



## Whitepaper

# Quality, quantity or maybe both? A look behind the curtain of MALDI Biotyper® library entry creation

Author: Dr. Thomas Maier, Director Scientific Affairs Europe, Bruker Microbiology & Infection Diagnostics

Accurate microbial identification by MALDI-TOF MS depends not only on high-quality spectra, but equally on correct strain selection, robust species delineation, and nomenclature that reflects the current scientific consensus and internationally accepted rules. Here we are going to explain Bruker's concept for: (1) selecting and measuring candidate isolates for new MALDI Biotyper library entries; (2) establishing reliable taxonomic assignments using a polyphasic framework; (3) validating and applying correct, validly published species names (with attention to clinically relevant naming stability); and (4) maintaining a controlled, MALDI-aware naming convention that aligns taxonomic evidence with the practical resolution limits of MALDI. A meticulous approach to maintain high library entry quality, while going for large species coverage.

### Strain selection

When new species shall be added to the MALDI Biotyper reference library, the general objective is to obtain as many isolates of this species as possible and necessary for comprehensive species identification (generally in the range of 10 to 20, not hundreds). Ideally, these isolates should be as diverse as possible. "Diverse" refers to efforts to include isolates originating from different geographical regions or from different hosts and variation in MALDI spectra. Isolates that have been pre-characterized using various typing methods and are known to represent diverse lineages are particularly valuable.

If several isolates are available, each isolate is cultivated on its preferred medium, followed by optimized sample preparation and MALDI measurement. If many isolates of a species are available, the final selection of reference strains for inclusion in the library is based exclusively on MALDI-specific criteria. For example, if isolates from different hosts (animals and humans) show identical MALDI spectra, only one representative isolate is selected. Conversely, if two isolates isolated from the same patient show high MALDI diversity, both isolates are included in the library. Thus, the initial biological or geographical considerations serve only as a basis for assembling a broad candidate set; the ultimate selection is determined solely by MALDI performance and spectral diversity.

After completing the strain selection based on MALDI spectral diversity, the next essential step is a reliable species delineation for each individual strain. The following chapter provides a brief excursus addressing the concepts of "taxonomy" and "nomenclature" and the process of determining the correct species name. These aspects form the basis for consistent species identification within the MALDI Biotyper library.

## Taxonomy

Microbial taxonomy nowadays commonly follows a “polyphasic approach”. This means that the taxonomic characterization of a strain is no longer based on a single identification method; instead, a combination of very different methods is applied to achieve a reliable species assignment.

The following section provides a summary of polyphasic taxonomy as derived from the publication by Vandamme *et al.*, 1996<sup>1</sup>.

Polyphasic taxonomy, introduced conceptually by Colwell in 1970, represents an integrative approach in bacterial systematics that combines phenotypic, genotypic, and phylogenetic evidence to establish consensus-based taxonomic units. Over the past decades, this framework has become a major milestone in modern bacterial taxonomy, emphasizing that species concepts remain dynamic and will continue to evolve with technological progress rather than reaching a definitive endpoint.

Taxonomy - often used synonymously with systematics - is traditionally divided into three interrelated components: (i) **classification**, the organization of organisms into groups; (ii) **nomenclature**, the formal naming of these groups; and (iii) **identification**, the assignment of unknown isolates to the established taxa. Within this structure, the bacterial species is defined as a group of strains, including the type strain, that share high genomic similarity, historically expressed as  $\geq 70\%$  DNA–DNA hybridization with  $\leq 5^\circ\text{C } \Delta T_m$ , supported by consistent phenotypic and chemotaxonomic traits.

However, DNA-DNA hybridization methods vary technically and quantitatively, making the traditional thresholds approximate rather than absolute. Alternative phylogenetic approaches, such as defining species boundaries through 16S rRNA gene similarity, have been considered but remain imperfect substitutes for genome wide comparisons.

Polyphasic identification applies these principles to the practical classification of isolates. While most routine laboratory identifications can be performed using single method (monophasic) phenotypic or molecular schemes, their reliability depends on the quality and taxonomic accuracy of reference databases. Truly stable and well defined taxonomic entries typically require polyphasic evaluation. Consequently, new, rare, or atypical isolates often necessitate a polyphasic approach that integrates multiple complementary data types, thereby ensuring more robust species delineation and enabling accurate downstream monophasic identification workflows.

