

Accepted Manuscript

The gene tree delusion

Mark S. Springer, John Gatesy

PII: S1055-7903(15)00222-5

DOI: <http://dx.doi.org/10.1016/j.ympev.2015.07.018>

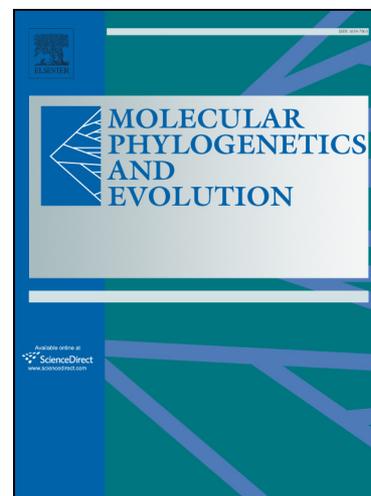
Reference: YMPEV 5253

To appear in: *Molecular Phylogenetics and Evolution*

Received Date: 19 March 2015

Revised Date: 4 June 2015

Accepted Date: 22 July 2015



Please cite this article as: Springer, M.S., Gatesy, J., The gene tree delusion, *Molecular Phylogenetics and Evolution* (2015), doi: <http://dx.doi.org/10.1016/j.ympev.2015.07.018>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

The gene tree delusion

Mark S. Springer*, John Gatesy*

Department of Biology, University of California, Riverside, CA 92521, USA

E-mail address: mark.springer@ucr.edu (M. S. Springer); john.gatesy@ucr.edu (J. Gatesy)

ABSTRACT. Higher-level relationships among placental mammals are mostly resolved, but several polytomies remain contentious. Song et al. (2012) claimed to have resolved three of these using shortcut coalescence methods (MP-EST, STAR) and further concluded that these methods, which assume no within-locus recombination, are required to unravel deep-level phylogenetic problems that have stymied concatenation. Here, we reanalyze Song et al.'s (2012) data and leverage these re-analyses to explore key issues in systematics including the recombination ratchet, gene tree stoichiometry, the proportion of gene tree incongruence that results from deep coalescence versus other factors, and simulations that compare the performance of coalescence and concatenation methods in species tree estimation. Song et al. (2012) reported an average locus length of 3.1 kb for the 447 protein-coding genes in their phylogenomic dataset, but the true mean length of these loci (start codon to stop codon) is 139.6 kb. Empirical estimates of recombination breakpoints in primates, coupled with consideration of the recombination ratchet, suggest that individual coalescence genes (c-genes) approach ~12 bp or less for Song et al.'s (2012) dataset, three to four orders of magnitude shorter than the c-genes reported by these authors. This result has general implications for the application of coalescence methods in species tree estimation. We contend that it is illogical to apply coalescence methods to complete protein-coding sequences. Such analyses amalgamate c-genes with different evolutionary histories (i.e., exons separated by >100,000 bp), distort true gene tree stoichiometry that is required for accurate species tree inference, and contradict the central rationale for applying coalescence methods to difficult phylogenetic problems. In addition, Song et al.'s (2012) dataset of 447 genes includes 21 loci with switched taxonomic names, eight duplicated loci, 26 loci with non-homologous sequences that are grossly misaligned, and numerous loci with >50% missing data for taxa that are misplaced in their gene trees. These problems were compounded by inadequate tree searches with nearest neighbor interchange branch swapping and inadvertent application of substitution models that did not account for among-site rate heterogeneity. Sixty-six gene trees imply unrealistic deep coalescences that exceed 100 million years (MY). Gene trees that were obtained with better justified models and search parameters show large increases in both likelihood scores and congruence. Coalescence analyses based on a curated set of 413 improved gene trees and a superior coalescence method (ASTRAL) support a Scandentia (treeshrews) + Glires (rabbits, rodents) clade, contradicting one of the three primary systematic conclusions of Song et al. (2012). Robust support for a Perissodactyla + Carnivora clade within Laurasiatheria is also lost, contradicting a second major conclusion of this study. Song et al.'s (2012) MP-EST species tree provided the

Download English Version:

<https://daneshyari.com/en/article/5918710>

Download Persian Version:

<https://daneshyari.com/article/5918710>

[Daneshyari.com](https://daneshyari.com)