

## Definitive Assignment by Multigenome Analysis of the Gammaproteobacterial Genus *Thermithiobacillus* to the Class *Acidithiobacillia*

COREY M. HUDSON<sup>1</sup>, KELLY P. WILLIAMS<sup>1</sup> and DONOVAN P. KELLY<sup>2,\*</sup>

<sup>1</sup>Systems Biology Department, Sandia National Laboratories, Livermore, CA 94551, USA

<sup>2</sup>School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK

Submitted 13 March 2014, accepted 14 April 2014

### Abstract

The class *Acidithiobacillia* was established using multiprotein phylogenetic analysis of all the available genomes of the genus *Acidithiobacillus* (comprising Family I, the *Acidithiobacillaceae*, of the *Acidithiobacillales*, the order created for *Bergey's Manual of Systematic Bacteriology*), and for representative genomes of all available bacterial orders. The *Acidithiobacillales* contain a second family, the *Thermithiobacillaceae*, represented by *Thermithiobacillus tepidarius*, and created on the basis of nearest neighbour 16S ribosomal RNA gene sequence similarities. This could not be included in the original multiprotein analysis as no genome sequence for *Thermithiobacillus* was available. The publication of the genome sequence of *Thermithiobacillus tepidarius* in 2013 has enabled phylogenetic assessment of this organism by comparative multigenome analysis. This shows definitively that *Thermithiobacillus* is a member of the class *Acidithiobacillia*, distinct from the *Acidithiobacillus* genus, and confirms it to represent a second family within the *Acidithiobacillia*.

**Key words:** *Thermithiobacillus*, *Acidithiobacillus*, *Acidithiobacillia*, multigenome phylogeny; multiprotein analysis

*Thermithiobacillus tepidarius* is an obligately aerobic, moderately thermophilic, neutrophilic, obligately chemolithotrophic proteobacterium, initially classified as a strain of the betaproteobacterial genus *Thiobacillus* on the basis of its morphology and biochemistry (Wood and Kelly, 1985, 1986). It was subsequently redefined as the type species of the genus *Thermithiobacillus* (Kelly and Wood, 2000), and the type (and only) genus of the family *Thermithiobacillaceae*. This family was created using the proteobacterial phylogeny based on 16S ribosomal RNA gene sequencing employed for the 2<sup>nd</sup> edition of *Bergey's Manual of Systematic Bacteriology* (Garrity, Bell and Lilburn, 2005; Kelly and Wood, 2014a). It became Family II of the order *Acidithiobacillales* in the *Gammaproteobacteria* (Garrity, Bell and Lilburn, 2005). The earliest study of 16S rRNA relationships among sulfur-oxidizing chemolithotrophs, using a partial sequence for *Thermithiobacillus tepidarius*, placed it midway between *Thiobacillus neapolitanus* and the extremely acidiphilic *Thiobacillus thiooxidans* in terms of evolutionary distances (Lane *et al.*, 1992)). The adoption of 16S rRNA gene sequence analysis as a major phylogenetic tool led to the reclassification of numerous species of *Thiobacillus* into a range of new genera, including *Thermithiobacillus*, or assignment to existing

genera such as *Thiomonas*, *Paracoccus* and *Acidiphilium* (Kelly and Wood, 2000, 2005, 2014a, b; Kelly, Wood and Stackebrandt, 2005). While *T. tepidarius* was physiologically more like neutrophilic species of *Thiobacillus*, *Halothiobacillus*, and *Thiomicrospira*, or even the thermophile *Thermothrix*, its 16S rRNA gene sequence shared only 85% identity with that of *Thiobacillus thioparus*, 82–85% with *Halothiobacillus neapolitanus*, *H. halophilus*, and *H. kellyi*, 79% with *Thiomicrospira crunogena*, and 74–79% with *Thermothrix thiopara* and *Tx. azorensis*, unequivocally excluding it from those genera. The least distant phylogenetic relatives of *Thermithiobacillus tepidarius* are *Acidithiobacillus* species (comprising Family I, the *Acidithiobacillaceae*, of the *Acidithiobacillales*), with 90–91% 16S rRNA gene sequence identity to *At. thiooxidans* and *At. ferrooxidans*. On the basis of nearest-neighbour 16S rRNA relationship, the parent order for *Thermithiobacillus* was the *Acidithiobacillales*, but it differed from these acidophiles in being thermophilic and neutrophilic, with only *Acidithiobacillus caldus* also being moderately thermophilic (Kelly and Wood, 2014b). It is not uncommon to find differences among 16S rRNA gene sequences of species of the same genus or family: for example, the range among *Halothiobacillus* species is 91–99% identity, but

