

Simultaneously Mapping and Superimposing Landmark Configurations with Parsimony as Optimality Criterion

S. A. CATALANO^{1,2,*} AND P. A. GOLOBOFF^{1,3}

¹Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina; ²Fundación Miguel Lillo, Miguel Lillo 251, 4000 S.M. de Tucumán, Argentina; and ³Facultad de Ciencias Naturales e Instituto Miguel Lillo, Miguel Lillo 205, 4000 S.M. de Tucumán, Argentina;

*Correspondence to be sent to: Fundación Miguel Lillo, Miguel Lillo 251, 4000 S.M. de Tucumán, Argentina; E-mail: sacatalano@gmail.com.

Received 25 May 2011; reviews returned 1 August 2011; accepted 29 August 2011

Associate Editor: Norman MacLeod

Abstract.—All methods proposed to date for mapping landmark configurations on a phylogenetic tree start from an alignment generated by methods that make no use of phylogenetic information, usually by superimposing all configurations against a consensus configuration. In order to properly interpret differences between landmark configurations along the tree as changes in shape, the metric chosen to define the ancestral assignments should also form the basis to superimpose the configurations. Thus, we present here a method that merges both steps, map and align, into a single procedure that (for the given tree) produces a multiple alignment and ancestral assignments such that the sum of the Euclidean distances between the corresponding landmarks along tree nodes is minimized. This approach is an extension of the method proposed by Catalano et al. (2010. Phylogenetic morphometrics (I): the use of landmark data in a phylogenetic framework. *Cladistics*. 26:539–549) for mapping landmark data with parsimony as optimality criterion. In the context of phylogenetics, this method allows maximizing the degree to which similarity in landmark positions can be accounted for by common ancestry. In the context of morphometrics, this approach guarantees (heuristics aside) that all the transformations inferred on the tree represent changes in shape. The performance of the method was evaluated on different data sets, indicating that the method produces marked improvements in tree score (up to 5% compared with generalized superimpositions, up to 11% compared with ordinary superimpositions). These empirical results stress the importance of incorporating the phylogenetic information into the alignment step. [Landmark configuration; mapping; optimization; parsimony; phylogeny; superimposition.]

The use of landmark data in phylogenetics has been surrounded by a long-lasting controversy, centered on different topics: the definition of character; the proper way to handle continuous varying characters; the dependence issue; the appropriateness of different phylogenetic methods to analyze these data, etc. (e.g., Smith 1990; Bookstein 1994; Zelditch et al. 1995, 1998; Naylor 1996; Adams and Rosenberg 1998; Rohlf 1998, 2002; Monteiro 2000; Bookstein 2002; MacLeod 2002; Stone 2003). In recent years, the debate has been renewed with the publication of new approaches for the phylogenetic treatment of landmark data (González-José et al. 2008, 2011; Catalano et al. 2010; Jones and Moriarty 2010; Klingenberg and Gidaszewski 2010; Adams et al. 2011). Catalano et al. (2010, p. 544) briefly discussed how to superimpose (match, align) landmark configurations when these are to be analyzed in a phylogenetic framework—an aspect that had not been part of the debate before. This issue is essential, given that different multiple alignments may potentially produce different ancestral landmark configurations and, concomitantly, different inferences of shape change along the tree.

The criterion most commonly used for the superimposition of 2 landmark configurations is minimizing the sum of squared distances between corresponding landmarks (Rohlf 1990; Rohlf and Slice 1990) mainly because it allows subsequent sophisticated statistical analyses (Slice 2005). However, alternative criteria have been proposed: Siegel and Benson (1982) and Benson et al. (1982) developed the method of resistant-fit theta-rho analysis (RFTRA) based on repeated medians, which is less sensitive to outliers; Slice (1996) described an extension of

