

Review

Not peer-reviewed version

Challenges in the Identification of Environmental Bacterial Isolates from the Pharmaceutical Industry Facility by 16S rRNA Gene Sequences

[Juliana Nunes Ramos](#)*, [Luciana Veloso da Costa](#), [Verônica Viana Vieira](#), [Marcelo Luiz Lima Brandao](#)

Posted Date: 31 March 2025

doi: 10.20944/preprints202503.2316.v1

Keywords: 16S rRNA gene sequencing; housekeeping genes; MALDI-TOF MS; genomic taxonomy analysis; environmental bacterial isolates



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Review

Challenges in the Identification of Environmental Bacterial Isolates from the Pharmaceutical Industry Facility by 16S rRNA Gene Sequences

Juliana Nunes Ramos ^{1,*}, Luciana Veloso da Costa ¹, Verônica Viana Vieira ²
and Marcelo Luiz Lima Brandão ¹

¹ Institute of Technology in Immunobiologicals, Oswaldo Cruz Foundation, Rio de Janeiro 21040-900, Brazil

² Interdisciplinary Laboratory of Medical Research, Oswaldo Cruz Foundation, Rio de Janeiro 21040-900, Brazil

* Correspondence: jnr.uerj@gmail.com (J.N.R.)

Abstract: Background: Microbial contamination is a significant challenge for the pharmaceutical sector, especially for heat-sensitive sterile products. This contamination can alter the physical and chemical properties of pharmaceutical products, compromising quality and safety. Correct bacterial identification is essential for tracing sources of contamination and implementing preventive and corrective measures. Matrix-Assisted Laser Desorption Ionization-Time of Flight/Mass Spectrometry (MALDI-TOF MS) technology has revolutionized microbial identification, but there are limitations in the databases, requiring additional analyses, such as sequencing of the 16S rRNA and housekeeping genes and/or whole genome sequencing. **Objectives:** This review explores the challenges of identifying bacterial contaminants in the pharmaceutical industry using 16S rRNA gene sequences. Additional sequencing of housekeeping genes can be useful for differentiating bacteria at the species level. Advances in DNA sequencing technology have expanded genomic taxonomy, allowing for more accurate bacterial classification. This study aims to demonstrate how the combination of these methods increases the accuracy of bacterial identification. **Conclusions and future directions:** MALDI-TOF MS is widely used in the pharmaceutical sector for bacterial identification. A public database of spectral profiles of environmental bacterial isolates would be essential for bacterial identification and information exchange between institutions. Complementary methods, such as 16S rRNA and housekeeping genes sequencing, can provide reliable bacterial identification and allow the expansion of the MALDI-TOF MS database. Advances in genomic taxonomy have enabled the development of genomic taxonomy tools to improve bacterial identification. The combination of multiple methods provides greater precision and overcomes the limitations of single-gene approaches.

Keywords: 16S rRNA gene sequencing; housekeeping genes; MALDI-TOF MS; genomic taxonomy analysis; environmental bacterial isolates

1. Background

Microbial contamination is one of the biggest obstacles facing the pharmaceutical industry, especially for heat-sensitive sterile products such as immunobiologicals that cannot be terminally sterilized. Microbial contamination can alter the physical and chemical properties of pharmaceutical products and excipients, affecting product quality and consumer safety. Therefore, the performance of microbiological testing is essential for the quality control of these products [1,2]. Good Manufacturing Practices (GMPs) must be followed to reduce the risk of microbial contamination in pharmaceutical production environments, in order to ensure that biological products meet, among other parameters, acceptable limits for microorganisms [3].

Several groups of bacteria may be present as contaminants in clean areas, but aerobic endospore-forming bacteria have been described as one of the most important groups of bacteria isolated from these environments, due to their ability to produce spores that are resistant to temperature variations and to the sanitizers used in industry, allowing them to persist in the environment for long periods [1,3]. Identification of the detected microbiome is essential to investigate the source of contamination and consequently to take preventive and corrective measures [2,4,5].

A bacterial species is characterized as a group of strains, including the type strain, that share more than 70% similarity in DNA-DNA hybridization (DDH), maximum of 2% G+C span, values above 98.7% in 16S rRNA gene sequence identity, and distinct chemotaxonomic and phenotypic characteristics [6]. Several identification methods are used in clinical or industrial laboratories to identify the bacterial species, which are described in more detail below.

Matrix-Assisted Laser Desorption Ionization–Time of Flight/Mass Spectrometry (MALDI-TOF MS) technology has been widely used to identify microorganisms contaminating pharmaceutical production environments because it offers several advantages over other identification methods, particularly in terms of speed and greater specificity and sensitivity [1,3,7]. Therefore, if the MALDI-TOF database fails to identify a bacterial isolate, it is necessary to sequence the 16S rRNA gene and some housekeeping genes according to the suspected species or genera, or even the whole genome to perform genomic taxonomy analyses [1,7,8]. The biggest challenge is when the strain is a potential new species and the genes and/or genomes of all the closest species in the genus are not deposited in the databases for the purpose of comparing similarities and reconstructing the phylogenetic tree.

This review discusses the difficulties encountered in identifying environmental bacterial isolates from pharmaceutical facilities using MALDI-TOF MS technology, 16S rRNA and housekeeping gene sequence analyses, which often require genomic taxonomy analyses using genome sequencing data.

2. Phenotypic Identification

Phenotypic identification using commercial biochemical systems can be performed before the bacterial isolate is submitted for identification by MALDI-TOF MS or gene sequencing. Biochemical identification of environmental isolates using commercial biochemical systems such as API® system (bioMérieux, Craponne, France) and VITEK® 2 Compact System (bioMérieux, Craponne, France) is still used in the pharmaceutical industry [9–11].

The API® system consists of a series of biochemical tests based on the fermentation of sugars (carbohydrates), assimilation of other carbon sources and the production of unique enzymes and metabolites. These tests are used to identify Gram-positive, Gram-negative bacteria and yeasts. The type of kit is selected based on colony morphology and staining results. The API profiles obtained are identified using the APIWEB™ database [12,13].

The VITEK® 2 system is used for microbial identification and antimicrobial susceptibility testing. It consists of a card filling/sealing system, an incubator/reader, coupled to a computer. Microbial identification is performed using cards containing dehydrated biochemical substrates that require no additional reagents. As with the API® system, card selection is based on colony morphology and staining results. VITEK® identification includes microbial species of clinical and industrial importance [14].

Both systems have potential limitations, such as (i) the difficulty in determining phenotypic variation between strains; (ii) some may show different results in repeated tests; (iii) limited databases; (iv) small changes in test performance can give false results. In addition, non-fermenting bacteria can be problematic due to their phenotypic variations and slower growth rates [15].

For microorganisms of environmental origin and pharmaceutical products, some studies show the need for molecular methods to complete the identification of different bacterial groups. It is important to note that despite the ineffectiveness of these methods in phenotypic identification, the results of biochemical assays can be useful in differentiating some species [10,11].

