

Research Article

The Interaction between Base Compositional Heterogeneity and Among-Site Rate Variation in Models of Molecular Evolution

Nathan C. Sheffield

Program in Computational Biology and Bioinformatics, Institute for Genome Sciences and Policy, Duke University, 101 Science Drive, Box 3382, Durham, NC 27708, USA

Correspondence should be addressed to Nathan C. Sheffield; nathan.sheffield@duke.edu

Received 1 November 2012; Accepted 27 November 2012

Academic Editors: J. Fan, G. Glöckner, and J. A. Norman

Copyright © 2013 Nathan C. Sheffield. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Many commonly used models of molecular evolution assume homogeneous nucleotide frequencies. A deviation from this assumption has been shown to cause problems for phylogenetic inference. However, some claim that only extreme heterogeneity affects phylogenetic accuracy and suggest that violations of other model assumptions, such as variable rates among sites, are more problematic. In order to explore the interaction between compositional heterogeneity and variable rates among sites, I reanalyzed 3 real heterogeneous datasets using several models. My Bayesian inference recovers accurate topologies under variable rates-among-sites models, but fails under some models that account for compositional heterogeneity. I also ran simulations and found that accounting for rates among sites improves topology accuracy in compositionally heterogeneous data. This indicates that in some cases, models accounting for among-site rate variation can improve outcomes for data that violates the assumption of compositional homogeneity.

1. Introduction

Recent phylogenetic studies have explored the effect of compositional heterogeneity on phylogenetic methods. Compositional heterogeneity can arise in a dataset as a result of nonstationary evolution (when the substitution pattern is not uniform across an evolutionary tree). If two nonsister subtrees have similar substitution bias, this can lead to a convergence in nucleotide composition (CNC). The taxa may then look similar due to convergent evolution rather than common ancestry, which can mislead phylogenetic analysis. There are several methods to detect and quantify the level of compositional heterogeneity in a dataset, including chi-squared tests (e.g., [1]), Disparity Index [2], and relative-rates tests [3]. When found, the presence of compositional heterogeneity is often assumed to cause problems for both parametric and nonparametric phylogenetic methods [4]. However, this assumption has been challenged; Conant and Lewis [5] claimed that “extreme amounts of heterogeneity must be present before it can mislead phylogenetics” and Rosenberg and Kumar [6] “did not find a significant interaction between phylogenetic accuracy and substitution pattern heterogeneity among lineages.”

Another commonly studied modeling question is the variation of substitution rates among sites. It has been established that accounting for among-site rate variation is important in phylogenetics [7]. This is most commonly done by assuming the substitution rates among sites vary according to a discrete gamma distribution with a fixed number of categories. Models incorporating this idea are usually annotated with “+G.” This allows the model to account for some nucleotides evolving faster than others. In a slight variation, some models allow sites to be assigned to an invariable category (“+I”), which works in roughly the same way but fixes the substitution rate at 0, instead of varying according to a gamma parameter.

Several studies, including those mentioned above, have suggested that violations of the assumption of constant rates among sites are more problematic than that of compositional homogeneity [7–10]. This question is important because it directs us to consider the models we use. Model choice requires a balance between model complexity, computational feasibility, and data availability. As Box and Draper [11] famously wrote, “all models are wrong, but some are useful.” In an evolutionary context, any model will necessarily simplify reality, but if a model can yield appropriate relationships,

it retains utility. In statistical terms, models that “work” despite small deviations from assumptions are “robust.” The question outlined above is essentially a question of robustness: is a model that accounts for among-site rate variation robust to violations of the assumption of compositional homogeneity?

In order to further explore the interaction between base compositional heterogeneity and among-site rate variation, I selected 3 recent empirical datasets that exhibit base compositional heterogeneity and re-analyzed them using both gamma-distributed rates models (+G and +I) and constant rate models. I also ran a simulation study to see if accounting for variable rates among sites helps infer accurate topologies from compositionally heterogeneous data.

Gruber et al. [12] sequenced the RAG1 locus for 41 species of didelphid marsupials. They combined these data with two other genes for phylogenetic inference. They found high levels of base compositional heterogeneity in the RAG1 third codon position and reported that this “convergence in base composition is strong enough to overwhelm phylogenetic signal from other genes” [12]. This dataset appears to be one of the most compelling examples of compositional bias, particularly due to the low rank of the taxa under scrutiny. They analyzed the data using several inference methods and showed that most analyses reported a particular erroneous clade (clade B + I). Among the phylogenetic reconstruction methods they employed two are designed to compensate for base compositional heterogeneity: LogDet [13] and *nhml* [14]. LogDet uses matrix determinants to construct a distance metric that is more robust to datasets with compositional heterogeneity; *nhml* introduces a “GC-content” parameter for each branch on the tree. Neither of these methods was able to overcome the artifact caused by compositional bias. The analysis by Gruber et al. [12] thus comprehensively addressed the level of base compositional heterogeneity. They also accounted for among-site rate variation by reporting the result of a maximum likelihood (ML) analysis using the GTR+I+G model for the RAG1 locus; however, they did not report any results that accounted for among-site rate heterogeneity on the whole dataset.

