*Molecular Phylogeny of Microorganisms*, Chapter Seven
http://www.horizonpress.com/
Please do not cite or reproduce this preprint without permission. The final version will be available in 2010 via the Horizon website.

# Construction and deconstruction: The influence of lateral gene transfer on the evolution of the Tree of Life

Maureen O'Malley, Egenis, University of Exeter, UK
M.A.O'Malley@exeter.ac.uk

## Abstract

Efforts to construct the tree of life take their conceptual motivation from Charles Darwin's theory of evolution. Until the advent of molecular biology, however, a universal tree of life was well beyond the scope of the data and methods of traditional organismal phylogeny. The rapid development of these methods and bodies of genetic sequence from the 1970s onwards resulted in major reclassifications of life and revived ambitions to represent all organismal lineages by one true tree of life. Subsequent realization of the significance of lateral gene transfer and other non-vertical processes has subtly reconceptualized and reoriented attempts to construct this universal phylogeny. This chapter sets out these shifts of construction, deconstruction and reconstruction, with an eye towards understanding the future of the tree of life.

## Outline

## Introduction

> 'Old prejudices tend to inhibit, distort, or otherwise shape new ideas, and historical analysis helps to eliminate much of the negative impact of the status quo' (Woese, 1987: 222).

The Tree of Life is a powerful symbol of the unity of evolutionary process and pattern. From branches of vertical descent emerge species bifurcations, which go on to further bifurcate or end in extinction. Proposed by Charles Darwin as both the phenomenon to be explained by evolutionary theory, as well as proof of evolution by natural selection (Doolittle and Bapteste, 2007; Doolittle, 2009a), the Tree of Life for today's evolutionary biologists is both a fact and a logical necessity (e.g., Cracraft and Donaghue, 2004; www.tolweb.org; Eldredge, 2005).

Although its history has in fact much deeper roots than Darwin (Ragan et al., 2009; Archibald, 2009; Pallen, 2009), there is little in-depth examination of what such a Tree has meant to the communities that have employed it. One scientific area in which the metaphor of the Tree of Life has reemerged and been closely examined is the triumphant molecular microbial phylogeny of the last few decades. This chapter will outline this recent history, examine why a Tree was so central to the three-domain proposal of life, and why challenges to such a Tree structure are so vigorously contested right up to the present day. The analysis will conclude with an outline of the future prospects of gaining knowledge of evolutionary history through the Tree metaphor.

## 1. Darwin as a basis for the Tree of Life

Anyone who has thought at all about the Tree of Life, and everyone who has examined Darwin's texts, has taken note of Darwin's sole diagram in *On the Origin of Species* and its accompanying interpretation.

> 'The affinities of all the beings of the same class have sometimes been represented by a great tree. I believe this simile largely speaks the truth. The green and budding twigs may represent existing species; and those produced during each former year may represent the long succession of extinct species' (1859: 120).

Darwin expanded on this metaphor in the 6th edition of the *Origin*:

> 'As limbs give rise by growth to fresh buds, and these, if vigorous, branch out and overtop on all sides many a feebler branch, so by generation I believe it has been with the great Tree of Life, which fills with its dead and broken branches the crust of the earth, and covers the surface with its ever branching and beautiful ramifications'
> (1872: 171-2)

For biologists, the Tree did not provide a central organizing metaphor immediately after Darwin, despite or perhaps because of Ernst Haeckel's attempts to draw detailed trees of life shortly after Darwin (Haeckel, 1866; Dayrat, 2003). Haeckel, known for his unorthodox appropriation of Darwinian thinking, may have tainted the legitimacy of endeavours to understand the history of all evolution in one general mapping process. Moreover, the terminology of 'tree of life' had strong religious connotations, both from Biblical references and a variety of other cultural sources (Hacking, 2007). Classification of the time was still largely taxonomical as opposed to evolutionary, and branching patterns, if they were suggested, were derived straightforwardly from existing taxonomical schema (Stevens, 1984; de Queiroz, 1988; Mayr, 1942). 'Our phylogenies are invented to account for our taxonomic facts or theories', complained botanist Harry Allan, as he discussed the 'new systematics' of the 1940s (Allan, 1940).

All of this would change with the advent of contemporary phylogenetic methods, especially as formalized by cladism in the 1960s and 70s. But these transformations applied primarily to animals, plants and occasionally fungi (Hennig, 1966; Mayr, 1982; Futuyma, 2004). Unicellular organisms known as bacteria, and even unicellular eukaryotes, were still difficult to classify, let alone to make them divulge their evolutionary histories. Nevertheless, influential microbiologists of the 1940s and 50s saw the development of a ‘natural’ evolution-based classification system as an imperative for the scientific advancement of microbiology (Stanier and van Niel, 1941; Sapp, 2009). But even a sketch of the Tree of Life, necessarily rooted in the microbial world, was beyond the grasp of the methods and means of data collection until well into the second half of the twentieth century (Fernholm et al., 1989). And for most botanists and zoologists, incorporating the evolution of microbes into a universal representation of speciating lineages was not a pressing or relevant task: trees of angiosperms or arthropods were demanding enough. More general talk of ‘The Tree of Life’ thus fell outside the disciplinary commitments of most evolutionary biologists and phylogeneticists.

## 2. Constructing the Tree of Life

The ambition to represent, at least schematically, the evolutionary relationships of all organismal lineages found its succour in the 1970s. Despite the limited integration of microbiology and evolutionary biology, and the lower institutional status of evolutionary understandings of microbes than of pandas and orchids, the unifying urge spread *from* microbiology into the rest of biology. In some respects this should not be surprising, because it is undeniable that deep phylogeny will always be concerned with microbes. But most zoologists and botanists had little compulsion or ability to do such deep phylogeny, even if they made vague speculations about basal eukaryotes or prokaryotes or their properties. One such example comes from the work of ornithologist and co-architect of the modern synthesis of evolution, Ernst Mayr. In his arguments about evolutionary phylogeny and the nature of species, he frequently claimed that the original organisms on the earth must have been sexual reproducers, and that asexuality was therefore a derived and not a primitive condition (Mayr, 1963). He could only make this argument theoretically, however. What the Tree of Life shift required was a method for substantiating such speculations, and that method was found in molecular approaches to phylogeny.

Emile Zuckerkandl and Linus Pauling provided the rationale for such methods by arguing and demonstrating the efficacy of using molecular sequences as repositories of evolutionary records. Building on several earlier efforts to construct animal phylogenies from protein and antiserum data, Zuckerkandl and Pauling declared that amino acid changes could serve as a molecular clock, and that comparison of such changes across lineages would allow accurate and objective measurements of evolutionary distance (1965; Pauling and Zuckerkandl, 1963; Zuckerkandl, 1987). They thus amplified the suggestions of earlier molecular biologists, such as Frederick Sanger, Francis Crick, and Emanuel Margoliash, who had anticipated the creation of sequence-based taxonomies and the evolutionary interpretations such taxonomies would enable (Harris et al., 1956; Crick, 1958; Margoliash, 1963; Sibley, 1962). In the same paper that outlined his early views on the sequence hypothesis and central dogma, Crick foretold that,

> ‘before long we shall have a subject that might be called “protein taxonomy” – the study of the amino acid sequences of the proteins of an organism and the comparison of them between species. It can be argued that these sequences are the most delicate expression possible of the phenotype of an organism, and that vast amounts of evolutionary information may be hidden away within them’ (1958: 142).

But in even more fundamentally transformative ways, Margaret Dayhoff and colleagues opened up the possibility of constructing ‘a biologically comprehensive phylogenetic tree’ (Dayhoff and Schwartz, 1981: 92; Dayhoff et al., 1974). They provided a nucleic acid and protein sequence database, which they advertised with exhortations to combine multiple gene analyses through computational methods. The ‘phylogenetic tree of all life’ (Dayhoff et al., 1974: 325) must be a ‘composite tree’, argued Robert Schwartz and Dayhoff (1978: 397), because single gene trees could not be expected to depict fully major evolutionary events and relationships between many different lineages. A composite tree could sketch out an evolutionary framework that could be constantly expanded with new sequence data, they argued, and this could be collectively compiled by the sequencing community.

Carl Woese, who was developing his own database of oligonucleotide sequences, took this advice to heart but came ultimately to rely on particular macromolecules, small subunit ribosomal RNAs and their genes, as the primary determinants of evolutionary history. Everybody reading this textbook will recognize the revelatory impact Woese had on microbiology, microbial phylogenetics, and organismal classification in general. His most obvious achievement was to challenge fundamentally a previously popular scheme depicting a five-kingdom division of lifeforms that was focused on ‘modes of nutrition’ for multicellular organisms, and unicellularity for everything left over (Whittaker, 1959; 1969; Woese, 1987).[1] Woese used signature sequences followed by biochemical investigation to show that a previously undistinguished form of unicellular life was apparently different enough to constitute its own domain (a new Woesian hierarchical level above that of kingdoms): that of archaebacteria, later Archaea (Woese and Fox, 1977; Woese, 2005). While earlier it had been taken for granted that prokaryotes and eukaryotes represented the most fundamental division in cellular organization and evolutionary history (e.g., Stanier and van Niel, 1962; Mayr, 1982; Doolittle and Brown, 1994), Woese’s three-fold division of the living world immediately required a more sophisticated historical narrative and representation than could be obtained by a ‘simple to complex’ story of evolution.

