### FAQ

#### Frequently asked questions

 $\mathscr{O}$ 

Here we have compiled general answers to questions users had about LPSN. For explanations of terms and abbreviations and for information on specific topics please see the LPSN glossary.

#### Is there an introduction into LPSN?

ଡ

Yes, please see the LPSN introduction page.

#### Who has contributed to LPSN?

B

Please see the LPSN acknowledgements page.

#### How should LPSN requests be made?

6

We are asking users to report LPSN requests *via* the form for taxonomic submissions whenever possible and *via* the contact form in all other cases, using an appropriate subject. Requests may be questions, error reports, feature requests or other kinds of statements. An appropriate subject assists us in properly dealing with your request. The subject categories are as follows.

**Graphical user interface.** This category should be chosen if you are satisfied with some LPSN content but you opine that this content should be displayed in a different manner, or that a distinct kind of presentation of the content should additionally be offered. These would be feature requests. Alternatively, you can send us an error report if you believe that the presentation of the content leads to misunderstandings or is unsatisfactory in some other manner.

Missing taxon name. You should not normally choose this category but use the form for taxonomic submissions if you fail to find a certain taxon name in LPSN. Only if the request does not fit into this form an ordinary LPSN request should be made. Prior to sending us an error report about a missing name, please make sure you understand which taxon names are to be expected in LPSN and since which point in time. This is explained in a separate FAQ entry. (In contrast, the taxonomic submission form can immediately be used.) Asking us for including a certain taxon name may in many cases rather be regarded a a feature request, which is also explained in a distinct FAQ entry. Requesting LPSN to prefer a certain taxonomic opinion over another one, i.e. to regard another name a the correct name, should not be assigned to the "missing taxon name" category but to one described below.

**Missing statement.** This category should be chosen if you believe some relevant information to be missing in LPSN but it is not a missing taxon name and not a missing synonymy or child-parent relationship. (The form for taxonomic submissions should be used in these cases.) Depending on the kind of information a message in this category could be understood as error report or feature request. Please

see below for information that can be temporarily missing and should thus not immediately be reported as a problem. User-defined non-mandatory statements can well be included in LPSN, usually as rotes. For instance, there is a separate FAQ entry about how we can include your publication in LPSN.

**Inaccurate statement.** This LPSN category is frequently confused with the next one. In a literature-driven database such as LPSN, virtually all statements are linked to a literature source. If an LPSN statement is linked to a literature source that does not make that statement, LPSN is inaccurate, irrespective of whether or not the linked statement is accurate. A correction of LPSN is needed in that case. If an LPSN statement is linked to a literature source that actually makes that statement, LPSN is accurate, irrespective of whether or not the linked statement is accurate. If the linked statement is inaccurate, LPSN should also be modified, of course, but in that case by adding a hint regarding the erroneous statement found in the literature. Such an explicit clarification is preferable to completely removing some inaccurate literature statement, as researchers may still independently stumble over that error in the sources. LPSN does a lot of checking but we cannot detect all errors ourselves. An error *in* LPSN is not necessarily an error *of* LPSN – although errors made by LPSN can certainly occur.

These considerations are also of relevance for the taxonomy preferred by LPSN. The International Code of Nomenclature of Prokaryotes (ICNP) regulates nomenclature but does not regulate classification. For this reason, LPSN is not forced to prefer a certain taxonomic solution over another one, as long as both are compatible with the ICNP. In particular, LPSN does not need to adopt the most recently published taxonomic arrangement. Users are urged to take this into account and to not report deviating taxonomic opinions as an inaccuracy. Instead, the next category of requests should be used in such situations.

Alternative taxonomic opinion. This is virtually always a feature request, as opposed to an error report. As detailed above, LPSN does not need to prefer certain taxonomic solutions over others just because they were more recently proposed. Users are encouraged to ask for modifications of the taxonomic classification used by LPSN. However, users are also asked to not confuse the occurrence of alternative taxonomic views with actual inaccuracies. LPSN preferences of certain taxonomic arrangements over others may benefit from explicit justifications. This can be done in LPSN notes and may be a valuable outcome even if the LPSN maintainers chose to not immediately modify the LPSN classification in response to a user request. Note that the form for taxonomic submissions should be used to request the inclusion of child-parent relationships and synonymy relationships whenever possible.

