

Reconstructing Phylogenetic Networks Using Maximum Parsimony

Luay Nakhleh Guohua Jin Fengmei Zhao John Mellor-Crummey
Department of Computer Science, Rice University
Houston, TX 77005
{nakhleh, jin, fzhao, johnmc}@cs.rice.edu

Abstract

Phylogenies—the evolutionary histories of groups of organisms—are one of the most widely used tools throughout the life sciences, as well as objects of research within systematics, evolutionary biology, epidemiology, etc. Almost every tool devised to date to reconstruct phylogenies produces trees; yet it is widely understood and accepted that trees oversimplify the evolutionary histories of many groups of organisms, most prominently bacteria (because of horizontal gene transfer) and plants (because of hybrid speciation).

Various methods and criteria have been introduced for phylogenetic tree reconstruction. Parsimony is one of the most widely used and studied criteria, and various accurate and efficient heuristics for reconstructing trees based on parsimony have been devised. Jotun Hein suggested a straightforward extension of the parsimony criterion to phylogenetic networks. In this paper we formalize this concept, and provide the first experimental study of the quality of parsimony as a criterion for constructing and evaluating phylogenetic networks. Our results show that, when extended to phylogenetic networks, the parsimony criterion produces promising results. In a great majority of the cases in our experiments, the parsimony criterion accurately predicts the numbers and placements of non-tree events.

Keywords: *phylogenetic networks, maximum parsimony, horizontal gene transfer.*

1 Introduction

Phylogenies—the evolutionary histories of groups of organisms—play a major role in representing the interrelationships among biological entities. Many methods for reconstructing and studying such phylogenies have been proposed, almost all of which assume that the underlying history of a given set of species can be represented by a binary tree. Although many biological processes can be effectively modeled and summarized in this fashion, others cannot: re-

combination, hybrid speciation, and horizontal gene transfer result in networks of relationships rather than trees of relationships. While this problem is widely appreciated, there has been comparatively little work on computational methods for estimating and studying evolutionary networks (see [15, 17] for surveys of work on phylogenetic networks).

Due to the pervasiveness of phylogenies, many different methods have been proposed for their reconstruction. Parsimony has been studied and used extensively for phylogenetic trees [22]; it is based on a minimum-information principle (similar to Occam’s razor): in absence of information to the contrary, the best explanation for the observed data is that involving the smallest number of manipulations—in the case of evolutionary histories, that involving the fewest evolutionary events. As pointed out by Hein [10, 11], parsimony can be extended to phylogenetic networks: he observed that each individual site in a set of sequences labeling a network evolves down a tree contained in the network (i.e., a tree whose edges are edges of the network). In consequence, the obvious extension is to define the parsimony score of a network as the sum, over all sites, of the parsimony score of the best possible tree contained within the network for each site.

This generalization suffers from a major flaw: adding reticulation events (in the form of additional edges) to the network can only lower the parsimony score, since introducing a new edge cannot increase the score of any site, but may help lower the score of some. In consequence, this criterion may lead to a gross overestimation of the number of reticulation events needed to explain the data. In this paper, we generalize Hein’s parsimony concept to *blocks* of sites, rather than individual sites, and study the performance of the generalized criterion. This generalization is a more realistic reflection of the biology of reticulate evolution: blocks of sites, or genes, are inherited as single units from their “parents”.

Reconstructing maximum parsimony phylogenetic networks is NP-hard, since it is a generalization of the maximum parsimony problem on phylogenetic trees, which is NP-hard [4, 8]. Further, whereas computing the parsimony

score of a given phylogenetic tree can be done in time linear in the number of tree nodes, using Fitch’s algorithm [6, 9], computing the parsimony score of phylogenetic networks is a hard problem (possibly NP-hard). However, the problem is fixed-parameter tractable (FPT) [5], as we show.

Since we are interested in assessing the quality of the parsimony criterion for phylogenetic networks (rather than heuristics), and due to the absence of any efficient algorithms for solving the problem, we implemented an exhaustive search method that traversed the entire space of networks, and considered the parsimony score of every network in the space. We considered a version of the phylogenetic network reconstruction problem that applies to horizontal gene transfer: given an organismal (species) tree, compute an additional set of edges whose addition to the tree explains the horizontal gene transfer events that occurred during the evolutionary history of the sequences. Since these events are unknown, we apply the parsimony criterion and seek the solution that is optimal with respect to this criterion. Our experimental results show that the parsimony criterion, when used carefully, can be very promising in both reconstructing phylogenetic networks, and quantifying their quality (in terms of capturing the true evolutionary events).

