Optimization of Multilocus Sequence Analysis for Identification of Species in the Genus *Vibrio*

Michael W. Gabriel, George Y. Matsui, Robert Friedman, and Charles R. Lovell

Department of Biological Sciences

University of South Carolina

Columbia, SC 29208

*Corresponding author contact information:

Mailing Address: Dept. of Biological Sciences, 715 Sumter St., Columbia, SC, 29208
Phone: 803-777-5084
FAX: 803-777-4002
Email: lovell@biol.sc.edu

Michael Gabriel, gabriemw@email.sc.edu
George Matsui, matsui2003@gmail.com
Robert Friedman, bobf@biol.sc.edu

Running title: Improved MLSA of vibrios

Key Words: Multilocus Sequence Analysis, Phylogenetic Analysis, *Vibrio*
Abstract.

Multilocus Sequence Analysis (MLSA) is an important method for identification of taxa that are not well differentiated by 16S rRNA gene sequences alone. In this procedure, concatenated sequences of selected genes are constructed and then analyzed. The effects that the number and the order of genes used in MLSA have on reconstruction of phylogenetic relationships were examined. The recA, rpoA, gapA, 16S rRNA gene, gyrB, and ftsZ sequences from 56 species of the genus Vibrio were used to construct molecular phylogenies and these were evaluated individually and using various gene combinations. Phylogenies from two-gene sequences employing recA and rpoA, in both possible gene orders were different. The addition of the gapA gene sequence, producing all six possible concatenated sequences, reduced the differences in phylogenies to degrees of statistical (bootstrap) support for some nodes. The overall statistical support for the phylogenetic tree, assayed on the basis of a reliability score (calculated from the number of nodes having bootstrap values $\geq 80$ divided by the total number of nodes) increased with increasing numbers of genes used, up to a maximum of four. No further improvement was observed from addition of the fifth gene sequence (ftsZ) and addition of the sixth gene (gyrB) resulted in lower proportions of strongly supported nodes. Reductions in the numbers of strongly supported nodes were also observed when maximum parsimony was employed for tree construction. Use of a small number of gene sequences in MLSA resulted in accurate identification of Vibrio species.
Introduction.

The genus *Vibrio* consists of over one hundred validly described species of Gram negative, mainly marine bacteria. Many additional candidate species have been noted (e.g. 1), but have not yet been described in the literature (2, 3). Most vibrios are versatile and fast growing chemoheterotrophs. Several species are also diazotrophic, contributing combined nitrogen to marine ecosystems (4, 5, 6). In addition to their participation in nutrient cycling, many vibrios engage in very close relationships with higher organisms. These interactions range from the bioluminescence symbiosis of *Vibrio fischeri* with the Hawaiian bobtail squid, *Euprymna scolopes* (7, 8, 9), to the many pathogenic interactions between a variety of *Vibrio* species and marine fauna. For example, *Vibrio alginolyticus* and *Vibrio splendidus* are bivalve associated pathogens (10, 11, 12), *Vibrio vulnificus* causes vibriosis in eels (13, 14, 15), *Vibrio ordalii* is a pathogen of fishes (16, 17), and *V. harveyi* and *Vibrio cambelli* are pathogenic to shrimp (18, 19). Several *Vibrio* species are also important opportunistic human pathogens. The best known of these are *Vibrio cholerae* (20, 21, 22), *Vibrio parahaemolyticus* (23, 24), *Vibrio vulnificus* (15, 25), but *Vibrio cincinnatiensis* (26), *Vibrio fluvialis* (27, 28, 29), *Vibrio furnissii* (30), *Vibrio metschnikovii* (31, 32, 33), and *Vibrio mimicus* (34, 35, 36) can also cause infections in human hosts. These organisms are obviously of great interest, as are additional *Vibrio* species that have been shown to carry an array of virulence-related genes (1).