---

[1] This challenge was, of course, resisted (e.g., Mayr, 1990; Margulis and Guerrero, 1991; Margulis and Schwartz, 1998), but even these attempts to hold on to the five-kingdom view of the tree of life accommodated Archaea as an important group of evolutionarily and biologically distinct organisms.

With the incorporation of Archaea as one of these primary domains of life, a new universal model of evolutionary pattern became necessary for synthesizing and reinforcing these basic divisions. Woese attempted to establish a universal phylogeny that showed the fundamental domains of life as given by Nature and her history, rather than devised by human aims and interests. He argued that Bacteria, Archaea and Eukarya constituted the 'primary tripartite division of the living world' (Woese et al., 1990: Woese, 1987). All life was included and the totality of evolutionary relationships could be more effectively unraveled once this basic division was understood, claimed Woese. Basic questions about the nature of the first organisms and major evolutionary transitions could be treated scientifically once a universal phylogeny had been constructed (Fox et al., 1980; Olsen et al., 1994). In other words, the acceptance of the three domains and domain-level classification opened up a broad outlook on how evolutionary history and life could be represented, and this was the molecular realization of 'Darwin's dream, a phylogenetic map covering all life' (Woese, 1996: 1061; Wheelis et al., 1992: 2930).

It is at this point that two transformations come together fully: the use of molecules and molecular databases to understand all evolutionary history, and a global representation of that history. But despite the comprehensive view of evolutionary relationships that had now become possible and desirable, any search of PubMed for the term 'tree of life' turns up very few instances of its use until the early 1990s (see Figure One). Woese may have been drawing on the Tree of Life as a guiding narrative in which to understand his findings, but he persistently called his own representation and goal 'the universal phylogenetic tree' or 'the universal tree' that 'encompasses all extant life'(e.g., 1987: 231; Woese et al., 1990; 2000). In one of his few uses of 'tree of life', Woese contrasted his universal phylogeny with Darwin's more timid 'genealogies' of life. 'Perhaps even Darwin dared not dream the synthesis of the great kingdoms into the universal tree of life', suggested Woese (1994a: 1). Dayhoff and colleagues likewise emphasized the comprehensiveness of such a tree, but they too shied away from using the more colloquial term of 'tree of life'.

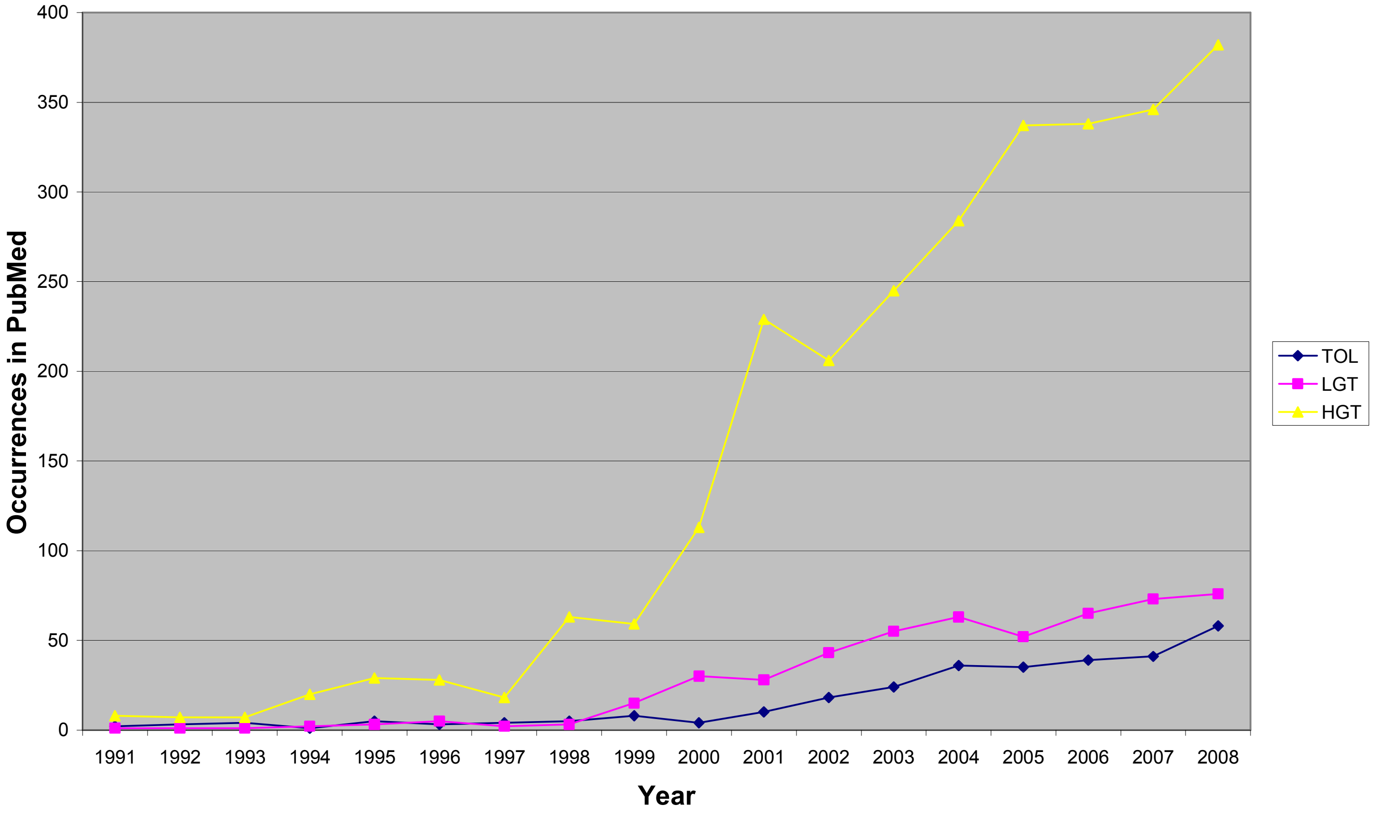

**Figure One: Occurrences of ToL, LGT and HGT in PubMed abstracts and titles, 1991-2008**
Uses of terms begin in 1982 (HGT), 1989 (ToL) and 1991 (LGT) and continue thereafter at very low rates, which are not represented on this graph until 1991. 2009 data has been excluded due to incompleteness, but indications from data gathered in mid-August, 2009, are for 2008 rates of occurrence to increase. Prior to the terminology of HGT and LGT, common expressions included ‘genetic transfer’ ‘genetic exchange’, ‘chromosome transfer’ and ‘gene transfer’ (the last the most prolific, due to being employed for animal genetics in laboratory situations). These broader terms begin in the 1950s and increase in frequency in the 1960s and 70s (and onwards), but are only represented here when prefaced by ‘horizontal’ or ‘lateral’ in relation to microorganismal genetics from 1991 onwards.
Graph constructed by Cheryl Sutton, Egenis, University of Exeter.

One of the earliest references to a universal ‘tree of life’ can be found in the work of Allan Wilson, a preeminent pioneer of molecular approaches to the interpretation of evolutionary histories.[2] Although most well known for his once controversial interpretations of molecular evidence of human and other primate evolution (e.g., Wilson and Sarich, 1969), Wilson more broadly conceived of a universal phylogeny that could accurately relate all organisms to one another, no matter the extent of their evolutionary distance or the depth of those relationships in time (Sidow and Wilson, 1990; Wilson et al., 1977). In 1990, he (along with Arend Sidow) began a tradition of using the term ‘Tree of Life’ to cover all three domains as he attempted to resolve deep branching order in the microbial origins of all life (Sidow and Wilson, 1990).

Although Wilson may have been an early advocate of the ‘tree of life’, it was Peter Gogarten and colleagues who popularized the term in the early 1990s, as they reflected on the implications of horizontal gene transfer (HGT) for a universal tree (Linkkila and Gogarten, 1991; Hilario and Gogarten, 1993[3]). Their use of ‘tree of life’ occurred very shortly after they had rooted it in accordance with Dayhoff’s suggestions for using primeval gene duplications to trace back the earliest divergences between lineages (Gogarten et al., 1989; Iwabe et al., 1989; Dayhoff and Schwartz, 1981). Gogarten comments that his choice of ‘tree of life’ terminology was not a particularly conscious or determined decision at the time. He simply thought that Woese’s term, ‘rooting the universal phylogeny of life’, had much less rhetorical and metaphorical appeal than ‘rooting the Tree of Life’ (2009, personal communication). As well as sounding more attractive, Gogarten believes that his affinity to the tree image was likely to have been influenced by his training as a botanist, during which he worked on water transport in trees.