\* Corresponding author: D.P. Kelly, School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK; e-mail: D.P.Kelly@warwick.ac.uk

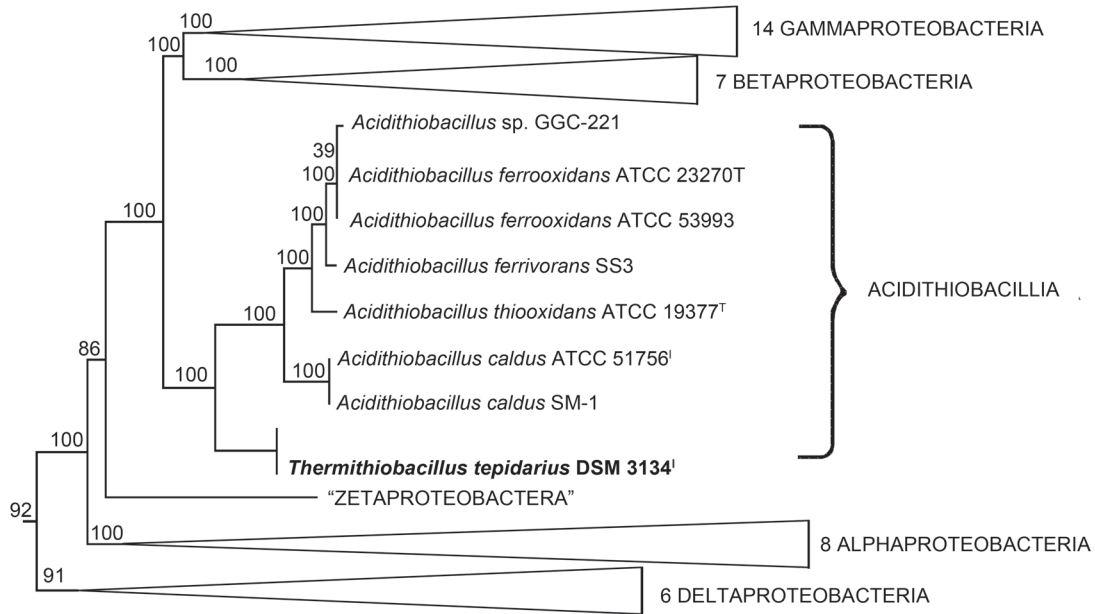


Fig. 1. Phylogenetic placement of *Thermithiobacillus* among the *Acidithiobacillia*. This portion of the full genome-scale maximum-likelihood analysis of 98 protein families, including each available member of the *Acidithiobacillia*, depicts several additional classes in collapsed form, using representative (“diagnostic”) genera of *Alpha*-, *Beta*-, and *Gammaproteobacteria*, and roots this portion of the tree with the remaining bacteria and archaeal taxa, as shown in the full tree previously presented for the *Acidithiobacillus* species only (Williams and Kelly, 2013). Details of the 98 protein families and 135 genomes used in the analyses were listed previously (Williams and Kelly, 2013). Bootstrap support percentages are shown.

accompanied by significant physiological similarities. The 16S rRNA similarity of *Thermithiobacillus* to, but physiological divergence from *Acidithiobacillus* thus supported its retention as a separate family within the *Acidithiobacillales*. Clearly a more robust confirmation of this taxonomic relationship was desirable.

The *Acidithiobacillales* were classed as *Gammaproteobacteria* until genome-scale phylogenetic analyses of *Acidithiobacillus* showed that this genus could not be assigned to any previously defined class (Williams *et al.*, 2010; Williams and Kelly, 2013). Our expanded multigenome study, using all the available genomes of *Acidithiobacillus* species, led to the creation of a seventh proteobacterial class, the *Acidithiobacillia* (Williams and Kelly, 2013). The initial study trialled 1891 prokaryotic genomes, including 104 gammaproteobacterial genomes: 356 protein families were analyzed and compared with those encoded by alpha- and betaproteobacterial genomes, resulting in the exclusion of *Acidithiobacillus* from the *Gammaproteobacteria* (Williams *et al.*, 2010). A study of the ubiquitous ribosomal proteins of 995 bacteria by Yutin *et al.* (2012) also placed *At. ferrooxidans* outside 251 species of *Gamma*-, 75 species of *Beta*-, and 180 species of *Alpha*-, *Delta*-, and *Epsilon*-proteobacteria, and to be located in the root of the *Beta*- and *Gammaproteobacteria*. When the class *Acidithiobacillia* was created, no genome sequence was available for *Thermithiobacillus*, so its membership of the class could not be definitively tested. The genome sequence of

