Available online at www.sciencedirect.com
ScienceDirect
Systematic and Applied Microbiology 31 (2008) 241–250
SYSTEMATIC AND APPLIED MICROBIOLOGY
www.elsevier.de/syapm

The All-Species Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains

Pablo Yarza[a], Michael Richter[a], Jörg Peplies[b], Jean Euzeby[c], Rudolf Amann[d], Karl-Heinz Schleifer[e], Wolfgang Ludwig[e,**], Frank Oliver Glöckner[d,f,**], Ramon Rosselló-Móra[a,*]

[a]*Marine Microbiology Group, Institut Mediterrani d'Estudis Avançats (CSIC-UIB), C/ Miquel Marqués 21, E-07190 Esporles, Illes Balears, Mallorca, Spain*
[b]*Ribocon GmbH, D-28359 Bremen, Germany*
[c]*Société de Bactériologie Systématique et Vétérinaire (SBSV) & École Nationale Vétérinaire de Toulouse (ENVT), F-31076 Toulouse Cedex 03, France*
[d]*Max Planck Institute for Marine Microbiology, D-28359 Bremen, Germany*
[e]*Lehrstuhl für Mikrobiologie, Technische Universität München, D-85350 Freising, Germany*
[f]*Jacobs University Bremen, D-28759 Bremen, Germany*

Abstract

The signing authors together with the journal Systematic and Applied Microbiology (SAM) have started an ambitious project that has been conceived to provide a useful tool especially for the scientific microbial taxonomist community. The aim of what we have called "The All-Species Living Tree" is to reconstruct a single 16S rRNA tree harboring all sequenced type strains of the hitherto classified species of *Archaea* and *Bacteria*. This tree is to be regularly updated by adding the species with validly published names that appear monthly in the Validation and Notification lists of the International Journal of Systematic and Evolutionary Microbiology. For this purpose, the SAM executive editors, together with the responsible teams of the ARB, SILVA, and LPSN projects (www.arb-home.de, www.arb-silva.de, and www.bacterio.cict.fr, respectively), have prepared a 16S rRNA database containing over 6700 sequences, each of which represents a single type strain of a classified species up to 31 December 2007. The selection of sequences had to be undertaken manually due to a high error rate in the names and information fields provided for the publicly deposited entries. In addition, from among the often occurring multiple entries for a single type strain, the best-quality sequence was selected for the project. The living tree database that SAM now provides contains corrected entries and the best-quality sequences with a manually checked alignment. The tree reconstruction has been performed by using the maximum likelihood algorithm RAxML. The tree provided in the first release is a result of the calculation of a single dataset containing 9975 single entries, 6728 corresponding to type strain gene sequences, as well as 3247 additional high-fquality sequences to give robustness to the reconstruction. Trees are dynamic structures that change on the basis of the quality and availability of the data used for their calculation.

**Also for correspondence. Max Planck Institute for Marine Microbiology, D-28359 Bremen, Germany. Tel.: +49 421 2028 970; fax: +49 421 2028 580.
*Corresponding authors. Tel: +34 971 611 826; fax: +34 971 611 761 (R. Rosselló-Móra), Tel.: +49 8161 71 5451; fax: +49 8161 71 5475 (W. Ludwig).
E-mail addresses: ludwig@mikro.biologie.tu-muenchen.de (W. Ludwig), fog@mpi-bremen.de (F.O. Glöckner), rossello-mora@uib.es (R. Rosselló-Móra).

0723-2020/$ - see front matter © 2008 Elsevier GmbH. All rights reserved.
doi:10.1016/j.syapm.2008.07.001

Therefore, the addition of new type strain sequences in further subsequent releases may help to resolve certain branching orders that appear ambiguous in this first release.

On the web sites: www.elsevier.de/syapm and www.arb-silva.de/living-tree, the All-Species Living Tree team will release a regularly updated database compatible with the ARB software environment containing the whole 16S rRNA dataset used to reconstruct "The All-Species Living Tree". As a result, the latest reconstructed phylogeny will be provided. In addition to the ARB file, a readable multi-FASTA universal sequence editor file with the complete alignment will be provided for those not using ARB. There is also a complete set of supplementary tables and figures illustrating the selection procedure and its outcome. It is expected that the All-Species Living Tree will help to improve future classification efforts by simplifying the selection of the correct type strain sequences.