Whatever Gogarten’s personal reasons, his imagery resonated with the existing community of microbial and other phylogeneticists. The metaphor caught on rapidly. The Darwinian affinities of constructing such trees was taken for granted, but Darwin was seldom if ever interrogated for his ideas about this metaphor until very recently (Doolittle and Bapteste, 2007). By the late-1990s, the reintroduction of ‘tree of life’ terminology to modern biology had been successfully achieved (see Figure One), as it was extensively propagated by researchers trying to piece together a global understanding of evolutionary relationships especially in regard to rooting this universal tree of life or detecting the origin of eukaryotes (e.g., Sogin, 1991; Sogin et al., 1993; Margulis and Guerrero, 1991; Forterre et al., 1993; Brown and Doolittle, 1995; Saccone et al., 1995; Baldauf et al., 1996; Gribaldo and Cammarano, 1998). Converts to the project of generating a

---

[2] Paleobiologist Michael Benton, making a commentary on the meeting ‘Major Evolutionary Radiations’, used the metaphor of ‘pruning’ the tree of life to express how diversity has shaped and pruned by adaptation, extinction and opportunism (Benton, 1989). The meeting was concerned primarily with animals and other organisms that had left paleobiological traces.

[3] This paper marks the first intersection of ToL and HGT as terms used in the same paper.

universal understanding of life were by no means exclusively microbial geneticists or microbial phylogeneticists, although the most fundamental efforts were driven by those communities (e.g., Embley et al., 1994; Forterre et al., 1992; Benachehou-Lahfa et al., 1993).[4]

But accompanying these very earliest uses of ToL were increasingly sophisticated discussions of HGT and lateral gene transfer (LGT).[5] Even the very first papers putting such terms together (e.g., Linkkila and Gogarten, 1991; Hilario and Gogarten, 1993) pointed out the implications of LGT and how it could undermine the very idea of a single universal tree. Once thought of as a process that occurred due to laboratory manipulations, microbial phylogeneticists and other biologists quickly became convinced that LGT existed and was important in natural environments, such as the human body (Anderson, 1968; Jones and Sneath, 1970; Reanney, 1977; Coughter and Stewart, 1989). Cases of cross-lineage genetic exchange rapidly gained a great deal of attention, suggested one commentator, because 'they challenge common perceptions about inheritance and the sanctity of species … [they have the] tantalizing … aura of heresy' (Sprague, 1991: 530, 531). One of the very first proposals that a tree-of-life representation of evolutionary history was not only challenged but deeply problematized was the Gogarten paper suggesting that HGT might imply a 'web of life' structure to evolutionary history rather than a tree (Hilario and Gogarten, 1993). But for most microorganismal tree builders, a discovery of gene transfer was an anomaly that might need to be reported but would not fundamentally endanger the basic tree structure (e.g., Wilson et al., 1977; Schwartz and Dayhoff, 1978; Dayhoff and Schwartz, 1981; Wheelis et al., 1992; Woese, 1987; 2000; Woese et al., 1980).

By the late 1990s a steady stream of papers on HGT and LGT had become a flood, with the frequency of the term HGT quickly overwhelming that of LGT (Figure One). But the ToL concept grew along with them, even if not at the same pace. Two publications that contributed to cementing the term 'Tree of Life' into general use were W. Ford Doolittle's influential papers that drew attention to the implications of LGT and predicted the demise of the Tree. One of them was in a popular publication, *American Scientist* (Doolittle, 2000); the other a highly cited paper in *Science* (Doolittle, 1999[6]). This key publication summed up the issues facing microbial phylogeny and suggested that the new Woesian paradigm was besieged. Its diagnosis as well as aim was the 'uprooting' of the Tree of Life. That message was reinforced by a further powerful visualization in another

---

[4] Many of the later uses of 'tree of life', especially in the 2000s, are partial, applying to particular groups of organisms (e.g., 'avian tree of life', 'angiosperm tree of life').

[5] HGT and LGT are synonymous terms, although the former is now more commonly used (see Figure One). Although there are occasional remarks made that LGT is favoured by more radical evolutionary microbiologists (O'Malley and Boucher, 2005), some of the strongest challenges to a single Tree of Life come from researchers who use HGT as their preferred term (e.g., Gogarten, 1995).

[6] This paper marks the first intersection of ToL and LGT in the same paper.

article, by Bill Martin, who also aimed to convince a broader readership of the deeply perturbing consequences of LGT (Martin, 1999). These papers and a slew of associated publications not only deconstructed in important ways the project of microbial phylogeny but also began to reconstruct it in ways that are still now setting the agenda, whatever 'side' participants are on.

## 3. The deconstruction of the Tree of Life

Throughout the 80s and 90s, and into the 2000s, the understanding of what molecules could do for phylogeny made a huge epistemological shift: from being tools which could test existing theories about organismal relationships to becoming the very source of novel hypotheses about the evolutionary history of organismal groups (Doolittle, 1996). Numerous papers, both original research and reviews, set out scenarios of rampant and promiscuous LGT (e.g., Syvanen, 1987; Doolittle and Brown, 1994; Lan and Reeves, 1996; Martin, 1999; Gogarten et al., 2002). The more sequence data that was accumulated, the less resolution to messy branching patterns they seemed to offer – quite contrary to the earlier expectations of Woese and colleagues (e.g., Woese et al., 1980). There were increasing reports of conflict between phylogenies from different genes or sets of genetic data, and between genetic data and accepted organism-based classifications. These incongruent findings indicated that different evolutionary processes, including LGT, had been at work at different levels of biology, and that rRNA trees were disappointingly inadequate for the purpose of constructing a representative tree of life.

An additional complication to LGT between species (and genera, families, phyla and even domains) came from deepening recognition of the intra-species recombination of genetic material in prokaryotes (mostly bacteria). In a number of taxa, a single prokaryote species label was found to cover a multitude of genomically differentiated strains due to the homologous recombination of acquired DNA fragments. Although these acquisitions within species are more similar than those gained from evolutionarily distant groups, the acquired and recombined DNA nevertheless further reticulates the evolutionary history of the organisms involved and refuses the straightforward mapping of the histories of genes, genomes and organisms onto one another (Spratt et al., 2001; Feil et al., 2001; Feil and Spratt, 2001; Lawrence, 2002; Lawrence and Retchless, 2009; Maynard Smith et al., 2000). As troublesome as this reticulation (inter- and intra-species) seemed with limited genetic datasets, it became even worse with the advent of genomics in the 1990s.

Genomics, which has transformed microbiology in numerous ways (Koonin, 2009; Ward and Fraser, 2005; Gogarten et al., 2009), has made extensive contributions to comparative evolutionary analyses of microbes – possibly the technology's greatest success story so far. Analyses of single microbial genomes have revealed many to have mosaic or patchwork genomes. Acquisitions of DNA from other lineages (often other domains) comprise at least 17% of the *Escherichia coli* K12 genome, 24% of the genome of the thermophilic bacterium, *Thermotoga maritima*, and 34% of the genome of the mesophilic archeon, *Methanosarcina mazei* (Lawrence and Ochman, 1998; Nelson et al., 1999 Deppenmeier et al., 2002). Recently, a detailed analysis of four Thermotogales genomes in addition to *T. maritima* has found that only a tiny proportion of the genes of these organisms has *not* been transferred at some point in their evolutionary histories (Zhaxybayeva et al., 2009a). Not only do such transfers cross large phylogenetic distances, but many make functionally crucial contributions to the lifestyle of the recipient. Furthermore, strains in some prokaryote taxa have been found to vary hugely in gene content and to have wide-ranging phenotypic differences, with more variation expected in every genome sequenced from the same taxon (Welch et al., 2002; Medini et al., 2005; Tettelin et al., 2008; Lapierre and Gogarten, 2008).

Many prokaryote to eukaryote transfers have also been detected (Andersson et al., 2006; Alsmark et al., 2009; Loftus et al., 2005; Hotopp et al., 2007), and a few eukaryote to prokaryote donations (Keeling and Palmer, 2008). There also appear to be some eukaryote to eukaryote exchanges, that have crossed small and large evolutionary distances (Busslinger et al., 1982; Lang, 1984; Syvanen, 1984;[7] Andersson, 2005; 2009; Richards et al., 2009). Ambitious LGT searches have even ventured into the human genome analyses and found – quite erroneously – bacterial transfers into the vertebrate lineage (Stanhope et al., 2001; Salzberg et al., 2001; Andersson et al., 2001). But despite some excesses, LGT and HGT findings and credibility flourished (Figure One), accompanied by growing doubts about the very project of constructing a single universal tree of life.

