
- [135] A. Willems, "The taxonomy of rhizobia: an overview," *Plant and Soil*, vol. 287, no. 1–2, pp. 3–14, 2006.
- [136] B. Frank, "Ueber die Pilzsymbiose der Leguminosen," *Berichte der Deutschen Botanischen Gesellschaft*, vol. 7, pp. 332–346, 1889.
- [137] E. Velázquez, P. García-Fraile, M. H. Ramírez-Bahena, R. Rivas, and E. Martínez-Molina, "Current status of the taxonomy of bacteria able to establish nitrogen-fixing legume symbiosis," *Microbes for Legume Improvement*, Springer, Berlin, Germany, 2nd edition, 2017.
- [138] I. L. Baldwin and E. B. Fred, "Nomenclature of the root nodule bacteria of the Leguminosae," *Journal of Bacteriology*, vol. 17, no. 2, pp. 141–150, 1929.
- [139] P. A. Dangeard, "Recherches sur les tubercules radicaux des Légumineuses," *Botaniste*, vol. 16, pp. 1–275, 1926.
- [140] M. M. Eckhardt, I. L. Baldwin, and E. B. Fred, "Studies on the root-nodule organism of *Lupinus*," *Journal of Bacteriology*, vol. 21, no. 4, pp. 273–285, 1931.
- [141] R. E. Buchanan, "What names should be used for the organisms producing nodules on the roots of leguminous plants?" *Proceedings of the Iowa Academy of Science*, vol. 33, 1926.
- [142] E. B. Fred, I. L. Baldwin, and E. Mc Coy, *Root Nodule Bacteria and Leguminous Plants*, Madison, Madison, Wisconsin, 1932.
- [143] J. K. Wilson, "Over five hundred reasons for abandoning the cross-inoculation groups of legumes," *Soil Science*, vol. 58, no. 1, pp. 61–70, 1944.
- [144] F. Zakhia and P. De Lajudie, "Taxonomy of rhizobia," *Agronomie*, vol. 21, 2001.
- [145] P. H. Graham, "The application of computer techniques to the taxonomy of the root-nodule bacteria of legumes," *Journal of General Microbiology*, vol. 35, no. 3, pp. 511–517, 1964.
- [146] P. H. Graham, "Antigenic affinities of the root-nodule bacteria of legumes," *Antonie van Leeuwenhoek*, vol. 29, no. 1, pp. 281–291, 1963.
- [147] J. De Ley and A. Rassel, "DNA base composition, flagellation and taxonomy of the genus *Rhizobium*," *Journal of General Microbiology*, vol. 41, no. 1, pp. 85–92, 1965.
- [148] M. L. Moffett and R. R. Colwell, "Adansonian analysis of the Rhizobiaceae," *Journal of General Microbiology*, vol. 51, no. 2, pp. 245–266, 1968.
- [149] D. C. Jordan, "Transfer of *Rhizobium japonicum* Buchanan 1980 to *Bradyrhizobium* gen. nov., a genus of slow-growing, root nodule bacteria from leguminous plants," *International Journal of Systematic Bacteriology*, vol. 32, no. 1, pp. 136–139, 1982.
- [150] B. Dreyfus, J. L. Garcia, and M. Gillis, "Characterization of *Azorhizobium cauhodans* gen. nov., sp. nov., a stem-nodulating nitrogen-fixing bacterium isolated from *Sesbania rostrata*," *International Journal of Systematic Bacteriology*, vol. 38, no. 1, pp. 89–98, 1988.
- [151] W. X. Chen, G. H. Yan, and J. L. Li, "Numerical taxonomic study of fast-growing soybean rhizobia and proposal that *Rhizobium fredii* be assigned to *Sinorhizobium* gen. nov.," *International Journal of Systematic Bacteriology*, vol. 38, no. 4, pp. 392–397, 1988.
- [152] H. J. Conn, "Validity of the genus *Alcaligenes*," *Journal of Bacteriology*, vol. 44, no. 3, pp. 353–360, 1942.
- [153] A. Peix, M. H. Ramírez-Bahena, E. Velázquez, and E. J. Bedmar, "Bacterial associations with legumes," *Critical Reviews in Plant Sciences*, vol. 34, no. 1–3, pp. 17–42, 2015.
