

THE EFFECT OF THE GUIDE TREE ON MULTIPLE SEQUENCE ALIGNMENTS AND SUBSEQUENT PHYLOGENETIC ANALYSES

S. NELESEN, K. LIU, D. ZHAO, C. R. LINDER, AND T. WARNOW *

The University of Texas at Austin

Austin, TX 78712

E-mail: {serita, kliu, wzhao, tandy}@cs.utexas.edu, rlinder@mail.utexas.edu

Many multiple sequence alignment methods (MSAs) use guide trees in conjunction with a progressive alignment technique to generate a multiple sequence alignment but use differing techniques to produce the guide tree and to perform the progressive alignment. In this paper we explore the consequences of changing the guide tree used for the alignment routine. We evaluate four leading MSA methods (ProbCons, MAFFT, Muscle, and ClustalW) as well as a new MSA method (FTA, for “Fixed Tree Alignment”) which we have developed, on a wide range of simulated datasets. Although improvements in alignment accuracy *can* be obtained by providing better guide trees, in general there is little effect on the “accuracy” (measured using the SP-score) of the alignment by improving the guide tree. However, RAxML-based phylogenetic analyses of alignments based upon better guide trees tend to be *much* more accurate. This impact is particularly significant for ProbCons, one of the best MSA methods currently available, and our method, FTA. Finally, for very good guide trees, phylogenies based upon FTA alignments are more accurate than phylogenies based upon ProbCons alignments, suggesting that further improvements in phylogenetic accuracy may be obtained through algorithms of this type.

1. Introduction

Although methods are available for taking molecular sequence data and simultaneously inferring an alignment and a phylogenetic tree, the most common phylogenetic practice is a sequential, two-phase approach: first an alignment is obtained from a multiple sequence alignment (MSA) program and then a phylogeny is inferred based upon that alignment. The two-phase approach is usually preferred over simultaneous alignment and

*This work was supported by NSF under grants ITR-0331453, ITR-0121680, ITR-0114387 and EIA-0303609.

tree estimation because, to date, simultaneous methods have either been restricted to a very limited number of taxa (less than about 30) or have been shown to produce less accurate trees than the best combinations of alignment and tree inference programs, e.g., alignment with ClustalW¹⁷ or one of the newer alignment methods such as MAFFT⁵, ProbCons¹ or Muscle², followed by maximum likelihood methods of tree inference such as RAxML¹³. Many of the best alignment programs use dynamic programming to perform a progressive alignment, with the order of the progressive alignment determined by a guide tree. All methods produce a default guide tree and some will also accept one input by the user.

Whereas much effort has been made to assess the accuracy of phylogenetic tree reconstruction using different methods, models and parameter values, and much attention has been paid to the progressive alignment techniques, far less effort has gone into determining how different guide trees influence the quality of the alignment *per se* and the subsequent phylogeny. A limited study by Roshan et al.¹⁰ looked at improving maximum parsimony trees by iteratively improving the guide trees used in the alignment step. However, they showed little improvement over other techniques.

We address this question more broadly in a simulation study, exploring the performance of five MSA methods (ClustalW, Muscle, ProbCons, MAFFT and FTA - a new method, which we present here) on different guide trees. We find that changes in the guide tree generally do not impact the accuracy of the estimated alignments, as measured by SP-score (Section 3.1 defines this score). *However*, some RAxML-based phylogenies, obtained using alignments estimated on more accurate guide trees, were *much* more accurate than phylogenies obtained using MSA methods on their default guide trees. Muscle and ClustalW were impacted the least by the choice of guide tree, and ProbCons and FTA were impacted the most. The improvement produced for ProbCons is particularly relevant to systematists, since it is one of the two best MSA methods currently available.

Finally, we find that using FTA as an alignment technique results in even more accurate trees than available using ProbCons when a highly accurate guide tree is input, showing the potential for even further improvements in multiple sequence alignment and phylogeny estimation.

The organization of the rest of the paper is as follows. Section 2 provides background on the multiple alignment methods we study, and includes a discussion of the design of FTA. Section 3 describes the experimental study, and the implications of these results. Finally, we discuss future work in Section 4.

