
Preface

From a 'Protoplasmal Primordial Atomic Globule' to a Wealth of 'Scientific Names of Beings Animalculous'

Don't mention it. I am, in point of fact, a particularly haughty and exclusive person, of pre-Adamite ancestral descent. You will understand this when I tell you that I can trace my ancestry back to a protoplasmal primordial atomic globule. Consequently, my family pride is something inconceivable.

This statement was made by Poo-Bah, the 'Lord High Everything Else' of the Town of Titipu in *The Mikado*, the comic opera with lyrics by William Schwenck Gilbert and music by Arthur Sullivan, first produced in 1885. It is a surprising statement, as in that period hardly anybody thought about the phylogeny of those ancient 'protoplasmal primordial atomic globules' that may have had some resemblance to the microorganisms found in nature today. Charles Darwin, who introduced the concepts of phylogeny and evolution to the scientific world, was indeed optimistic that some time in the future, the secrets of phylogeny will be revealed. 'The time will come, I believe, though I shall not live to see it, when we shall have very fairly true genealogical trees of each great kingdom of Nature', thus Darwin wrote in a letter to Thomas Henry Huxley (26 September 1859). It is also noteworthy that the only illustration to appear in *The Origin of Species* was a tree-like representation showing how life may have diverged from one or more common ancestors to result in the current diversity of living organisms on Earth. However, it is interesting to note that Darwin did not often mention the microbial world in his writings. Darwin surely knew about the diversity of bacteria, protozoa,

etc., known since the days when Antonie van Leeuwenhoek saw the first bacteria through the lens of this primitive microscope. The diversity of protozoa was documented in detail in the first decades of the nineteenth century by Christian Ehrenberg and others. What Darwin did not know was how genetic information was inherited, though Mendel sent him a manuscript of his work with peas. Despite the fact that Fredrick Griffith's 1928 experiments demonstrating DNA as the molecule of genetic information was clearly a case of lateral gene transfer, acceptance of this mode of inheritance as an import mechanism for the evolution of prokaryotes had to wait until the creation of inexpensive DNA sequencing and computational power. Today, lateral gene transfer is acknowledged as a relevant source of genetic variation in each domain of life, but the debate still rages as to whether phylogenetic reconstruction of the 'tree of life', especially for prokaryotes, represents organism or gene evolution, which has an enormous impact on all facets of evolution and taxonomy. Here, we explore phylogeny and its usefulness with respect to microorganisms.

The first chapter by Aharon Oren provides an historical overview of how scientists looked at the phylogeny of microbes, and especially of prokaryotic microorganisms, in the past – from the nineteenth century morphology-based schemes to schemes based on physiological properties and information on the cell wall, to the 16S rRNA gene-based phylogenetic trees used nowadays by nearly all workers in the field to establish order in the complex world of the prokaryotes. This molecular approach, pioneered by Carl Woese

in the 1970s, leading to the three-domain model (Archaea–Bacteria–Eucarya), has revolutionized our thinking about evolution in the microbial world.

The wealth of sequence information that is accumulating at an ever-increasing speed can only be used for phylogenetic reconstruction when appropriate computational tools are available to use the information to reconstruct phylogenetic trees. Chapter 2 by Jongsik Chun and Soon Gyu Hong provides an overview of the methods and programs for the calculation of phylogenetic relationships and trees. Even though we may never know the true phylogeny, phylogenetic analysis provides best assumptions, thereby providing a framework for various disciplines in microbiology. Due to the technologically innovative burst of modern molecular biology and the rapid advancement in computational science, inferring the true phylogeny of a gene or organism seems possible in the near future. The chapter reviews major steps in phylogenetic analysis and introduces relevant computer software tools with respect to their accuracy, efficiency and availability.

Chapter 3 by Pablo Vinuesa presents a review of the critical factors that have to be considered and evaluated in multilocus sequence analysis, a technique increasingly used in order to make robust estimates of bacterial species phylogenies when the 16S rRNA alone does not provide the full information. The chapter presents a number of case studies, reviews criteria for marker selection and provides practical advice on the computational aspects, potential pitfalls, and software choices available for each step in multilocus sequence analysis, emphasizing the importance of using multiple isolates per species/lineage, proper model selection and thorough tree searches to get a good estimate of a multispecies phylogeny.

In Chapter 4, Wolfgang Ludwig asks whether the small subunit rRNA sequence is still a useful marker to reconstruct the molecular phylogeny of microorganisms. The current taxonomy of prokaryotes as well as modern probe and chip based identification methods are mainly based upon rRNA derived phylogenetic conclusions. The significance of single gene based phylogenetic inference was evaluated by including alternative global markers such as elongation and initiation factors, RNA polymerase subunits, DNA gyrase,

heat shock and recA proteins. Although the comparative analyses are hampered by the generally low phylogenetic information content, a domain and prokaryotic phyla concept is globally supported.

