

The relative sensitivity of different alignment methods and character codings in sensitivity analysis

Mark P. Simmons^{a,*}, Kai F. Müller^b and Colleen T. Webb^a

^aDepartment of Biology, Colorado State University, Fort Collins, CO 80523, USA; ^bNees-Institut für Biodiversität der Pflanzen, Rheinische Friedrich-Wilhelms-Universität Bonn, Meckenheimer Allee 170, Bonn, D-53115, Germany

Accepted 15 May 2008

Abstract

Sensitivity analysis provides a way to measure robustness of clades in sequence-based phylogenetic analyses to variation in alignment parameters rather than measuring their branch support. We compared three different approaches to multiple sequence alignment in the context of sensitivity analysis: progressive pairwise alignment, as implemented in MUSCLE; simultaneous multiple alignment of sequence fragments, as implemented in DCA; and direct optimization followed by generation of the implied alignment(s), as implemented in POY. We set out to determine the relative sensitivity of these three alignment methods using rDNA sequences and randomly generated sequences. A total of 36 parameter sets were used to create the alignments, varying the transition, transversion, and gap costs. Tree searches were performed using four different character-coding and weighting approaches: the cost function used for alignment or equally weighted parsimony with gap positions treated as missing data, separate characters, or as fifth states. POY was found to be as sensitive, or more sensitive, to variation in alignment parameters than DCA and MUSCLE for the three empirical datasets, and POY was found to be more sensitive than MUSCLE, which in turn was found to be as sensitive, or more sensitive, than DCA when applied to the randomly generated sequences when sensitivity was measured using the averaged jackknife values. When significant differences in relative sensitivity were found between the different ways of weighting character-state changes, equally weighted parsimony, for all three ways of treating gapped positions, was less sensitive than applying the same cost function used in alignment for phylogenetic analysis. When branch support is incorporated into the sensitivity criterion, our results favour the use of simultaneous alignment and progressive pairwise alignment using the similarity criterion over direct optimization followed by using the implied alignment(s) to calculate branch support.

© The Willi Hennig Society 2008.

Sensitivity analysis provides a way to measure robustness of clades in sequence-based phylogenetic analyses to variation in alignment parameters rather than measuring their branch support (Wheeler, 1995; Giribet, 2002, 2003; Goloboff et al., 2003; Giribet and Wheeler, 2007; but see Farris, 2004; Grant and Kluge, 2005), as with the bootstrap (Felsenstein, 1985) and jackknife (Farris et al., 1996) for an alignment created using a single set of alignment parameters. The less sensitive a clade is to variation in alignment parameters the better. For example, Wheeler (1995, p. 328) stated that: “If a high fraction of the total analysis space supports a group, the group is generally supported by the data

because most combinations of analytical parameters will yield that clade, especially if the areas of support are contiguous. If, however, the areas in which the clade is supported are broken up and distributed over the space, this group (however general) would be unstable because small perturbations in analysis would lead to a new result”. Giribet (2003, p. 559) asserted that “Stability under different parameters/models may well become the preferred criterion for taxonomic revision”.

In the present study, we used rDNA sequences as well as randomly generated sequences to determine the relative sensitivity (as measured by congruence of trees inferred by the alignments) of different multiple-alignment methods to variation in alignment parameters. We compared the following three approaches to global multiple sequence alignment: progressive pairwise

*Corresponding author: Fax +1 970 491 0649.
E-mail address: mark.simmons@colostate.edu

alignment (Feng and Doolittle, 1987), as implemented in MUSCLE (Edgar, 2004a,b); simultaneous multiple alignment of sequence fragments (Tönges et al., 1996), as implemented in DCA (Stoye, 1998); and direct optimization (Wheeler, 1996) followed by generation of implied alignments (Wheeler, 2003), as implemented in POY (Wheeler et al., 2003).

In progressive pairwise alignment, the most similar pair(s) of sequences is aligned first, followed by progressively less similar pairs or sets of sequences until all sequences are aligned in a single multiple alignment. Lake (1991) demonstrated that the order used in progressive pairwise alignment can undesirably determine the topology of the phylogeny inferred from the aligned sequences (see also Thorne et al., 1991; Thorne and Kishino, 1992). Lake (1991) suggested that the solution to this problem would be to restrict phylogenetic inference to using characters from those regions for which the alignment is independent of any particular progressive alignment order. Alternatively, simultaneous alignment may potentially be used to eliminate any biases that may be caused by use of any particular progressive alignment order. The alignment criterion for both progressive pairwise alignment and simultaneous alignment is similarity.

Direct optimization differs from the other methods examined in that its alignment criterion is likelihood or parsimony, rather than similarity (Wheeler, 1996, 2006). The alignment(s) that produces the tree(s) with the highest likelihood or fewest steps is favoured over alternative alignments (but note that direct optimization does not necessarily create any single multiple alignment; Wheeler, 1996). This is accomplished by integrating alignment and phylogenetic tree search into a single step, after which an implied alignment may then be produced. Note that implied alignments are equivalent to secondary, not primary, homology statements (de Pinna, 1991), in contrast to DCA and MUSCLE alignments.

Our *a priori* hypothesis on the relative sensitivity of the three alignment methods to variation in alignment parameters (when branch support is incorporated into the sensitivity measure) was that direct optimization would be more sensitive than progressive pairwise alignment, which would be more sensitive than simultaneous alignment. This hypothesis was based on the use of a single order (or in some cases two or more orders for direct optimization when equally optimal alignments are found) in which entire sequences are aligned (in the case of progressive pairwise alignment) or optimized (in the case of direct optimization), in contrast to simultaneous alignment. The use of a single tree (or subset of possible trees) to determine alignment or optimization order produces alignments from which phylogenetic trees are derived that are subject to the artifacts described by Lake (1991). Direct optimization optimizes

sequences at internal nodes. In contrast, progressive pairwise alignment relies on pairwise comparisons among each sequence in one set of sequences being aligned relative to all sequences in the other set of sequences that are being aligned (e.g., see Thompson et al., 1994, fig. 2). As such, direct optimization is more effective at minimizing the number of substitutions and indels required to change from one sequence to another on the most parsimonious tree(s). Hence, the implied alignment(s) derived from direct optimization is more closely linked to its associated tree topology (or topologies, in some cases where two or more equally optimal alignments are reported), and their bias in favour of this tree topology will be more pronounced than a progressive pairwise alignment that used the same topology to determine the pairwise alignment order. This comes into play for sensitivity analysis because the tree that is used to guide optimization or pairwise alignment is expected to vary when different costs are assigned to the various alignment parameters. Whichever tree(s) is chosen will set the course for the alignments to be biased in favour of that tree.

