# Genomic analysis of *Hyphomonas neptunium* contradicts 16S rRNA gene-based phylogenetic analysis: implications for the taxonomy of the orders '*Rhodobacterales*' and *Caulobacterales*

Jonathan H. Badger, Jonathan A. Eisen and Naomi L. Ward

The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA

Correspondence
Jonathan H. Badger
jbadger@tigr.org

*Hyphomonas neptunium* is a marine prosthecate α-proteobacterium currently classified as a member of the order '*Rhodobacterales*'. Although this classification is supported by 16S rRNA gene sequence phylogeny, 23S rRNA gene sequence analysis, concatenated ribosomal proteins, HSP70 and EF-Tu phylogenies all support classifying *Hyphomonas neptunium* as a member of the *Caulobacterales* instead. The possible reasons why the 16S rRNA gene sequence gives conflicting results in this case are also discussed.

## INTRODUCTION

*Hyphomonas neptunium* is a prosthecate (having an appendage or 'stalk') α-proteobacterium that was isolated from sea water from the harbour at Barcelona, Spain, and was originally described as *Hyphomicrobium neptunium* (Liefson, 1964). This description was later emended to the current *Hyphomonas neptunium* on the basis of DNA–DNA hybridization information (Moore *et al.*, 1984), which showed a closer relationship with *Hyphomonas polymorpha* (Pongratz, 1957), a marine prosthecate bacterium isolated from a diver with a severe sinus infection, than with other members of the genus *Hyphomicrobium*. *Hyphomonas neptunium* also lacks, as does *Hyphomonas polymorpha*, the ability to utilize $C_1$ molecules as carbon sources, whereas recognized members of *Hyphomicrobium* have this ability (Moore *et al.*, 1984).

Members of *Hyphomonas* have an unusual reproductive cycle for prosthecate bacteria; daughter cells are formed on the distal side of the stalk, indicating that DNA, proteins and other cellular components must traverse the stalk (Hirsch, 1974). This trait is shared with numerous marine bacteria originally classified as members of the genus *Caulobacter*, and the closer relationship between these caulobacters and *Hyphomonas* to the exclusion of the freshwater caulobacters is also supported by 16S rRNA gene sequence phylogeny (Strömpl *et al.*, 2003; Abraham *et al.*, 1999; Stahl *et al.*, 1992). However, to our knowledge, there have been no studies suggesting a close relationship between freshwater members of *Caulobacter* (such as *Caulobacter crescentus* CB15) and *Hyphomonas*. Currently, *Hyphomonas* is classified as a member of the order '*Rhodobacterales*' (Garrity *et al.*, 2005), whereas the caulobacters are considered members of the eponymous order *Caulobacterales* (Henrici & Johnson, 1935). In this paper we show that, although 16S rRNA gene sequence analysis supports the current classification, phylogenies based on other markers, such as the 23S rRNA gene and many protein sequences, support grouping *Hyphomonas* as a member of the *Caulobacterales*. The implications for the taxonomy of the '*Rhodobacterales*' and *Caulobacterales* are discussed, as recent taxonomic recommendations (Stackebrandt *et al.*, 2002) support taking into account phylogenetic analyses from multiple genes.

Published online ahead of print on 3 December 2004 as DOI 10.1099/ijs.0.63510-0.

Abbreviations: JGI, Joint Genome Institute; TIGR, The Institute for Genomic Research.

Newick tree files and FASTA-format sequence alignments used to generate the trees are available as supplementary information in IJSEM Online.

