HYPOTHESIS

# Evidence for an early prokaryotic endosymbiosis

James A. Lake[1,2,3,4]

**Endosymbioses have dramatically altered eukaryotic life, but are thought to have negligibly affected prokaryotic evolution. Here, by analysing the flows of protein families, I present evidence that the double-membrane, Gram-negative prokaryotes were formed as the result of a symbiosis between an ancient actinobacterium and an ancient clostridium. The resulting taxon has been extraordinarily successful, and has profoundly altered the evolution of life by providing endosymbionts necessary for the emergence of eukaryotes and by generating Earth's oxygen atmosphere. Their double-membrane architecture and the observed genome flows into them suggest a common evolutionary mechanism for their origin: an endosymbiosis between a clostridium and actinobacterium.**

Endosymbioses have affected almost every aspect of eukaryotic evolution[1], but appreciation of the pervasiveness of this mechanism has not been achieved without controversy. Virtually every proposed symbiotic event in the origin of the eukaryotic cell has been vigorously debated[2–7]. Nevertheless, today there is general agreement that the chloroplast and mitochondria have endosymbiotic origins, and there is a developing recognition that eukaryotic cells have hosted diverse endosymbiotic guests during their existence. These include numerous and varied prokaryotes, and even other eukaryotes complete with nucleus and organelles.

Acting as functional modules, endosymbionts have the ability to dramatically transform the metabolic and architectural properties of their hosts[8]. Endosymbiotic guests bring with them novel metabolic and signalling capabilities. These new capabilities can be provided through the introduction of new genes, or through the generation of additional membrane-bounded compartments formed during endosymbioses. The ability to form new cellular compartments distinguishes endosymbioses from repeated episodes of gene transfer involving conjugation, transformation and transduction[8–10].

These are exciting times for microbiology, with important discoveries being made almost every week[11–13]. Here I explore what is known about prokaryotic symbioses and endosymbioses, and provide evidence that a significant fraction of prokaryotic diversity has been affected by symbioses, and probably by endosymbioses.

## Exploring the possibility of prokaryotic endosymbioses

It is widely accepted that eukaryotes have frequently hosted endosymbionts during the course of their evolution. But can prokaryotes themselves also host other endosymbiotic guests? Apparently so, at least within eukaryotic cells. There is good evidence for a eukaryote containing a prokaryotic host and guest in the secondary endosymbionts of certain aphids[14]. Specifically, the β-proteobacterial *Buchnera* endosymbionts themselves contain γ-proteobacterial endosymbionts, generating endosymbionts within other endosymbionts, much like a set of Russian nesting (*matryoshka*) dolls. Resembling eukaryotic endosymbioses, the γ-proteobacterial endosymbionts are present within the cytoplasm of their β-proteobacterial hosts (Fig. 1a).

Exclusively prokaryotic endosymbioses are unknown, although the phototrophic consortium '*Chlorochromatium aggregatum*' comes pretty close. It consists of two partners; a motile, heterotrophic, central β-proteobacterium surrounded by 10–60 peripheral photosynthetic green sulphur bacteria[15]. The green sulphur bacteria, or epibionts, labelled GSB in Fig. 1b, are attached to the central bacterium through periplasmic tubules formed by extensions from the outer membrane of the flagellated central β-proteobacterium. This allows the central β-proteobacterium, labelled β in Fig. 1b, to move the consortium to favourable anaerobic habitats where sulphides and sufficient light for photosynthesis are available. Structural studies of the consortium suggest that, "the two partner bacteria may actually share a common periplasmic space"[16]. If so, then both partners would be enclosed by a common periplasm, but not by a common membrane. Although this symbiosis does not provide a precedent for a prokaryotic endosymbiosis, it suggests how one might have started.

Could it be possible that primary endosymbioses exclusively involving prokaryotes have shaped the evolution of life at some time in the past?