2. Basics

Phylogeny and alignment estimation methods. Although there are many phylogeny estimation methods, our studies (and those of others) suggest that maximum likelihood analyses of aligned sequences produce the most accurate phylogenies. Of the various software programs for maximum likelihood analysis, RAxML and GARLI²¹ are the two fastest and most accurate methods. We used RAxML for our analyses.

Of the many MSA methods, ClustalW tends to be the one most frequently used by systematists, although several new methods have been developed that have been shown to outperform ClustalW with respect to alignment accuracy. Of these, we included ProbCons, MAFFT, and Muscle. ProbCons and MAFFT are the two best performing MSA methods, and Muscle is included because it is very fast.

We also developed and tested a new MSA method, which we call FTA for “Fixed Tree Alignment.” FTA is a heuristic for the “Fixed-Tree Sankoff Problem”, which we now define.

The Sankoff problem. Over 30 years ago, David Sankoff proposed an approach for simultaneous estimation of trees and alignments based upon minimizing the total edit distance, which we generalize here to allow for an arbitrary edit distance function $f(\cdot, \cdot)$ as part of the input, thus defining the “Generalized Sankoff Problem”¹²:

- *Input:* A set S of sequences and a function $f(s, s')$ for the cost of an optimal alignment between s and s' .
- *Output:* A tree T , leaf-labeled by the set S , and with additional sequences labelling the internal nodes of T , so as to minimize tree-length, $\sum_{(v,w) \in E} f(s_v, s_w)$, where s_v and s_w are the sequences assigned to nodes v and w , respectively, and E is the edge set of T .

The problem thus depends upon the function $f(\cdot, \cdot)$. In this paper we follow the convention that all mismatches have unit cost, and the cost of a gap of length k is affine (i.e., equals $c_0 + c_1 * k$ for some choice of c_0 and c_1) (see^{6,4}). The constants c_0 and c_1 are the “gap-open” cost and the “gap-extend” cost, respectively. The Generalized Sankoff problem is NP-hard since the special case where $c_0 = \infty$ is the maximum parsimony (MP) problem, which is NP-hard. (The problem is also called the “Generalized Tree Alignment” problem in the literature.)

In the fixed-tree version of the Sankoff problem, the tree T is given as part of the input, and the object is to assign sequences to the internal

nodes of T so as to produce the minimum total edit distance. This problem is also NP-hard¹⁹. Exact solutions⁶ which run in exponential time have been developed, but these are computationally too expensive to be used in practice. Approximation algorithms for the problem have also been developed^{18,20}, but their performance guarantees are not good enough for algorithms to be reliable in practice.

The FTA (“fixed tree alignment”) technique. We developed a fast heuristic for the Fixed-Tree Sankoff problem^a. We make an initial assignment of sequences to internal nodes and then attempt to improve the assignment until a local optimum is reached. To improve the assignment, we iteratively replace the sequence at an internal node by a new sequence if the new sequence reduces the total edit distance on the tree. To do this, we estimate the “median” of the sequences labelling the neighbors of v . Formally, the “median” of three sequences A , B , and C with respect to an edit distance function $f(\cdot, \cdot)$ is a sequence X such that $f(X, A) + f(X, B) + f(X, C)$ is minimized. This can be solved exactly, but the calculation takes $O(k^3)$ time⁶, where k is the maximum sequence length. Since we estimate medians repeatedly ($\Omega(n)$ times for each n -leaf tree analyzed), we needed a faster estimator than these exact algorithms permit.

We designed a heuristic that is not guaranteed to produce optimal solutions for estimating the median of three sequences. The technique we picked is a simple, two-step procedure, where we compute a multiple alignment using some standard MSA technique, and then compute the majority consensus of the multiple alignment. If replacing the current sequence at the node with the consensus reduces the total treelength, then we use the new sequence; otherwise, we keep the original sequence for the node.

We tested several MSA methods (MAFFT, DCA¹⁴, Muscle, ProbCons, and ClustalW) for use in the median estimator, and examined the performance of FTA under a wide range of model conditions and affine gap penalties. Medians based upon DCA generally gave the best performance with respect to total edit distances as well as SP-error; MAFFT-based medians were second best but less accurate. Because of the improvement in accuracy, we elected to work with DCA-based medians even though they sometimes took twice as long as MAFFT-based medians.

