Disentangling Methodological and Biological Sources of Gene Tree Discordance on Oryza (Poaceae) Chromosome 3

2014

Posted by: webmaster

Derrick J. Zwickl1,5, Joshua C. Stein2, Rod A. Wing3, Doreen Ware2,4 and Michael J. Sanderson1

1Department of Ecology and Evolutionary Biology, University of Arizona, Tucson AZ 85721
2Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724
3School of Plant Sciences, University of Arizona, Tucson, AZ 85721
4Robert W. Holley Center for Agriculture and Health, United States Department of Agriculture-Agricultural Research Service, Ithaca, NY 14853

Abstract

We describe new methods for characterizing gene tree discordance in phylogenomic datasets, which screen for deviations from neutral expectations, summarize variation in statistical support among gene trees, and allow comparison of the patterns of discordance induced by various analysis choices. Using an exceptionally complete set of genome sequences for the short arm of chromosome 3 in Oryza (rice) species, we applied these methods to identify the causes and consequences of differing patterns of discordance in the sets of gene trees inferred using a panel of 20 distinct analysis pipelines. We found that discordance patterns were strongly affected by aspects of data selection, alignment, and alignment masking. Unusual patterns of discordance evident when using certain pipelines were reduced or eliminated by using alternative pipelines, suggesting that they were the product of methodological biases rather than evolutionary processes. In some cases, once such biases were eliminated, evolutionary processes such as introgression could be implicated. Additionally, patterns of gene tree discordance had significant downstream impacts on species tree inference. For example, inference from supermatrices was positively misleading when pipelines that led to biased gene trees were used. Several results may generalize to other data sets: we found that gene tree and species tree inference gave more reasonable results when intron sequence was included during sequence alignment and/or tree inference, the alignment software PRANK was used, and/or detectable "block-shift" alignment artifacts were removed. We discuss our findings in the context of well-established relationships in Oryza and continuing controversies regarding the domestication history of O. sativa.

http://sysbio.oxfordjournals.org/cont ... 09/sysbio.syu027.abstract