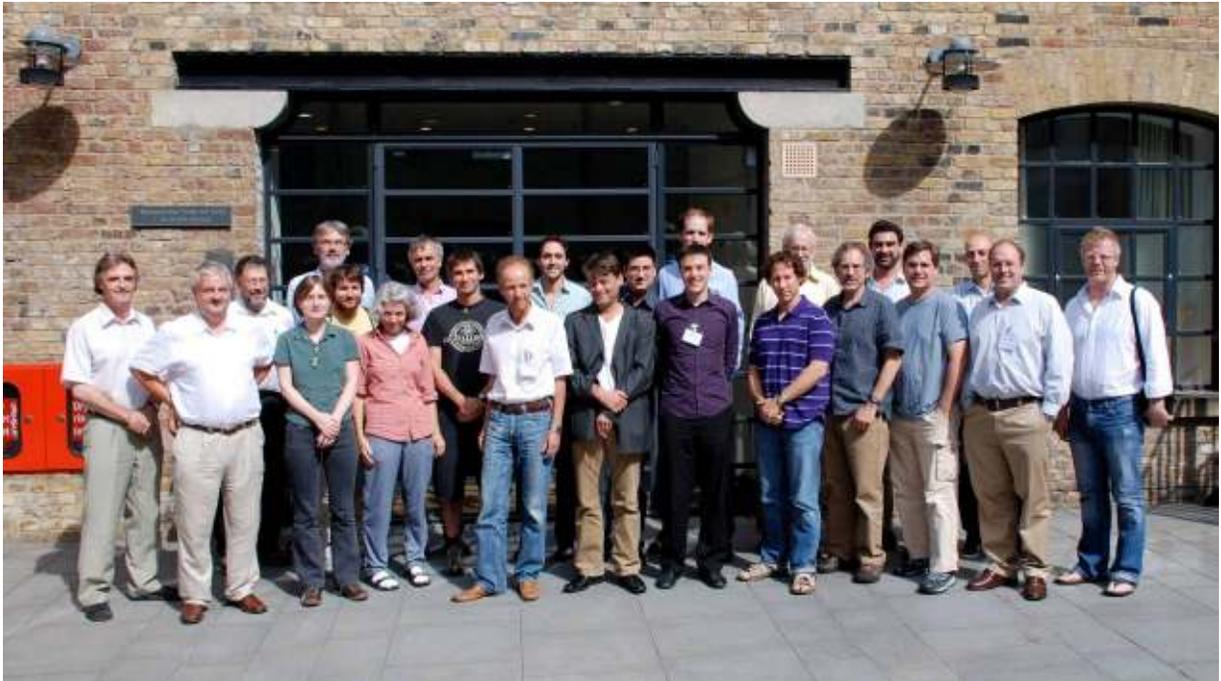


Report on the final meeting of Questioning the Tree of Life, a network sponsored by the Leverhulme Trust

Beyond the Tree of Life, London, July 10-11th, 2010



L-R: Ulrich Kutschera, John Dupré, Dick Burian, Laura Franklin-Hall, Peter Gogarten (back), Eric Bapteste, Maureen O'Malley, Jim Mallet, François-Joseph Lapointe, Eugene Koonin, Yan Boucher, Staffan Müller-Wille, Joel Velasco, Thomas Pradeu (front), Rob Beiko, Pierre-Olivier Méthot (front), Peter Stevens, Marc Ereshefsky, Nick Lane, Jeffrey Lawrence, Davide Vecchi, Greg Morgan, James McInerney. Photograph courtesy of Maria Gogarten.

The final meeting of the Questioning the Tree of Life network, funded by the Leverhulme Trust, took place in London on July 10th and 11th, 2010. The theme of this meeting was 'Beyond the Tree of Life', and its aim was to move constructively beyond criticisms of the traditional tree of life (TOL). A majority of the network was able to attend and additional participants were brought in for new perspectives. One of the most satisfying outcomes of the meeting was the lively and far-reaching discussion that followed the presentations and engaged all the disciplinary perspectives represented at the meeting. Presentations were organized into four sessions, alternating between groups of scientists followed by philosophical commentators, and groups of philosophers and historians followed by scientific commentaries or elaborations. In addition, the opening session included an artist-designer, Daisy Ginsberg, whose work on representing phylogenetically the 'synthetic kingdom of life' has gained wide recognition. Ginsberg outlined her synthetic kingdom perspective and collected numerous comments and alternative representations with which to refine her image making.¹