This hypothesis is generally consistent with results from previous studies that have conducted sensitivity analyses using different alignment methods. Terry and Whiting (2005), Pons and Vogler (2006), and Sharkey et al. (2006) (see also Shull et al., 2001; Caterino and Vogler, 2002; Ogden and Whiting, 2003) compared progressive pairwise alignments generated by Clustal (Thompson et al., 1994) with direct optimization in POY. Unfortunately, however, it is impossible to set identical alignment parameters between Clustal and POY (Ogden and Whiting, 2003). With that qualification in mind, Clustal alignments were found to be less sensitive than POY implied alignments as measured by the incongruence length difference (Mickeych and Farris, 1981) by Terry and Whiting (2005) and taxonomic congruence (Nelson, 1979) by Pons and Vogler (2006) and Sharkey et al. (2006).

A further difficulty when using taxonomic congruence to compare trees inferred from progressive pairwise alignments with trees generated by direct optimization is that character-state changes are generally weighted differently. Whereas POY implements the “logically consistent” (LC) approach advocated by Wheeler (1994; see also Giribet and Wheeler, 1999; Phillips et al., 2000) wherein the same cost function used in alignment is also used for phylogenetic analysis, both Pons and Vogler (2006) and Sharkey et al. (2006) used equally weighted parsimony to infer trees based on their Clustal alignments. Unlike Wheeler (1994), Simmons and Ochoterena (2000, pp. 369–370; see also Simmons, 2004) considered alignment and tree search to be logically independent of each other, and therefore did not consider it necessary to use the same cost function in alignment and phylogenetic analysis.

In addition to examining differences in sensitivity among alignment methods, we also tested for differences in sensitivity to alignment parameters when using different ways of coding gaps and weighting character-state changes for phylogenetic inference. This part of the study required that we use implied alignments from POY rather than relying on internal POY tree searches and jackknife analyses. We performed tree searches from aligned sequences using four approaches: (i) treating gaps as missing data, (ii) scoring gaps as separate characters (for internal gaps flanked by aligned bases on both ends) using simple indel coding (SIC; Simmons and Ochoterena, 2000) as implemented in SeqState (Müller, 2005), (iii) scoring gaps as 5th states for each character regardless of gap length (e.g. Giribet and Wheeler, 1999), and (iv) using the LC approach advocated by Wheeler (1994).

Our *a priori* hypothesis on the relative sensitivity of two of the four alternative approaches to variation in alignment parameters (when branch support is incorporated into the sensitivity measure) was that the LC approach would be more sensitive than equally weighted parsimony with gaps scored using SIC—when applied to implied alignments output by direct optimization and progressive pairwise alignments. The basis for this hypothesis was that use of the same cost function in phylogenetic analyses as that used in alignment would reinforce the biases inherent in tree-based alignment methods that are described above. We did not formulate any *a priori* hypotheses for the other two alternative approaches to variation in alignment parameters.

Materials and methods

Datasets examined

Three empirical rDNA datasets were sampled from different portions of the nuclear rDNA repeat [208 18S rDNA sequences from metazoans: G. Giribet, pers. commun., which represents expanded taxon sampling based on Giribet (2002); 746 internal transcribed spacer (ITS) sequences from the plant family Brassicaceae: Bailey et al. (2006); 206 28S rDNA sequences sampled from 194 fungi and 12 eukaryote outgroups: James et al. (2006)]. These three empirical examples were selected so as to sample different lineages, taxonomic “levels”, and types of rDNA regions. These datasets were also selected as challenging empirical examples of divergent sequences from loci at which both indels and substitutions occur, rather than straightforward-to-align loci such as *rbcL* (e.g. Källersjö et al., 1998). As rDNA sequences, all three sets of sequences need to be aligned using DNA sequences rather than amino acids.

Of the 208 metazoan 18S sequences, the following 54 sequences were deleted: 38 sequences with one or more

“N”s at the 5′ and/or 3′ ends, 15 of the remaining sequences with three or more internal “N”s as well as those with many subset ambiguities, and the *Brachionus* sequence, which was found to be identical to the *Brachionus plicatilis* sequence by using MacClade ver. 4.03 (Maddison and Maddison, 2001). There were 154 unique 18S sequences remaining.

Of the 746 ITS sequences in Bailey et al.’s (2006) “Matrix 1”, the following 201 sequences were removed: 36 sequences with missing data for the 5.8S region, followed by 37 sequences with one or more “N”, and 128 sequences that were identified as identical by MacClade. After these three steps, 545 unique ITS sequences remained.

Of the 8874 aligned positions in James et al.’s (2006) data matrix of 28S sequences, only those bases aligned from position 561 to 4735 were used in the present study. These positions were selected because they occur in conserved sequence blocks and were sampled for the majority of the sequences. Of the 206 sequences, 33 that were not completely sequenced for the region sampled were eliminated, as well as 11 sequences with three or more internal “N”s as well as those with many subset ambiguities. Of the 162 remaining sequences, none was identified as identical by MacClade.

One hundred five-sequence replicates were then created for each of the three empirical datasets. Sequences were randomly reordered using MacClade, and contiguous blocks of five sequences were then sampled based on the output order. Because 545 sequences were sampled from Bailey et al. (2006), no duplicate sequences were included in the 100 five-sequence replicates. For the 154 18S rDNA and 162 28S rDNA sequences, we attempted to minimize redundant sequence-sampling overlap among replicates by using the following strategy. Each set of sequences was randomly reordered four times in MacClade. The sequences were saved after each random reordering, and five-sequence replicates were created from these saved files until 100 replicates were obtained.

In addition to the three empirical datasets, one set of randomly generated sequences was used. Random sequences were generated using Perl scripts. Each random sequence is 1000 bp long with equal nucleotide frequencies (on average). These are the same sequences used in the eight-sequence simulations of M.P. Simmons, K.F. Müller and A.P. Norton (unpublished data), albeit with the last three sequences removed. All 100 replicates were created separately from each other. All 400 five-sequence replicates are provided as online Supporting Information at: <http://www.biology.colostate.edu/Research/>.

Replicates of five sequences each were chosen for three reasons. First, if only four sequences were sampled, gapped positions would be parsimony uninformative when gaps are treated as missing data.

In contrast, this is not necessarily the case when five sequences are sampled. Second, DCA and POY alignments are often time consuming, and running 14 400 five-sequence DCA and POY analyses was computationally tractable. Third, there are 15 possible unrooted bifurcating tree topologies for five sequences, whereas there are only three trees for four sequences and 105 trees for six sequences (Felsenstein, 1978). The greater number of alternative five-sequence topologies relative to four sequences allows more topological differences to be detected when comparing trees, without the complexity entailed by the large number of possible six-sequence topologies.