In another study, Ho and Jermini [4] reported the results of an analysis of beta-tubulin in Metazoa. This 6-taxon dataset included 3 mammals, 2 insects (moth and fruit fly), and an octopus. Using a relative-rates test [15], they detected significant rate differences among lineages that has led to a convergence in nucleotide composition between the moth (*Heliothis*) and octopus. They ran a series of analyses using the F84 model of nucleotide substitution and reported that the moth and octopus universally grouped together to the exclusion of the fruit fly. This problem could be partially alleviated with RY-coding of the third codon sites, presumably because RY-coding extricates the erroneous composition signals. These authors acknowledged the effect of among-site rate variation, but did not further explore any models that account for it.

Mallatt et al. [16] found a similar issue in a 43-taxon dataset of a quite divergent group of ecdysozoans (molting animals). Using a chi-squared test, they found significant

nonstationary nucleotide frequency ($P < 5 \times 10^{-9}$). In the face of this nonstationarity, they tried LogDet+I, as well as a Bayesian method that accounts for among-site rate variation (using GTR+I+G). One may expect LogDet to outperform a stationary method in such a dataset; however, Mallatt et al. [16] reported that the Bayesian analysis outperformed the LogDet+I analysis, which incorrectly placed several taxa with long branches. These results illustrate that, despite the base compositional bias, the Bayesian GTR+I+G model succeeded. Does this model succeed because it accounts for among-site rate variation? What would happen to the inference under the GTR model alone?

In each of these 3 studies, the authors found significant compositional heterogeneity in a dataset. They ascribed failure of many phylogenetic methods to the confounding signal of the convergent nucleotide composition, but the first 2 reports did not thoroughly explore models that account for among-site rate variation. In the third study, Mallatt et al. [16] showed a GTR+I+G to work despite compositional heterogeneity, but whether the improvement is due to accounting for among-site rate variation is unknown, as an analysis with GTR alone, was not presented.

To elucidate the effects of among-site rate heterogeneity in datasets with compositional heterogeneity, I re-analyzed these 3 datasets using Bayesian phylogeny software MrBayes version 3.1.2 under several different evolutionary models. My results show that when accounting for among-site rate heterogeneity, Bayesian inference helps in each study; in each case, I find the GTR+I+G, GTR+I and GTR+G models to outperform the GTR model alone.

2. Methods

I used the data matrix of Gruber et al. [12] as provided by these authors, and downloaded the datasets of Ho and Jermini [4] and Mallatt et al. [16] directly from their websites. I ran each of the 3 datasets through a series of Bayesian analyses using MrBayes v. 3.1.2 [17] under the following conditions: 4 runs with 4 chains each for 10 million generations, sampling every 1000 trees. I divided the dataset of Gruber et al. [12] into 4 partitions (RAG1, DMP1, IRBP, and morphology). For each dataset, I applied GTR, GTR+I, GTR+G, and GTR+I+G models in 4 separate analyses. I also tried the F81 and HKY models on the Ho and Jermini [4] dataset to confirm their results. To assess convergence, I monitored log likelihood traces in R and posterior clade probabilities in AWTY [18]. I discarded 1-2 million generations as burn-in in each case. I also confirmed some of the results using LogDet and LogDet+I as implemented in PAUP* 4b10 [19]. I used FigTree v1.1.1 (<http://tree.bio.ed.ac.uk/software/figtree/>) to produce the topology figures.

In order to more thoroughly explore this interaction, I constructed a simulation test using p4 [20]. I simulated data under varying degrees of compositional heterogeneity and then reconstructed phylogenies under models that either do or do not account for among-site rate variation. I first defined a model (a 7-taxon tree, 1000 nucleotides, plus composition vectors and substitution matrix with equal rates),

with 3 separate parameter sets. I included rates-among-sites variation (gamma parameter = .2). The tree, rates-among-sites parameter, and substitution matrix were held constant. I varied the composition vectors (G+C = .8, .7, or .5) to yield a dataset with either high, low, or no convergence in nucleotide composition. I then generated 100 simulated datasets for each of these 3 parameter sets. I used PhyML [21] to find the maximum likelihood tree under two models: either GTR or GTR+G. To evaluate the results, I calculated the Robinson-Foulds distance (implemented in p4) between each PhyML tree and the true tree used to generate the data [22]. I used a one-tailed Mann-Whitney test to check for significant difference between the distribution of Robinson-Foulds distances for each model.