For queries, information updates, remarks on the dataset or tree reconstructions shown, a contact email address has been created (living-tree@arb-silva.de). This provides an entry point for anyone from the scientific community to provide additional input for the construction and improvement of the first tree compiling all sequenced type strains of all prokaryotic species for which names had been validly published.

© 2008 Elsevier GmbH. All rights reserved.

Keywords: 16S rRNA; SSU; Phylogeny; Species; Living tree; ARB; Taxonomy; Classification; Taxa boundaries

The need for a curated all-species tree

Thirty years ago, the systematics of prokaryotes experienced an important breakthrough when attempts were made to establish the first genealogical relationships by using comparative cataloguing of the primary sequence of the small subunit (SSU) of the ribosome [8]. At that time, systematicists were already aware that the new tool for inferring genealogies would have an important impact on the way the taxonomy of prokaryotes developed [9]. However, the establishment of a phylogenetic backbone for the classification of prokaryotes has required the important task of validation for the tree topologies in comparison with other molecular clocks [19]. Nevertheless, nowadays, it is clear that the 16S rRNA gene sequence analysis applied to bacterial systematics is of paramount relevance. Nearly all descriptions of taxa are accompanied by relevant sequence information and reconstruction of their relationships based on the sequence of the SSU of the ribosome. Furthermore, it has been recommended that the inclusion of a high-quality sequence should be mandatory in the future [30]. Actually, the current overview of the classification of prokaryotes is mainly based on genealogical affiliations [11], and the circumscription of any new taxon with a higher hierarchy than species (i.e. genus and above categories) is based on genealogical relationships. The single category for which SSU sequence divergences cannot provide a sharp resolution is species [26]. In this respect, identical or nearly identical SSU sequences cannot guarantee that two organisms belong to the same species following the criteria traditionally used to define and circumscribe this category [10]. Despite the fuzziness of the resolution power of the SSU at the species level, it has been observed that, in general, two organisms with sequence divergence above a 3% nucleotide identity may not belong to the same species [1,31], and, for the same reason, lower divergences may be tested by DNA–DNA hybridization analysis. Currently, it is recommended that the hybridization is to be done when identity values are below 98.7–99% [29]. Nevertheless, SSU analysis is important for inferring monophyly [30], and this is one of the most important premises for circumscribing a prokaryotic species.

One of the main controversial issues concerning the validity of SSU gene analysis is whether this single gene really represents the genealogy of the organism that harbors it. Phenomena such as genetic crossover of ribosomal genes [27] or horizontal gene transfer (HGT, [6]) have been referred to as being responsible for blurring the validity of SSU to represent organismal genealogy. Today, whole genome comparisons provide unprecedented insights. On the one hand, and in the light of the current knowledge of the genetic content of prokaryotes, a large HGT occurrence has been hypothesized [14], whereas, on the other hand, there are severe criticisms of how data are interpreted [15]. In any case, it has been hypothesized that an organism's genome may contain a certain set of genes which would be largely excluded from HGT, and would be responsible for what an organism is and thus for its identification [16]. In general, large phylogenetic studies with different sets of housekeeping genes based on comparative genomics provide strong support for the genealogies based on SSU analysis [4,28]. Altogether, the comparisons indicate that, for classification purposes, SSU tree reconstructions may be the most parsimonious and accurate way to establish genealogical relationships.

Despite the criticisms, comparative sequence analysis of the SSU rRNA has been established as the gold standard for reconstructing phylogenetic relationships among prokaryotes for classification purposes [18]. As a consequence, the number of SSU sequences deposited in public databases has increased exponentially by about three orders of magnitude in approximately 15 years

Download English Version:

https://daneshyari.com/en/article/2063914

Download Persian Version:

https://daneshyari.com/article/2063914

Daneshyari.com