Alignments

Wheeler (1995, p. 325) used 36 alternative sets of alignment parameters, with six gap/transversion cost ratios (0.5 : 1, 1 : 1, 2 : 1, 4 : 1, 8 : 1, 16 : 1) and six transversion/transition cost ratios (0.5 : 1, 1 : 1, 2 : 1, 4 : 1, 8 : 1, 1 : 0). With few exceptions (e.g., Ogden and Whiting, 2005; Terry and Whiting, 2005), the original costs used by Wheeler (1995) encompass the range actually applied in most empirical analyses (e.g., Giribet and Wheeler, 1999; Frost et al., 2001; Giribet et al., 2002; Giribet, 2003; Aagesen, 2004; Boyer et al., 2005; Scott, 2005; Schelly et al., 2006; Sorensen et al., 2006). As such, these same costs were used in this study [but see Aagesen (2005) and Kjer et al. (2007) for other parameters that may be varied to create the “alignment space” (*sensu* Pons and Vogler, 2006)]. The order of the 36 alternative sets of alignment parameters is listed in Appendix 1. Note that these 36 alternative sets of alignment parameters did not include gap extension costs, which have been shown to improve taxonomic congruence with an accepted tree (Pons and Vogler, 2006).

Alignments of the nucleotide sequences were performed using DCA ver. 1.1, MUSCLE ver. 3.6, and POY ver. 3.0.11. MUSCLE was used instead of Clustal, which has been frequently used in comparisons between progressive pairwise alignment and direct optimization, because it is impossible to set identical alignment parameters between Clustal and POY. Identical alignment parameters, to the degree possible, were specified for DCA, MUSCLE, and POY for each of the 36 alternative sets of alignment parameters. The programs differ in how they treat terminal (i.e. leading and trailing) gaps. In DCA, terminal gaps were assigned a cost of zero (although there must be some heuristic method used to prevent trivial alignments), following the program’s default settings. In MUSCLE, terminal gaps were assigned half the cost of internal gaps, following the program’s default settings (Edgar, 2004b, p. 7). In POY, terminal gaps were treated using the “-noleading” command, wherein “... leading and trailing gaps are counted and accounted for during builds

and refinements (to prevent trivial alignments) but discounted when determining tree length” (De Laet and Wheeler, 2003, p. 39). POY searches were performed using 100 replicates, with 50 builds per replicate and a maximum of ten trees held.

To lower the potential confounding effect of differential treatment of terminal gaps by the different alignment methods, blocks of invariant bases were added to the 5’ and 3’ ends of all sequences following Simmons and Freudenstein (2003, p. 446). The blocks inserted at each end of the sequences were ten repeats of “ACGT”. Despite these blocks of invariant bases, errors in DCA and MUSCLE alignments were observed in both leading (DCA: 1324 total, 1.8%; MUSCLE: 1825 total, 2.5%; POY 3787 total, 5.0%) and trailing (DCA: 1041 total, 1.4%; MUSCLE: 599 total, 0.8%; POY: 0 total, 0%) blocks of invariant bases.

In DCA, costs between ambiguous entries (e.g., “R” for “A” or “G”) are set separately. Because there are ambiguities in some replicates from the empirical matrices, these costs needed to be specified in the DCA cost matrices. The costs were set to match the lowest cost assigned for transitions or transversions for each of the 36 parameter sets separately from each other. All 36 of the DCA, MUSCLE, and POY cost matrices are provided as online Supporting information, as are all of the alignments.

Tree searches

Aligned sequences were analysed by using the four alternative approaches described above: (i) treating gaps as missing data, (ii) scoring gaps as separate characters (for internal gaps flanked by aligned bases on both ends) using SIC, (iii) scoring gaps as 5th states for each character regardless of gap length, and (iv) using the LC approach advocated by Wheeler (1994) wherein the same cost function used in alignment is also used for phylogenetic analysis. The step matrices (Sankoff and Rousseau, 1975) used in PAUP* for the LC approach are reported as online Supporting information. Note that the first parameter set (gap cost: 1, transversion cost: 2, transition cost: 4) violates the triangle inequality (Wheeler, 1993) because the gap cost is less than half the transition cost. This violation of the triangle inequality was automatically adjusted by PAUP* (Swofford, 2001), which decreases the transformation cost that violates the triangle inequality to the point where it is no longer violated. Implied alignments were output by POY for tree searches conducted by PAUP*. By using the LC approach, the same most parsimonious tree topology (or topologies) should be inferred by both POY and PAUP* (heuristics aside; Wheeler, 2003).

Tree searches were performed using parsimony in PAUP* ver. 4.0b10. All most parsimonious tree(s) were found using an exhaustive search, and the strict

consensus was calculated. Also, 50%-deletion jackknife analyses were performed using 10 000 jackknife replicates with branch-and-bound searches within each replicate; 50, 75, 88, 94, 97, 98, 99, and 100% majority rule jackknife trees were then calculated.

Unlike DCA and MUSCLE, POY can report multiple equally optimal alignments. Two, three, or four equally optimal implied alignments were reported by POY for 552 (3.8%) of the 14 400 separate analyses performed. In these 552 cases, tree searches were performed on all equally optimal alignments independently of one another. If congruent tree topologies and identical jackknife support values for both clades (when applicable) were obtained, then no changes were necessary. If congruent tree topologies but different jackknife support values for one or both clades (when applicable) were obtained, then the lowest jackknife value(s) were used. If totally or partially incongruent tree topologies were inferred, the strict consensus (i.e., conflicting clades were unresolved) was used.

Quantification of results

Among the 15 possible unrooted bifurcating tree topologies for five sequences, there are ten possible pairs of clades that can be resolved. These are: (i) ((1,2,3)(4,5)), (ii) ((1,2,4)(3,5)), (iii) ((1,3,4)(2,5)), (iv) ((1,4,5)(2,3)), (v) ((1,3,5)(2,4)), (vi) ((1,2,5)(3,4)), (vii) ((1,2)(3,4,5)), (viii) ((1,3)(2,4,5)), (ix) ((1,4)(2,3,5)), and (x) ((1,5)(2,3,4)). The presence or absence of each of these ten clades (after arbitrarily rooting with the first sequence) was determined in each PAUP* tree by using PEST ver. 2.2 (Zujko-Miller and Miller, 2003) with each of the ten partially resolved reference trees in turn.

All possible pairwise comparisons were then performed among the most parsimonious trees (or the strict consensus trees, when applicable) inferred from the 36 alignments for each of the 100 five-sequence replicates. Differences among topologies were calculated using Hamming (1950) distance. For example, if two fully resolved trees share no clades in common, the Hamming distance is four because they differ for four of the ten possible clades. If a fully resolved tree is compared with an unresolved bush, the Hamming distance is two because they differ for two of the ten possible clades. If two unresolved bushes are compared with each other, the Hamming distance is 0. Note that the Hamming distance is a conservative measure of taxonomic congruence when applied to unresolved bushes given that the same strict consensus tree may be formed by non-overlapping sets of most parsimonious trees. The Hamming distance between the most parsimonious trees inferred for each of the 36 alignments was used to measure taxonomic congruence irrespective of branch support.