3. Results and Discussion

3.1. Empirical Datasets. In all 3 cases, I find that accounting for among-site rate variation with the GTR+I+G model in Bayesian inference (as implemented in MrBayes) recovers the expected relationships, despite the model violation of compositional heterogeneity. Assuming the “expected” relationships are actually the correct relationships (which seems reasonable in these datasets), in these datasets, the GTR+I+G model is robust to this assumption violation. This is not to say that the GTR+I+G model is the “correct” model for the data; on the contrary, the demonstrated heterogeneity violates the assumptions of the model. However, it does indicate that the model is robust to this violation, at least to the extent that it yields the accepted relationships. In this case, I have not considered the branch lengths, which are more difficult to assess. To conclude, in these 3 datasets, it is not necessary to account for nonstationary evolution to recover the established relationships. It is necessary, instead, to account for among-site rate variation.

For the Gruber et al. dataset, the Bayesian analysis under models GTR+I+G, GTR+I, and GTR+G all converged on an identical topology with the expected clade grouping (Figure 1). This topology is remarkably consistent with the “correct” topology presented by Gruber et al. [12]. It matches closely the topology they obtained using IRBP, DMP1, and morphology, to the exclusion of RAG1. Clades B, C, H, and I match Figure 1 from Gruber et al. [12]. Under the GTR model (no among-site rate heterogeneity), however, MrBayes reported a topology that closely matched that reported by Gruber et al. [12] for parsimony, LogDet, and *nhml* using the combined dataset (Figure 2). This topology has clade B sister to clade I, which are the two high-GC% clades. Bayesian analysis under GTR, therefore, recovered the erroneous clade just as the other methods that do not account for among-site rate heterogeneity.

For the dataset of Ho and Jermini [4], I recovered the original erroneous tree when using either the F81 or HKY models. However, I find that using a GTR model almost solves the problem; this analysis recovers the correct relationship ((moth + fly) + octopus), though the posterior probability is rather low (57%, Figure 3(b)). Under the GTR+I+G model, posterior support for the correct clade increases to 100%

(Figure 3(c)). In this case, it appears that the slightly more complex evolutionary model (GTR versus HKY/F84) is able to solve the issue to some extent, but accounting for among-site rate variation is key to recovering the correct relationship with high support.

According to the results of Mallatt et al. [16], two insect taxa (*Drosophila* and *Aedes*) incorrectly group with a clade of shrimp, rendering Hexapoda paraphyletic; and *Hanseniella* (a myriopod) incorrectly falls outside Myriapoda. The result is the same whether using LogDet alone or LogDet+I. I confirmed that the Bayesian analysis under GTR+I+G correctly places these “problem taxa” (Figure 4(a)). When I reran the analysis using the GTR model (without gamma-distributed among-site rate variation), I recovered a topology with the incorrect groupings (Figure 4(b)). Therefore, correcting for among-site rate variation appears to be necessary for this dataset as well (although some of the internal Hexapoda relationships are more realistic in the GTR tree).

In all of these datasets, the authors demonstrated significant deviation in base frequency; however, the results of the Bayesian analysis show that nonstationarity is not the prevalent model violation in these datasets. Rather, using the GTR+I+G model in MrBayes to account for among-site rate variation is sufficient to recover the more accurate result. In fact, accounting for either invariable sites (+I) or gamma-distributed rate variation (+G) alone was sufficient to recover the correct topology (in the dataset of Ho and Jermini, they both gave increases to the posterior support of the correct grouping). However, using both invariable-sites and gamma-distributed rates consistently gave the best results (as reflected in higher posterior probabilities for “correct” clades). These two models account for positional rate variation in a similar manner [7]. Where invariable sites models allow only two rate categories, zero, or positive variation, gamma-distributed rates models group characters according to how much they change (it is possible that one of the groups could have a zero rate, degenerating into an invariable category). That each can work without the other is not too surprising; researchers have cited invariable site models as a reasonable approximation for the gamma-distribution model [9, 23]. Including both is most often the most realistic for real sequences.

In each case, there is compelling evidence for nonstationary evolutionary rates, but in each case, correct relationships are only recovered with high support under models that account for among-site rate variation. These examples appear to support the conclusion of Conant and Lewis [5]: “convergence in nucleotide composition alone is insufficient to cause any commonly used methods to fail, and accounting for other evolutionary factors (such as site-to-site rate heterogeneity) can give a correct inference without accounting for CNC.” It is now clearer why LogDet and *nhml* were not able to recover correct relationships in these datasets. Without incorporating among-site rate variation, it is less surprising that they recover the incorrect topology. It is interesting that LogDet+I, which does model invariable sites, also fails in the study of Mallatt et al. [16]; probably the simplistic distance-measure framework of LogDet is simply not enough to overcome the misleading signal. Other studies have had success using LogDet+I [1,