*Thermithiobacillus tepidarius* DSM 3134 became available from the Joint Genome Institute in 2013 ([http://img.jgi.doe.gov/cgi-bin/w/main.cgi?section=TaxonDetail&page=taxonDetail&taxon\\_oid=2523533554](http://img.jgi.doe.gov/cgi-bin/w/main.cgi?section=TaxonDetail&page=taxonDetail&taxon_oid=2523533554)), and shown to contain 2.96 Mb, which is within the range of 2.83–3.02 Mb for *Acidithiobacillus* species (Kelly and Wood, 2014b). This has now enabled us to include *Thermithiobacillus* in the multigenome dataset that was used to create the *Acidithiobacillia* class.

The protocol for the present multigenome study was as described previously (Williams *et al.*, 2010; Williams and Kelly, 2013), and began with the “phase 2” dataset used in the demonstration of the *Acidithiobacillia* (Williams and Kelly, 2013). This dataset comprised 11,949 protein sequences from 98 broadly-distributed protein families, from 135 taxa, including six *Archaea*, seven *Acidithiobacillus* strains, and 122 diverse additional bacteria (Williams and Kelly, 2013). For each of the 98 protein families, a *Thermithiobacillus tepidarius* member was identified using the hidden Markov model approach implemented in HMMER v. 3.0 (Finn, Clements and Eddy, 2011). The protein family sequences were freshly aligned, masking ambiguously-aligned regions, and concatenated into a partitioned amino-acid sequence supermatrix that was subjected to a maximum-likelihood tree search with 1000 bootstraps as described previously (Williams *et al.*, 2010; Williams and Kelly, 2013). This analysis (Fig. 1) unambiguously placed *Thermithiobacillus tepidarius* into the *Acidithiobacillia*, distinct from

the *Acidithiobacillus* strains, and recapitulated the previous placement of the class, as a sister to the combined *Gammaproteobacteria* and *Betaproteobacteria*, and subtended by the ‘*Zetaproteobacteria*’ (Williams and Kelly, 2013). We conclude that *Thermithiobacillus* is a valid member of the class *Acidithiobacillia*, and should be retained in the class as the type genus of the family *Thermithiobacillaceae*. From the data currently available, we do not propose further division of the class, although the *Acidithiobacillus* and *Thermithiobacillus* species might in due course justify separation at the order level. Moreover, the confirmation of the taxonomic location of *Thermithiobacillus* by this multigenome analysis illustrates both the reliability of 16S rRNA phylogeny as a diagnostic tool, and the power of the multigenome analysis procedure at the genus, family and class levels. The multigenome method could become a useful new tool, especially for discrimination among families, orders, and classes of prokaryotes. This would require the eventual establishment of a public database of multiprotein sequences and computational tools.

Currently, there is only one established species of *Thermithiobacillus* (DSM 3134), but strain NCIMB 8349 (isolated as *Thiobacillus thioparus* strain ParkerM) is also a strain of *Thermithiobacillus* (Boden *et al.*, 2012; Kelly and Wood, 2014a), showing 99.5% 16S rRNA gene sequence identity to that of the type strain (accession numbers HM173631 and AJ459801): the genome of this strain is currently being sequenced (R. Boden, personal communication). Three other strains have been reported from Korea (a moderate thermophile from mine wastewater; Chang *et al.*, 1997) and China (from a sewage treatment plant; K. Gao, H.L. Yang and S.S. Feng, unpublished data). The 16S rRNA gene sequences of these strains (AF02364, KC478673, KC493561) show 98.3–99.4% identity to that of the type strain and 99–100% identity to that of NCIMB 8349, suggesting that the family represented by this genus is widely distributed and may contain several species.

#### Acknowledgements

Work at Sandia was supported by a Laboratory Directed Research and Development award. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly-owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-AC04-94AL85000.

We thank Ann Wood (King’s College London), Rich Boden (Plymouth University), and Nikos Kyrpides (DOE Joint Genome Institute) for advice about *Thermithiobacillus*.