RFTRA to superimpose 3-dimensional (3D) landmark configurations; Dryden and Walker (1999) proposed a matching procedure based on least median of squares estimators; more recently, Larsen (2008) proposed an approximate method to superimpose landmark configurations minimizing the sum of Euclidean distances between landmarks. Although any of these criteria can be applied directly to the superimposition of 2 configurations, this is not so when the comparison involves several configurations. That is, except for trivial cases, it is not possible to sort out the differences that are not due to changes in shape for all possible pairs of configurations simultaneously. For instance, if configurations A and B are superimposed against configuration C by minimizing the sum of squared distances between corresponding landmarks, the resulting superimposition between configurations A and B may not be optimal for that criterion. Additional considerations (which may depend on the kind of analysis) are needed to produce reasonable multiple superimpositions. The generalized Procrustes analysis (Gower 1975; modified by Rohlf and Slice 1990) produces a multiple alignment of configurations by minimizing the sum of the squared distances between corresponding landmarks of each configuration and a consensus configuration. The solution given by a generalized superimposition procedure may be reasonable if the goal is to extract shape variables to be used subsequently in statistical analyses. However, alternative approximations seem more appealing in a phylogenetic context.

We have recently proposed a method (Catalano et al. 2010) that establishes ancestral landmark configurations

by minimizing the differences in landmark position along the nodes of the tree. The tree score (S) for all landmarks of a configuration is calculated (for landmarks in 2D) as

$$S = \sum_{i=0}^{L-1} \sum_{j=0}^{N-1} \sqrt{(x_{i,j} - x_{i,\text{an}(j)})^2 + (y_{i,j} - y_{i,\text{an}(j)})^2},$$

where

- x = coordinate of the landmark along X -axis
- y = coordinate of the landmark along Y -axis
- N = number of nodes (internal and terminal) in the tree excluding the root
- L = number of landmarks in the configuration
- $\text{an}(j)$ = ancestor of node j

Hence, the difference in position for each landmark on each branch is calculated as the Euclidean distance between the position of the landmark in the ancestor and the position of the landmark in the descendant. The total score for each landmark is obtained by summing up those distances over all nodes. Finally, the score for the whole configuration on the tree is obtained by summing the scores over all landmarks. The choice of the Euclidean metric to calculate the score of each landmark is not arbitrary: it was chosen because this allows extending the parsimony criterion to the analysis of continuous characters that change in more than one dimension (Catalano et al. 2010). Given that the tree score for a given configuration is calculated by summing up the contribution of each landmark, the corresponding metric for the tree score is a Manhattan-like distance. This combination of an Euclidean distance to evaluate the changes between configurations in the position of each landmark and a Manhattan metric to calculate the total score by summing up the contribution of each landmark has the consequence that neither the tree score nor the inferences of shape change are affected by the original orientation of the specimens (see Catalano et al. 2010).

Our method (Catalano et al. 2010) to map landmark data infers the assignments to the internal nodes considering an a priori superimposition that remains unmodified during the mapping procedure (this is common to all previously proposed methods to map landmark data onto a tree). However, the tree score may be improved (concomitantly maximizing the amount of similarity in landmark positions that can be explained by common ancestry) if the superimposition among configurations is modified during the mapping procedure. This approach is not novel within systematics: in molecular sequence analysis this idea has more than 30 years, grounded on the seminar work by Sankoff and collaborators (Sankoff et al. 1973; Sankoff 1975; Sankoff and Rousseau 1975; Sankoff and Cedergren 1983), and has been further developed more recently by Wheeler and collaborators (Wheeler 1996, 2003; Wheeler et al. 2006; Varón et al. 2010). The problem of finding internal node assignments with an implied multiple alignment minimizing tree alignment score, given a tree, is known as

the *tree alignment problem* (Sankoff and Cedergren 1983, Schwikowski and Vingron 1997). In the present study, we describe a method to obtain a multiple superimposition of landmark configurations and internal node assignments using parsimony as optimality criterion. Throughout this paper, a *dynamic* alignment denotes one produced by simultaneously superimposing and mapping on a tree and a *static* alignment one produced by aligning landmark configurations prior to the mapping process.