¹ <http://synthetickingdom.wordpress.com/2010/07/10/redesigning-the-tree-of-life/>

In the first scientific session, Peter Gogarten (Connecticut) also began with an outline of TOL imagery in which he discussed Darwin's corals and trees, contrasting them with the 'potato of life' and its problematic representation of the evolution of the three domains. He suggested an alternative 'watershed of life' in which two sources (rivers of genes) merge but do not become indistinguishable. Using the notion of 'homeoalleles' (genes that can swap in and out of lineages due to the presence of alternative genes), Gogarten showed how biased transfer was indistinguishable from shared ancestry. In this complex, non-coalescing history, he suggested that the 16S tree is in fact a better reference tree than a tree of averaged genomes, and it can be rooted by using the 'echo' in ribosomal proteins of the ancient expansion of the genetic code. 'The ribosome reflects the history of the main currents of the watershed of life', he argued, as long as it was not presented as the tree of ALL life but as the tree of 16S — an important pragmatic distinction.

Jeffrey Lawrence (Pittsburgh) presented his efforts to work out what influences the rate of gene loss and gain during the evolution of genomes. For tractability's sake, rather than focusing on prokaryotes, he examined phage genes. Because capsid size is constant, genome increase is lethal. Lawrence modelled genome divergence and found that genes do not change at the same rate, and that a power law model of gene interactions allowed predictions to be made about core genome size and exchange rates. These insights into the dynamics of genome content are now being extended to prokaryotes in order to gain an enriched and subtle understanding of genome evolution.

Jim Mallet (UCL) focused his talk on species as the units that are at the root of TOL problems, and whether claims that humans are 'hardwired' to recognize species instinctively (e.g., Carol Yoon; John Maynard Smith) can be effectively refuted. Using the notion of species as equivalence sets, Mallet showed first how eukaryote and prokaryote species might be considered to be very different (treelike and non-treelike), but that this conclusion was false once approximate equivalence and less-than-equivalence were recognized in respect to gene flow. If restricted gene flow communities are seen as equivalence relations, then LGT, introgression, and population structure violate such an assumption. Both for prokaryotes and eukaryotes there is a log-linear (i.e., continuous) decline of compatibility with genetic distance, indicating that there is no way to measure equivalence setoids and nothing special about species-level groupings. While clusters exist, argued Mallet, their phylogeny is fuzzy and the TOL with its crisp branches does not exist in either the prokaryote or eukaryote realm.

Peter Stevens (Missouri, St Louis) brought in botanical insights with his outline of plant phylogeny and its vicissitudes. Major issues persist, particularly in regard to dating ancestral divergences and even quantifying the number of taxa in a group. Representing hybridizing lineages is problematic, especially recent hybridizations. In addition, many plants function in ancient symbiotic relationships with prokaryotes and viruses, for nitrogen fixation, secondary metabolite production and toxin

biosynthesis. This coevolutionary cooperation means that the tree, with its problematic notion of individual, is a highly imperfect analogy. Furthermore, when it comes to naming branches, folk taxonomies persist because they allow communication. They cannot be dismissed on the grounds of not being supported by a common underlying evolutionary history. Ranks and groups have to be accepted simply as conventions, with no single classification representing the totality of information. The philosophy of taxonomy has thus to work around the biology of classification.

Dick Burian (Virginia Tech) summarized these four presentations and facilitated a lengthy discussion on how and why prokaryote groups were defined and identified. With more attention being paid to underlying evolutionary processes, and the interactions on which natural selection works, new tools had to be found with which to analyse such processes. Against this, concerns were raised that genealogical relationships were indeed special, and that biological entities needed to be named the same way to facilitate effective scientific inquiry. Questions of whether local systems of communication would suffice, and whether recognition that taxonomy was useful but not 'true', were reflected on relation to whether a tree is needed for such claims or not.

Marc Ereshefsky (Calgary) opened the afternoon session with a confession that he once thought the word 'species' should be abandoned. Now, however, he thought that species is an ambiguous term that does not necessitate abandonment – simply great care in its use. Ereshefsky set out the case for species pluralism and a possible consequence (suggested by Doolittle and Zhaxybayeva, 2009²) that the word species will simply 'disappear from scientific literature'. A possible replacement, suggested by Baptiste and Boucher (2008³), of coherent and composite evolutionary units, was also too ambiguous, said Ereshefsky. While it must be concluded that a final, unambiguous set of terms was unattainable and that there are no definitive species in nature, we cannot deny, he argued, that species is still a useful term that has to be disambiguated use by use in order to work.