2 Phylogenetic Networks

When events such as horizontal gene transfer occur, the evolutionary history of the set of organisms may not be modeled by phylogenetic trees; in this case, *phylogenetic networks* provide the correct model. In horizontal gene transfer (HGT), genetic material is transferred from one lineage to another; see the phylogenetic network in Figure 1(a). In an evolutionary scenario involving horizontal transfer, certain sites (specified by a specific substring within the DNA sequence of the species into which the horizontally transferred DNA was inserted) are inherited through horizontal transfer from another species (as in Figure 1(c)), while all others are inherited from the parent (as in Figure 1(b)). Thus, *each site evolves down one of the trees induced by the network*.

We adopt the general model of phylogenetic networks formalized by Moret *et al.* [18]

Definition 1 A *phylogenetic network* $N = (V, E)$ with a set $L \subseteq V$ of n leaves, is a directed acyclic graph in which exactly one node, the root, has no incoming edges, and all other nodes have either one incoming edge—tree nodes—or two incoming edges—reticulation nodes.

In this paper, we focus on *binary networks*, i.e., networks in which the outdegree of a tree node is 2 and the outdegree of a reticulation node is 1. A tree T is *contained* (or *induced*) inside a network N , if T can be obtained from N by

removing exactly one of the two edges incoming into each reticulation node in N and using any applicable forced contractions. We denote by $\mathcal{T}(N)$ the set of all trees induced by a network N . While a phylogenetic network models the evolutionary history of a set of organisms, the evolutionary histories of *individual* genes are still trees [16], which are the trees contained inside the network.

Reticulation events impose time constraints on the phylogenetic network. [18] A phylogenetic network $N = (V, E)$ defines a partial order on the set V of nodes. If we associate time $t(u)$ with node u of N , then, if there exists a directed path p from u to some other node v such that p contains at least one tree edge, we must have $t(u) < t(v)$ in order to respect the time flow; moreover, if $e = (u, v)$ is a network edge, then we must have $t(u) = t(v)$, because hybridization is, at the scale of evolution, an instantaneous process. Given a network N , we say that p is a *positive-time directed path* from u to v , if p is a directed path from u to v and p contains at least one tree edge. Given a network N , two nodes u and v *cannot co-exist in time* if there exists a sequence $P = \langle p_1, p_2, \dots, p_k \rangle$ of paths such that: (1) p_i is a positive-time directed path, for every $1 \leq i \leq k$, (2) u is the tail of p_1 , and v is the head of p_k , and (3) for every $1 \leq i \leq k - 1$, there exists a network node whose two parents are the head of p_i and the tail of p_{i+1} . Since events such as horizontal gene transfer occur between two lineages (nodes in the network) that co-exist in time [16, 19, 18], a phylogenetic network N must satisfy the following property:

- If two nodes x and y cannot co-exist in time, then they cannot participate in a reticulation event, that is, the network cannot include either of the two edges (x, y) and (y, x) .

It is important to note that the aforementioned model and temporal constraints may be violated in practice, due to incomplete sampling, extinction, etc. Moret *et al.* addressed these issues and extended the phylogenetic network models to capture these scenarios, whenever they may occur [18].

3 The Parsimony Criterion

Parsimony is one of the most popular methods used for phylogenetic tree reconstruction. Roughly this method is based on the assumption that “evolution is parsimonious”, i.e., the best evolutionary trees are the ones that minimize the number of changes along the edges of the tree. We now formulate this concept.

Definition 2 The *Hamming distance* between two equal-length sequences x and y , denoted by $H(x, y)$, is the number of positions j such that $x_j \neq y_j$.

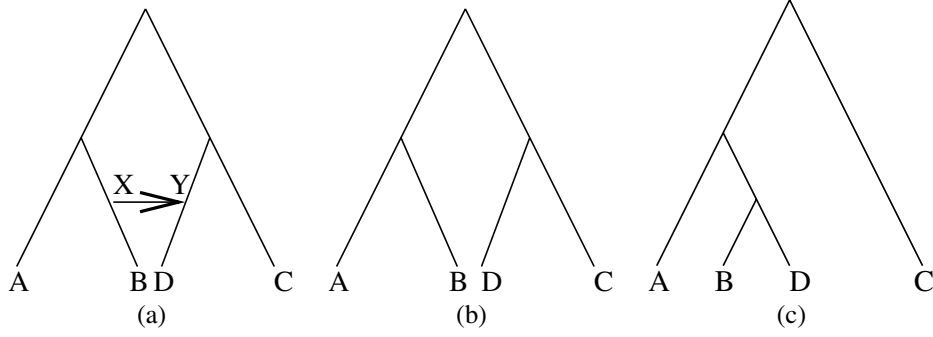


Figure 1. (a) A phylogenetic network that consists of an organismal tree and an additional edge that corresponds to an HGT event. The two (gene) trees contained inside the network are shown in (b) and (c).