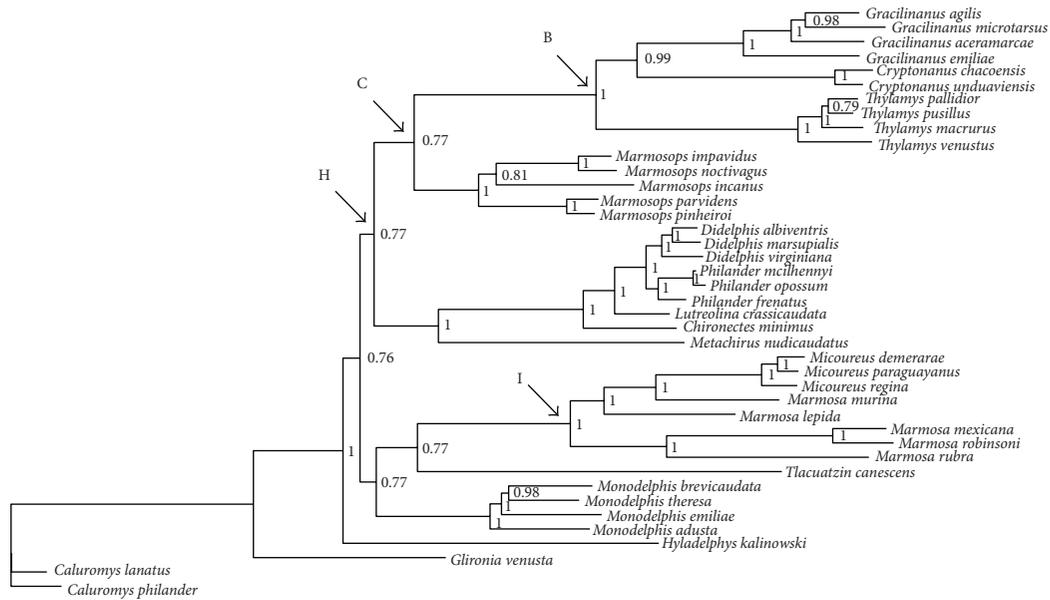


FIGURE 1: Dataset of Gruber et al. [12]. Consensus topology reported by MrBayes using the GTR+I+G model. Clade C is the established relationship not recovered unless among-site rate variation is included in the model.

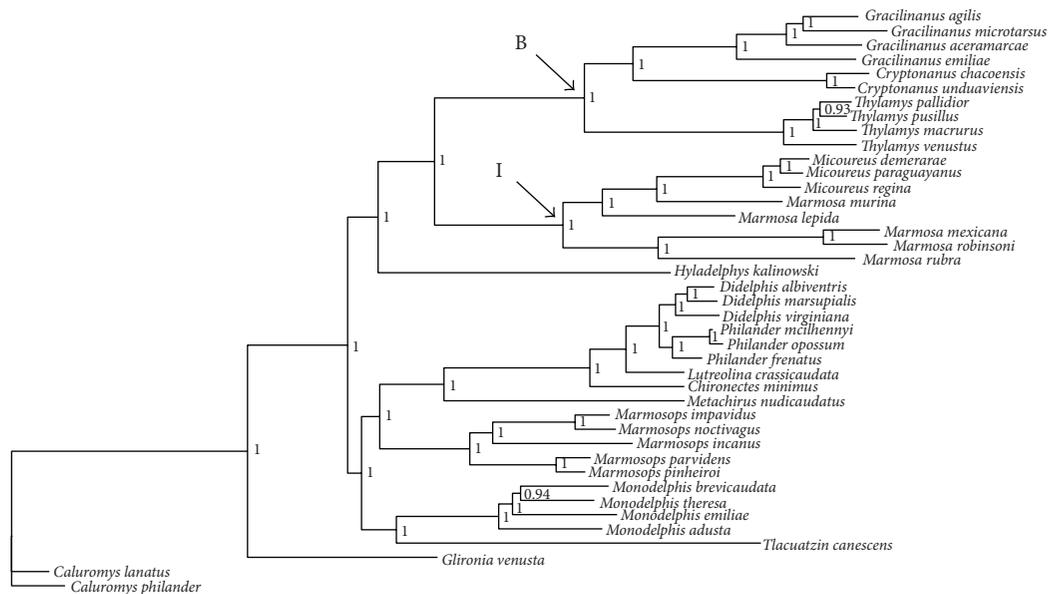


FIGURE 2: Dataset of Gruber et al. [12]. Consensus topology reported by MrBayes using the GTR model. Clades I and B are high AT% sister clades, corresponding to the parsimony topology of Gruber et al. [12].

24]. There are also newer models that build among-site rate variation more appropriately into the nonstationary models; *nhml* has been superseded by *nhPhyML* [25], and LogDet can be adjusted to incorporate gamma-distributed among-site rate heterogeneity [10, 26].

