Taxonomic Outline of the Bacteria and Archaea

(Formerly the Taxonomic Outline of the Prokaryotes)

Release 7.7 March 6, 2007

George M. Garrity

Timothy G. Lilburn

James R. Cole

Scott H. Harrison

Jean Euzéby

Brian J. Tindall

George M. Garrity, Sc.D.
Microbiology & Molecular Genetics
Michigan State University
East Lansing, MI 48824
and NamesforLife, LLC
Okemos, MI 48864, USA

Timothy G. Lilburn, Ph.D.
American Type Culture Collection
Manasas, VA 20110, USA

James R. Cole, Ph.D.
The Ribosomal Database Project II
Center for Microbial Ecology
Michigan State University
and NamesforLife, LLC
Okemos, MI 48864, USA

Scott H. Harrison, Ph.D.
Microbiology & Molecular Genetics
Michigan State University
East Lansing, MI 48824, USA

J. Euzéby, DVM
Laboratoire de Bactériologie
École Nationale Vétérinaire
23, chemin des Capelles
B.P. 87614
31076, Toulouse cedex 03, France

Brian J. Tindall, Ph. D.
DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
Inhoffenstrasse 7B
D-38124 Braunschweig, Germany

Taxonomic Outline of the Bacteria and Archaea, Release 7.7 March 6, 2007.
George M. Garrity, Timothy G. Lilburn, James R. Cole,
Scott H. Harrison, Jean Euzéby, and Brian J. Tindall

Introduction

Following a prolonged hiatus, we have resumed our efforts to periodically distribute a comprehensive taxonomy of the *Bacteria* and *Archaea*. Release 7.7 of the Taxonomic Outline of *Bacteria* and *Archaea* (TOBA 7.7) was created specifically to provide a mapping of ongoing genome sequencing efforts onto an updated version of the taxonomy in advance of a meeting at the University of Georgia on March 8-9, 2007. The focus of that meeting was to discuss how a large-scale project could be organized that would result in sequencing the genomes of all the type strains of the species recognised by a combination of the Rules of the Bacteriological Code and current taxonomic interpretation. A second goal was to draw up a "short list" of public collections that could serve as the source of type strains within the framework of this project. The criteria for selecting only one collection per type strain were outlined before the meeting. In future releases of TOBA we will resume our practice of distributing a comprehensive list of deposits.

As with earlier releases of the taxonomy, the arrangement of taxa at the family level and above is based on the phylogeny of the 16S rRNA gene (8-11, 14, 15). Placement of named species into genera is based on the nomenclature used in the most recent description of each genus or species. This is done to produce a human-readable *vs*. machine-readable document. TOBA 7.7 differs from earlier versions in that synonyms are resolved, with placement of basonyms and earlier homotypic and heterotypic synonyms into the same genus as the corresponding new combination or later synonym. The rationale for this was to provide a more comprehensive linking of validly published names to genome projects. TOBA 7.7 also differs from earlier versions in that only a single deposited strain and corresponding 16S sequence are presented. This was done to provide an initial list of potential suppliers of type strains to the proposed large-scale sequencing effort. Two other differences in the information content of TOBA 7.7 are also noteworthy. We have reincorporated RDP identifiers (2-4, 19-21) for the auto-aligned 16S rRNA sequences we used in creating and maintaining the taxonomy. We also include *NamesforLife* name-

ids (N4Lids) to provide direct, persistent links to content provided by that project. N4Lids are suffices of digital object identifiers (DOIs) that resolve to individual *NamesforLife* information objects that contain more detailed information about the nomenclature, taxonomy, and members of higher taxa and additional strain identifiers, sequences, and other information about the type strains and higher taxa. N4Lids preceded by the "DOI:" prefix will resolve to web pages that are part of Release 6.0 of the Taxonomic Outline.