The averaged support for each clade resolved was calculated by using the jackknife trees. By using a 50%-deletion jackknife analysis and averaging across the eight different (50–100%) output jackknife trees, the amount of support for each resolved clade was scaled to that provided by one to eight uncontradicted synapomorphies (Simmons and Webb, 2006). Instead of 0 versus 0, 0 versus 1, or 1 versus 1 comparisons for each of the ten possible clades as with the Hamming distance, the range of support for any given clade ranged from 0 to 1 in increments of 0.125. For example, the distance between a tree with the equivalent of two uncontradicted synapomorphies of support for a single clade of (4,5) and a tree with the equivalent of four uncontradicted synapomorphies of support for a single clade of (3,5) would be: $(0.25 - 0)^2 + (0.5 - 0)^2 = 0.3125$. This measure is essentially Euclidean distance, which assigns lower sensitivities to more, but smaller differences than it does to fewer, larger differences. In using this approach, character-based congruence (Mickevich and Farris, 1981) is integrated with taxonomic congruence by quantifying the amount of resolution and branch support.

Statistical analyses

A multivariate analysis of variance (ANOVA) approach was used to test our hypotheses, and models were implemented using JMP IN ver. 5.1 (SAS Institute, Inc., Cary, NC). Two sets of ANOVAs were used. For the first set of models, the response variable was either sensitivity measured using the average Hamming distance (most parsimonious trees) across the 36 parameter sets or sensitivity measured using the average Euclidean-like distance (averaged jackknife trees) across the 36 parameter sets. The independent variables were: (i) dataset, a nominal categorical variable indicating which dataset was used (i.e. 18S, ITS, 26S rDNA, or random sequences); (ii) character coding, a nominal categorical variable indicating which way gapped positions were treated and how character-state changes were weighted (i.e. treating gaps as missing data, scoring gaps as separate characters using SIC, scoring gaps as 5th states, and using the LC approach to character-state weighting); and (iii) method, a nominal categorical variable indicating which alignment method was used [i.e. simultaneous alignment as implemented in DCA, progressive pairwise alignment as implemented in MUSCLE, or direct optimization followed by generation of implied alignment(s) in POY]. All pairwise interactions and the three-way interaction were included. All independent variables were treated as fixed effects.

For the second set of models, we used only the random-sequence dataset. This second set of models was used to test the generality of Simmons et al.'s (unpublished data) results on the amount of resolution and

support obtained from DCA, MUSCLE, and POY alignments when applied to randomly generated sequences. The response variable was the amount of resolution, based on the most parsimonious trees or averaged jackknife trees. The independent variables included character coding and method, as described above. An additional independent variable was parameter, a nominal categorical variable indicating which alignment parameter set (1–36) was used. All pairwise interactions and the three-way interaction were included. All independent variables were treated as fixed effects.

For all of our reported analyses, the basic assumptions of parametric statistics were met. Residuals were significantly non-normally distributed, but this is common when working with large datasets. The residual distribution was unimodal, and ANOVA is quite robust to the assumption of normally distributed residuals when they are unimodally distributed (Kutner et al., 2004), indicating that ANOVA using the untransformed data was appropriate. Least squares mean estimates of the independent variables were obtained, and independent contrasts on the least squares means were performed where needed. Groups of analyses with multiple tests were Bonferroni-corrected in order to control for spurious significant results that could be the result of large numbers of comparisons.

Results

The relative sensitivity of each of the alignment methods to variation in alignment parameters as quantified using least squares means from the most parsimonious trees is shown in Table 1, while corresponding results based on the jackknife trees are presented in Table 2. Regardless of how gaps were treated and character-state changes were weighted, no significant differences among adjacent pairs were found among the alignment methods for any of the four datasets when sensitivity was measured using the most parsimonious trees (Table 1). In contrast, many significant differences among alignment methods were observed when sensitivity was measured using the averaged jackknife values (Table 2). For the three empirical datasets, POY was found to be as sensitive, or more sensitive, to variation in alignment parameters than DCA and MUSCLE. No significant differences were found between DCA and MUSCLE. For the random data, POY was found to be more sensitive than MUSCLE, which in turn was found to be as sensitive as, or more sensitive than, DCA.

The relative sensitivity of each of the different ways of treating gapped positions and weighting character-state changes to variation in alignment parameters as quantified using least squares means from the most

Table 1

The relative sensitivity of each of the alignment methods to variation in alignment parameters as quantified using least squares means from the most parsimonious trees (no significant differences were found between adjacent pairs)

	Bailey et al. (2006) ITS	Giribet 18S	James et al. (2006) 28S	Random
Gaps = missing				
DCA	0.073	0.101	0.110	0.271
MUSCLE	0.061	0.088	0.101	0.266
POY implied	0.055	0.099	0.094	0.278
Gaps = SIC				
DCA	0.056	0.087	0.070	0.284
MUSCLE	0.056	0.093	0.072	0.261
POY implied	0.060	0.104	0.079	0.279
Gaps = 5th states				
DCA	0.049	0.084	0.067	0.282
MUSCLE	0.053	0.089	0.071	0.260
POY implied	0.057	0.102	0.080	0.279
Log. consistent				
DCA	0.085	0.123	0.084	0.290
MUSCLE	0.088	0.134	0.083	0.257
POY implied	0.090	0.128	0.091	0.283

parsimonious trees is shown in Table 3, while corresponding results based on averaged jackknife trees are presented in Table 4. No significant differences among different ways of treating gaps with equally weighted parsimony were detected using least squares means from the most parsimonious trees (Table 3). In the ITS and 18S rDNA datasets using the MUSCLE alignments and the ITS dataset using the POY implied alignments, the LC approach of using the same cost function in both alignment and phylogenetic analysis was found to be more sensitive to variation in alignment parameters than any of the three ways of treating gapped positions with equally weighted parsimony. No other significant differences between adjacent pairs were observed.

When sensitivity was measured using the averaged jackknife values, no significant differences were observed among the different ways of treating gapped positions and weighting character-state changes to variation in alignment parameters using the DCA or MUSCLE alignments for the ITS, 28S rDNA, or random-sequence datasets (Table 4). For the 18S rDNA dataset using the DCA and MUSCLE alignments, the LC approach to weighting character-state changes was found to be more sensitive to variation in alignment parameters than any of the three ways of treating gapped positions with equally weighted parsimony. This finding also applied to all four datasets when the POY implied alignments were used. Using the POY implied alignments, coding gaps as 5th states or using SIC were found to be as sensitive (ITS), or more sensitive (18S, 28S rDNA, random sequences), to variation in

Table 2

The relative sensitivity of each of the alignment methods to variation in alignment parameters as quantified using least squared means of the averaged jackknife values

	Bailey et al. (2006) ITS	Giribet 18S	James et al. (2006) 28S	Random
Gaps = missing				
DCA	0.017	0.020* 2nd	0.037	0.005 2nd
MUSCLE	0.012	0.016 2nd	0.037	0.012* 2nd
POY implied	0.014	0.037 1st	0.047	0.126 1st
Gaps = SIC				
DCA	0.010	0.013 2nd	0.031 2nd	0.005* 3rd
MUSCLE	0.011	0.020* 2nd	0.036* 2nd	0.018* 2nd
POY implied	0.021	0.060 1st	0.059 1st	0.169 1st
Gaps = 5th states				
DCA	0.009 2nd	0.012 2nd	0.029 2nd	0.005* 3rd
MUSCLE	0.010† 2nd	0.018* 2nd	0.032* 2nd	0.021* 2nd
POY implied	0.023 1st	0.059 1st	0.061 1st	0.163 1st
Log. consistent				
DCA	0.019* 2nd	0.032 2nd	0.045* 2nd	0.005* 3rd
MUSCLE	0.018 2nd	0.036* 2nd	0.045 2nd	0.026* 2nd
POY implied	0.037 1st	0.101 1st	0.078 1st	0.209 1st

*Significant at $P = 0.01$ after Bonferroni correction relative to the next highest mean.