## Prokaryotic diversity

There is currently much discussion of the prokaryotic 'tree of life', but there are few points of agreement regarding its topology, except

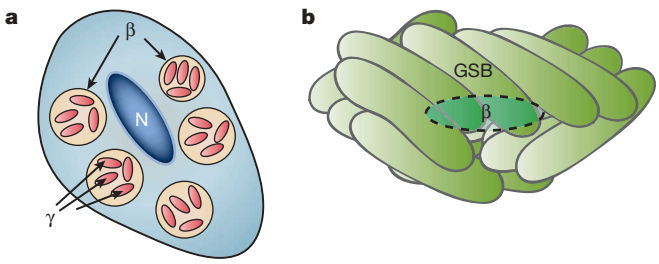


**Figure 1 | Schematic diagrams illustrating prokaryotic symbionts. a**, A diagram of a mealybug cell containing β-proteobacterial endosymbionts that contain γ-proteobacterial endosymbionts within them[14]. The mealybug nucleus, N, is surrounded by β-proteobacterial endosymbionts, labelled β, that are themselves hosts to γ-proteobacterial endosymbionts, labelled γ. **b**, A diagram of the prokaryotic consortium, *Chlorochromatium aggregatum*[16], consisting of peripheral green sulphur bacteria, labelled GSB, surrounding a central, spindle shaped, β-proteobacterium, shown by a dashed line and labelled β.

[1]Department of Molecular, Cellular and Developmental Biology, [2]Molecular Biology Institute, [3]Department of Human Genetics, [4]UCLA Astrobiology Institute, University of California, Los Angeles, California 90095, USA.

that it is not a tree[17–25]. However, if we blur our taxonomic vision and just look at the low-resolution outlines of life, then it is generally agreed that all prokaryotes belong to one of five groups[26–29]. These five natural, phylogenetically well separated, and relatively homogeneous taxa[30–32] are: the Archaea, the Bacilli and relatives, the Clostridia and relatives, the Actinobacteria, and the double membrane, or Gram negative, prokaryotes. The Archaea contain extreme halophiles, methanogens, hyperthermophiles and other unique phenotypes. The Bacilli and the Clostridia are firmicutes, characterized by their low-guanine-cytosine (low-GC) genomic compositions, although not exclusively[33]. The Clostridia are unique among the single-membrane prokaryotes for including photosynthetic organisms as well as fermenting ones. The Actinobacteria, characterized by high-GC genomic compositions, are morphologically diverse and include many human pathogens, such as those that cause leprosy and tuberculosis. The double-membrane, Gram-negative, prokaryotes encompass approximately 42 phyla. (For a detailed list of the phyla, including some recent taxonomic changes[26–28], see Supplementary Information section 1.) Together, these five super-taxa contain all known prokaryotic life[34].

## The double-membrane prokaryotes

Consider the double-membrane prokaryotes in a little more detail. This taxon contains an amazingly diverse, inordinately speciose, and possibly primitively photosynthetic group of organisms. It encompasses more proteomic diversity as measured by protein families (4,756) than all other prokaryotes combined (4,000). Furthermore, they have profoundly altered life on Earth, producing 20% of the Earth's atmosphere and providing endosymbionts critical for the emergence of eukaryotes. It includes the photosynthetic Cyanobacteria, α-, β-, γ-, δ- and ε-proteobacteria, Chloroflexi and Chlorobi, and numerous other intriguing taxa, such as the Spirochaetes, Planctomycetes and Aquificales.

Before the advent of molecular sequencing, the double-membrane clan, an unrooted clade[35], was loosely distinguished from single-membrane prokaryotic groups by its Gram-negative membrane staining properties[36]. Subsequently, its membership was defined more accurately when phyla within the double-membrane clan were identified on the basis of the presence of characteristic indels (that is, insertions or deletions) contained within proteins[37]. Most recently, indel rooting studies have shown that the double-membrane clan is a clade, based on heat-shock protein Hsp70 (ref. 31) and on enzymes involved in pyrimidine-biosynthesis (PyrD protein) and histidine-biosynthesis (HisA and HisF proteins)[29]. It is interesting that these last two indels also showed that the double-membrane prokaryotes are closely related to the Actinobacteria. This point will be discussed further below.