3. MALDI-TOF MS

MALDI-TOF MS is considered a high-throughput technology, based on the acquisition of unique molecular signatures that are representative of a wide range of proteins and can clearly distinguish the differences between two closely related species. Its database contains spectra of microorganisms that are compared to the mass spectrum of the microorganism to be identified to find a closer match. Regardless of which system is used, be it the MALDI Biotyper (Bruker Daltonics, France) or the Vitek MS (bioMérieux, France), each of them requires different preparations and has different databases and algorithms [16,17].

This method has undoubtedly revolutionized the microbial identification system, reducing the time and cost of identification. Accuracy and speed in reporting results are crucial for batch release of pharmaceutical products. It is important to note that the MALDI-TOF MS databases were initially created using mainly clinically relevant strains for species identification. The application of technology for the identification of bacterial isolates from other sources, such as soil, industry, freshwater, etc., required new databases with relevant corresponding species. Due to its rapid identification method and low cost per isolate, the technology attracted the interest of several specialist groups who expanded the MALDI-TOF MS databases. Species not detected by MALDI-TOF MS, but identified by 16S rRNA gene sequences, were added to the database to study microbial communities from poorly understood locations, making this technology more widely used beyond clinical laboratories [18]. Costa et al. [1] added 24 *Bacillus* spp. strains and related genera from the pharmaceutical industry to the MALDI-TOF MS database, after a careful genotypic characterization of the strains using 16S rRNA and *rpoB* gene sequences. Miranda et al. [10] reported for the first time the isolation of six *Sutcliffiella horikoshii* strains from an immunobiological pharmaceutical facility that were not identified by MALDI-TOF MS. After gene sequencing, the MALDI-TOF MS database was expanded, and the strains were correctly identified as *S. horikoshii* by MALDI-TOF MS. In 2017, an aerobic, spore-forming Gram-positive rod isolated from an air monitoring sample from an immunobiological production unit was not identified by MALDI-TOF MS. After physiological and genotypic characterization and biochemical tests, the *Bacillus lumeideiriae* species was described and the MALDI-TOF database was expanded with the spectra profiles of the proposed new species [8].

It is possible to increase the identification capacity of MALDI-TOF by including the spectral profiles in the database [19]. For this purpose, the bacterial isolate must first be identified by analysis of the 16S ribosomal gene sequence at the genus level, since this gene has limitations in differentiating species. Sequences of genes encoding highly conserved proteins, called housekeeping genes, can be used in conjunction with 16S rRNA gene sequences to try to reach the species level [1,19].

It is important to mention that there is no public database of environmental bacterial isolates for comparing spectra obtained with MALDI-TOF MS, which makes bacterial identification and the exchange of information between researchers and institutions difficult. For highly pathogenic bacteria (biosafety level 3), such as *Bacillus anthracis*, *Yersinia pestis*, *Burkholderia mallei*, *Burkholderia pseudomallei* and *Francisella tularensis*, as well as their related species, the Robert Koch German Institute has developed a database with MALDI-TOF MS mass spectra, which serves as a reference for diagnosing these bacteria using microbial identification software. The spectra are available in a zip file containing the original mass spectra in the data format used by Bruker Daltonics [20].

4. Analysis of the 16S Ribosomal Gene Sequences

Sequencing of the 16S ribosomal gene is widely used in bacterial identification [21]. The 16S rRNA gene is approximately 1500 base pairs (bp) in size and is located in the small subunit (30S) of the prokaryotic ribosome [22] (**Figure 1**). Some species may have shorter or longer sequences. The 16S rRNA gene does not encode proteins, but in addition to having a structural role, it is crucial for protein synthesis. Although rare, horizontal transfer of the 16S rRNA gene can also occur, but only at the intragenus or intraspecific level [23,24].

The 16S rRNA gene is considered an important molecular marker, present in all members of Bacteria and Archaea, highly conserved in bacteria and evolving slowly, it is the target widely used for phylogenetic studies of bacteria. In this sense, its sequences are well used for taxonomic

classification at the genus level. The presence of multiple copies in the bacterial genome with sequence differences and low polymorphism are some of the limitations that must be considered for classification at the species level [25].

The 16S rRNA gene contains highly conserved, variable and hypervariable regions that are unevenly distributed, being nine hypervariable regions, designated V1 to V9, that vary in length, position and taxonomic discrimination. Such variation is conducive to infer phylogenetic relationships between phyla, while also being used in comparison of interests [23,24,26].

Although considered the gold standard for bacterial identification, amplification of the 16S rRNA gene is still costly and impractical in the routine of some laboratories [27]. For more accurate molecular identification, complete sequencing of the gene (~1500 bp) is required [25]. The 16S rRNA gene sequences should be compared with available databases such as Basic Local Alignment Search Tool (Blast, <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) or EzBioCloud (<https://www.ezbiocloud.net/>). If a new bacterial species is suspected, it is important to limit the comparison of similarities to the sequences of type strains [28].

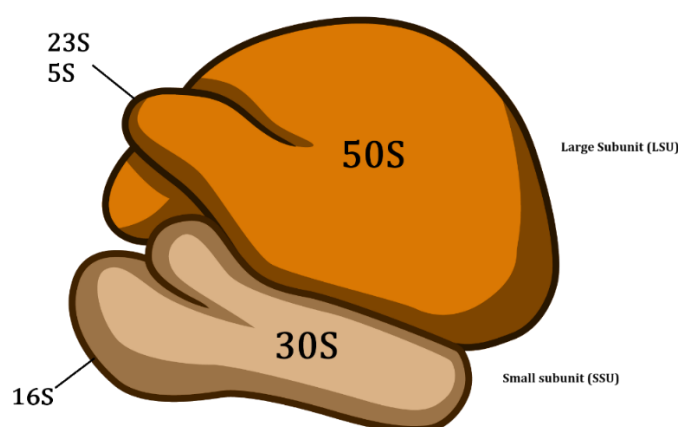


Figure 1. Structure of the prokaryotic ribosome. Image adapted from NIAID Visual & Medical Arts, 2024..

Previously, to be considered from the same species, bacterial isolates should share a similarity in the 16S rRNA gene of more than 97%, based on a relationship with 70% DDH, which is considered the gold standard method for delimiting bacterial species [30]. According to Stackebrandt & Ebers (2006), the similarity criterion for the 16S rRNA gene would be greater than 98.7%. Identity values in the 16S rRNA gene sequence below 95% with the phylogenetically closest species with a validated name may indicate that the isolate is a representative of a new genus [31].

Although they provide valuable phylogenetic information, 16S ribosomal gene sequences are not always useful for distinguishing closely related species due to their highly conserved nature. When several species of the same genus share >98.7% identity in the 16S gene sequence, the sequencing of other genes, called essential or constitutive, is recommended [8,22].

5. Analysis of the Housekeeping Gene Sequences

Housekeeping genes, also called essential genes, encode enzymes that are essential for maintaining cellular function. Examples include the *recA* gene, which encodes a recombinase protein; *rpoB*, which encodes the beta subunit of RNA polymerase; and *gyrB*, which encodes the B subunit of the DNA gyrase protein. Housekeeping genes are highly conserved and accumulate mutations more rapidly than rRNA genes, making them useful for differentiating bacteria at the species level, since the 16S rRNA gene shows high genetic homology within certain genera and may not be useful alone for distinguishing closely related species [21,24,31–33].

