PSG

SILENCE

journals.panorama-sg.com/index.php/Silence

ISSN (Online) 3054-4386
Publisher: Panorama Scholarly Group

Vol. 1, No.1, 62-68, 2026
DOI:10.5281/zenodo.19122073

UNWANTED REVIEWS

Boosting Branch Supports: A 'Pragmatic' Self-Aid Guide for the Data-Poor Graduate Students

Anonymous Bioinformatic Loser[1] **and Kal'tsit**[2*]

[1]*State Key Laboratory of Data Manipulation Research, College of Science Regression, Alienation University, Hinterland in Bioinformatics, Yan*
[2]*Rhodes Island Pharmaceutical Inc.*
[*]*Corresponding author*

Corresponding author:
ama_10@foxmail.com

Received: 22 February 2026
Accepted: 01 March 2026
Available online: 20 March 2026

Abstract
The continuous advancement of sequencing technologies has precipitated the proliferation of molecular phylogenetics far beyond the boundaries of traditional systematics. Despite its methodological ubiquity, novice postgraduate researchers frequently grapple with the systemic pressure to generate universally high branch support values, often constrained by suboptimal empirical data and limited institutional resources. This article provides a systematic discussion, encompassing a theoretical and empirical review of established branch support metrics. Grounded in these computational findings, we introduce a pragmatic 'guide' detailing methodological strategies designed to artificially inflate these metrics. We argue that when robust branch support devolves from a statistical diagnostic tool into a strict prerequisite for publication, it catalyzes a metric-optimization culture that prioritizes numerical aesthetics over epistemological rigor. By deconstructing these optimization practices, this work functions as a structural critique of the incentive mechanisms within contemporary data-driven science, highlighting the crisis of authentic representation under severe academic performance pressures.

Keywords: Academic Irony; Data Manipulation; Data Hacking; Publish or Perish; Methodological Critique; Molecular Phylogenetics

1. Introduction

Molecular phylogenetics operationalizes the diversification of biomolecules by testing falsifiable evolutionary hypotheses through tree or network topologies grounded in distinct mathematical and statistical principles. Validating the 'stability' or 'robustness' of a specified topology is fundamental to establishing the model's epistemological validity. A branch support value serves to quantify the robustness of corresponding internal nodes within a reconstructed phylogeny [1]. Among various computational approaches, bootstrap resampling, introduced by Felsenstein [2], remains the methodological gold standard for assessing branch support. Elevated support values indicate that a phylogenetic branch is statistically fortified by the underlying dataset and resilient against stochastic randomized deviations. Consequently, high numerical support is critical for corroborating the evolutionary assumptions underpinning specific clades, and is routinely mandated as a baseline requirement by academic publishers and peer reviewers.

However, this ubiquitous diagnostic mandate necessitates critical epistemological scrutiny. As Nei and Kumar [1] establish, even topologies exhibiting low branch supports often represent the optimal algorithmic solution and likely reflect accurate evolutionary branching, a premise substantiated by extensive computational simulations. Conversely, topologies exhibiting high statistical support can fundamentally contradict true evolutionary histories due to structural model misspecification [3,4]. Therefore, the ritualistic overemphasis on maximizing numerical thresholds undermines the inherent diagnostic intent of phylogenetic testing, transforming scientific inference into an academic ceremony characterized by intellectual inertia and conceptual oversimplification.

Novice researchers frequently encounter naturally low branch supports when analyzing typical genetic markers with constrained sequence lengths, rendering the 'optimization' of these values a ubiquitous technical inquiry within academic communication networks (Figure 1). Phenomenologically, limited informative sites, substitution saturation, and latent evolutionary heterogeneity inherently produce diminished

Cite as: Anonymous Bioinformatic Loser, Kal'tsit(2026). Boosting Branch Supports: A 'Pragmatic' Self-Aid Guide for the Data-Poor Graduate Students. Panorama Scholarly Group | SILENCE, 1(1) (2026) 62-68. https://doi.org/10.5281/zenodo.19122073

support values—a theoretically sound and empirically standard outcome. A pervasive systemic pathology within contemporary peer review and institutional evaluation is the uncritical mandate for high branch support metrics—often arbitrarily standardized above 90% or 95%—irrespective of the actual biological context or phylogenetic diagnostic utility. Consequently, naturally occurring low support values are routinely conflated with methodological inadequacy or potential academic misconduct, engendering a coercive research environment heavily burdened by performance anxiety.