In general, relatively little divergence of 16S rRNA gene sequences occurs among many *Vibrio* species (37), complicating species identification. Molecular phylogenetics is particularly problematic in the case of *Vibrio* species that are known or potential pathogens. All such species are very closely related to more benign species, often making correct identification of pathogenic
isolates difficult (1, 38, 39, 40). Genes other than 16S rRNA genes, including the recombinase alpha subunit \((\text{recA})\) (41, 42), have been employed to differentiate species and to construct phylogenies, but the phylogenetic resolution that can be obtained from any single gene is perforce limited. This has led to the widespread use of Multilocus Sequence Analysis (MLSA), which employs a number of housekeeping genes, joined end to end to construct concatenated sequences for phylogenetic analysis (42, 43, 44). The advantages of this approach include extensive databases of useful reference sequences, the low cost of DNA sequencing relative to that of detailed physiological and immunological characterizations, and the speed, ease, and accuracy of data collection and analysis. Typically, “housekeeping” genes that encode proteins essential to cellular reproduction are employed, as these genes are sufficiently conservative to allow accurate and facile sequence comparisons, while also encompassing sufficient sequence variation to provide the needed resolution of different species (43, 44). Genes commonly employed in MLSA analyses include \(\text{recA}\) (41, 42), RNA polymerase alpha subunit \((\text{rpoA})\) (42, 43), glyceraldehyde-3-phosphate dehydrogenase alpha subunit \((\text{gapA})\) (42, 45), cell division protein \((\text{ftsZ})\) (46), and DNA gyrase beta subunit \((\text{gyrB})\) (47), along with several others (42, 44, 48).

A broad range of gene sequences are available and, as sequencing of whole genomes continues to expand, many more are becoming available. Thus it becomes necessary to address how many gene sequences should be employed and how they should be employed to produce efficient, economical, and robust MLSA results. The assumption implicit in many \textit{Vibrio} phylogenetic studies is that addition of more genes to the analysis results in more accurate representation of the relationships of species, but this assumption has not yet been subjected to
rigorous testing (42, 49, 50, 51). In addition to the genes used, the impact of gene order has not been established. The objective of this study was to determine the impacts of gene numbers and orders in the concatenated sequences on the accuracy and precision of MLSA of vibrios. We have established that the sensitivity of MLSA saturates for concatenated sequences only a few genes in length and addition of more sequences can in fact compromise the reliability of the method.

Materials and Methods

Gene sequences for the 16S rRNA gene, ftsZ, gapA, gyrB, recA, and rpoA were downloaded from the NCBI GenBank database before February 2012. Sequences of a given gene that were 65% shorter than the mean sequence length (Supplemental Table S1) were excluded from the analysis. Sequences from each Vibrio species, unless otherwise noted, are from the type strain or from a well characterized reference strain.

Individual gene sequences were validated by alignment using ClustalW in MEGA5 with the default alignment parameters (52). The alignments were checked manually and any alignment and translation errors were corrected. Gene sequences were then combined to produce concatenated sequences for MLSA. All possible combinations of two or three genes (2 and 6 combinations, respectively) were constructed, then gene sequences were added to the 6 three gene concatamers to yield higher order concatamers. Concatenated sequences were aligned and checked for alignment errors in the same manner as the individual genes. The lengths of sequences for a given gene varied, as each gene sequence set included both full-length sequences and shorter sequences derived from PCR amplicons (Table 1). Since a particular gene sequence
might not be available for all species, the addition of gene sequences to concatamers reduced the number of species that could be included in the analysis. The number of species for which single gene sequences were available varied from 89 species for the 16S rRNA gene to 58 species for gyrB (Supplemental Table 1). Seventy species could be included in analyses based on the initial two-gene concatamers, but only 41 species were represented in the 6-gene concatamers (Supplemental Table 1).

Phylogenetic trees were constructed using the maximum likelihood, maximum parsimony, minimum evolution, and neighbor joining methods using MEGA5 (52). Settings utilized in the construction of the phylogenetic trees for the individual methods are given in Table 1. The clades and tree topologies from the various methods were compared for commonalities and differences. Bootstrap analysis was employed to quantitatively compare the trees. The use of bootstrap analysis in the reconstruction of phylogenetic trees provides a statistical validation through random resampling of the topology presented in the consensus tree (53, 54, 55). The number of nodes having bootstrap values greater than 80 were tallied and divided by the total number of nodes present in the tree to yield a reliability score for that tree. Statistical evaluation of variability in these scores employed the Ryan-Einot-Gabriel-Welsh F test (REGW F) in the statistical package SPSS (56), and differences were deemed statistically significant for $P \leq 0.05$.