It is also intriguing that they have this double-membrane structure, whereas other prokaryotic groups are surrounded by single membranes. In fact, the membrane organization of double-membrane prokaryotes fundamentally differs from that found in single-membrane prokaryotes. In the former, the peptidoglycan layer is sandwiched between the outer and inner membranes, so that it surrounds the inner membrane: in contrast, in the latter there is no inner membrane, and the peptidoglycan layer, located outside the cell, surrounds the outer membrane. Also, double-membrane prokaryotes contain their flagellar motors in the inner membrane, whereas single-membrane prokaryotes contain their flagellar motors in the outer membrane. And the photosynthetic apparatus in double-membrane prokaryotes is in the inner membrane, rather than in the outer membranes as in single-membrane prokaryotes. In other words, the organization of the inner membrane of the double-membrane prokaryotes resembles that of the outer membranes of typical single-membrane prokaryotes. The inner membranes of double-membrane prokaryotes are organized almost as if they were derived from the outer membrane of an engulfed single-membrane prokaryote.

## Phylogenetic testing

If the double-membrane prokaryotes were produced by an endosymbiosis, then perhaps the genomic signatures of the endosymbiosis might still be strong enough to be detected in phylogenetic reconstructions. Here I reconstruct the evolution of the double-membrane prokaryotes using methods capable of discriminating between rings, caused by endosymbioses, and trees, caused by clonal evolution. These methods[38,39] distinguish trees from rings by analysing gene, or protein family, presences and absences. In many ways they are similar to, but in other ways quite different from, tree reconstruction algorithms that use nucleotide and amino acid changes.

If the double-membrane prokaryotes were the result of an endosymbiosis, then the flow of genes from the two donors into the proto-double-membrane-prokaryote would not be consistent with any tree, as branches on a tree must diverge from a single node representing a common ancestral organism. In contrast, an endosymbiosis, representing the genomic merger of two distinct organisms into one, proceeds from two separate nodes corresponding to the host and the guest. Thus endosymbioses produce rings rather than trees, as in Fig. 2a, similar to that observed for the eukaryotic ring of life[20]. If the present studies were to reconstruct a tree rather than a ring, then an endosymbiosis would be rejected, thereby providing a first test for an endosymbiosis.

But if the double membranes are signs of a past endosymbiosis, then which prokaryotes might be the culprits—that is, the host and the guest? Fortunately, because genes can flow into the new organism from two different donors during an endosymbiosis, phylogenetic analyses can also identify the donors. In the case of an endosymbiosis, the two donor taxa surrounding the clade of double-membrane prokaryotes (D) would correspond to the donor taxa (Fig. 2b). Thus if the taxa adjacent to D could be identified, this would provide evidence for specific donors, and represent a second level of testing for an endosymbiosis.

And finally, how could one determine which taxon within the ring would be the endosymbiont? For example, if a ring of five taxa is reconstructed, at first glance it might seem that any of the five could be the endosymbiont. As demonstrated elsewhere[38], presence-absence analyses are able to detect the directions of gene flow. In an endosymbiosis, one expects to see genes flowing from both donors into the endosymbiont (Fig. 2c). Thus if genes were found to flow into the double-membrane prokaryotes from both sides, then this pattern of gene flow, and only this pattern, would be consistent with a double-membrane endosymbiotic origin. Accordingly, gene presence-absence studies can specifically identify the endosymbiont and thereby provide a third level of proof for an endosymbiosis.

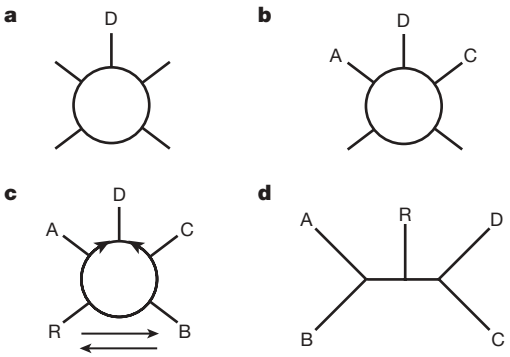


**Figure 2 | An illustration of three steps required for the identification of an endosymbiont from gene flow data, and the tree of life that best fits the gene flow data.** **a**, The endosymbiont, D, must be part of a ring. **b**, The genome donors, A and C, must surround the endosymbiont and be fully resolved. **c**, The direction of gene flow must be from both donors to the endosymbiont, D. **d**, The best tree of life. A, Actinobacteria; B, Bacilli; C, Clostridia; D, Double-membrane prokaryotes; R, Archaea.