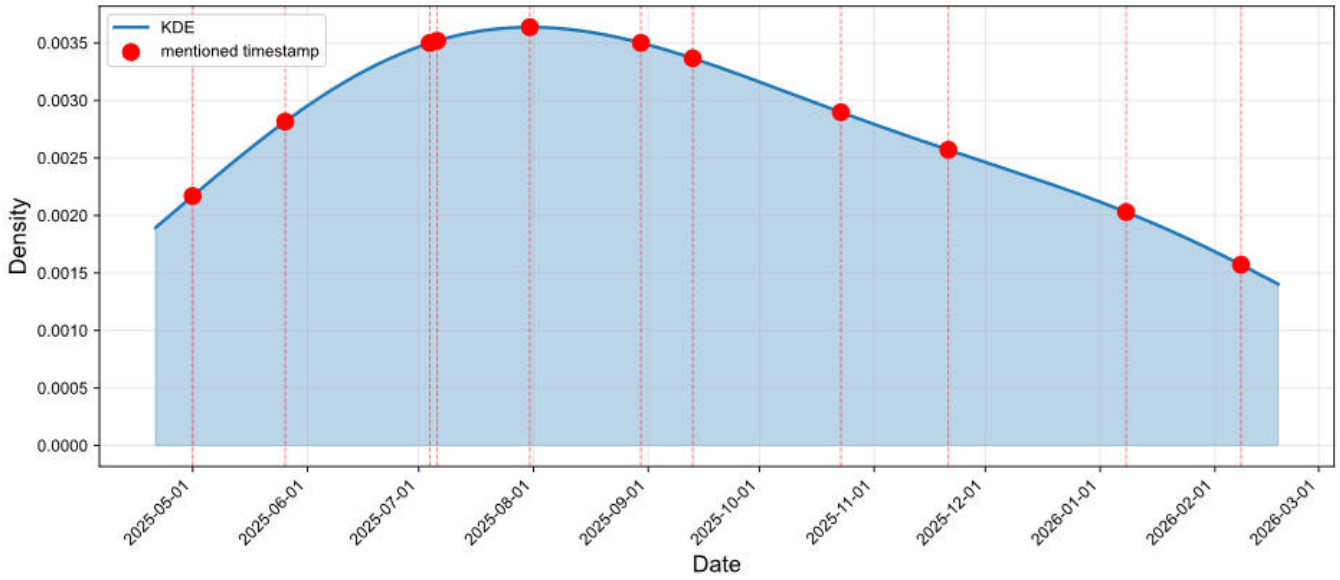

Figure 1: Message Statics of the Topic 'How to Solve Poor Branch Supports' in a Rednote Molecular Phylogenetics Mutual-Aid Group

Rather than introducing novel inferential technologies, this article navigates the ethical deep-water zone of academia, dissecting computational strategies explicitly engineered to inflate branch supports across phylogenetic topologies—a subject systematically marginalized in conventional methodological literature. We first review predominant branch support categories, subsequently detailing computationally 'effective' optimization strategies substantiated by mathematical theory and simulated empirical evidence. Adopting the rhetorical framework of an ironic 'self-help guide,' this work dissects the academic pathology of 'metric worship' driven by the 'publish or perish' paradigm. These delineated strategies reveal how the institutional pursuit of superficial robustness systematically distorts methodology, exploits algorithmic vulnerabilities, sanitizes complex scientific presentations, and ultimately prunes intricate evolutionary histories to satisfy the sanitized aesthetic preferences of academic journals.

2. Phylogenetic Test Methods

2.1. Likelihood Ratio Test (LRT)

The Likelihood Ratio Test (LRT) functions as a universal statistical method comparing the likelihood of an inferred phylogenetic hypothesis against a null hypothesis to ascertain structural divergence. Within phylogenetics, the prevalent iteration for branch support generation is the Shimodaira-Hasegawa-like approximate likelihood ratio test (SH-aLRT) developed by Shimodaira and Hasegawa [5]. The null hypothesis of this framework posits an internal branch length of zero, collapsing the contiguous nodes into an unresolved polytomy and inherently nullifying the monophyly of the terminal clades. If the likelihood of this null hypothesis exhibits no statistically significant deviation from the fully resolved topology under identical evolutionary parameters, the branch is deemed poorly supported by the empirical data.

2.2. Resamplings (Bootstrap and Jackknife)

Resampling techniques simulate stochastic variation by computationally perturbing the original alignment via randomized sampling. The algorithms subsequently reconstruct topologies from these resamples, calculating the topological reproduction frequency of each specific branch across the replicate set to derive its support percentage. The fundamental methodological divergence between Bootstrap and Jackknife algorithms lies in replacement strategy. For an alignment comprising N sites, bootstrap generates pseudo-replicates of length N, inherently permitting redundant sampling of specific loci. In contrast, the jackknife algorithm executes sampling without replacement, systematically discarding a fraction of the data to generate sub-sampled alignments shorter than N.

2.3. Bayesian Posterior Probability (BPP) and Approximate Bayesian Probability (aBayes)

Bayesian Posterior Probability (BPP) is derived utilizing Bayesian inference coupled with Markov Chain Monte Carlo (MCMC) stochastic sampling. This approach iteratively evaluates hundreds of thousands of MCMC generations, proposing topological and parametric modifications to maximize the overall posterior probability. Proposals are probabilistically accepted or rejected based on established prior distributions. Upon reaching analytical convergence, the posterior probabilities of the sampled tree topologies stabilize near an optimal solution space with

minimal stochastic variation. Following the discarding of early non-convergent generations (burn-in), the residual samples are consolidated into a majority-rule consensus topology, where branch posterior probabilities are defined by the frequency of local topological support within the post-burn-in distribution [6].

The Approximate Bayesian test [7] serves as a computationally economical approximation of MCMC mechanics, utilizing Laplace approximation and Bayes factors. This algorithm rapidly calculates posterior-probability equivalents within alternative reconstructive frameworks, such as Maximum Likelihood, bypassing the temporal constraints of standard Bayesian algorithms.