### Why is a taxon name missing from LPSN?

 $\mathscr{O}$ 

This depends on the name and on the point in time. At a given time point a taxon name can be missing from LPSN for one of the following reasons.

- The name is the name of a eukaryote and thus beyond the scope of LPSN.
- The name is the name of a cyanobacterium validly published under the Botanical Code (ICN). Such
  names are of interest for LPSN although they do not belong to the core data set. Users are encouraged
  to send us requests for the inclusion of such names. However, LPSN does not guarantee that such
  names are incorporated after a given maximum time period.
- The name is not revalidly published under the revalid ICNP and not validly published under the ICN. Such names are of interest for LPSN although they do not belong to the core data set. Users are encouraged to send us requests for the inclusion of such names. However, LPSN does not guarantee that such names are incorporated after a given maximum time period.
- The name has only recently been validly published under the ICNP. This should not normally be reported. However, the form for taxonomic submissions can be used at any time.
- The name is validly published under the IN ICNP since quite some time. This may well be an error and should better be reported using the form for taxonomic submissions. (Note that we deliberately do not define here what "since quite some time" means, as this is supposed to change over time.)

For including taxon names from your own publications please see the FAQ entry below.

# Why is a taxon name that has been included in a Validation List marked as not validly published?

 $\mathscr{O}$ 

Taxon names that are validly published under the ICNP by means of inclusion in a Validation List are not marked as such until the Validation List has been imported into LPSN. If the Validation List is not mentioned on the LPSN page for the taxon name, the Validation List may simply not have been imported yet. Unless you find the same Validation List cited elsewhere on LPSN, such cases should not be reported.

On very rare occasions, names on a Validation List are found not to be validly published. (Note that inclusion on a Validation List is a sufficient, not a necessary, condition for valid publication). If this is the case, a note will describe the situation.

#### Can you include my publication in LPSN?

ଡ

Publications listed in LPSN are either original proposals of taxon names, formal emendations of descriptions of taxon names, or publications related to notes. A publication with a proposal of a taxon name or emendation can be added only if:

- it proposed a taxon name including Candidatus or any other name that is not validly published that is not yet listed in LPSN;
- it makes an emendation that is not yet listed in LPSN.

Publications with proposals of taxon names or emendations will be exchanged only if the wrong publication is given in LPSN.

A publication related to a note can be added at any time provided a statement is given that can serve as an informative new note, refers to a specific taxon name and originates from that publication. If you are aware of such a publication please send us the taxon name, the suggested text of the note if applicable, and the DOI or a link to the PDF file of that publication.

Note that a description of a strain that is not accompanied by the proposal of a taxon name will not be considered for inclusion. Also note that a request made *via* the form for taxonomic submissions will be processed much faster.

### Why and how does LPSN assign the status "correct name"?

6

What is the "correct name"? Some context based on the International Code of Nomenclature of Prokaryotes (ICNP) is needed to properly answer that question.

Principle 5 of the ICNP defines that "The correct name of a taxon is based upon valid publication, legitimacy and priority of publication", whilst Principle 8 adds that "Each phylum or taxon of a lower rank with a given circumscription, position, and rank can bear only one correct name, i.e., the earliest that is in accordance with the Rules of this Code.", whilst Note 2 to the Principle defines circumscription, position and rank. Rule 23a of the ICNP states that "Each taxon above and including species, up to and including order, with a given circumscription, position, and rank can bear only one correct name, i.e., the earliest that is in accordance with the Rules of this Code."

So how does LPSN deal with the issue of the "correct name"?

During the previous eras of LPSN a particular name was not designated as the correct name. At that time LPSN simply displayed the series of new combinations (or, rarely, nomina nova), if any. In practice this led some users to believe that the most recent name needs to be preferred. The LPSN curator at that time then added a disclaimer, but it remains unclear how much this helped.

Another issue was the number of species placed within a genus. This was shown on LPSN during all eras. However, is this number mainly relevant as purely nomenclatural information, indicating how many validly published species names are there within a genus? Or are users instead mainly interested in how many species are currently taxonomically placed in a genus? A purely nomenclatural count comes across each set of homotypic synonyms several times. This may not yield the information of interest.

When establishing the current implementation of LPSN, it seemed beneficial to also provide the functionality of DSMZ's former PNU service. PNU always selected one name as the correct name.

