



When Complex Models Fit the Wrong Mechanistic Complexity in Phylogenomic Analysis

Liang Liu¹ · David A. Liberles²

Received: 3 May 2026 / Accepted: 12 May 2026 / Published online: 23 May 2026
© The Author(s) 2026

Abstract

Substitution model selection is central to phylogenetic inference and is commonly treated as a problem of identifying the substitutional complexity required to describe sequence evolution along a single tree. This framework implicitly assumes a shared genealogy across all sites, an assumption that is routinely violated in phylogenomic data by incomplete lineage sorting and other sources of gene-tree discordance. Recent work by Lozano et al. (2026) demonstrates that unmodeled genealogical heterogeneity can systematically distort substitution model selection, creating spurious support for parameter-rich models even when the underlying substitution process is simple. In this perspective, we examine the consequences of this confounding for phylogenetic estimation, divergence-time inference, and the biological interpretation of substitution model parameters, highlighting the need to more explicitly integrate genealogical heterogeneity into model selection and adequacy assessment in phylogenomics. This perspective reframes substitution model choice not only as a tool for describing molecular evolution, but also as a potential diagnostic for violations of shared-genealogy assumptions that are central to modern systematics.

Keywords Phylogenomics · Coalescent · Substitution model · Model selection

The choice of an appropriate substitution model is a foundational component of molecular evolutionary inference (Anisimova et al., 2013; Spielman, 2020; Del Amparo and Arenas, 2023). For decades, substitution model selection has been framed as a statistical optimization problem: identifying the stochastic model of nucleotide change that best explains an observed sequence alignment, typically conditional on a fixed tree topology. This practice has become so deeply embedded in phylogenetic workflows that it is often treated as a routine technical step rather than as a potential source of conceptual and inferential uncertainty (Anisimova et al., 2013; Spielman, 2020; Del Amparo and Arenas, 2023).

At the core of conventional model selection frameworks lies an implicit assumption of genealogical homogeneity that all sites within an alignment share a single underlying evolutionary history. Within this framework, heterogeneity among sites is attributed primarily to variation in the substitution process itself, such as differences in evolutionary rates or nucleotide frequencies driven by biochemical constraints, functional roles, or selective pressures (Echave et al., 2016). To accommodate these forms of substitutional heterogeneity while maintaining a single-tree model, phylogenetic methods have progressively adopted increasingly parameter-rich formulations, including gamma-distributed rate variation across sites (Yang, 1993), proportions of invariant sites (Hasegawa et al., 1985), and highly flexible rate matrices (see Woodhams et al., 2015).

The widespread adoption of genome-scale data has, however, called the validity of this single-tree assumption into question. Concatenated multilocus datasets routinely combine sequences drawn from distinct genomic regions that may follow disparate evolutionary histories (Boussau, 2013; Palmer et al., 2019). Processes such as incomplete lineage sorting (Degnan and Rosenberg, 2009; Tan et al., 2023), introgression (Hibbins and Hahn, 2022), hybridization

Handling editor: Peter Chi.

✉ Liang Liu
lliu@uga.edu

¹ Department of Statistics, Institute of Bioinformatics, University of Georgia, Athens, GA, USA

² Department of Biology, Temple University, Philadelphia, PA, USA

(Kubatko, 2009; Wang et al., 2021), and gene duplication and loss (Rasmussen and Kellis, 2012; Smith and Hahn, 2021; Soewongsono et al., 2025) routinely generate topological discordance among loci, indicating that genealogical heterogeneity is often the biological norm rather than an exceptional complication. Despite this, standard substitution model selection procedures typically disregard underlying conflicts among gene trees. As a result, the selected model is frequently interpreted as a faithful representation of the molecular substitution process, even though it may instead reflect a statistical compromise imposed by unmodeled genealogical discordance.

In this issue, Lozano et al. (2026) explicitly interrogate this long-standing disconnect between model assumptions and empirical reality. By demonstrating that genealogical heterogeneity can systematically bias substitution model selection toward increasingly complex models, their study challenges conventional interpretations of what substitution models capture in concatenated analyses. Their results suggest that, under realistic evolutionary scenarios, model selection criteria may be more sensitive to hidden topological conflict than to true substitutional heterogeneity. This work prompts a reassessment of substitution model selection as an inferential step in phylogenetics and underscores the need to integrate concepts of genealogical heterogeneity more explicitly into systematic practice.