2.4. Benchmark of Each Method

Our simulation of standard quartet evolution, under constraints lacking evolutionary and substitution rate heterogeneity, demonstrates that all branch support metrics rapidly ascend toward 100% proportionally with increases in corresponding internal branch lengths. Comparatively, the mean values of BPP and aBayes are the earliest to achieve mathematical saturation, maintaining a systematic numerical advantage over ultrafast bootstrap (UFBoot) [8] and SH-aLRT (Figure 2A). This dynamic reflects the diminishing impact of randomized deviation as the aggregate number of substitutions increases. Because BPP algorithms operate without fundamentally altering the underlying dataset through resampling, the resulting support remains remarkably robust following MCMC convergence. The variance (standard deviation) of all observed support values inversely correlates with simulated sequence length (Figure 2B), as expanded data matrices concurrently introduce variable and invariable sites, thereby diluting the susceptibility to randomized stochastic noise.

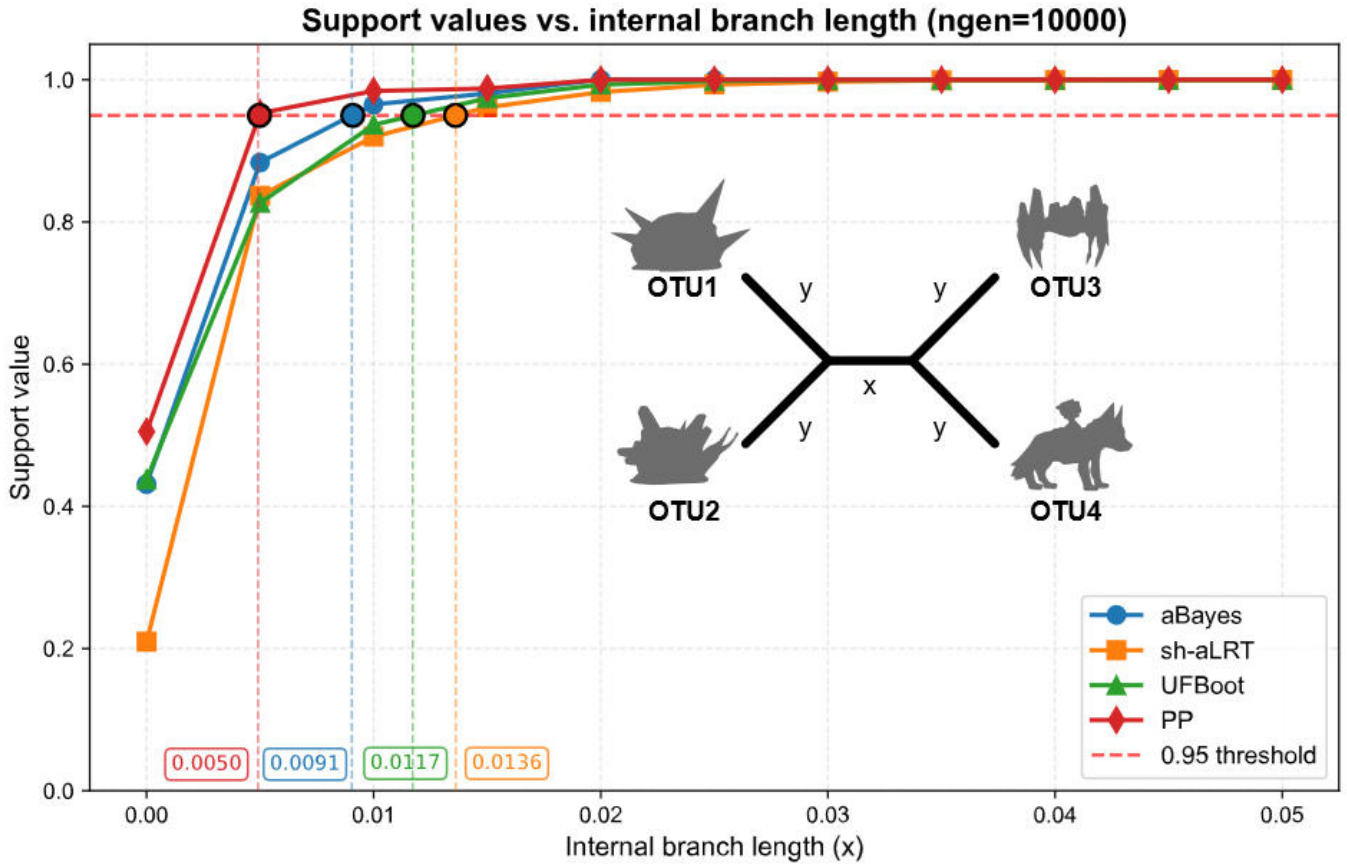

Figure 2: The Variations of 4 Branch Support Values with the Increase of Quartet Tree's Internal Branch Length, and a Diagram of a Quartet

To quantify the structural impact of topological discordance, we executed a secondary simulation utilizing mixed gene trees. The empirical results (Figure 3) illustrate that heterogenous datasets incorporating conflicting topological signals exhibit negligible deterioration in aggregate branch supports, provided the divergent signals are not in perfect equilibrium (e.g., 50% vs. 50%) and total data volume is sufficient. This resilience occurs because algorithmic inference reliably converges upon the dominant phylogenetic signal. Notably, Bayesian methods (BPP and aBayes) demonstrated extreme resistance to topological conflict within our simulation framework, averaging above 95% support even under conditions of perfectly matched signal conflict. This pronounced resilience is intrinsic to the MCMC architecture, which refrains from dataset perturbation, thereby locking the sampling distribution onto the optimal topology post-convergence.