Selecting an affine gap penalty. We investigated the effect of affine gap penalties on alignment accuracy, using a wide range of model conditions (number of taxa, rates of indels and site substitutions, and gap length

^aAll modified and developed software is available upon request.

distributions). Although the best affine gap penalty (assessed by the SP-error of the alignment) varied somewhat with the model conditions, we found some gap penalties that had good performance over a wide range of model conditions. Based upon these experiments (data not shown), we chose an affine gap penalty for our analyses, with gap-open cost of 2, mismatch cost of 1, and gap-extend cost of 0.5.

3. Experimental study

Overview. We performed a simulation study to evaluate the performance of the different MSA methods we studied on each of several guide trees. We briefly describe how simulation studies can be used to evaluate two-phase techniques and give an overview of our approach.

First, a stochastic model of sequence evolution is selected (e.g., GTR, HKY, K2P, etc.³), and a model tree is picked or generated. A sequence of specified length is placed at the root of the tree T and evolved down the tree according to the parameters of the evolutionary process. At the end of this process, each leaf of the tree has a sequence. In addition, the true tree and the true alignment are both known and can be used later to assess the quality of alignment and phylogeny inference. The sequences are then aligned by a MSA technique and passed to the phylogeny estimation technique, thus producing an estimated alignment and an estimated tree which are scored for accuracy. If desired, the phylogeny estimation method can also be provided the true alignment, to see how it performs when alignment estimation is perfect.

In our experiment, we evolved DNA sequence datasets using the ROSE¹⁵ software (because it produces sequences that evolve with site substitutions and also indels) under 16 different model conditions, half for 100 taxon trees and half for 25 taxon trees. For each model condition, we generated 20 different random datasets, and analyzed each using a variety of techniques. We then compared the estimated alignments and trees to the true alignments and trees, recording the SP-error and missing edge rates. The details of this experiment are described below.

3.1. *Experimental design.*

Model Trees. We generated birth-death trees of height 1.0 using the program `r8s`¹¹ with 100 and 25 taxa. We modified branch lengths to deviate the tree moderately from ultrametricity, using the technique used by Moret *et al.*⁸ with deviation factor c set to 2.0.

Sequence Evolution. We picked a random DNA sequence of length 1000 for the root. We evolved sequences according to the K2P+Indel+ Γ model of sequence evolution. For all our model trees, we set the transition/transversion ratio to 2.0, and had all sites evolve at the same rate. We varied the model conditions between experiments by varying the remaining parameters for ROSE: the mean substitution rate, the gap length distribution, and the indel rate. We set the mean substitution rate such that the edgewise average normalized Hamming distance was (approximately) between 2% and 7%. We used two single-gap-event length distributions, both geometric with finite tails. Our “short” single-gap-event length distribution had average gap length 2.00 and a standard deviation of 1.16. Our “long” single-gap-event length distribution had average gap event length 9.18 and a standard deviation of 7.19. Finally, we set insertion and deletion probabilities so as to produce different degrees of gappiness (S-gaps in the table).

The table in Figure 1 shows the parameter settings for each of the 16 model conditions, and the resultant statistics for the model conditions (MNHD=maximum normalized Hamming distance, E-ANHD=average normalized Hamming distance on the edges, S-gaps=percent of the true alignment matrix occupied by gaps, and E-gaps = average gappiness per edge); the standard error is given parenthetically.