Fourteen years after Vandamme *et al.* 1996, the publication by Moore *et al.*, 2010<sup>3</sup> further refined the microbial systematics and taxonomy. In addition to clarifying the concepts of “characterization,”

“classification” and “nomenclature”, this work also marked the first appearance of the MALDI method in a taxonomic publication. Presented below is a non exhaustive summary of the publication.

Taxonomy comprises three interdependent components - characterization, classification, and nomenclature - each essential for a coherent microbial taxonomy. Species represent the foundational unit, defined as genomically coherent groups of strains, including the type strain, that share high similarity across multiple independently assessed characteristics under standardized conditions.

Nomenclature provides formal names for taxa that have been characterized and classified. While characterization and classification follow general recommendations, the naming of prokaryotes is strictly regulated by the International Code of Nomenclature of Prokaryotes (ICNP, “Bacteriological Code”). A name becomes validly published only when it complies with the pertinent rules of the Code and is published in the *International Journal of Systematic and Evolutionary Microbiology* (IJSEM), which serves as the official registry.

Standardization of prokaryotic nomenclature was established in 1980 with the *Approved Lists of Bacterial Names*, consolidating 2,335 names of “adequately described” taxa. Names not included lost standing in nomenclature. Since then, new names become validly published only through IJSEM publication, either directly or by validation, whereas names published otherwise remain “effectively published”. The continuously updated online registry of validly published names is therefore essential for tracking nomenclatural changes.

Revisions must follow the rules of the Code, such as retaining species epithets when transferring species to new genera. Additionally, practical considerations may influence naming decisions, especially when taxonomic changes affect clinically important organisms. Such changes have downstream implications for medical microbiology, including the need to update commercial identification systems and ensure robust strain authentication in culture collections.

Some conclusive bullet points for taxonomical species delineation:

### ■ Definitions in **Taxonomy**:

- I. Characterization (genotypic, phenotypic, polyphasic **properties**).
- II. Classification (arrangement in taxonomic **groups**).
- III. Nomenclature (**labeling** of the units defined in II.).
- IV. Identification of unknown organisms (process of **determining** whether an organism belongs to one of the units defined in II. and labeled in III)

- **Taxonomy results** from different polyphasic methods can be **different**.
- **Taxonomic resolution** (strain, species, genus, family, etc.) of the different methods can be **different**.
- **“Species”** is the basic unit.
- **Distances between species** are different within the phylogenetic tree.
- There is **no official, formally regulated taxonomy** and there will never be a definitive classification of bacteria.
- **Nomenclature:** names should be validly published under the ICNP.

### Application to Bruker MALDI Biotyper library

After a potential new strain for the MALDI Biotyper library has been successfully measured, the next step is its taxonomic characterization within a polyphasic framework. For bacteria, this process typically begins with sequencing of the first 500 bp of the 16S rRNA gene. The sequence is compared against several curated reference databases. If, according to the CLSI MM18-A guidelines, this DNA sequence is sufficient to assign the isolate to a species, the polyphasic approach stops already at this stage.

If the 16S rRNA gene sequence does not provide sufficient taxonomical resolution and the MALDI results indicate that the isolate can be assigned to one of several closely related species, additional protein genes will be used to confirm the species-level identification of the isolate. If DNA-based methods are generally insufficient or unsuitable for a certain group of species, other methods like biochemical profiling can be used as well.

Nowadays also whole genome sequencing (WGS) with own species-specific thresholds (e.g. ANI>95% or dDDH>70%) is used for the species delineation of MALDI Biotyper reference entries. The advantage of WGS is its very high species resolution. But this is at the same time also its biggest disadvantage: the high species resolution provided by WGS exceeds the discriminatory capacity of MALDI. This is in the same way true for the 16S rRNA based species resolution which cannot be reflected by WGS as well.

After obtaining the initial species assignment based on curated DNA sequence databases, the next step is to determine a validly published name for the strain. It will be checked if the pre-identification fits the current nomenclature rules. The nomenclature used in the Bruker MALDI Biotyper follows, to the greatest extent possible, the ICNP nomenclature rules basically described (e.g.) by Moore *et al.*, 2010. The primary reference for validly published names is the “List of Prokaryotic

names with Standing in Nomenclature (LPSN)”, hosted at the DSMZ (<https://lpsn.dsmz.de>). This expert curated resource serves as the basis for determining whether a name is validly published, correctly formed, and considered the “correct name,” i.e. the name to be applied to a taxon. A key component of LPSN is the so-called “List of Recommended Names for bacteria of medical importance” (LoRN) which provides standardized names intended for consistent use within the microbial community. LoRN offers guidance when validly published species names undergo orthographic correction, reclassification, or replacement by names deemed more appropriate according to nomenclatural rules. Thus, LoRN effectively “translates” validly published names into the recommended forms that should be applied in routine practice.