3.2. *Simulation Results.* The results of my simulations are consistent with the results from the empirical data. I find that the GTR+G model outperforms the GTR model only on the

data with significant compositional heterogeneity (Figure 5). In the simulation, when the composition vectors are uniform throughout the tree (so there is no convergence in nucleotide composition), the GTR and the GTR+G models perform similarly. There is no significant difference in the distribution of Robinson-Foulds distances between the reconstructed tree and the true tree. This indicates that the low level of among-site rate variation is not enough to mislead the GTR model. However, when I adjust the composition vectors in the generative model to create a convergence of nucleotide

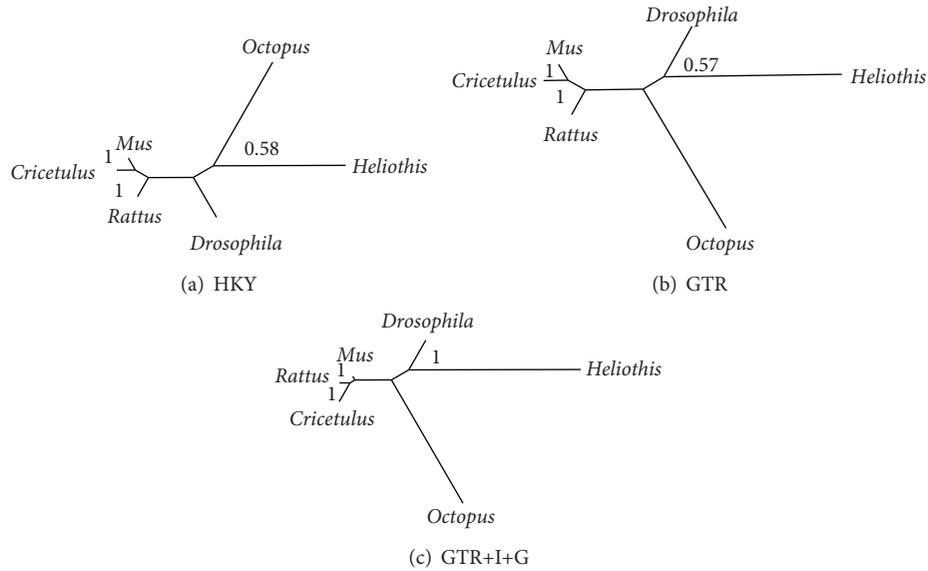


FIGURE 3: Dataset of Ho and Jermin [4]. Consensus topology reported by MyBayes under 3 different nucleotide substitution models: (a) HKY, (b) GTR, (c) GTR+I+G.