†Significant at $P = 0.05$ after Bonferroni correction relative to the next highest mean.

Table 3

The relative sensitivity of each of the different ways of treating gapped positions and weighting character-state changes to variation in alignment parameters as quantified using least squares means from the most parsimonious trees

	Bailey et al. (2006) ITS	Giribet 18S	James et al. (2006) 28S	Random
DCA				
Missing	0.073	0.101	0.110	0.271
SIC	0.056	0.087	0.070	0.284
5th states	0.049	0.084	0.067	0.282
Log. consistent	0.085	0.123	0.084	0.290
MUSCLE				
Missing	0.061† 2nd	0.088 2nd	0.101	0.266
SIC	0.056 2nd	0.093* 2nd	0.072	0.261
5th states	0.053 2nd	0.089 2nd	0.071	0.260
Log. consistent	0.088 1st	0.134 1st	0.083	0.257
POY implied				
Missing	0.055 2nd	0.099	0.094	0.278
SIC	0.060† 2nd	0.104	0.079	0.279
5th states	0.057 2nd	0.102	0.080	0.279
Log. consistent	0.090 1st	0.128	0.091	0.283

*Significant at $P = 0.01$ after Bonferroni correction relative to the next highest mean.

†Significant at $P = 0.05$ after Bonferroni correction relative to the next highest mean.

alignment parameters than treating gapped positions as missing data.

Least squares means among alignment methods differed in the number of clades resolved on most parsimonious trees and the averaged number of clades resolved on jackknife trees taken across the 36 sets of alignment parameters for the randomly generated sequences and are presented in Table 5 and supplementary Figs S1–S5. Based on the most parsimonious trees (or the strict consensus trees, when applicable), MUSCLE was found to provide as much,

or significantly more, resolution than POY and DCA. Based on the averaged number of clades resolved on the jackknife trees, POY implied alignments were found to provide more supported resolution than MUSCLE, which in turn was found to provide more supported resolution than DCA, regardless of how gaps were treated or character-state changes were weighted.

Note that parameter cost sets 1 and 2, wherein the cost of an internal gap was less than or equal to half the cost of either a transversion or a transition,

Table 4

The relative sensitivity of each of the different ways of treating gapped positions and weighting character-state changes to variation in alignment parameters as quantified using least squares means of the averaged jackknife values

	Bailey et al. (2006) ITS	Giribet 18S	James et al. (2006) 28S	Random
DCA				
Missing	0.017	0.020† 2nd	0.037	0.005
SIC	0.010	0.013 2nd	0.031	0.005
5th states	0.009	0.012 2nd	0.029	0.005
Log. consistent	0.019	0.032 1st	0.045	0.005
MUSCLE				
Missing	0.012	0.016 2nd	0.037	0.012
SIC	0.011	0.020* 2nd	0.036	0.018
5th states	0.010	0.018 2nd	0.032	0.021
Log. consistent	0.018	0.036 1st	0.045	0.026
POY implied				
Missing	0.014 2nd	0.037* 3rd	0.047† 3rd	0.126* 3rd
SIC	0.021 2nd	0.060* 2nd	0.059 2nd	0.169* 2nd
5th states	0.023* 2nd	0.059 2nd	0.061* 2nd	0.163 2nd
Log. consistent	0.037 1st	0.101 1st	0.078 1st	0.209 1st

*Significant at $P = 0.01$ after Bonferroni correction relative to the next highest mean.

†Significant at $P = 0.05$ after Bonferroni correction relative to the next highest mean.

Table 5

Least squares means among alignment methods in the number of clades resolved on most parsimonious trees and the averaged number of clades resolved on averaged jackknife trees taken across the 36 sets of alignment parameters for the randomly generated sequences

	No. of clades	Averaged no. of clades
Gaps = missing		
DCA	1.756* 2nd	0.147* 3rd
MUSCLE	1.853 1st	0.235* 2nd
POY implied	1.861 1st	0.949 1st
Gaps = SIC		
DCA	1.868 2nd	0.155* 3rd
MUSCLE	1.901 1st	0.301* 2nd
POY implied	1.875† 2nd	1.174 1st
Gaps = 5th states		
DCA	1.872	0.159* 3rd
MUSCLE	1.895	0.318* 2nd
POY implied	1.874	1.139 1st
Log. consistent		
DCA	1.909* 2nd	0.165* 3rd
MUSCLE	1.943 1st	0.376* 2nd
POY implied	1.894 2nd	1.431 1st

*Significant at $P = 0.01$ after Bonferroni correction relative to the next highest mean.

†Significant at $P = 0.05$ after Bonferroni correction relative to the next highest mean.

resulted in trivial DCA alignments such that no substitutions were inferred. These trivial alignments account for the lack of any resolution observed in Figs S1 and S5 for those two parameter cost sets when gaps were treated as missing data. In contrast, non-trivial alignments were recovered by

MUSCLE and POY for both of these parameter cost sets.

Discussion

Our results generally corroborated our *a priori* hypotheses. Our first hypothesis was that direct optimization (POY) would be more sensitive than progressive pairwise alignment (MUSCLE), which would be more sensitive than simultaneous alignment (DCA), when branch support is incorporated into the sensitivity measure. No significant differences among adjacent pairs were found among the alignment methods for any of the four datasets when sensitivity was measured using the most parsimonious trees (Table 1). However, POY implied alignments were found to be as sensitive, or more sensitive, to variation in alignment parameters than DCA and MUSCLE for the three empirical datasets, and POY implied alignments were found to be more sensitive than MUSCLE alignments, which in turn were found to be as sensitive as, or more sensitive than, DCA alignments when applied to the randomly generated sequences when sensitivity was measured using the averaged jackknife values (Table 2).

Our second hypothesis was that Wheeler's (1994) LC approach, wherein the same cost function used in alignment is also used for phylogenetic analysis, would be more sensitive (when branch support is incorporated into the sensitivity measure) than equally weighted parsimony with gaps scored using SIC, when applied to implied alignments output by direct optimization and progressive pairwise alignments. Indeed, the LC approach to character-state-change

weighting was found to be significantly more sensitive than equally weighted parsimony with gaps scored using SIC for all four datasets when POY implied alignments were used and branch support was incorporated into the sensitivity measure (Table 4). This difference in sensitivity was less pronounced when applied to the MUSCLE alignments, with a significant difference between the LC and SIC approaches in only one (18S rDNA) of the four datasets. Unexpectedly, this same result for the 18S rDNA dataset also held for the DCA alignments.

