Dating nodes in a phylogeny using inferred horizontal gene transfers

Tatiana Giraud¹ and Toni Gabaldón²

¹ Laboratoire Ecologie, Systématique et Evolution, Univ. Paris Sud, CNRS, Université Paris-Saclay – 91400 Orsay, France
² Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology -- Barcelona, Spain
Correspondence to Tatiana Giraud (tatiana.giraud@u-psud.fr)
doi: 10.24072/pci.evolbiol.100037

Cite as: Giraud T. and Gabaldón T. 2017. Dating nodes in a phylogeny using inferred horizontal gene transfers. Peer Community in Evolutionary Biology, 100037. doi: 10.24072/pci.evolbiol.100037


Dating nodes in a phylogeny is an important problem in evolution and is typically performed by using molecular clocks and fossil age estimates [1]. The manuscript by Chauve et al. [2] reports a novel method, which uses lateral gene transfers to help ordering nodes in a species tree. The idea is that a lateral gene transfer can only occur between two species living at the same time, which indirectly informs on node relative ages in a phylogeny: the donor species cannot be more recent than the recipient species. Horizontal gene transfers are increasingly recognized as frequent, even in eukaryotes, and especially in microorganisms that have little fossil records [3-7]. Yet, such an important source of information has been very rarely used so far for inferring relative node ages in phylogenies. In this context, the method by Chauve et al. [2] represents an innovative and original approach to a difficult problem. An obvious limitation of the approach is that it relies on inferences of horizontal transfers, which detection is in itself a difficult problem. Incomplete taxon sampling, or the extinction of the true donor lineage may render patterns difficult to interpret in a temporary fashion. Yet, for clades with no fossils this may be the only piece of information
we have at hand, and the growing amount of sequence data is likely to minimize issues derived from incomplete sampling.

The developed method, MaxTiC (for Maximal Time Consistency) [2], represents a very nice application of theoretical developments on the well-known « Feedback Arc Set » computer science problem to the evolutionary question of ordering nodes in a phylogeny. MaxTiC uses as input a species tree and a set of time constraints based on lateral gene transfers inferred using other softwares, and minimizes conflicts between node ordering and these time constraints. The application of MaxTiC on simulated datasets indicated that node ordering was fairly accurate [2]. MaxTiC is implemented in a freely available software, which represents original and relevant contribution to the field of evolutionary biology.

References


Appendix

Reviews by Mukul Bansal, Alexandros Stamatakis and an anonymous referee:
http://dx.doi.org/10.24072/pci.evolbiol.100037