

THE ALL SPECIES LIVING TREE PROJECT RELEASE LTP_s100, 28th September 2009

The ribosomal SSU sequence database of type strains, which is available online at <http://arb-silva.de/living-tree/>, has been upgraded with the new species validly published up to the August 2009 issue of the IJSEM journal.

The ARB database accounts for a total of 7,710 sequences belonging to various type strains and all the sequences can be found in SILVA SSURef100. Please read the summary below concerning changes and new features of the current release.

Changes in the dataset.

A total of 63 species omitted from the last release have now been included in the database. IJSEM issues 58(7) to 59(8) were tracked in order to build a list of new species with their corresponding type strain SSU sequence. Among the 742 species that have been validly published between July (2008)-August (2009), 14 species had to be omitted due to several reasons* and 41 were new combinations of species already present in the tree, and their names only needed to be updated. The remaining 701 species have been included *de novo* in the LTP database.

In total, 750 new distinct species ([table1.pdf](#)) have been incorporated into the database and tree. A manual crosscheck was undertaken between the sequence-associated information provided by SILVA and the up-to-date LPSN (<http://www.bacterio.cict.fr>). Around 50% of the sequences carried mistakes in important fields for taxonomy, such as the species name and/or the strain numbers ([table2.pdf](#)). Consequently, the corrected and/or updated species name now appears in the ARB database in the field called 'fullname_ltp'.

Later heterotypic synonyms were excluded from the database as they may generate confusion. An updated correspondence of heterotypic synonyms is reported in [table3.pdf](#).

Tree reconstruction.

The LTPs95 alignment was used as the reference for automatically aligning the new 750 sequences by using the SINA aligner (implemented in SILVA) followed by manual inspection of misplaced bases. As many as 3,240 additional sequences belonging to non-type strains were selected in order to complement the total number of 7,710 type strain sequences and balance the number of over/under-represented groups in the tree. A maximum likelihood-based reconstruction was calculated in a single run using the RAxML (Stamatakis, 2006) algorithm with all the 10,950 sequences, following the same procedures as in Yarza *et al.* (2008).

In previous LTP releases, the particular subset of non-type strain sequences, added initially to stabilize the tree, caused confusion among the user community. Consequently, they were removed from the tree and the ARB database.

Remarks on the LTP_s100 tree topology.

The current tree was compared in detail against a large collection of taxa-specific and broad-range trees made using different approaches, subsets of sequences and alignments. Since they were in strong agreement, the discrepancies found were within acceptable significance ranges.

In our tree, as in other complete trees, there is a low significance of the relative branching order of phyla, as indicated by relatively short branch lengths. Whereas some phyla may 'jump' in the topology, depending on treeing procedures, other putative 'super-phyla' always appear as highly significant and stable clusters.

Other analyses do not clearly support a monophyletic structure of all *Proteobacteria*, and a tendency exists to separate *Deltaproteobacteria* and *Epsilonproteobacteria* from their counterparts. For the *Actinobacteria*, both in our tree as in others, there is a low significance of the relative branching order at intra-class and intra-order sublevels. There is no support for a clear monophyletic structure of the orders *Kineosporiineae* and *Frankineae*. In addition, special attention needs to be paid to the phylum *Fusobacteria*, whose comparison with other trees shows that this phylum has unstable rooting and is related to *Firmicutes*. In the case of *Firmicutes*, the remote position of *Thermoanaerobacterales* is well supported by other analyses.

Features of the ARB database.

For practical reasons, we still maintain the use of an ARB color to distinguish the type from the non-type species. However, a new field called 'type_ltp' has been created and every type species of a genus now receives the label 'type sp.'.

The 10%-50% conservational filters have been updated, taking into account the current 7,710 type strain sequences included in the database.

Supplementary files.

[LTP_s100.opt.arb](#): ARB-database.

[LTP_description_of_fields.pdf](#): Description of ARB fields introduced by the Living Tree Project.

[fasta_LTPs100.eft](#): Fasta-format export filter used to extract entries from the ARB database.

[typestrains_LTPs100.fasta](#) and [typestrains_LTPs100_unaligned.fasta](#): Multifasta alignments of type strains. The header of the sequences stands for, in order, the following information: accession number, start and stop position, length, type of sequence, fullname_ltp, hi_tax_ltp.

[LTP_s100_tree.pdf](#): The complete All-Species Living Tree as a pdf file.

[LTP_s100_tree.svg](#): An overview of the tree in svg format.

[LTP_s100_tree.newick](#): The complete tree in newick format.

“Sequencing the orphans”: The Living Tree Project leads a new international collaboration.

In the course of curating our type strains database, we found that about 9% of the hitherto validly published species of Bacteria and Archaea do not have at least one ribosomal SSU sequence of enough quality for in depth phylogenetic studies. These type strains have either never been sequenced (we call them orphans, see [table4.pdf](#)) or their sequence is too bad to be included in the LTP ([table5.pdf](#)).

Consequently, nine international partners (DSMZ, CCUG, NBRC, CIP, LMG, CECT, ATCC, JCM and NCCB) have already joined us in a novel initiative for type strain sequencing. The aim, for the first time, is to have the 16S rRNA gene sequenced for all validly published species of Bacteria and Archaea which are available from these culture collections.

Acknowledgements

The team wants to acknowledge the contributions of Prof. Dr. Peter Dawyndt, Wim De Smet, Prof. Dr. Hans Peter Klenk, Dr. Pierre Nicolas and Dr. Renzo Kottmann.

Contact us

We acknowledge all the requests and comments received to date regarding any aspect of the All-Species Living Tree Project. These suggestions have either already been implemented in this new release or are currently being dealt with. The additional input from the scientific community is helping us to develop our product through the updates. Therefore, please do not hesitate to contact us at [living-tree\(at\)arb-silva.de](mailto:living-tree(at)arb-silva.de).

*The following 14 species, having been validly published up to issue 59(8) of the IJSEM, were omitted from this release due to the following reasons:

i.- Four species whose SSU sequences were submitted after the SILVA 100 deadline will be included in the next LTP release: EU740416, *Salinisphaera hydrothermalis* (Crespo-Medina *et al.*, 2009, sp. nov.); EF488759, *Enterobacter oryzae* (Peng *et al.*, 2009, sp. nov.); EU868814, *Brevibacterium massiliense* (Roux and Raoult, 2009, sp. nov.); AB445007, *Paraoerskovia marina* (Khan *et al.*, 2009, sp. nov.: type species of the genus).

ii.- Eight species only have SSU sequences below the LTP quality criteria: *Blautia hansenii* (Holdeman and Moore, 1974; Liu *et al.*, 2008, comb. nov.); *Afifella pfennigii* (Caumette *et al.*, 2007; Urdian *et al.* 2009, comb. nov.); *Actinomyces johnsonii* (Henssge *et al.*, 2009, sp. nov.); *Actinomyces oris* (Henssge *et al.*, 2009, sp. nov.); *Providencia burhodogranariea* (Juneja and Lazzaro, 2009, sp. nov.); *Providencia sneebia* (Juneja and Lazzaro, 2009, sp. nov.); *Bifidobacterium bombi* (Killer *et al.*, 2009, sp. nov.); *Natronaerobius trueperi* (Mesbah and Wiegel, 2009, sp. nov.).

iii.- Two species lack SSU entries in public repositories: *Burkholderia contaminans* (Vanlaere *et al.*, 2009, sp. nov.); *Corynebacterium ilicis* (Mandel *et al.*, 1961, species).