Model Selection Under a Shared Genealogy

Classical nucleotide substitution models were developed within an explicitly genealogical framework, in which sequence evolution is modeled along a single, fixed tree. Early formulations, including the Jukes–Cantor (Jukes and Cantor, 1969) and Kimura two-parameter models (Kimura, 1980), describe stochastic character change along branches of a shared genealogy, with later extensions accommodating empirical features of molecular evolution such as base-frequency bias (Felsenstein, 1981), transition–transversion and other base conversion asymmetry (Lanave et al., 1984; Woodhams et al., 2015), and rate heterogeneity among sites (Yang, 1993). Importantly, these extensions were motivated by mechanistic and biological considerations pertaining to the substitution process itself, not by differences in underlying tree topology or branch lengths among sites.

Within this framework, likelihood-based and information-theoretic approaches to model selection are conceptually well defined. A more parameter-rich model is preferred when its improvement in fit outweighs the penalty imposed by increased complexity. Under the assumption of a shared genealogy, such improvement is naturally interpreted as evidence that a more realistic description of the substitution

process, such as heterogeneous rates or asymmetric substitution patterns, is required to explain the observed data.

Crucially, this interpretation depends on the validity of the shared-tree assumption. When all sites genuinely evolve along the same genealogy, increases in likelihood can be meaningfully attributed to improvements in modeling the substitution process. When this assumption is violated, however, the source of improved model fit becomes ambiguous. Additional parameters may instead compensate for unmodeled genealogical heterogeneity, allowing a single-tree model to better approximate data generated under multiple underlying histories. Under such conditions, model selection no longer provides a clear basis for distinguishing substitutional complexity from violations of shared ancestry.

Networks vs. Trees

The goal in building a phylogeny is to capture the evolutionary history of the sequences (and in the species context, the organisms they represent). When there isn't a single topology representing the origin of the samples, one alternative to assuming a fixed genealogy is to model evolution with a phylogenetic network instead of a phylogenetic tree (Bryant and Moulton 2004; Wen et al. 2018). The biological underpinnings of a phylogenetic network is that every nucleotide is represented by a single tree that is embedded in the network such that the network structure is a union of such trees but contains an evolutionary history that is not relevant in all paths to any given nucleotide. That isn't necessarily a problem. However, methods to build networks using complex model-based approaches and to perform inference on networks, including tests for selection or ancestral state reconstruction are much less developed than for trees. In building methods that are robust to samples with multiple underlying topologies, expanding the methodological toolbox for phylogenetic network construction and analysis is one fruitful direction for the community.

Genealogical Heterogeneity as a Hidden Axis of Variation

Gene-tree discordance introduces an axis of variation in sequence data that is fundamentally distinct from, yet formally confounded with, traditional substitution model parameters. Even when loci evolve under identical substitution processes, their underlying genealogies may differ in both topology and branch lengths (see for example, Degnan and Rosenberg 2009). When such loci are concatenated and analyzed under a single-tree framework, the resulting

alignment represents a mixture of evolutionary histories that no single genealogy can faithfully capture.

Lozano et al. (2026) explicitly examine the consequences of this mismatch between data-generating processes and model assumptions. Using simulations in which all sites evolve under the simplest possible substitution model without rate heterogeneity, compositional bias, or other forms of substitutional complexity, the authors vary only the number of distinct gene trees contributing to a concatenated alignment. As genealogical heterogeneity increases, standard likelihood-based model selection procedures increasingly favor more parameter-rich substitution models, particularly those incorporating among-site rate variation.

This result is notable not because it is unexpected, but because it renders explicit a phenomenon likely pervasive in empirical analyses yet rarely acknowledged. Variation in branch lengths and topologies among gene trees generates patterns of site-wise divergence that, when projected onto a single inferred tree, closely resemble classic signatures of rate heterogeneity. In response, substitution models accommodate this variation by inflating rate-distribution parameters, despite the absence of genuine differences in substitutional tempo among sites.

From this perspective, commonly invoked parameters such as the gamma shape parameter or the proportion of invariant sites assume a dual statistical role. Beyond modeling authentic biochemical or functional heterogeneity in substitution rates, these parameters can also serve as effective summaries of unmodeled genealogical discordance. This duality complicates biological interpretation of model choice and calls into question the extent to which selected substitution models can be straightforwardly construed as descriptions of the underlying molecular evolutionary process.

Reinterpreting Substitution Model Complexity

The interpretation of substitution model complexity as direct evidence for biological complexity in the substitution process is deeply entrenched in molecular systematics (Steel 2005). Parameter-rich models are commonly viewed as biologically informative, invoked to support inferences about heterogeneous selective pressures, functional constraints, or other forms of site-specific evolutionary variation. The results discussed here urge caution against such straightforward interpretations, particularly in the context of concatenated phylogenomic analyses.