Central to the Tree doubters' assault is the claim that there is no universal tree, whether this is called a universal phylogeny or more dramatically The Tree of Life. Trees exist for doubters, to be sure, but as partial representations of a much more complex evolutionary process, especially in the prokaryotic world. But these LGT-based challenges to the very possibility of a tree of life came up against an entrenched opposition. A certain amount of the interest in LGT, as depicted in Figure One, was aimed at showing such findings to have been put to the service of tenuous and overinterpreted theoretical conclusions. These more conservative analyses of genetic and phylogenetic discordance consisted of a variety of attempts to save the concept of a universal tree and to develop methods that would deal effectively with incongruences. For at least one group of microbial phylogeneticists, species and the universal tree are crucial for the development of a truly Darwinian microbiology, and too many theoretical concessions to the existence of LGT are merely destructive (e.g.: Kurland et al., 2003). From the point of view of these commentators, LGT may occasionally blur vertical patterns (the tree-like structure), but its effects should not be exaggerated to the extent that the tree disappears (Snel et al., 2005). And the fact that LGT is usually detected against the background of a reference tree, particularly the 16S

---

[7] Not all of the earlier cases have withstood subsequent scrutiny.

rRNA tree (Gogarten, 1995), seemed contradictory to tree supporters, who saw no way forward without the framework a universal tree provided.

However LGT is acknowledged, the pursuit of a global phylogeny – or at least a phylogeny that imposes evolutionary order on early Archaea, Bacteria and Eukarya – is a major undertaking. Problematic in any endeavour to construct this tree is finding sufficient signal of the process of bifurcation so that it can be recovered in a biologically meaningful representation (Beiko and Ragan, 2009). All methods that attempt to tease out this pattern from LGT, gene duplication and differential loss, poor signal and phylogenetic artefact have to rationalize numerous exclusions and layers of interpretations to arrive at a tree, which may, in the end, be neither the history of any single gene nor the history of the organism itself (Swithers et al., 2009; Haggerty et al., 2009; Lawrence and Retchless, 2009). Nevertheless, it is undeniable that traces of vertical signal are often found through a variety of analyses, and the question then becomes one of whether such signal should be taken as the central truth of the evolutionary past or just one amongst several measures and representations of genetic relatedness. 'Highways' of gene exchange may, in fact, have more to say than vertical descent about major events in the evolutionary history of many groups of organisms (Beiko et al., 2005; Huang and Gogarten, 2006). LGT may be in many cases the creative force that structures organismal relationships and the patterns detected by phylogenetic analyses (Zhaxybayeva et al., 2009b). Presuming coherence and congruence to be produced only by vertical descent will result in an inadequate understanding of evolutionary history and organismal relationships.

Different approaches to preserving the tree, with different degrees of recognition of the evolutionary importance of LGT, have developed over the last decade (Ragan and Beiko, 2009; Brown, 2003; House, 2009). For conceptual convenience we can label some of the more recent efforts 'core genome' approaches (concerned with a biologically existent stable core of genes that can be taken to represent the organismal lineage) and 'central trend' approaches (focused on finding methods for weaving together vertical signals that may not agree in all of their histories). The minimal phylogenetic core approach seeks to identify genes that have a wide representation and also, that produce congruent phylogenetic signal (e.g., Lerat et al., 2003; Daubin et al., 2003). Selection of such genes is crucial. There is little doubt that groups of genes can represent particular and recent evolutionary histories – usually at the genus or family level, sometimes class – but much more scepticism that sufficient genes can be found to constitute the entire tree of life. One well known core analysis examined 191 species genomes from all three domains of life and was able to identify 31 universal genes (Ciccarelli et al., 2006). But because the total number of genes constituting each prokaryote numbers (conservatively) around 3000, the tree constructed by 31 genes is a very limited tree – 'a tree of 1%' at the most, and less if much larger eukaryote genomes are considered (Dagan and Martin, 2006).

Although a flippant response to this problem might be that phylogeneticists are very lucky to have even a small core (Gribaldo and Brochier, 2009), in general, the objection that all existing genes in prokaryote genomes have undergone at least one LGT in the 3.5 billion year history of cellular genomes means that no pure untransferred core exists, and no tree uncontaminated by LGT is constructable (McInerney et al., 2008). A very practical reinterpretation of the core approach is to relax the ubiquity requirement for the categorization of genes as ‘core’, and aim instead for very common and broadly distributed genes (Charlebois and Doolittle, 2004). This strategy has some conceptual overlap with what is sometimes called the central trend approach, in which the aim is to piece together whatever signal there is in a large body of data and see how much of it constitutes a universal tree.

A conceptually intriguing example of a central trend approach by Eugene Koonin and colleagues formulates its method as an effort to highlight vertical ‘tree’ patterns against the ‘forest’ of life (Koonin et al., 2009; Puigbò et al., 2009). This study thoroughly acknowledges the pervasiveness of LGT but nevertheless finds a central trend of vertical descent from the consensus trees of highly conserved genes. Each tree is assessed against all other trees (the forest) for inconsistency, and the most consistent trees put in a category of ‘nearly universal trees’ or NUTs. The central trend, composed of NUTs, is very faint at deep phylogenetic levels, except for the signal of bifurcation between archaea and bacteria. The overall conclusions of this analysis are that recovery of a universal tree-like structure is possible from some parts of genomes and for some part of life’s history (Puigbò et al., 2009; Koonin et al., 2009). The authors urge their study to be seen not as a failure to recover a complete Tree, but as a success. Not only has it given further reason to pursue such methodologies, but the nature of limitations of the Tree itself are now understood more clearly. Is such a tree a total representation of evolutionary history? No. Does this analysis imply that LGT events should be removed from the picture? Again, no: this is an example of an integrated approach that opens up a constructive route of ongoing inquiry and shows its current limits.

> ‘Whether or not this central trend is denoted a tree of life could be a matter of convention and convenience, but the nature of this trend as well as the other trends that can be discerned in the forest merit investigation’ (Puigbò et al., 2009: 12).

In other words, the authors leave open the ontological status of the tree (the nature of its existence), in favour of advancing its usefulness as an epistemological tool – something that generates knowledge.

## 4. What are the current state and future prospects of the Tree of Life?

If we think further about conceptions of the Tree of Life as a core or central trend, it is clear that these notions are already only conceptual cousins to the early Tree of Life conceptions, in which vertical descent by modification was supposedly revealed straightforwardly by most of the data. This basic understanding of evolution has been profoundly enriched by insights into dynamic genome-based evolutionary processes, in which a range of entities play major adaptive roles (Norman et al., 2009; Jain et al., 2003; Gogarten and Townsend, 2005; Ragan and Beiko, 2009; Brüssow, 2009; Bapteste and Boucher, 2009). But such conceptual transitions imply, according to some analyses, that the microbial phylogeny community should see the Tree of Life as a ladder, and that it should now be kicked away because it has taken the community to the top of its usefulness and can go no further (Doolittle and Bapteste, 2007). Simply clinging to it for security is not a good scientific option, runs this argument.

For many tree analysts, however, challenges to the Tree of Life from LGT have led not to its abandonment but to a different ontologically based shift in conceptualizing the tree. Rather than seeing it as a tree of *species*, based on a tree of genes or genomes, the Tree has been reinvented as a tree of *organisms*, or equivalently, a tree of cells (e.g., Woese, 2002). One reason for taking the latter option is that the history of life is in fact a history of bifurcating cell divisions and genome replication: The 'Tree of Cells' (Puigbò et al., 2009). For some modellers, a tree structure simply represents the history of cells, and it is the organismal backbone on which the web-like structure of genome evolution can be draped (Gribaldo and Brochier, 2009). 'The species tree could still [therefore] be a useful concept even if incongruent with *every* gene tree' (Daubin and Galtier, 2008 – emphasis added). What phylogeny is doing from this perspective, it is claimed, is taking the true history of organisms and not being deceived by the wayward history of some (or all) parts of the constitutive genomes (Gribaldo and Brochier, 2009: 3). The rationale seems to be that the tree of organisms is (somehow) available as a reference tree, even though any tree of cell divisions has to be interpreted as a tree of species in order to make any evolutionary sense (a tree of individual organisms would not be enough), and for most unicellular organisms the evidence is still going to be gathered almost exclusively from genomes. None of the problems of the traditional bifurcating tree of species is avoided in this conceptual reconstruction, but it constitutes an interesting appeal to the existence of evolutionarily necessary processes.

A major philosophical question that arises here is the distinction between the epistemology (or methodology) of trees and their ontology or the nature of their existence in the world. For many commentators the two are conflated: the postulated existence of the Tree means that it must be knowable. Richard Dawkins sums up this sort of conflation very aptly:

> 'For there is, after all, one true tree of life, the unique pattern of evolutionary branchings that actually happened. It exists. It is in principle knowable. We don't know it all yet. By 2050 we should – or if we do not, we shall have been defeated only at the terminal twigs, by the sheer number of species' (2003: 112).