| MC | Model Condition Parameters | | | | True Align Statistics | | | |
|----|----------------------------|--------|--------|----------|-----------------------|-----------|-----------|-----------|
| | Taxa | P(gap) | P(sub) | Gap dist | MNHD | e-ANHD | S-gaps | E-gaps |
| 1 | 100 | 0.0001 | 0.005 | long | 37.5 (.2) | 1.9 (.02) | 40.8 (.3) | .72 (.01) |
| 2 | 100 | 0.0001 | 0.01 | long | 56.9 (.3) | 3.2 (.04) | 43.7 (.6) | .69 (.01) |
| 3 | 100 | 0.0005 | 0.005 | long | 38.0 (.3) | 2.4 (.04) | 81.9 (.3) | 4.1 (.07) |
| 4 | 100 | 0.0005 | 0.01 | long | 57.4 (.4) | 4.9 (.03) | 83.0 (.1) | 4.9 (.04) |
| 5 | 100 | 0.0005 | 0.005 | short | 36.9 (.2) | 1.9 (.04) | 42.6 (.6) | .76 (.01) |
| 6 | 100 | 0.0005 | 0.01 | short | 56.7 (.2) | 4.1 (.04) | 46.4 (.2) | .86 (.01) |
| 7 | 100 | 0.0025 | 0.005 | short | 38.6 (.3) | 2.2 (.04) | 81.4 (.2) | 4.2 (.07) |
| 8 | 100 | 0.0025 | 0.01 | short | 56.3 (.2) | 4.6 (.07) | 82.4 (.2) | 4.6 (.07) |
| 9 | 25 | 0.0001 | 0.004 | long | 32.2 (.2) | 2.9 (.05) | 22.8 (.3) | 1.2 (.02) |
| 10 | 25 | 0.0001 | 0.008 | long | 51.3 (.2) | 5.9 (.09) | 25.0 (.4) | 1.4 (.02) |
| 11 | 25 | 0.0005 | 0.004 | long | 31.3 (.2) | 2.2 (.03) | 55.1 (.5) | 4.7 (.10) |
| 12 | 25 | 0.0005 | 0.008 | long | 50.2 (.3) | 5.2 (.08) | 61.3 (.5) | 5.9 (.12) |
| 13 | 25 | 0.0005 | 0.004 | short | 30.0 (.1) | 2.6 (.08) | 26.1 (.4) | 1.4 (.04) |
| 14 | 25 | 0.0005 | 0.008 | short | 50.0 (.1) | 6.4 (.07) | 28.5 (.2) | 1.7 (.02) |
| 15 | 25 | 0.0025 | 0.004 | short | 31.1 (.4) | 3.2 (.06) | 66.7 (.4) | 7.7 (.14) |
| 16 | 25 | 0.0025 | 0.008 | short | 50.2 (.5) | 5.2 (.07) | 62.4 (.4) | 6.6 (.08) |

Figure 1. Model condition parameters and true alignment statistics.

Methods for estimating multiple alignments and trees. We used five multiple sequence alignment programs to create alignments from raw sequences: ClustalW, Muscle, MAFFT, ProbCons and FTA. ClustalW, Muscle, MAFFT and ProbCons are publicly available, while FTA is a method we developed (see Section 2 for a description of this method). For this

study, ClustalW, Muscle, MAFFT and ProbCons were each run using their default guide trees as well with guide trees that we provided. We modified ProbCons to allow it to use an input guide tree, and the authors of MAFFT provided us with a version that accepts guide trees as input. FTA does not have a default guide tree, and therefore was run using only the computed guide trees and the true tree. MAFFT has multiple alignment strategies built in, and we used each of L-INS-i, FFT-NS-i and FFT-NS-2. However, when there were difference between variants of MAFFT, FFT-NS-2 usually performed best, so we only show results using this variant. We used RAxML in its default setting.

User-input guide trees. We tested performance on four user-input guide trees. We included the true tree, and three other guide trees that we computed. The first two of the computed guide trees are UPGMA trees based upon different distance matrices. For the first UPGMA guide tree (“upgma1”), we computed a distance matrix based upon optimal pairwise alignments between all pairs of sequences, using the affine gap penalty with $gap-open = 0$, $gap-extend = 1$ and $mismatch = 1$. For the second (“upgma2”), we computed the distance matrix based upon optimal pairwise alignments between all pairs of sequences for the affine gap penalty with $gap-open = 2$, $gap-extend = 0.5$ and $mismatch = 1$. In both cases, we used custom code based on the Needleman-Wunsch algorithm with the specified gap penalty to compute the distance matrices and PAUP*¹⁶ to compute the UPGMA trees. The third guide tree (“probtree”) was obtained as follows. We used the upgma1 guide tree as input to ProbCons to estimate an alignment that was then used to estimate a tree using RAxML.

Error rates for phylogeny reconstruction methods. We used the missing edge rate, which is the percentage of the edges of the true tree that are missing in the estimated tree (also known as the false negative rate). The “true tree” is obtained by contracting the zero-event edges in the model tree; it is usually binary, but not always.