Taxonomic coverage

TOBA 7.7 provides coverage of the validly published named species and higher taxa of *Bacteria* and *Archaea* through October 1, 2006, including all those names included on Validation Lists through No. 111 (1). In addition, TOBA 7.7 contains a limited number of well known taxa of *Cyanobacteria* that were included in earlier releases (8-11, 14, 15), the myxobacterial taxa described by Reichenbach for which duplicate deposits had not been confirmed at the time of publication (22-37), and a number of provisional names of higher taxa that were used as placeholders in previous releases. A breakdown of the totals at each taxonomic rank in Releases 5.0, 6.0 (published as information objects in the *NamesforLife* prototype accessible at DOI: 10.1601/tx.0) and TOBA 7.7 are shown in Table 1.

	Release		
Rank	**5.0**	**6.0**	**7.7**
Domain	2	2	2
Phylum	26	27	27
Class	42	42	44
Subclass	5	5	5
Order	94	94	97
Suborder	10	10	17
Family	244	249	260
Genus	1253	1320	1553
Species	6747	7158	8233
Total	8423	8907	10238s

Table 1. Number of taxa included in Releases 5.0, 6.0 and 7.7 of the Taxonomic Outline.

TOBA 7.7 differs slightly from other sources of nomenclatural information (5, 16). We do not treat subspecies as a separate level within the taxonomic hierarchy as this leads to an over estimation of the true species count. Since subspecies of the type of the species are automatically created under Rule 40d of the Code of Prokaryotic Nomenclature, we regard the default condition for all species to implicitly exist as an undeclared subspecies.

Progress in 16S rRNA sequencing
Although high-quality (>1200 nts, < 1% ambiguity) 16S rRNA sequences (HQS) are routinely provided with all contemporary taxonomic proposals dealing with *Bacteria* or *Archaea*, there remain a large number of type strains for which such data were not yet available in the public repositories. In so much as the nomenclatural record reflects the number of proposals that have been validly published, independent of the type material, it overstates the true number of unique biological entities (exemplars *sensu* Namesfor-Life (12, 13)) that have been named one or more times. To accurately assess the degree of completeness of the 16S rRNA sequencing effort, one must factor in the number of exemplars that exist as unique named biological entities or as a basonyms, new combinations, or homotypic synonyms. A breakdown of this information is provided in Table 2.

	Total	HQS
Exemplars	7031	5407
Single name	5965	4592
Two names	930	714
Three names	123	93
Four names	10	7
Five names	1	1

Table 2. Number of unique biological entities (exemplars) appearing in Release 7.7 and instances of simple or complex synonymies. HQS column indicates the number of exemplars for which high quality 16S sequences are available in the public repositories .

At present, approximately 17.5% of the validly published species names of *Bacteria* and *Archaea* are synonyms. The vast majority of these names exist in simple 1:1 relationships. However, a small number (< 2%) of the type strains are known by three or more names. Based on this approach, we find that approximately 23% of

the type strains cannot be associated with a high-quality 16S rRNA sequence.

To assist in the placement of species for which HQS data is not available, we have attempted to locate surrogate sequences derived from non-type strains. A listing of the type strains for which either no sequences or only surrogate sequences are available is presented in Appendices 1 and 2.

Taxonomic coverage of genome sequencing
We have used a similar approach to determine the status of ongoing genome sequencing efforts. A review of the GOLD (17, 18), GCAT (6, 7), and Entrez genome databases reveals that 973 (703 exemplars) of the completed or ongoing projects could be mapped to validly published bacterial or archaeal names. Of those projects, only 285 genomes can be confirmed as derived from type strains. The remaining genomes are derived from non-type strains. To further examine the extent of taxonomic coverage, we also mapped those genomes into the higher taxa. These results are presented in Table 3.

Rank	Type	Non-Type	Combined
Domain	2	2	2
Phylum	26	21	27
Class	42	35	43
Order	81	66	86
Family	130	125	172
Genus	222	243	390

Table 3. Taxonomic coverage of ongoing genome sequencing efforts. Combined reflects the total coverage provided by type and non-type sequences and is not additive.