As noted above, a major misconception regarding the correct name is that, among a series of homotypic synonyms, the most recent validly published and legitimate name must be selected as the correct name. Such homotypic synonyms are mostly generated by proposing new combinations for expressing the affiliation of a species to another genus. However, the ICNP does not indicate any preference whatsoever for the last name among a set of validly published and legitimate names. In contrast, the ICNP emphasizes the goal of taxonomic freedom. Hence, both Principle 8 and Rule 23a refer to a correct name being contextualized by a "given circumscription, position, and rank". Position and circumscription of a taxon are matters of taxonomic opinion, on which the ICNP does not rule. One can assign only one correct name to a taxon at a given time. However, distinct authors can choose a distinct correct name, depending on their taxonomic opinion on circumscription and position of the taxon. Thus, the name considered by one party as the correct name may not so be considered by another party and yet both parties may be in accordance with the ICNP.

An example may illustrate this issue. Both *Klebsiella terrigena* Izard et al. 1981 and *Raoultella terrigena* (Izard et al. 1981) Drancourt et al. 2001 are validly published and legitimate names. *Klebsiella terrigena* is the correct name if the taxonomist's opinion is that the species should be classified in the genus *Klebsiella*. However, *Raoultella terrigena* is the correct name if the taxonomist's opinion is that the species should be classified in the genus *Raoultella*.

Another misconception is that a validly published and legitimate name that is not regarded as correct name would be "incorrect". In fact, such a name is just a synonym. The ICNP defines validly published and not validly published names as well as legitimate and illegitimate names but it does not define incorrect names.

One may now argue that whereas the selection of the most recent validly published and legitimate name as the correct name may not be based on the ICNP, one could still assume that later taxonomic studies were based on better data. This is often true. However, better data may not explain the difference in taxonomic opinions expressed in distinct studies. Data alone do not yield a taxonomic classification. One also needs a certain taxonomic concept. The difference between the outcomes from two subsequent taxonomic studies may just be due to a difference in the underlying taxonomic concepts.

#### **Examples**

- The dissection of the genus *Arcobacter* into six genera was opposed by other researchers. The six genera appeared monophyletic in the trees inferred from genome-scale data presented in the study that proposed the dissection. However, so did the genus *Arcobacter* sensu lato. For this reason, the proposal to dissect the genus is just an example for splitting vs. lumping and not an attempt to solve an instance of non-monophyly.
- The dissection of the genus *Borrelia* into two genera was opposed by other researchers. The two genera appeared monophyletic in the trees inferred from genome-scale data presented in the study that proposed the dissection. However, so did the genus *Borrelia* sensu lato. For this reason, the proposal to dissect the genus is just an example for splitting vs. lumping and not an attempt to solve an instance of non-monophyly.
- The dissection of the genus *Mycobacterium* into five genera was opposed by other researchers. The five genera appeared monophyletic in the trees inferred from genome-scale data presented in the study

that proposed the dissection. However, so did the genus *Mycobacterium* sensu lato. For this reason, the proposal to dissect the genus is just an example for splitting vs. lumping and not an attempt to solve an instance of non-monophyly.

Solving non-monophyly is an attempt to reconcile the taxonomic classification with a phylogenetic tree. However, each instance of non-monophyly could be solved by either splitting or lumping. The need to repair non-monophyletic taxa is a stronger argument for a taxonomic reclassification than just a distinct view on how divergent a taxon of the considered rank should be.

The counter-argument against any reclassification is, of course, taxonomic conservatism. Use of an older classification would not be arbitrary but just conservative (if it could also be based on a literature source, of course). Taxonomic conservatism can be overdone. For instance, the mere concern that a reclassification introduces new names is not in agreement with the ICNP.

For this reason, the best solution may well be to be conditionally conservative, i.e. to accept most instances of reclassification but not all of them. Importantly, a certain taxonomy that is skipped at first can be reconsidered at any later time point after observing the further developments in the literature.

The issue of consistency most likely arises anyway. LPSN applies a "concept consistency" approach. This approach aims at either incorporating a certain taxonomic concept in its entirety or not at all. In the case of a given taxonomic revision regarding a genus, all new combinations (or, rarely, nomina nova) that yielded a validly published and legitimate name are treated as correct name, or none of them. The only possible exception is the result of an even newer taxonomic revision. In that manner, taxonomic consistency is guaranteed while also taking taxonomic freedom into account.

