AEM
Journals.ASM.org

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

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

Department of Biological Sciences, University of South Carolina, Columbia, South Carolina, USA

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 of ≥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.

The genus *Vibrio* consists of more than 100 validly described species of Gram-negative, mainly marine bacteria. Many additional candidate species have been noted (e.g., see reference 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 (4–6). 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* (7–9), 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 (10–12), *Vibrio vulnificus* causes vibriosis in eels (13–15), *Vibrio ordalii* is a pathogen of fishes (16, 17), and *V. harveyi* and *Vibrio campbellii* are pathogenic to shrimp (18, 19). Several *Vibrio* species are also important opportunistic human pathogens. The best known of these are *Vibrio cholerae* (20–22), *Vibrio parahaemolyticus* (23, 24), and *Vibrio vulnificus* (15, 25), but *Vibrio cincinnatiensis* (26), *Vibrio fluvialis* (27–29), *Vibrio furnissii* (30), *Vibrio metschnikovii* (31–33), and *Vibrio mimicus* (34–36) 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, 38–40). Genes other than 16S rRNA genes, including the recombinase alpha subunit (*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 (42–44). 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 include *recA* (41, 42) and genes for RNA polymerase alpha subunit (*rpoA*) (42, 43), glyceraldehyde-3-phosphate dehydrogenase alpha subunit (*gapA*) (42, 45), cell division protein (*ftsZ*) (46), and DNA gyrase beta subunit (*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 *Vibrio* phylogenetic studies is that addition of more genes to the analysis results in more

Received 10 April 2014 **Accepted** 16 June 2014
Published ahead of print 20 June 2014
Editor: C. A. Elkins
Address correspondence to Charles R. Lovell, lovell@biol.sc.edu.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.01206-14.

doi:10.1128/AEM.01206-14

TABLE 1 Settings used in MEGA5 to construct phylogenetic trees for the maximum likelihood, minimum parsimony, minimum evolution, and neighbor joining methods[a]

Model	No. of bootstrap replicates	Substitution model	Model	Substitutions to include	Rates among sites	Gaps/missing data treatment	Select codon positions	Heuristic method(s)
Minimum evolution	1,000	Nucleotide	K-2	TT	$\gamma = 4$	Complete deletion	All sites	NA
Neighbor joining	1,000	Nucleotide	K-2	TT	$\gamma = 4$	Complete deletion	All sites	NA
Maximum likelihood	1,000	Nucleotide	K-2	NA	$\gamma = 4$	Complete deletion	All sites	NNI NJ-BioNJ VS NT = 1
Maximum parsimony	1,000	Nucleotide	NA	NA	NA	Complete deletion	All sites	SPR NTRA = 10 SL = 1 MT = 100

[a] 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; MT, maximum number of trees to retain.

accurate representation of the relationships of species (2, 3, 42, 50), but this assumption has not yet been subjected to rigorous testing. 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 that the sensitivity of MLSA saturates for concatenated sequences only a few genes in length and that 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 (see Table S1 in the supplemental material) 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 (51). 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, and then gene sequences were added to the 6 three-gene concatemers to yield higher-order concatemers. 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 concatemers 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* (see Table S1 in the supplemental material). Seventy species could be included in analyses based on the initial two-gene concatemers, but only 41 species were represented in the 6-gene concatemers (see Table S1).

Phylogenetic trees were constructed using the maximum likelihood, maximum parsimony, minimum evolution, and neighbor joining methods using MEGA5 (51). 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 (52–54). 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 (55), and differences were deemed statistically significant for a P value of ≤ 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. (2, 3) (Fig. 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 scores from each method employing a single concatemer were 0.623, 0.604, 0.623, and 0.558 for maximum likelihood (16S rRNA gene-*gapA-recA-rpoA*), neighbor joining (16S rRNA gene-*gapA-recA-rpoA* and 16S rRNA gene-*rpoA-gapA-recA*), minimum evolution (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 concatemer also influenced the reliability score. The addition of genes to the concatemers, 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 the sixth gene decreased the reliability of the phylogenetic analysis. The reliability scores for trees constructed using