## METHODS

**Data.** The complete genome sequence of *Hyphomonas neptunium* ATCC 15444$^T$, comprising a single circular chromosome of 3 705 611 nt (J. H. Badger and others, unpublished), was sequenced by The Institute for Genomic Research (TIGR) by means of the whole genome shotgun method (Fleischmann *et al.*, 1995). Gene predictions were provided by GLIMMER (Delcher *et al.*, 1999) and functional assignments were produced according to Tettelin *et al.* (2001). The following complete (or nearly complete) genomes of α-proteobacteria were used as sources of sequences for phylogenetic analyses: *Agrobacterium tumefaciens* C58 (Wood *et al.*, 2001), *Anaplasma phagocytophilum* HZ (TIGR, unpublished), *Bradyrhizobium japonicum* USDA 110 (Kaneko *et al.*, 2002), *Brucella suis* 1330 (Paulsen *et al.*, 2002), *C. crescentus* CB15 (Nierman *et al.*, 2001), *Ehrlichia chaffeensis* Arkansas$^T$ (TIGR, unpublished), *Mesorhizobium loti* MAFF303099 (Kaneko *et al.*, 2000), *Neorickettsia sennetsu* Miyayama (TIGR, unpublished), *Novosphingobium aromaticivorans* DSM 12444$^T$ [Joint Genome Institute (JGI), unpublished], *Rhodobacter capsulatus* SB1003 (Integrated Genomics, unpublished), *Rhodopseudomonas palustris* CGA009 (Larimer *et al.*, 2004), *Rhodospirillum*

*rubrum* ATCC 11170$^T$ (JGI, unpublished), *Rickettsia conorii* Malish 7$^T$ (Ogata *et al.*, 2001), *Silicibacter pomeroyi* DSS-3$^T$ (Moran *et al.*, 2004), *Sinorhizobium meliloti* 1021 (Capela *et al.*, 2001) and *Wolbachia pipientis w*Mel (Wu *et al.*, 2004). Additionally, the genome of *Escherichia coli* K-12 MG1655 (Blattner *et al.*, 1997) was used as a source of outgroup sequences. The data from the published genomes were obtained from GenBank; the unpublished data can be obtained from TIGR (http://www.tigr.org/tdb/mdb/mdbinprogress.html), JGI (http://genome.jgi-psf.org/microbial/) and Integrated Genomics (http://ergo.integratedgenomics.com/R_capsulatus.html).

**Phylogenetic analysis.** Five multiple sequence alignments (see supplementary information available in IJSEM Online) were created for the purpose of phylogenetic inference. These alignments were of: (i) the 16S rRNA gene sequence, (ii) the 23S rRNA gene sequence, (iii) 30 concatenated ribosomal proteins (totalling approximately 4000 amino acids), (iv) HSP70 proteins and (v) EF-Tu proteins. The rRNA sequences were aligned and masked using the ALIGN sequence tool of the Ribosomal Database Project (Cole *et al.*, 2003), and the protein sequences were aligned using MUSCLE (Edgar, 2004). For all the alignments, bootstrapped neighbour-joining (Saitou & Nei, 1987) trees were created using the program QUICKTREE (Howe *et al.*, 2002). For the rRNA alignments, bootstrapped maximum-likelihood (Felsenstein, 1981) trees were created using the DNAML program from PHYLIP 3.6b (Felsenstein, 2004), with a Γ-distribution ($\alpha = 0{\cdot}5$) of rates over four categories of variable sites. For the protein alignments, PROML (also from PHYLIP 3.6b) was used to create maximum-likelihood trees, applying the JTT (Jones *et al.*, 1992) model of substitution, again with a Γ-distribution ($\alpha = 0{\cdot}5$) of rates over four categories of variable sites. The resulting consensus trees for the protein and rRNA trees were fed into the appropriate program (PROML or DNAML) as user trees in order to obtain the branch lengths. In addition, APIS (J. H. Badger, unpublished), an automated pipeline for phylogenetic inference, was run on all predicted proteins in the *Hyphomonas neptunium* genome, generating bootstrapped neighbour-joining trees of each protein and its homologues.