Given a fully-labeled tree T , i.e., a tree in which each node v is labeled by a sequence s_v over some alphabet Σ , we define the Hamming distance of an edge $e \in E(T)$, denoted by $H(e)$, to be $H(s_u, s_v)$, where u and v are the two endpoints of e . We now define the parsimony score of a tree T .

Definition 3 *The parsimony score of a fully-labeled tree T , is $\sum_{e \in E(T)} H(e)$. Given a set S of sequences, a maximum parsimony tree for S is a tree leaf-labeled by S and assigned labels for the internal nodes, of minimum parsimony score.*

Given a set S of sequences, the parsimony problem is to find a maximum parsimony phylogenetic tree T for the set S . Unfortunately, this problem is NP-hard, even when the sequences are binary [4, 8]. One approach that is used in practice is to look at as many leaf-labeled trees as possible, and choose one with a minimum parsimony score. The problem of computing the parsimony score of a fixed leaf-labeled tree is solvable in polynomial time [6, 9].

3.1 Parsimony on Phylogenetic Networks

The evolutionary history of a site i in a set S of sequences that evolved on a network N is captured by one of the trees contained inside the network N . Therefore, a natural way to extend the tree-based parsimony score to fit a dataset that evolved on a network is to define the parsimony score for each site as the minimum parsimony score of that site over all tree contained inside the network. This extension was first introduced by Hein [10, 11] in the context of meiotic recombination. We now formalize a slightly more general definition of parsimony.

Definition 4 *The parsimony score of a network N leaf-labeled by a set S of taxa, is*

$$NCost(N, S) := \sum_{b_i \in B} (\min_{T \in \mathcal{T}(N)} TCost(T, b_i))$$

where B is a set of blocks of equal length that partition that sequences, $TCost(T, b_i)$ is the number of changes of block b_i on tree T , and $\mathcal{T}(N)$ denotes the set of trees contained inside network N .

Based on this criterion, we would want to reconstruct a phylogenetic network whose parsimony score is minimized. In the case of horizontal gene transfer, Lerat *et al.* have observed that the underlying organismal tree is reconstructible [14]. Hence, the problem of reconstructing phylogenetic networks in this case becomes one of computing a set of edges whose addition to the organismal tree “explains” the horizontal gene transfer events. This is formalized as follows.

Definition 5 **FIXED-TREE MP ON PHYLOGENETIC NETWORKS (FTMPPN):**

Input: *a set S of aligned biomolecular sequences, a tree T leaf-labeled by S , and a bound B .*

Output: *a phylogenetic network N consisting of T and at most B additional non-tree edges so as to minimize $NCost(N, S)$.*

Notice that based on Definition 4, whenever we add an extra non-tree edge to a given network (or tree, which is a special case of networks), the parsimony score of the network either remains the same or drops, which necessitates bounding the number of extra non-tree edges. Although this bound is not known *a priori*, we show experimentally that the rate of decrease in the parsimony score (as a function of the number of added edges) can be used in most cases to infer the bound and provide a stopping rule. Solving the FTMPPN problem requires computing the parsimony score of a fixed network. This is formalized as follows.

Definition 6 **PARSIMONY SCORE OF PHYLOGENETIC NETWORKS (PSPN):**

Input: a set S of aligned biomolecular sequences, and a phylogenetic network N leaf-labeled by S .

Output: $NCost(N, S)$.

The PSPN problem is of unknown complexity, yet, it is straightforward to show that the problem is solvable in polynomial time for a fixed B .

Theorem 1 *The PSPN problem is solvable in polynomial time when B is constant.*

Proof: Let N be a phylogenetic network leaf-labeled by a set S of n sequences over an alphabet Σ , each of length k , and let B be the number of hybrids in N . The number of trees induced by N is $O(2^B)$. To compute the parsimony score of a site i on a tree T with n leaves it takes $O(|\Sigma|n)$ (Fitch’s algorithm [6]). Therefore, to compute the parsimony score of N on the set S of sequences it takes $O(2^B|\Sigma|nk)$, which is polynomial when B is fixed. \square

3.2 Related Work on the Parsimony of Networks

In a series of papers [2, 1, 3, 12], the authors proposed “median networks” as a model for representing the evolutionary histories in the presence of non-treelike processes, and used an approach that combined Kruskal’s algorithm for finding minimum spanning trees with Farris’ maximum parsimony heuristic algorithm. This work does not extend the parsimony criterion beyond trees, nor does it give a model for phylogenetic inference; rather, it gives a model for graphical representation of datasets when trees fail to be the appropriate model.

