



REPLY TO GATESY AND SPRINGER:

Claims of homology errors and zombie lineages do not compromise the dating of placental diversification

Liang Liu^{a,b,c}, Jin Zhang^d, Frank E. Rheindt^e, Fumin Lei^f, Yanhua Qu^f, Yu Wang^{g,h}, Yu Zhang^{g,h}, Corwin Sullivanⁱ, Wenhui Nie^j, Jinhuan Wang^j, Fengtang Yang^k, Jinping Chen^l, Scott V. Edwards^{a,m,n,1}, Jin Meng^o, and Shaoyuan Wu^{a,p,1}

Gatesy and Springer (1) consider 3 out of 89 nodes in our “preferred STAR tree” (2) unusual, raising suspicions that underlying alignment errors have generated these and other perceived misestimations in our analysis. As in their other critiques of our work, their claims are based on subjective and unrepeatable logic. We acknowledge that our alignments can be improved; in particular, we neglected to align and trim our loci based on more conserved amino acid alignments. However, our alignments still contain substantial phylogenetic information, and our protocols correctly extracted individual codon positions for analysis. Application of a suite of repeatable best practices for quality control in phylogenomics (3, 4) suggests that, after trimming, about 2.5% of individual sequences—a better measure of contamination than percentage of whole alignments—might be compromised. However, analysis of an improved 60-gene subset of our data yields divergence times that correlate strongly (0.9997) with our reported results (95% CI on slope of old versus new dates: 0.9995–0.9998), rejecting the notion that misalignments invalidate our analysis. As expected (5), our interordinal dates are influenced most strongly by the particular fossils we chose as calibrations, rather than sequence features.

Gatesy and Springer assert our data are plagued by “clear-cut homology errors.” “Clear-cut” is not an objective measure of quality, varies among investigators, and is therefore unscientific. Claims of errors in homology must include quantification and a repeatable protocol, neither of which Gatesy and Springer provide. We agree (6) that the manual data curation that Gatesy and Springer favor (1) has become “unfeasible” (3) and is subjective and neither repeatable nor sustainable, rendering its role in phylogenomics highly questionable (7, 8).

Other analyses, including STAR trees of C3 and CDS partitions, and concatenation trees of C12, C3, and CDS partitions, were consistent with the topological arrangements advocated by Gatesy and Springer (1). Our “preferred” tree was preferred not because it captured clades deemed correct by Gatesy and Springer, but because it possessed stability across analyses and broad congruence with previous work. The three clades in question are nowhere presented as new findings of our study and represent a minor issue in what is recognized by others (9, 10) as our paper’s (2) broader context. The claim that ASTRAL results differ from ours is insupportable: one cannot compare the local posterior probabilities produced by ASTRAL (11), which do not

^aJiangsu Key Laboratory of Phylogenomics and Comparative Genomics, School of Life Sciences, Jiangsu Normal University, Xuzhou 221116, Jiangsu, China; ^bDepartment of Statistics, University of Georgia, Athens, GA 30606; ^cInstitute of Bioinformatics, University of Georgia, Athens, GA 30606; ^dKey Laboratory of High Performance Computing and Stochastic Information Processing of the Ministry of Education of China, Department of Computer Science, College of Mathematics and Computer Science, Hunan Normal University, Changsha, Hunan 410081, China; ^eDepartment of Biological Sciences, National University of Singapore, Singapore 117543; ^fKey Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China; ^gState Key Laboratory of Reproductive Biology, Chinese Academy of Sciences, Beijing 100101, China; ^hInstitute of Zoology, Chinese Academy of Sciences, Beijing 100101, China; ⁱKey Laboratory of Vertebrate Evolution and Human Origins of the Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China; ^jState Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan 650223, China; ^kWellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom; ^lGuangdong Public Laboratory of Wild Animal Conservation and Utilization, Guangdong Key Laboratory of Integrated Pest Management in Agriculture, Guangdong Institute of Applied Biological Resources, Guangzhou 510260, China; ^mDepartment of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138; ⁿMuseum of Comparative Zoology, Harvard University, Cambridge, MA 02138; ^oDivision of Paleontology, American Museum of Natural History, New York, NY 10024; and ^pDepartment of Biochemistry and Molecular Biology, 2011 Collaborative Innovation Center of Tianjin for Medical Epigenetics, Tianjin Key Laboratory of Medical Epigenetics, School of Basic Medical Sciences, Tianjin Medical University, Tianjin 300070, China

Author contributions: L.L., S.V.E., and S.W. designed research; L.L., F.E.R., S.V.E., J.M., and S.W. performed research; F.L., Y.Q., Y.W., Y.Z., C.S., W.N., J.W., F.Y., and J.C. contributed new reagents/analytic tools; L.L., J.Z., F.E.R., and S.W. analyzed data; and L.L., F.E.R., S.V.E., and S.W. wrote the paper.

The authors declare no conflict of interest.

Published under the [PNAS license](#).

¹To whom correspondence may be addressed. Email: shaoyuanwu@outlook.com or sedwards@fas.harvard.edu.

take gene tree error into account, to the measures of support we used.

Gatesy and Springer are concerned that “zombie lineages” compromise our conclusions. We acknowledged zombie lineages as a reasonable concern and discussed such discrepancies and their likely causes at some length in our study (2). At the same time, our analysis is an advance because many more fossil and molecular divergences, particularly ordinal divergences, are now better reconciled. Hard bounds on priors can work but are also more likely to

mislead (5) than the soft bounds we used. Even sophisticated approaches can misestimate divergences in some cases, while uncertainties in the phylogenetic placement and dating of fossils may often yield false assumptions about fossil ages used for calibration.

Acknowledgments

This work was supported by funds from the Priority Academic Program Development of Jiangsu Higher Education Institutions, the National Natural Science Foundation of China (Grant 31470111) (to S.W.).

- 1 Gatesy J, Springer MS (2017) Phylogenomic red flags: Homology errors and zombie lineages in the evolutionary diversification of placental mammals. *Proc Natl Acad Sci USA* 114:E9431–E9432.
- 2 Liu L, et al. (2017) Genomic evidence reveals a radiation of placental mammals uninterrupted by the KPg boundary. *Proc Natl Acad Sci USA* 114:E7282–E7290.
- 3 Irisarri I, et al. (2017) Phylotranscriptomic consolidation of the jawed vertebrate timetree. *Nat Ecol Evol* 1:1370–1378.
- 4 Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25:1972–1973.
- 5 Yang Z, Rannala B (2006) Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol Biol Evol* 23:212–226.
- 6 Edwards SV, et al. (2016) Implementing and testing the multispecies coalescent model: A valuable paradigm for phylogenomics. *Mol Phylogenet Evol* 94:447–462.
- 7 Morrison DA (2009) Why would phylogeneticists ignore computerized sequence alignment? *Syst Biol* 58:150–158.
- 8 Philippe H, et al. (2011) Resolving difficult phylogenetic questions: Why more sequences are not enough. *PLoS Biol* 9:e1000602.
- 9 Bruno M (2017) *Nature Research Highlights*. Available at <https://www.nature.com/articles/d41586-017-02706-z>. Accessed August 22, 2017.
- 10 Carter AM (August 24, 2017) Interesting hypothesis. Available at <https://f1000.com/prime/728639209#eval793535590>. Accessed October 6, 2017.
- 11 Sayyari E, Mirarab S (2016) Fast coalescent-based computation of local branch support from quartet frequencies. *Mol Biol Evol* 33:1654–1668.