However, this still might not distinguish an endosymbiosis from alternative types of symbioses, but that point will be discussed subsequently.

## Protein family analyses

To test for prokaryotic endosymbioses, I analyse here the five natural taxa previously described: the Archaea (R), the Actinobacteria (A), the Bacilli and relatives (B), the Clostridia and relatives (C), and the double-membrane prokaryotes (D). Using absence-presence analyses of protein families obtained from >3,000 diverse prokaryotes, I reconstruct a low-resolution graph of ancient prokaryotic evolution, to determine whether it is a tree, a ring, or some other graph.

The information required to reconstruct the 'tree/graph of life' is contained in ten parsimoniously informative patterns of protein family absences and presences[38,39]. These patterns are listed in the first five columns of Table 1, and consist of three gene presences ('+'), and two gene absences (blanks). In the first row, three protein families present in Archaea (R), Actinobacteria (A) and Bacilli (B), and absent from all Clostridia (C) and all double-membrane prokaryotes (D) are analysed. The most parsimonious tree or graph is determined by computing the character state patterns allowed by graphical models, as described elsewhere[38], and then comparing these 'allowed patterns' with the observed numbers of protein families that support each pattern.

As it might happen that the data have little resolving power and are unable to discriminate readily between tree-like evolution and endosymbiotic evolution, I explicitly compare the graph that best fits the data with the tree that best fits the data to see if trees can be resolved from rings. The number of protein families observed for each pattern is shown in the sixth and seventh columns of Table 1. Bold font with asterisk indicates that large numbers of counts are predicted, and standard fonts indicate that small numbers of counts are predicted[38]. Predictions based on the best graph (Fig. 2c) are shown in the sixth column, and those based on the best tree (Fig. 2d) are shown in the seventh column.

It is immediately clear that the data fit the predictions calculated from the best graph. The five allowed patterns coincide with the five largest scores, 62–174, and the five disallowed patterns coincide with the five smallest scores, 0–15. But might they, through some statistical quirk, also fit the best tree? For the best tree, the data conflict with the predictions for three patterns. The conflicts are: large scores are predicted by the best tree for the patterns in the first and fourth rows and yet they correspond to small counts, 3 and 15; and small scores are predicted for the pattern in row five and yet it corresponds to a large count, 62. Clearly, the significantly better agreement found for the best graph compared with the best tree indicates that graphs are needed to explain prokaryotic evolution.

**Table 1 | Protein family support for the best ring and for the best tree of prokaryotic life**

| Character state patterns | | | | | Best graph protein families | Best tree protein families |
|---|---|---|---|---|---|---|
| R | A | B | C | D | | |
| + | + | + | | | 3 | **3*** |
| + | + | | + | | 0 | 0 |
| + | | + | + | | 8 | 8 |
| | + | + | + | | 15 | **15*** |
| + | + | | | + | **62*** | 62 |
| + | | + | | + | 15 | 15 |
| | + | + | | + | **91*** | **91*** |
| + | | | + | + | **99*** | **99*** |
| | + | | + | + | **73*** | **73*** |
| | | + | + | + | **174*** | **174*** |

The topology corresponding to the 'best ring' of prokaryotic life is shown in Fig. 2c and that corresponding to the 'best tree' is shown in Fig. 2d. Protein families are calculated using the Pfam (v. 22.0) database as implemented on the Janelia Farms website (available at http://pfam.janelia.org/search). Allowed character state patterns are indicated in the table by bold type and are calculated according to procedures published elsewhere[38,39]. Specifically, the five allowed patterns predicted by the best ring of life, shown in Fig. 2c, are those in which A and D are both '+', and/or those in which C and D are both '+'. The six allowed patterns predicted by the best tree of life, shown in Fig. 2d, are those in which A and B are both '+' and/or those in which C and D are both '+'. Explicit definitions of the prokaryotic groups included within the Archaea, R, the Actinobacteria, A, the Bacilli and relatives, B, the Clostridia and relatives, C, and the double-membrane prokaryotes, D, are provided in Supplementary Table 1. Detailed listings of the protein families that are taxonomically distributed according to each of the ten character state patterns are provided in Supplementary Information section 2.