In [7], the author defines a “most parsimonious network” as an alternative to consensus trees. In studying population data, it is common to have many equally possible parsimonious trees. The author proposes most parsimonious networks as a model that reconciles all the trees in one network such that the network contains all most parsimonious trees and no other trees. The author also proposes a simple set of loop-breaking rules that should find all the “resident” most parsimonious trees and no other trees should be obtainable from the rules. There are two main differences between our approach and the approach in [7]: first, our phylogenetic networks do represent evolutionary histories where the internal nodes of the network represent ancestral taxa, whereas in [7] the networks are graphical representation of the data that do not reflect phylogenetic inference. Second, we extend the parsimony criterion from trees to networks, whereas in [7] the parsimony criterion is not defined for networks.

4 Experiments

To generate networks with different number of HGT edges, we used an incremental network generation algorithm (outlined in Figure 2). To handle the huge number of networks (over 10 million networks were inspected for the case of three HGT events), we developed two parallel implementations of our network generation algorithm: a static scheduling algorithm (outlined in Figure 3) and a dynamic scheduling algorithm (outlined in Figure 4).

4.1 Settings

We used the `r8s` tool [21] to generate a random birth-death phylogenetic tree on 10 taxa (shown in Figure 5(a)). The `r8s` tool generate molecular clock trees; we deviated the tree from this hypothesis by multiplying each edge in the tree by a number randomly drawn from an exponential distribution. The expected evolutionary diameter (longest path between any two leaves in the tree) is 0.2. We then considered two HGT events: *close* HGT event, as shown in Figure 5(b), involves two closely related species, and *divergent* HGT event, as shown in Figure 5(c), involves two divergent species. Figure 5(d) shows the phylogenetic network that consists of the organismal tree and the two HGT events. Since divergent events span a larger distances than close ones, we expect their detection to be easier.

For each of the three phylogenetic networks shown in Figure 5, we used the `Seq-gen` tool [20] to evolve 30 datasets of DNA sequences of length 1500 down the organismal tree and DNA sequences of length 500 down the other tree contained inside the network. Both sequence datasets were evolved under the $K2P+\gamma$ model of evolution, with shape parameter 1 [13]. Finally, we concatenated the two datasets, solved the FTMPPN exhaustively on the organismal tree and concatenated sequence dataset, and compared the resulting network against the model network. In our experiments, we used blocks of 500 nucleotides. To compute the parsimony score of a leaf-labeled phylogenetic tree, we used the `PAUP*` tool [23].

4.2 Results

We investigate three main questions. (1) Can we infer, based on parsimony, the number of HGT events that occurred during the evolution of a set of sequences? (2) How does the parsimony score of the best phylogenetic network computed compare to the parsimony score of the model network? (3) How does the parsimony criterion perform in reconstructing the actual HGT events?

Figures 6 and 7 answer the first question. In Figure 6(a), we observe that, for each of the 30 datasets, the optimal parsimony scores are almost identical, regardless of the num-

```

Tree2NetTrees( $N, numH$ )
   $bestScore = maxVal$ ;
  Net2Nets( $N, 0, numH, bestScore$ );

Net2Nets( $N, curH, numH, bestScore$ )
  for each pair of edges ( $e1, e2$ )
    if (IsLegalHGT( $N, e1, e2$ )  $\wedge$  HasEffect( $N, e1, e2$ ))
      CreateNetwork( $N, N'$ );
      AddHGT( $N', e1, e2$ );
      if ( $curH == numH$ )
        Net2Trees( $N', numH, score$ );
        if ( $newScore < bestScore$ )  $bestScore = newScore$ ;
      else
        Net2Nets( $N', curH + 1, numH, bestScore$ );

Net2Trees( $N, numH, score$ )
  for ( $i = 0; i < 2^{numH}; i++$ )
    TreeGen( $N, T, numH$ );
     $TS = TS \cup T$ ;
     $score = \text{CompNetMPScore}(TS, S)$ ;

```

Figure 2. **Tree2NetTrees** generates all possible networks from a given tree. For networks with more than one HGT edge, the **Net2Nets** produces networks by adding edges to networks. The **Net2Trees** procedure computes the set of all trees contained inside a network, computes the parsimony of a set of sequences on each tree (by calling PAUP*), and then computes the parsimony of the network. Procedure **IsLegalHGT** verifies that the resulting network does not violate the time co-existence property, defined in Section 2. Procedure **HasEffect** excludes edges whose addition would not result in new trees.

ber of edges added, which implies that no HGT events are inferred. Indeed, these 30 datasets evolved in their entirety down the organismal trees, and hence no HGT events were present.

