

Abstracts of papers for the special issue on the Tree of Life *Biology and Philosophy*

Maureen A. O'Malley, William Martin, John Dupré

The Tree of Life: Introduction to an evolutionary debate

The 'Tree of Life' is intended to represent the pattern of evolutionary processes that result in bifurcating species lineages. Often justified in reference to Darwin's discussions of trees, the Tree of Life has run up against numerous challenges especially in regard to prokaryote evolution. This special issue examines scientific, historical and philosophical aspects of debates about the Tree of Life, with the aim of turning these criticisms towards a reconstruction of prokaryote phylogeny and even some aspects of the standard evolutionary understanding of eukaryotes. These discussions have arisen out of a multidisciplinary collaboration of people with an interest in the Tree of Life, and we suggest that this sort of focused engagement enables a practical understanding of the relationships between biology, philosophy and history.

Ford Doolittle

The attempt on the life of the tree of life: Science, philosophy and politics

Lateral gene transfer (LGT), the exchange of genetic information between (primarily prokaryotic) lineages, not only makes construction of a universal Tree of Life (TOL) difficult to achieve, but calls into question the utility and meaning of any result. Here I review the science of prokaryotic LGT, the philosophy of the TOL as it figured in Darwin's formulation of the Theory of Evolution, and the politics of the current debate within the discipline over how threats to the TOL should be represented outside it. We could encourage a more realistic and supportive public understanding of evolution by admitting that what we believe in is not a unified meta-theory but a versatile and well-stocked *explanatory toolkit*.

Olivier Rieppel

The series, the network and the tree: Changing metaphors of order in nature

The history of biological systematics documents a continuing tension between classifications in terms of nested hierarchies congruent with branching diagrams (the 'Tree of Life') versus reticulated relations. The recognition of conflicting character distribution led to the dissolution of the *scala naturae* into reticulated systems, which were then transformed into phylogenetic trees by the addition of a vertical axis. The cladistic revolution in systematics resulted in a representation of phylogeny as a strictly bifurcating pattern (cladogram). Due to the ubiquity of character conflict - at the genetic or morphological level, or at any level in between - some characters will necessarily have to be discarded (*qua* noise) in favor of others in support of a strictly bifurcating phylogenetic tree. Pattern analysts will seek maximal congruence in the distribution of characters (ultimately of any kind) relative to a branching tree-topology; process explainers will call such tree-topologies into question by reference to incompatible evolutionary processes. Pattern analysts will argue

that process explanations must not be brought to bear on pattern reconstruction; process explainers will insist that the reconstructed pattern requires a process explanation to become scientifically relevant, i.e., relevant to evolutionary theory. The core question driving the current debate about the adequacy of the 'Tree of Life' metaphor seems to be whether the systematic dichotomization of the living world is an adequate representation of the complex evolutionary history of global biodiversity. In 'Questioning the Tree of Life', it seems beneficial to draw at least four conceptual distinctions: *pattern reconstruction versus process explanation* as different epistemological approaches to the study of phylogeny; *open versus closed systems* as expressions of different kinds of population (species) structures; *phylogenetic trees versus cladograms* as representations of evolutionary processes versus patterns of relationships; and *genes versus species* as expressions of different levels of causal integration and evolutionary transformation.

Jim Mallet

What was Darwin's view of species?

Historians and philosophers of science agree that Darwin had an understanding of species which led to a workable theory of their origins. To Darwin species did not differ essentially from 'varieties' within species, but were distinguishable in that they had developed gaps in formerly continuous morphological variation. Similar ideas can be defended today after updating them with modern population genetics. Why then, in the 1930s and 1940s, did Dobzhansky, Mayr and others argue that Darwin failed to understand species and speciation? Mayr and Dobzhansky argued that reproductively isolated species were more distinct and 'real' than Darwin had proposed. Believing species to be inherently cohesive, Mayr inferred that speciation normally required geographic isolation, an argument that he believed, incorrectly, Darwin had failed to appreciate. Also, before the sociobiology revolution of the 1960s and 1970s, biologists often argued that traits beneficial to whole populations would spread. Reproductive isolation was thus seen as an adaptive trait to prevent disintegration of species. Finally, genetic markers did not exist, and so a presumed biological function of species, reproductive isolation, seemed to delimit cryptic species better than character-based criteria like Darwin's. Today, abundant genetic markers are available, and are widely used to delimit species, for example using assignment tests: genetics has replaced a Darwinian reliance on morphology for detecting gaps between species. In the 150th anniversary of *The Origin of Species*, we appear to be returning to more Darwinian views on species, and to a fuller appreciation of what Darwin meant.