- [154] B. D. W. Jarvis, P. van Berkum, W. X. Chen et al., "Transfer of *Rhizobium loti*, *Rhizobium huakuii*, *Rhizobium ciceri*, *Rhizobium mediterraneum*, and *Rhizobium tianshanense* to *Mesorhizobium* gen. nov.," *International Journal of Systematic and Evolutionary Microbiology*, vol. 47, no. 3, pp. 895–898, 1997.
- [155] P. de Lajudie, E. Laurent-Fulele, A. Willems et al., "*Allo-rhizobium undicola* gen. nov., sp. nov., nitrogen-fixing bacteria that efficiently nodulate *Neptunia natans* in Senegal," *International Journal of Systematic Bacteriology*, vol. 48, no. 4, pp. 1277–1290, 1998.
- [156] L. E. J. Casida, "*Ensifer adhaerens* gen nov. sp. nov.: a bacterial predator of bacteria in soil," *International Journal of Systematic Bacteriology*, vol. 32, no. 3, pp. 339–345, 1982.
- [157] P. van Berkum and B. D. Eardly, "Molecular evolutionary systematics of the Rhizobiaceae," *The Rhizobiaceae*, Elsevier, Dordrecht, Netherlands, 1998.
- [158] M. Yanagi and K. Yamasato, "Phylogenetic analysis of the family Rhizobiaceae and related bacteria by sequencing of 16S rRNA gene using PCR and DNA sequencer," *FEMS Microbiology Letters*, vol. 107, no. 1, pp. 115–120, 1993.
- [159] T. E. Patt, G. C. Cole, and R. S. Hanson, "*Methylobacterium*, a new genus of facultatively methylotrophic bacteria," *International Journal of Systematic Bacteriology*, vol. 26, no. 2, pp. 226–229, 1976.
- [160] A. Sy, E. Giraud, P. Jourand et al., "Methylotrophic *Methylobacterium* bacteria nodulate and fix nitrogen in symbiosis with legumes," *Journal of Bacteriology*, vol. 183, no. 1, pp. 214–220, 2001.
- [161] H. Sawada, H. Ieki, H. Oyaizu, and S. Matsumoto, "Proposal for rejection of *Agrobacterium tumefaciens* and revised descriptions for the genus *Agrobacterium* and for *Agrobacterium radiobacter* and *Agrobacterium rhizogenes*," *International Journal of Systematic Bacteriology*, vol. 43, no. 4, pp. 694–702, 1993.
- [162] J. M. Young, L. D. Kuykendall, E. Martínez-Romero, A. Kerr, and H. Sawada, "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*," *International Journal of Systematic and Evolutionary Microbiology*, vol. 51, no. 1, pp. 89–103, 2001.
- [163] R. Rivas, E. Velázquez, A. Willems et al., "A new species of *Devosia* that forms a unique nitrogen-fixing root-nodule symbiosis with the aquatic legume *Neptunia natans* (L.f.) Druce," *Applied and Environmental Microbiology*, vol. 68, no. 11, pp. 5217–5222, 2002.
- [164] A. Valverde, E. Velázquez, F. Fernández-Santos et al., "*Phyllobacterium trifolii* sp. nov., nodulating *Trifolium* and *Lupinus* in Spanish soils," *International Journal of Systematic and Evolutionary Microbiology*, vol. 55, no. 5, pp. 1985–1989, 2005.
- [165] B. Holmes, M. Popoff, M. Kiredjian, and K. Kersters, "*Ochrobactrum anthropi* gen. nov., sp. nov. from human clinical specimens and previously known as group Vd," *International Journal of Systematic Bacteriology*, vol. 38, no. 4, pp. 406–416, 1988.

- [166] K. F. Meyer and E. B. Shaw, "A comparison of the morphologic, cultural and biochemical characteristics of *B. abortus* and *B. melitensis* from cattle. Studies on the genus *Brucella* nov.," *Journal of Infectious Diseases*, vol. 27, no. 3, pp. 173–184, 1920.