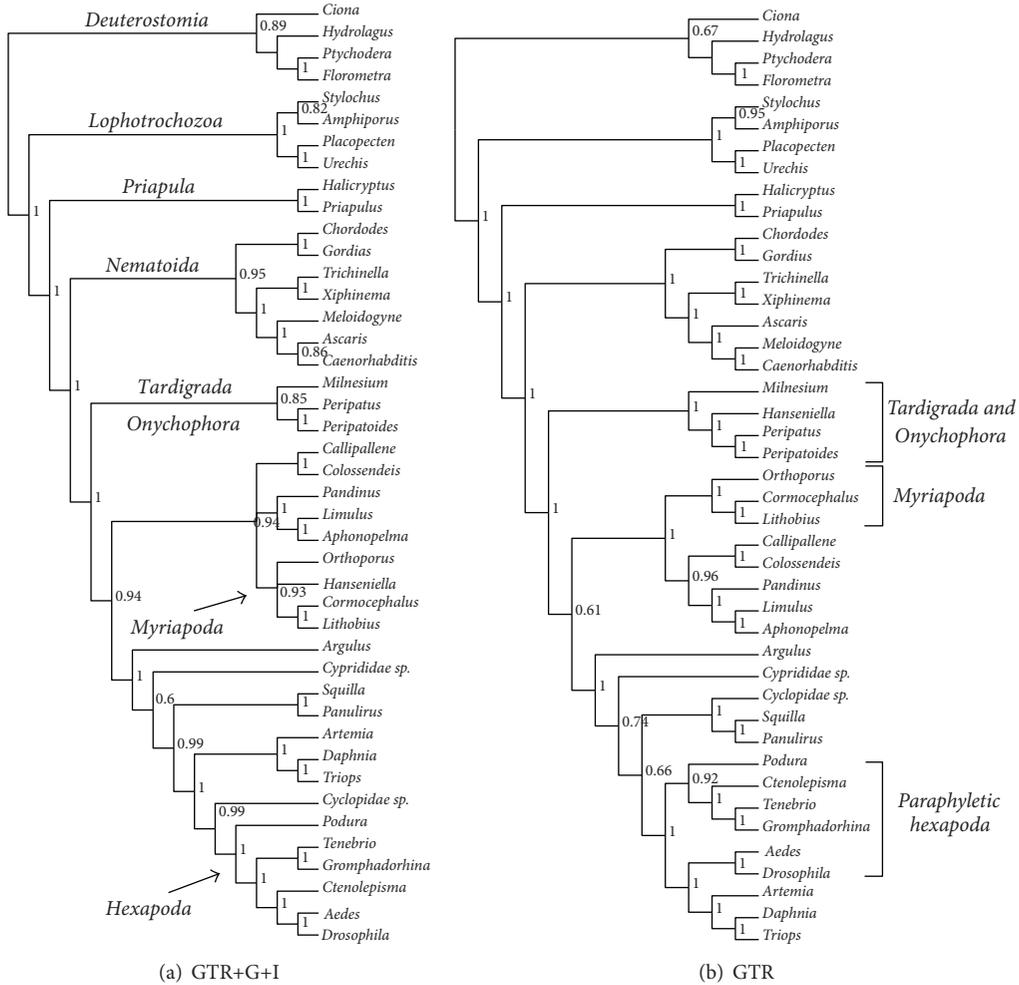


FIGURE 4: Dataset of Mallatt et al. [16]. Consensus topology reported by MrBayes under (a) GTR+I+G and (b) GTR. Two insects, *Drosophila* and *Aedes*, and one myriapod, *Hanseniella*, group incorrectly under the GTR model.

