CEB Journal Club: Andam et al. (2010)

James E. Allen

13th January 2011

Members of the Computational and Evolutionary Biology (CEB) group at the University of Manchester participate in a monthly journal club, where a paper of broad interest is discussed. Here, I briefly describe the paper and its context, and summarize our conclusions about the methodology and results presented. (I have attempted to represent the discussion and consensus of the group, but any inaccuracies are my own.)

[If referring to this document, please cite its location on the Monkeyshines website: http://www.monkeyshines.co.uk/blog/archives/347]

Biased gene transfer mimics patterns created through shared ancestry.
Cheryl P. Andam, David Williams, and J. Peter Gogarten (2010) *PNAS* 107: 23, 10679-10684. PubMed: 20495090 (Presented by James Allen at Jabez Clegg, 28th July 2010)

The paper in a sentence: The authors describe a specific case of a gene that makes a bacterial enzyme, which has been horizontally transferred between species in a biased manner, such that the molecular evidence resembles that of a gene transferred by descent from parent to offspring.

Background: Until relatively recently, genetic information was thought largely to have been transferred from parent to offspring, analogous to a branching tree structure. The applicability of this analogy for all forms of life is under debate, however, given the discovery of the extent of other mechanisms for gene transfer in bacteria and other single-celled organisms. Horizontal gene transfer (HGT) refers to the process where genetic data from one organism is transferred to another which is not necessarily related, nor even necessarily the same species; the prevalence of HGT calls into question not only the 'tree of life' metaphor (suggesting, perhaps, that a network analogy is more appropriate), but also the (already rather labile) concept of species.

The paper in detail: The authors present one key result, which is supplemented by evidence from three other sources which would not be convincing in isolation, but here provide valuable circumstantial support. The results are based on a particular enzyme, which has the important property (for this analysis) that it has two distinct types. The main result is that the tree in figure 1 in the paper, generated by looking solely at this enzyme, has two distinct sub-trees, representing each of the the two types. Each one of these sub-trees closely resembles the tree that most likely characterizes the vertical inheritance of genetic data, i.e. the 'species tree' in figure 2. It is not easy to quantify whether one tree structure resembles another, particularly with the number of species used here; the authors look at the distances along the tree branches that separate all pairs of species, which discards information about some of the tree structure, but does not prevent them from convincingly demonstrating that the sub-trees for each type resemble the species tree. Moreover, in the species tree, the species with the same type of enzyme are grouped together within broader groupings at the phylum or class level; i.e. there are patches of red and green branches (representing the two types) in figure

2. This is evidence for biased HGT because it shows that HGT occurs not in a random fashion, but more often between more closely related species.

Another line of evidence presented is that a scenario of gene gain and loss that would explain the trees is far less likely than one where some degree of HGT occurs; the authors gloss over the fact that this demonstrates that HGT, rather than biased HGT, has most likely occurred. Additionally, the genes that surround the enzyme's gene are found to be similar for both types, which would not be the case if the genes were being repeatedly gained and lost; again, this is evidence for HGT, not necessarily biased HGT.

The final piece of supporting evidence comes via simulations of biased and unbiased HGT, which result in data that resembles the real data. Some of the choices for the simulations are questionable, in particular the modelling of reciprocal transfer events, meaning that genes from two species are swapped. This does not reflect the biological reality, where the transfer generally happens in one direction only. Also, an extreme bias is modelled, using an exponential function, so that transfers are likely to occur between only the most closely related species - this may well be realistic, but the use of this particular model is not justified by the authors. Finally, the unbiased and biased transfers are simulated sequentially, which was perhaps done as it is often easier to show that something is changing, rather than staying the same, but is an uncommon approach that makes it difficult to interpret the results.

Journal club conclusion: While not wholly convinced by some of the evidence presented, particularly the approach to simulation, we believe that the main conclusions of the paper are valid: in the case of this particular enzyme, the horizontal gene transfer is biased, such that transfer is more likely between more similar species, and thus the molecular data provides the same signal as transmission through vertical inheritance. It remains to be shown how widespread this phenomenon is; if HGT generally reinforces, rather than contradicts, vertical inheritance of genetic material, then the tree of life analogy may well be useful for practical purposes, even if does not reflect the true evolutionary history.