Alignment error rates. To measure alignment accuracy, we used the SP (sum-of-pairs) error rate (the complement of the SP accuracy measure), which we now define. Let $S = \{s_1, s_2, \dots, s_n\}$, and let each s_i be a string over some alphabet Σ (e.g., $\Sigma = \{A, C, T, G\}$ for nucleotide sequences). An alignment on S inserts spaces within the sequences in S so as to create a matrix, in which each column of the matrix contains either a dash or an element of Σ . Let s_{ij} indicate the j^{th} letter in the sequence s_i . We identify the alignment A with the set $Pairs(A)$ containing all pairs $(s_{ij}, s_{i'j'})$ for which some column in A contains s_{ij} and $s_{i'j'}$. Let A^* be the true

alignment, and let \hat{A} be the estimated alignment. Then the SP-error rate is $\frac{|Pairs(A^*) - Pairs(\hat{A})|}{|Pairs(A^*)|}$, expressed as a percentage; thus the SP-error is the percentage of the pairs of truly homologous nucleotides that are unpaired in the estimated alignment. However, it is possible for the SP-error rate to be 0, and yet have different alignments.

3.2. Results.

We first examine the guide trees with respect to their topological accuracy. As shown in Figure 2, the accuracy of guide trees differs significantly, with the ProbCons default tree generally the least accurate, and our “probtree” guide tree the most accurate; the two UPGMA guide trees have very similar accuracy levels.

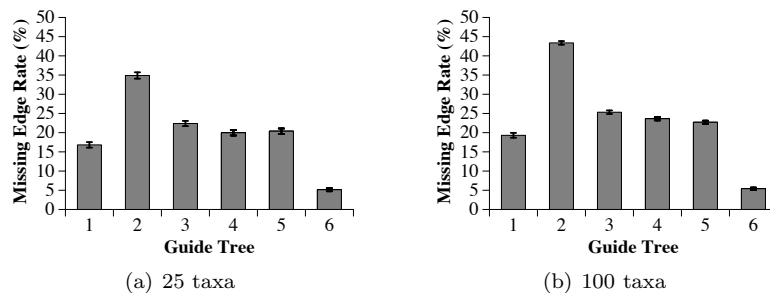


Figure 2. Guide tree topological error rates, averaged over all model conditions and replicates. (1) ClustalW default, (2) ProbCons default, (3) Muscle default, (4) upgma1, (5) upgma2, and (6) probtree.

In Figure 3 we examine the accuracy of the alignments obtained using different MSA methods on these guide trees. Surprisingly, despite the large differences in topological accuracy of the guide trees, alignment accuracy (measured using SP-error) for a particular alignment method varies relatively little between alignments estimated from different guide trees. For example, two ClustalW alignments or two Muscle alignments will have essentially the same accuracy scores, independent of the guide tree. The biggest factor impacting the SP-error of the alignment is the MSA method. Generally, ProbCons is the most accurate and ClustalW is the least.

We then examined the impact of changes in guide tree on the accuracy of the resultant RAxML-based phylogeny (see Figure 4). In all cases, for a given MSA method, phylogenetic estimations obtained when the guide

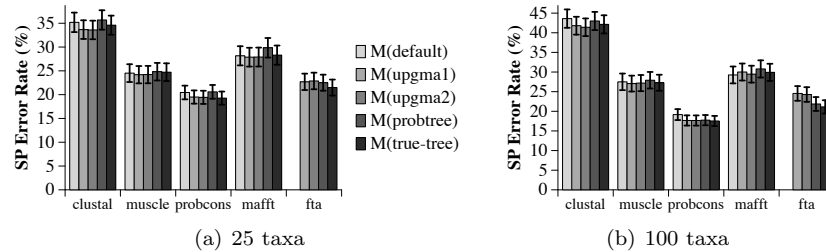


Figure 3. SP-error rates of alignments. M(guide tree) indicates multiple sequence alignment generated using the indicated guide tree.