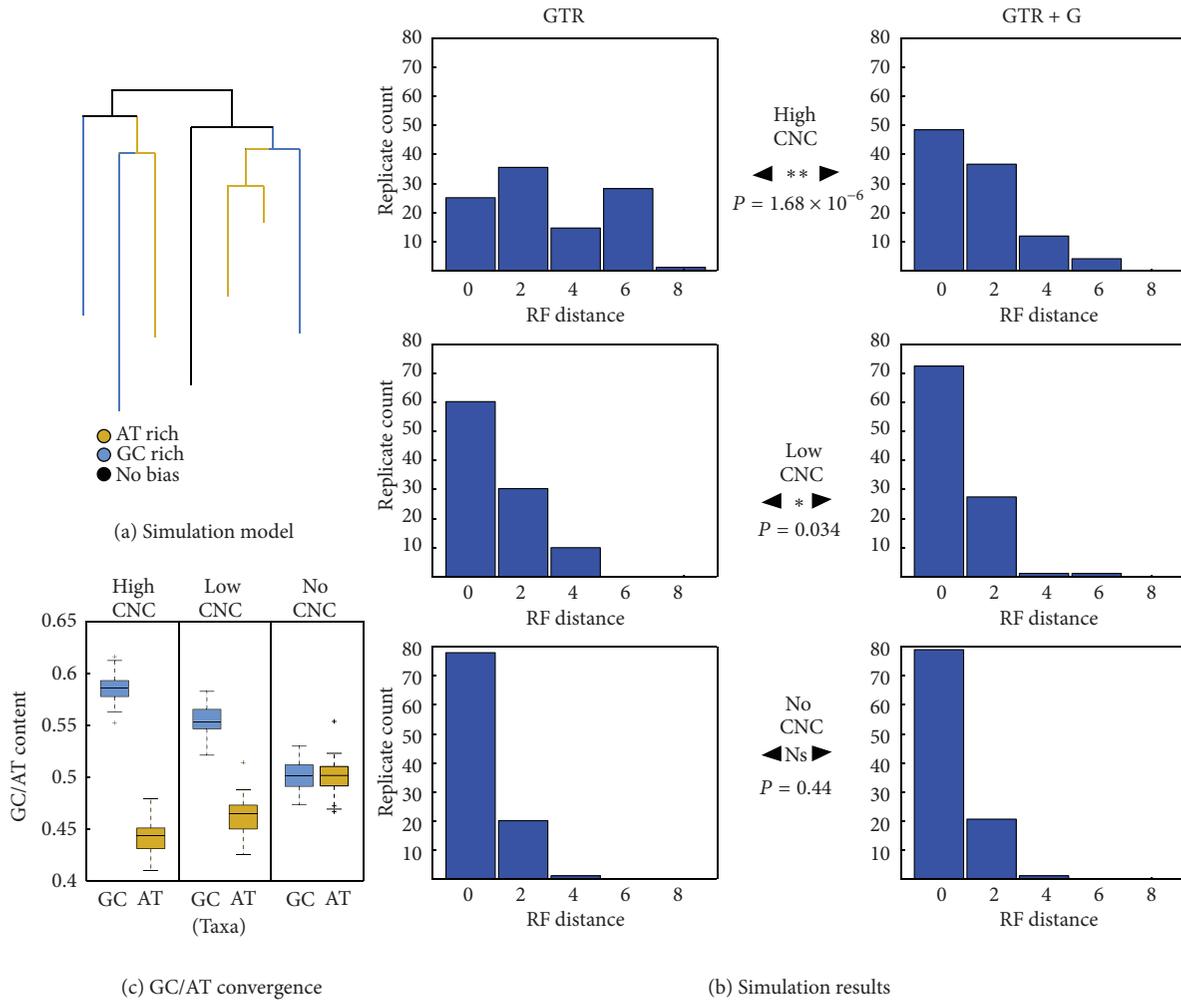


FIGURE 5: Simulation model and results. (a) Tree from which the sequences were simulated. The colored branches show the different composition vectors, which drove the simulated sequences toward a convergence in nucleotide composition. I ran this simulation with 3 different levels of bias; each level had 100 simulation replicates. (b) Histograms showing the distribution of Robinson-Foulds distances (RF distance) between the maximum likelihood tree and the true tree. The maximum likelihood tree was generated by PhyML for 100 simulated replicates for each of the 3 parameter settings (high, low, and no CNC). The difference between the models (GTR versus GTR+G) grows as the level of CNC increases (**: highly significant; *: marginally significant; NS: not significant). In the simulation with no CNC, the GTR model performs as well as the GTR+G model. (c) Boxplot showing the distribution of bias introduced by the different composition vectors, under each of the 3 simulation parameter settings.

composition, the GTR+G model scores significantly better than the GTR model. In the final case, with extreme convergence of nucleotide composition, the improvement is highly significant. This illustrates a clear trend; the more convergence in nucleotide composition, the better then GTR+G model performs relative to the GTR model. Accounting for among-site rate variation has a clear benefit in combating convergence of nucleotide composition.

How could accounting for among-site rate variation contribute to correct inferences in heterogeneous data? If the model is able to assign the sites with convergent signal to the highest rate categories, these sites would be downweighted in the likelihood and contribute less to the tree inference. In my simulation study, there was a clear trend; the most convergent sites were those that were assigned to the highest

rate category. The +G model improves the results by assigning these sites a higher substitution rate. In this way, accounting for among-site rate variation is able to contribute to reducing the effect of compositional heterogeneity.

The importance of accounting for among-site rate variation shown here is not new; however, its importance has perhaps been forgotten as recent studies focus on other issues. There has been growing interest in nonstationary evolution and how it affects phylogenetic methods [10]. Several software packages [20, 27, 28] have been developed that implement models that address this issue. While it is clear that this problem is real and problematic for some datasets [29], it is still unclear how this factor interacts with other evolutionary phenomena, as illustrated here. Identifying compositional heterogeneity in a dataset is not

proof that it is the cause of erroneous phylogenetic results. Unfortunately, it is not so straightforward to discover and decode reasons for method failure. The take-home message is this: model violations interact in complex ways, and phylogenetic methods may be more robust to some model violations than to others. Because we will never have a model that captures reality completely, it is important to understand how robust our models are to violated assumptions. Future studies should be able to quantify this more accurately as more datasets are presented and analyzed.

Acknowledgments

Thanks to authors of the original data for providing their original data matrices. The author would also like to thank Hojun Song and Michael Whiting for their guidance and advice, and the BYU Department of Biology for computational resources.