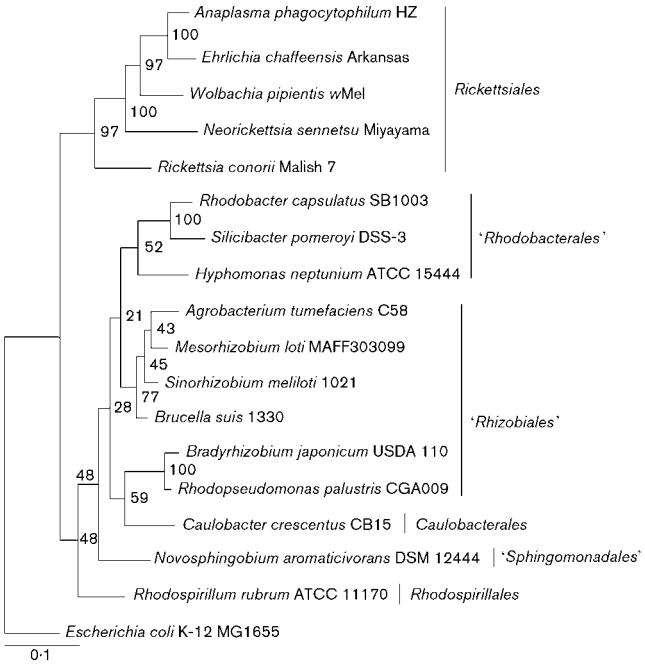


**Fig. 1.** Maximum-likelihood tree based on 16S rRNA gene sequences from sequenced α-proteobacteria. The node labels are bootstrap values (100 replicates). Note the grouping of *Hyphomonas neptunium* among the '*Rhodobacterales*'. See Table 1 for the GenBank GI numbers and ranges used from published genomes.

## RESULTS AND DISCUSSION

### Results of phylogenetic analysis

Maximum-likelihood analysis of the 16S rRNA gene sequences (Fig. 1; see Table 1 for the GenBank GI numbers and ranges used from published genomes) supports the current classification of *Hyphomonas neptunium* as a member of the order '*Rhodobacterales*', and indeed a similar analysis was probably the reason behind this classification. However, none of the other commonly used phylogenetic markers, including the 23S rRNA gene sequence (Fig. 2a), concatenated ribosomal proteins (Fig. 2b), HSP70 proteins

**Table 1.** GenBank GI numbers and sequence ranges (if applicable) from published genomes used in this study

Unpublished genome data were also used for other organisms not listed here.

| Organism | rRNA genes | | | HSP70 | EF-Tu |
|---|---|---|---|---|---|
| | GenBank GI no. | 16S rRNA range | 23S rRNA range | GenBank GI no. | GenBank GI no. |
| *Agrobacterium tumefaciens* | 17936711 | 1304386–1305691 | 1307300–1309747 | 17934041 | 17935838 |
| *Bradyrhizobium japonicum* | 27375111 | 1528226–1529715 | 1530524–1533397 | 27375790 | 27380513 |
| *Brucella suis* | 23499767 | 1108162–1109615 | 1587832–1584925 | 23502973 | 17987025 |
| *Caulobacter crescentus* | 16124256 | 3770203–3771641 | 3766708–3769496 | 16124266 | 16125489 |
| *Escherichia coli* | 49175990 | 4033554–4035095 | 4035542–4038446 | 26245936 | 26249935 |
| *Mesorhizobium loti* | 13470324 | 2758991–2757518 | 2756598–2753751 | 13473986 | 13470532 |
| *Rickettsia conorii* | 15891923 | 884601–886108 | 281797–284557 | 15892156 | 15892931 |
| *Rhodopseudomonas palustris* | 39933080 | 5249983–5251464 | 5246346–5249235 | 39933410 | 39936346 |
| *Silicibacter pomeroyi* | 56694928 | 261989–263268 | 264483–267129 | 56676708 | 56680057 |
| *Sinorhizobium meliloti* | 15963753 | 81767–83250 | 84406–87280 | 15963935 | 15965107 |
| *Wolbachia pipientis* | 42519920 | 1167943–1169389 | 182428–185173 | 42520750 | 42519935 |