Maureen O'Malley

Ernst Mayr, the Tree of Life and philosophy of biology

Ernst Mayr's influence on philosophy of biology has given the field a particular perspective on evolution, phylogeny and life in general. Using debates about the tree of life as a guide, I show how Mayrian evolutionary biology excludes numerous forms of life and many important evolutionary processes. Hybridization and lateral gene transfer are two of these processes, and they occur frequently, with important outcomes in all domains of life. Eukaryotes appear to have a more tree-like history because successful lateral events tend

to occur among more closely related species, or at a lower frequency, than in prokaryotes, but this is a difference of degree rather than kind. Although the tree of life is especially problematic as a representation of the evolutionary history of prokaryotes, it can function more generally as an illustration of the limitations of a standard evolutionary perspective. Moreover, for philosophers, questions about the tree of life can be applied to the Mayrian inheritance in philosophy of biology. These questions make clear that the dichotomy of life Mayr suggested is based on too narrow a perspective. An alternative to this dichotomy is a multidimensional continuum in which different strategies of genetic exchange bestow greater adaptiveness and evolvability on prokaryotes and eukaryotes.

Marc Ereshefsky

Microbiology and the species problem

This paper examines the species problem in microbiology and its implications for the species problem more generally. Given the different meanings of 'species' in microbiology, the use of 'species' in biology is more multifarious and problematic than commonly recognized. So much so, that recent work in microbial systematics casts doubt on the existence of a prokaryote species category in nature. It also casts doubt on the existence of a general species category for all of life (one that includes both prokaryotes and eukaryotes). Prokaryote biology also undermines recent attempts to save the species category, such as the suggestion that species are metapopulation lineages and the idea that 'species' is a family resemblance concept.

Jeffrey Lawrence & Adam Retchless

The myth of bacterial species and speciation

The Tree of Life hypothesis frames the evolutionary process as a series of events whereby lineages diverge from one another, thus creating the diversity of life as descendent lineages modify properties from their ancestors. This hypothesis is under scrutiny due to the strong evidence for lateral gene transfer between distantly-related bacterial taxa, thereby providing extant taxa with more than one parent. As a result, one argues, the Tree of Life becomes confounded as the original branching structure is gradually superseded by reticulation, ultimately losing its ability to serve as a model for bacterial evolution. Here we address a more fundamental issue: is there a Tree of Life that results from bacterial evolution without considering such lateral gene transfers? Unlike eukaryotic speciation events, lineage separation in bacteria is a gradual process that occurs over tens of millions of years, whereby genetic isolation is established on a gene-by-gene basis. As a result, groups of closely-related bacteria, while showing robust genetic isolation as extant lineages, were not created by an unambiguous series of lineage-splitting events. Rather, a temporal fragmentation of the speciation process results in cognate genes showing different genetic relationships. We argue that lineage divergence in bacteria does not produce a tree-like framework, and inferences drawn from such a framework have the potential to be incorrect and misleading. Therefore, the Tree of Life is an inappropriate paradigm for bacterial evolution regardless of the extent of gene transfer between distantly related taxa.