Results and Discussion

The construction of phylogenetic trees from the selected gene sequences using different models resulted in some differences in phylogenies. The most extensive changes occurred between the minimum evolution and neighbor joining methods. These changes included
differences in the bifurcations present in the phylogenetic trees and most frequently occurred at
bifurcations supported by bootstrap values less than 80. Phylogenetic trees constructed using
maximum likelihood included many of the clades previously described by Sawabe et al. (49, 50)
(Figure 1), with additional species added to some clades.

The model employed to reconstruct phylogenetic trees also had an impact on the
phylogram mean reliability score. The maximum likelihood, minimum evolution, and neighbor
joining models produced mean reliability scores that were not significantly different (Table 2)
and phylogenies that had greater reliability than those of the maximum parsimony trees. The
highest reliability score from each method employing a single concatamer was 0.623, 0.604,
0.623 and 0.558 for maximum likelihood (16S rRNA gene-*gapA-recA-rpoA*), neighbor joining
(16S rRNA gene-*gapA-rpoA-recA*), and maximum parsimony (16S rRNA gene-*gapA-recA-
rpoA-gyrB*) respectively.

The number of gene sequences present in the concatamer also influenced the reliability
score. The addition of genes to the concatamers, followed by use of the maximum likelihood,
minimum evolution, and neighbor joining models, resulted in increased reliability through the
addition of the fourth gene (Table 2). Addition of the fifth gene did not significantly increase
reliability and addition of a sixth gene decreased the reliability of the phylogenetic analysis. The
reliability scores for trees constructed using maximum parsimony increased with addition of
genes to the concatamers up to a maximum of five genes. The addition of the sixth gene caused
the score to decrease to that obtained using four genes.

The phylogenetic trees constructed with two genes, *recA* and *rpoA*, included thirteen of
the fourteen clades previously described by Sawabe et al. (49) and this was not affected by the
order of the two genes. Both phylogenetic reconstructions had similar topologies, although there were changes in bifurcations having bootstrap support lower than 80. The addition of \textit{gapA} to the concatamers, in all six gene order combinations with \textit{recA} and \textit{rpoA}, had minimal impact on the tree topology. Thirteen of the fourteen clades were retained in phylogenetic trees using these concatamers with the Vulnificus clade decomposing and the placement of \textit{V. navarrensis} in the Gazogenes clade when three of the six concatamers were employed (\textit{gapA-recA-rpoA}, \textit{recA-gapA-rpoA}, and \textit{rpoA-recA-gapA}). The Vulnificus clade had low bootstrap support when it was observed. The addition of the 16S rRNA gene to the six order combinations of \textit{gapA}, \textit{recA}, and \textit{rpoA} had minimal effect on tree topology, with changes occurring only at bifurcations having low support. The addition of the 16S rRNA gene to the previous AASs retained all fourteen clades, although three clades (Orientalis, Nigripulchritudo, and Vulnificus) had low bootstrap support. The addition of the fourth gene to the concatamers allowed the closely related species \textit{V. harveyi}, \textit{V. campbellii}, and \textit{V. rotiferianus} to be differentiated with significant bootstrap values in all models except for maximum parsimony.

Thirteen of the fourteen clades were resolved following addition of the \textit{gyrB} gene to the AASs containing the 16S rRNA gene, \textit{gapA}, \textit{recA}, and \textit{rpoA}. The Orientalis clade decomposed with \textit{V. tubiashii} grouping with the Coralliilyticus clade (with significant bootstrap support) in five of the concatamers. Bootstrap support for the Coralliilyticus clade was not significant for the tree produced using the 16S rRNA gene-\textit{gapA-rpoA-recA-gyrB} concatamer. Minor topological changes occurred at bifurcations having low bootstrap support. The addition of the \textit{ftsZ} gene resulted in decomposition of the Orientalis clade, as well as minor topological changes at nodes having low bootstrap support. Clade support decreased with the addition of \textit{ftsZ} in three of the six combinations, resulting in decreases below a bootstrap value of 80 for the Cholerae.
and Orientalis clades. These concatamers produced the only phylogenetic reconstructions in which bootstrap support for the Cholerae clade was less than 80.