One of the main implications of a dynamic alignment approach is that the optimal superimposition depends on the tree (see below). A simple example (Fig. 1) illustrates how the evaluation of shape change along the tree is improved by dynamic alignment, with a case comprising 7 species mapped on 2 different trees that differ in the position of species B. In one of the trees (left), B is placed next to A. In the other tree (right), B is separated from A by several taxa with different parts of the configuration successively translated. The configurations mapped with a static (top) and a dynamic approach (bottom) are displayed for each of the trees. The static approach used for this example is an ordinary alignment that minimizes the sum of the Euclidean distances between landmarks and takes configuration C as reference. However, the point illustrated in this example also applies to any static alignment strategy.

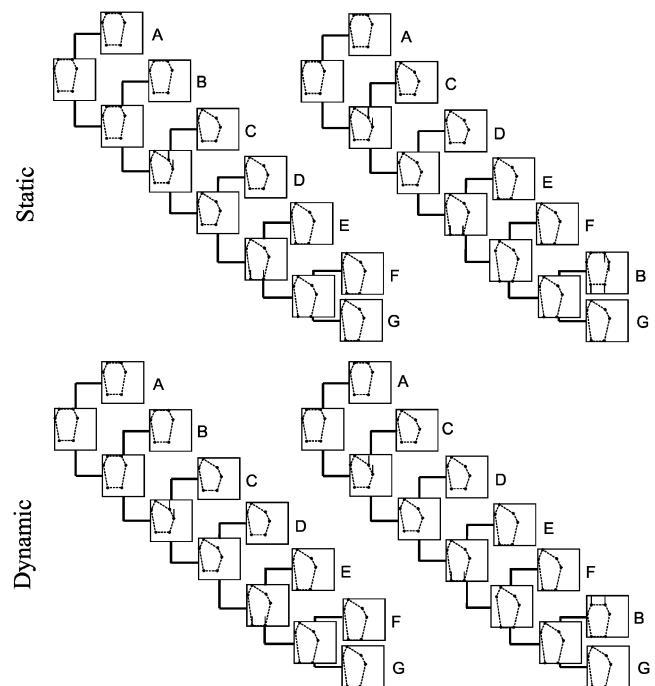


FIGURE 1. Improvement in shape change inference obtained with the dynamic approach. The inference of shape change is the same in static (top) and dynamic (bottom) approaches in the left tree where A is next to B. However, the approaches give different inferences of shape change when these species are separated in the tree (right tree). The dynamic approach recognizes that the similarity between A and B is due to convergence and hence does not force these configurations to be perfectly superimposed (as the static approach does) and improves 25% the tree score by repositioning shape B differently from A.

Both approaches, static and dynamic, produce the same alignment and conclusions of shape changes for the left tree but not for the right tree. In the right tree, the difference arises because the multiple alignment derived from a static approach will perfectly superimpose the configurations of A and B, but the distribution of shapes on the tree indicates that the similarity between configurations A and B is due to convergence and not to common ancestry. The dynamic alignment thus lowers tree score (25% compared with the static approach) by repositioning shape B differently from A, as implied by the sequencing of shapes on the tree. In contrast, in the left tree A is next to B, thus making it possible to perfectly account for the shape similarity in A and B in terms of common ancestry. The dynamic alignment, taking this into account, positions A and B identically—which static alignments always do, regardless of what the tree and the rest of the observations may indicate. Also note that, since both trees suggest a different scenario of how the differences in shape arise, then it is not possible for a single alignment to provide the best framework for mapping both trees at the same time—each tree should ideally be considered under a different alignment, showing that the idea of using a single static alignment for phylogenetic studies is misguided.

Given that the dynamic approach produces more appropriate evaluations of how a tree fits the observations, it can also be used for tree selection, considering either landmark data alone or in combination with other data such as DNA sequences and qualitative morphological data. Using the method for tree selection only requires that each tree be evaluated during a search routine (Wagner tree, subtree pruning regrafting [SPR], tree bisection reconnection [TBR], etc.) considering its optimal alignment.