- [167] M. E. Trujillo, A. Willems, A. Abril et al., "Nodulation of *Lupinus albus* by strains of *Ochrobactrum lupini* sp. nov.," *Applied and Environmental Microbiology*, vol. 71, pp. 1318–1327, 2005.
- [168] J. L. Zurdo-Piñeiro, R. Rivas, M. E. Trujillo et al., "*Ochrobactrum cytisi* sp. nov., isolated from nodules of *Cytisus scoparius* in Spain," *International Journal of Systematic and Evolutionary Microbiology*, vol. 57, no. 4, pp. 784–788, 2007.
- [169] D. X. Lin, E. T. Wang, H. Tang et al., "*Shinella kummerowiae* sp. nov., a symbiotic bacterium isolated from root nodules of the herbal legume *Kummerowia stipulacea*," *International Journal of Systematic and Evolutionary Microbiology*, vol. 58, no. 6, pp. 1409–1413, 2008.
- [170] D. S. An, W. T. Im, H. C. Yang, and S. T. Lee, "*Shinella granulii* gen. nov., sp. nov., and proposal of the reclassification of *Zoogloea ramigera* ATCC 19623 as *Shinella zoogloeoides* sp. nov.," *International Journal of Systematic and Evolutionary Microbiology*, vol. 56, no. 2, pp. 443–448, 2006.
- [171] M. J. Estrella, S. Muñoz, M. J. Soto, O. Ruiz, and J. Sanjuán, "Genetic diversity and host range of rhizobia nodulating *Lotus tenuis* in typical soils of the Salado River Basin (Argentina)," *Applied and Environmental Microbiology*, vol. 75, no. 4, pp. 1088–1098, 2009.
- [172] S. Kalso and B. K. C. Patel, "*Microvirga subterranea* gen. nov., sp. nov., a moderate thermophile from a deep subsurface Australian thermal aquifer," *International Journal of Systematic and Evolutionary Microbiology*, vol. 53, no. 2, pp. 401–406, 2003.
- [173] J. K. Ardley, M. A. Parker, S. E. De Meyer et al., "*Microvirga lupini* sp. nov., *Microvirga lotononidis* sp. nov. and *Microvirga zambiensis* sp. nov. are alphaproteobacterial root-nodule bacteria that specifically nodulate and fix nitrogen with geographically and taxonomically separate legume hosts," *International Journal of Systematic and Evolutionary Microbiology*, vol. 62, no. Pt_11, pp. 2579–2588, 2012.
- [174] V. Radl, J. L. Simões-Araújo, J. Leite et al., "*Microvirga vignae* sp. nov., a root nodule symbiotic bacterium isolated from cowpea grown in semi-arid Brazil," *International Journal of Systematic and Evolutionary Microbiology*, vol. 64, no. Pt_3, pp. 725–730, 2014.
- [175] L. Moulin, A. Munive, B. Dreyfus, and C. Boivin-Masson, "Nodulation of legumes by members of the beta-subclass of Proteobacteria," *Nature*, vol. 411, no. 6840, pp. 948–950, 2001.
- [176] E. Yabuuchi, Y. Kosako, H. Oyaizu et al., "Proposal of burkholderia gen. Nov. and transfer of seven species of the genus pseudomonas homology group II to the new genus, with the type species *Burkholderia cepacia* (palleroni and holmes 1981) comb. Nov.," *Microbiology and Immunology*, vol. 36, 1992.
- [177] E. Yabuuchi, Y. Kosako, I. Yano, H. Hotta, and Y. Nishiuchi, "Transfer of two burkholderia and an alcaligenes species to ralstonia gen. nov.: proposal of *Ralstonia pickettii* (ralston, palleroni and doudoroff 1973) comb. nov., *Ralstonia solanacearum* (Smith 1896) comb. nov. and *Ralstonia eutropha* (Davis 1969) comb. nov.," *Microbiology and Immunology*, vol. 39, 1995.
- [178] N. S. Makkar and L. E. Casida Jr, "*Cupriavidus necator* gen. nov., sp. nov.: a nonobligate bacterial predator of bacteria in soil," *International Journal of Systematic Bacteriology*, vol. 37, no. 4, pp. 323–326, 1987.