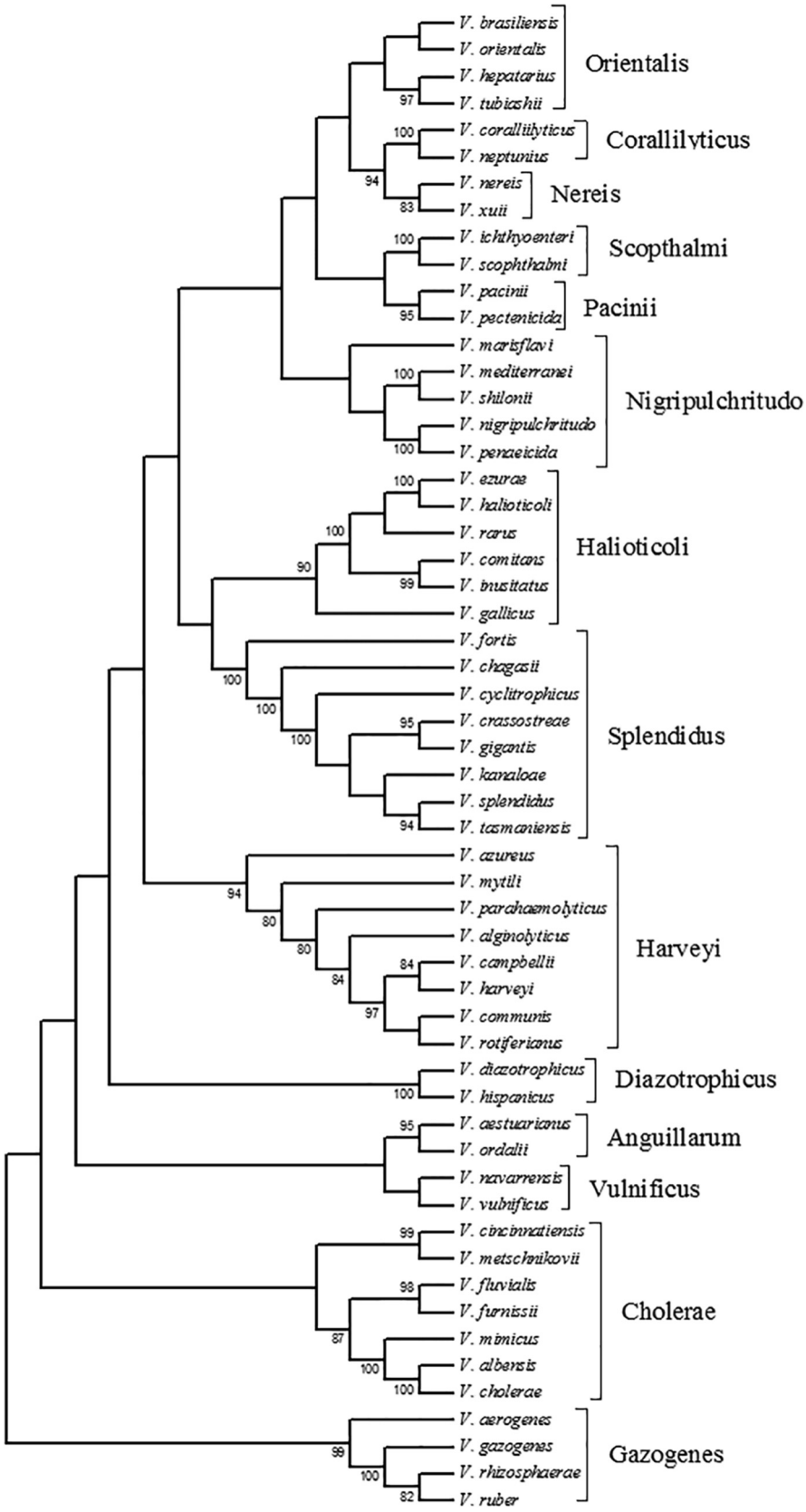

FIG 1 Unrooted phylogenetic reconstruction of the genus *Vibrio* constructed using the Kimura 2-parameter model with the maximum likelihood method, bootstrapped 1,000 times, with concatenated gene sequence order of 16S rRNA gene-*gapA*-*recA*-*rpoA*. Bootstrap values below 80 are not shown. Clades shown were defined by Sawabe et al. (2).