Interpretation of entries
We continue to use the American style for rendering Latinized names, in which all names appear in italics, regardless of rank. The superscript AL indicates that the name appears on the Approved Lists of Bacterial Names. The superscript VP indicates that the name was validly published in either the *International Journal of Systematic and Evolutionary Microbiology* or the *International Journal of Systematic Bacteriology.* Names appearing in double quotes have no standing in the nomenclature, even if those names have been effectively published elsewhere. A dagger (*†*) preceding a name designation indicates that the name is deprecated and that the preferred name should generally be used,

unless the reader disagrees with that author's particular taxonomy. To aid readers, deprecated names are rendered in gray text.

As noted above, basonyms (appearing in red text, preceded by "<==") and earlier homotypic synonyms (appearing in blue, preceded by "==") now appear in the same location within the taxonomy as the new combinations (appearing in green text, preceeded by "==>") and later homotypic synonyms (appearing in blue, preceded by "=="). Earlier and later heterotypic synonyms also appear in blue type but are preceded by "≈"). The rationale for this approach is to provide a taxonomy that is in better agreement with our phylogenetic models

Taxonomic names are followed by the names of the authority/authorities who proposed that name and the year in which the name was validly published. Following the taxonomic authority appears the most recent synonym (if any), the proposed source of type material for genome sequencing, the corresponding RDP and Genbank identifiers for the high quality 16S rRNA sequence used in our taxonomic models, and information pertaining to the genome sequence derived from the GOLD, GCAT, and Entrez databases. These data are followed by the name identifier for the corresponding *NamesforLife* name object.

Resolution of NamesforLife DOIs

More complete information and about taxonomic names, taxa, and exemplars is available through the *NamesforLife* project using various tools that resolve DOIs. On PCs, a plug-in is available for the Firefox browser. The tool supports direct entry of the DOI string in the navigation bar. For other browsers, end-users can use the DOI proxy server (e.g., the International DOI Foundation's server using the prefix http://dx.doi.org/ followed by the N4Lid (e.g., 10.1601/nm.5468) or the Handle proxy server (http://hdl.handle.net/ followed by the N4Lid).

Citation

Garrity, G.M., Lilburn, T.G., Cole, J.R., Harrison, S.H., Euzeby, J., and Tindall, B.J. 2007 *Taxonomic Outline of the* Bacteria *and* Archaea*, Release 7.7* Michigan State University Board of Trustees. DOI: 10.1601/TOBA7.7

References

1. 2006. Validation list no. 111. List of new names and new combinations previously effectively, but not validly, published. Int J Syst Evol Microbiol **56:**2025-2027.
2. **Cole, J. R., B. Chai, R. J. Farris, Q. Wang, S. A. Kulam, D. M. McGarrell, G. M. Garrity, and J. M. Tiedje.** 2005. The Ribosomal Database Project (RDP-II): sequences and tools for high-throughput rRNA analysis. Nucleic Acids Res **33:**D294-6.
3. **Cole, J. R., B. Chai, R. J. Farris, Q. Wang, A. S. Kulam-Syed-Mohideen, D. M. McGarrell, A. M. Bandela, E. Cardenas, G. M. Garrity, and J. M. Tiedje.** 2006. The ribosomal database project (RDP-II): introducing myRDP space and quality controlled public data. Nucleic Acids Res.
4. **Cole, J. R., B. Chai, T. L. Marsh, R. J. Farris, Q. Wang, S. A. Kulam, S. Chandra, D. M. McGarrell, T. M. Schmidt, G. M. Garrity, and J. M. Tiedje.** 2003. The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. Nucleic Acids Res **31:**442-3.
5. **Euzéby, J.** 1998-2007. List of prokaryotic names with standing in nomenclature. Société de Bactériologie Systématique et Vétérinaire.
6. **Field, D., G. M. Garrity, N. Morrison, J. Selengut, P. Sterk, T. Tatusov, and N. Thomson.** 2005. eGenomics: Cataloguing Our Complete Genome Collection. Comp. Funct. Genom. **6:**363-368.
7. **Field, D., N. Morrison, J. Selengut, and P. Sterk.** 2006. Meeting report: eGenomics: Cataloguing our Complete Genome Collection II. Omics **10:**100-4.
8. **Garrity, G. M., J. Bell, and T. G. Lilburn** 2003, posting date. Taxonomic Outline of the Procaryotes, Bergey's Manual of Systematic