Laura Franklin-Hall (NYU) extended the discussion of how to represent lineage evolution with an analysis of cultural, linguistic and biological evolution. These three fields constitute a group of sciences with parallel issues and problems. TOL issues are, she argued, universal to all systems of complex lineage-forming entities. While they share evolutionary processes of transfer, endogenous innovation and replication, they differ in regard to causal integration and their informational elements. Inference, representation and explanation are addressed in several similar ways by these complex lineage sciences, but for cultural and linguistic lineages, no metanarrative such as a 'Tree of Culture' is used to unite all

² Doolittle, W.F., and Zhaxybayeva, O. (2009). On the origin of prokaryotic species. *Genome Research*, 19: 744-756.

³ Baptiste, E., and Boucher, Y. (2008). Lateral gene transfer challenges principles of microbial systematics. *Trends in Microbiology*, 16: 200-207

the small trees. In regard to shared explanatory problems of vertical versus horizontal inheritance, all these sciences sought to identify degrees of integration of components, mechanisms of transfer and similarity between donor and recipient. What needs to be developed in all these sciences are the dimensions that *explain* tree-likeness.

Staffan Müller-Wille (Exeter) outlined a brief history of tree construction before focusing on Carl Linnaeus and his parody of the scale of nature. Linnaeus set out instead a scale of 'perversion' in contrast to the one of perfection. Mapping horizontal relationships was not an uncommon practice in the eighteenth and nineteenth centuries. Linnaeus's efforts to construct a classification of plants were built on a process of identifying 'real' boxes in which to place specimens and information relating to them. Müller-Wille described this as 'metaphysics in action', whereby the category exists before the essence. This movement towards a new mode of classification has a political context in which the 'corruption' of the nobility can be seen as the ideological backdrop to new representations of nature and her history.

Maureen O'Malley (Exeter) examined how the tree of life is used in scientific practice and whether the tree's persistence had anything to do with this use. Working from Doolittle and Baptiste's (2007⁴) suggestion that the TOL should be treated as a hypothesis, she outlined several ways in which the tree might be understood: as an axiom, as a hypothesis, as a myth or mistake, as a model and as a heuristic. Because of the limited benefits in seeing the TOL as the first three, she introduced Peter Godfrey-Smith's idea of models as concrete fictions, and considered the TOL as an exemplar of what he called a 'hub' fiction. Models as fictions have many problems, however, and it is tempting overall to see the tree as a heuristic. The heuristic works both by producing positive insights and systematic errors, thus producing a form of 'negative' knowledge that nevertheless advances inquiry.

François-Joseph Lapointe (Montreal), instead of commenting directly on the preceding talks, introduced a discussion of how to identify and represent treelikeness in evolutionary relationships and other entities (dendritic fields and river deltas). He argued that the dominance of monophyly biased many phylogeneticists to think about clades where they did not exist. Clades are relevant only when trees are rooted; if trees are unrooted, 'clans' are the relevant objects of analysis. The emerging endeavour to understand evolution without monophyly is called clanistics, and new technologies are needed to match the patterns clanistics identifies, such as slices and clips. This new vocabulary for analysing unrooted trees is desperately needed for more accurate discussion, he concluded, as are new metrics to compare such unrooted trees in order to identify underlying evolutionary processes.

⁴ Doolittle, W.F., and Baptiste, E. (2007). Pattern pluralism and the Tree of Life hypothesis. *Proceedings of the National Academy of Sciences USA*, 104: 2043-2049.

Extensive discussion followed each of these presentations, and continued into the final discussion session of the day. A general consensus seemed to be that the vocabulary of clanistics was necessary and useful, even though traditional talk of clades still tended to dominate where it shouldn't.

The second day of the meeting was opened by Eugene Koonin (NCBI), who set out a range of positions on the TOL and positioned his own approach as one of informed but constructive scepticism (i.e., not total rejection). He outlined a way in which the 'forest of life' (FOL) can be analysed for central trends of 'nearly universal trees' (NUTs). These NUTs, although a tiny proportion of trees in the forest, could represent a core evolutionary trend, despite such signal fading the deeper the tree reaches in evolutionary time. He showed that this analysis rejects the 'big bang' hypothesis of deepest branches in favour of a 'compressed cladogenesis' model of particular lineages; it also encourages a 'core-cloud-shell' conceptualization of genes. And finally, the NUTs also refused the strong 'Gogarten–Doolittle–Lawrence conjecture' (shortened to the Gogarten conjecture) that the appearance of trees could be caused by biased HGT. Instead, Koonin's analysis accommodated the weak form of the Gogarten conjecture, which is that HGT bias exists especially between closely related organisms. The FOL analysis indicates that the central trend is stronger than any particular route of HGT. The central trend can thus be understood as a mixture of tree signal and highways of gene sharing in which HGT reinforces the tree pattern (rather than creating it as in the strong Gogarten conjecture). These suggestions provoked considerable discussion.