A critical aspect in species designation for the MALDI Biotyper library is generally the “MALDI perspective”. Each final species delineation of an isolate is generally compared with the taxonomic MALDI resolution within the respective species group. If the MALDI Biotyper is unable to reliably distinguish two closely related species, this limitation is reflected in the naming convention applied to reference entries and/or in the generation of Matching Hints (MH). The final decisions regarding species names within the MALDI Biotyper libraries are made by Bruker and are determined individually for each single species group. There is no automatic rule like “if the sequence shows a certain property, the name is automatically determined”. Instead, species naming follows a controlled decision-making process that integrates taxonomic evidence, nomenclatural requirements, and the practical resolution limits of the MALDI method. Consequently, the final naming convention rests under the authority of Bruker and ensures consistency between taxonomic understanding and MALDI-based identification performance.

As one example the species “*Pseudomonas paraeruginosa*” was published in 2022<sup>6</sup> as a new species. However, its current nomenclatural status requires careful interpretation. According to the LPSN (<https://lpsn.dsmz.de>) and LoRN (<https://lpsn.dsmz.de/species/pseudomonas-paraeruginosa>) the use of this name is presently not recommended: “*This name is on the List of Recommended Names for bacteria of medical importance (LoRN) as a synonym of the name to be applied to this species. The name is on LoRN because of the recorded pathogenicity of the species itself. This name is taxonomically suspended (and therefore not recommended for medical use for now) until no later than 2027 in favour of Pseudomonas aeruginosa.*” That means that currently all MALDI Biotyper reference strains which would be finally determined as *P. paraeruginosa* by a reference method (e.g., with WGS) could be introduced as reference entry with the name “*Pseudomonas aeruginosa*” into the MALDI Biotyper library. Renaming is ultimately desired, but the naming of *P. paraeruginosa* will be discussed carefully and possibly the name remains “*P. aeruginosa*” for a different time period.

The Bruker reference entry naming convention in short bullet points:

- Reference names orient strongly on the LPSN names.
- The List of Recommended Names (LoRN) is considered.
- Renaming is desired, but the exact time point is internally discussed.
- Renaming is not an automatism. The entire library is not scanned automatically for new names.
- Renaming is not mandatory and many old names can be further used (according to the guidelines of the "Ad Hoc Committee on Mitigating Changes in Prokaryotic Nomenclature").
- "Double names"; "triple names" (combining two or three species) or even names at species group level can be used.
- If a strain does not correspond to any validly published name, a not validly published name can be used as an alternative name.

The taxonomic „species identification“ of an isolate is always an approximation to an existing and described species. Consequently, the species assignment within any reference library - including the MALDI Biotyper library - also represents an approximation based on the best available data and the discriminatory capabilities of the used method.

In contrast to the polyphasic species identification applied to isolates used (e.g.) for the creation of a reference library, species identification in routine diagnostic practice is typically **monophasic**. In everyday workflows, only one characteristic of the unknown isolate (e.g. biochemistry, 16S rRNA gene sequence or the MALDI spectrum) is determined by the respective identification system. As a result, routine identifications are inherently limited by the discriminatory power of the single method employed.

The considerations outlined above primarily apply to prokaryotes; however, the general conceptual framework is, to some extent, also applicable to yeasts and filamentous fungi. Their taxonomy is considerably more complex and relies on a broader set of phenotypic, biochemical, and molecular markers for species delineation. For yeasts, loci such as the D1/D2 domains of the LSU rRNA gene, ITS regions, and protein coding markers are routinely employed, while the classification of filamentous fungi often requires multilocus sequence analysis (MLSA) combined with micromorphological assessments. In addition, many fungal taxa exhibit extensive intraspecific variability and polyphyletic historical groupings, further complicating nomenclatural stability.

Within the MALDI Biotyper library, species designation for yeasts and filamentous fungi is likewise not automated. It results from an individual and case specific decision making process that integrates the established principles of fungal taxonomy with the methodological constraints of MALDI.