When significant differences in relative sensitivity were found between the different ways of weighting character-state changes, equally weighted parsimony, for all three ways of treating gapped positions, was less sensitive than applying the same cost function used in alignment for phylogenetic analysis, regardless of whether sensitivity was quantified using least squares means of the most parsimonious trees (Table 3) or the averaged jackknife values (Table 4).

Based on the differences in least squares means among alignment methods in the number of clades resolved on most parsimonious trees and the averaged number of clades resolved on jackknife trees taken across the 36 sets of alignment parameters for the randomly generated sequences, the results of Simmons et al. (unpublished data) do appear to be generalizable, rather than being specific to the alignment parameters that they used. In our results, based on the averaged number of clades resolved on the jackknife trees, POY implied alignments were found to provide more supported resolution than MUSCLE alignments, which in turn was found to provide more supported resolution than DCA alignments, regardless of how gaps were treated or character-state changes were weighted.

The results based on averaged jackknife values are specific to implied alignments rather than internal POY bootstrap or jackknife analyses using dynamic homology, wherein the alignment of different regions may vary among replicates. However, implied alignments have been used to calculate branch support for many empirical datasets. Implied alignments from direct optimization have been used to conduct jackknife analyses (e.g., Smith and Wheeler, 2004, 2006; Faivovich et al., 2005; Stiassny and Sparks, 2006; Smith and Craig, 2007), conduct bootstrap analyses (e.g., Jarvis et al., 2005; Worsaae, 2005; Gómez-Zurita et al., 2006), conduct Bremer-support analyses (e.g., Sparks et al., 2005; Frost et al., 2006; Grant et al., 2006; Schelly et al., 2006), and calculate posterior probabilities in Bayesian MCMC analyses (e.g., Giribet et al., 2006; Sorensen and Giribet, 2006; Whiting et al., 2006; Schulze et al., 2007). Of those 16 papers, only Jarvis et al. (2005) and Frost et al. (2006) noted that using implied alignments may lead to higher branch-support values than expected using

different alignment/tree-search methods or dynamic homology, respectively.

As noted in the Introduction, sensitivity analysis provides a way to measure robustness of clades in sequence-based phylogenetic analyses to variation in alignment parameters. The less sensitive a clade is to variation in alignment parameters the better. So as to help avoid "... assumption-specific unstable conclusions ..." (Wheeler, 1995, p. 331), all else being equal, a method that infers divergent trees and/or support values based on the choice of parameters selected is less desirable than an alternative method that does not show this behaviour to the same degree. When branch support is incorporated into the sensitivity criterion, our results favour the use of simultaneous alignment and progressive pairwise alignment using the similarity criterion over direct optimization followed by using the implied alignment(s) to calculate branch support. The generality of our results based on limited taxon and character sampling may be tested by empirical and/or simulation studies with greater taxon and/or character sampling.

Acknowledgements

We thank Lone Aagesen, an anonymous reviewer, and the Associate Editor for constructive criticisms that helped to improve the manuscript; Bob Edgar, Kurt Pickett, and Jens Stoye for clarifications regarding MUSCLE, POY, and DCA, respectively; Donovan Bailey for assistance with the Bailey et al. (2006) sequences; Gonzalo Giribet for permission to use his 18S rDNA sequences; Ross Madden for assistance with Condor; John Freudenstein, Chris Randle, and Pat Reeves for helpful discussions; and the Deutsche Telekom Stiftung for a scholarship to K.F.M.

References

- Aagesen, L., 2004. The information content of an ambiguously alignable region, a case study of the *trnL* intron from the Rhamnaceae. *Org. Divers. Evol.* 4, 35–49.
- Aagesen, L., 2005. Direct optimization, affine gap costs, and node stability. *Mol. Phylogenet. Evol.* 36, 641–653.
- Bailey, C.D., Koch, M.A., Mayer, M., Mummenhoff, K., O’Kane, S.L., Warwick, S.I., Windham, M.D., Al-Shehbaz, I.A., 2006. Toward a global phylogeny of the Brassicaceae. *Mol. Biol. Evol.* 23, 2142–2160.
- Boyer, S.L., Karaman, I., Giribet, G., 2005. The genus *Cyphophthalmus* (Arachnida, Opiliones, Cyphophthalmi) in Europe: a phylogenetic approach to Balkan Peninsula biogeography. *Mol. Phylogenet. Evol.* 36, 554–567.
- Caterino, M.S., Vogler, A.P., 2002. The phylogeny of the Histeroidea (Coleoptera: Staphyliniformia). *Cladistics* 18, 394–415.
- De Laet, J., Wheeler, W., 2003. POY Version 3.0.11 Command Line Documentation. distributed by the authors, New York.