3. Strategies for Branch Support Optimization and Their Implications

Grounded in the empirical mechanics revealed by our simulation data, we present seven distinct methodological strategies frequently utilized to engineer enhanced branch support values. The objective herein is not algorithmic endorsement, but rather the systematic deconstruction of the mechanisms and epistemological costs incurred by catering to the institutionalized demand for artificial 'certainty'.

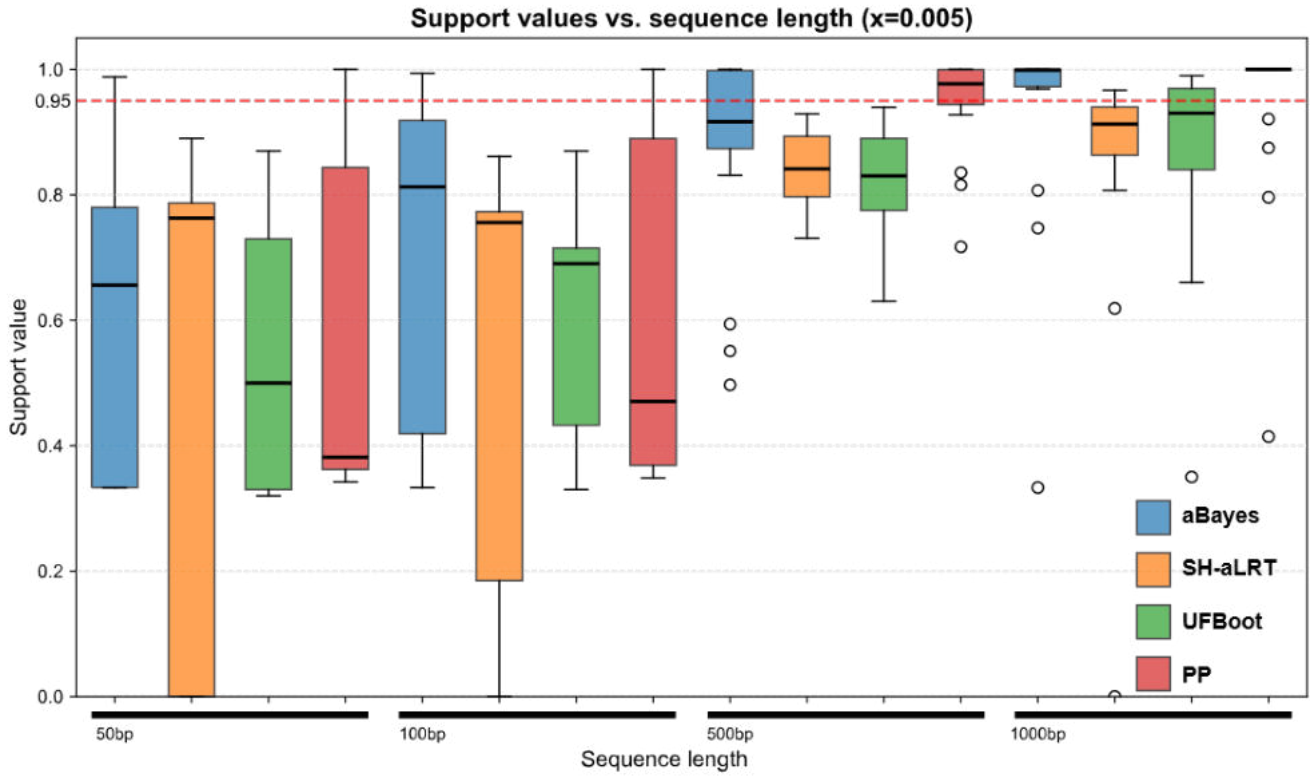

Figure 3: The Variations of 4 Branch Support Values when x is fixed at 0.005 with the Increase of Simulated Sequence Length