For anyone who has thought at all about prokaryote evolution is it clear that Dawkins may be right about a tree of animals (Dagan and Martin, 2009), but he is unlikely to be right about a tree of all life (Eldredge, 2005). And even though many microbiologists might be willing to acknowledge that some evolutionary processes form a fundamentally tree-like pattern, they are just as likely to accompany this acknowledgement with questions about its *knowability* in relation to prokaryote evolution.

But Dawkins does express quite aptly the pervading background assumptions to the idea of a universal tree of life. These are echoed by more abstract statements by philosophers of evolution that Dobzhansky's famous dictum should be interpreted to read, 'Nothing in biology makes any sense except in the context of its place in phylogeny, its context in the tree of life … reconstructing that tree is critical to understanding the living world' (Sterelny and Griffiths, 1999: 379). What they are articulating is a deeply held intuitive conviction that of course all evolution involves groups arising out of groups, and that every organism should belong to one of those groups. But the claim does not address how much of this process can be known, and whether in fact it is the primary thing to be known about evolutionary processes. As Ragan and fellow authors put it,

> 'LGT is a central modality of genome evolution, and treating it purely as a distraction from vertical (parent-to-offspring) transmission hinders us from appreciating the plurality of mechanism and pattern beyond a unitary tree of life' (Ragan et al., 2009: 2171).

Not only does the conventional schema of a tree need to be accurately supplemented by LGT processes – another methodological complication (Zhaxybayeva, 2009; Poptsova, 2009) – but it also needs to accommodate other patterns from major evolutionary events, such as endosymbiosis, hybridization, co-evolving symbioses and other such instances of lineage fusion and innovation (Gogarten and Townsend, 2005; Dagan and Martin, 2006; Martin and Müller, 2007; Lake, 2009; Fournier et al., 2009; Foster et al., 2009; McInerney et al., 2008; see also Archibald and Simpson, this volume). All these processes are of major evolutionary importance, and their exclusion by strict tree conceptions would seem to be far more problematic than recognizing the 'universal' tree as a representation of some but not all evolutionary history. An additional issue is whether to separate representations and theories of eukaryote and prokaryote evolution, due to the different tempos, modes and outcomes involved (Dagan and Martin, 2009; Bapteste et al., 2009).

For many purposes, a tree of life is still a valuable ambition. It serves as a general metaphor of evolutionary relatedness, even if those relationships cannot be captured by a strictly bifurcating pattern. It accepts fundamental distinctions in cell type and physiology that are central to the evolution of life on earth. It provides a way in which to order biological knowledge for both scientific and broader social purposes (e.g., www.tolweb.org). Even some pro-Tree commentators are able to agree that any construction of a universal tree is in fact a human-made conceptual tool that is useful for *some* relationships and for heuristically imposing order on the world.

> 'In my view, a tree is just a human-made conceptual tool that we might decide to adopt if it *means* something to us, like any other graphical

> representation, irrespective of its "existence" in the real world' (Galtier, 2009).

In reaching this conventionalist viewpoint, a conceptual rapprochement becomes possible between those who would persist with the idea of a universal tree, and those who have long been arguing against the unilateral deployment of trees to represent prokaryote evolution:

> 'I have no objection to the continued use of an rRNA tree (or of any other agreed upon averaging or gene core-based TOCD&S [tree of cell division and speciation]) as a conventional framework for ***classification***, provided everyone knows that that is all that it might be, a conventional taxonomic framework, not the TOL with all its baggage. Other ways of classifying microbes (for instance by gene content or ecological role or indeed by relative position in a multidimensional network) might well have more predictive value, but still this relatively stable hierarchical scheme would serve a very useful organizing function. In fact, I think this is the posture that many microbiologists have already accepted' (Doolittle, 2009b).

This pragmatic stance, adopted by microbiologists in a variety of forms, may explain why the Tree of Life is thriving as never before despite all the challenges to its realness and epistemological legitimacy. Hundreds of biologists participate in an online attempt to understand the totality of genealogical relationships (www.tolweb.org), and as Figure One shows, the trend towards increasing citation of the term 'tree of life' continues into the end of the first decade of this millennium. Part of this popularity is inspired by the 'Darwin year' of 2009, of course (the 200th anniversary of his birth, and the 150th of the publication of the *Origin*), but much is probably due to the unifying capacity of this metaphor. We live in times of both increasingly fine-grained and high-volume data on biodiversity, which at a genetic level indicate far more complex evolutionary relationships than there are clear visualizations for. And, as science becomes more specialized and technical, the existence of a unifying metaphor for at least some biological knowledge is of deep appeal – both inside and outside science. As Woese noted, the tree of life might also unite disciplines, notably microbiology:

> 'provided at last with a phylogenetic articulating framework, microbiology can now grow to become a complete biological discipline' (Woese, 1994b: 1602).

## Conclusions

As theories and disciplines mature, the idea that a single framework is vital to success becomes increasingly questioned. From some points of view in microbial phylogeny, contemporary understandings of LGT mean that it is time, finally, to give up on any 'unifying metanarrative' such as the tree (Doolittle and Bapteste, 2007: 2048), or that it may be appropriate to find another one, such as a Web of Life (Doolittle, 2009a; Dagan and Martin, 2009) or even a Ring of Life (Lake et al., 2009). The more general question that might be raised at the end of such an overview is whether the Tree itself still relevant. Does it continue to provoke challenges, defences and an ongoing parade of valuable findings, both for and against its existence? From a practical perspective, the Tree of Life can be conceived as a central illustration of the process of scientific inquiry. As a heuristic, it has advanced understanding, despite its many (now obvious) inadequacies. And as a metaphor and theory, it functions not only to guide further investigation but to orient and integrate a multiplicity of communities and technologies. As in all enduring and powerful scientific avenues of inquiry, the pertinent question is not the abstract one of ‘Is there a tree?’, but the practical ones of ‘What is achieved by thinking about trees? What is not achieved by thinking about trees?’ As an emblem of how science works, the universal tree is not and has never been purely about wrongness and rightness, but about its practical knowledge-producing value. As long as the notion of a Tree stimulates inquiry it is valuable; the moment it is taken for granted it loses its scientific power and becomes just an everyday assumption or an unchallengeable metaphysical assertion. As a deeply contested concept, informing a battery of different approaches and a variety of biological goals, the Tree’s demise is still far off in the future. But it too has a history that has branched into different conceptualizations, methods, and representations, so the future is inevitably one of trees and not a Tree.

## Acknowledgements

Thanks to Ford Doolittle, Peter Gogarten and Bill Martin for information, insights and encouragement, and to the Leverhulme Trust for funding the Questioning the Tree of Life Network. The research for this chapter was carried out as part of the programme of Egenis, the ESRC Centre for Genomics in Society at the University of Exeter, UK.

## REFERENCES

Allan, H.H. (1940). Natural hybridization in relation to taxonomy. In The New Systematics, J. Huxley, ed. (Oxford: Clarendon), pp. 515-528.

Alsmark, U.C., Sicheritz-Ponten, T., Foster, P.G., Hirt, R.P., and Embley, T.M. (2009). Horizontal gene transfer in eukaryotic parasites: A case study of *Entamoeba histolytica* and *Trichomonas vaginalis*. In Horizontal Gene Transfer: Genomes in Flux, M.B. Gogarten, J.P. Gogarten, and L.C. Olendzenski, eds. (NY: Humana), pp. 489-500.

Anderson, E.S. (1968). The ecology of transferable drug resistance in the enterobacteria. Annu. Rev. Microbiol. 22, 131-180.

Andersson, J.O. (2005). Lateral gene transfer in eukaryotes. Cell. Mol. Life Sci. *62*, 1182-1197.

Andersson, J.O. (2009). Horizontal gene transfer between microbial eukaryotes. In Horizontal Gene Transfer: Genomes in Flux, M.B. Gogarten, J.P. Gogarten, and L.C. Olendzenski, eds. (NY: Humana), pp. 473-487.

Andersson, J.O., Doolittle, W.F., and Nesbø, C.L. (2001). Are there bugs in our genome? Science *292*, 1848-50.

Andersson, J.O., Hirt, R.P., Foster, P.G., and Roger, A.J. (2006). Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. Curr. Biol.*13*, 94-104.

Archibald, J.D. (2009). Edward Hitchcock's pre-Darwin (1840) "Tree of life." J. Hist. Biol. *42*, 561-592.

Baldauf, S.L., Palmer, J.D., and Doolittle, W.F. (1996). The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny. Proc. Natl. Acad. Sci. U.S.A. *93*, 7749-7754.

Bapteste, E., and Boucher, Y. (2009). Epistemological aspects of horizontal gene transfer on classification in microbiology. In Horizontal Gene Transfer: Genomes in Flux, M.B. Gogarten, J.P. Gogarten, and L.C. Olendzenski, eds. (NY: Humana), pp. 55-72.

Bapteste, E., O’Malley, M.A., Beiko, R.M., Ereshefsky, M., Gogarten, J.P., Franklin-Hall, L., Lapointe, F.J., Dupré, J., Dagan, T., Boucher, Y., and Martin, W. (2009). Prokaryote evolution and the tree of life are two different things. Biol. Direct 4:34 doi:10.1186/1745-6150-4-34.