The analysis of the sequences of a certain number of housekeeping genes, called multilocus sequence analysis (MLSA), incorporates similarity values to differentiate species, and is considered a

phylogenetic tool to support and clarify the delimitation of bacterial species with a higher resolution than studies based on 16S rRNA genes. The gene sequences are used to construct a phylogenetic tree to infer phylogenies [34,35]. MLSA is based on multilocus sequence typing (MLST), a method for typing pathogenic bacteria for epidemiological and population genetic purposes, first described by Maiden et al, 1998 [35,36].

The choice of genes for each taxon analyzed is critical to the reliability of the analysis. Housekeeping genes should be considered because they are more stable in relation to rapid genetic change and are present in all species of a genus. It is also recommended that they be single-copy genes, distributed throughout the entire genome. The *ad hoc* Committee recommends the use of at least five housekeeping genes for the re-evaluation of the species definition in bacteriology, although most studies use seven genes. Some species require the use of more genes for better differentiation [35,37,38]. Before amplifying and sequencing of the selected housekeeping genes of interest for comparison and construction of a phylogenetic tree, it is very important to consult the different sequence databases, such as National Center for Biotechnology Information (NCBI, <https://www.ncbi.nlm.nih.gov/>), European Molecular Biology Laboratory's – European Bioinformatics Institute (EMBL-EBI, <https://www.ebi.ac.uk/>), DNA Data Bank of Japan (DDBJ, <https://www.ddbj.nig.ac.jp/index-e.html>), Bacterial Diversity Metadatabase (BacDive, <https://bacdive.dsmz.de/>), Joint Genome Institute (JGI, <https://genome.jgi.doe.gov/portal/>), among others. If the sequences of the housekeeping genes of interest or the complete genome of the closest type strain are not available, an alternative would be to purchase the type strain from a collection, such as the American Type Culture Collection (ATCC) or the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ), for example, and perform the sequencing yourself. Some sequences may be in the possession of researchers who have not yet published their data. Contacting these researchers may also be an alternative.

Gene sequencing for MLSA involves time-consuming and laborious steps. With the advancement of high-throughput sequencing and bioinformatics tools in recent years, MLSA can be performed *in silico* due to the substantial increase of whole-genome data in public databases, allowing gene sequences to be extracted directly from genomes [38].

6. Genomic Taxonomy Tools

Improvements in DNA sequencing technologies have resulted in a significant increase in the amount of genomic data generated, combined with a reduction in the cost of sequencing [34]. Whole genome sequencing (WGS) data and the development of bioinformatics tools have allowed the establishment of taxonomic schemes based on evolutionary information contained in genome sequences, such as digital DNA-DNA hybridization (dDDH), Average Amino Acid Identity (AAI), supertrees, and others [39]. Such taxonomic schemes, which are described in more detailed below, can be used not only to describe a new bacterial species, but also to confirm the identification of the bacterial isolate, especially when it comes to environmental isolates belonging to the genus *Bacillus* and/or related genera, which have a large number of described species, and genotypic characterization by sequencing the 16S rRNA and housekeeping genes is not always possible.

Genomic taxonomy is defined based on an integrated comparative genomics approach with the goal of extracting taxonomic information from genomes. The main goal of genomic taxonomy is to extract taxonomic information from genomes that can be used to provide a solid framework for identifying and classifying prokaryotic species and even populations. These new tools mentioned above have led to new understanding of genetic relationships that the 16S rRNA gene can only approximate [6,39,40].

DDH indirectly measures the degree of genetic similarity between two genomes, one of which is the genome of the type strain, and has been the “gold standard” for bacterial species delimitation [41]. In brief, the heated DNA strands are dissociated and immediately reassociated, and hybridization occurs. The degree of relationship between the two genomes is verified, and the two genomes are considered to belong to the same species if the DDH value is greater than 70% [34,38].

However, this value is not sufficient to distinguish, for example, the species *Rickettsia rickettsii*, *Rickettsia conorii*, *Rickettsia sibirica* and *Rickettsia montanensis*, in other words, the DDH limit used is not applicable to all genera [31]. Few laboratories in the world perform this methodology because it is a laborious, slow, and expensive technique that requires specialized personnel. Another disadvantage is that the results may vary depending on the protocol used, which can lead to experimental errors and the comparison of results obtained by different methodologies is not recommended [30,31,38].

With the advent of high-throughput DNA sequencing and the various genomes deposited in public repositories, dDDH was then proposed [34,42], and some authors suggested replacing DDH as the “gold standard” in prokaryotic taxonomy by pairwise genomic sequence-derived similarity [42,43]. The analysis of dDDH consists of the local alignment between two genomes and intergenomic correspondences are generated, which are later used to calculate the distance matrix, whose values are analogous to DDH (> 70%) [38]. The Genome-to-Genome Distance Calculator (GGDC) is one of the most popular online tools for calculating in silico DDH values, provided free of charge by the German bacterial collection DSMZ. The GGDC allows the comparison of bacterial genomes to determine their similarity on the same scale as the DDH, aiding in the identification and classification of bacterial species. Moreover, the GGDC reports the difference in G+C content, which can be used for species delineation [43,44].

Over the years, there has been an increasing use of the Average Nucleotide Identity (ANI) tool for the classification and identification of bacterial species [45,46], which shows a strong correlation with the DDH [47]. ANI was proposed in 2005 and consisted in the local alignment between two genomes, calculating the average identity of the nucleotides of the shared open reading frames (ORFs) instead of the whole genome [43]. In 2007, the comparison between two whole genomes was implemented, allowing ANI results to be directly compared with the DDH [48]. Two prokaryotic genomes can be considered to belong to the same species if they share an ANI value $\geq 96\%$ [31]. Furthermore, ANI shows a strong correlation with 16S rRNA gene sequence similarity [49]. Although a robust tool for bacterial species delimitation, ANI has been shown to have low resolution at higher taxonomic levels. Therefore, other metrics should be used, such as AAI [50,51], another widely used tool for bacterial species classification and identification, which is based on the calculation of conserved protein-coding genes between a pair of genomes determined by pairwise comparison of whole genomes using the BLAST algorithm. Briefly, all protein-coding genes in one genome are searched for in relation to all protein-coding genes in the other genome. The AAI of all conserved genes between the pair of genomes is then measured from the genetic relatedness between a pair of genomes and a value of ~95% to 96% was established, with a strong correlation to the similarity of the 16S rRNA gene sequence [38,39,52]. The AAI has been shown to be a very useful metric for genera delimitation (>60-65%) [49]. Based on the discontinuous distribution of its AAI values, several studies have developed boundaries to delimit genera, such as *Chryseobacterium* [53], *Prochlorococcus* [54] and *Lactobacillus* [55].

Complex phylogenetic relationships between different taxa can be captured by genome-based phylogenies, or phylogenomics. In 2019, the online platform TYGS (Type Strain Genome Server) was developed by the German collection DSMZ for the classification and identification of microorganisms based on their genomes, combining genomic data with robust phylogenetic approaches to provide accurate and reproducible classification. The Type Strain Genome Server (TYGS) contains a comprehensive database that is constantly updated and revised to ensure accuracy and reliability, and is maintained by the Leibniz Institute/DSMZ. This server allows the analysis of dDDH indices, and the construction of phylogenies based on the 16S rRNA gene and the Genome BLAST Distance Phylogeny (GBDP) method, using as reference the next closest phylogenetic matches previously identified by Mash genomic distances and 16S rRNA gene data. In addition, TYGS is integrated with the database that powers the List of Prokaryotic Names with Standing in Nomenclature (LPSN), providing information on the most recent updates in nomenclature and taxonomic literature. The

platform also performs a preliminary species-level classification to assist in the identification of possible new species [56,57].