Cheryl Andam, David Williams and Peter Gogarten

Natural taxonomy in light of horizontal gene transfer

We discuss the impact of horizontal gene transfer (HGT) on phylogenetic reconstruction and taxonomy. We review the power of HGT as a creative force in assembling new metabolic pathways, and we discuss the impact that HGT has on phylogenetic reconstruction. On one hand, shared derived characters are created through transferred genes that persist in the recipient lineage, either because they were adaptive in the recipient lineage or because they resulted in a functional replacement. On the other hand, taxonomic patterns in microbial phylogenies might also be created through biased gene transfer. The agreement between different molecular phylogenies has encouraged interpretation of the consensus signal as reflecting organismal history or as the tree of cell divisions; however, to date the extent to which the consensus reflects shared organismal ancestry and to which it reflects highways of gene sharing and biased gene transfer remains an open question. Preferential patterns of gene exchange act as a homogenizing force in creating and maintaining microbial groups, generating taxonomic patterns that are indistinguishable to those created by shared ancestry. To understand the evolution of higher bacterial taxonomic units, concepts usually applied in population genetics need to be applied.

Greg Morgan

Evaluating Maclaurin and Sterelny's conception of biodiversity in cases of frequent, promiscuous lateral gene transfer

The recent conception of biodiversity proposed by James Maclaurin and Sterelny was developed mostly with macrobiological life in mind. They suggest that we measure biodiversity by dividing life into natural units (typically species) and quantifying the differences among units using phenetic rather than phylogenetic measures of distance. They identify problems in implementing quantitative phylogenetic notions of difference for non-prokaryote species. I suggest that if we focus on microbiological life forms that engage in frequent, promiscuous lateral gene transfer (LGT), and their associated reticulated phylogenies, we need to rethink the notion of species as the natural unit, and we discover additional problems with phylogenetic notions of distance. These problems suggest that a phenetic approach based on morphospaces has just as much appeal, if not more, for microbes as they do for multi-cellular life. Facts about LGT, however, offer no new insight into the additional challenge of reconciling units and differences into a single measure of biodiversity.

Frédéric Bouchard

Symbiosis, lateral function transfer, and the (many) saplings of life

One of intuitions driving the acceptance of a neat structured tree of life is the assumption that organisms and the lineages they form have somewhat stable spatial and temporal boundaries. The phenomenon of symbiosis shows us that such 'fixist' assumptions does not correspond to how the natural world actually works. The implications of lateral gene transfer (LGT) have been discussed elsewhere; I wish to stress a related point. I will focus on lateral function transfer (LFT) and will argue, using examples of what many would call 'superorganisms', that the emergence of symbiotic individuals revives the

importance of functional and adaptationist thinking in how we conceptualize the lineages of biological individuals. The consequence of the argument is that, if we really want to hold on to tree of life thinking, we had better accept that new saplings appear and disappear all the time.

Christophe Malaterre

Lifeness signatures and the roots of the tree of life

Do trees of life have roots? What do these roots look like? In this contribution, I argue that research on the origins of life might offer glimpses on the topology of these very roots. More specifically, I argue (1) that the roots of the tree of life go well below the level of the commonly mentioned 'ancestral organisms' down into the level of much simpler, minimally living entities that might be referred to as 'protoliving systems', and (2) that further below, one finds a system of roots that gradually dissolve into non-living matter along several functional dimensions. In between non-living and living matter, one finds physico-chemical systems that I propose to characterize by a 'lifeness signature'. In turn, this 'lifeness signature' might also account for a diverse range of biochemical entities that are found to be 'less-than-living' yet 'more-than-non-living'.

Rob Beiko

Gene sharing and genome evolution: Networks in trees and trees in networks

Frequent lateral genetic transfer undermines the existence of a unique "tree of life" that relates all organisms. Vertical inheritance is nonetheless of vital interest in the study of microbial evolution, and knowing the "tree of cells" can yield insights into ecological continuity, the rates of change of different cellular characters, and the evolutionary plasticity of genomes. Notwithstanding within-species recombination, the relationships most frequently recovered from genomic data at shallow to moderate taxonomic depths are likely to reflect cellular inheritance. At the same time, it is clear that several types of 'average signals' from whole genomes can be highly misleading, and the existence of a central tendency must not be taken as *prima facie* evidence of vertical descent. Phylogenetic networks offer an attractive solution, since they can be formulated in ways that mitigate the misleading aspects of hybrid evolutionary signals in genomes. But the connections in a network typically show genetic relatedness without distinguishing between vertical and lateral inheritance of genetic material. The solution may lie in a compromise between strict tree-thinking and network paradigms: build a phylogenetic network, but identify the set of connections in the network that are potentially due to vertical descent. Even if a single tree cannot be unambiguously identified, choosing a subnetwork of putative vertical connections can still lead to drastic reductions in the set of candidate vertical hypotheses.