Figures 6(b) and 6(c) show the results on datasets whose evolution involves a single HGT event, between two closely related species in the former, and distantly related in the latter. Figure 6(c) show a much sharper decrease in the optimal parsimony score when adding the first edge, compared to the decrease when adding a second edge. In this case, stopping after the first edge is easily determined. The contrast between the effects of adding the first and second edges is not as clear when the HGT event is between two closely related species, which is expected (shown in Figure 6(b)). Figure 6(d) shows a similar trend, yet the decrease in the parsimony score after adding the first edge is not very large, which is a reflection of the hardness of detecting HGT events between closer organisms. In this case, the parsimony criterion may underestimate the number of HGT events. However, we predict that if both HGT events were between divergent organisms, we would see a sharper decrease in the parsimony score when adding the second

edge.

Figure 7 summarize the previous results in a clearer picture. When the sequences evolve with no HGT, there is an extremely slow decrease in the parsimony score. In the case of a single HGT event between two divergent organisms, there is a sharp decrease after adding one edge, and then the decrease starts leveling. Similar results are observed for the case of the single HGT event between closely related organisms. For the case of two HGT events, the decrease starts leveling at the second edge (from partial results that we obtained on adding a third edge, the decrease in the parsimony score after adding the second edge is negligible). Therefore, depending on the location of the HGT events, a clear threshold in most cases can be drawn so as to estimate the correct number of HGT events.

Figures 8 and 9 answer the second question, namely how the optimal parsimony scores computed compare to the parsimony score of the model network. The results show that these two scores are identical in almost all cases. These results clearly exhibit the quality of parsimony as a criterion for reconstructing phylogenetic networks.

To answer the third question, and since the underlying

```

if (myrank = 0) read and process command line;
if (myrank = 0) send input parameters to all by broadcasting;
if (myrank ≠ 0) receive input parameters from Processor 0;
read in tree and create networks;
for (i = 1; i ≤ totNets; i + +)
    if (myrank = (j - 1) % nprocs) create network Ni;

```

Figure 3. A static scheduling algorithm.

```

SCHEDULER:
    for each way of adding the first HGT edge h
        wait request from the next available worker;
        assign the worker to generate networks by first adding h;
    notify all workers that all work is assigned;
WORKERS:
    send request for work to the SCHEDULER;
    receive assignment (adding h) from the SCHEDULER
    generate networks by adding h as the first HGT edge;

```

Figure 4. A dynamic scheduling algorithm.

organismal tree in fixed in our experiments, we compared the HGT events inferred by the parsimony criterion against the HGT events in the model networks as follows. Let X_1 and Y_1 be the sets of taxa under the source and target nodes, respectively, of the HGT events in the model network, and let X_2 and Y_2 be the sets of taxa under the source and target nodes, respectively, of the HGT events in the inferred network. We define

$$\Delta_{HGT} = \frac{|X_1 \Delta X_2|}{|X_1 \cup X_2|} + \frac{|Y_1 \Delta Y_2|}{|Y_1 \cup Y_2|},$$

where $A \Delta B$ denotes the symmetric difference between the two sets A and B . If the two HGT events are identical, we have $\Delta_{HGT} = 0$; when they are totally different, we have $\Delta_{HGT} = 1$. Figure 10 shows that the parsimony criterion did not compute a network that is identical to the model network in only four out of 60 cases, which is very impressive. The figures also show that the parsimony criterion in some cases return more than one optimal network, but usually one of them is identical to the model network (this is similar to the behavior of parsimony on trees as well).

5 Conclusions

In this paper we provided the first empirical proof of the appropriateness of parsimony as a criterion for both reconstructing and evaluating the quality of phylogenetic networks. Now that the quality of the criterion has been established, efficient algorithms and heuristics for solving the problems outlined in the paper are needed.