Bacteriology, Second Edition, Release 4.0. October, 2003. Springer-Verlag. [Online.]

9. **Garrity, G. M., J. Bell, and T. G. Lilburn** 2004, posting date. Taxonomic Outline of the Procaryotes, Bergey's Manual of Systematic Bacteriology, Second Edition, Release 5.0. May, 2004. Springer-Verlag. [Online.]

10. **Garrity, G. M., K. L. Johnson, J. Bell, and D. B. Searles** 2002, posting date. Taxonomic Outline of the Procaryotes, Bergey's Manual of Systematic Bacteriology, Second Edition, Release 3.0. July, 2002. Springer-Verlag. [Online.]

11. **Garrity, G. M., J. Leventhal, and T. G. Lilburn** 2005, posting date. Taxonomic Outline of the Procaryotes, Bergey's Manual of Systematic Bacteriology, Second Edition, Release 6.0. May, 2005. Springer. [Online.]

12. **Garrity, G. M., and C. Lyons.** 2003. Future-proofing biological nomenclature. OMICS **7:**31-31.

13. **Garrity, G. M., and C. Lyons.** 04, August 2005 2005. Systems and methods for resolving ambiguity between names and entities.

14. **Garrity, G. M., M. Winters, A. W. Kuo, and D. B. Searles** 2002, posting date. Taxonomic Outline of the Procaryotes, Bergey's Manual of Systematic Bacteriology, Second Edition, Release 2.0. January, 2002. Springer-Verlag. [Online.]

15. **Garrity, G. M., M. Winters, and D. B. Searles** 2001, posting date. A Taxonomic Outline of the Procaryotes, Bergey's Manual of Systematic Bacteriology, Second Edition. Springer-Verlag, NY Release 1.0. [Online.]

16. **Kracht, M., and B. J. Tindall.** 2007. Bacterial nomenclature up-to-date. Deutsche Sammlung vor Mikrooganismen and Zellkulturen GmbH.

17. **Kyrpides, N. C.** 1999. Genomes OnLine Database (GOLD 1.0): a monitor of complete and ongoing genome projects world-wide. Bioinformatics **15:**773-4.

18. **Liolios, K., N. Tavernarakis, P. Hugenholtz, and N. C. Kyrpides.** 2006. The Genomes On Line Database (GOLD) v.2: a monitor of genome projects worldwide. Nucleic Acids Res **34:**D332-4.

19. **Maidak, B. L., J. R. Cole, T. G. Lilburn, C. T. Parker, Jr., P. R. Saxman, R. J. Farris, G. M. Garrity, G. J. Olsen, T. M. Schmidt, and J. M. Tiedje.** 2001. The RDP-II (Ribosomal Database Project). Nucleic Acids Res **29:**173-4.

20. **Maidak, B. L., J. R. Cole, T. G. Lilburn, C. T. Parker, Jr., P. R. Saxman, J. M. Stredwick, G. M. Garrity, B. Li, G. J. Olsen, S. Pramanik, T. M. Schmidt, and J. M. Tiedje.** 2000. The RDP (Ribosomal Database Project) continues. Nucleic Acids Res **28:**173-4.