- Edgar, R.C., 2004a. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 22, 1792–1797.
- Edgar, R.C., 2004b. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* 5, 113.
- Faivovich, J., Haddad, C.F.B., Garcia, P.C.A., Frost, D.R., Campbell, J.A., Wheeler, W.C., 2005. Systematic review of the frog family Hylidae, with special reference to Hylinae: phylogenetic analysis and taxonomic revision. *Bull. Am. Mus. Nat. Hist.* 294, 6–228.
- Farris, J.S., 2004. Sensitivity analysis: does it measure robustness? *Cladistics* 20, 82.
- Farris, J.S., Albert, V.A., Källersjö, M., Lipscomb, D., Kluge, A.G., 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12, 99–124.
- Felsenstein, J., 1978. The number of evolutionary trees. *Syst. Zool.* 27, 27–33.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.
- Feng, D.-F., Doolittle, R.F., 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* 25, 351–360.
- Frost, D.R., Etheridge, R., Janies, D., Titus, T.A., 2001. Total evidence, sequence alignment, evolution of polychrotid lizards, and a reclassification of the Iguania (Squamata: Iguania). *Am. Mus. Novit.* 3343, 1–38.
- Frost, D.R., Grant, T., Faivovich, J., Bain, R.H., Haas, A., Haddad, C.F.B., de Sá, R.O., Channing, A., Wilkinson, M., Donnellan, S.C., Raxworthy, C.J., Campbell, J.A., Blotto, B.L., Moler, P., Drewes, R.C., Nussbaum, R.A., Lynch, J.D., Green, D.M., Wheeler, W.C., 2006. The amphibian tree of life. *Bull. Am. Mus. Nat. Hist.* 297, 1–370.
- Giribet, G., 2002. Relationships among metazoan phyla as inferred from 18S rRNA sequence data: a methodological approach. In: DeSalle, R., Giribet, G., Wheeler, W. (Eds.), *Molecular Systematics and Evolution: Theory and Practice*. Birkhäuser Verlag, Basel, pp. 85–101.
- Giribet, G., 2003. Stability in phylogenetic formulations and its relationship to nodal support. *Syst. Biol.* 52, 554–564.
- Giribet, G., Wheeler, W.C., 1999. On gaps. *Mol. Phylogenet. Evol.* 13, 132–143.
- Giribet, G., Wheeler, W.C., 2007. The case for sensitivity: a response to Grant and Kluge. *Cladistics* 23, 294–296.
- Giribet, G., Edgecombe, G.D., Wheeler, W.C., Babbitt, C., 2002. Phylogeny and systematic position of Opiliones: a combined analysis of chelicerate relationships using morphological and molecular data. *Cladistics* 18, 5–70.
- Giribet, G., Okusu, A., Lindgren, A.R., Huff, S.W., Schrödl, M., Nishiguchi, M.K., 2006. Evidence for a clade composed of molluscs with serially repeated structures: monoplacophorans are related to chitons. *Proc. Natl Acad. Sci. U.S.A.* 103, 7723–7728.
- Goloboff, P.A., Farris, J.S., Källersjö, M., Oxelman, B., Ramírez, M.J., Szumik, C.A., 2003. Improvements to resampling measures of group support. *Cladistics* 19, 324–332.
- Gómez-Zurita, J., Funk, D.J., Vogler, A.P., 2006. The evolution of unisexuality in *Calligrapha* leaf beetles: molecular and ecological insights on multiple origins via interspecific hybridization. *Evolution* 60, 328–347.
- Grant, T., Kluge, A.G., 2005. Stability, sensitivity, science and heuristic. *Cladistics* 21, 597–604.
- Grant, T., Frost, D.R., Caldwell, J.P., Gagliardo, R., Haddad, C.F.B., Kok, P.J.R., Means, D.B., Noonan, B.P., Schargel, W.E., Wheeler, W.C., 2006. Phylogenetic systematics of dart-poison frogs and their relatives (Amphibia: Athesphatanura: Dendrobatidae). *Bull. Am. Mus. Nat. Hist.* 299, 6–292.
- Hamming, R.W., 1950. Error detecting and error correcting codes. *AT&T Tech. J.* 26, 147–160.
- James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J. et al., 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443, 818–822.
- Jarvis, K.J., Haas, F., Whiting, M.F., 2005. Phylogeny of earwigs (Insecta: Dermaptera) based on molecular and morphological evidence: reconsidering the classification of Dermaptera. *Syst. Entomol.* 30, 442–453.
- Källersjö, M., Farris, J.S., Chase, M.W., Bremer, B., Fay, M.F., Humphries, C.J., Petersen, G., Seberg, O., Bremer, K., 1998. Simultaneous parsimony jackknife analysis of 2538 *rbcL* DNA sequences reveals support for major clades of green plants, land plants, seed plants, and flowering plants. *Plant Syst. Evol.* 213, 259–287.
- Kjer, K.M., Gillespie, J.J., Ober, K.A., 2007. Opinions on multiple sequence alignment, and an empirical comparison of repeatability and accuracy between POY and structural alignment. *Syst. Biol.* 56, 133–146.
- Kutner, M.H., Nachsheim, C.J., Neter, J., Li, W., 2004. *Applied Linear Statistical Models*, 5th edn. McGraw Hill, Columbus, OH.
- Lake, J.A., 1991. The order of sequence alignment can bias the selection of tree topology. *Mol. Biol. Evol.* 8, 378–385.
- Maddison, D.R., Maddison, W.P., 2001. *MacClade: Analysis of Phylogeny and Character Evolution*, Version 4.03. Sinauer Associates, Sunderland, MA.
- Mickevich, M.F., Farris, J.S., 1981. The implications of congruence in *Menidia*. *Syst. Zool.* 30, 351–370.
- Müller, K., 2005. SeqState – primer design and sequence statistics for phylogenetic DNA data sets. *Appl. Bioinformatics* 4, 65–69.
- Nelson, G., 1979. Cladistic analysis and synthesis: principles and definitions, with a historical note on Adanson's *Familles des plantes* (1763–1764). *Syst. Zool.* 28, 1–21.
- Ogden, T.H., Whiting, M.F., 2003. The problem with “the Paleoptera Problem:” sense and sensitivity. *Cladistics* 19, 432–442.
- Ogden, T.H., Whiting, M.F., 2005. Phylogeny of Ephemeroptera (mayflies) based on molecular evidence. *Mol. Phylogenet. Evol.* 37, 625–643.
- Phillips, A., Janies, D., Wheeler, W., 2000. Multiple sequence alignment in phylogenetic analysis. *Mol. Phylogenet. Evol.* 16, 317–330.
- de Pinna, M.C.C., 1991. Concepts and tests of homology in the cladistic paradigm. *Cladistics* 7, 367–394.
- Pons, J., Vogler, A.P., 2006. Size, frequency, and phylogenetic signal of multiple-residue indels in sequence alignment of introns. *Cladistics* 22, 144–156.
- Sankoff, D., Rousseau, P., 1975. Locating the vertices of a steiner tree in an arbitrary metric space. *Math. Program.* 9, 240–246.
- Shelly, R., Salzburger, W., Koblmüller, S., Duftner, N., Sturmbauer, C., 2006. Phylogenetic relationships of the lamprologine cichlid genus *Lepidiolamprologus* (Teleostei: Perciformes) based on mitochondrial and nuclear sequences, suggesting introgressive hybridization. *Mol. Phylogenet. Evol.* 38, 426–438.
- Schulze, A., Cutler, E.B., Giribet, G., 2007. Phylogeny of sipunculan worms: a combined analysis of four gene regions and morphology. *Mol. Phylogenet. Evol.* 42, 171–192.
- Scott, E., 2005. A phylogeny of ranid frogs (Anura: Ranoidea: Ranidae), based on a simultaneous analysis of morphological and molecular data. *Cladistics* 21, 507–574.
- Sharkey, M.J., Laurence, N.M., Sharanowski, B., Quicke, D.L.J., Murray, D., 2006. Revision of the Agathidinae (Hymenoptera: Braconidae) with comparisons of static and dynamic alignments. *Cladistics* 22, 546–567.
- Shull, V.L., Vogler, A.P., Baker, M.D., Maddison, D.R., Hammond, P.M., 2001. Sequence alignment of 18S ribosomal RNA and the basal relationships of adephagan beetles: evidence for monophyly