References

- [1] H.-J. Bandelt, P. Forster, and A. Röhl. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.*, 16(1):37–48, 1999.
- [2] H.-J. Bandelt, P. Forster, B. Sykes, and M. Richards. Mitochondrial portraits of human populations using median networks. *Genetics*, 141:743–753, 1995.
- [3] H.-J. Bandelt, V. Macaulay, and M. Richards. Median networks: speedy construction and greedy reduction, one simulation, and two case studies from human mtDNA. *Molecular Phylogenetics and Evolution*, 16(1):8–28, 2000.
- [4] W. Day. Computationally difficult parsimony problems in phylogenetic systematics. *Journal of Theoretical Biology*, 103:429–438, 1983.
- [5] R. Downey and M. Fellows. Fixed parameter tractability and completeness I: basic theory. *SIAM Journal of Computing*, 24:873–921, 1995.
- [6] W. Fitch. Toward defining the course of evolution: minimum change for a specified tree topology. *Syst. Zool.*, 20:406–416, 1971.
- [7] W. Fitch. Networks and viral evolution. *J. Mol. Evol.*, 44(Suppl 1):S65–S75, 1997.
- [8] L. Foulds and R. Graham. The steiner problem in phylogeny is NP-Complete. *Adv. Appl. Math.*, 3:43–49, 1982.
- [9] J. Hartigan. Minimum mutation fits to a given tree. *Biometrics*, 29:53–65, 1973.
- [10] J. Hein. Reconstructing evolution of sequences subject to recombination using parsimony. *Mathematical Biosciences*, 98:185–200, 1990.
- [11] J. Hein. A heuristic method to reconstruct the history of sequences subject to recombination. *J. Mol. Evol.*, 36:396–405, 1993.

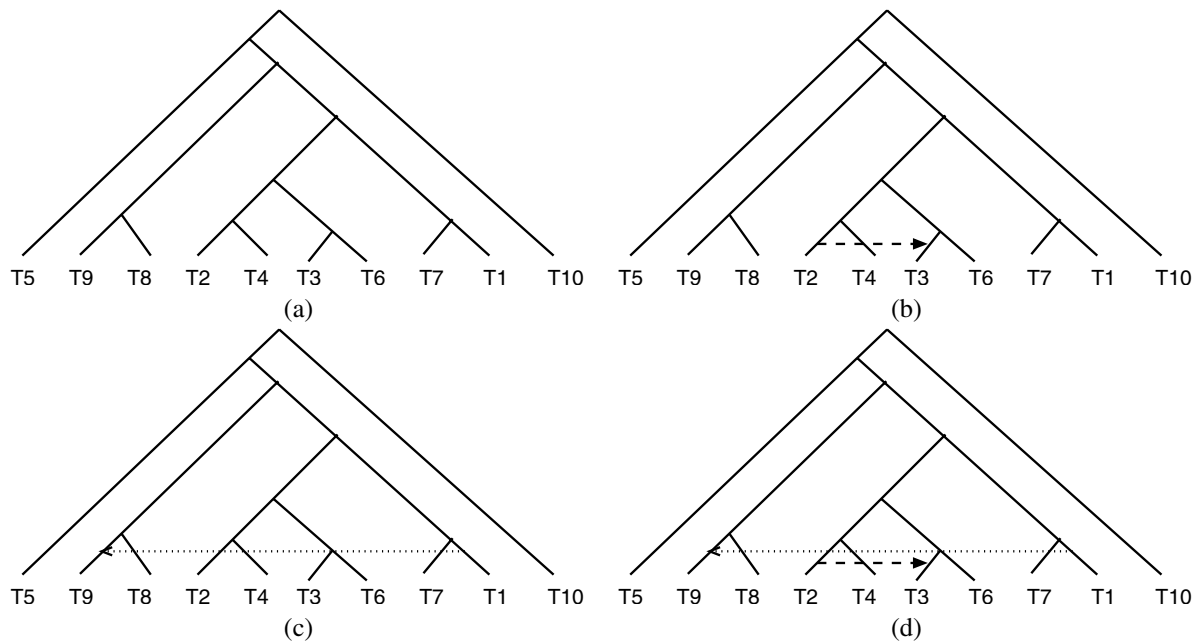


Figure 5. (a) The “organismal” tree we used in the experiments. (b) The organismal tree with a single *close* HGT event. (c) The organismal tree with a single *divergent* HGT event. (d) The organismal tree with two HGT events.