If genealogical heterogeneity alone can drive the preferential selection of complex substitution models, even when the underlying substitution process is homogeneous, then

model complexity cannot be uncritically equated with substitutional realism. Rather than indicating a more accurate description of molecular evolution, increased parameterization may instead reflect a mismatch between the assumed genealogical framework and the true data-generating process. Under such circumstances, substitution models absorb variation that originates not from biochemical processes, but from unmodeled discordance among gene trees.

This observation aligns with broader discussions in phylogenetics concerning the distinction between model fit and model adequacy. A model may outperform alternatives according to standard selection criteria while remaining conceptually inadequate if it accommodates violations of key assumptions in unintended ways (Steel 2005; Liberles et al. 2013). Viewed through this lens, increasingly complex substitution models may function as statistical surrogates for missing components of the evolutionary model, most notably, an explicit treatment of genealogical heterogeneity. Recognizing this distinction is essential for interpreting both model choice and downstream phylogenetic inferences in genome-scale analyses.

Model Adequacy and Model Selection

Model selection is performed using various information criteria that compare the fit of the data by the model with the number of parameters used to do so (see for example, Susko and Roger 2020 for a discussion). Model selection presents a comparison between alternative models that are proposed, as a formalization of scientific hypothesis testing (Liberles et al. 2013). This comes from the perspective that modeling is not meant to just fit data but is an activity that enables discovery of scientific processes. But what happens when all of the models generated are mis-specified? Model selection at this stage still presents the best scientific explanation for the data among hypotheses that have been generated, but one still should know that the model doesn't explain data well and that some new thinking about process and the model is needed.

How is model adequacy evaluated? The most standard approach in phylogenetics is to perform posterior predictive simulation (Duchene et al. 2019; see McCarthy et al. 2023). However, what happens if a complex model fits the data well but is still mis-specified in that parameters are not fitting what they were mechanistically designed to fit? One may not know that this is the case except when one needs overly complex mixture models to fit data that might provide a red flag. The solution to diagnosing this as a problem is using sets of simulations that make different sets of assumptions and examining the fit of the model to understand when parameters are mopping up complexity that the model isn't

designed to fit. There isn't a standardized answer to doing this in the absence of specific hypotheses about the biological generation of the data. In this case, the hypothesis that does present itself is the absence of a single underlying tree.

One diagnostic that has been presented in the literature is to ask whether the observed nucleotide site-pattern frequencies are compatible with a homogeneous gene tree model, in which all sites are assumed to evolve along the same underlying genealogy under a specified substitution process (Liu and Yu, 2010). This can be evaluated using classical goodness-of-fit tests, such as the Pearson chi-square test or likelihood ratio tests, which compare observed site-pattern counts to their expectations under the fitted model. In this framework, systematic deviations between observed and expected frequencies indicate that the assumed single-tree model fails to capture important structure in the data. Notably, such lack of fit can arise even when the substitution process itself is correctly specified, if heterogeneous gene genealogies generate site-pattern distributions that cannot be represented by any single tree. Consequently, rejection of the homogeneous gene tree model by these tests should be interpreted cautiously: it may reflect unmodeled genealogical heterogeneity rather than inadequacy of the substitution model *per se*. Used in this way, goodness-of-fit tests provide a useful diagnostic for violations of shared-genealogy assumptions, complementing information-criterion-based model selection and posterior predictive approaches.

Implications for Phylogenetic Inference

The implications of genealogically driven substitution model complexity extend well beyond the model selection step itself. Substitution models shape virtually all components of phylogenetic inference, influencing estimates of tree topology, branch lengths, divergence times, and ancestral character states. If selected models are, in part, absorbing genealogical conflict rather than capturing features of the substitution process, then downstream inferences may inherit systematic, but difficult-to-detect, biases.

Inflation of among-site rate heterogeneity parameters provides a clear illustration. Apparent increases in rate variation can distort branch-length estimates, thereby affecting interpretations of relative evolutionary tempo among lineages. Divergence time estimation may be particularly sensitive to such effects, as many molecular clock models interact directly with assumptions about rate heterogeneity across sites (Soubrier et al. 2012; Mello and Schrago 2024). Likewise, tests for molecular adaptation or selection that rely on fitted substitution models may be confounded if inferred complexity reflects genealogical discordance rather

than true heterogeneity in selective regimes (Diekmann and Pereira-Leal 2016).