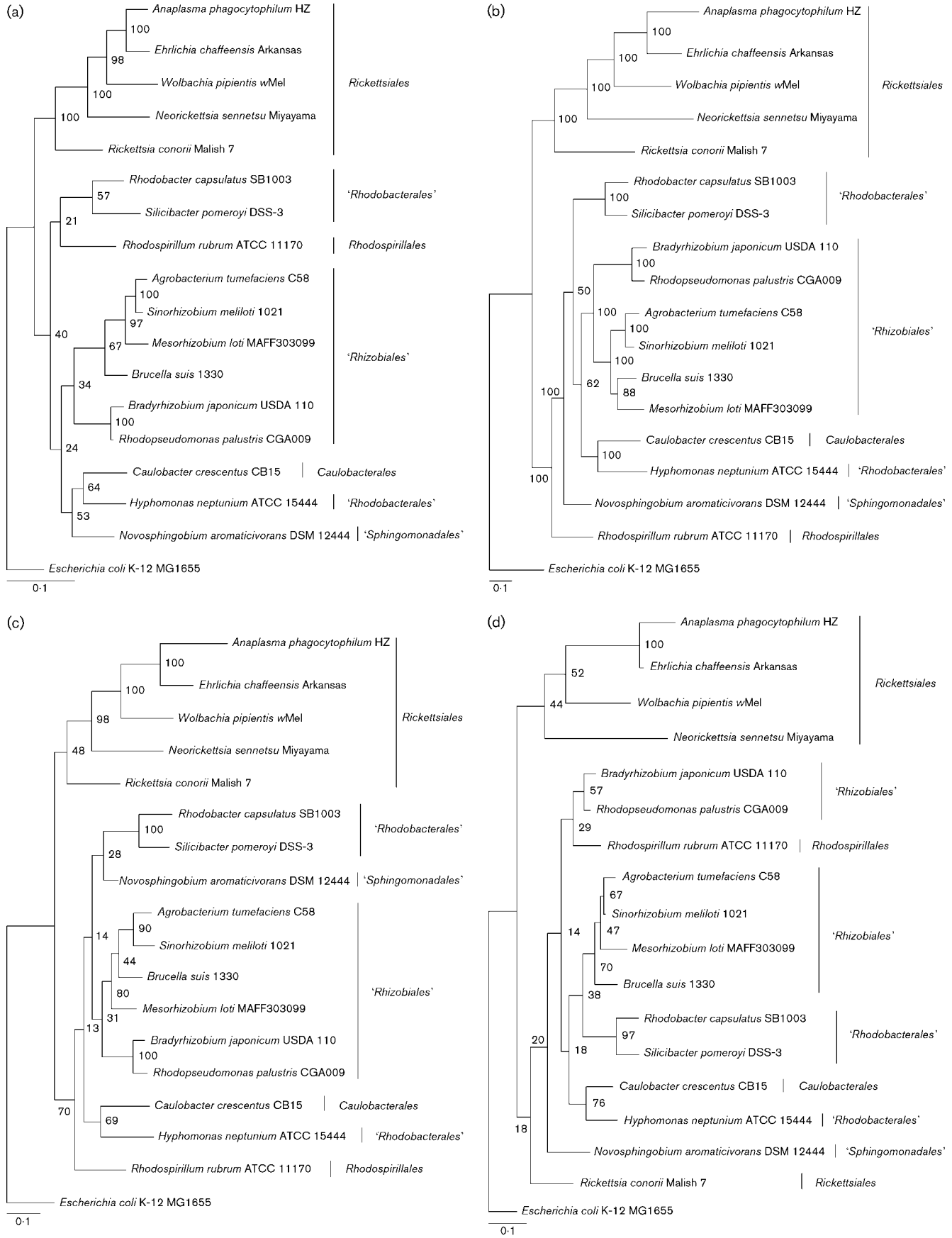
(a)
Anaplasma phagocytophilum HZ
Ehrlichia chaffeensis Arkansas
Wolbachia pipientis wMel
Neorickettsia sennetsu Miyayama
Rickettsia conorii Malish 7
Rickettsiales
Rhodobacter capsulatus SB1003
Silicibacter pomeroyi DSS-3
'Rhodobacterales'
Rhodospirillum rubrum ATCC 11170
Rhodospirillales
Agrobacterium tumefaciens C58
Sinorhizobium meliloti 1021
Mesorhizobium loti MAFF303099
Brucella suis 1330
Bradyrhizobium japonicum USDA 110
Rhodopseudomonas palustris CGA009
'Rhizobiales'
Caulobacter crescentus CB15
Caulobacterales
Hyphomonas neptunium ATCC 15444
'Rhodobacterales'
Novosphingobium aromaticivorans DSM 12444
'Sphingomonadales'
Escherichia coli K-12 MG1655
0·1
(b)
Anaplasma phagocytophilum HZ
Ehrlichia chaffeensis Arkansas
Wolbachia pipientis wMel
Neorickettsia sennetsu Miyayama
Rickettsia conorii Malish 7
Rickettsiales
Rhodobacter capsulatus SB1003
Silicibacter pomeroyi DSS-3
'Rhodobacterales'
Bradyrhizobium japonicum USDA 110
Rhodopseudomonas palustris CGA009
Agrobacterium tumefaciens C58
Sinorhizobium meliloti 1021
Brucella suis 1330
Mesorhizobium loti MAFF303099
'Rhizobiales'
Caulobacter crescentus CB15
Caulobacterales
Hyphomonas neptunium ATCC 15444
'Rhodobacterales'
Novosphingobium aromaticivorans DSM 12444
'Sphingomonadales'
Rhodospirillum rubrum ATCC 11170
Rhodospirillales
Escherichia coli K-12 MG1655
0·1
(c)
Anaplasma phagocytophilum HZ
Ehrlichia chaffeensis Arkansas
Wolbachia pipientis wMel
Neorickettsia sennetsu Miyayama
Rickettsia conorii Malish 7
Rickettsiales
Rhodobacter capsulatus SB1003
Silicibacter pomeroyi DSS-3
'Rhodobacterales'
Novosphingobium aromaticivorans DSM 12444
'Sphingomonadales'
Agrobacterium tumefaciens C58
Sinorhizobium meliloti 1021
Brucella suis 1330
Mesorhizobium loti MAFF303099
Bradyrhizobium japonicum USDA 110
Rhodopseudomonas palustris CGA009
'Rhizobiales'
Caulobacter crescentus CB15
Caulobacterales
Hyphomonas neptunium ATCC 15444
'Rhodobacterales'
Rhodospirillum rubrum ATCC 11170
Rhodospirillales
Escherichia coli K-12 MG1655
0·1
(d)
Anaplasma phagocytophilum HZ
Ehrlichia chaffeensis Arkansas
Wolbachia pipientis wMel
Neorickettsia sennetsu Miyayama
Rickettsiales
Bradyrhizobium japonicum USDA 110
Rhodopseudomonas palustris CGA009
'Rhizobiales'
Rhodospirillum rubrum ATCC 11170
Rhodospirillales
Agrobacterium tumefaciens C58
Sinorhizobium meliloti 1021
Mesorhizobium loti MAFF303099
Brucella suis 1330
'Rhizobiales'
Rhodobacter capsulatus SB1003
Silicibacter pomeroyi DSS-3
'Rhodobacterales'
Caulobacter crescentus CB15
Caulobacterales
Hyphomonas neptunium ATCC 15444
'Rhodobacterales'
Novosphingobium aromaticivorans DSM 12444
'Sphingomonadales'
Rickettsia conorii Malish 7
Rickettsiales
Escherichia coli K-12 MG1655
0·1