Gene order had differing effects on the reconstruction of phylogenetic relationships, based on the model used. Phylogenetic trees reconstructed using the minimum evolution method appeared to be most susceptible to changes in gene order when using two genes. In the recA-rpoA concatamer the number of nodes having bootstrap scores greater than 80 was 29, and in the rpoA-recA order the number of strongly supported nodes was only 25. Another notable change occurred when using the maximum parsimony model. In the case of concatamers containing three genes in the order rpoA-recA-gapA, 28 nodes had bootstrap support greater than or equal to 80, while the average number of nodes with values greater than or equal to 80 for the remaining permutations was 23.1 \pm 2.5. The order of genes used in the maximum parsimony model appears to have an impact on the number of nodes that are statistically supported. The variation driven by gene order appears to be dampened with an increase in the number of genes utilized, with most variation occurring in the two gene concatamers and that variation decreasing with each gene addition, with the previously stated exceptions. The variation of the reliability scores yielded by all four methods suggests that gene order can have an impact on the outcome of the analysis, with greatest impact occurring in the maximum parsimony method.

The reliability of phylogenetic reconstructions employing concatamers having different numbers of genes varied depending on the model employed. The REGW F test, when applied to reliability scores for trees built from concatamers employing from three genes up to six genes, defined two subsets at $P \leq 0.05$. The first subset included only trees built using the maximum parsimony method, which produced mean reliability scores that were significantly lower in all cases than those from the other models. The second subset included the maximum likelihood,
minimum evolution and neighbor joining models, which were not significantly different from each other (Table 3).

The optimum number of genes employed in the concatamers is influenced by the phylogenetic model used (Table 4). The REGW F test results grouped the maximum likelihood and neighbor joining products into three subsets. Concatamers utilizing five or four genes yielded higher reliability scores than those containing two, three and six genes. The grouping of the four and five gene concatamers indicates that these two means are not significantly different at P = 0.05. The minimum evolution model produced four subsets, with the four gene concatamers grouped separately, indicating that the highest reliability score occurs in the minimum evolution model with four genes. Maximum parsimony grouped the concatamers into three subsets that overlapped extensively and the highest reliability score was obtained using four genes.

This analysis showed that the specific genes employed in MLSA, the numbers of these genes, the orders of genes in concatamers, and the model used to reconstruct phylogenetic trees for species in the genus *Vibrio* all have impacts on the reliability of the analysis. The order of genes appeared to matter most when either very short concatamers or the maximum parsimony model was used. The maximum parsimony model removes from the analysis all sites that do not affect tree topology, therefore reducing the amount of data that is available for phylogenetic reconstruction (57) and likely contributing to the impact of gene order. The number of genes used in the concatamers appears to have a strong impact on the reliability of the final phylogenetic tree, but not necessarily the impact expected. For the maximum likelihood, minimum evolution, maximum parsimony, and neighbor joining models, the optimum number of genes that should be employed is four genes, as four gene concatamers yielded the highest mean
reliability scores. Including more genes in concatamers in an effort to increase accuracy of the
MLSA method may not yield the expected result, but does require more sequencing, a greater
time commitment to concatamer construction and quality control, and is more computationally
intensive. It should be noted that the concatamers used in this study yielded the most reliable
phylogenetic reconstructions for species of *Vibrio*, but the use of different genes may yield
different results and use of the genes employed here may yield different results for other
bacterial genera. When looking at genera that include more divergent species, the number of
genes required to accurately define species could be fewer than in genera consisting of less
divergent species. The model employed can also impact the overall reliability of the MLSA
phylogenetic reconstructions.