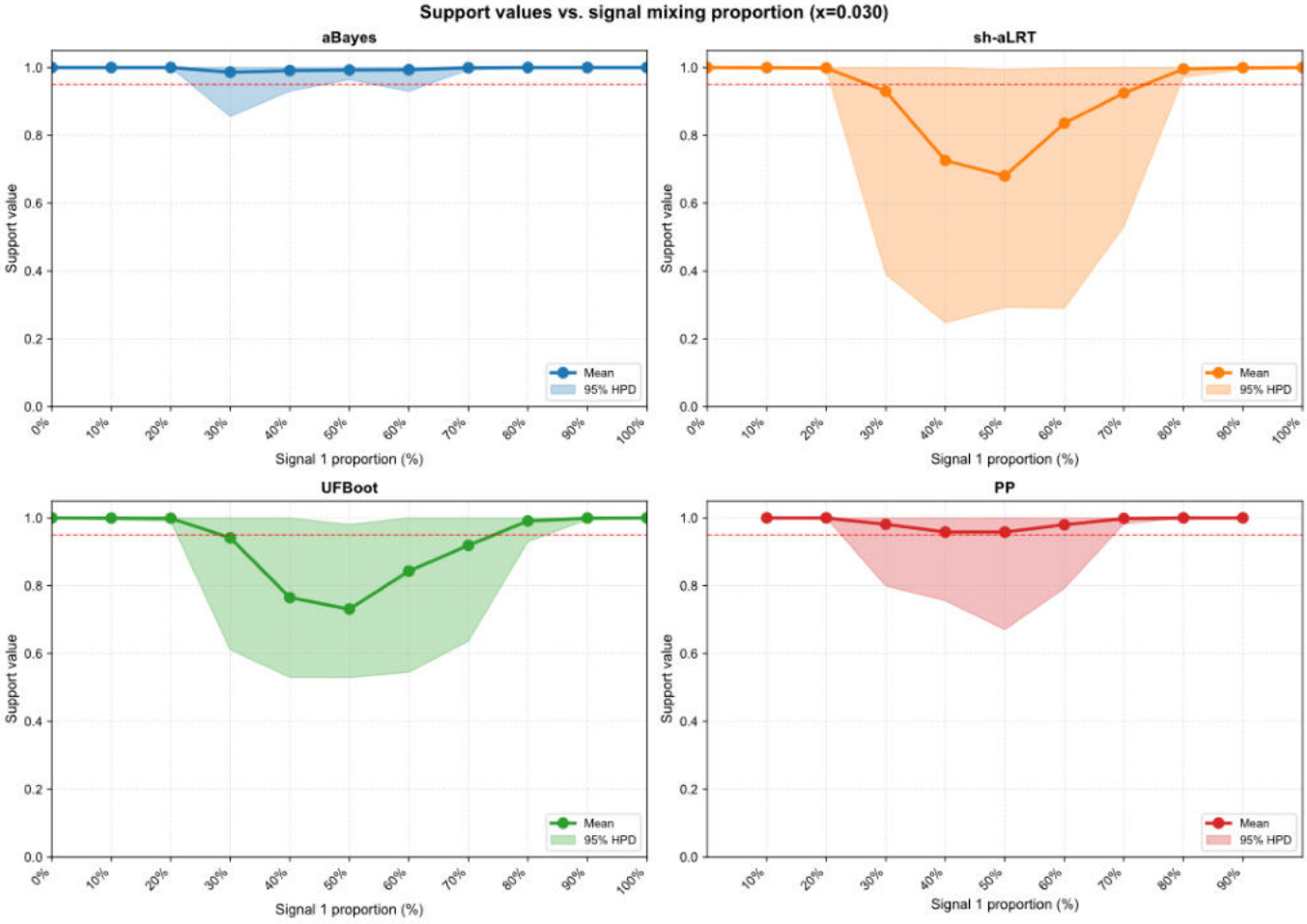

Figure 4: The Variations of 4 Branch Support Values when x is fixed at 0.030 with the Variation of Mixed Topology Signals

3.1. Delete Closely Related OTUs

Initial simulations substantiate that internal branch elongation consistently elevates support metrics; consequently, the deliberate excision of closely related Operational Taxonomic Units (OTUs)—such as recently diverged taxa or members of cryptic species complexes—represents a prevalent strategy to simultaneously sanitize the visual topology and artificially inflate clade support. This heuristic, widely propagated among novice researchers utilizing automated alignment platforms (e.g., NCBI BLASTN), seeks to circumvent the low support values inevitably generated by highly homologous sequences lacking distinct phylogenetic signals. However, the indiscriminate algorithmic purging of proximate taxa represents a manifestation of intellectual inertia, wherein empirical data is structurally manipulated to silence numerical warnings rather than prompting a rigorous re-evaluation of the study's fundamental biological rationale.

3.2. Switch to Non-perturbation Methods (aBayes and BPP)

Our computational benchmarks establish that BPP and aBayes yield systematically elevated support values relative to SH-aLRT and ultrafast bootstraps. Ergo, when Maximum Likelihood algorithms yield suboptimal metrics, researchers are incentivized to transition to Bayesian inference to effectively exploit its inherent statistical leniency. Paradoxically, this optimization pathway remains underutilized; researchers lacking adequate Bayesian expertise are incapable of executing it, while competent bioinformaticians either reject the premise or remain insulated from the unsophisticated evaluative criteria imposed by uncritical reviewers and editors. This methodological misalignment acutely exposes the absurdity of a metrics-driven academic culture: it actively promotes the superficial, utilitarian exploitation of analytical tools over a rigorous comprehension of their statistical boundaries. Consequently, the differential application of Bayesian methods serves as both a mirror reflecting the distortion of methodological integrity by academic incentives, and an indictment of training paradigms that prioritize technical checklists over foundational epistemological comprehension.