TABLE 2 Mean reliability scores in phylogenetic reconstructions[a]

No. of genes in concatemers	Reliability score, mean ± SD			
	Maximum likelihood	Minimum evolution	Neighbor joining	Maximum parsimony
1	0.27 ± 0.09	0.31 ± 0.10	0.31 ± 0.10	0.22 ± 0.09
2	0.43 ± 0.01	0.40 ± 0.04	0.42 ± 0.01	0.41 ± 0.01
3	0.53 ± 0.02	0.54 ± 0.01	0.53 ± 0.02	0.44 ± 0.04
4	0.59 ± 0.02	0.60 ± 0.02	0.59 ± 0.02	0.47 ± 0.02
5	0.57 ± 0.01	0.57 ± 0.01	0.57 ± 0.01	0.51 ± 0.03
6	0.55 ± 0.01	0.56 ± 0.01	0.55 ± 0.01	0.47 ± 0.01

[a] Each reliability score is the number of nodes with bootstrap values greater than or equal to 80 divided by the total number of nodes. Standard deviations represent 1 standard deviation from the mean. The genes used in the concatemers are as follows: 16S rRNA gene, *ftsZ*, *gapA*, *gyrB*, *recA*, or *rpoA* in 1-gene concatemers, *recA* and *rpoA* in 2-gene concatemers, *recA*, *rpoA*, and *gapA* in 3-gene concatemers, 16S rRNA gene, *recA*, *rpoA*, and *gapA* in 4-gene concatemers, 16S rRNA gene, *recA*, *rpoA*, *gapA*, and *gyrB* in 5-gene concatemers, and 16S rRNA gene, *recA*, *rpoA*, *gapA*, *gyrB*, and *ftsZ* in 6-gene concatemers.

TABLE 3 Intermodel comparison of mean reliability scores[a]

No. of genes	Model	Reliability score for subset	
		1	2
3	Maximum parsimony	0.437	
	Maximum likelihood		0.538
	Minimum evolution		0.528
	Neighbor joining		0.541
4	Maximum parsimony	0.475	
	Maximum likelihood		0.588
	Minimum evolution		0.597
	Neighbor joining		0.585
5	Maximum parsimony	0.512	
	Maximum likelihood		0.574
	Minimum evolution		0.570
	Neighbor joining		0.574
6	Maximum parsimony	0.467	
	Maximum likelihood		0.549
	Minimum evolution		0.557
	Neighbor joining		0.545

[a] Each reliability score is the 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* value of 0.05 for confidence level using the Ryan-Einot-Gabriel-Welsch F test. The genes used in the concatemers are as follows: 16S rRNA gene, *ftsZ*, *gapA*, *gyrB*, *recA*, or *rpoA* in 1-gene concatemers, *recA* and *rpoA* in 2-gene concatemers, *recA*, *rpoA*, and *gapA* in 3-gene concatemers, 16S rRNA gene, *recA*, *rpoA*, and *gapA* in 4-gene concatemers, 16S rRNA gene, *recA*, *rpoA*, *gapA*, and *gyrB* in 5-gene concatemers, and 16S rRNA gene, *recA*, *rpoA*, *gapA*, *gyrB*, and *ftsZ* in 6-gene concatemers.

maximum parsimony increased with addition of genes to the concatemers 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 13 of the 14 clades previously described by Sawabe et al. (2), and this was not affected by the order of the two genes. The two phylogenetic reconstructions had similar topologies, although there were changes in bifurcations having bootstrap support lower than 80. The addition of *gapA* to the concatemers, in all six gene order combinations with *recA* and *rpoA*, had minimal impact on the tree topology. Thirteen of the 14 clades were retained in phylogenetic trees using these concatemers, with the Vulnificus clade decomposing and the placement of *V. navarrensis* in the Gazogenes clade when three of the six concatemers were employed (*gapA-recA-rpoA*, *recA-gapA-rpoA*, and *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 *gapA*, *recA*, and *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 concatemers retained all 14 clades, although three clades (Orientalis, Nigripulchritudo, and Vulnificus) had low bootstrap support. The addition of the fourth gene to the concatemers allowed the closely related species *V. harveyi*, *V. campbellii*, and *V. rotiferianus* to be differentiated with significant bootstrap values in all models except for maximum parsimony.