Just as the dynamic approach can be seen as an extension of the optimization method of Catalano et al. (2010), it can at the same time be considered as an extension of the superimposition methods for the case of landmark configurations linked by a tree. Within geometric morphometrics, differences in shape are defined as those differences in landmark configurations that remain after sorting out rotation, size, and translation (Kendall 1977). This definition of differences in shape is completely tied to criteria to superimpose, with alternative criteria producing alternative evaluations of shape differences. Changes in shape between configurations are quantified by a metric that is a function of the differences in position of individual landmarks once the configurations are optimally superimposed. In the analysis of shape change on a tree, in order to properly interpret differences between landmark configurations along the tree as changes in shape, the metric should also form the basis to choose the ancestral assignments. In other words, the metric to superimpose and map must be the same. In our parsimony-based method (parsimony sensu Farris 1983), this metric is the sum of the Euclidean distances between corresponding landmarks.

In the following sections, we describe new algorithms to produce optimal multiple alignments and ancestral landmark configurations by minimizing the sum of the Euclidean distances between corresponding landmarks. In addition, we evaluate the performance of the method in several data sets. The methods described have been implemented in the program TNT (Tree Analysis using New Technologies; Goloboff et al. 2003, 2008).

THE APPROACH

Throughout this paper, the term “position” when referring to a configuration indicates not only its location in a x, y, z coordinate system but also its orientation. Obtaining an exact solution for the tree alignment problem for the case of landmark configurations would require evaluating every combination of positions for the configurations observed in the terminal nodes and every possible landmark configuration for each internal node and choosing that combination which minimizes the tree score. As this task seems computationally impossible even for the smallest data sets, a heuristic method to find suitable solutions is needed. The method presented here allows obtaining improvements in tree score by rotating and translating configurations but not by resizing. As the score of a multiple alignment depends on the size of the configurations, the score will always decrease if all the configurations are shrunk. Hence, configurations should be adjusted in size before the analysis (e.g., scaling to centroid size or performing an ordinary alignment). There are possible modifications of the method presented here that would allow analyzing landmarks configurations despite the differences in size, allowing changes in size to be considered as an additional source of phylogenetic evidence. This possible approach will be included in future contributions.

HEURISTICS

The problem can in principle be decomposed in 2 parts: how to obtain the positions of the configurations assigned to the terminals and how to obtain the optimal ancestral assignments, both subproblems requiring simultaneous resolution. The spatial optimization (Catalano et al. 2010; Goloboff and Catalano 2011) allows establishing ancestral landmark configurations from a multiple superimposition of the terminals. Hence, the problem can be solved if the score of every possible multiple alignment is calculated by means of spatial optimization and the one that minimizes the score is kept as the best solution, but the process can be made more time efficient in several ways.

The approach starts from a preliminary multiple alignment and an initial assignment of ancestral landmark configurations obtained by spatial optimization (Goloboff and Catalano 2011). A good starting point can be an ordinary superimposition minimizing the sum of Euclidean distances between corresponding landmarks. Given that all the costs are symmetric, it is irrelevant whether the tree is rooted or not. For the sake

of simplicity, the method will be presented for an unrooted tree. The procedure starts visiting internal nodes (for the time being, no particular order is required). At each internal node N , the tree is decomposed into as many subtrees as nodes are connected to N (3 in the case of binary trees), with their terminals forming an equal number of sub-alignments (Fig. 2). Better multiple alignments can be obtained by modifying the positions of the configurations for each sub-alignment in tandem, relative to the rest of the multiple alignment. The score for the new multiple alignment is calculated by a complete spatial optimization; the alignment is kept if better, discarded otherwise.