Another computational tool that supports the phylogenomic classification of bacteria based on genomic data is the Genome Taxonomy Database Toolkit (GTDB-Tk), which compares genomes to the Genome Taxonomy Database (GTDB) taxonomy using a pre-calculated phylogenomic tree and metrics such as ANI and Relative Evolutionary Divergence (RED). GTDB-Tk allows the assignment of taxa to new genomes, facilitating the identification of new species and the reclassification of organisms within a standardized phylogenomic taxonomy. GTDB-Tk is an independent tool linked to GTDB that inserts the genome into a previously calculated multigene-based phylogenies (MBP) phylogenomic tree, calculates RED indices and performs species assignment based on ANI, when viable. This process allows the identification of new taxa, both at the species level and in broader taxonomic categories. The GTDB and backbone trees are updated regularly, including a major annual revision [56,58].

Another approach worth mentioning is Ribosomal Multilocus Sequence Typing (rMLST, <https://pubmlst.org/species-id>), which was proposed to overcome the limitations of current methods for bacterial typing and phylogenetic reconstruction by using 53 ribosomal genes (*rps* genes) that are present in all bacteria and distributed along bacterial chromosomes. They also encode ribosomal proteins that are highly conserved. This technique provides greater phylogenetic resolution than traditional methods, such as using the 16S rRNA gene, and allows bacteria to be accurately classified at multiple taxonomic levels, including domain, phylum, class, order, family, genus, and species. The rMLST database is an extensible, web-accessible database containing complete genomic data from thousands of bacterial isolates, enabling rapid and computationally efficient identification of the phylogenetic position of any bacterial sequence at multiple taxonomic levels [59].

7. Conclusions and Future Directions

Many laboratories in the pharmaceutical industry use MALDI-TOF MS for bacterial identification because of its rapid results and low cost, which are essential for batch release of pharmaceutical products. Here, we suggest some methods for the identification of environmental bacteria, from MALDI-TOF MS to genomic taxonomy analysis, and the advantages and disadvantages of each method (**Figure 2**). Initially, MALDI-TOF MS databases focused on clinically relevant strains, but their application has expanded to environmental and industrial samples, requiring new references. Species not detectable by MALDI-TOF MS have been included after identification by 16S rRNA and housekeeping genes sequencing. Identification can be improved by adding spectral profiles to the database.

The most commonly used method to assess the phylogenetic position of a prokaryote is the comparison of the 16S rRNA gene sequence. Sequencing of the 16S rRNA gene is widely used for bacterial identification due to its universal presence in Bacteria and Archaea, high conservation and low rate of evolution, being an essential marker for phylogenetic studies. Nucleotide variations within multiple rRNA operons in a single genome and the possibility of 16S rRNA genes being derived from horizontal gene transfer can distort relationships between taxa in phylogenetic trees. Despite its relevance, the 16S rRNA gene does not always distinguish closely related species due to its high conservation; its phylogeny is robust at the genus level and above, making it necessary to complement the analysis with the sequencing of housekeeping genes, which have been used in MLSA as an alternative for more accurate identification of microorganisms.

Advances in DNA sequencing technologies have increased the availability of genomic data and reduced their costs, allowing the creation of taxonomic schemes based on in silico analysis of genomes. Genomic taxonomy uses tools such as dDDH, ANI and AAI to identify and classify bacterial species. The traditional DDH method, previously considered the gold standard for species delimitation, has been replaced by computational methods, such as the GGDC, which calculates the genomic similarity between organisms. ANI, which compares the average nucleotide identity between two genomes is widely used to define species with values $\geq 96\%$, while AAI, based on the

identity of conserved coding proteins, helps to delimit genera (>60-65%). In addition, phylogenomic tools such as TYGS and GTDB-Tk provide robust approaches for microbial classification using continuously updated databases and phylogenetic analyses based on multiple genes. In addition to all the bioinformatic tools described above for the identification of the bacterial isolate, we also suggest the rMLST approach, which is a web-accessible database with support for an online database hosted by PubMLST. The approach is based on the analysis of 53 highly conserved ribosomal genes for the identification and classification of a wide range of bacterial species, and can be applied to poorly characterized or undescribed species. The rMLST database is compatible with whole-genome sequencing data or metagenomes, facilitating the analysis of complex microbial communities.

Approaches that combine multiple analyses and metrics are strongly encouraged, as they provide greater accuracy in bacterial identification and overcome the limitations of 16S rRNA gene sequencing in distinguishing closely related species.