Thirteen of the 14 clades were resolved following addition of the *gyrB* gene to the concatemers containing the 16S rRNA gene, *gapA*, *recA*, and *rpoA*. The Orientalis clade decomposed with *V. tubiashii* grouping with the Coralliilyticus clade (with significant bootstrap support) in five of the concatemers. Bootstrap support for the Coralliilyticus clade was not significant for the tree produced using the 16S rRNA gene-*gapA-rpoA-recA-gyrB* concatemer. Minor topological changes occurred at bifurcations having low bootstrap support. The addition of the *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 *ftsZ* in three of the six combinations, resulting in decreases below a bootstrap value of 80 for the Cholerae and Orientalis clades. These concatemers 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* concatemer 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 concatemers 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 ± 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 concatemers 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 the greatest impact occurring in the maximum parsimony method.

The reliability of phylogenetic reconstructions employing concatemers having different numbers of genes varied depending on the model employed. The REGW F test, when applied to reliability

TABLE 4 Intramodel comparison of mean reliability scores[a]

Model	No. of genes	Reliability score for subset			
		1	2	3	4
Maximum likelihood	2	0.425			
	3		0.538		
	6		0.549		
	5			0.574	
	4			0.588	
Neighbor joining	2	0.418			
	3		0.528		
	6		0.545		
	5			0.574	
	4			0.585	
Minimum evolution	2	0.403			
	3		0.541		
	6		0.557	0.557	
	5			0.570	
	4				0.597
Maximum parsimony	2	0.410			
	3	0.437	0.437		
	6	0.467	0.467	0.467	
	4		0.475	0.475	
	5			0.512	

[a] Each reliability score is the 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 is listed from lowest to highest, with a *P* value of 0.05 for confidence level using the Ryan-Einot-Gabriel-Welsh F test. The genes used in the concatemers are as follows: 16S rRNA gene, *ftsZ*, *gapA*, *gyrB*, *recA*, or *rpoA* in 1-gene concatemers, *recA* and *rpoA* in 2-gene concatemers, *recA*, *rpoA*, and *gapA* in 3-gene concatemers, 16S rRNA gene, *recA*, *rpoA*, and *gapA* in 4-gene concatemers, 16S rRNA gene, *recA*, *rpoA*, *gapA*, and *gyrB* in 5-gene concatemers, and 16S rRNA gene, *recA*, *rpoA*, *gapA*, *gyrB*, and *ftsZ* in 6-gene concatemers.

scores for trees built from concatemers employing from three genes up to six genes, defined two subsets at *P* values of ≤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 concatemers 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. Concatemers 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 concatemers indicates that these two means are not significantly different at a *P* value of 0.05. The minimum evolution model produced four subsets, with the four-gene concatemers grouped separately, indicating that the highest reliability score occurs in the minimum evolution model with four genes. Maximum parsimony grouped the concatemers 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 concatemers, 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 concatemers 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 (56) and likely contributing to the impact of gene order. The number of genes used in the concatemers 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 concatemers yielded the highest mean reliability scores. Including more genes in concatemers in an effort to increase the accuracy of the MLSA method may not yield the expected result but does require more sequencing and a greater time commitment to concatemer construction and quality control, and it is more computationally intensive. It should be noted that the concatemers 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 for genera consisting of less divergent species. The model employed can also impact the overall reliability of the MLSA phylogenetic reconstructions.

ACKNOWLEDGMENTS

This work was supported by South Carolina SeaGrant Sub-Award N156 to C.R.L. M.W.G. acknowledges support from an American Society for Microbiology undergraduate research fellowship.

We thank Jay Pinckney for statistical assistance.