The results of Lozano et al. (2026) also highlight meaningful differences among commonly used information criteria. More conservative criteria penalize over-parameterization more strongly and may therefore be comparatively resistant to model inflation under moderate discordance. Nevertheless, no criterion is immune when genealogical heterogeneity is extreme. This underscores that model selection is not a purely mechanical procedure but reflects trade-offs between goodness of fit and parsimony that interact in nontrivial ways with data structure and model misspecification.

Concatenation vs. Coalescent

These findings contribute to a growing body of work reexamining the widespread use of concatenation relative to coalescent-based approaches in phylogenomics (Liu et al. 2015). Although concatenation can yield statistically consistent estimates of species relationships under certain conditions, coalescent theory makes clear that heterogeneous gene genealogies are an expected outcome of population-level processes such as incomplete lineage sorting (Liu et al. 2009). The interaction between concatenation and substitution models is therefore more complex than is often acknowledged. The present results suggest that even when concatenation produces a reasonable estimate of species-level topology, it may nonetheless distort inferences about the substitution process itself.

This distortion does not imply that concatenation is inherently invalid, nor that coalescent approaches are universally superior. Rather, it highlights the importance of matching modeling assumptions to the data-generating process. When loci with heterogeneous genealogies, explicitly modeled under the multispecies coalescent, are instead combined into a single alignment and analyzed under a shared-tree assumption, the resulting data violate the generative assumptions of standard substitution models. Under such conditions, improved model fit achieved through increased parameterization may reflect compensation for unmodeled genealogical heterogeneity rather than enhanced realism of the substitution process. Consequently, biologically meaningful interpretation of substitution model parameters requires caution in concatenated analyses lacking an explicit coalescent framework.

Toward Integrative Modeling Approaches

A central question raised by these findings is how best to accommodate genealogical heterogeneity in phylogenetic inference. One clear solution is greater reliance on models that explicitly account for variation in gene genealogy, such as multispecies coalescent frameworks (Rannala and Yang 2003). By modeling genealogical and substitutional processes separately, these approaches avoid conflating historical discordance with molecular evolutionary dynamics. However, their computational demands and limited integration with flexible substitution model selection currently restrict their routine application to genome-scale datasets.

An alternative direction lies in the development of hybrid or mixture models that incorporate aspects of genealogical heterogeneity while remaining computationally tractable (Liu et al. 2010a, b). Such approaches may capture some of the benefits of coalescent-based modeling without abandoning the concatenation framework entirely. Emerging machine-learning-based tools for model selection and adequacy assessment may also play a role by detecting when increased model complexity reflects deeper violations of shared-genealogy assumptions rather than true substitutional heterogeneity (Liu and Yu, 2010).

From a phylogenetic perspective, for methods not based upon the coalescent, formal gene tree-species tree reconciliation can be undertaken with a model that assumes a particular process is involved in the reconciliation. For example, Rasmussen and Kellis (2012) and Soewongsono et al. (2025) present models that formally characterize the duplication and loss process in reconciling individual gene trees with a species tree, or in the generation of a species tree.

Equally important is a shift in how substitution model selection is interpreted. Rather than viewing model choice solely as a quest for biological realism, it may be more productive to treat it as a diagnostic step. The repeated selection of highly parameterized models, particularly in large concatenated datasets, may serve as an indicator of underlying genealogical conflict, prompting closer examination of data structure and modeling assumptions.

Broader Implications for Molecular Evolution

Beyond its immediate methodological relevance, this work speaks to a broader challenge in molecular evolution: disentangling multiple, overlapping sources of heterogeneity in genomic data. As phylogenetic datasets increase in scale, they increasingly reflect not only variation in substitution processes, but also complex population-level histories shaped by coalescence, introgression, and lineage sorting.

Models developed for simpler evolutionary scenarios often respond by absorbing these complexities into substitution parameters, thereby obscuring the distinction between evolutionary process and statistical pattern.

Recognizing when apparent substitutional complexity is instead a signature of deeper genealogical processes is essential for robust evolutionary interpretation. Lozano et al. (2026) provide a clear and empirically grounded demonstration of this problem, along with a valuable conceptual framework for addressing it.

Concluding Remarks

By demonstrating that genealogical heterogeneity can systematically bias substitution model selection toward increased complexity, Lozano and colleagues challenge conventional interpretations of model choice in phylogenetics. Their results underscore the importance of aligning modeling assumptions with the structure of genomic data and caution against equating parameter richness with biological realism.