#### Literature

- Boden R., D. Cleland, P.N. Green, Y. Katayama, Y. Uchino, J.C. Murrell and D.P. Kelly.** 2012. Phylogenetic assessment of culture collection strains of *Thiobacillus thioparus*, and definitive 16S rRNA gene sequences for *T. thioparus*, *T. denitrificans*, and *Halotheiobacillus neapolitanus*. *Arch. Microbiol.* 194: 187–195.
- Chang S.Y., J.S. Yoon, Y.K. Shin, Y.H. Park, J.Y. Park, S.S. Yang, M-J. Koh, S.M. Yoon, J.S. Lee, I.H. Lee and S.Y. Kim.** 1997. Isolation, characterization, and phylogenetic position of a new sulfur-oxidizing bacterium. *J. Microbiol. (Korea)* 35: 165–171.
- Finn R.D., J. Clements and S.R. Eddy.** 2011. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res.* 39: W29–37.
- Garrity G.M., J.A. Bell and T. Lilburn.** 2005. Order II. *Acidithiobacillales* ord. nov. p. 60. In: Bergey’s Manual of Systematic Bacteriology, vol. 2B, 2<sup>nd</sup> edition. Brenner, D.J., N.R. Krieg, J.T. Staley and G.M. Garrity (eds). New York, Springer.
- Kelly D.P. and A.P. Wood.** 2000. Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halotheiobacillus* gen. nov. and *Thermithiobacillus* gen. nov. *Int. J. Syst. Evol. Microbiol.* 50: 511–516.
- Kelly D.P. and A.P. Wood** (2005) Genus I. *Thermithiobacillus* Kelly and Wood. pp. 62–63. In: Bergey’s Manual of Systematic Bacteriology, vol. 2B, 2<sup>nd</sup> edition. Brenner D.J., N.R. Krieg, J.T. Staley and G.M. Garrity (eds). New York, Springer.
- Kelly D.P. and A.P. Wood** (2014a) The family *Thermithiobacillaceae*. In: The Prokaryotes, vol. 9, 4<sup>th</sup> edition. Stackebrandt E., E. Rosenberg, E. Delong, S. Lory and F. Thompson (eds). New York and Heidelberg, Springer (in press).
- Kelly D.P. and A.P. Wood** (2014b) The family *Acidithiobacillaceae*. In: The Prokaryotes, vol. 9, 4<sup>th</sup> edition. Stackebrandt, E., E. Rosenberg, E. Delong, S. Lory and F. Thompson (eds). New York and Heidelberg, Springer (in press).
- Kelly D.P., A.P. Wood and E. Stackebrandt.** 2005. Genus II. *Thiobacillus* Beijerinck 1904. pp. 764–769. In: Bergey’s Manual of Systematic Bacteriology, vol. 2C, 2<sup>nd</sup> edition. Brenner, D.J., N.R. Krieg, J.T. Staley and G.M. Garrity (eds). New York, Springer.
- Lane D.J., A.P. Harrison, D. Stahl, B. Pace, S.J. Giovannoni, G.J. Olsen and N.R. Pace.** 1992. Evolutionary relationships among sulfur- and iron-oxidizing eubacteria. *J. Bacteriol.* 174: 269–278.
- Williams K.P. and D.P. Kelly.** 2013. Proposal for a new class within the phylum *Proteobacteria*, *Acidithiobacillia* classis nov., with the type order *Acidithiobacillales*, and emended description of the class *Gammaproteobacteria*. *Int. J. Syst. Evol. Microbiol.* 63: 2901–2906.
- Williams K.P., J.J. Gillespie, B.W. Sobral, E.K. Nordberg, E.E. Snyder, J.M. Shallom and A.W. Dickerman.** 2010. Phylogeny of Gammaproteobacteria. *J. Bacteriol.* 192: 2305–2314.
- Wood A.P. and D.P. Kelly.** 1985. Physiological characteristics of a new thermophilic obligately chemolithotrophic *Thiobacillus* species, *Thiobacillus tepidarius*. *Int. J. Syst. Bacteriol.* 35: 434–437.
- Wood A.P. and D.P. Kelly.** 1986. Chemolithotrophic metabolism of the newly-isolated moderately thermophilic, obligately autotrophic *Thiobacillus tepidarius*. *Arch. Microbiol.* 144: 71–77.
- Yutin N., P. Puigbò, E.V. Koonin and Y.I. Wolf.** 2012. Phylogenomics of prokaryotic ribosomal proteins. *PLoS ONE* 7(5) e36972. (doi:10.1371/journal.pone.0036972).