Acknowledgements

This work was supported by South Carolina SeaGrant Sub-Award N156 to CRL. MG
acknowledges support from an American Society for Microbiology Undergraduate Research
Fellowship. The authors would also like to thank Dr. Jay Pinckney for statistical assistance.

References


shrimp from south India and establishment of its pathogenic potential in an *Artemia* model. Microbiology 157:179-188.


---

**Figure Legend.**

Figure 1: Unrooted phylogenetic reconstruction of the genus *Vibrio* constructed using the Kimura 2-Parameter model with the maximum likelihood method, bootstrapped 1000 times, with concatenated gene sequence order of 16S rRNA gene-*gapA-recA-rpoA*. Bootstrap values below 80 are not shown. Clades shown were defined by Sawabe et al. (47)
Table 1: Settings used in MEGA5 to construct phylogenetic trees for the maximum likelihood, minimum parsimony, minimum evolution, and neighbor joining methods\(^1\).

<table>
<thead>
<tr>
<th>Model</th>
<th>Bootstrap Replicates</th>
<th>Substitution Model</th>
<th>Model</th>
<th>Substitutions to include</th>
<th>Rates among sites</th>
<th>Gaps/ Missing data treatment</th>
<th>Select Codon Positions</th>
<th>Heuristic Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum Evolution</td>
<td>1000</td>
<td>Nucleotide</td>
<td>K-2</td>
<td>TT</td>
<td>(\gamma = 4)</td>
<td>Complete Deletion</td>
<td>All Sites</td>
<td>NA</td>
</tr>
<tr>
<td>Neighbor Joining</td>
<td>1000</td>
<td>Nucleotide</td>
<td>K-2</td>
<td>TT</td>
<td>(\gamma = 4)</td>
<td>Complete Deletion</td>
<td>All Sites</td>
<td>NA</td>
</tr>
<tr>
<td>Maximum Likelihood</td>
<td>1000</td>
<td>Nucleotide</td>
<td>K-2</td>
<td>NA</td>
<td>(\gamma = 4)</td>
<td>Complete Deletion</td>
<td>All Sites</td>
<td>NNI</td>
</tr>
<tr>
<td>Maximum Parsimony</td>
<td>1000</td>
<td>Nucleotide</td>
<td>NA</td>
<td>NA</td>
<td>Complete Deletion</td>
<td>All Sites</td>
<td>SPR</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)NA = Not Applicable, K-2 = Kimura 2-Parameter, TT = Transitions and Transversions, NNI = Nearest Neighbor Interchange, NJ-BioNJ = Make initial tree automatically (default NJ/BioNJ), VS = Branch Swap Filter (Very Strong), NT = Number of Threads, SPR = Subtree Pruning Regrafting, SL = Search Level, and MT = Max Number of Trees to Retain
Table 2: Mean reliability scores (number of nodes with bootstrap values greater than or equal to 80 divided by the total number of nodes) in phylogenetic reconstructions using the maximum likelihood, minimum evolution, neighbor joining, and maximum parsimony methods. Standard deviations represent one standard deviation from the mean. The genes used in the concatamers are as follows: (16S rRNA gene, ftsZ, gapA, gyrB, recA, rpoA) in 1 gene concatamers, (recA and rpoA) in 2 gene concatamers, (recA, rpoA, gapA) in 3 gene concatamers, (16S rRNA gene, recA, rpoA, gapA, gyrB) in 4 gene concatamers, (16S rRNA gene, recA, rpoA, gapA, gyrB, ftsZ) in 5 gene concatamers, and (16S rRNA gene, recA, rpoA, gapA, gyrB, ftsZ, recA) in 6 gene concatamers.