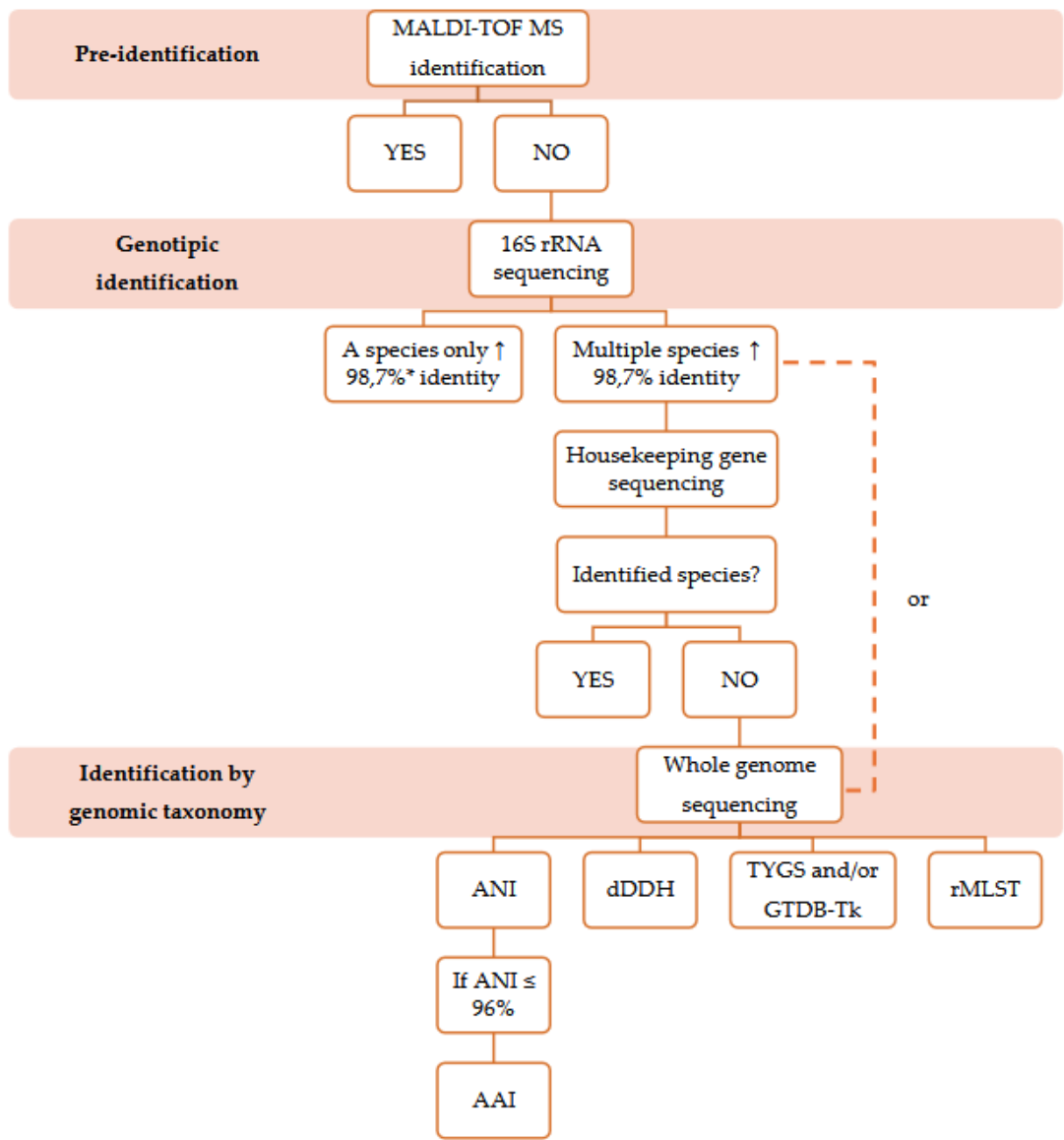


Figure 2. Steps of identification of environmental bacterial isolates, from MALDI-TOF MS to genomic taxonomy analyses. * According to Stackebrandt, Ebbers (2006).

Author Contributions: Conceptualization, M.L.L.B.; formal analysis, J.N.R., L.V.C., V.V.V. and M.L.L.B.; data curation, L.V.C. and M.L.L.B.; writing-original draft preparation, J.N.R.; writing-review and editing, J.N.R., L.V.C., V.V.V. and M.L.L.B.; supervision, M.L.L.B. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro – FAPERJ (E-26/200.546/2025). The funding body played no role in the design of the study and writing the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AAI	Average Amino Acid Identity
ANI	Average Nucleotide Identity
ATCC	American Type Culture Collection
BacDive	Bacterial Diversity Metadatabase
BLAST	Basic Local Alignment Search tool
bp	Base pairs
DDBJ	DNA Data Bank of Japan
dDDH	digital DNA-DNA hybridization
DDH	Hibridization DNA-DNA
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen
EMBL-EBI	European Molecular Biology Laboratory – European Bioinformatics Institute
G+C	Guanine and cytosine
GBDP	Genome BLAST Distance Phylogeny
GGDC	Genome-to-Genome Distance Calculator
GMPs	Good manufacturing practices
GTDB	Genome Taxonomy Database
GTDB-Tk	Genome Taxonomy Database Toolkit
JGI	Joint Genome Institute
LPSN	List of Prokatiotic Names with Standing in Nomenclature
MALDI-TOF MS	Matrix-Assisted Laser Desorption Ionization–Time of Flight/Mass Spectrometry
MBP	multigene-based phylogenies
MLSA	multilocus sequence analysis
MLST	multilocus sequence typing
NCBI	National Center for Biotechnology Information
ORFs	open reading frames
RED	Relative Evolutionary Divergence
rMLST	Ribosomal Multilocus Sequence Typing
rRNA	Ribosomal RNA
TYGS	Type Strain Genome Server
WGS	Whole-genome sequencing

References

- Costa, L. V. da; Miranda, R. V. da S. L. de; Reis, C. M. F. dos; Andrade, J. M. de; Cruz, F. V.; Frazão, A. M.; Fonseca, E. L. da; Ramos, J. N.; Brandão, M. L. L.; Vieira, V. V. MALDI-TOF MS Database Expansion for Identification of Bacillus and Related Genera Isolated from a Pharmaceutical Facility. *J Microbiol Methods* **2022**, 203. <https://doi.org/10.1016/J.MIMET.2022.106625>.
- United States Pharmacopeial Convention. *The United States Pharmacopeia 43rd ed.* The United States Pharmacopeia.
- Caldeira, N. G. S.; de Souza, M. L. S.; de Miranda, R. V. da S. L.; da Costa, L. V.; Forsythe, S. J.; Zahner, V.; Brandão, M. L. L. Characterization by MALDI-TOF MS and 16S RRNA Gene Sequencing of Aerobic Endospore-Forming Bacteria Isolated from Pharmaceutical Facility in Rio de Janeiro, Brazil. *Microorganisms* **2024**, Vol. 12, Page 724 2024, 12 (4), 724. <https://doi.org/10.3390/MICROORGANISMS12040724>.
- Food and Drug Administration - FDA. Guidance for Industry Sterile Drug Products Produced by Aseptic Processing –Current Good Manufacturing Practice,. ed by Aseptic Processing –Current Good Manufacturing Practice, .
- European Medicines Agency. The Rules Governing Medicinal Products in the European Union. Volume 4: European Union Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use. Annex 1: Manufacture of Sterile Medicinal Products. . The Rules Governing Medicinal