**Fig. 2.** Maximum-likelihood trees based on 23S rRNA gene sequences (a), 30 concatenated ribosomal proteins (L2, L3, L4, L5, L13, L14, L15, L16, L17, L20, L21, L22, L23, L24, L27, S2, S3, S4, S6, S7, S8, S10, S11, S12, S13, S14, S15, S16, S17 and S19) (b), HSP70 proteins (c) and EF-Tu proteins (d) from sequenced α-proteobacteria. Node labels are bootstrap values (100 replicates). Note the grouping of *Hyphomonas neptunium* with *C. crescentus* in each tree. See Table 1 for the GenBank GI numbers and ranges used from published genomes.

(Fig. 2c) and EF-Tu proteins (Fig. 2d), supports this classification. Instead, they support a relationship between *Hyphomonas neptunium* and *C. crescentus*. A similar relationship was seen in the trees generated by APIS, in which over 30 % of the *Hyphomonas neptunium* proteins had a protein from *C. crescentus* as their closest relative, as opposed to only 6 % that grouped with a member of the '*Rhodobacterales*'. Most notably, the flagellar and other chemotaxis proteins tend to show a closer relationship to those of *Silicibacter pomeroyi* than to those of *C. crescentus*, although this may be because the *Hyphomonas neptunium* versions of these proteins are quite divergent from even their closest known homologues.

The bootstrap support values for the clades of interest in these trees vary. The 16S rRNA gene sequence tree (Fig. 1) shows only weak (52 %) support for the currently accepted grouping of *Hyphomonas neptunium* among the '*Rhodobacterales*', and the 23S rRNA gene sequence tree (Fig. 2a) shows only somewhat stronger (64 %) support for the alternative classification among the *Caulobacterales*. The concatenated ribosomal protein tree (Fig. 2b), however, shows excellent support (100 %) for this alternative classification, and levels of support from the HSP70 (Fig. 2c) and EF-Tu (Fig. 2d) trees for the alternative classification are strong as well (69 and 76 %, respectively).