- [179] W. M. Chen, E. K. James, A. R. Prescott, M. Kierans, and J. I. Sprent, "Nodulation of *Mimosa* spp. by the beta-proteobacterium *Ralstonia taiwanensis*," *Molecular Plant-Microbe Interactions*®, vol. 16, no. 12, pp. 1051–1061, 2003.
- [180] P. Vandamme and T. Coenye, "Taxonomy of the genus *Cupriavidus*: a tale of lost and found," *International Journal of Systematic and Evolutionary Microbiology*, vol. 54, no. 6, pp. 2285–2289, 2004.
- [181] A. Sawana, M. Adeolu, and R. S. Gupta, "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," *Frontiers in Genetics*, vol. 5, 2014.
- [182] A. P. Dobritsa and M. Samadpour, "Transfer of eleven *Burkholderia* species to the genus *Paraburkholderia* and proposal of *Caballeronia* gen. nov., a new genus to accommodate twelve species of *Burkholderia* and *Paraburkholderia*," *International Journal of Systematic and Evolutionary Microbiology*, vol. 66, no. 8, pp. 2836–2846, 2016.
- [183] S. A. Mousavi, J. Österman, N. Wahlberg et al., "Phylogeny of the Rhizobium–Allorhizobium–Agrobacterium clade supports the delineation of Neorhizobium gen. nov.," *Systematic and Applied Microbiology*, vol. 37, 2014.
- [184] S. A. Mousavi, A. Willems, X. Nesme, P. de Lajudie, and K. Lindström, "Revised phylogeny of Rhizobiaceae: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations," *Systematic & Applied Microbiology*, vol. 38, no. 2, pp. 84–90, 2015.
- [185] C. Fagorzi, A. Ilie, F. Decorosi et al., "Symbiotic and non-symbiotic members of the genus *Ensifer* (syn. *Sinorhizobium*) are separated into two clades based on comparative genomics and high-throughput phenotyping," *Genome Biology and Evolution*, vol. 12, no. 12, pp. 2521–2534, 2020.
- [186] N. Kuzmanović, C. Fagorzi, A. Mengoni, F. Lassalle, and G. C. DiCenzo, "Taxonomy of Rhizobiaceae revisited: proposal of a new framework for genus delimitation," *International Journal of Systematic and Evolutionary Microbiology*, vol. 72, no. 3, 2022.
- [187] Y. Benhiziia, H. Benhizia, A. Benguedouar, R. Muresu, A. Giacomini, and A. Squartini, "Gamma Proteobacteria can nodulate legumes of the genus *Hedysarum*," *Systematic & Applied Microbiology*, vol. 27, no. 4, pp. 462–468, 2004.
- [188] A. Shiraishi, N. Matsushita, and T. Hougetsu, "Nodulation in black locust by the gammaproteobacteria *Pseudomonas* sp. and the Betaproteobacteria *Burkholderia* sp.," *Systematic & Applied Microbiology*, vol. 33, no. 5, pp. 269–274, 2010.
- [189] M. A. Rogel, E. Ormeño-Orrillo, and E. Martínez Romero, "Symbiovars in rhizobia reflect bacterial adaptation to legumes," *Systematic & Applied Microbiology*, vol. 34, no. 2, pp. 96–104, 2011.
- [190] J. R. M. Delamuta, R. A. Ribeiro, E. Ormeño-Orrillo et al., "*Bradyrhizobium tropiciagri* sp. nov. and *Bradyrhizobium embrapense* sp. nov nitrogen-fixing symbionts of tropical forage legumes," *International Journal of Systematic and Evolutionary Microbiology*, vol. 65, no. Pt_12, pp. 4424–4433, 2015.
- [191] M. Hungria, G. W. O'Hara, J. E. Zilli, R. S. Araujo, R. Deaker, and J. G. I. Howieson, "Isolation and growth of rhizobia,"

- Working with Rhizobia*, ACIAR, Canberra, ACT, Australia, 2016.
- [192] G. C. diCenzo and T. M. Finan, "The divided bacterial genome: structure, function, and evolution," *Microbiology and Molecular Biology Reviews*, vol. 81, no. 3, 2017.