- Products in the European Union. Volume 4: European Union Guidelines for Good Manufacturing Practice for Medicinal Products for Human and Veterinary Use. Annex 1: Manufacture of Sterile Medicinal Products.
6. Thompson, C. C.; Vidal, L.; Salazar, V.; Swings, J.; Thompson, F. L. Microbial Genomic Taxonomy. In *Trends in the systematics of bacteria and fungi*; CABI: UK, **2021**; pp 168–178. <https://doi.org/10.1079/9781789244984.0168>.
 7. Moreira, F. M.; Pereira, P. de A.; Miranda, R. V. da S. L. de; Reis, C. M. F. dos; Braga, L. M. P. da S.; de Andrade, J. M.; do Nascimento, L. G.; Mattoso, J. M. V.; Forsythe, S. J.; da Costa, L. V.; Brandão, M. L. L. Evaluation of MALDI-TOF MS, Sequencing of D2 LSU rRNA and Internal Transcribed Spacer Regions (ITS) for the Identification of Filamentous Fungi Isolated from a Pharmaceutical Facility. *J Pharm Biomed Anal* **2023**, 234. <https://doi.org/10.1016/J.JPBA.2023.115531>.
 8. Costa, L. V. da; Ramos, J. N.; Albuquerque, L. de S.; Miranda, R. V. da S. L. de; Valadão, T. B.; Veras, J. F. C.; Vieira, E. M. D.; Forsythe, S.; Brandão, M. L. L.; Vieira, V. V. *Bacillus Lumedeiriae* Sp. Nov., a Gram-Positive, Spore-Forming Rod Isolated from a Pharmaceutical Facility Production Environment and Added to the MALDI Biotyper® Database. *Microorganisms* **2024**, 12 (12). <https://doi.org/10.3390/MICROORGANISMS12122507>.
 9. Obasi, A.; Nwachukwu, S.; Ugoji, E.; Kohler, C.; Göhler, A.; Balau, V.; Pfeifer, Y.; Steinmetz, I. Extended-Spectrum β -Lactamase-Producing *Klebsiella Pneumoniae* from Pharmaceutical Wastewaters in South-Western Nigeria. *Microbial Drug Resistance* **2017**, 23 (8), 1013–1018. <https://doi.org/10.1089/MDR.2016.0269>.
 10. da Silva Lage de Miranda, R. V.; da Costa, L. V.; de Sousa Albuquerque, L.; dos Reis, C. M. F.; da Silva Braga, L. M. P.; de Andrade, J. M.; Ramos, J. N.; Mattoso, J. M. V.; Forsythe, S. J.; Brandão, M. L. L. Identification of *Sutcliffiella Horikoshii* Strains in an Immunobiological Pharmaceutical Industry Facility. *Lett Appl Microbiol* **2023**, 76 (5). <https://doi.org/10.1093/LAMBIO/OVAD056>.
 11. Costa, L. V. da; Miranda, R. V. da S. L. de; Fonseca, E. L. da; Gonçalves, N. P.; Reis, C. M. F. dos; Frazão, A. M.; Cruz, F. V.; Brandão, M. L. L.; Ramos, J. N.; Vieira, V. V. Assessment of VITEK® 2, MALDI-TOF MS and Full Gene 16S rRNA Sequencing for Aerobic Endospore-Forming Bacteria Isolated from a Pharmaceutical Facility. *J Microbiol Methods* **2022**, 194, 106419. <https://doi.org/10.1016/J.MIMET.2022.106419>.
 12. Sala-Comorera, L.; Vilaró, C.; Galofré, B.; Blanch, A. R.; García-Aljaro, C. Use of Matrix-Assisted Laser Desorption/Ionization–Time of Flight (MALDI-TOF) Mass Spectrometry for Bacterial Monitoring in Routine Analysis at a Drinking Water Treatment Plant. *Int J Hyg Environ Health* **2016**, 219 (7), 577–584. <https://doi.org/10.1016/J.IJHEH.2016.01.001>.
 13. Vithanage, N. R.; Yeager, T. R.; Jadhav, S. R.; Palombo, E. A.; Datta, N. Comparison of Identification Systems for Psychrotrophic Bacteria Isolated from Raw Bovine Milk. *Int J Food Microbiol* **2014**, 189, 26–38. <https://doi.org/10.1016/J.IJFOODMICRO.2014.07.023>.
 14. Biomérieux. VITEK® 2. VITEK® 2. Fully integrated Identification and Antimicrobial Susceptibility Testing.
 15. Bosshard, P. P.; Zbinden, R.; Abels, S.; Böddinghaus, B.; Altwegg, M.; Böttger, E. C. 16S rRNA Gene Sequencing versus the API 20 NE System and the VITEK 2 ID-GNB Card for Identification of Nonfermenting Gram-Negative Bacteria in the Clinical Laboratory. *J Clin Microbiol* **2006**, 44 (4), 1359–1366. <https://doi.org/10.1128/JCM.44.4.1359-1366.2006/ASSET/A29BD8EB-F9F4-48EA-B418-11F2E6172524/ASSETS/GRAPHIC/ZJM0040663480001.JPEG>.
 16. Seuylemezian, A.; Aronson, H. S.; Tan, J.; Lin, M.; Schubert, W.; Vaishampayan, P. Development of a Custom MALDI-TOF MS Database for Species-Level Identification of Bacterial Isolates Collected From Spacecraft and Associated Surfaces. *Front Microbiol* **2018**, 9 (MAY). <https://doi.org/10.3389/FMICB.2018.00780>.
 17. Zasada, A. A.; Mosiej, E. Contemporary Microbiology and Identification of *Corynebacteria* Spp . Causing Infections in Human. *Lett Appl Microbiol* **2018**, 66 (11), 472–483. <https://doi.org/10.1111/lam.12883>.
 18. Shah, H. N.; Shah, A. J.; Belgacem, O.; Ward, M.; Dekio, I.; Selami, L.; Duncan, L.; Bruce, K.; Xu, Z.; Mkrtchyan, H. V.; Cave, R.; Shah, L. M. N.; Gharbia, S. E. MALDI-TOF MS and Currently Related Proteomic Technologies in Reconciling Bacterial Systematics. In *Trends in the systematics of bacteria and fungi*; CABI: UK, 2021; pp 93–118. <https://doi.org/10.1079/9781789244984.0093>.