REFERENCES

1. **Klein SL, Guitierrez West CK, Mejia DM, Lovell CR.** 2014. Genes similar to the *Vibrio parahaemolyticus* virulence-related genes *tdh*, *tlh*, and *vscC2* occur in other *Vibrionaceae* species isolated from a pristine estuary. Appl. Environ. Microbiol. **80:**595–602. http://dx.doi.org/10.1128/AEM.02895-13.
2. **Sawabe T, Kita-Tsukamoto K, Thompson FL.** 2007. Inferring the evolutionary history of vibrios by means of multilocus sequence analysis. J. Bacteriol. **189:**7932–7936. http://dx.doi.org/10.1128/JB.00693-07.
3. **Sawabe T, Ogura Y, Matsumura Y, Feng G, Amin AKMR, Mino S, Nakagawa S, Sawabe T, Kumar R, Fukui Y, Santomi M, Matsushima R, Thompson FL, Gomez-Gil B, Christian R, Maruyama F, Kurokawa K, Hayashi T.** 2013. Updating the Vibrio clades defined by multilocus sequence phylogeny: proposal of eight new clades, and the novel description of *Vibrio tritonius* sp. nov. Front. Microbiol. **4:**414. http://dx.doi.org/10.3389/fmicb.2013.00414.
4. **Brown MM, Friez MJ, Lovell CR.** 2003. Expression of *nifH* genes by diazotrophic bacteria in the rhizosphere of short form *Spartina alterniflora*. FEMS Microbiol. Ecol. **43:**411–417. http://dx.doi.org/10.1111/j.1574-6941.2003.tb01081.x.
5. **Criminger JD, Hazen TH, Sobecky PA, Lovell CR.** 2007. Nitrogen fixation by *Vibrio parahaemolyticus* and its implications for a new ecological niche. Appl. Environ. Microbiol. **73:**5959–5961. http://dx.doi.org/10.1128/AEM.00981-07.
6. **Urdaci MC, Stal LJ, Marchand M.** 1988. Occurrence of nitrogen fixation among *Vibrio* spp. Arch. Microbiol. **150:**224–229. http://dx.doi.org/10.1007/BF00407784.
7. **Miyashiro T, Ruby EG.** 2012. Shedding light on bioluminescence regulation in *Vibrio fischeri*. Mol. Microbiol. **84:**795–806. http://dx.doi.org/10.1111/j.1365-2958.2012.08065.x.