Beiko, R.G., and Ragan, M.A. (2009). Untangling hybrid phylogenetic signals: Horizontal gene transfer and artifacts of phylogenetic reconstruction. In Horizontal Gene Transfer: Genomes in Flux, M.B. Gogarten, J.P. Gogarten, and L.C. Olendzenski, eds. (NY: Humana), pp. 241-256.

Beiko, R.G., Harlow, T.J., and Ragan, M.A. (2005). Highways of gene sharing in prokaryotes. Proc. Natl. Acad. Sci. U.S.A. *102*, 14332-14337.

Benachenhou-Lahfa, N., Forterre, P., and Labedan, B. (1993). Evolution of glutamate dehydrogenase genes: Evidence for two paralagous protein families and unusual branching patterns of the archaebacteria in the universal tree of life. J. Mol. Evol. *36*, 335-346.

Benton, M.J. (1989). Pruning the tree of life. Nature *342*, 129-130.

Brown, J.R. (2003). Ancient horizontal gene transfer. Nat. Rev. Genet. *4*, 121-132.

Brown, J.R., and Doolittle, W.F. (1995). Root of the universal tree of life based on an ancient aminoacyl-tRNA synthase gene duplications. Proc. Natl. Acad. Sci. U.S.A. *92*, 2441-2445.

Brüssow, H. (2009). The not so universal tree of life *or* the place of viruses in the living world. Phil. Trans. R. Soc. B *364*, 2263-2274.

Busslinger, M., Rusconi, S., and Birnstiel, M.L. (1982). An unusual evolutionary behaviour of a sea urchin histone cluster. EMBO J. *1*, 27-33.

Charlebois, R.L., and Doolittle, W.F. (2004). Computing prokaryotic gene ubiquity: Rescuing the core from extinction. Genome Res. *14*, 2469-2477.

Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C.J., Snel, B., and Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. Science 311: 1283-1287.

Coughter, J.P., and Stewart, G.J. (1989). Genetic exchange in the environment. Antonie van Leeuwenhoek *55*, 15-22.

Cracraft, J., and Donoghue, M.J. eds. 2004. Assembling the Tree of Life (NY: OUP).

Crick, F.H.C. (1958). On protein synthesis. Symp. Soc. Exp. Biol. *12*, 138-163.
Dagan, T., and Martin, W. (2006). The tree of one percent. Genome Biol. *7*, 118 doi:10.1186/gb-2006-7-10-118

Dagan, T., and Martin, W. (2009). Getting a better picture of microbial evolultion en route to a network of genomes. Phil. Trans. R. Soc. B *364*, 2187-2196.

Darwin, C. (1859). On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life, 1st ed. (London: John Murray).

Darwin, C. (1872). On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life, 6th ed. (London: John Murray).

Daubin, V., Moran, N.A., and Ochman, H. (2003). Phylogenetics and the cohesion of bacterial genomes. Science *301*, 829-832.

Dawkins, R. (2003). A Devil’s Chaplain (NY: Mariner).

Dayhoff, M.O., Barker, W.C., and McLaughlin, P.J. (1974). Inferences from protein and nucleic acid sequences: Early molecular evolution, divergence of kingdoms and rates of change. Orig. Life *5*, 311-330.

Dayhoff, M.O., and Schwartz, R.M. (1981). Evidence on the origin of eukaryotic mitochondria from protein and nucleic acid sequences. Ann. N. Y. Acad. Sci. *361*, 92-104.

Dayrat, B. (2003). The roots of phylogeny: How did Haeckel build his trees? Syst. Biol. *52*, 515-527.

Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R.A., Martinez-Arias, R., Henne, A., Wiezer, A., Bäumer, S., Jacobi, C., Brüggemann, H., Lienard, T., Christmann, A., Bömeke, M., Steckel, S., Bhattacharyya, A., Lykidis, A., Overbeek, R., Klenk, H.P., Gunsalus, R.P., Fritz, H.J., and Gottschalk, G. (2002). The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. J. Mol. Microbiol. Biotechnol. *4*, 453-61.

de Queiroz, K. (1988). Systematics and the Darwinian revolution. Phil. Sci. *55*, 238-259.

Doolittle, W.F. (1996). At the core of the Archaea. Proc. Natl. Acad. Sci. U.S.A. *93*, 8797-8799.

Doolittle, W.F. (1999). Phylogenetic classification and the universal tree. Science *284*, 2124-2128.

Doolittle, W.F. (2000). Uprooting the tree of life. Sci. Am. *282*, 90-95.

Doolittle, W.F. (2009a.) The practice of classification and the theory of evolution, and what the demise of Charles Darwin's tree of life hypothesis means for both of them. Phil. Trans. R. Soc. B *364*, 2221-2228.

Doolittle, W.F. (2009b). Response to Eric Bapteste et al., Biol. Direct 4:34 doi:10.1186/1745-6150-4-34.

Doolittle, W. F., and Bapteste, E. (2007). Pattern pluralism and the Tree of Life hypothesis. Proc. Natl. Acad. Sci. U.S.A. *104*, 2043-2049.

Doolittle, W.F., and Brown, J.R. (1994). Tempo, mode, the progenote, and the universal root. Proc. Natl. Acad. Sci. U.S.A. *91*, 6721-6728.

Eldredge, N. 2005. Darwin: Discovering the Tree of Life (NY: Norton).

Embley, T.M., Hirt, R.P., and Williams, D.M. (1994). Biodiversity at the molecular level: The domains, kingdoms and phyla of life. Phil. Trans. R. Soc. B *345*, 21-33.

Feil, E.J., and Spratt, B.G. (2001). Recombination and the population structures of bacterial pathogens. Annu. Rev. Microbiol. *55*, 561-90.

Feil, E.J., Holmes, E.C., Bessen, D.E., Chan, M.S., Day, N.P., Enright, M.C., Goldstein, R., Hood, D.W., Kalia, A., Moore, C.E., Zhou. J., and Spratt, B.G. (2001). Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. Proc. Natl. Acad. Sci. U.S.A. *98*, 182-187.

Fernholm, B., Bremer, K., and Jörnvall, H. eds. (1989). The Hierarchy of Life: Molecules and Morphologies in Phylogenetic Analysis (Amsterdam: Elsevier).

Forterre, P., Benachenhou-Lahfa, N., and Confalonieri, F. (1992). The nature of the last universal ancestor and the root of the tree of life, still open questions. Biosystems *28*, 15-32.

Forterre, P., Benachenhou-Lahfa, N., and Labedan, B. (1993). Universal tree of life. Nature *362*, 795.

Foster, P.G., Cox, C.J., and Embley, T.M. (2009). The primary divisions of life: A phylogenomic approach employing composition-heterogeneous methods. Phil. Trans. R. Soc. B *364*, 2197-2207.

Fournier, G.P., Huang, J., and Gogarten, J.P. (2009). Horizontal gene transfer from extinct and extant lineages: Biological innovation and the tree of life. Phil. Trans. R. Soc. B *364*, 2229-2239.

Fox, G.E., Stackebrandt, E., Hespell, R.B., Gibson, J., Maniloff, J., Dyer, T.A., Wolfe, R.S., Balch, W.E., Tanner, R.S., Magrum, L.J., Zablen, L.B., Blakemore, R., Gupta, R., Bonen, L., Lewis, B.J., Stahl, D.A., Luehrsen, K.R., Chen, K.N., and Woese, C.R. (1980). The phylogeny of prokaryotes. Science 209, 457-463.

Futuyma, D. (2004). The fruit of the tree: Insights into evolution and ecology. In J. Cracraft, and M.J. Donoghue, eds. Assembling the Tree of Life (NY: OUP), pp. 25-39.

Galtier, N. (2009). Comment on Bapteste et al., Biol. Direct 4:34 doi:10.1186/1745-6150-4-34.

Galtier, N., and Daubin, V. (2008). Dealing with incongruence in phylogenetic analysis. Phil. Trans. R. Soc. B *363*, 4023-4029.

Gogarten, J.P. (1995). The early evolution of cellular life. Trends Ecol. Evol. *10*, 147-151.

Gogarten, J.P., and Townsend, J.P. (2005). Horizontal gene transfer, genome innovation and evolution. Nat. Rev. Microbiol. *3*, 679-687.

Gogarten, J.P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E.J., Bowman, B.J., Manolson, M.F., Poole, R.J., Date, T., and Oshima, T. (1989). Evolution of the vacuolar $H^+$-ATPase: Implications for the origins of the eukaryotes. Proc. Natl. Acad. Sci. U.S.A. *86*, 6661-6665.

Gogarten, J.P., Doolittle, W.F., and Lawrence, J.G. (2002). Prokaryotic evolution in light of gene transfer. Mol. Biol. Evol. *19*, 2226-2238.