In addition, the LPSN liaises with the taxonomic subcommittees of the ICSP on classification matters and attempts to incorporate information from these subcommittees whenever it deems appropriate.

#### Further reading

- ICNP
- Judicial Opinion 122
- Paper criticizing last-name-wins approach
- Misunderstanding the ICNP Arcobacter Borrelia Mycobacterium

# Why is the full information on the species within a genus not displayed on the page of the genus any more?

The new design was chosen to prepare for the future of LPSN and to create a sustainable structure for collecting and displaying an increasing amount of information over time. Having one page per taxon name throughout irrespective of the taxonomic category should make navigation through LPSN easier for users.

Individual LPSN pages for genus names used to list the full information for all their species names (and subspecies names, if any). For navigation downwards one needed to scroll through potentially a lot of text. But because of past, present and future additions over time (new species in same genus, emendations of genus or species, synonyms, notes etc.) it seemed advisable to generate more pages instead of longer pages. It also appeared to be a straightforward and consistent design to create one page per taxon name throughout. Because the pages for all individual taxon names now have the same structure irrespective of the taxonomic category it should be rather easy to learn how to navigate through LPSN. However, all the species belonging to a genus are listed under the child taxa section of the genus entry.

In addition, it is now possible to move from one species of a certain genus to another species of that genus by using the "siblings" button in the upper right corner. Having each species on its own page means more clicking but having all species of a genus on a single page would mean more scrolling. Navigation using the "parent" and "children" buttons should also be straightforward and intuitive.

Information on the hierarchical classification is available as:

- list of parent taxa, in order, visible on the category pages (Domain to Species) if you click on the arrowhead on the left side or on "Open all lineages" (unless there is no immediate parent taxon);
- list of child taxa visible on the category pages (Domain to Species) if you click on "Show [child category] list ..." or on "Open all [child category] lists" (unless there are no child taxa);
- entry for the parent taxon on each page for a single taxon name (unless there is no parent taxon);
- list of child taxa on each page for a single taxon name (unless there are no child taxa);
- navigation option "parent" in the upper right corner on each page for a single taxon name (unless there is no parent taxon);
- navigation option "siblings" in the upper right corner on each page for a single taxon name (unless there are no sibling taxa);
- navigation option "children" in the upper right corner on each page for a single taxon name (unless there are no child taxa).

See also the next entry.

#### What are "child taxa", "parent taxa", and "siblings"?

 $\mathscr{O}$ 

An in-depth explanation is provided in the LPSN glossary. In brief, the term "child taxon" means that some taxon belongs to the category below a given taxon in terms of a hierarchical classification from domain down to subspecies. So for a family, the child taxa would be its constituent genera; for a genus, the child taxa would be its constituent species, etc. Conversely, the "parent taxon" of a species would be a genus, the parent taxon of a genus would by a family, etc. The "siblings" of a species would by other species in the same genus, the siblings of a genus would be other genera in the same family, etc.

The navigation menu at the top right of each taxon entry has "parent – siblings – children" - so, for a genus "parent" would be the link to the entry of the family to which it belongs, "siblings" would be list of links to the entries for other genera in the same family, and "children" would be the list of links to the entries for the species in that genus.

### What does the formatting on the category pages mean?

6

The category pages (Domain to Species) used for browsing the hierarchical classification as well as other pages, such as the result of an advanced search, apply various text formatting options such as bold text, quotation marks, and colouring. An in-depth explanation of the meaning of the formatting is provided in the introduction to LPSN. Note that LPSN also uses substant placeholders in certain situations. These are enclosed in square brackets.

## Where can I find the geographic origin of strains?

 $\mathscr{O}$ 

The geographic origin and other source information for strains can be found in BacDive either by searching there directly or by following the BacDive link in the LPSN species entry when available ("See detailed strain information at BacDive").

#### Is it possible that LPSN genera lack species?

6

Yes, temporarily. LPSN attempts to display information as soon as possible once it has achieved a certain degree of completeness and underwent a certain number of checks. This may yield entries for genera that are not yet linked to an entry for a species, families that show no genera, orders that show no families, etc. Such entries should not be read as an indication of really missing child taxa. For analogous reasons LPSN generates stubs or placeholders for parent taxa.