8. **Ruby EG.** 1996. Lessons from a cooperative, bacterial-animal association: the *Vibrio fischeri-Euprymna scolopes* light organ symbiosis. Annu. Rev. Microbiol. **50:**591–624. http://dx.doi.org/10.1146/annurev.micro.50.1.591.
9. **Stabb EV.** 2006. The *Vibrio fischeri-Euprymna scolopes* light organ symbiosis, p 204–218. *In* Thompson FL, Austin B, Swings J (ed), The biology of vibrios. ASM Press, Washington DC.
10. **Gómez-León J, Villamil L, Lemos ML, Novoa B, Figueras A.** 2005. Isolation of *Vibrio alginolyticus* and *Vibrio splendidus* from aquacultured carpet shell clam (*Ruditapes decussatus*) larvae associated with mass mortalities. Appl. Environ. Microbiol. **71:**98–104. http://dx.doi.org/10.1128/AEM.71.1.98-104.2005.
11. **Sugumar G, Nakai T, Hirata Y, Matsubara D, Muroga K.** 1998. *Vibrio splendidus* biovar II as the causative agent in bacillary necrosis of Japanese oyster *Crassostrea gigas* larvae. Dis. Aquat. Org. **33:**111–118. http://dx.doi.org/10.3354/dao033111.
12. **Luna-González A, Maeda-Martinez AN, Sainz JC, Ascencio-Valle F.** 2002. Comparative susceptibility of veliger larvae of four bivalve mollusks to a *Vibrio alginolyticus* strain. Dis. Aquat. Org. **49:**221–226. http://dx.doi.org/10.3354/dao049221.
13. **Amaro C, Biosca EG.** 1996. *Vibrio vulnificus* biotype 2, pathogenic for eels, is also an opportunistic pathogen for humans. Appl. Environ. Microbiol. **62:**1454–1457.
14. **Tison DL, Nishibuchi M, Greenwood JD, Seidler RJ.** 1982. *Vibrio vulnificus* biogroup 2: new biogroup pathogenic for eels. Appl. Environ. Microbiol. **44:**640–646.
15. **Jones MK, Oliver JD.** 2009. *Vibrio vulnificus*: disease and pathogenesis. Infect. Immun. **77:**1723–1733. http://dx.doi.org/10.1128/IAI.01046-08.
16. **Toranzo AE, Magariños B, Romalde JL.** 2005. A review of the main bacterial fish diseases in mariculture systems. Aquaculture **246:**37–61. http://dx.doi.org/10.1016/j.aquaculture.2005.01.002.
17. **Schiewe MH, Trust TJ, Crosa JH.** 1981. *Vibrio ordalii* sp. nov.: a causative agent of vibriosis in fish. Curr. Microbiol. **6:**343–348. http://dx.doi.org/10.1007/BF01567009.
18. **Soto-Rodriguez SA, Gomez-Gil B, Lozano R.** 2010. 'Bright-red' syndrome in Pacific white shrimp *Litopenaeus vannamei* is caused by *Vibrio harveyi*. Dis. Aquat. Org. **92:**11–19. http://dx.doi.org/10.3354/dao02274.
19. **Haldar S, Chatterjee S, Sugimoto N, Das S, Chowdhury N, Hinenoyu A, Asakura M, Yamasaki S.** 2011. Identification of *Vibrio campbellii* isolated from diseased farm-shrimp from south India and establishment of its pathogenic potential in an *Artemia* model. Microbiology **157:**179–188. http://dx.doi.org/10.1099/mic.0.041475-0.
20. **Colwell RR, Brayton PR, Grimes DJ, Roszak DB, Huq SA, Palmer LM.** 1985. Viable but non-culturable *Vibrio cholerae* and related pathogens in the environment: implications for release of genetically engineered microorganisms. Nat. Biotechnol. **3:**817–820. http://dx.doi.org/10.1038/nbt0985-817.
21. **Cottingham KL, Chiavelli DA, Taylor RK.** 2003. Environmental microbe and human pathogen: the ecology and microbiology of *Vibrio cholerae*. Front. Ecol. Environ. **1:**80–86. http://dx.doi.org/10.1890/1540-9295(2003)001[0080:EMAHPT]2.0.CO;2.
22. **Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos JJ.** 2002. Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. Proc. Natl. Acad. Sci. U. S. A. **99:**1556–1561. http://dx.doi.org/10.1073/pnas.042667999.
23. **Daniels NA, MacKinnon L, Bishop R, Altekruse S, Ray B, Hammond RM, Thompson S, Wilson S, Bean NH, Griffin PM, Slutsker L.** 2000. *Vibrio parahaemolyticus* infections in the United States, 1973–1998. J. Infect. Dis. **181:**1661–1666. http://dx.doi.org/10.1086/315459.
24. **Velazquez-Roman J, León-Sicairos N, Flores-Villaseñor H, Villafaña-Rauda S, Canizalez-Roman A.** 2012. Association of pandemic *Vibrio parahaemolyticus* O3:K6 present in the coastal environment of northwest Mexico with cases of recurrent diarrhea between 2004 and 2010. Appl. Environ. Microbiol. **78:**1794–1803. http://dx.doi.org/10.1128/AEM.06953-11.
25. **Strom MS, Paranjpye RN.** 2000. Epidemiology and pathogenesis of *Vibrio vulnificus*. Microbes Infect. **2:**177–188. http://dx.doi.org/10.1016/S1286-4579(00)00270-7.
26. **Brayton PR, Bode RB, Colwell RR, MacDonell MT, Hall HL, Grimes DJ, West PA, Bryant TN.** 1986. *Vibrio cincinnatiensis* sp. nov., a new human pathogen. J. Clin. Microbiol. **23:**104–108.