3.3. Increase Overall Data Volume of Available Sites

Simulations dictate that expanding total data volume exerts a stabilizing, positive influence on branch support metrics. Therefore, integrating matrices rich in informative loci constitutes a logically sound approach. Conversely, inexperienced researchers frequently execute excessive manual alignment trimming—applying constraints far more aggressive than automated algorithms such as TrimAl [9] or GBlocks [10]—under the misguided assumption that hyper-conservative filtering automatically optimizes inference. This is demonstrably false, as regions containing minor alignment gaps frequently harbor critical phylogenetic signals. The persistence of this detrimental behavior warrants sociological inquiry. It is largely attributable to the pedagogical dominance of specific graphical software (e.g., MEGA) and the uncritical dissemination of unofficial tutorials that advocate manual sequence ablation devoid of any cited mathematical rationale. Because such software inherently incorporates configurable gap-handling algorithms, redundant manual trimming guarantees a compounding loss of phylogenetic information. This phenomenon illustrates a profound deficiency in methodological transmission: the ritualistic mimicry of software interfaces entirely supersedes the comprehension of algorithmic principles and scientific intent. The unchecked proliferation of these erroneous protocols within academic ecosystems represents an epistemological failure of the highest order.

3.4. Producing "Clean" Trees: The Use of Consensus Thresholds to Hide Uncertainty

The majority-rule consensus algorithm (frequently utilizing a 50% operational cutoff) was initially formalized to synthesize divergent maximum parsimony topologies by collapsing nodes exhibiting substandard support [1]. Epistemologically, this operation is justified as generating 'soft polytomies'—legitimate visual representations of regions lacking sufficient phylogenetic signal to support strict bifurcation. Collapsing these uncertain nodes supposedly renders a 'conservative' topology. However, within the 'publish or perish' ecosystem, this statistical summarization tool is aggressively repurposed as an instrument of narrative curation. Imposing an artificially elevated cutoff threshold (e.g., ¿90%) functionally erases all graphical and narrative traces of underlying data uncertainty. The outcome is a hyper-sanitized phylogenetic diagram engineered specifically for audiences fixated solely on support metrics, wherein every visible node projects an illusion of absolute stability. When cutoff thresholds are determined by anticipated peer-review palatability rather than the authentic statistical distribution of the dataset, consensus building devolves into post-hoc data trimming. This rhetorical strategy flourishes because it perfectly satisfies the systemic evaluative preference for absolute, unequivocal certainty, entirely bypassing the requirement to honestly articulate complexity. It empowers researchers to transform statistically ambiguous regions into authoritative, definitive illustrations. Consequently, defining a consensus threshold is rarely an objective mathematical decision; it is a calculated rhetorical compromise brokered between scientific transparency and professional survival.

3.5. Concatenate Multiple Datasets (Regardless of Heterogeneity)

Because singular genetic markers typically possess constrained phylogenetic utility resulting in marginal support values, researchers universally adopt data concatenation, amalgamating multiple genomic loci into a single supermatrix to mathematically force highly-supported topologies. This approach relies on the sweeping assumption that all concatenated partitions adhere to a unified evolutionary history. As confirmed by our simulation data, the sheer volumetric expansion of informative sites aggressively suppresses the variance of subsequent support estimates. While rigorous methodology dictates the preliminary diagnosis of phylogenetic heterogeneity to prevent discordant concatenation, researchers incentivized by metric optimization rely on the principle of dominant signal suppression. The aggregate statistical weight of a concatenated matrix efficiently marginalizes discordant stochastic noise, forcibly inflating support values for the dominant topology regardless of underlying conflict. This specific obfuscation is exceptionally difficult to audit, as numerous studies actively concatenate fundamentally divergent evolutionary markers (e.g., nuclear ribosomal DNA and mitochondrial coding sequences) while tactically burying the irreconcilable single-gene topologies in supplementary files. This curatorial framing ensures that reviewer scrutiny remains anchored to the hyper-supported supermatrix. Consequently, support values derived from uncritical concatenation are not merely statistically inflated; they are rhetorically weaponized to project an artificial consensus that entirely glosses over true evolutionary complexity. When weaponized to satisfy publication mandates, concatenation functions as

an instrument of manufactured certainty, trading biological nuance for editorial clarity.

3.6. Beyond Analysis: The Newick String as a Site of Final Curatorial Intervention

When legitimate computational optimization and structural data manipulation are exhausted, and critical nodes remain stubbornly below arbitrary publication thresholds (e.g., stranded at 87% or 89%), researchers confront an ultimate epistemological boundary. At this juncture, the most extreme intervention materializes: the direct, manual manipulation of the Newick or Nexus text strings defining the phylogenetic output. Because these plain-text formats utilize nested parentheses for topology and simple text labels for support metrics, they are inherently vulnerable to unfettered editorial alteration via any standard text editor. Furthermore, popular graphical user interfaces targeted at novice users (e.g., MEGA [11], DAMBE [12]) structurally facilitate this breach by failing to generate immutable analytical logs or auditable computational trails. When institutional performance pressures cross a critical threshold, Goodhart's Law dictates that researchers will exploit this specific void of unauditability to manufacture the requisite metrics.