References

- [1] J. Mallatt and J. Sullivan, "28S and 18S rDNA sequences support the monophyly of lampreys and hag fishes," *Molecular Biology and Evolution*, vol. 15, no. 12, pp. 1706–1718, 1998.
- [2] S. Kumar and S. R. Gadagkar, "Disparity index: a simple statistic to measure and test the homogeneity of substitution patterns between molecular sequences," *Genetics*, vol. 158, no. 3, pp. 1321–1327, 2001.
- [3] S. V. Muse and B. S. Weir, "Testing for equality of evolutionary rates," *Genetics*, vol. 132, no. 1, pp. 269–276, 1992.
- [4] S. Y. W. Ho and L. S. Jermini, "Tracing the decay of the historical signal in biological sequence data," *Systematic Biology*, vol. 53, no. 4, pp. 623–637, 2004.
- [5] G. C. Conant and P. O. Lewis, "Effects of nucleotide composition bias on the success of the parsimony criterion in phylogenetic inference," *Molecular Biology and Evolution*, vol. 18, no. 6, pp. 1024–1033, 2001.
- [6] M. S. Rosenberg and S. Kumar, "Heterogeneity of nucleotide frequencies among evolutionary lineages and phylogenetic inference," *Molecular Biology and Evolution*, vol. 20, no. 4, pp. 610–621, 2003.
- [7] Z. Yang, "Among-site rate variation and its impact on phylogenetic analyses," *Trends in Ecology and Evolution*, vol. 11, no. 9, pp. 367–372, 1996.
- [8] R. A. van den Bussche, R. J. Baker, J. P. Huelsenbeck, and D. M. Hillis, "Base compositional bias and phylogenetic analyses: a test of the "flying DNA" hypothesis," *Molecular Phylogenetics and Evolution*, vol. 10, no. 3, pp. 408–416, 1998.
- [9] M. Steel, D. Huson, and P. J. Lockhart, "Invariant sites models and their use in phylogeny reconstruction," *Systematic Biology*, vol. 49, no. 2, pp. 225–232, 2000.
- [10] N. Rodríguez-Ezpeleta, H. Brinkmann, B. Roure, N. Lartillot, B. F. Lang, and H. Philippe, "Detecting and overcoming systematic errors in genome-scale phylogenies," *Systematic Biology*, vol. 56, no. 3, pp. 389–399, 2007.
- [11] G. E. P. Box and N. R. Draper, *Empirical Model-Building and Response Surfaces*, John Wiley & Sons, New York, NY, USA, 1987.
- [12] K. F. Gruber, R. S. Voss, and S. A. Jansa, "Base-compositional heterogeneity in the RAG1 locus among didelphid marsupials: implications for phylogenetic inference and the evolution of GC content," *Systematic Biology*, vol. 56, no. 1, pp. 83–96, 2007.
- [13] P. J. Lockhart, M. A. Steel, M. D. Hendy, and D. Penny, "Recovering evolutionary trees under a more realistic model of sequence evolution," *Molecular Biology and Evolution*, vol. 11, no. 4, pp. 605–612, 1994.
- [14] N. Galtier and M. Gouy, "Inferring pattern and process: maximum-likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis," *Molecular Biology and Evolution*, vol. 15, no. 7, pp. 871–879, 1998.
- [15] W. H. Li, C. I. Wu, and C. C. Luo, "A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes," *Molecular Biology and Evolution*, vol. 2, no. 2, pp. 150–174, 1985.
- [16] J. M. Mallatt, J. R. Garey, and J. W. Shultz, "Ecdysozoan phylogeny and Bayesian inference: first use of nearly complete 28S and 18S rRNA gene sequences to classify the arthropods and their kin," *Molecular Phylogenetics and Evolution*, vol. 31, no. 1, pp. 178–191, 2004.
- [17] F. Ronquist and J. P. Huelsenbeck, "MrBayes 3: Bayesian phylogenetic inference under mixed models," *Bioinformatics*, vol. 19, no. 12, pp. 1572–1574, 2003.
- [18] J. A. A. Nylander, J. C. Wilgenbusch, D. L. Warren, and D. L. Swofford, "AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics," *Bioinformatics*, vol. 24, no. 4, pp. 581–583, 2008.
- [19] D. L. Swofford, *PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods), Version 4*, Sinauer Associates, Sunderland, Mass, USA, 2003.
- [20] P. G. Foster, "Modeling compositional heterogeneity," *Systematic Biology*, vol. 53, no. 3, pp. 485–495, 2004.
- [21] S. Guindon, J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel, "New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0," *Systematic Biology*, vol. 59, no. 3, pp. 307–321, 2010.
- [22] D. F. Robinson and L. R. Foulds, "Comparison of phylogenetic trees," *Mathematical Biosciences*, vol. 53, no. 1-2, pp. 131–147, 1981.
- [23] J. Adachi and M. Hasegawa, "Improved dating of the human/chimpanzee separation in the mitochondrial DNA tree: heterogeneity among amino acid sites," *Journal of Molecular Evolution*, vol. 40, no. 6, pp. 622–628, 1995.
- [24] C. J. Winchell, J. Sullivan, C. B. Cameron, B. J. Swalla, and J. Mallatt, "Evaluating hypotheses of deuterostome phylogeny and chordate evolution with new LSU and SSU ribosomal DNA data," *Molecular Biology and Evolution*, vol. 19, no. 5, pp. 762–776, 2002.
- [25] B. Boussau and M. Gouy, "Efficient likelihood computations with nonreversible models of evolution," *Systematic Biology*, vol. 55, no. 5, pp. 756–768, 2006.
- [26] M. Tholleson, "LDDist: a Perl module for calculating LogDet pair-wise distances for protein and nucleotide sequences," *Bioinformatics*, vol. 20, no. 3, pp. 416–418, 2004.
- [27] S. Blanquart and N. Lartillot, "A Bayesian compound stochastic process for modeling nonstationary and nonhomogeneous sequence evolution," *Molecular Biology and Evolution*, vol. 23, no. 11, pp. 2058–2071, 2006.

- [28] V. Gowri-Shankar and M. Rattray, "A reversible jump method for Bayesian phylogenetic inference with a nonhomogeneous substitution model," *Molecular Biology and Evolution*, vol. 24, no. 6, pp. 1286–1299, 2007.
- [29] N. C. Sheffield, H. Song, S. L. Cameron, and M. F. Whiting, "Nonstationary evolution and compositional heterogeneity in beetle mitochondrial phylogenomics," *Systematic Biology*, vol. 58, no. 4, pp. 381–394, 2009.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