<table>
<thead>
<tr>
<th>Number of Genes</th>
<th>Maximum Likelihood</th>
<th>Minimum Evolution</th>
<th>Neighbor Joining</th>
<th>Maximum Parsimony</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.27 ± 0.09</td>
<td>0.31 ± 0.10</td>
<td>0.31 ± 0.10</td>
<td>0.22 ± 0.09</td>
</tr>
<tr>
<td>2</td>
<td>0.43 ± 0.01</td>
<td>0.40 ± 0.04</td>
<td>0.42 ± 0.01</td>
<td>0.41 ± 0.01</td>
</tr>
<tr>
<td>3</td>
<td>0.53 ± 0.02</td>
<td>0.54 ± 0.01</td>
<td>0.53 ± 0.02</td>
<td>0.44 ± 0.04</td>
</tr>
<tr>
<td>4</td>
<td>0.59 ± 0.02</td>
<td>0.60 ± 0.02</td>
<td>0.59 ± 0.02</td>
<td>0.47 ± 0.02</td>
</tr>
<tr>
<td>5</td>
<td>0.57 ± 0.01</td>
<td>0.57 ± 0.01</td>
<td>0.57 ± 0.01</td>
<td>0.51 ± 0.03</td>
</tr>
<tr>
<td>6</td>
<td>0.55 ± 0.01</td>
<td>0.56 ± 0.01</td>
<td>0.55 ± 0.01</td>
<td>0.47 ± 0.01</td>
</tr>
</tbody>
</table>
Table 3: Inter-model comparison of mean reliability scores (number of nodes with bootstrap values greater than or equal to 80 divided by total number of nodes) of phylogenetic trees reconstructed using the maximum parsimony, maximum likelihood, minimum evolution, and neighbor joining methods grouped into subsets (groups of homogenous means) with a P = 0.05 confidence level using the Ryan-Einot-Gabriel-Welsch F test.


<table>
<thead>
<tr>
<th>Number of Genes</th>
<th>Model</th>
<th>Subset</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Maximum Parsimony</td>
<td>0.437</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum Likelihood</td>
<td>0.538</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimum Evolution</td>
<td>0.528</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neighbor Joining</td>
<td>0.541</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Maximum Parsimony</td>
<td>0.475</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum Likelihood</td>
<td>0.588</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimum Evolution</td>
<td>0.597</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neighbor Joining</td>
<td>0.585</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Maximum Parsimony</td>
<td>0.512</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum Likelihood</td>
<td>0.574</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimum Evolution</td>
<td>0.570</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neighbor Joining</td>
<td>0.574</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Maximum Parsimony</td>
<td>0.467</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Maximum Likelihood</td>
<td>0.549</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimum Evolution</td>
<td>0.537</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Neighbor Joining</td>
<td>0.545</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Intra-model comparison of mean reliability scores (number of nodes with bootstrap values greater than or equal to 80 divided by total number of nodes) of phylogenetic trees reconstructed using the maximum likelihood, neighbor joining, minimum evolution, and maximum parsimony methods grouped into subsets, a group of homogenous means listed from lowest to highest, with a P = 0.05 confidence level using the Ryan-Einot-Gabriel-Welsh F test. The genes used in the concatamers are as follows: (16S rRNA gene, ftsZ, gapA, gyrB, recA, rpoA) in 1 gene concatamers, (recA and rpoA) in 2 gene concatamers, (recA, rpoA, gapA) in 3 gene concatamers, (16S rRNA gene, recA, rpoA, gapA) in 4 gene concatamers, (16S rRNA gene, recA, rpoA, gapA, gyrB) in 5 gene concatamers, and (16S rRNA gene, recA, rpoA, gapA, gyrB, ftsZ) in 6 gene concatamers.

<table>
<thead>
<tr>
<th>Model</th>
<th>Number of Genes</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maximum Likelihood</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.425</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.538</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>0.549</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>0.574</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.588</td>
</tr>
<tr>
<td><strong>Neighbor Joining</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.418</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.528</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td>0.545</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>0.574</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.585</td>
</tr>
<tr>
<td><strong>Minimum Evolution</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.403</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.541</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.557</td>
<td>0.557</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>0.570</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>0.597</td>
<td></td>
</tr>
<tr>
<td><strong>Maximum Parsimony</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.410</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0.437</td>
<td>0.437</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>0.467</td>
<td>0.467</td>
<td>0.467</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>0.475</td>
<td>0.475</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