Rob Beiko (Dalhousie) started out by admitting that averaging signals did not reflect either trend (vertical or horizontal) truly. Using a misleading diagram from a standard microbiology textbook (Brock, 7th edition), Beiko tackled the problems of taking one genome as a proxy for close relatives, and argued that a 'representative' genome would miss many genetic and evolutionary differences, as in the textbook diagram. However, with huge numbers of genomes, these are many problems of analysis and often methods can be chosen simply because they are faster. Taking as a maxim, 'Never compare anything to everything else', Beiko showed how his lab was making progress not by cutting corners, nor by trying to construct the perfect tree, but by pragmatic decision-making about comparability, feasibility and tree agreement. Further improvements seem possible, especially in regard to producing more comprehensible visualizations of the evolutionary relationships between large numbers of genes.

Yan Boucher (Alberta) went well beyond the traditional way of thinking about trees with his outline of gene ecology. Generally, ecology is considered to be about organisms in their environments, but he argued that the focus could be very usefully shifted to genes. Mobile genes in particular are not telling the evolutionary story of the organism, and yet they are of huge ecological importance. Not only does analysing genes ecologically tell scientists about the environment, but it also reveals how the organism manages to exist in such an environment. Boucher then set out a history of metagenomics versus that of the study of environmental single-

genes (especially 16S), and how these histories have merged. He proposed that this merger should be understood as ‘targeted metagenomics’, and that it could have considerable medical application, especially for tracking epidemics and predicting the movement of, for example, virulence factors between populations around the globe.

Nick Lane (UCL) emphasized the need to understand branch fusion and not just bifurcation in the TOL. He outlined the central role of the mitochondrion, the product of a major endosymbiotic event in evolutionary history, in producing complex life forms. He showed how larger cells need the mitochondrion to allow respiration to spread over the cell’s volume, and that ultimately, this process results in larger genome size too. Giant prokaryotes, which have genomes that are extremely polyploid but no extranuclear inheritance, did not make the sorts of evolutionary steps the mitochondrion enabled.

James McInerney (NUI, Maynooth), after an analysis of the mosaicism of halophile genomes, proposed the existence of the ‘great gene distribution anomaly’ (of rapidly evolving families with little similarity and no ‘normal’ distribution). McInerney showed how a tree-like phylogeny could not be produced, even at the tips of the branches. This finding, he suggested, required us to think of genes as public goods: they don’t care who replicates, transcribes or translates them, just as long as those processes are achieved. This biological nonchalance did not produce a purely selfish outcome, however, but one of mutual advantage for the genes and the entities (be they individual organisms or communities) who took advantage of those genes. Harking back to previous discussions of human-machine hybrids, he concluded with a quotation from Flann O’Brien’s *The Third Policeman* in which a hybridization of bicycle and human is amusingly discussed.

Eric Bapteste (UPMC, Paris) opened his talk with a discussion of bias, especially in regard to the propensity to see one particular tree instead of the forest. He urged that a broader focus on evolutionary processes would enable scientists and others to see that evolutionary units did not fit onto single branches, and that mobile elements and the genetic partnerships they formed was a more valuable sort of entity to examine. Using weighted networks instead of trees for the sake of his demonstration, he showed how such relationships could be represented to reveal a far more dynamic and richly populated account of evolutionary history. In many cases, he suggested, lifestyle was more important to genes than phylogeny, and that modules of genes ended up being traded evolutionarily as enablers of, for example, thermophilic lifestyles.

Joel Velasco (Cornell) commented on all these papers with a defence of tree thinking. Using Robert O’Hara’s claim that ‘tree-thinking’ was the right way of approaching systematics, Velasco argued that trees were special, and that trees were in many cases the necessary background against which to understand other evolutionary processes and patterns (he used serial plastid endosymbioses as the example). He claimed that this was true, despite the observations that not all evolutionary processes are tree-like and that only thinking about trees is often

misleading. Velasco urged an integrated approach that took local solutions into account, but then tried to produce something more general in order to be able to explain and predict effectively. This defence generated a lot of discussion, especially in regard to whether 'thinking historically' meant thinking with trees (because of the 'rule' this involves, of 'ignore all other parents').