A drawback of the above procedure is the time required to re-optimize the whole tree each time the multiple alignment is modified. However, it is possible to evaluate the quality of a candidate multiple alignment without making a complete re-optimization. To see why this is so, consider Figure 2. The tree score (i.e., the sum of the landmarks displacements) can be decomposed into 6 terms: the sum of the scores for all the branches belonging to each of the 3 sub-alignments (S_1, S_2, S_3) and the score along the branches that connect each sub-alignment to the central node (B_1, B_2, B_3). The total tree score L can hence be expressed as $L = \sum S_i + \sum B_i$. When a sub-alignment i is modified relative to the rest, if all its configurations (terminal and internal) are modified in the same way (i.e., maintaining the relative positions of the configurations within the sub-alignment), its score S_i remains unaltered. So do the S 's and the B 's corresponding to the unmodified sub-alignments. The only score that changes is B_i for the branch that subtends the modified sub-alignment. If the modification produces a smaller B_i , L will be necessarily smaller, and a better multiple alignment is obtained. This score for the new multiple alignment, however, may be an

overestimation of the score that would be established if a new spatial optimization is conducted, updating the assignments for the rest of the nodes. This procedure is repeated for all internal nodes, and after that, a new spatial optimization is performed considering the new multiple alignment, obtaining new assignments to the internal nodes. At this point, the procedure can be either restarted, visiting all the internal nodes in a second round of improvements, or interrupted if no further improvement was observed.

A rearrangement among the sub-alignments that produces a smaller B_i necessarily improves the score of the multiple alignment. However, some rearrangements worsening B_i may also produce better multiple alignment. This is the cost to be paid for saving time in not fully updating the ancestral assignments when a rearrangement is evaluated, and it will decrease the effectiveness of the method to find better global solutions. A way to reduce the rejection rate of better solutions maintaining time efficiency is to update the assignments only in the nearby region of the node that subtends the modified sub-alignment. Every time a new rearrangement is evaluated, a new assignment to the central node is established considering the new position for the neighbor node that subtends the modified sub-alignment and the remaining (unmodified) neighbor nodes. The new assignment to the central node is determined by calculating for each landmark the position that minimizes the sum of distances with homologous landmarks on neighbor nodes (see Catalano et al. 2010). The decision of rejecting or accepting the new rearrangement will be now based on the new $\sum B_i$ and not only on the score of the branch leading to the modified sub-alignment.

Further Time-Saving Shortcuts

Every time a new rearrangement among the sub-alignments is evaluated, all the configurations belonging to this subtree must be modified to maintain their relative positions unaltered. For instance, if the modification to be evaluated involves a translation of 0.01 units to the right, this requires recalculating the coordinates of all the configurations (those of internal nodes and terminals) considering this new position. This procedure can take considerable time, especially in the case of rotations. An alternative procedure (the one actually used in the implementation of this method in TNT) is to modify only the node connected to the central node, leaving the rest of the nodes unmodified (Fig. 3). This requires visiting tree nodes in an orderly fashion, making sure that unmodified descendant nodes will be visited subsequently. The modification of the multiple superimposition is hence obtained when the procedure visit an internal node which has a terminal node as neighbor.

The details of the procedure for choosing the optimal position for the neighbor nodes are included as Supplementary Material (available from <http://www.sysbio.oxfordjournals.org>). Succinctly, the algorithm modifies in turn the orientation, position in x , y , and z of each of the neighbor configurations using a binary-like

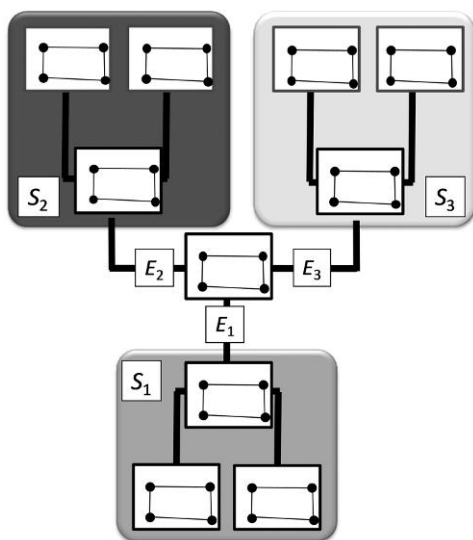


FIGURE 2. Decomposition of the tree score into 6 components: the score of each sub-alignment (S_i) and the score of each branch connecting the central node and each sub-alignment (E_i).