3.7. The Punished Diagnostician: Avoiding Heterogeneity-Sensitive Metrics

Contemporary diagnostic metrics, exemplified by the site concordance factor (sCF) [13], are specifically engineered to unmask and quantify topological discordance driven by biological phenomena such as incomplete lineage sorting and introgression. In stark contrast to legacy metrics designed to enforce global conformity, these modern indices explicitly target and measure evolutionary heterogeneity. However, this intrinsic capability to "reveal uncertainty," combined with the profound methodological inertia of established academic networks, severely disadvantages these metrics in an economy driven by universally high support values. Within an evaluation paradigm fixated on absolute topological stability, any metric that honestly depresses numerical output in the presence of conflicting data is structurally penalized as a "liability" rather than celebrated as a "diagnostic instrument". Consequently, the strategic evasion of heterogeneity-sensitive indicators represents a highly rational, albeit epistemologically bankrupt, adaptation to ensure compliance with the narrative demand for universal certainty. Despite emerging protocols advocating for the integration of sCF [14], an empirical survey of 79 recent articles detailing chloroplast phylogenies—datasets notorious for introgression and topological conflict—revealed a complete, systematic omission of sCF reporting. This selective implementation exposes a fundamental systemic paradox: as analytical methodologies advance to accurately capture the chaotic reality of evolutionary history, the academic publication apparatus actively penalizes researchers who deploy these instruments honestly, preferring streamlined, fictitious narratives over the authentic presentation of biological complexity.

4. Conclusion and Discussion

This systematic review computationally evaluates prevailing methodologies utilized to artificially "boost" branch support values within molecular phylogenetics. The empirical data confirm that targeted algorithmic selection (e.g., Bayesian integration), structural data curation (e.g., exclusionary filtering, uncritical concatenation), and narrative suppression (e.g., inflated consensus thresholds) effectively engineer significant numerical enhancements. However, the primary epistemological objective of this treatise is not to validate these operations, but to expose the structural consequences of transmuting a statistical diagnostic tool into a definitive publication prerequisite. This shift catalyzes a metric-optimization culture that fundamentally corrupts methodological integrity and reporting transparency.

Beneath the superficial manipulation of branch supports lies a profound systemic failure. Dry-lab data analytics and molecular phylogenetics constitute rigorous scientific disciplines requiring extensive theoretical and mathematical synthesis; they are not mere computational labor intended for mechanical execution. Current institutional incentive structures frequently prioritize the rapid generation of 'publishable' topologies, reducing graduate training to a process of methodological deskilling. Rather than cultivating critical scientific literacy and epistemological depth, this environment implicitly incentivizes the mass production of algorithmic technicians, trained primarily to execute software pipelines and optimize arbitrary indicators at the expense of genuine intellectual inquiry. The persistence of rudimentary data manipulation strategies indicates a systemic pedagogical failure to transmit foundational mathematical logic and biological data architecture to early-career researchers.

The institutionalized demand for enhanced branch supports renders advisors, peer reviewers, and journal editors fully culpable in the degradation of phylogenetic integrity. By degrading a nuanced measure of statistical uncertainty into a binary gauge of academic 'validity,' the evaluative apparatus fosters a cycle of intellectual decay. The spectrum of manipulative strategies detailed herein—from selective data curation to direct epistemological breaches (text-editing)—thrives precisely because the prevailing review culture assesses numerical outcomes while remaining entirely agnostic to the analytical logic utilized to generate them. Consequently, academic publishing degenerates into the mechanized batch production of mediocrity, actively suppressing genuine scientific innovation.

For academic mentors and institutional leaders, the proliferation of these practices should prompt severe introspection regarding the operational strategies they incentivize, and the distinct lack of intellectual vitality those strategies yield compared to tier-one research environments.

5. Materials & Methods

All quartet phylogenetic tree inferences were simulated utilizing AliSim [15] within the IQ-Tree 3 framework [16], governed by the HKY85 evolutionary model with equal base frequencies. Robustness was ensured via 20 replicates per simulation. Maximum likelihood and Bayesian inference topologies were reconstructed using IQ-Tree 3 and MrBayes-MPI 3.2.7a [17], respectively, utilizing the BEAGLE 3 [18] library for double-precision likelihood acceleration. Approximate Bayesian tests (aBayes), SH-aLRT, and ultrafast bootstraps (1000 replicates) were executed via IQ-Tree. Bayesian posterior probabilities (BPP) were consolidated by MrBayes via a 50% majority-rule consensus protocol. Two primary simulation branches were executed: 1) balanced quartets maintaining constant total divergence time but exhibiting variable internal branch lengths; 2) fixed topologies (x=0.005) subject to variable matrix sequence lengths.

Data Availability

All desensitized computational simulation results and accompanying Python execution scripts are archived within a compressed tar file as supplementary material.

Statement

Statement of Epistemological Hazard: While the methodological manipulations detailed herein are computationally reproducible, their practical application is strictly intended as a structural critique of academic publishing metrics. Readers are unequivocally advised against implementing these optimization strategies in empirical research. The workflows are presented *as-is*, devoid of any warranty—implied or explicit—regarding their scientific integrity, MERCHANTABILITY, or FITNESS FOR A PARTICULAR PURPOSE, and serve solely to expose the fragility of indicator-centric evaluation systems.