The final assignment of species names in the MALDI Biotyper reference library is determined by Bruker, but always in close consultation with experts in yeast and fungal taxonomy and experts in clinical microbiology. This collaboration ensures that nomenclatural decisions reflect current phylogenetic evidence, adhere to accepted taxonomic practice, and remain consistent with the analytical capabilities and technical limitations of MALDI-based identification while considering diagnostic requirements.

## Links

- LPSN - List of prokaryotic names with standing in Nomenclature: <https://lpsn.dsmz.de/>
- Introduction: <https://lpsn.dsmz.de/text/introduction>
- Nomenclature: <https://lpsn.dsmz.de/text/nomenclature>
- FAQ: <https://lpsn.dsmz.de/text/faq#where-can-i-find-the-list-of-recommended-names-for-bacteria-of-medical-importance>
- Glossary: <https://lpsn.dsmz.de/text/glossary##list-of-recommended-names-for-bacteria-of-medical-importance>
- Guidelines of the ad hoc committee on mitigating changes in prokaryotic nomenclature: <https://zenodo.org/records/15586964>

## References

- 1 **Vandamme P, Pot B, Gillis M, Vos PD, Kersters K, et al.** Polyphasic Taxonomy, a Consensus Approach to Bacterial Systematics. *MICROBIOL REV*;60.
- 2 **Colwell RR.** Polyphasic Taxonomy of the Genus *Vibrio*: Numerical Taxonomy of *Vibrio cholerae*, *Vibrio parahaemolyticus*, and Related *Vibrio* Species. *J Bacteriol* 1970;104:410–433.
- 3 **Moore ERB, Mihaylova SA, Vandamme P, Krichevsky MI, Dijkshoorn L.** Microbial systematics and taxonomy: relevance for a microbial commons. *Res Microbiol* 2010;161:430–438.
- 4 **Tindall BJ, Kämpfer P, Euzéby JP, Oren A.** Valid publication of names of prokaryotes according to the rules of nomenclature: past history and current practice. *Int J Syst Evol Microbiol* 2006;56:2715–2720.
- 5 **Skerman VDB, McGowan Vicki, Sneath PHA.** Approved Lists of Bacterial Names. *Int J Syst Evol Microbiol* 1980;30:225–420.
- 6 **Rudra B, Duncan L, Shah AJ, Shah HN, Gupta RS.** Phylogenomic and comparative genomic studies robustly demarcate two distinct clades of *Pseudomonas aeruginosa* strains: proposal to transfer the strains from an outlier clade to a novel species *Pseudomonas paraeruginosa* sp. nov. *Int J Syst Evol Microbiol*;72. Epub ahead of print 30 November 2022. DOI: 10.1099/ijsem.0.005542.
- 7 **Göker M, Christensen H, Fingerle V, Kostovski M, Margos G, et al.** List of Recommended Names for bacteria of medical importance: report of the Ad Hoc Committee on Mitigating Changes in Prokaryotic Nomenclature. *Int J Syst Evol Microbiol*;75. Epub ahead of print 23 October 2025. DOI: 10.1099/ijsem.0.006943.
- 8 **Sneath PH, Brenner DJ.** 'Official' nomenclature list. *Nomencl Lists Am Soc Microbiol News* 1992;58:175.
- 9 **Buchanan RE, Gibbons NE.** *Bergey's Manual of Determinative Bacteriology. Eighth Edition.* Baltimore: The Williams & Wilkins Company; 1974.
- 10 **Goodfellow M, O'Donnell AG.** *Handbook of new bacterial systematics.* London ; San Diego: Academic Press; 1993.
- 11 **Arahal DR, Bull CT, Busse H-J, Christensen H, Chuvochina M, et al.** Guidelines for interpreting the International Code of Nomenclature of Prokaryotes and for preparing a Request for an Opinion. *Int J Syst Evol Microbiol*;73. Epub ahead of print 15 March 2023. DOI: 10.1099/ijsem.0.005782.

 **Bruker Daltonics GmbH & Co. KG**

Bremen · Germany  
Phone +49 (0) 421-2205-0

[info.md@bruker.com](mailto:info.md@bruker.com)

**Bruker Scientific LLC**

Billerica, MA · USA  
Phone +1 (978) 663-3660

**Online information**  
[bruker.com/microbiology](https://bruker.com/microbiology)