Gogarten, M.B., Gogarten, J.P., and Olendzenski, L.C. eds. (2009). Horizontal Gene Transfer: Genomes in Flux, (NY: Humana).

Gribaldo, S., and Brochier, C. (2009). Phylogeny of prokaryotes: Does it exist and why should we care? Res. Microbiol. epub ahead of print, doi:10.1016/j.resmic.2009.07.006

Gribaldo, S., and Cammarano, P. (1998). The root of the universal tree of life inferred from anciently duplicated genes encoding components of the protein-targeting machinery. J. Mol. Evol. *47*, 508-516.

Hacking, I. (2007). Root and branch. The Nation, October 8. http://www.thenation.com/doc/20071008/hacking

Haeckel, E. (1866). Generelle Morphologie der Organismen (Berlin: Reimer).

Haggerty, L.S., Martin, F.J., Fitzpatrick, D.A., and McInerney, J.O. (2009). Gene and genome trees conflict at many levels. Phil. Trans. R. Soc. B *364*, 2209-2219.

Harris, J.I., Sanger, F., and Naughton, M.A. (1956). Species differences in insulin. Arch. Biochem. Biophys. *65*, 427-438.

Hennig, W. (1966). Phylogenetic Systematics (Urbana, Illinois: University of Illinois Press).

Hilario, E., and Gogarten, J.P. (1993). Horizontal transfer of ATPase genes – the tree of life becomes a net of life. BioSystems *31*, 111-119.

Hotopp, J.C.D, Clark, M.E., Oliveira, D.C., Foster, J.M., Fischer, P., Torres, M.C., Giebel, J.D., Kumar, N., Ishmael, N., Wang, S., Ingram, J., Nene, R.V., Shepard, J., Tomkins, J., Richards, S., Spiro, D.J., Ghedin, E., Slatko, B.E., Tettelin, H., and Werren, J.H. (2007). Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. Science *317*, 1753-1756.

House, C.H. (2009). The tree of life viewed through the contents of genomes. In Horizontal Gene Transfer: Genomes in Flux, M.B. Gogarten, J.P. Gogarten, and L.C. Olendzenski, eds. (NY: Humana), pp. 141-161.

Huang, J., and Gogarten, J.P. (2006). Ancient horizontal gene transfer can benefit phylogenetic reconstruction. Trends Genet. *22*, 361-366.

Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., and Miyata, T. (1989). Evolutionary relationship of archaebacteria, eubacteria and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. U.S.A. *86*, 9335-9359.

Jain, R., Rivera, M.C., Moore, J.E., and Lake, J.A. (2003). Horizontal gene transfer accelerates genome innovation and evolution. Mol. Biol. Evol. *20*, 1598-1602.

Jones, D., and Sneath, P.H.A. (1970). Genetic transfer and bacterial taxonomy. Bacteriol. Rev. *34*, 40-81.

Koonin, E.V. (2009). Darwinian evolution in light of genomics. Nucleic Acids Res. *37*, 1011-1034.

Koonin, E.V., Wolf, Y.I., and Puigbò, P. (2009). The phylogenetic forest and the quest for the elusive tree of life. Cold Spring Harb. Symp. Quant. Biol. *74*, doi: 10.1101/sqb.2009.74.006

Keeling, P.J., and Palmer, J.D. (2008). Horizontal gene transfer in eukaryotic evolution. Nat. Rev. Genet. *9*, 605-618.

Kurland, C.G., Canback, B., and Berg, O.G. (2003). Horizontal gene transfer: A critical view. Proc. Natl. Acad. Sci. U.S.A. *100*, 9658-9662.

Lake, J.A. (2009). Evidence for an early prokaryotic endosymbiosis. Nature *460*, 967-971.

Lake, J.A., Skophammer, R.G., Herbold, C.W., and Servin, J.A. (2009). Genome beginnings: Rooting the tree of life. Phil. Trans. R. Soc. B *364*, 2177-2185.

Lan, R., and Reeves, P.R. (1996). Gene transfer is a major factor in bacterial evolution. Mol. Biol. Evol. *13*, 47-55.

Lang, B.F. (1984). The mitochondrial genome of the fission yeast *Schizosaccharomyces pombe*: Highly homologous introns are inserted at the same position of the otherwise less conserved cox1 genes in *Schizosaccharomyces pombe* and *Aspergillus nidulans*. EMBO J. *3*, 2120-2136.

Lapierre, P., and Gogarten, J.P. (2008). Estimating the size of the bacteria pan-genome. Trends Genet. *25*, 107-110.

Lawrence, J.G. (2002). Gene transfer in bacteria: speciation without species. Theor. Popul. Biol. *61*, 449-460.

Lawrence, J.G., and Ochman, H. (1998). Molecular archaeology of the *Escherichia coli* genome. Proc. Natl. Acad. Sci. U.S.A. *95*, 9413-9417.

Lawrence, J.G., and Retchless, A.C. (2009). The interplay of homologous recombination and horizontal gene transfer in bacterial speciation. In Horizontal Gene Transfer: Genomes in Flux, M.B. Gogarten, J.P. Gogarten, and L.C. Olendzenski, eds. (NY: Humana), pp. 29-53.

Lerat, E., Daubin, V., and Moran, N.A. (2003). From gene trees to organismal phylogeny in prokaryotes: The case of the $\gamma$-proteobacteria. PLoS Biol. *1*, 101-109.

Linkkila, T.P., and Gogarten, J.P. (1991). Tracing origins with molecular sequences: Rooting the tree of life. Trends Biochem. Sci. *16*, 287-288.

Loftus, B., Anderson, I., Davies, R., Alsmark, U.C., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J., Nozaki, T., Suh, B., Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M.A., Rabbinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guillén, N., Gilchrist, C., Stroup, S.E., Bhattacharya, S., Lohia, A., Foster, P.G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayed, N.M., Petri, W.A. Jr, Clark, C.G., Embley, T.M., Barrell, B., Fraser, C.M., and Hall, N. (2005). The genome of the protist parasite *Entamoeba histolytica*. Nature *433*, 865-868.

Margoliash, E. (1963). Primary structure and evolution of cytochrome c. Proc. Natl. Acad. Sci. U.S.A. *50*, 672-679.

Margulis, L., and Guerrero, R. (1991). Kingdoms in turmoil. New Sci. *1761*, 46-50.

Margulis, L., and Schwartz, K.V. (1998). Five Kingdoms: An Illustrated Guide to the Phyla of Life on Earth, 3rd ed. (NY: W.H. Freeman).

Martin, W. (1999). Mosaic bacterial chromosomes: a challenge en route to a tree of genomes. BioEssays *21*, 99-104.

Martin, W. (2005). Molecular evolution: lateral gene transfer and other possibilities. Heredity *94*, 565-566.

Martin, W.F., and Müller, M. (2007). Origin of Mitochondria and Hydrogenosomes (NY: Springer).

Maynard Smith, J., Feil, E.J., and Smith, N.H. (2000). Population structure and evolutionary dynamics of pathogenic bacteria. BioEssays *22*, 1115-1122.

Mayr, E. (1942). Systematics and the Origin of Species: From the Viewpoint of a Zoologist (NY: Columbia University Press).

Mayr, E. (1963). Animal Species and Evolution (Cambridge, MA: Harvard University Press).

Mayr, E. (1982). The Growth of Biological Thought: Diversity, Evolution, and Inheritance (Cambridge, MA: Harvard University Press).

Mayr, E. (1990). A natural system of organisms. Nature *348*, 491.

McInerney, J.O., Cotton, J.A., and Pisani, D. (2008). The prokaryotic tree of life: Past, present … and future? Trends Ecol. Evol. *23*, 276-281.

Medini, D., Donati, Tettelin, H., Masignani, V., and Rappuoli, R. (2005). The microbial pan-genome. Curr. Opin. Genet. Dev. *15*, 589-594.

Nelson, K.E., Clayton, R.A., Gill, S.R., Gwinn, M.L., Dodson, R,J,, Haft, D.H., Hickey, E.K., Peterson, J.D., Nelson, W.C., Ketchum, K,A,, McDonald. L,, Utterback, T.R., Malek, J.A., Linher, K.D., Garrett, M.M., Stewart, A.M., Cotton, M.D., Pratt, M.S., Phillips, C.A., Richardson, D., Heidelberg, J., Sutton, G.G., Fleischmann, R.D., Eisen, J.A., White, O., Salzberg, S.L., Smith, H.O., Venter, J.C., and Fraser, C.M. (1999). Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. Nature *399*, 323-329.

Norman, A., Hansen, L.H., and Sørensen, S.J. (2009). Conjugative plasmids: Vessels of the communal gene pool. Phil. Trans. R. Soc. B *364*, 2275-2289.

Olsen, G.J., Woese, C.R., and Overbeek, R. (1994). The winds of (evolutionary) change: Breathing new life into microbiology. J. Bacteriol. *176*, 1-6.