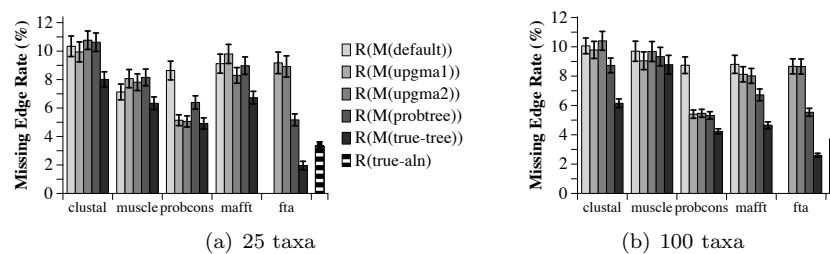


Figure 4. Missing edge rate of estimated trees. R(M(guide tree) indicates RAxML run on the alignment generated by the multiple sequence alignment method using the guide tree indicated. R(true-aln) indicates the tree generated by RAxML when given the true alignment.

tree is the true tree are more accurate than for all other guide trees. However, MSA methods otherwise respond quite differently to improvements in guide trees. For example, Muscle responded very little (if at all) to improvements in the guide tree, possibly because it computes a new guide tree after the initial alignment on the input guide tree. ClustalW also responds only weakly to improvement in guide tree accuracy, often showing - for example - worse performance on the probtree guide tree compared to the other guide trees. On the other hand, ProbCons and FTA both respond positively and significantly to improvements in guide trees. This is quite interesting, since the alignments did *not* improve in terms of their SP-error rates! Furthermore, ProbCons improves quite dramatically as compared to its performance in its default setting. The performance of FTA is intriguing. It is generally worse than ProbCons on the UPGMA guide trees, but comparable to ProbCons on the probtree guide tree, and better than ProbCons on the true tree.

In fact, trees estimated using the alignment produced by FTA using the true guide tree are even better than trees estimated from the true alignment. There are several possible explanations for this phenomenon, but further study is required.

The graphs we show in Figures 3 and 4 have values that have been averaged over all model conditions and replicates (for the given number of taxa). The relative performance of the methods shown in the averages holds (with few exceptions) for each model condition. However, the magnitudes of the actual errors and amount of improvement based on a given guide tree vary. Graphs for individual model conditions are available here: <http://www.cs.utexas.edu/users/serita/pubs/psb08-aux/>

3.3. Conclusions.

Except for FTA, MSA accuracy (as measured using SP-error) is not strongly correlated with guide tree accuracy. Further, for most of these MSA methods, phylogenetic accuracy is not directly predicted by the accuracy of the guide tree (except again, in the case of FTA). Although it is common to evaluate alignments purely in terms of criteria like SP (or column score), these experiments provide clear evidence that not all errors are of equal importance, at least in terms of phylogenetic consequences. This is not completely surprising, since when Ogden and Rosenberg⁹ studied the influence of tree shape on alignment and tree reconstruction accuracy they too found that alignment error did not always have a large impact on tree accuracy. Thus, although FTA alignments are often “worse” with respect to the SP-error, trees estimated from FTA alignments can be more accurate than trees estimated from other alignments with lower SP-error rates. Finally, it is important to realize that although alignments may have similar SP-error rates as compared to a true alignment, they can still be very different from each other.

The experiments show clearly that tree estimation can be improved through the use of improved guide trees, though only some alignment methods seem to be able to take advantage of these improved guide trees. It is also clear that these improvements require some additional computational effort. Is it worth it? Consider the following different methods, which we will call “Good” and “Better”.

- Good: Run ProbCons in its default setting, followed by RAxML.
- Better: Run ProbCons on one of the UPGMA guide trees, followed by RAxML. (Note that this method produces the “probtree” guide

tree, if the upgma1 guide tree is used.)

How much time do these methods take in our experiments?

In our experiments, run using a distributed system via Condor ⁷, alignment using ProbCons was the most expensive step in terms of running time. The Good technique took approximately 8 minutes on 25 taxa and slightly more than 2 hours for 100 taxa, while Better took under 9 minutes on 25 taxa and 2.5 hours for 100 taxa. In other words, for a very minimal increase in running time, substantial improvements in topological accuracy are obtainable.