The result of the 16S rRNA sequence based phylogenetic tree reconstruction is summarized in Chapter 5, in which Aharon Oren presents the major phyla of prokaryotes, based on cultured and named species as well as on 'environmental' sequences recovered from DNA isolated from different environments. To obtain the organisms harbouring these sequences and to study their properties is a major challenge of microbiology today. This type of classification, in which the now over 8200 species of prokaryotes with validly published names are classified in respectively 27 phyla of Bacteria and two phyla of Archaea (as of November 2009), was adopted in the latest edition of *Bergey's Manual of Systematic Bacteriology*. It should, however, be stressed that there is no official classification of prokaryotes. For the higher taxa there even is no official nomenclature: the rules of the International Code of Nomenclature of Prokaryotes do not cover taxa above the rank of class. Alternative classifications have been proposed, based, for example, on the structure of the cell wall.

The approaches towards the rooting of the universal tree of life are discussed in Chapter 6 by Greg Fournier. Identifying the location of the root of the tree corresponding to the most recent common ancestor is a challenging and distinct problem that has yet to be completely solved. To date, many investigations have proposed various roots, using a wide diversity of biological data and techniques. A survey of the most promising of these models illustrates the difficulty faced in reaching a scientific consensus on the issue, as well as the additional philosophical complications posed by our emerging understanding of the role of horizontal gene transfer in genome evolution.

Comparative analysis of genome sequences is leading to discovery of large numbers of novel molecular markers that are proving very helpful in understanding many important aspects of microbial phylogeny. Of these molecular markers, conserved inserts or deletions (indels) in protein sequences provide a means for identifying different groups of microbes in molecular terms and for

understanding how they have may have evolved from a putative common ancestor. Chapter 7 by Radhey Gupta discusses the applications of conserved indels for analysing microbial phylogeny, and shows how genetic and biochemical studies of these markers should also lead to identification of novel properties that are unique to different groups of microbes.

Elucidation of an ever increasing number of genome sequences of Bacteria and Archaea shows how important lateral gene transfer has been in the evolution of the prokaryotic world. Realization of the significance of lateral gene transfer and other non-vertical processes has reconceptualized and reoriented attempts to construct the universal phylogeny. Chapter 8 by Maureen O'Malley discusses the occurrence of paradigm shifts as construction, deconstruction and reconstruction of the tree of life changed, based on our increasing understanding of lateral gene transfer phenomena.

The impact of gene transfer on the formation of groups of organisms is described in further depth in Chapter 9 by David Williams, Cheryl Andam and Peter Gogarten. Gene transfer can make it more difficult to define and determine relationships. In those cases where many genes have been transferred between partners, the majority of genes in a genome may reflect biased gene acquisition, and as a consequence, if a coherent signal is detected, one nevertheless might not be sure that the signal is due to shared organismal ancestry. Lateral gene transfer may have positive aspects as well when trying to reconstruct phylogenetic trees. The presence of a particular transferred gene was shown, in several cases, to constitute a shared derived character useful in classification. Gene transfer can put together new metabolic pathways that open up new ecological niches, and consequently, the transfer of an adaptive gene might create a new group of organisms.

In Chapter 10, finally, Christopher Lane and Dion Durnford discuss the role of endosymbiosis and the evolution of plastids. Photosynthesis is one of the most successful energy production strategies on the planet and has been co-opted numerous times throughout evolutionary history via the uptake and retention of photosynthetic

cells by non-photosynthetic eukaryotic heterotrophs. The result of this process is clear, but the mode and tempo of plastid movement among eukaryotes, particularly plastids of red algal derivation, are largely unknown. Recent changes in our understanding of the relationships between eukaryotic supergroups have only served to complicate the picture further.

The different approaches applied today to elucidate the molecular phylogeny of the prokaryotes (and of eukaryotic protists as well) have taken us a long way toward illuminating the complex world of the microbes that Linnaeus described under a single taxon, 'Chaos infusorium'. The methods available to the microbiologist today will rapidly place a newly isolated strain or an environmental small subunit rRNA gene sequence within the existing framework of the universal tree, even if some details of the exact topology of the tree may still be unknown.

The number of isolated, characterized, and named species of prokaryotic and eukaryotic microorganisms – presumably all descendants of the above-mentioned 'protoplasmal primordial atomic globule' that supposedly was the beginning of all life of Earth – is ever increasing. A proper understanding of the diversity, systematics and nomenclature of the microbes is important in many branches of biological science. Having opened this preface with a quote from a Gilbert & Sullivan comic opera, we close with another quotation from their work, this time from *The Pirates of Penzance* (1879), that shows that knowledge of microbial systematics, nomenclature and phylogeny can be of importance to many professionals, even to a Major-General:

I'm very good at integral and differential calculus;

I know the scientific names of beings animalculous:

In short, in matters vegetable, animal, and mineral,

I am the very model of a modern Major-General.

Aharon Oren and R. Thane Papke