O'Malley, M.A., and Boucher, Y. (2005). Paradigm change in evolutionary microbiology. Stud. Hist. Philos. Biol. Biomed. Sci. *36*, 183-208.

Pallen, M. (2009). Lamarck, Darwin and the Tree of Life. http://roughguidetoevolution.blogspot.com/2009/08/lamarck-darwin-and-tree-of-life.html (accessed 18.09.09).

Pauling, L., and Zuckerkandl, E. (1963). Chemical paleogenetics: molecular ‘restoration studies’ of extinct forms of life. Acta Chem. Scand. *17*, S9-S16.

Poptsova, M. (2009). Testing phylogenetic methods to identify horizontal gene transfer. In Horizontal Gene Transfer: Genomes in Flux, M.B. Gogarten, J.P. Gogarten, and L.C. Olendzenski, eds. (NY: Humana), pp. 227-240.

Puigbò, P., Wolf, Y.I., and Koonin, E.V. (2009). Search for a ‘Tree of Life’ in the thicket of the phylogenetic forest. J. Biol. *8*, 59 doi: 10.1186/jbiol159

Ragan, M.A., and Beiko, R.G. (2009). Lateral genetic transfer: Open issues. Phil. Trans. R. Soc. B *364*, 2241-2251.

Ragan, M.A., McInerney, J.O., and Lake, J.A. (2009). The network of life: Genome beginnings and evolution. Phil. Trans. R. Soc. B. *364*, 2169-2175.

Reanney, D. (1977). Gene transfer as a mechanism of microbial evolution. BioScience *27*, 340-344.

Richards, T.A., Soanes, D.M., Foster, P.G., Leonard, G., Thornton, C.R., and Talbot, N.J. (2009). Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. Plant Cell *21*, 1897-1911.

Saccone, C., Gissi, C., Lanave, C., and Pesole, G. (1995). Molecular classification of living organisms. J. Mol. Evol*.* *40*, 273-279.

Salzberg, S.L., White, O., Peterson, J., and Eisen, J.A. (2001). Microbial genes in the human genome: lateral transfer or gene loss? Science *292*, 1903-1906.

Sapp, J. (2009). The New Foundations of Evolution: On the Tree of Life (Oxford: Oxford University Press).

Schwartz, R.M., and Dayhoff, M.O. (1978). Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts: A perspective is derived from protein and nucleic acid sequence data. Science *199*, 395-403.

Sibley, C.G. (1962). The comparative morphology of protein molecules as data for classification. Syst. Zool. *11*, 108-118.

Sidow, A., and Wilson, A.C. (1990). Compositional statistics: An improvement of evolutionary parsiomony and its application to deep branches in the tree of life. J. Mol. Evol*.* *31*, 51-68.

Snel, B., Huynen, M.A., and Dutilh, B.E. (2005). Genome trees and the nature of genome evolution. Annu. Rev. Microbiol. *59*, 191-209.

Sogin, M.L. 1991. Early evolution and the origin of eukaryotes. Curr. Opin. Genet. Dev. *1*, 457-463.

Sogin, M.L., Hinkle, G., and Leipe, D.D. (1993). Universal tree of life. Nature *362*, 795.

Sprague, G.F. Jr. (1991). Genetic exchange between kingdoms. Curr. Opin. Genet. Dev. *1*, 530-533.

Spratt, B.G., Hanage, W.P., and Feil, E.J. (2001). The relative contributions of recombination and point mutation to the diversification of bacterial clones. Curr. Opin. Microbiol. 4, 602-606.

Stanhope, M.J., Lupas, A., Italia, M.J., Kovetke, K.K., Volker, C., and Brown, J.R. (2001). Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. Nature *411*, 940-44.

Stanier, RY., and van Niel, C.B. (1941). The main outlines of bacterial classification. J. Bacteriol. *42*, 437-466.

Stanier, R.Y., and van Niel, C.B. (1962). The concept of a bacterium. Arch. Mikrobiol. *42*, 17-35.

Sterelny, K., and Griffiths, P.E. (1999). Sex and Death: An Introduction to Philosophy of Biology (Chicago: Chicago University Press).

Stevens, P.F. (1984). Metaphors and typology in the development of botanical systematics 1690-1960, or the art of putting new wine in old bottles. Taxon *33*, 169-211.

Swithers, K.S., Gogarten, J.P., and Fournier, G.P. (2009). Trees in the web of life. J. Biol. *8*:54: doi:10.1186/jbiol160

Syvanen, M. (1987). Molecular clocks and evolutionary relationships: Possible distortions due to horizontal gene flow. J. Mol. Evol. *26*, 16-23.

Syvanen, M. (1984). The evolutionary implications of mobile genetic elements. Annu. Rev. Genet. *18*, 271-293.

Tettelin, H., Riley, D., Cattuto, C., and Medini, D. (2008). Comparative genomics: The bacterial pan-genome. Curr. Opin. Microbiol. *12*, 472-477.

Ward, N., and Fraser, C.M. (2005). How genomics has affected the concept of microbiology. Curr. Opin. Microbiol. *8*, 564-571.

Welch, R.A., Burland, V., Plunkett, G., 3rd, Redford, P., Roesch, P., Rasko, D., Buckles, E.L., Liou, S.R., Boutin, A., Hackett, J., Stroud, D., Mayhew, G.F., Rose, D.J., Zhou, S., Schwartz, D.C., Perna, N.T., Mobley, H.L., Donnenberg, M.S., and Blattner, F.R. (2002). Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. Proc. Natl. Acad. Sci. U.S.A. *99*, 17020-17024.

Wheelis, M.L., Kandler, O., and Woese, C.R. (1992). On the nature of global classification. Proc. Natl. Acad. Sci. U.S.A. *89*, 2930-2934.

Whittaker, R.H. (1959). On the broad classification of organisms. Q. Rev. Biol. *34*, 210-226.

Whittaker, R.H. (1969). New concepts of kingdoms of organisms. Science *163*, 150-160.

Wilson, A.C., Carlson, S.S., and White, T.J. (1977). Biochemical evolution. Ann. Rev. Biochem. *46*, 573-639.

Wilson, A.C., and Sarich, V.M. (1969). A molecular time scale for human evolution. Proc. Natl. Acad. Sci. U.S.A. *63*, 1088-1093.

Woese, C.R. (1987). Bacterial evolution. Microbiol. Rev. *51*, 221-271.

Woese, C.R. (1994a). There must be a prokaryote somewhere: Microbiology’s search for itself. Microbiol. Rev. *58*, 1-9.

Woese, C.R. (1994b). Microbiology in transition. Proc. Natl. Acad. Sci. U.S.A. *91*, 1601-1603.

Woese, C.R. (1996). Phylogenetic trees: Whither microbiology? Curr. Biol*.* *6*, 1060-1063.

Woese, C.R. (2000). Interpreting the universal phylogenetic tree. Proc. Natl. Acad. Sci. U.S.A. *97*, 8392-8396.

Woese, C.R. (2002). On the evolution of cells. Proc. Natl. Acad. Sci. U.S.A. *99*, 8742-8747.

Woese, C.R. (2005). The archaeal concept and the world it lives in: A retrospective. In Discoveries in Photosynthesis, Govindjee, J.T. Beatty, H. Gest, and J.F. Allen, eds. (Dordrecht: Springer), pp. 1109-1120.

Woese, C.R., and Fox, G.E. (1977). Phylogenetic structure of the prokaryotic domain: The primary kingdoms. Proc. Natl. Acad. Sci. U.S.A. *74*, 5088-5090.

Woese, C.R., Gibson, J., and Fox, G.E. (1980). Do genealogical patterns in purple photosynthetic bacteria reflect interspecific gene transfer? Nature *283*, 212-214.

Woese, C.R., Kandler, O., and Wheelis, M.L. (1990). Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. U.S.A. *87*, 4576-4579.

Zhaxybaeva, O. (2009). Detection and quantitative assessment of horizontal gene transfer. In Horizontal Gene Transfer: Genomes in Flux, M.B. Gogarten, J.P. Gogarten, and L.C. Olendzenski, eds. (NY: Humana), pp. 195-213.

Zhaxybayeva, O., Swithers, K.S., Lapierre, P., Fournier, G.P., Bickhart, D.M., DeBoy, R.T., Nelson, K.E., Nesbø, C.L., Doolittle, W.F., Gogarten, J.P., and Noll, K.M. (2009a). On the chimeric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. Proc. Natl. Acad. Sci. U.S.A. *106*: 5865-5870.

Zhaxybayeva, O., Doolittle, W.F., Papke, R.T., and Gogarten, J.P. (2009b). Intertwined evolutionary histories of marine *Synechococcus* and *Prochlorococcus marinus*. Genome Biology and Evolution *1*, 325-329.

Zuckerkandl, E. (1987). On the molecular evolutionary clock. J. Mol. Evol. *26*, 34-46.

Zuckerkandl, E., and Pauling, L. (1965). Molecules as documents of evolutionary history. J. Theor. Biol. *8*, 357-366.