John Dupré (Exeter) opened the afternoon's philosophical session on microbial evolution with an argument that seeing tree branches as the boundaries of units of selection was immensely problematic. Interactors are not bound by single branches, as several presentations had already noted in regard to gene-sharing ecologically defined communities. Instead, polygenomic systems in which branches twine and mingle are the appropriate units of selective attention, and the evolutionary trajectory of a single branch has to be understood as significantly influenced by the other branches with which it is entangled. A philosophical framework called 'Developmental Systems Theory' may be helpful to guide understanding of these entanglements, suggested Dupré.

Thomas Pradeu (Sorbonne) extended this theme by focusing on how microbes and other organisms could be understood as biological individuals through immunological analysis. An organism viewed immunologically is in fact a local concentration of evolutionary, ecological and developmental interfaces, and the TOL ignores the multiplicity of organisms involved in producing such concentrations. Using the idea of tolerance, Pradeu argued that the unit of evolution had to be understood in terms of host-symbiont relationships, and symbiosis as a major source of evolutionary innovation. This recognition, he argued, has implications for how natural selection is understood because of how symbiosis modifies understandings of the causality of selection.

Davide Vecchi (Universidad de Santiago de Chile) elaborated further on microbial evolution with a discussion of the multiple processes and dynamics of evolution at play in the microbial world. Microbial evolution, once successfully fenced off from the rest of evolutionary biology, has now exploded into it, perhaps as a consequence of all the molecular knowledge gained in the genomic era about microbial interactions and their histories. Vecchi argued for a broader model of evolution he called 'compositional' evolution. Modularity, mobile genetic elements, and generative processes (including LGT and endosymbiosis) play major roles in this model. In addition, the interaction between selective, neutral and compositional evolutionary processes need considerable theorizing, as do the relationships between 'within lineage' and 'between lineage' evolution. Michael Lynch's idea of evolution being neutral at the genomic level but adaptive at the phenotypic level could be relevant to the production of an enlarged evolutionary synthesis, suggested Vecchi.

Greg Morgan (Stevens, NJ) began his talk with an image of the *New Scientist* cover that trumpeted 'Darwin was wrong!' Morgan argued that such assertions were unethical, not so much because they were false in various ways (as the Dawkins, Dennett and Myers letter to the *New Scientist* claimed), but because of

how they were incomplete and provide potentially misleading material in important social and political debates. What Darwin was wrong about was decontextualized on the *New Scientist* cover, even though the editors cautioned inside that such claims had to be understood contextually. There are ethical obligations in regard to how information is presented, said Morgan, and journalists cannot ignore the consequences of what they do when they write up potentially inflammatory articles. Morgan stressed that this conclusion applied to science journalism, not journalism in general.

Ulrich Kutschera (Kassel) concluded this session by discussing a view in which endosymbiosis, plate tectonics and directional selection are the key elements of a revised evolutionary understanding. He argued for the need to understand the 'dynamic Earth' in order to comprehend evolution fully, and suggested that this integrated account would explain macroevolution in a five-kingdom model. This broader framework would take up issues Darwin had only hinted at, thus complementing Darwin's ideas with a broad range of knowledge gained since the publication of the *Origin*.

Following discussion of all these points, plans were made for future activities and publications. The Leverhulme Trust funding runs out this year, and this meeting is its last event. Efforts will be made to find other sources of funding to continue the



network, possibly from EU funding sources that will allow non-EU participants. The recent Society for Molecular Biology and Evolution meeting in Lyon, just before the London workshop, had a tree of life session that Bill Martin organized, and that put philosophers on the programme for the first time. Other opportunities like this will be pursued.

(Photo: After the SMBE session, courtesy of Maria Gogarten.)

Three publication plans were proposed (two in journals and the other an edited book). A plan to pursue the favoured first option (journal thematic issue) will be implemented this summer (2010).

As the organizers of this meeting, we were delighted with how well it went, and are very pleased to anticipate further publication. The Halifax 2009 meeting publication is already online (15 papers in a special issue of *Biology and Philosophy*); the new and fascinating insights produced in the London meeting also need to be shared with a wider audience. While this was the last meeting of the network, we hope very much that collaborative activities will continue in various forms, and that this interdisciplinary dialogue continues to bear intellectual fruit far into the future.

Staffan Müller-Wille and Maureen O'Malley, University of Exeter