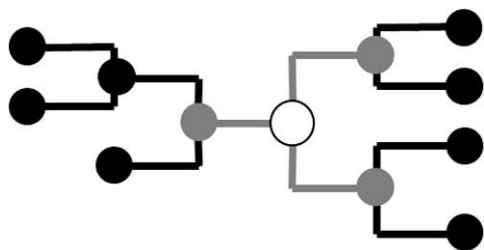


FIGURE 3. The method works by visiting each internal node, choosing the assignment for the central node (empty circle) and the position of the configurations (orientation and location in x -, y -, and z -axes) in neighbor nodes (gray circles) which minimize the landmark displacements along the reduced tree defined by these nodes (gray branches). The multiple alignment is modified when the neighbor nodes are represented by terminals.

algorithm trying to iteratively minimize the sum of the Euclidean distance for all landmarks in all branches connecting the internal node to its neighbors.

Escaping Local Optima

The procedure previously described can be trapped in a local optimum for several reasons: (i) in our heuristics, each rearrangement evaluated represents a modification only in either the orientation of the configuration or its position on x -, y -, or z -axis. This is more time efficient, but it is possible that a best solution can only be achieved if the modification involves a combination of these changes; (ii) the specific search algorithm used to improve the score by modifying the x , y , z position and the orientation at each internal node can be trapped into a local optimum (see Supplementary Material 1); (iii) when a node is visited, the neighbor nodes are modified one at a time to improve the score but achieving a better alignment may require that several neighbors are simultaneously modified; (iv) the optimal multiple alignment might only be obtained if nodes further apart on the tree are simultaneously modified. Hence, techniques to overcome local optima are needed.

Which local optimum the method reaches depends on the multiple alignment used as a starting point. A possible solution to overcome local optima is trying different starting points in the hopes that one of them ends up in a global optimum. This approach has the disadvantage that all the computational effort of each run is thrown away every time a new starting point is evaluated. An alternative solution, the one followed here, is to take the multiple alignment obtained after a single run, slightly perturb it, and use it as a new starting point. The approach we implemented includes a (user-defined) number of cycles of improvement/perturbation. In the perturbation phase, the position of some terminal configurations is randomly modified. After that, a spatial optimization is conducted and a new round of improvement on the multiple alignment is performed. If the score obtained is better than the best so far, the new alignment is accepted; otherwise the best alignment is randomized again. The positions of the configurations

are randomized in such a way that the structure of the tree is taken into account: the positions of all the configurations belonging to a clade are randomly modified, not only individual configurations. In each randomization cycle, a certain number of clades (one-third of the total in our implementation) are randomly chosen to be modified. If an internal node is chosen, the configurations of all the terminals belonging to that node are modified. Although this approach does not guarantee reaching a global optimum, in the long run it greatly increases the chances of doing so.

A summary of the main steps of the whole procedure is

0. Generate a static multiple alignment.
 1. Perform a spatial optimization.
 2. Travel along all internal tree nodes; for each node:
 - 2.1. Generate a reduced tree formed by the neighbors of the node. Calculate its score.
 - 2.2. Try to improve the score of the reduced tree by changing in turn the position of the configurations assigned to the neighbor nodes. If a better score is obtained, update the assignment for the middle node and reposition existing configurations for the neighbor nodes (for further details, see Supplementary Material 1).
3. Repeat step 2 until no improvement in the score on any internal node is obtained or the number of rounds equals a predefined limit.
4. Perform a new spatial optimization. If no improvement in tree score is obtained or the number of optimization cycles equals a predefined limit go to 5, else restart at 1.
5. If the number of perturbation cycles equals the user-defined number, stop the procedure and keep the multiple alignment; otherwise, perturb the alignment and restart at 1.

EMPIRICAL ANALYSES

The data sets analyzed comprise configurations representing different biological structures, different taxonomic groups, and different sizes in terms of both number of landmarks and number of terminals (Table 1). The trees used as guide were derived from previously published phylogenies and, except for the hominid data set, were obtained from alternative sources of evidence. Unless indicated, all the analyses were conducted in TNT (Goloboff et al. 2008).