- [193] B. A. Geddes, J. Kearsley, R. Morton, G. C. diCenzo, and T. M. Finan, "The genomes of rhizobia," *Regulation of Nitrogen-Fixing Symbioses in Legumes*, Elsevier, Amsterdam, Netherlands, 2020.
- [194] G. C. diCenzo, A. Checcucci, M. Bazzicalupo et al., "Metabolic modelling reveals the specialization of secondary replicons for niche adaptation in *Sinorhizobium meliloti*," *Nature Communications*, vol. 7, no. 1, Article ID 12219, 2016.
- [195] G. Dávila and R. Palacios, "Origins of genomics in nitrogen-fixation research," *Genomes and Genomics of Nitrogen-Fixing Organisms*, Springer, Berlin, Germany, 2005.
- [196] F. Galibert, T. M. Finan, S. R. Long et al., "The composite genome of the legume symbiont *Sinorhizobium meliloti*," *Science*, vol. 293, no. 5530, pp. 668–672, 2001.
- [197] J. T. Sullivan and C. W. Ronson, "Evolution of rhizobia by acquisition of a 500-kb symbiosis island that integrates into a phe-tRNA gene," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 9, pp. 5145–5149, 1998.
- [198] S. Wang, B. Hao, J. Li et al., "Whole-genome sequencing of *Mesorhizobium huakuii* 7653R provides molecular insights into host specificity and symbiosis island dynamics," *BMC Genomics*, vol. 15, no. 1, p. 440, 2014.
- [199] M. Black, P. Moolhuijzen, B. Chapman et al., "The genetics of symbiotic nitrogen fixation: comparative genomics of 14 rhizobia strains by resolution of protein clusters," *Genes*, vol. 3, no. 1, pp. 138–166, 2012.
- [200] S. E. de Meyer, L. Briscoe, P. Martínez-Hidalgo et al., "Symbiotic *Burkholderia* species show diverse arrangements of *nif*/*fix* and nod genes and lack typical high-affinity cytochrome *cbb3* oxidase genes," *Molecular Plant-Microbe Interactions*, vol. 29, no. 8, pp. 609–619, 2016.
- [201] P. Remigi, J. Zhu, J. P. W. Young, and C. Masson-Boivin, "Symbiosis within symbiosis: Evolving nitrogen-fixing legume symbionts," *Trends in Microbiology*, vol. 24, no. 1, pp. 63–75, 2016.
- [202] E. Giraud, L. Moulin, D. Vallenet et al., "Legumes symbioses: absence of *nod* genes in photosynthetic bradyrhizobia," *Science*, vol. 316, no. 5829, pp. 1307–1312, 2007.
- [203] X. Perret, C. Staehelin, and W. J. Broughton, "Molecular basis of symbiotic promiscuity," *Microbiology and Molecular Biology Reviews*, vol. 64, pp. 180–201, 2000.
- [204] G. E. D. Oldroyd, J. D. Murray, P. S. Poole, and J. A. Downie, "The rules of engagement in the legume-rhizobial symbiosis," *Annual Review of Genetics*, vol. 45, no. 1, pp. 119–144, 2011.
- [205] D. Patra and S. Mandal, "Nod-factors are dispensable for nodulation: a twist in bradyrhizobia-legume symbiosis," *Symbiosis*, vol. 86, no. 1, pp. 1–15, 2022.
- [206] L. H. Madsen, L. Tirichine, A. Jurkiewicz et al., "The molecular network governing nodule organogenesis and infection in the model legume *Lotus japonicus*," *Nature Communications*, vol. 1, no. 1, p. 10, 2010.
- [207] P. W. Harrison, R. P. Lower, N. K. Kim, and J. P. W. Young, "Introducing the bacterial 'chromid': not a chromosome, not a plasmid," *Trends in Microbiology*, vol. 18, no. 4, pp. 141–148, 2010.
- [208] J. S. Silva Batista, M. Hungria, F. G. Barcellos, M. C. Ferreira, and I. C. Mendes, "Variability in *Bradyrhizobium japonicum* and *B. elkanii* seven years after introduction of both the exotic microsymbiont and the soybean host in a Cerrados soil," *Microbial Ecology*, vol. 53, no. 2, pp. 270–284, 2007.