19. Stackebrandt, E.; Ebers, J. Taxonomic Parameters Revisited: Tarnished Gold Standards – ScienceOpen. *Microbial Today*. **2006**, p 152. <https://www.scienceopen.com/document?vid=0cf4b084-5683-4ef4-a80c-c0df44a135dc> (accessed 2025-03-17).
20. Lasch, P.; Stämmeler, M.; Schneider, A. A MALDI-TOF Mass Spectrometry Database for Identification and Classification of Highly Pathogenic Microorganisms from the Robert Koch-Institute (RKI). <https://doi.org/10.5281/ZENODO.163517>.
21. Caamaño-Antelo, S.; Fernández-No, I. C.; Böhme, K.; Ezzat-Alnakip, M.; Quintela-Baluja, M.; Barros-Velázquez, J.; Calo-Mata, P. Genetic Discrimination of Foodborne Pathogenic and Spoilage *Bacillus* Spp. Based on Three Housekeeping Genes. *Food Microbiol* **2015**, *46*, 288–298. <https://doi.org/10.1016/J.FM.2014.08.013>.
22. Rajendhran, J.; Gunasekaran, P. Microbial Phylogeny and Diversity: Small Subunit Ribosomal RNA Sequence Analysis and Beyond. *Microbiol Res* **2011**, *166* (2), 99–110. <https://doi.org/10.1016/J.MICRES.2010.02.003>.
23. Church, D. L.; Cerutti, L.; Gürtler, A.; Griener, T.; Zelazny, A.; Emler, S. Performance and Application of 16S rRNA Gene Cycle Sequencing for Routine Identification of Bacteria in the Clinical Microbiology Laboratory. *Clin Microbiol Rev* **2020**, *33* (4), e00053-19. <https://doi.org/10.1128/CMR.00053-19>.
24. Madigan, M.; Martinko, J.; Bender, K.; Buckley, D.; Stahl, D. *Brock Biology of Microorganisms*, 14th ed.; Benjamin Cummings, **2015**.
25. Mahato, N. K.; Gupta, V.; Singh, P.; Kumari, R.; Verma, H.; Tripathi, C.; Rani, P.; Sharma, A.; Singhvi, N.; Sood, U.; Hira, P.; Kohli, P.; Nayyar, N.; Puri, A.; Bajaj, A.; Kumar, R.; Negi, V.; Talwar, C.; Khurana, H.; Nagar, S.; Sharma, M.; Mishra, H.; Singh, A. K.; Dhingra, G.; Negi, R. K.; Shakarad, M.; Singh, Y.; Lal, R. Microbial Taxonomy in the Era of OMICS: Application of DNA Sequences, Computational Tools and Techniques. *Antonie Van Leeuwenhoek* **2017**, *110* (10), 1357–1371. <https://doi.org/10.1007/S10482-017-0928-1>.
26. D'Amore, R.; Ijaz, U. Z.; Schirmer, M.; Kenny, J. G.; Gregory, R.; Darby, A. C.; Shakyia, M.; Podar, M.; Quince, C.; Hall, N. A Comprehensive Benchmarking Study of Protocols and Sequencing Platforms for 16S rRNA Community Profiling. *BMC Genomics* **2016**, *17* (1), 55. <https://doi.org/10.1186/S12864-015-2194-9>.
27. Rodrigues, N. M. B.; Bronzato, G. F.; Santiago, G. S.; Botelho, L. A. B.; Moreira, B. M.; Coelho, I. da S.; Souza, M. M. S. de; Coelho, S. de M. de O. The Matrix-Assisted Laser Desorption Ionization–Time of Flight Mass Spectrometry (MALDI-TOF MS) Identification versus Biochemical Tests: A Study with Enterobacteria from a Dairy Cattle Environment. *Brazilian Journal of Microbiology* **2016**, *48* (1), 132. <https://doi.org/10.1016/J.BJM.2016.07.025>.
28. Stackebrandt, E.; Mondotte, J. A.; Fazio, L. L.; Jetten, M. Authors Need to Be Prudent When Assigning Names to Microbial Isolates. *Curr Microbiol* **2021**, *78* (12), 4005–4008. <https://doi.org/10.1007/S00284-021-02678-4/METRICS>.
29. NIAID Visual & Medical Arts. *Ribosome*. NIAID NIH BIOART Source.
30. Tindall, B. J.; Rosselló-Móra, R.; Busse, H. J.; Ludwig, W.; Kämpfer, P. Notes on the Characterization of Prokaryote Strains for Taxonomic Purposes. *Int J Syst Evol Microbiol* **2010**, *60* (Pt 1), 249–266. <https://doi.org/10.1099/IJS.0.016949-0>.
31. Sentausa, E.; Fournier, P. E. Advantages and Limitations of Genomics in Prokaryotic Taxonomy. *Clinical Microbiology and Infection* **2013**, *19* (9), 790–795. <https://doi.org/10.1111/1469-0691.12181/ATTACHMENT/53D5D670-64BB-4218-A52D-6219AF9C4287/MMC2.JPG>.
32. Vlach, J.; Javůrková, B.; Karamonová, L.; Blažková, M.; Fukal, L. Novel PCR-RFLP System Based on RpoB Gene for Differentiation of *Cronobacter* Species. *Food Microbiol* **2017**, *62*, 1–8. <https://doi.org/10.1016/J.FM.2016.08.004>.
33. Payne, G. W.; Vandamme, P.; Morgan, S. H.; LiPuma, J. J.; Coenye, T.; Weightman, A. J.; Jones, T. H.; Mahenthalingam, E. Development of a RecA Gene-Based Identification Approach for the Entire *Burkholderia* Genus. *Appl Environ Microbiol* **2005**, *71* (7), 3917–3927. <https://doi.org/10.1128/AEM.71.7.3917-3927.2005>.
34. Chun, J.; Rainey, F. A. Integrating Genomics into the Taxonomy and Systematics of the Bacteria and Archaea. *Int J Syst Evol Microbiol* **2014**, *64* (PART 2), 316–324. <https://doi.org/10.1099/IJS.0.054171-0/CITE/REFWORKS>.

35. Glaeser, S. P.; Kämpfer, P. Multilocus Sequence Analysis (MLSA) in Prokaryotic Taxonomy. *Syst Appl Microbiol* **2015**, *38* (4), 237–245. <https://doi.org/10.1016/J.SYAPM.2015.03.007>.
36. Maiden, M. C. J.; Bygraves, J. A.; Feil, E.; Morelli, G.; Russell, J. E.; Urwin, R.; Zhang, Q.; Zhou, J.; Zurth, K.; Caugant, D. A.; Feavers, I. M.; Achtman, M.; Spratt, B. G. Multilocus Sequence Typing: A Portable Approach to the Identification of Clones within Populations of Pathogenic Microorganisms. *Proc Natl Acad Sci U S A* **1998**, *95* (6), 3140–3145. <https://doi.org/10.1073/PNAS.95.6.3140>.
37. Stackebrandt, E.; Frederiksen, W.; Garrity, G. M.; Grimont, P. A. D.; Kämpfer, P.; Maiden, M. C. J.; Nesme, X.; Rosselló-Mora, R.; Swings, J.; Trüper, H. G.; Vauterin, L.; Ward, A. C.; Whitman, W. B. Report of the Ad Hoc Committee for the Re-Evaluation of the Species Definition in Bacteriology. *Int J Syst Evol Microbiol* **2002**, *52* (3), 1043–1047. <https://doi.org/10.1099/00207713-52-3-1043>.
38. Hayashi Sant'Anna, F.; Bach, E.; Porto, R. Z.; Guella, F.; Hayashi Sant'Anna, E.; Passaglia, L. M. P. Genomic Metrics Made Easy: What to Do and Where to Go in the New Era of Bacterial Taxonomy. *Crit Rev Microbiol* **2019**, *45* (2), 182–200. <https://doi.org/10.1080/1040841X.2019.1569587>.
39. Thompson, C. C.; Chimetto, L.; Edwards, R. A.; Swings, J.; Stackebrandt, E.; Thompson, F. L. Microbial Genomic Taxonomy. *BMC Genomics* **2013**, *14* (1). <https://doi.org/10.1186/1471-2164-14-913>.
40. Land, M.; Hauser, L.; Jun, S. R.; Nookaew, I.; Leuze, M. R.; Ahn, T. H.; Karpinets, T.; Lund, O.; Kora, G.; Wassenaar, T.; Poudel, S.; Ussery, D. W. Insights from 20 Years of Bacterial Genome Sequencing. *Funct Integr Genomics* **2015**, *15* (2), 141. <https://doi.org/10.1007/S10142-015-0433-4>.
41. Wayne, L. G.; Brenner, D. J.; Colwell, R. R.; Grimont, P. A. D.; Kandler, O.; Krichevsky, M. I.; Moore, L. H.; Moore, W. E. C.; Murray, R. G. E.; Stackebrandt, E.; Starr, M. P.; Truper, H. G. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. *Int J Syst Evol Microbiol* **1987**, *37* (4), 463–464. <https://doi.org/10.1099/00207713-37-4-463>.
42. Chun, J.; Oren, A.; Ventosa, A.; Christensen, H.; Arahal, D. R.; da Costa, M. S.; Rooney, A. P.; Yi, H.; Xu, X. W.; De Meyer, S.; Trujillo, M. E. Proposed Minimal Standards for the Use of Genome Data for the Taxonomy of Prokaryotes. *Int J Syst Evol Microbiol* **2018**, *68* (1), 461–466. <https://doi.org/10.1099/IJSEM.0.002516/CITE/REFWORKS>.
43. Gosselin, S.; Fullmer, M. S.; Feng, Y.; Gogarten, J. P. Improving Phylogenies Based on Average Nucleotide Identity, Incorporating Saturation Correction and Nonparametric Bootstrap Support. *Syst Biol* **2022**, *71* (2), 396–409. <https://doi.org/10.1093/SYSBIO/SYAB060>.
44. Meier-Kolthoff, J. P.; Carbasse, J. S.; Peinado-Olarte, R. L.; Göker, M. TYGS and LPSN: A Database Tandem for Fast and Reliable Genome-Based Classification and Nomenclature of Prokaryotes. *Nucleic Acids Res* **2022**, *50* (D1), D801–D807. <https://doi.org/10.1093/NAR/GKAB902>.
45. Wang, J.; ran, Q.; Du, X.; Wu, S.; Wang, J.; Sheng, D.; Chen, Q.; Du, Z.; Li, Y. zhong. Two New Polyangium Species, *P. Aurulentum* Sp. Nov. and *P. Jinanense* Sp. Nov., Isolated from a Soil Sample. *Syst Appl Microbiol* **2021**, *44* (6). <https://doi.org/10.1016/J.SYAPM.2021.126274>.
46. Cuny, H.; Offret, C.; Boukerb, A. M.; Parizadeh, L.; Lesouhaitier, O.; Le Chevalier, P.; Jégou, C.; Bazire, A.; Brillet, B.; Fleury, Y. *Pseudoalteromonas Ostreae* Sp. Nov., a New Bacterial Species Harboured by the Flat Oyster *Ostrea Edulis*. *Int J Syst Evol Microbiol* **2021**, *71* (11). <https://doi.org/10.1099/IJSEM.0.005070>.
47. Colston, S. M.; Fullmer, M. S.; Beka, L.; Lamy, B.; Peter Gogarten, J.; Graf, J. Bioinformatic Genome Comparisons for Taxonomic and Phylogenetic Assignments Using *Aeromonas* as a Test Case. *mBio* **2014**, *5* (6). <https://doi.org/10.1128/MBIO.02136-14>.
48. Goris, J.; Konstantinidis, K. T.; Klappenbach, J. A.; Coenye, T.; Vandamme, P.; Tiedje, J. M. DNA-DNA Hybridization Values and Their Relationship to Whole-Genome Sequence Similarities. *Int J Syst Evol Microbiol* **2007**, *57* (Pt 1), 81–91. <https://doi.org/10.1099/IJS.0.64483-0>.
49. Konstantinidis, K. T.; Tiedje, J. M. Genomic Insights That Advance the Species Definition for Prokaryotes. *Proc Natl Acad Sci U S A* **2005**, *102* (7), 2567–2572. https://doi.org/10.1073/PNAS.0409727102/SUPPL_FILE/09727FIG6.PDF.
50. Qin, Q. L.; Xie, B. Bin; Zhang, X. Y.; Chen, X. L.; Zhou, B. C.; Zhou, J.; Oren, A.; Zhang, Y. Z. A Proposed Genus Boundary for the Prokaryotes Based on Genomic Insights. *J Bacteriol* **2014**, *196* (12), 2210. <https://doi.org/10.1128/JB.01688-14>.

