A Web of Prokaryotic Life

Stefan R. Henz,[1] Daniel H. Huson,[2] Alexander F. Auch,[2] Vincent Moulton, [3] and Stephan C. Schuster[1]

Keywords: molecular evolution, tree of life, web of life, prokaryotes, consensus methods

1   Introduction.

The evolution of species is generally believed to be a *branching process* [3] and, thus, the ultimate goal of phylogenetic analysis is to compute the "Tree of Life". However, there are a number of evolutionary mechanisms such as hybridization, recombination or swapping of genes, that may imply that evolutionary history is perhaps more accurately described by a "Web of Life" [2], that is, by a network that generalizes the concept of a phylogenetic tree [1, 7, 6].

Traditionally, in molecular phylogeny, trees are built from singular phylogenetic markers such as 16S rRNA [8]. In Figure 1(a) we show such a tree for 91 prokaryotes. More recently, given the rapidly rising abundance of whole genome sequences, it has become possible to base such phylogenies on many different molecular markers, such as different shared genes, gene-order or gene-content, for example. In Figure 1(b) we show a phylogenetic "consensus network" [5] for the same 91 prokaryotes based on their 30 most conserved genes [4]. This graph shows every split that occurs in two or more of the 30 trees. By increasing this threshold, one obtains increasingly less cluttered graphs (not shown here). This phylogenetic network gives a good indication of which parts of the phylogeny shown in Figure 1(a) are supported by the 30 genes, and which parts potentially have multiple evolutionary histories.

References

[1] H.-J. Bandelt and A.W.M. Dress. A canonical decomposition theory for metrics on a finite set. *Advances in Mathematics*, 92:47–105, 1992.

[2] F. W. Doolittle. Phylogenetic classification and the universal tree. *Science*, 284(5423):2124–2128, 1999.

[3] Ernst Haeckel. Monophyletischer Stammbaum der Organismen, 1866.

[4] S. R. Henz, D. H. Huson, A.F. Auch, and S. C. Schuster. Computing the web of prokarayotic life. In preparation, 2004.

[5] B. Holland and V. Moulton. Consensus networks: A method for visualizing incompatibilities in collections of trees. In G. Benson and R. Page, editors, *Proceedings of "Workshop on Algorithms in Bioinformatics"*, volume 2812 of *LNBI*, pages 165–176. Springer, 2003.

[6] D. H. Huson and D. Bryant. jSplits - a framework for phylogenetic trees and networks. `http://www-ab.informatik.uni-tuebingen.de/software/jsplits/`, 2004.

[7] D.H. Huson. SplitsTree: A program for analyzing and visualizing evolutionary data. *Bioinformatics*, 14(10):68–73, 1998.

[8] C. R. Woese. Bacterial evolution. *Microbiol. Rev.*, 51:221–272, 1987.

[1]Max-Planck-Institute for Developmental Biology, 72076 Tübingen, Germany.
E-mail: `stefan.henz@tuebingen.mpg.de`, `stephan.schuster@tuebingen.mpg.de`
[2]Center for Bioinformatics, Tübingen University, Sand 14, 72076 Tübingen, Germany.
E-mail: `huson@informatik.uni-tuebingen.de`, `auch@informatik.uni-tuebingen.de`
[3]The Linnaeus Centre for Bioinformatics, Uppsala University, Sweden.
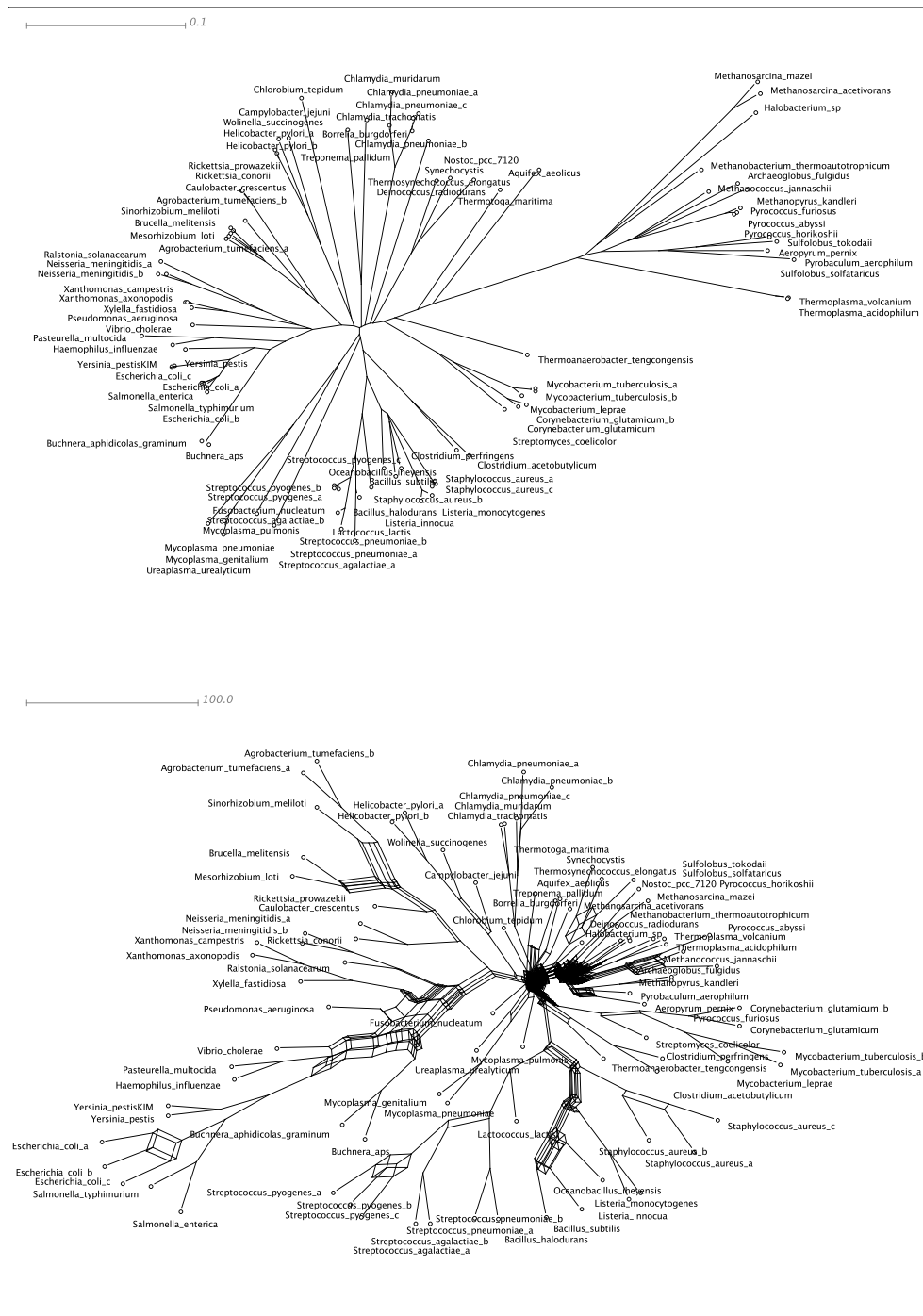E-mail: `vincent.moulton@lcb.uu.se`

Figure 1: (a) A tree of prokaryotic life based on 16S rRNA. (b) A "web" of prokaryotic life based on the 30 most conserved genes.