- [209] L. C. F. Helene, J. R. M. Delamuta, R. A. Ribeiro, and M. Hungria, "*Bradyrhizobium mercantei* sp. nov., a nitrogen-fixing symbiont isolated from nodules of *Deguelia costata* (syn. *Lonchocarpus costatus*)," *International Journal of Systematic and Evolutionary Microbiology*, vol. 67, no. 6, pp. 1827–1834, 2017.
- [210] L. C. Ferraz Helene, J. R. Marçon Delamuta, R. Augusto Ribeiro et al., "*Bradyrhizobium viridifuturi* sp. nov., encompassing nitrogen-fixing symbionts of legumes used for green manure and environmental services," *International Journal of Systematic and Evolutionary Microbiology*, vol. 65, no. Pt_12, pp. 4441–4448, 2015.
- [211] J. P. W. Young, "The phylogeny and evolution of nitrogenases," *Genomes and Genomics of Nitrogen-Fixing Organisms*, Springer, Berlin, Germany, 2005.
- [212] M. Hungria, P. Menna, and J. R. M. Delamuta, "Bradyrhizobium, the ancestor of all rhizobia: phylogeny of house-keeping and nitrogen-fixation genes," *Biological Nitrogen Fixation*, John Wiley & Son, Hoboken, NJ, USA, 2015.
- [213] S. L. Turner and J. P. W. Young, "The glutamine synthetases of rhizobia: phylogenetics and evolutionary implications," *Molecular Biology and Evolution*, vol. 17, no. 2, pp. 309–319, 2000.
- [214] P. Menna and M. Hungria, "Phylogeny of nodulation and nitrogen-fixation genes in *Bradyrhizobium*: supporting evidence for the theory of monophyletic origin, and spread and maintenance by both horizontal and vertical transfer," *International Journal of Systematic and Evolutionary Microbiology*, vol. 61, no. 12, pp. 3052–3067, 2011.
- [215] C. S. Smillie, M. B. Smith, J. Friedman, O. X. Cordero, L. A. David, and E. J. Alm, "Ecology drives a global network of gene exchange connecting the human microbiome," *Nature*, vol. 480, no. 7376, pp. 241–244, 2011.
- [216] J. Ling, H. Wang, P. Wu et al., "Plant nodulation inducers enhance horizontal gene transfer of *Azorhizobium caulinodans* symbiosis island," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 48, pp. 13875–13880, 2016.
- [217] C. Masson-Boivin, E. Giraud, X. Perret, and J. Batut, "Establishing nitrogen-fixing symbiosis with legumes: how many rhizobium recipes?" *Trends in Microbiology*, vol. 17, no. 10, pp. 458–466, 2009.
- [218] D. Bulgarelli, M. Rott, K. Schlaeppli et al., "Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota," *Nature*, vol. 488, no. 7409, pp. 91–95, 2012.
- [219] J. Edwards, C. Johnson, C. Santos-Medellín et al., "Structure, variation, and assembly of the root-associated microbiomes of rice," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 112, no. 8, 2015.
- [220] F. Ibanez, J. Angelini, T. Taurian, M. L. Tonelli, and A. Fabra, "Endophytic occupation of peanut root nodules by opportunistic gammaproteobacteria," *Systematic & Applied Microbiology*, vol. 32, no. 1, pp. 49–55, 2009.
- [221] A. M. Chibeba, C. S. Pereira, J. E. L. Antunes et al., "Polyphasic characterization of nitrogen-fixing and co-resident bacteria in nodules of *Phaseolus lunatus* inoculated with soils from Piauí State, northeast Brazil," *Symbiosis*, vol. 80, no. 3, pp. 279–292, 2020.
- [222] S. M. B. Rocha, L. W. Mendes, L. M. D. S. Oliveira et al., "Nodule microbiome from cowpea and lima bean grown in composted tannery sludge-treated soil," *Applied Soil Ecology*, vol. 151, Article ID 103542, 2020.