21. **Maidak, B. L., J. R. Cole, C. T. Parker, Jr., G. M. Garrity, N. Larsen, B. Li, T. G. Lilburn, M. J. McCaughey, G. J. Olsen, R.**

Overbeek, S. Pramanik, T. M. Schmidt, J. M. Tiedje, and C. R. Woese. 1999. A new version of the RDP (Ribosomal Database Project). Nucleic Acids Res **27:**171-3.

22. **Reichenbach, H.** 2005. Genus I. *Cystobacter* Schroeter 1886, 170[AL], p. 1086-1096. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-, Beta-, Delta-,* and *Epsilonproteobacteria*. Springer, New York.

23. **Reichenbach, H.** 2005. Genus I. *Kofleria* gen. nov., p. 1143-1144. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-, Beta-, Delta-,* and *Epsilonproteobacteria*. Springer, New York.

24. **Reichenbach, H.** 2005. Genus I. *Myxococcus* Thaxter 1892, 403[AL], p. 1074-1079. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-, Beta-, Delta-,* and *Epsilonproteobacteria*. Springer, New York.

25. **Reichenbach, H.** 2005. Genus I. *Nannocystis* Reichenbach 1970, 137[AL], p. 1137-1142. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-, Beta-, Delta-,* and *Epsilonproteobacteria*. Springer, New York.

26. **Reichenbach, H.** 2005. Genus I. *Polyangium* Link 1809, 42[AL], p. 1110-1118. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-, Beta-, Delta-,* and *Epsilonproteobacteria*. Springer, New York.

27. **Reichenbach, H.** 2005. Genus II. *Archangium* Jahn 1924, 66[AL], p. 1097-1099. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-, Beta-, Delta-,* and *Epsilonproteobacteria*. Springer, New York.

28. **Reichenbach, H.** 2005. Genus II. *Byssophaga* gen. nov., p. 1118-1121. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-, Beta-, Delta-,* and *Epsilonproteobacteria*. Springer, New York.

29. **Reichenbach, H.** 2005. Genus II. *Corallococcus* gen. nov. (Chondrococcus Jahn 1924, 85), p. 1079-1082. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-, Beta-, Delta-,* and *Epsilonproteobacteria*. Springer, New York.

30. **Reichenbach, H.** 2005. Genus III. *Chondromyces* Berkeley and Curtis in Berkeley 1874, 64[AL], p. 1121-1129. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of System-

atic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*. Springer, New York.

31. **Reichenbach, H.** 2005. Genus III. *Hyalangium* gen. nov., p. 1099-1101. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*. Springer, New York.

32. **Reichenbach, H.** 2005. Genus III. *Pyxicoccus* gen. nov., p. 1083-1084. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*. Springer, New York.

33. **Reichenbach, H.** 2005. Genus IV. *Haploangium* (ex Peterson 1959) gen.nov., nom. rev., p. 1129-1130. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*. Springer, New York.

34. **Reichenbach, H.** 2005. Genus IV. *Melittangium* Jahn 1924, 7[AL], p. 1101-1104. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-*, *Beta-*,

Delta-, and *Epsilonproteobacteria*. Springer, New York.

35. **Reichenbach, H.** 2005. Genus V. *Jahnia* gen. nov., p. 1130-1131. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*. Springer, New York.

36. **Reichenbach, H.** 2005. Genus V. *Stigmatella* Berkeley and Curtis *in* Berkeley 1857, 313, p. 1104-1108. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*. Springer, New York.

37. **Reichenbach, H.** 2005. Genus VI. *Sorangium* (Jahn 1924) gen. nov., nom. rev., p. 1132-1136. *In* Brenner, Krieg, Staley, and Garrity (ed.), Bergey's Manual of Systematic Bacteriology, vol. 2, The *Proteobacteria*, Part C, *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*. Springer, New York.

38. **Sneath, P. H. A.** 1992. Code of Nomenclature of Bacteria (1990 Revision). American Society for Microbiology, Washington, D.C.