Furthermore, there is unusually strong statistical support for the best graph. Using the numbers of allowed protein families to estimate the probabilities of each of the ten informative character state sets provides significant posterior bootstrap support[38,40] for the graph in Fig. 2c, $P > 0.9999$, whereas there is no support for the tree in Fig. 2d, $P < 0.0001$.

Indeed, when the evidence based on the flow of protein families is examined, it is certainly plausible that it was an endosymbiosis. In Fig. 3, the taxa meet all of the criteria supporting an endosymbiosis. (1) The putative endosymbiont, the double-membrane taxon, is part of the ring. (2) The double-membrane prokaryotes are surrounded on the left by the Actinobacteria, and on the right by the Clostridia, indicating that these two taxa are the donors. (3) The arrows on either side of the double-membrane prokaryotes indicate that genes flow from the Actinobacteria and from the Clostridia, implicating them as the gene donors, and flow into the double-membrane prokaryotes, specifically identifying them as the recipient.

The topology of the graph in Fig. 3 is supported by some recent, formerly perplexing observations. Analyses of the distribution of two slowly evolving indels found in enzymes involved in pyrimidine- and histidine-biosynthesis, PyrD and HisA, demonstrated that the double-membrane prokaryotes and the Actinobacteria form a statistically supported clade[29]. This supports the upper-left portion of the prokaryotic ring of life, where genes flow from the Actinobacteria into the double-membrane prokaryotes. Furthermore, as the PyrD and HisA indels were analysed with respect to paralogous outgroups, they established that genes flowed from the root into the Actinobacteria and the double-membrane prokaryotes, consistent with the direction of gene flow shown in Fig. 3. Thus these indel analyses provide statistically significant, independent support for the left half of the symbiosis.

The graph in Fig. 3 also starts to make sense of the perplexing phylogenetic distribution of the photosynthetic machinery. Complex, highly integrated systems involving many genes are thought to be more difficult to transfer horizontally/laterally[10,41]. As the photosynthetic apparatus involves scores of genes, this makes the presence or absence of photosynthesis a reasonable phylogenetic marker for the study of ancient prokaryotic divergences. In prokaryotes, photosynthesis is found only in the double-membrane prokaryotes and in the Clostridia[42]. This distribution has been hard to reconcile with trees. But it now starts to fit with Fig. 3, because at the upper right-hand side of the figure the double-membrane prokaryotes are adjacent to the Clostridia. Thus, consistent with photosynthesis originating in the Clostridia, this graph parsimoniously explains the transfer of

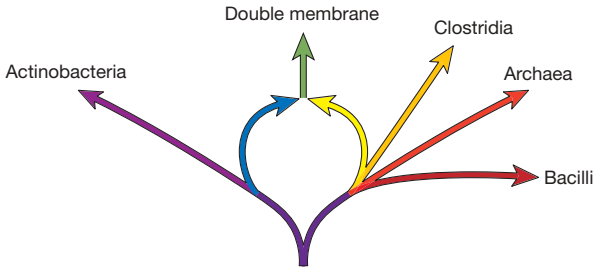


**Figure 3 | A schematic diagram illustrating the prokaryotic ring of life.** The actinobacterial genome donor, at the left (blue), and the clostridial genome donor, at the right (yellow), transfer their genomes to form the double-membrane prokaryotes at the top of the ring (green). The protein family data identify the Actinobacteria and the Clostridia as donors, and the double-membrane prokaryotes as the fusion organism, but cannot fully resolve the relationship between the Bacilli (dark red) and the Archaea (bright red)[38]. This ring partially explains why a prokaryotic 'tree of life' is not a tree[23]. The rooting shown is the new root of life[29,32,51].

photosynthetic genes to the double-membrane prokaryotes. This, of course, does not resolve the many questions related to the evolution of photosystems I and II[43], but it is a start.