Acknowledgements

We extend our gratitude to researchers who engage with the theoretical architecture of phylogenetic diagnostics, yet rigorously abstain from applying the manipulative methodologies exposed within this text.

References

[1] M. Nei and S. Kumar, *Molecular Evolution and Phylogenetics*. New York, NY: Oxford University Press, 2000. [Online]. Available: https://doi.org/10.1093/oso/9780195135848.001.0001
[2] J. Felsenstein, "Confidence limits on phylogenies: An approach using the bootstrap," *Evolution*, vol. 39, no. 4, pp. 783–791, 1985. [Online]. Available: https://doi.org/10.1111/j.1558-5646.1985.tb00420.x
[3] R. Zardoya and A. Meyer, "Phylogenetic performance of mitochondrial protein-coding genes in resolving relationships among vertebrates," *Molecular Biology and Evolution*, vol. 13, no. pp. 933–942, 1996. [Online]. Available: https://doi.org/10.1093/oxfordjournals.molbev.a025661
[4] M. Cummings, S. Otto and J. Wakeley, "Sampling properties of DNA sequence data in phylogenetic analysis." *Molecular biology and evolution*, vol. 12, no. 5, pp. 814–822. [Online]. Available: https://doi.org/10.1093/oxfordjournals.molbev.a040258
[5] H. Shimodaira and M. Hasegawa, "Multiple comparisons of log-likelihoods with applications to phylogenetic inference," *Molecular Biology and Evolution*, vol. 16, no. 8, p. 1114, 1999. [Online]. Available: https://doi.org/10.1093/oxfordjournals.molbev.a026201
[6] Z. Yang, Computational Molecular Evolution. Oxford: Oxford University Press, 2006. [Online]. Available: https://doi.org/10.1093/acprof:oso/9780198567028.001.0001
[7] M. Anisimova, M. Gil, J.-F. Dufayard, C. Dessimoz, and O. Gascuel, "Survey of branch support methods demonstrates accuracy, power, and robustness of fast likelihood-based approximation schemes," *Systematic Biology*, vol. 60, no. 5, pp. 685–699, 2011. [Online]. Available: https://doi.org/10.1093/sysbio/syr041
[8] D. T. Hoang, O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh, "UFBoot2: Improving the ultrafast bootstrap approximation," *Molecular Biology and Evolution*, vol. 35, no. 2, pp. 518–522, 2018. [Online]. Available: https://doi.org/10.1093/molbev/msx281
[9] S. Capella-Gutiérrez, J. M. Silla-Martínez, and T. Gabaldón, "trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses," *Bioinformatics*, vol. 25, no. 15, pp. 1972–1973, 2009. [Online]. Available: https://doi.org/10.1093/bioinformatics/btp348
[10] G. Talavera and J. Castresana, "Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments," *Systematic Biology*, vol. 56, no. 4, pp. 564–577, 2007. [Online]. Available: https://doi.org/10.1080/10635150701472164
[11] S. Kumar, G. Stecher, M. Suleski, M. Sanderford, S. Sharma, and K. Tamura, "MEGA12: Molecular Evolutionary Genetics Analysis version 12 for adaptive and green computing," *Molecular Biology and Evolution*, vol. 41, no. 12, p. msae263, 2024. [Online]. Available: https://doi.org/10.1093/molbev/msae263
[12] X. Xia, "DAMBE7: New and improved tools for data analysis in molecular biology and evolution," *Molecular Biology and Evolution*, vol. 35, no. 6, pp. 1550–1552, 2018. [Online]. Available: https://doi.org/10.1093/molbev/msy073
[13] B. Q. Minh, M. W. Hahn, and R. Lanfear, "New methods to calculate concordance factors for phylogenomic datasets," *Molecular Biology and Evolution*, vol. 37, no. 9, pp. 2727–2733, 2020. [Online]. Available: https://doi.org/10.1093/molbev/msaa106
[14] J. Winn, A. Bester-Van Der Merwe and S. Maduna, "Annotated Bioinformatic Pipelines for Phylogenomic Placement of Mitochondrial Genomes," *Bio-protocol*, vol. 15, no. 5, e5232, 2025. [Online]. Available: https://doi.org/10.21769/BioProtoc.5232
[15] N. Ly-Trong, S. Naser-Khdour, R. Lanfear, and B. Q. Minh, "AliSim: A fast and versatile phylogenetic sequence simulator for the genomic era," *Molecular Biology and Evolution*, vol. 39, no. 5, p. msac092, 2022. [Online]. Available: https://doi.org/10.1093/molbev/msac092
[16] T. K. F. Wong et al., "IQ-TREE 3: Phylogenomic inference software using complex evolutionary models," *Molecular Biology and Evolution*, 2025. [Online]. Available: https://doi.org/10.32942/X2P62N
[17] F. Ronquist et al., "MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space," *Systematic Biology*, vol. 61, no. 3, pp. 539–542, 2012. [Online]. Available: https://doi.org/10.1093/sysbio/sys029
[18] D. L. Ayres et al., "BEAGLE 3: Improved performance, scaling, and usability for a high-performance computing library for statistical phylogenetics," *Systematic Biology*, vol. 68, no. 6, pp. 1052–1061, 2019. [Online]. Available: https://doi.org/10.1093/sysbio/syz020