51. Kim, D.; Park, S.; Chun, J. Introducing EzAAI: A Pipeline for High Throughput Calculations of Prokaryotic Average Amino Acid Identity. *J Microbiol* **2021**, *59* (5), 476–480. <https://doi.org/10.1007/S12275-021-1154-0>.
52. Rodriguez-R., L. M.; Konstantinidis, K. T. Bypassing Cultivation To Identify Bacterial Species: Culture-Independent Genomic Approaches Identify Credibly Distinct Clusters, Avoid Cultivation Bias, and Provide True Insights into Microbial Species. *Microbe Magazine* **2014**, *9* (3), 111–118. <https://doi.org/10.1128/MICROBE.9.111.1>.
53. Nicholson, A. C.; Gulvik, C. A.; Whitney, A. M.; Humrighouse, B. W.; Bell, M. E.; Holmes, B.; Steigerwalt, A. G.; Villarma, A.; Sheth, M.; Batra, D.; Rowe, L. A.; Burroughs, M.; Pryor, J. C.; Bernardet, J. F.; Hugo, C.; Kämpfer, P.; Newman, J. D.; McQuiston, J. R. Division of the Genus *Chryseobacterium*: Observation of Discontinuities in Amino Acid Identity Values, a Possible Consequence of Major Extinction Events, Guides Transfer of Nine Species to the Genus *Epilithonimonas*, Eleven Species to the Genus *Kaistella*, and Three Species to the Genus *Halpernia* Gen. Nov., with Description of *Kaistella Daneshvariae* Sp. Nov. and *Epilithonimonas Vandammei* Sp. Nov. Derived from Clinical Specimens. *Int J Syst Evol Microbiol* **2020**, *70* (8), 4432–4450. <https://doi.org/10.1099/IJSEM.0.003935>.
54. Walter, J. M.; Coutinho, F. H.; Dutilh, B. E.; Swings, J.; Thompson, F. L.; Thompson, C. C. Ecogenomics and Taxonomy of Cyanobacteria Phylum. *Front Microbiol* **2017**, *8* (NOV). <https://doi.org/10.3389/FMICB.2017.02132>.
55. Zheng, J.; Wittouck, S.; Salvetti, E.; Franz, C. M. A. P.; Harris, H. M. B.; Mattarelli, P.; O'toole, P. W.; Pot, B.; Vandamme, P.; Walter, J.; Watanabe, K.; Wuyts, S.; Felis, G. E.; Gänzle, M. G.; Lebeer, S. A Taxonomic Note on the Genus *Lactobacillus*: Description of 23 Novel Genera, Emended Description of the Genus *Lactobacillus* Beijerinck 1901, and Union of *Lactobacillaceae* and *Leuconostocaceae*. *Int J Syst Evol Microbiol* **2020**, *70* (4), 2782–2858. <https://doi.org/10.1099/IJSEM.0.004107>.
56. Riesco, R.; Trujillo, M. E. Update on the Proposed Minimal Standards for the Use of Genome Data for the Taxonomy of Prokaryotes. *Int J Syst Evol Microbiol* **2024**, *74* (3). <https://doi.org/10.1099/IJSEM.0.006300>.
57. Meier-Kolthoff, J. P.; Göker, M. TYGS Is an Automated High-Throughput Platform for State-of-the-Art Genome-Based Taxonomy. *Nat Commun* **2019**, *10* (1), 2182. <https://doi.org/10.1038/S41467-019-10210-3>.
58. Parks, D. H.; Chuvochina, M.; Rinke, C.; Mussig, A. J.; Chaumeil, P. A.; Hugenholtz, P. GTDB: An Ongoing Census of Bacterial and Archaeal Diversity through a Phylogenetically Consistent, Rank Normalized and Complete Genome-Based Taxonomy. *Nucleic Acids Res* **2022**, *50* (D1), D785–D794. <https://doi.org/10.1093/NAR/GKAB776>.
59. Jolley, K. A.; Bliss, C. M.; Bennett, J. S.; Bratcher, H. B.; Brehony, C.; Colles, F. M.; Wimalaratna, H.; Harrison, O. B.; Sheppard, S. K.; Cody, A. J.; Maiden, M. C. J. Ribosomal Multilocus Sequence Typing: Universal Characterization of Bacteria from Domain to Strain. *Microbiology (Reading)* **2012**, *158* (Pt 4), 1005–1015. <https://doi.org/10.1099/MIC.0.055459-0>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.