- [12] K. Huber, V. Moulton, P. Lockhart, and A. Dress. Pruned median networks: a technique for reducing the complexity of median networks. *Molecular Phylogenetics and Evolution*, 19(2):302–310, 2001.
- [13] M. Kimura. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.*, 16:111–120, 1980.
- [14] E. Lerat, V. Daubin, and N. Moran. From gene trees to organismal phylogeny in prokaryotes: The case of the γ -proteobacteria. *PLoS Biology*, 1(1):1–9, 2003.
- [15] C. Linder, B. Moret, L. Nakhleh, and T. Warnow. Network (reticulate) evolution: biology, models, and algorithms. In *The Ninth Pacific Symposium on Biocomputing (PSB)*, 2004. A tutorial.
- [16] W. Maddison. Gene trees in species trees. *Syst. Biol.*, 46(3):523–536, 1997.
- [17] V. Makarenkov, D. Kevorkov, and P. Legendre. Phylogenetic network reconstruction approaches. *Applied Mycology and Biotechnology (Genes, Genomics and Bioinformatics)*, 6, 2005. To appear.
- [18] B. Moret, L. Nakhleh, T. Warnow, C. Linder, A. Tholse, A. Padolina, J. Sun, and R. Timme. Phylogenetic networks: modeling, reconstructibility, and accuracy. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 1(1):13–23, 2004.
- [19] R. Page and M. Charleston. Trees within trees: Phylogeny and historical associations. *Trends in Ecol. and Evol.*, 13:356–359, 1998.
- [20] A. Rambaut and N. C. Grassly. Seq-gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. *Comp. Appl. Biosci.*, 13:235–238, 1997.
- [21] M. Sanderson. r8s software package. Available from <http://loco.ucdavis.edu/r8s/r8s.html>.
- [22] M. Sanderson, B. Baldwin, G. Bharathan, C. Campbell, D. Ferguson, J. M. Porter, C. V. Dohlen, M. Wojciechowski, and M. Donoghue. The growth of phylogenetic information and the need for a phylogenetic database. *Syst. Biol.*, 42:562–568, 1993.
- [23] D. Swofford. PAUP*: Phylogenetic analysis using parsimony (and other methods), 1996. Sinauer Associates, Sunderland, Massachusetts, Version 4.0.

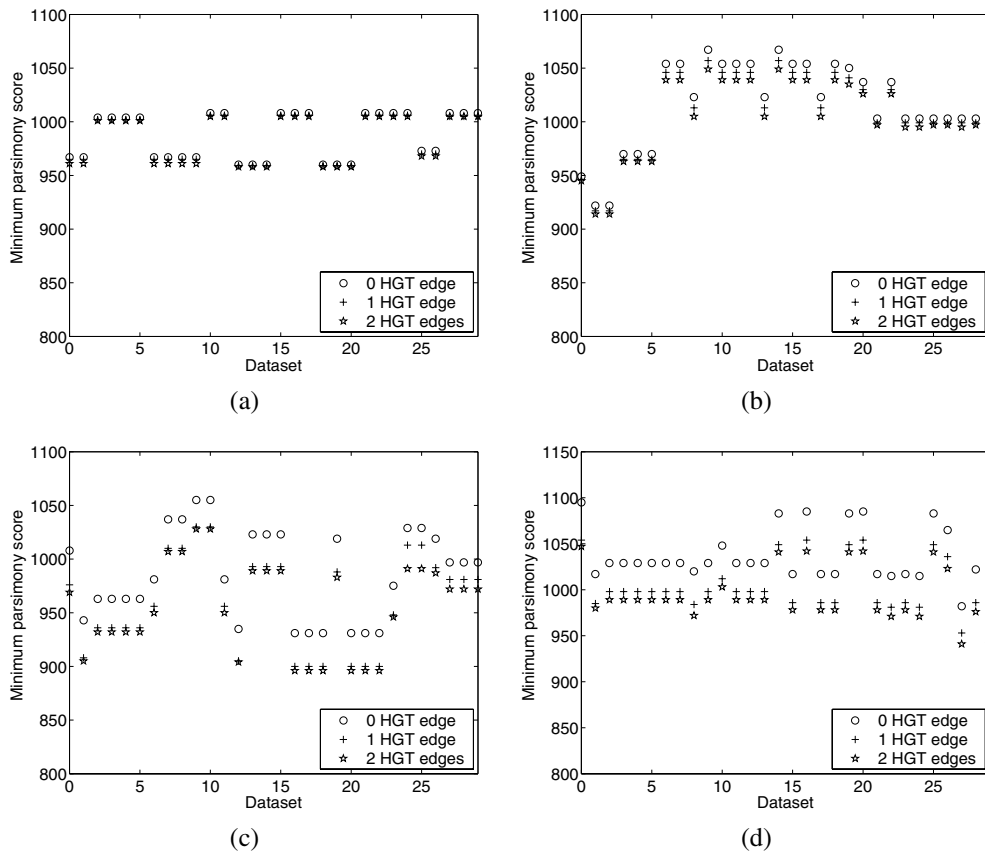


Figure 6. The minimum parsimony score found by adding no additional edges, 1 edge, and 2 edges. The results in (a)–(d) were obtained on datasets that evolved down the networks in Figures 5(a)–(d), respectively.