27. **Igbinosa EO, Okoh AL.** 2010. *Vibrio fluvialis*: an unusual enteric pathogen of increasing public health concern. Int. J. Environ. Res. Public Health **7:**3628–3643. http://dx.doi.org/10.3390/ijerph7103628.
28. **Chakraborty R, Chakraborty S, De K, Sinha S, Mukhopadhyay AK, Khanam J, Ramamurthy T, Takeda Y, Bhattacharya SK, Nair GB.** 2005. Cytotoxic and cell vacuolating activity of *Vibrio fluvialis* isolated from paediatric patients with diarrhoea. J. Med. Microbiol. **54:**707–716. http://dx.doi.org/10.1099/jmm.0.45820-0.
29. **Chowdhury G, Pazhani GP, Dutta D, Guin S, Dutta S, Ghosh S, Izumiya H, Asakura M, Yamasaki S, Takeda Y, Arakawa E, Wantanabe H, Mukhpadhyay AK, Bhattacharya MK, Rajendran K, Nair GB, Ramamurthy T.** 2012. *Vibrio fluvialis* in patients with diarrhea, Kolkata, India. Emerg. Infect. Dis. **18:**1868–1871. http://dx.doi.org/10.3201/eid1811.120520.
30. **Brenner DJ, Hickman-Brenner FW, Lee JV, Steigerwalt AG, Fanning GR, Hollis DG, Farmer JJ, III, Weaver RE, Joseph SW, Seidler RJ.** 1983. *Vibrio furnissii* (formerly aerogenic biogroup of *Vibrio fluvialis*), a new species isolated from human feces and the environment. J. Clin. Microbiol. **18:**816–824.
31. **Linde H-J, Kobuch R, Jayasinghe S, Reischal V, Lehn N, Kaulfuss S, Beutin L.** 2004. *Vibrio metschnikovii*, a rare cause of wound infection. J. Clin. Microbiol. **42:**4909–4911. http://dx.doi.org/10.1128/JCM.42.10.4909-4911.2004.
32. **Dalsgaard A, Alarcon A, Lanata CF, Jensen T, Hansen HJ, Delgado F, Gil AI, Penny ME, Taylor D.** 1996. Clinical manifestations and molecular epidemiology of five cases of diarrhoea in children associated with *Vibrio metschnikovii* in Arequipa, Peru. J. Med. Microbiol. **45:**494–500. http://dx.doi.org/10.1099/00222615-45-6-494.
33. **Wallet F, Tachon M, Nseir S, Courcol RJ, Roussel-Delvallez M.** 2005. *Vibrio metschnikovii* pneumonia. Emerg. Infect. Dis. **11:**1641–1642. http://dx.doi.org/10.3201/eid1110.050177.
34. **Vieira VV, Teixeira LFM, Vicente ACP, Momen H, Salles CA.** 2001. Differentiation of environmental and clinical isolates of *Vibrio mimicus* from *Vibrio cholerae* by multilocus enzyme electrophoresis. Appl. Environ. Microbiol. **67:**2360–2364. http://dx.doi.org/10.1128/AEM.67.5.2360-2364.2001.
35. **Campos E, Bolaños H, Acuña MT, Diaz G, Matamoros MC, Raventós H, Sánchez LM, Sánchez O, Barquero C.** 1996. *Vibrio mimicus* diarrhea following ingestion of raw turtle eggs. Appl. Environ. Microbiol. **62:**1141–1144.
36. **Chitov T, Kirikaew P, Yungyune P, Ruengprapan N, Sontikun K.** 2009. An incidence of large foodborne outbreak associated with *Vibrio mimicus*. Eur. J. Clin. Microbiol. Infect. Dis. **28:**421–424. http://dx.doi.org/10.1007/s10096-008-0639-7.
37. **Dorsch M, Lane D, Stackebrandt E.** 1992. Towards a phylogeny of the Genus *Vibrio* based on 16S rRNA sequences. Int. J. Syst. Evol. Microbiol. **42:**58–63.
38. **Farmer JJ, III, Janda JM, Brenner FM, Cameron DN, Birkhead KM.** 2005. Genus I. Vibrio Pacini 1854, 411^{AL}, p 494–546. *In* Brenner DJ, Kreig NR, Staley JT, Garrity GM (ed), Bergey's manual of systematic bacteriology, 2nd ed, vol 2part B. Springer, New York, NY.
39. **Stackebrandt E, Goebel BM.** 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. Int. J. Syst. Evol. Microbiol. **44:**846–849.
40. **Ruimy R, Breittmayer V, Elbaze P, Lafay B, Boussemart O, Gauthier M, Christen R.** 1994. Phylogenetic analysis and assessment of the genera *Vibrio*, *Photobacterium*, *Aeromonas*, and *Plesiomonas* deduced from small-subunit rRNA sequences. Int. J. Syst. Evol. Microbiol. **44:**416–426.
41. **Thompson CC, Thompson FL, Vandemeulebroecke K, Hoste B, Dawyndt P, Swings J.** 2004. Use of *recA* as an alternative phylogenetic marker in the family *Vibrionaceae*. Int. J. Syst. Evol. Microbiol. **54:**919–924. http://dx.doi.org/10.1099/ijs.0.02963-0.
42. **Thompson FL, Gomez-Gil B, Vasconcelos ATR, Sawabe T.** 2007. Multilocus sequence analysis reveals that *Vibrio harveyi* and *V. campbellii* are distinct species. Appl. Environ. Microbiol. **73:**4279–4285. http://dx.doi.org/10.1128/AEM.00020-07.
43. **Thompson FL, Gevers D, Thompson CC, Dawyndt P, Naser S, Hoste B, Munn CB, Swings J.** 2005. Phylogeny and molecular identification of vibrios on the basis of multilocus sequence analysis. Appl. Environ. Microbiol. **71:**5107–5115. http://dx.doi.org/10.1128/AEM.71.9.5107-5115.2005.
44. **Martens M, Dawyndt P, Coopman R, Gillis M, De Vos P, Willems A.** 2008. Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including