In order to explore further the degree of support that each tree has for the alternative hypotheses, Kishino–Hasegawa–Templeton tests (Kishino & Hasegawa, 1989; Templeton, 1983) were performed to determine whether each alignment preferred the 16S or the 23S rRNA gene sequence tree. For each alignment, if the mean of the log-likelihood differences between the 16S and 23S tree across the sites was greater than 1·96 standard deviations, then the more likely tree was judged to be significantly preferred. The 23S alignment and all protein alignments except for the EF-Tu alignment significantly preferred the 23S tree; although the 16S alignment preferred the 16S tree and the EF-Tu alignment preferred the 23S tree, they did not do so at a statistically significant level.

## Evolutionary implications

Although the discovery of conflict between 16S rRNA gene sequence and protein trees is not in itself a novel finding (e.g. Doolittle, 1999; Gupta & Golding, 1993), in general such studies either try to argue for the superiority over rRNA of a single favourite marker protein [as was done by Gupta & Golding (1993) for HSP70] or claim that rampant horizontal gene transfer has destroyed all phylogenetic signal (as in Doolittle, 1999). To our knowledge, this is the first study in which numerous proteins, together with the 23S rRNA gene, consistently yield a single alternative order-level classification for a bacterial species.

What can be the cause of this difference? One possibility is horizontal gene transfer of the 16S rRNA gene. Horizontal gene transfer of the 16S rRNA gene has been suggested as an explanation for patterns seen at the genus level (e.g. Schouls *et al.*, 2003; Parker *et al.*, 2002), and artificially induced transfer of the 16S and 23S rRNA genes between *Escherichia coli* and *Salmonella typhimurium* has been demonstrated experimentally (Asai *et al.*, 1999). The presence of only a single copy of the 16S rRNA gene in *Hyphomonas neptunium* would also make horizontal gene transfer of the 16S rRNA gene possibly easier than in most bacteria. Another possibility could be long-branch attraction (Felsenstein, 1978) in the tree based on 16S rRNA gene sequence analysis, but, as shown in Figs 1 and 2(a), the branch lengths appear not to be particularly long.

In addition to being supported by all the sequence data except that for the 16S rRNA gene, a classification of *Hyphomonas* as a member of the *Caulobacterales* also makes sense from the standpoint of phenotypic characters. Like *Caulobacter*, members of *Hyphomonas* are aerobic, dimorphic, prosthecate bacteria. In the current classification scheme, these traits either would have had to evolve independently in the '*Rhodobacterales*' or would have to have been present in a common ancestor of the '*Rhodobacterales*' and *Caulobacterales* and then been lost by the majority of the members of the '*Rhodobacterales*'.

Current guidelines for the rearrangement of higher order taxa preclude the transfer of a genus without analysis of the type species (Sneath, 1992). Given that the type species of *Hyphomonas* is *Hyphomonas polymorpha* rather than *Hyphomonas neptunium*, a transfer of the genus *Hyphomonas* is not presently possible. However, given the close phylogenetic relationship between these two species [according to the 16S rRNA gene sequence and DNA–DNA hybridization studies in Weiner *et al.* (2000) they are among the most closely related of the eight recognized *Hyphomonas* species], we expect that future work on *Hyphomonas polymorpha* will support such a transfer.

Additionally, there exist several genera of prosthecate budding bacteria (*Hirschia*, *Maricaulis* and *Oceanicaulis*) that are immediate relatives of *Hyphomonas* according to 16S rRNA gene sequence phylogeny (Strömpl *et al.*, 2003). Assuming that this is not an artefact of 16S rRNA gene

sequence phylogeny, these genera would have to be transferred into the *Caulobacterales* along with *Hyphomonas*. Further work, including genome sequencing of the type species of representatives of these genera, would provide valuable data that will help to clarify the relationships among the prosthecate α-proteobacteria, and possibly support the transfer of *Hyphomonas*.