These results are certainly consistent with the proposition that the double-membrane prokaryotes were formed as the result of an endosymbiosis between an ancient ancestral population of Clostridia and Actinobacteria. So far I have assumed that an endosymbiosis was responsible, but might other symbioses also be possible?

## Symbioses and genome transfers

Symbioses, including endosymbioses, do not happen in a single generation. They develop over long periods of time as symbiotic partners evolve, adapt, and exchange genes through the traditional mechanisms of gene transfer—that is, conjugation, transformation, and viral transduction[9,10]. During a symbiosis, frequent gene transfers occur between partners owing to physical proximity and other factors[10,19,44–46]. Thus it is thought that, over time, organisms in stable symbioses can transfer significant portions of their genomes to their partners, resulting in the nearly complete transfer of genomes, or 'genome transfer'[47]. On the basis of the protein family results and analyses presented here, it is also possible that various types of symbioses could have produced the protein family results in Table 1, provided that these symbioses are also consistent with those analyses. Thus any two-donor symbioses involving a clostridium and an actinobacterium donating their genes to form the double-membrane prokaryotes would fit these data. This very probably happened over an extended period of time.

We cannot say exactly how much time this symbiosis required, but it definitely did not happen in the past two billion years. We know this because the cyanobacterial double-membrane prokaryotes are responsible for producing the Earth's oxygen atmosphere. This implies that their diversification, and hence that of the double-membrane clade, started before the rise of oxygen in the atmosphere, approximately 2.4 billion years ago[48,49] or possibly earlier (2.7 billion years ago[48,50]). Because the double-membrane prokaryotes are descended from Actinobacteria and Clostridia, these two donor clades must trace their beginnings back to even earlier times. Thus there seem to be no obvious physical time constraints on how long the initial phase of the double-membrane symbiosis lasted. In the future, the double-membrane clade may provide a useful reference taxon for calibrating molecular clock studies.

## Symbiosis or endosymbiosis

I find it fascinating that this prokaryotic symbiosis could so profoundly shape the evolution of life, and thereby set the stage for the formation of an oxygen-rich atmosphere and for the emergence of eukaryotic organelles. The symbiosis of two disparate prokaryotes to form the double-membrane prokaryotes appears to have done more than just introduce a new combination of genes into a cell: it also generated a group of organisms surrounded by two membranes. In some ways, it appears as if the formation of the periplasm, the movement of the peptidoglycan layer from outside the cell into the periplasm, the movement of the flagella motor into the inner membrane, and the redesign of older transport proteins into those capable of spanning two membranes represents a radically new structural design.

I cannot help but notice that the existence of the double-membrane structure immediately suggests a possible mechanism for its formation and for the observed genome transfers from Clostridia and Actinobacteria to the double-membrane prokaryotes—endosymbioses. I believe that this agreement should not be ascribed to chance.