The only alternative is to postpone the display of parent taxa until all expected child taxa can be displayed, and to postpone the display of child taxa until their parent taxa have been entered. This approach would create problems of its own.

### Is it possible that LPSN taxon names lack nomenclatural types?

B

Yes. Most removed nomenclatural types are missing only temporarily. LPSN attempts to display information as soon as possible once it has achieved a certain degree of completeness and underwent a certain number of checks. This may yield entries for taxon names that do not yet list a nomenclatural type. Such entries should not be read as an indication of an actually missing type. If the nomenclatural type of a validly published name is actually missing (rendering the name is illegitimate), this will be explicitly indicated. Such indications may not be made for names that are not validly published.

The only alternative is to postpone the display of taxon names until the nomenclatural type can be displayed. This approach would create problems of its own.

## Is it possible that synonyms are not mentioned?

Ø

Yes, temporarily. LPSN attempts to display information as soon as possible once it has achieved a certain degree of completeness and underwent a certain number of checks. This may yield entries for taxon names that are not yet linked to an entry for a synonym although this entry is present in LPSN. The link will usually be made in short time but may initially be missing.

The only alternative is to postpone the display of taxon names until all synonyms have been linked. This approach would create problems of its own.

## Is it possible that the year of an authority changes?

 $\mathscr{O}$ 

Unfortunately, yes. Two aspects are of relevance here: A change of the year of the effective publication and a change cause by the validation of a taxon name.

The problem of a change of the year of the effective publication is related to "online first" or other preliminary publications that do not indicate the final volume, issue, and/or page numbers. In such cases the year of the first online publication may differ from the year of the final volume. Since the final year cannot be predicted with certainty from the first online publication, LPSN initially uses the year of the first online appearance of some paper and may later on be forced to update the year. This is annoying but in the view of the LPSN maintainers it is not a problem that is caused by LPSN. Also note that PubMed behaves in the same way, as publication years given in PubMed entries can get updated for apparently the same reason.

The only alternative is to postpone the display of a taxon name until the final version of the citation gets published by the journal. This approach would create problems of its own.

Such updates of publication details may also cause changes of volume, issue or page numbers. LPSN uses 0 as placeholder for volume, issue or page numbers that are as yet missing.

For names validly published under the ICNP, LPSN provides the year of valid publication. Among other reasons, this year determines the priority of the name. If LPSN has recorded the same name as not being validly published beforehand, the year of the authority may change upon validation in IJSEM. The year of valid publication is always the year of listing the name in IJSEM, either in a Validation List or directly in an effective publication in IJSEM.

IJSEM also publishes Notification Lists. These do not determine the year or status of valid publication. Also note that even the inclusion in a Validation List or directly in an effective publication in IJSEM is only a necessary condition but not a sufficient condition for a name to be validly published.

Finally, in the case of emendations there is no distinction between validly published ones and others. Hence, the year given is always the year of original publication of the emendation.

#### Which taxon names should be written in italics?

 $\mathscr{O}$ 

This can be found in the 2022 revision of the ICNP in **Chapter 4. Advisory Notes A. Suggestions for Authors and Publishers**, which states that "It is recommended that scientific names be printed in a different font, e.g. italics, or by some other device to distinguish them from the rest of the text". See also Rule 33a Note 2. Thus, for prokaryotes, names in all categories are printed in italics. However, in the case of *Candidatus* names, only the word "Candidatus" is italicised (see Appendix 11 of the ICNP).

# Is it possible that LPSN does not write taxon names in publication titles or etymologies in italics?

S

LPSN attempts to display information as soon as possible once it has achieved a certain degree of completeness and underwent a certain number of checks. This may yield publication titles that do not yet display taxon names in italics. The formatting will usually be made in short time but may initially be missing. The only alternative is to postpone the display of publications until all included taxon names have been formatted. This approach would create problems of its own.

Names of non-bacterial taxa and incorrectly spelled names of bacterial taxa are deliberately not written in italics in publication titles. We believe the cost-benefit ratio of such extensive formatting attempts to be too high.

Names of bacterial or other taxa in explanations of etymology entries are deliberately not written in italics. Here italics is reserved for the explained words.