## ACKNOWLEDGEMENTS

We thank Gary Olsen for valuable discussion, and Hervé Tettelin for the use of prepublication data from *Anaplasma phagocytophilum*, *Ehrlichia chaffeensis* and *Neorickettsia sennetsu*. We also thank the US Department of Energy Joint Genome Institute for the use of their sequence data from *Novosphingobium aromaticivorans* and *Rhodospirillum rubrum* prior to publication and Integrated Genomics for the use of their *Rhodobacter capsulatus* genome data. The sequencing and analysis of *Hyphomonas neptunium* was funded by National Science Foundation Award 0237224 to Timothy Hoover, Yves Brun and N. L. W. In addition, the phylogenetic analysis was supported in part by NSF Tree of Life Grant 0228651 to J. A. E., N. L. W. and Karen Nelson.

## REFERENCES

**Abraham, W. R., Strömpl, C., Meyer, H. & 8 other authors (1999).** Phylogeny and polyphasic taxonomy of *Caulobacter* species. Proposal of *Maricaulis* gen. nov. with *Maricaulis maris* (Poindexter) comb. nov. as the type species, and emended description of the genera *Brevundimonas* and *Caulobacter*. *Int J Syst Bacteriol* **49**, 1053–1073.

**Asai, T., Zaporojets, D., Squires, C. & Squires, C. L. (1999).** An *Escherichia coli* strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. *Proc Natl Acad Sci U S A* **96**, 1820–1822.

**Blattner, F. R., Plunkett, G., Bloch, C. A. & 14 other authors (1997).** The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474.

**Capela, D., Barloy-Hubler, F., Gouzy, J. & 25 other authors (2001).** Analysis of the chromosome sequence of the legume symbiont *Sinorhizobium meliloti* strain 1021. *Proc Natl Acad Sci U S A* **98**, 9877–9882.

**Cole, J. R., Chai, B., Marsh, T. L. & 8 other authors (2003).** The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res* **31**, 442–443.

**Delcher, A. L., Harmon, D., Kasif, S., White, O. & Salzberg, S. L. (1999).** Improved microbial gene identification with GLIMMER. *Nucleic Acids Res* **27**, 4636–4641.

**Doolittle, W. F. (1999).** Phylogenetic classification and the universal tree. *Science* **284**, 2124–2129.

**Edgar, R. C. (2004).** MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.

**Felsenstein, J. (1978).** Cases in which parsimony and compatibility methods will be positively misleading. *Syst Zool* **27**, 401–410.

**Felsenstein, J. (1981).** Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* **17**, 368–376.

**Felsenstein, J. (2004).** PHYLIP (Phylogeny Inference Package), version 3.6b. Distributed by the author. Department of Genetics, University of Washington, Seattle, USA.

**Fleischmann, R. D., Adams, M. D., White, O. & 37 other authors (1995).** Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 279–291.

**Garrity, G. M., Bell, J. A. & Lilburn, T. (2005).** Order *Rhodobacterales* ord. nov. In *Bergey's Manual of Systematic Bacteriology*, 2nd edn, vol. 2. New York: Springer (in press).

**Gupta, R. S. & Golding, G. B. (1993).** Evolution of HSP70 gene and its implications regarding relationships between archaebacteria, eubacteria, and eukaryotes. *J Mol Evol* **37**, 573–582.

**Henrici, A. T. & Johnson, D. (1935).** Stalked bacteria, a new order of *Schizomycetes*. *J Bacteriol* **29**, 3–4.

**Hirsch, P. (1974).** Budding bacteria. *Annu Rev Microbiol* **28**, 391–444.

**Howe, K., Bateman, A. & Durbin, R. (2002).** QUICKTREE: building huge neighbour-joining trees of protein sequences. *Bioinformatics* **18**, 1546–1547.

**Jones, D. T., Taylor, W. R. & Thornton, J. M. (1992).** The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**, 275–282.

**Kaneko, T., Nakamura, Y., Sato, S. & 21 other authors (2000).** Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res* **7**, 331–338.

**Kaneko, T., Nakamura, Y., Sato, S. & 14 other authors (2002).** Complete genomic sequence of nitrogen-fixing symbiotic bacterium *Bradyrhizobium japonicum* USDA 110. *DNA Res* **9**, 189–197.

**Kishino, H. & Hasegawa, M. (1989).** Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data and the branching order in Hominoidea. *J Mol Evol* **29**, 170–179.