As phylogenomics continues to expand in both scope and complexity, these insights are particularly timely. They serve as a reminder that improved statistical fit does not necessarily imply improved biological understanding, and that substitution models must be evaluated not only by their likelihoods, but also by the assumptions they encode. In this sense, substitution model selection should be viewed not merely as a technical prerequisite, but as a window into the deeper evolutionary structure of molecular data.

Acknowledgements David A. Liberles initially invited this perspective as Editor-in-Chief of the *Journal of Molecular Evolution* before being invited to join the work as a secondary author.

Author Contributions LL and DL conceived the study and wrote the manuscript.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing Interests Liang Liu is an Associate Editor for the *Journal of Molecular Evolution*. David Liberles is Editor-in-Chief of the *Journal of Molecular Evolution*. Neither was involved in the peer review of this work. No other competing interests are noted.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless

indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anisimova M, Liberles DA, Philippe H et al (2013) State-of the art methodologies dictate new standards for phylogenetic analysis. *BMC Evol Biol* 13:161. <https://doi.org/10.1186/1471-2148-13-161>
- Boussau B, Szöllosi GJ, Duret L, Gouy M, Tannier E, Daubin V (2013) Genome-scale coestimation of species and gene trees. *Genome Res* 23(2):323–30. <https://doi.org/10.1101/gr.141978.112>
- Bryant D, Moulton V (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol* 21(2):255–265. <https://doi.org/10.1093/molbev/msh018>
- Degnan JH, Rosenberg NA (2009) Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol Evol* 24(6):332–340. <https://doi.org/10.1016/j.tree.2009.01.009>
- Del Amparo R, Arenas M (2023) Influence of substitution model selection on protein phylogenetic tree reconstruction. *Gene* 865:147336. <https://doi.org/10.1016/j.gene.2023.147336>
- Diekmann Y, Pereira-Leal JB (2016) Gene tree affects inference of sites under selection by the Branch-Site Test of Positive Selection. *Evol Bioinform Online* 11(Suppl 2):11–17. <https://doi.org/10.4137/EBO.S30902>
- Duchene S, Bouckaert R, Duchene DA, Stadler T, Drummond AJ (2019) Phylodynamic model adequacy using posterior predictive simulations. *Syst Biol* 68(2):358–364. <https://doi.org/10.1093/sysbio/syy048>
- Echave J, Spielman SJ, Wilke CO (2016) Causes of evolutionary rate variation among protein sites. *Nat Rev Genet* 17(2):109–121. <https://doi.org/10.1038/nrg.2015.18>
- Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol* 17(6):368–76. <https://doi.org/10.1007/BF01734359>
- Hasegawa M, Kishino H, Yano T (1985) Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol* 22(2):160–174. <https://doi.org/10.1007/BF02101694>
- Hibbins MS, Hahn MW (2022) Phylogenomic approaches to detecting and characterizing introgression. *Genetics* 220(2):iyab173. <https://doi.org/10.1093/genetics/iyab173>
- Jukes TH, Cantor CH (1969) Evolution of protein molecules. In: Munro HN (ed) *Mammalian protein metabolism*, vol 3. Elsevier, pp 21–132. <https://doi.org/10.1016/b978-1-4832-3211-9.50009-7>
- Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* 16(2):111–20. <https://doi.org/10.1007/BF01731581>
- Kubatko L (2009) Identifying hybridization events in the presence of coalescence via model selection. *Syst Biol* 58(5):478–488. <https://doi.org/10.1093/sysbio/syp055>
- Lanave C, Preparata G, Saccone C, Serio G (1984) A new method for calculating evolutionary substitution rates. *J Mol Evol* 20:86–93. <https://doi.org/10.1007/BF02101990>
- Liberles DA, Teufel AI, Liu L, Stadler T (2013) On the need for mechanistic models in computational genomics and metagenomics. *Genome Biol Evol* 5(10):2008–2018. <https://doi.org/10.1093/gbe/evt151>
- Liu L, Yu L, Pearl DK, Edwards SV (2009) Estimating species phylogenies using coalescence times among sequences. *Syst Biol* 58(5):468–477. <https://doi.org/10.1093/sysbio/syp031>