- of aquatic families and the placement of Trachypachidae. *Syst. Biol.* 50, 945–969.
- Simmons, M.P., 2004. Independence of alignment and tree search. *Mol. Phylogenet. Evol.* 31, 874–879.
- Simmons, M.P., Freudenstein, J.V., 2003. The effects of increasing genetic distance on alignment of, and tree construction from, rDNA internal transcribed spacer sequences. *Mol. Phylogenet. Evol.* 26, 444–451.
- Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. *Syst. Biol.* 49, 369–381.
- Simmons, M.P., Webb, C.T., 2006. Quantification of the success of phylogenetic inference in simulations. *Cladistics* 22, 249–255.
- Smith, W.L., Craig, M.T., 2007. Casting the percomorph net widely: the importance of broad taxonomic sampling in the search for the placement of serranid and percid fishes. *Copeia* 2007, 35–55.
- Smith, W.L., Wheeler, W.C., 2004. Polyphyly of the mail-cheeked fishes (Teleostei: Scorpaeniformes): evidence from mitochondrial and nuclear sequence data. *Mol. Phylogenet. Evol.* 32, 627–646.
- Smith, W.L., Wheeler, W.C., 2006. Venom evolution widespread in fishes: a phylogenetic road map for the bioprospecting of piscine venoms. *J. Hered.* 97, 206–217.
- Sorensen, M.V., Giribet, G., 2006. A modern approach to rotiferan phylogeny: combining morphological and molecular data. *Mol. Phylogenet. Evol.* 40, 585–608.
- Sorensen, M.V., Sterrer, W., Giribet, G., 2006. Gnathostomulid phylogeny inferred from a combined approach of four molecular loci and morphology. *Cladistics* 22, 32–58.
- Sparks, J.S., Dunlap, P.V., Smith, W.L., 2005. Evolution and diversification of a sexually dimorphic luminescent system in ponyfishes (Teleostei: Leiognathidae), including diagnoses for two new genera. *Cladistics* 21, 305–327.
- Stiassny, M.L.J., Sparks, J.S., 2006. Phylogeny and taxonomic revision of the endemic Malagasy genus *Ptychochromis* (Teleostei: Cichlidae), with the description of five new species and a diagnosis for *Katria*, new genus. *Am. Mus. Novit.* 3535, 1–55.
- Stoye, J., 1998. Multiple sequence alignment with the divide-and-conquer method. *Gene* 211, GC45–GC56.
- Swofford, D.L., 2001. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods). Sinauer Associates, Sunderland, MA.
- Terry, M.D., Whiting, M.F., 2005. Comparison of two alignment techniques within a single complex data set: POY versus Clustal. *Cladistics* 21, 272–281.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.
- Thorne, J.L., Kishino, H., 1992. Freeing phylogenies from artifacts of alignment. *Mol. Biol. Evol.* 9, 1148–1162.
- Thorne, J.L., Kishino, H., Felsenstein, J., 1991. An evolutionary model for maximum likelihood alignment of DNA sequences. *J. Mol. Evol.* 33, 114–124.
- Tönges, U., Perry, S.W., Stoye, J., Dress, A.W.M., 1996. A general method for fast multiple sequence alignment. *Gene* 172, GC33–GC41.
- Wheeler, W.C., 1993. The triangle inequality and character analysis. *Mol. Biol. Evol.* 10, 707–712.
- Wheeler, W.C., 1994. Sources of ambiguity in nucleic acid sequence alignment. In: Schierwater, B., Wagner, G.P., DeSalle, R. (Eds), *Molecular Ecology and Evolution: Approaches and Applications*. Birkhäuser Verlag, Basel, pp. 323–352.
- Wheeler, W.C., 1995. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. *Syst. Biol.* 44, 321–331.
- Wheeler, W., 1996. Optimization alignment: the end of multiple sequence alignment in phylogenetics? *Cladistics* 12, 1–10.
- Wheeler, W.C., 2003. Implied alignment: a synapomorphy-based multiple-sequence alignment method and its use in cladogram search. *Cladistics* 19, 261–268.
- Wheeler, W.C., 2006. Dynamic homology and the likelihood criterion. *Cladistics* 22, 157–170.
- Wheeler, W., Gladstein, D., De Laet, J., 2003. POY, Version 3.0.11. Distributed by the authors, New York City.
- Whiting, A.S., Sites, J.W., Pellegrino, K.C.M., Rodrigues, M.T., 2006. Comparing alignment methods for inferring the history of the new world lizard genus *Mabuya* (Squamata: Scincidae). *Mol. Phylogenet. Evol.* 38, 719–730.
- Worsaae, K., 2005. Phylogeny of Nerillidae (Polychaeta, Annelida) as inferred from combined 18S rDNA and morphological data. *Cladistics* 21, 143–162.
- Zujko-Miller, C., Miller, J.A., 2003. *PEST: Precision Estimated by Sampling Traits*. Program distributed by the authors, <http://www.gwu.edu/~clade/spiders/pestDocs.htm>.

Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1. Number of clades resolved by POY, MUSCLE, and DCA across the 36 alignment-parameter sets when gaps were treated as missing data. Plus and minus one standard error is shown for each bar.

Fig. S2. Number of clades resolved by POY, MUSCLE, and DCA across the 36 alignment-parameter sets when gaps were coded using simple indel coding (SIC). Plus and minus one standard error is shown for each bar.

Fig. S3. Number of clades resolved by POY, MUSCLE, and DCA across the 36 alignment-parameter sets when gaps were coded as 5th states. Plus and minus one standard error is shown for each bar.

Fig. S4. Number of clades resolved by POY, MUSCLE, and DCA across the 36 alignment-parameter sets when characters were weighted using the “logically consistent” approach. Plus and minus one standard error is shown for each bar.

Fig. S5. Averaged number of clades resolved by POY, MUSCLE, and DCA across the 36 alignment-parameter sets when gaps were treated as missing data. Plus and minus one standard error is shown for each bar.

Fig. S6. Averaged number of clades resolved by POY, MUSCLE, and DCA across the 36 alignment-parameter sets when gaps were coded using simple indel coding (SIC). Plus and minus one standard error is shown for each bar.

Fig. S7. Averaged number of clades resolved by POY, MUSCLE, and DCA across the 36 alignment-parameter sets when gaps were coded as 5th states. Plus and minus one standard error is shown for each bar.

Fig. S8. Averaged number of clades resolved by POY, MUSCLE, and DCA across the 36 alignment-parameter sets when characters were weighted using the “logically consistent” approach. Plus and minus one standard error is shown for each bar.

Please note: Blackwell Publishing are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Appendix 1

The 36 alternative sets of alignment parameters used in sensitivity analyses.

Simulation	Gap cost	Transversion cost	Transition cost
1	1	2	4
2	1	2	2
3	1	2	1
4	2	4	1
5	4	8	1
6	1	2	0
7	1	1	2
8	1	1	1
9	2	2	1
10	4	4	1
11	8	8	1
12	1	1	0
13	2	1	2

Appendix 1 Continued

Simulation	Gap cost	Transversion cost	Transition cost
14	2	1	1
15	4	2	1
16	8	4	1
17	16	8	1
18	2	1	0
19	4	1	2
20	4	1	1
21	8	2	1
22	16	4	1
23	32	8	1
24	4	1	0
25	8	1	2
26	8	1	1
27	16	2	1
28	32	4	1
29	64	8	1
30	8	1	0
31	16	1	2
32	16	1	1
33	32	2	1
34	64	4	1
35	128	8	1
36	16	1	0