**Larimer, F. W., Chain, P., Hauser, L. & 16 other authors (2004).** Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. *Nat Biotechnol* **22**, 55–61.

**Liefson, E. (1964).** *Hyphomicrobium neptunium* sp. nov. *Antonie van Leeuwenhoek* **30**, 249–256.

**Moore, R. L., Weiner, R. M. & Gebers, R. (1984).** Genus *Hyphomonas* Pongratz 1957 nom. rev. emend., *Hyphomonas polymorpha* Pongratz 1957 nom. rev. emend., and *Hyphomonas neptunium* (Liefson 1964) comb nov. emend. (*Hyphomicrobium neptunium*). *Int J Syst Bacteriol* **34**, 71–73.

**Moran, M. A., Buchan, A., Gonzalez, J. M. & 32 other authors (2004).** Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* **432**, 910–913.

**Nierman, W. C., Feldblyum, T. V., Laub, M. T. & 35 other authors (2001).** Complete genome sequence of *Caulobacter crescentus*. *Proc Natl Acad Sci U S A* **98**, 4136–4141.

**Ogata, H., Audic, S., Renesto-Audiffren, P. & 8 other authors (2001).** Mechanisms of evolution in *Rickettsia conorii* and *R. prowazekii*. *Science* **293**, 2093–2098.

**Parker, M. A., Lafay, B., Burdon, J. J. & van Berkum, P. (2002).** Conflicting phylogeographic patterns in rRNA and *nifD* indicate regionally restricted gene transfer in *Bradyrhizobium*. *Microbiology* **148**, 2557–2565.

**Paulsen, I. T., Seshadri, R., Nelson, K. E. & 27 other authors (2002).** The *Brucella suis* genome reveals fundamental similarities between animal and plant pathogens and symbionts. *Proc Natl Acad Sci U S A* **99**, 13148–13153.

**Pongratz, E. (1957).** D'une bactérie pediculé isolé d'un pus de sinus. *Schweiz Z Allg Pathol Bakteriol* **20**, 593–608 (in French).

**Saitou, N. & Nei, M. (1987).** The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.

**Schouls, L. M., Schot, C. S. & Jacobs, J. A. (2003).** Horizontal transfer of segments of the 16S rRNA genes between species of the *Streptococcus anginosus* group. *J Bacteriol* **185**, 7241–7246.

**Sneath, P. H. A. (1992).** *International Code of Nomenclature of Bacteria (1990 Revision). Bacteriological Code.* Washington, DC: American Society for Microbiology.

**Stackebrandt, E., Frederiksen, W., Garrity, G. M. & 10 other authors (2002).** Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* **52**, 1043–1047.

**Stahl, D. A., Key, R., Flesher, B. & Smit, J. (1992).** The phylogeny of marine and freshwater caulobacters reflects their habitat. *J Bacteriol* **174**, 2193–2198.

**Strömpl, C., Hold, G. L., Lünsdorf, H., Graham, J., Gallacher, S., Abraham, W. R., Moore, E. R. B. & Timmis, K. N. (2003).** *Oceanicaulis alexandrii* gen. nov., sp. nov., a novel stalked bacterium isolated from a culture of the dinoflagellate *Alexandrium tamarense* (Lebour) Balech. *Int J Syst Evol Microbiol* **53**, 1901–1906.

**Templeton, A. R. (1983).** Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the evolution of humans and the apes. *Evolution* **37**, 221–244.

**Tettelin, H., Nelson, K. E., Paulsen, I. T. & 36 other authors (2001).** Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* **293**, 498–506.

**Weiner, R. M., Melick, M., O'Neill, K. & Quintero, E. (2000).** *Hyphomonas adhaerens* sp. nov., *Hyphomonas johnsonii* sp. nov. and *Hyphomonas rosenbergii* sp. nov., marine budding and prosthecate bacteria. *Int J Syst Evol Microbiol* **50**, 459–469.

**Wood, D. W., Setubal, J. C., Kaul, R. & 48 other authors (2001).** The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**, 2317–2323.

**Wu, M., Sun, L. V., Vamathevan, J. & 27 other authors (2004).** Phylogenomics of the reproductive parasite *Wolbachia pipientis* wMel: a streamlined genome overrun by mobile genetic elements. *PLoS Biol* **2**, E69.