1. Margulis, L. *Origin of the Eukaryotic Cells* (Yale Univ. Press, 1970).
2. Margulis, L. *Symbiosis in Cell Evolution* (Freeman and Company, 1981).
3. Lake, J. A. & Rivera, M. C. Was the nucleus the first endosymbiont? *Proc. Natl Acad. Sci. USA* **91,** 2880–2881 (1994).
4. Martin, W. & Muller, M. The hydrogen hypothesis for the first eukaryote. *Nature* **392,** 37–41 (1998).
5. Lang, B. F. & Gray, M. W. Mitochondrial genome evolution and the origin of eukaryotes. *Annu. Rev. Genet.* **33,** 351–397 (1999).
6. Poole, A., Jeffares, D. & Penny, D. Early evolution: prokaryotes, the new kids on the block. *Bioessays* **21,** 880–889 (1999).
7. Keeling, P. J. & Palmer, J. D. Horizontal gene transfer in eukaryotic evolution. *Nature Rev. Genet.* **9,** 605–614 (2008).
8. Archibald, J. M. & Keeling, P. J. Recycled plastids: a 'green movement' in eukaryotic evolution. *Trends Genet.* **18,** 577–584 (2002).
9. Syvanen, M. & Kado, C. I. (eds) *Horizontal Gene Transfer* (Chapman and Hall, 1998).
10. Jain, R., Rivera, M. C., Moore, J. E. & Lake, J. A. Horizontal gene transfer in microbial genome evolution. *Theor. Popul. Biol.* **61,** 489–495 (2002).
11. Bowler, C., Karl, D. M. & Colwell, R. R. Microbial oceanography in a sea of opportunity. *Nature* **459,** 180–184 (2009).
12. Fuhrman, J. A. Microbial community structure and its functional implications. *Nature* **459,** 193–199 (2009).
13. DeLong, E. F. The microbial ocean from genomes to biomes. *Nature* **459,** 200–206 (2009).
14. von Dohlen, C. D., Kohler, S., Alsop, S. T. & McManus, W. R. Mealybug β-proteobacterial endosymbionts contain γ-proteobacterial symbionts. *Nature* **412,** 433–436 (2001).
15. Froestl, J. & Overmann, J. Phylogenetic affiliation of the bacteria that constitute phototrophic consortia. *Arch. Microbiol.* **174,** 50–58 (2000).
16. Wanner, G., Vogl, K. & Overmann, J. Ultrastructural characterization of the prokaryotic symbiosis in "*Chlorochromatium aggregatum*". *J. Bacteriol.* **190,** 3721–3730 (2008).
17. Hilario, E. & Gogarten, J. P. Horizontal transfer of ATPase genes—the tree of life becomes a net of life. *Biosystems* **31,** 111–119 (1993).
18. Dagan, T. & Martin, W. The tree of one percent. *Genome Biol.* **7,** 1181–1187 (2006).
19. Jain, R., Rivera, M. C., Moore, J. E. & Lake, J. A. Horizontal gene transfer accelerates genome innovation and evolution. *Mol. Biol. Evol.* **20,** 1598–1602 (2003).
20. Rivera, M. C. & Lake, J. A. The ring of life: evidence for a genome fusion origin of eukaryotes. *Nature* **431,** 152–155 (2004).
21. Konstantinidis, K. T. & Tiedje, J. M. Towards a genome-based taxonomy for prokaryotes. *J. Bacteriol.* **187,** 6258–6264 (2005).
22. Sorek, R. *et al.* Genome-wide experimental determination of barriers to horizontal gene transfer. *Science* **318,** 1449–1452 (2007).
23. Doolittle, W. F. Phylogenetic classification and the universal tree. *Science* **284,** 2124–2128 (1999).
24. McInerney, J. O. & Pisani, D. Paradigm for life. *Science* **318,** 1390–1391 (2007).
25. Martin, W. & Embley, T. M. Early evolution comes full circle. *Nature* **431,** 134–136 (2004).
26. Ohno, M. *et al. Symbiobacterium thermophilum* gen. nov., sp. nov., a symbiotic thermophile that depends on co-culture with a *Bacillus* strain for growth. *Int. J. Syst. Evol. Microbiol.* **50,** 1829–1832 (2000).
27. Garrity, G. & Holt, J. G. in *Bergey's Manual of Systematic Bacteriology* (eds Boone, D. & Castenholz, R. W.) 119–141 (Springer, 2001).
28. Wu, M. *et al.* Life in hot carbon monoxide: the complete genome sequence of *Carboxydothermus hydrogenoformans* Z-2901. *PLoS Genet.* **1,** 563–574 (2005).
29. Lake, J. A., Servin, J. A., Herbold, C. W. & Skophammer, R. G. Evidence for a new root of the tree of life. *Syst. Biol.* **57,** 835–843 (2008).
30. Skophammer, R. G., Herbold, C. W., Rivera, M., Servin, J. A. & Lake, J. A. Evidence that the root of the tree of life is not within the Archaea. *Mol. Biol. Evol.* **23,** 1648–1651 (2006).
31. Lake, J. A., Herbold, C. W., Rivera, M. C., Servin, J. A. & Skophammer, R. G. Rooting the tree of life using non-ubiquitous genes. *Mol. Biol. Evol.* **24,** 130–136 (2007).
32. Skophammer, R. G., Servin, J. A., Herbold, C. W. & Lake, J. A. Evidence for a Gram positive, eubacterial root of the tree of life. *Mol. Biol. Evol.* **24,** 1761–1768 (2007).
33. Ueda, K. *et al.* Genome sequence of *Symbiobacterium thermophilum*, an uncultivable bacterium that depends on microbial commensalism. *Nucleic Acids Res.* **32,** 4937–4944 (2004).
34. Boone, D. & Castenholz, R. W. (eds) *Bergey's Manual of Systematic Bacteriology* 2nd edn Vol. 1, *The Archaea and the Deep Branching and Phototrophic Bacteria* (ed. Garrity, G. M.) (Springer, 2001).
35. Wilkinson, M., McInerney, J. O., Hirt, R. P., Foster, P. G. & Embley, T. M. Of clades and clans: terms for phylogenetic relationships in unrooted trees. *Trends Ecol. Evol.* **22,** 114–115 (2007).
36. Stanier, R. Y., Adelberg, E. A. & Ingraham, J. L. *The Microbial World* (Prentice-Hall, 1976).
37. Gupta, R. S. Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaebacteria, eubacteria, and eukaryotes. *Microbiol. Mol. Biol. Rev.* **62,** 1435–1491 (1998).
38. Lake, J. A. Reconstructing evolutionary graphs: 3D parsimony. *Mol. Biol. Evol.* **25,** 1677–1682 (2008).
39. Lake, J. A. & Rivera, M. C. Deriving the genomic tree of life in the presence of horizontal gene transfer: conditioned reconstruction. *Mol. Biol. Evol.* **21,** 681–690 (2004).
40. Lake, J. A. Calculating the probability of multitaxon evolutionary trees—bootstrappers gambit. *Proc. Natl Acad. Sci. USA* **92,** 9662–9666 (1995).