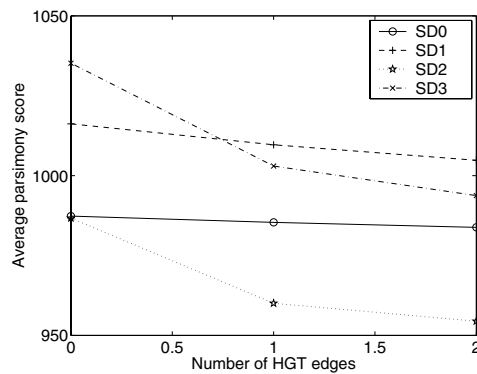
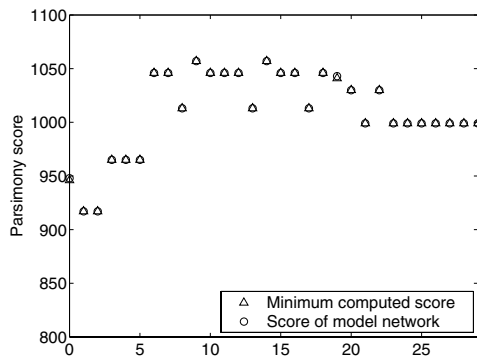
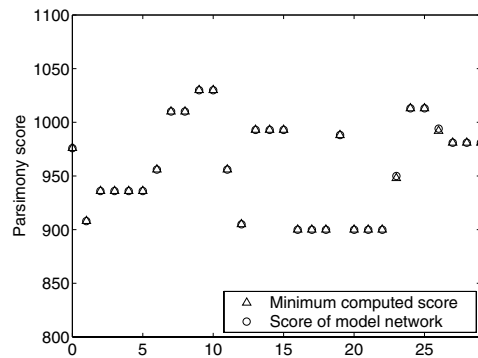


Figure 7. The rate of parsimony score decrease as a function of the number of edges added when solving the FTMPN problem on the sequence datasets evolved down each of the four networks in Figure 5. Each point in the graph is the average of 30 results. *SD0*–*SD3* correspond to datasets that evolved with zero, one (close), one (divergent), and two HGT events, respectively.



(a)



(b)

Figure 8. The minimum parsimony score computed, compared to the parsimony score of the model network. The results in (a) and (b) were obtained on datasets that evolved down the networks in Figures 5(b) and 5(c), respectively.

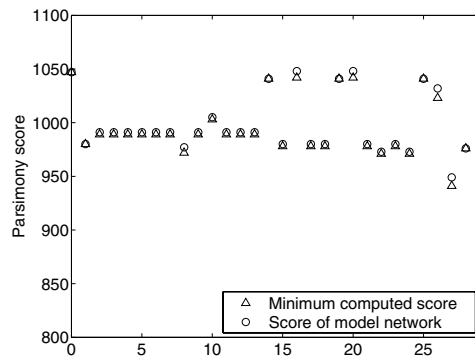


Figure 9. The minimum parsimony score computed, compared to the parsimony score of the model network. The results were obtained on datasets that evolved down the network in Figure 5(d).

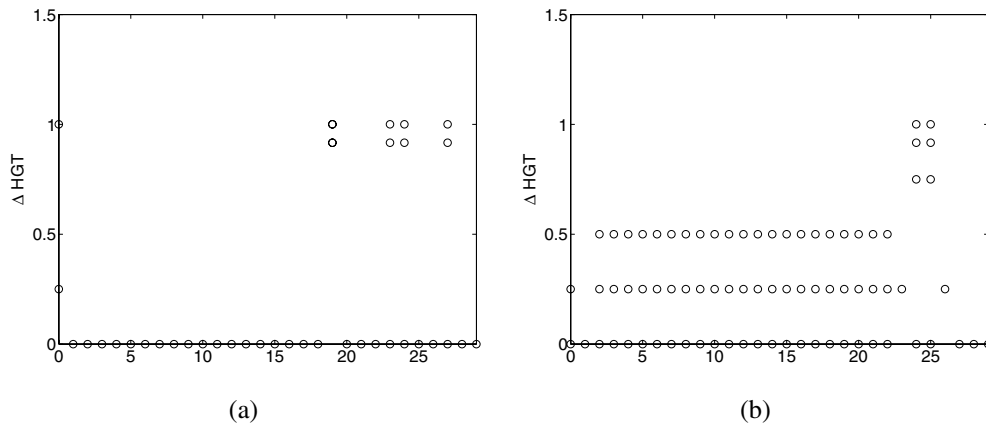


Figure 10. The accuracy of the HGT events, as measured by ΔHGT , inferred by the parsimony criterion. The results in (a) and (b) were obtained on datasets that evolved down the networks in Figures 5(b) and 5(c), respectively.