former *Sinorhizobium*). Int. J. Syst. Evol. Microbiol. **58:**200–214. http://dx.doi.org/10.1099/ijs.0.65392-0.
45. **Nishiguchi MK, Nair VS.** 2003. Evolution of symbiosis in the *Vibrionaceae*: a combined approach using molecules and physiology. Int. J. Syst. Evol. Microbiol. **53:**2019–2026. http://dx.doi.org/10.1099/ijs.0.02792-0.
46. **Tracz DM, Backhouse PG, Olson AB, McCrea JK, Walsh JA, Ng LK, Gilmour MW.** 2007. Rapid detection of *Vibrio* species using liquid microsphere arrays and real-time PCR targeting the *ftsZ* locus. J. Med. Microbiol. **56:**56–65. http://dx.doi.org/10.1099/jmm.0.46759-0.
47. **Le Roux F, Gay M, Lambert C, Nicolas JL, Gouy M, Berthe F.** 2004. Phylogenetic study and identification of *Vibrio splendidus*-related strains based on *gyrB* gene sequences. Dis. Aquat. Org. **58:**143–150. http://dx.doi.org/10.3354/dao058143.
48. **Gevers D, Vandepoele K, Simillion C, Van de Peer Y.** 2004. Gene duplication and biased functional retention of paralogs in bacterial genomes. Trends Micriobiol. **12:**148–154. http://dx.doi.org/10.1016/j.tim.2004.02.007.
49. Reference deleted.
50. **Pascual J, Macián MC, Arahal DR, Garay E, Pujalte MJ.** 2010. Multilocus sequence analysis of the central clade of the genus *Vibrio* by using *16S rRNA*, *recA*, *pyrH*, *rpoD*, *gyrB*, *rctB* and *toxR* genes. Int. J. Syst. Evol. Microbiol. **60:**154–165. http://dx.doi.org/10.1099/ijs.0.010702-0.
51. **Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S.** 2011. MEGA5: molecular evolutionary genetics using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. **28:**2731–2739. http://dx.doi.org/10.1093/molbev/msr121.
52. **Felsenstein J.** 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39:**783–791. http://dx.doi.org/10.2307/2408678.
53. **Alfaro ME, Zoller S, Lutzoni F.** 2003. Bayes or bootstrap? A simulation study comparing the performance of Bayesian Markov chain Monte Carlo sampling and bootstrapping in assessing phylogenetic confidence. Mol. Biol. Evol. **20:**255–266. http://dx.doi.org/10.1093/molbev/msg028.
54. **Hedges SB.** 1992. The number of replications needed for accurate estimation of the bootstrap P value in phylogenetic studies. Mol. Biol. Evol. **9:**366–369.
55. **IBM Corp.** 2012. IBM SPSS statistics for Windows, version 21.0. IBM Corp, Armonk, NY.
56. **Stewart CB.** 1993. The powers and pitfalls of parsimony. Nature **361:**603–607. http://dx.doi.org/10.1038/361603a0.