# Can LPSN supply strains?

ଚ

No, LPSN cannot supply strains. (Not even DSMZ strains.) However, LPSN does list type strains of taxa and where possible links directly to those strains in culture collections that have online catalogues – strains can be ordered directly from these collections.

# Is it possible that a type-strain deposit indicated on LPSN is not accessible any more?

 $\mathscr{O}$ 

LPSN primarily aims at accurately representing literature data. Alleged type-strain deposits found in the literature can later on well turn out to be not viable any more, or to not correspond to the description of their species or subspecies, or to be contaminated. Fixing such issues in a proactive way is beyond the scope of LPSN (and beyond our resources, for that matter). We are happy to include information from the according collection or from the scientific literature about the need to delete records of type-strain deposits. If you have proof for such cases, please let us know. This particularly holds if the lack of deposits puts a name in danger of not being validly published.

This issue also affects the links to web pages of deposits in culture collections. If none of the links to a certain culture collection works, they should indeed be deactivated. If single deposit numbers yield a malfunctioning link, however, this may have a variety of reasons, and deactivating the number of the deposit in LPSN may be the wrong choice. For instance, the deposit number may be listed in the literature but the collection did not yet publish the web page. Alternatively, the collection may have removed the page because it does not want the strain to be ordered although the deposit is fine. Both statuses may change at any time.

The main problem here is that the lack of a web page for a collection deposit does not indicate that the deposit is unavailable. This is unfortunate but LPSN can do little about it, even in the case of DSMZ deposits. LPSN would much prefer that all deposits in culture collections that were ever issued had a web page, even if this page just indicated that the deposit was abandoned.

Please do not send questions about the accessibility of deposits primarily to LPSN. Instead, please first contact the according collection or the authors who proposed the type strain. For instance, DSMZ provides a list of curators to contact about their deposits.

# Why are leading zeros sometimes missing from culture collection identifiers in LPSN?

B

The designator of a culture collection deposit usually comprises the acronym of the collection followed by a specific identifier (ID). One needs to determine whether the ID is intended as a number or as a character string that just happens to contain only digits. If the ID is a number, the leading zeros cannot be significant and are just a matter of formatting. In such situations the sources are often inconsistent regarding leading zeros. Hence, it makes sense to standardise the display of such IDs on LPSN by removing all leading zeros.

### Where can I find general information on Candidatus names?

S

In the LPSN so glossary. All *Candidatus* names are accessible using the advanced search while *Candidatus* phyla are compiled on a separate page.

# How can I find out which version of LPSN I am using?

B

The date of the last update of the public LPSN database can always be found on the LPSN numbers page. For the whole database this is almost as good as a version number. Multiple updates per day are very unlikely.

# Is there a list of taxonomic changes for a certain year or of the overall numbers of taxonomic proposals?

While there is no list of "taxonomic changes" as such, one can use the advanced search by selecting "validly published" and "correct name" and entering the year of interest into the authority field. The result comes pretty close to the list of "taxonomic changes" published in that year.

It is important to keep in mind that there is an official nomenclature of prokaryotes but not an official classification. The replace International Code of Nomenclature of Prokaryotes (ICNP) regulates which name must be applied to which taxon. However, the code does not regulate how prokaryotes are arranged into taxa. For the ICNP, apart from the occasional need to replace a name that contravenes a Rule, there are no name changes; there are only proposals of names. It is up to the user to select a taxonomic arrangement, and this selection then determines the name to apply.

The overall numbers of taxonomic proposals are found on the statistics page. They are also accessible as user-friendly graphics.

#### How does a name become validly published?

 $\mathscr{O}$ 

The valid publication of names of species and subspecies under the International Code of Nomenclature of Prokaryotes (ICNP) requires the deposit of pure cultures of the type strain in collections in at least two distinct countries. Authors of taxon names may not be aware that valid publication of names proposed in an effective publication outside the International Journal of Systematic and Evolutionary Microbiology (IJSEM) is not automatic. Instead, a request for validation must be sent to the IJSEM office, including the proposed name, the effective publication, and evidence for the deposit of the type strain in at least two culture collections in distinct countries.

Is there a way in the downloadable excel file to differentiate between homotypic and heterotypic synonyms?

Please see the page on LPSN download formats.

### How is LPSN governed?

6

Please see the introduction to LPSN for corresponding information.

This LPSN page was printed on 2024-05-03 17:45:42