- Liu L, Yu L, Edwards SV (2010a) A maximum pseudo-likelihood approach for estimating species trees under the coalescent model. *BMC Evol Biol* 10:302. <https://doi.org/10.1186/1471-2148-10-302>
- Liu L, Yu L (2010b) Phybase: an R package for species tree analysis. *Bioinformatics* 26(7):962–963. <https://doi.org/10.1093/bioinformatics/btq062>
- Liu L, Wu S, Xi Z, Davis CC, Edwards SV (2015) Estimating phylogenetic trees from genome-scale data. *Ann N Y Acad Sci* 1360:36–53. <https://doi.org/10.1111/nyas.12747>
- Lozano JR, Miya MS, Turner E, McKenna D, Adams R (2026) Concatenation, Conflict, and Complexity: Genealogical Heterogeneity Mimics Substitutional Heterogeneity for Nucleotide Model Selection. *J Mol Evol*. <https://doi.org/10.1007/s00239-026-10317-4>
- McCarthy CGP, Mulhair PO, Siu-Ting K, Creevey CJ, O'Connell MJ (2023) Improving orthologous signal and model fit in datasets addressing the root of the animal phylogeny. *Mol Biol Evol* 40(1):msac276. <https://doi.org/10.1093/molbev/msac276>
- Mello B, Schrago CG (2024) Modeling substitution rate evolution across lineages and relaxing the molecular clock. *Genome Biol Evol* 16(9):evae199. <https://doi.org/10.1093/gbe/evae199>
- Palmer M, Venter SN, McTaggart AR, Coetzee MPA, Van Wyk S, Avontuur JR, Beukes CW, Fourie G, Santana QC, Van Der Nest MA, Blom J, Steenkamp ET (2019) The synergistic effect of concatenation in phylogenomics: the case in *Pantoea*. *PeerJ* 7:e6698. <https://doi.org/10.7717/peerj.6698>
- Rannala B, Yang Z (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164(4):1645–1656. <https://doi.org/10.1093/genetics/164.4.1645>
- Rasmussen MD, Kellis M (2012) Unified modeling of gene duplication, loss, and coalescence using a locus tree. *Genome Res* 22(4):755–765. <https://doi.org/10.1101/gr.123901.111>
- Smith ML, Hahn MW (2021) New approaches for inferring phylogenies in the presence of paralogs. *Trends Genet* 37(2):174–187. <https://doi.org/10.1016/j.tig.2020.08.012>
- Soewongsono AC, Diao J, Stark T et al (2025) Matrix-analytic methods for the evolution of species trees, gene trees, and their reconciliation. *Methodol Comput Appl Probab* 27:9. <https://doi.org/10.1007/s11009-025-10135-z>
- Soubrier J, Steel M, Lee MS, Der Sarkissian C, Guindon S, Ho SY, Cooper A (2012) The influence of rate heterogeneity among sites on the time dependence of molecular rates. *Mol Biol Evol* 29(11):3345–58. <https://doi.org/10.1093/molbev/mss140>
- Spielman SJ (2020) Relative model fit does not predict topological accuracy in single-gene protein phylogenetics. *Mol Biol Evol* 37(7):2110–2123. <https://doi.org/10.1093/molbev/msaa075>
- Steel M (2005) Should phylogenetic models be trying to fit an elephant? *Trends Genet* 21(6):307–309. <https://doi.org/10.1016/j.tig.2005.04.001>
- Susko E, Roger AJ (2020) On the use of information criteria for model selection in phylogenetics. *Mol Biol Evol* 37(2):549–562. <https://doi.org/10.1093/molbev/msz228>
- Tan X, Qi J, Liu Z, Fan P, Liu G, Zhang L, Shen Y, Li J, Roos C, Zhou X, Li M (2023) Phylogenomics reveals high levels of incomplete lineage sorting at the ancestral nodes of the Macaque radiation. *Mol Biol Evol* 40(11):msad229. <https://doi.org/10.1093/molbev/msad229>
- Wang Y, Cao Z, Ogilvie HA, Nakhleh L (2021) Phylogenomic assessment of the role of hybridization and introgression in trait evolution. *PLoS Genet* 17(8):e1009701. <https://doi.org/10.1371/journal.pgen.1009701>

- Wen D, Yu Y, Zhu J, Nakhleh L (2018) Inferring phylogenetic networks using PhyloNet. *Syst Biol* 67(4):735–740. <https://doi.org/10.1093/sysbio/syy015>
- Woodhams MD, Fernández-Sánchez J, Sumner JG (2015) A new hierarchy of phylogenetic models consistent with heterogeneous substitution rates. *Syst Biol* 64(4):638–650. <https://doi.org/10.1093/sysbio/syv021>
- Yang Z (1993) Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol Biol Evol* 10:1396–1401
- Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.