4. Future Work

Our study shows clearly that improving the guide tree for MSA methods can improve estimated phylogenies, provided that appropriate multiple alignment methods are used. Furthermore, it shows that FTA can obtain better trees than the other methods tested when the guide tree is very good. Indeed, our data suggest that once the guide tree is within about 20% RF distance to the true tree, trees based upon FTA alignments will be highly accurate. Given these results, we will test an iterative approach to phylogeny and alignment estimation: begin with a good guide tree (e.g., probtree); compute FTA on the guide tree; and then compute a new guide tree for FTA by running RAxML on the resultant alignment (and then repeat the FTA/RAxML analysis). In the current experiments, RAxML and FTA were both very fast, even on the 100-taxon dataset, so the iterative approach may scale well to significantly larger numbers of taxa. Other future work will seek to develop new alignment-error metrics that better capture differences among alignments, specifically in terms of their ability to predict accuracy of subsequent phylogenetic inference.

References

1. C.B. Do, M.S.P. Mahabhashyam, M. Brudno, and S. Batzoglou. PROB-CONS: Probabilistic consistency-based multiple sequence alignment. *Genome Research*, 15:330–340, 2005.
2. R. C. Edgar. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, 5(113), 2004.
3. J. Felsenstein. *Inferring Phylogenies*. Sinauer Associates, Sunderland, Massachusetts, 2004.
4. J. Fredslund, J. Hein, and T. Scharling. A large version of the small parsimony problem. In Gary Benson and Roderic Page, editors, *Algorithms in*

- Bioinformatics: Third International Workshop, WABI 2003, LNCS*, volume 2812, pages 417–432, Berlin, 2003. Springer-Verlag.
5. K. Katoh, K. Kuma, H. Toh, and T. Miyata. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.*, 33(2):511–518, 2005.
 6. B. Knudsen. Optimal multiple parsimony alignment with affine gap cost using a phylogenetic tree. In G. Benson and R. Page, editors, *WABI 2003, LNBI 2812*, pages 433–446. Springer-Verlag, Berlin, 2003.
 7. Michael Litzkow. Remote unix - turning idle workstations into cycle servers. In *Usenix Summer Conference*, pages 381–384, 1987.
 8. B.M.E. Moret, U. Roshan, and T. Warnow. Sequence length requirements for phylogenetic methods. In *Proc. 2nd Int'l Workshop Algorithms in Bioinformatics (WABI'02)*, volume 2452 of *Lecture Notes in Computer Science*, pages 343–356. Springer-Verlag, 2002.
 9. T. Heath Ogden and Michael S. Rosenberg. Multiple sequence alignment accuracy and phylogenetic inference. *Systematic Biology*, 55(2):314–328, 2006.
 10. U. Roshan, D.R. Livesay, and S. Chikkagoudar. Improving progressive alignment for phylogeny reconstruction using parsimonious guide-trees. In *Proceedings of the IEEE 6th Symposium on Bioinformatics and Bioengineering*. IEEE Computer Society Press, 2006.
 11. M. J. Sanderson. r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, 19(2):301–302, 2003.
 12. D. Sankoff. Minimal mutation trees of sequences. *SIAM J. Appl. Math.*, 28(1):35 – 42, January 1975.
 13. Alexandros Stamatakis. Raxml-vi-hpc: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21):2688–2690, 2006.
 14. J. Stoye. Multiple sequence alignment with the divide-and-conquer method. *Gene*, 211:GC45–GC56, 1998.
 15. J. Stoye, D. Evers, and F. Meyer. Rose: generating sequence families. *Bioinformatics*, 14(2):157–163, 1998.
 16. D. Swofford. PAUP*: Phylogenetic analysis using parsimony (and other methods), version 4.0. 1996.
 17. J.D. Thompson, D.G. Higgins, and T.J. Gibson. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22:4673–4680, 1994.
 18. L. Wang and D. Gusfield. Improved approximation algorithms for tree alignment. *J. Algorithms*, 25:255–273, 1997.
 19. L. Wang and T. Jiang. On the complexity of multiple sequence alignment. *J. Comput. Biol.*, 1(4):337–348, 1994.
 20. L. Wang, T. Jiang, and D. Gusfield. A more efficient approximation scheme for tree alignment. *SIAM J. Comput.*, 30(1):283–299, 2000.
 21. D. Zwickl. GARLI download page. Website, 2006. <http://www.zo.utexas.edu/faculty/antisense/Garli.html>.