41. Jain, R., Rivera, M. C. & Lake, J. A. Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl Acad. Sci. USA* **96,** 3801–3806 (1999).
42. Gest, H. & Favinger, J. L. *Heliobacterium chlorum* gen. nov. sp. nov., an anoxygenic brownish-green photosynthetic bacterium containing a new form of bacteriochlorophyll. *Arch. Microbiol.* **136,** 11–16 (1983).
43. Raymond, J., Zhaxybayeva, O., Gogarten, J. P., Gerdes, S. Y. & Blankenship, R. E. Whole-genome analysis of photosynthetic prokaryotes. *Science* **298,** 1616–1620 (2002).
44. Boone, D. R., Castenholz, R. W., Johnson, R. L. & Liu, Y. Diffusion of the interspecies electron carriers $H_2$ and formate in methanogenic ecosystems and its implications in the measurement of $K_m$ for $H_2$ or formate uptake. *Appl. Environ. Microbiol.* **55,** 1735–1741 (1989).
45. Overmann, J. & Schubert, K. Phototropic consortia: model systems for symbiotic interrelations between prokaryotes. *Arch. Microbiol.* **177,** 201–208 (2002).
46. Tyson, G. W. *et al.* Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* **428,** 37–43 (2004).
47. Margulis, L. in *Genomes in Flux* (eds Gogarten, M. Olendzenski, L. & Gogarten, J. P.) (Humana, in the press).
48. Fischer, W. W. Life before the rise of oxygen. *Nature* **455,** 1051–1052 (2008).
49. Rasmussen, B., Fletcher, I. R., Brocks, J. J. & Kilburn, M. R. Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature* **455,** 1101–1104 (2008).
50. Ventura, G. T. *et al.* Molecular evidence of Late Archean archaea and the presence of a subsurface hydrothermal biosphere. *Proc. Natl Acad. Sci. USA* **104,** 14260–14265 (2007).
51. Servin, J. A., Herbold, C. W., Skophammer, R. G. & Lake, J. A. Evidence excluding the root of the tree of life from the Actinobacteria. *Mol. Biol. Evol.* **25,** 1–4 (2008).

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** I thank C. Herbold, R. Skophammer, A. Norman and J. Servin for advice and help. This work was supported by grants from the NSF and the UCLA NASA Astrobiology Institute to J.A.L.

**Author Information** Reprints and permissions information is available at www.nature.com/reprints. Correspondence should be addressed to J.A.L. (lake@mbi.ucla.edu).