Joel Velasco and Elliott Sober

Testing for treeness: Lateral gene transfer, phylogenetic inference, and model selection

A phylogeny that allows for lateral gene transfer (LGT) can be thought of as a strictly branching tree (all of whose branches are vertical) to which lateral branches have been added. Given that the goal of phylogenetics is to depict

evolutionary history, we should look for the best supported phylogenetic network and not restrict ourselves just to trees. However, the obvious extensions of popular tree-based methods such as Maximum Parsimony and Maximum Likelihood face a serious problem – if we judge networks by fit to data alone, networks that have lots of lateral branches will always fit the data at least as well as any network that restricts itself to vertical branches. This is analogous to the well-studied problem of overfitting data in the curve-fitting problem. Analogous problems often have analogous solutions and we propose to treat network inference as a case of model selection and use the Akaike Information Criterion (AIC). Strictly tree-like networks are more parsimonious than those that postulate lateral as well as vertical branches. This leads to the conclusion that we should not always infer LGT events whenever it would improve our fit-to-data, but should do so only when the improved fit is larger than the penalty for adding extra lateral branches.

Laura Franklin-Hall

Trashing the tree: Bad reasons and good reasons

Evidence of extensive lateral gene transfer among prokaryotes has lead many biologists to conclude that the history of prokaryotic species cannot be represented using a tree graph, such as the one traditionally considered to be a large part of the Tree of Life. Some researchers have suggested in response that the Tree of Life might represent something slightly different, such as the history of cellular lineages, called the Tree of Cells. This paper examines and evaluates reasons offered against this view of the Tree of Life. It argues that some such reasons are bad reasons, based either on a false attribution of essentialism, on a misunderstanding of the problem of lineage identity, or on a limited view of scientific representation. I suggest that debate about the Tree of Cells and other successors to the traditional Tree of Life should be formulated in terms of the purposes these representations may serve. In pursuing this strategy, we see that the Tree of Cells cannot serve one purpose suggested for it: as an explanation for the hierarchical nature of taxonomy. We then explore whether, instead, the tree may play an important role in the dynamic modeling of evolution. As highly-integrated complex systems, cells may influence which lineage components can successfully transfer into them and how they change once integrated. Only if they do in fact have a substantial role to play in this process might the Tree of Cells have some claim to be the Tree of Life.

Eric Baptiste and Dick Burian

On the need for integrative phylogenomics – and some steps toward its creation

Recently improved understanding of evolutionary processes suggests that tree-based phylogenetic analyses of evolutionary change cannot adequately explain the divergent evolutionary histories of a great many genes and gene complexes. In particular, genetic diversity in the genomes of prokaryotes, phages, and plasmids cannot be fit into classic tree-like models of evolution. These findings entail the need for fundamental reform of our understanding of molecular evolution and the need to devise alternative apparatus for integrated analysis of these genomes. We advocate the development of integrative phylogenomics for analyzing these genomes and their histories,

with tools suited to analyzing the importance of lateral gene transfer (LGT) and of DNA evolution in extra-cellular mobile genetic elements (e.g., viruses, plasmids). These phenomena greatly increase the complexity of relationships among interacting genetic partners, as they exchange functional genetic units. We examine the ontology of functional genetic units, interacting genetic partners, and emergent genetic associations, argue that these three categories of entities are required for a successful integrated phylogenomics. We conclude with arguments to suggest that the proposed new perspective and associated tools are suitable, and perhaps required, as a replacement for the bifurcating trees that have dominated evolutionary thinking for the last 150 years.