1

Testing the neutral hypothesis of phenotypic evolution

Jianzhi Zhang[1]

1. *University of Michigan, Ann Arbor, MI, United States*

It is generally accepted that a large fraction of genomic sequence variations within and between species are neutral or nearly so. Whether the same is true for phenotypic variations is a central question in biology. On the one hand, numerous phenotypic adaptations have been documented and even Kimura, the champion of the neutral theory of molecular evolution, believed in widespread adaptive phenotypic evolution. On the other hand, phenotypic studies are strongly biased toward traits that are likely to be adaptive, contrasting genomic studies that tend to be unbiased. It is thus desirable to test the neutral hypothesis of phenotypic evolution using traits irrespective of their potential involvement in adaptation. Here we present such a test for 210 morphological traits measured in multiple strains of the yeast Saccharomyces cerevisiae and two related species. Our test is based on the premise that, under neutrality, the rate of phenotypic evolution declines as the trait becomes more important to fitness, analogous to the neutral paradigm that functional genes evolve more slowly than functionless pseudogenes. Neutrality is rejected in favor of adaptation if important traits evolve faster than less important ones, parallel to the demonstration of molecular adaptation when a functional gene evolves faster than pseudogenes. After controlling the mutational size, we find faster evolution of more important morphological traits within and between species. By contrast, an analysis of 3466 yeast gene expression traits fails to reject neutrality. Further, intraspecific and interspecific variations in yeast gene expression conform to the phylogenetic relations of the strains rather than their ecological environments. Thus, yeast morphological evolution is largely adaptive, but the same does not apply to the transcriptome, suggesting that phenotypic variations at different levels are shaped by different evolutionary forces.

2

Branch-heterogeneous models of sequence evolution

Bastien Boussau[2, 1]

1. *LBBE, CNRS, Lyon, France*
2. *LBBE, Université Lyon 1, Université de Lyon, Lyon, France*

Early models of sequence evolution made several important regularity assumptions to describe the process of nucleotide or amino-acid evolution in a simple and manageable manner. Much of the methodological work in statistical phylogenetics over the past decades has been devoted to relaxing these assumptions. As a result, nowadays the most popular models of sequence evolution assume that not all transitions between nucleotides or amino-acids are equally likely, that different sites evolve independently, but that they evolve at different rates and even sometimes according to different transition matrices. These models however still make several patently unrealistic assumptions. In particular, they assume that sequence evolution has operated according to a unique transition matrix along the entire phylogeny, an assumption that is not realistic notably for all data sets in which nucleotide or amino-acid composition varies among sequences. In this talk I will review models of sequence evolution that relax this assumption, that I will globally call branch-heterogeneous models. I will explain why these models are particularly difficult to use in either the maximum likelihood or the Bayesian frameworks, and I will present our current efforts to improve their usability.

3

Methods to find similar sites in alignments

Rob Lanfear[1], **Brett Calcott**[2], **Paul Frandsen**[3]

1. *Macquarie University, Sydney*
2. *University of Sydney, Sydney*
3. *Smithsonian Institution, Washington, D.C.*

In this talk, I will give an overview of methods for choosing partitioning schemes for the phylogenetic analysis of genome-scale datasets. This involves moving from analyses in which users provide information on the heterogeneity of evolutionary processes among sites (e.g. by defining different loci or codon positions) to analyses in which the heterogeneity is determined automatically. To do this, we have to develop methods that are both fast and free of bias. Both of these things are challenging, and despite much progress in recent years, a number of open questions remain.

4

Among genes heterogeneity of the phylogenetic signal in genome data: causes, symptoms, and treatments

Omar Rota-Stabelli[1]

1. *Fondazione Edmund Mach, San Michele All'Adige,, TN, Italy*

Inference of phylogeny is currently based on the concatenation of many genes, a procedure that enables reducing the stochasticity associated with single gene phylogenies. All possible drawbacks of this approach are however not fully understood, particularly the among gene heterogeneity of the phylogenetic signal. I studied the distribution of phylogenetic signal in the model system *Drosophila* using two genome-scaled datasets. Although both datasets *apparently* resolve most of the relationships with high support when analysed at the nucleotide level, there are at least two types of among genes phylogenetic incongruences. First, the phylogenetic signal is not homogenously distributed among nuclear coding, mitochondrial coding, and non-coding genes, which robustly support competing topologies at some nodes, particularly close to tips. Second, the phylogenetic signal is not homogenously distributed among ontology classes, whereby nuclear genes involved with the metabolism tend to carry their own signal. Most, but not all of these incongruences, are due to substitutions at synonymous sites which I show being affected by different mutational pressures in different types of data. Counter intuitively, partitioning is not successful in alleviating these incongruences, which are instead revealed by using across-site heterogeneous model or by using a coalescent aware approach. These results advocate that extra care should be taken when interpreting high supports from the analysis of genome scaled phylogenies, and that signal associated with synonymous sites may be unreliable even at the genus level. Phylogenetic incongruences may be however extremely instructive in disentangling possible sources of systematic error, as well as in revealing peculiar aspects of species biology such as introgression or incomplete lineage sorting due to fast radiation.

5