

PHYLOGENETIC GENOMICS AND GENOMIC PHYLOGENETICS

SCOTT STANLEY, BENJAMIN A. SALISBURY
*Genaisance Pharmaceuticals, 5 Science Park,
New Haven, CT 06511, USA*

Phylogenetics has a long history, with many contributions to the field actually predating Darwin. Genomics, in contrast, is barely a decade old. While these two fields may be disparate in their ages, each has much to contribute to the further development of the other. Phylogenetics provides a rich source of methodologies able to facilitate and enhance genomic research. Genome structure and function analyses can be strengthened through examining evolutionary history, whereby patterns and rates of genomic evolution can be inferred. The analysis of evolutionary history is similarly empowered by a genomic perspective, providing both a wealth of characters ideally suited for phylogenetic analysis as well as interesting subject matter in the form of comparative genomics.

This session includes three examples of research that address issues important to both phylogenetics and genomics. Zilversmit et al. provide a case study where high throughput sequencing enables phylogenetics at a genomic scale. Wang et al. propose novel phylogenetic algorithms and compare their performance to other phylogenetic algorithms designed to analyze gene order. Page and Cotton address the problem in genomic evolution imposed by gene duplication. Together, these studies address several important areas in the new field of phylogenomics.

The genomic revolution has been both enabled and defined by high throughput DNA sequencing. Advances in DNA sequencing technologies have allowed for the generation of extremely large DNA sequence data sets, many including entire genomes. Zilversmit et al. demonstrate how high throughput sequencing efforts can be applied to phylogenetic questions, using the insect family Drosophilidae as an example. They outline a method that can be applied to other non-model organisms, enabling other researchers to apply high throughput techniques to phylogenetics. They discuss how the size of data sets affects the results of phylogenetic analyses, particularly as data sets grow to include many regions from throughout the genome. In this way, Zilversmit et al. present a study that is truly phylogenetic and genomic.

Wang et al. present a novel phylogenetic method designed for gene order data, allowing researchers to conduct phylogenetic analyses using gene order data from whole or partial genomes. Gene order represents an attractive alternative to DNA sequence data, particularly when rates of DNA sequence evolution are high, or when divergence time is great. While DNA sequence data have provided an important source of characters for phylogenetic analyses in recent years, limitations associated with high rates of evolution together with the limited number of character states

available have forced researchers to look to other sources of molecular data to overcome these problems. This is particularly true when ancient divergences are involved. Gene order is one such example, and Wang et al. have added another tool to the phylogeneticists' toolbox for working on difficult phylogenetic problems. Wang et al. compare distance-based and parsimony-based methods of analyzing gene order phylogenetically. While studies of gene order have generally been restricted to genomes from organelles, it is possible to apply the methods presented by Wang et al. to nuclear genomes as well. Wang et al. found that their novel method based on a coding scheme proposed by Bryant outperformed the distance-based methods as well as the other parsimony method. Perhaps more importantly, this paper highlights the need for continued development of alternative phylogenetic methods that can be applied to something other than DNA sequence data.

Page and Cotton provide a beautiful example of the complementarity of phylogenetics and genomics. They demonstrate the use of gene tree reconciliation and a new method of mapping duplications to elucidate vertebrate phylogeny and genome evolution. Gene phylogenies may differ from each other and from species phylogenies because of complex histories of gene duplication, lineage sorting and gene loss. Page and Cotton use genomic-scale data (118 gene families) for phylogeny estimation and then map inferred gene duplications onto the resulting phylogeny to test for hypothesized genome duplications. Remarkably, they find that minimizing gene duplications produces a tree highly concordant with traditional views (where individual gene trees would resemble that phylogeny estimate poorly). Two commonly hypothesized genome duplications early in vertebrate history, however, were not evident in this analysis.

Our hope is that these papers can help further the dialogue between phylogeneticists and those in the field of genomics. Many researchers in systematics remain unaware of the importance of genomics to their own work, and vice versa. We feel that each has much to offer the other. This session is intended to spotlight those synergies and establish "phylogenomics" as a bona fide and important new field of study.

Acknowledgments

We are grateful to all the authors who submitted papers to this session and to the generous individuals who carefully reviewed them.