Behavior of the Algorithms

The results obtained with the procedure just described may depend on the starting point. This is caused by the heuristic aspect of our procedure, and not by the criterion itself; if the full space of solutions were evaluated, the same (optimal) solution would always be obtained. To evaluate empirically the effect of the starting

TABLE 1. Information of the data sets used in comparison between dynamic and static approaches

Source of landmark data	Source of guide tree	Taxonomic group	Structure	Number of landmarks	Number of terminals
González-José et al. (2008)	Goloboff and Catalano (2011)	Hominids	Cranium	7	17
Astúa (2009)	Astúa (2009)	Didelphids	Scapula	14	32
Rohlf (2002)	Harbach and Kitching (1998)	Culicids	Wing	18	11
Morales M.M., Prevosti F.J., Catalano S.A., Giannini N.P. (unpublished data)	Johnson et al. (2006)	Felids	Cranium	24	47

point on the results, we analyzed the convergence in scores among runs starting from different multiple alignments. Two different types of starting points were considered. In a first analysis, each run started from an alignment generated by superimposing all the configurations against a single one (a so-called ordinary superimposition), considering as many starting points as configurations included in the data set. The superimposition criterion was the minimization of the sum of the Euclidean distances between corresponding landmarks (i.e., the same criterion used to measure alignment optimality on the tree; a heuristic procedure was developed to match a pair of configurations using this criterion, see Supplementary Material 1). In a second analysis, each run started from a multiple alignment obtained by randomizing the position of all the configurations in the data set. Ten different starting points were considered in this case. The level of modification was such that the tree scores derived from the randomized alignments were about 10 times the score derived from an ordinary superimposition. The convergence was evaluated comparing the coefficient of variation of tree scores from the original alignments (ordinary alignments or randomized alignments) with the coefficient of variation of the scores obtained after 1, 10, 100, and 1000 cycles of improvement/perturbation. Before superimposing with the ordinary or the randomize procedure, and in order to filter out the differences in size, the data sets were first superimposed by generalized procrustes analysis (GPA) using *tpsRelw* (Rohlf 2008) and no further changes in size were subsequently performed (see “The Approach” section).

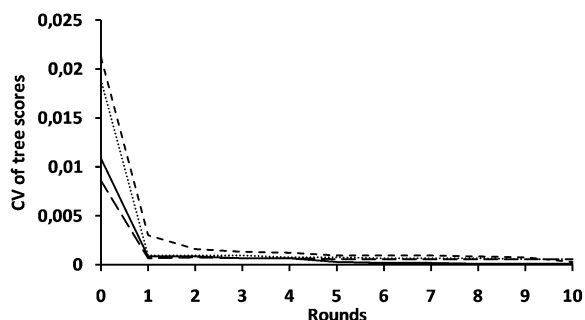


FIGURE 4. Convergence of tree scores (expressed as the coefficient of variation) among runs departing from different ordinary superimpositions along 10 cycles of improvement/perturbation. Type of line indicates data set: solid, culicids; long-dashed, felids; short-dashed, hominids; dotted line, didelphids.

The analyses showed that the different runs rapidly converge to similar scores, regardless of whether the starting point is from ordinary (Fig. 4) or random (not shown) alignments. In the former analysis, the coefficient of variation of tree scores was reduced after the first cycle between 6 and 23 times—depending on the data set—and diminished between 25 and 180 times after 100 cycles of improvement/perturbation. After 100 cycles, the coefficient of variation is lowered to values between 3.3×10^{-3} and 9.2×10^{-5} . These results strongly suggest that the final result is not significantly determined by the starting point.

The results also showed that most of the improvement in tree score was obtained in the first cycle, before entering the phase of improvement/perturbation (Fig. 5). However, the score continued improving as more cycles of improvement/perturbation are conducted.

Scores of Dynamic versus Static Approach

Although the approach presented here is theoretically sounder, in practice it might be the case that prior alignments provide a good enough approximation to the values and results obtained with the present method. This question can be evaluated empirically by examining whether the dynamic approach produces a marked improvement on tree scores compared with the non-phylogenetic alignments. Should the results be very similar, this would indicate that the static approach is a good and fast approximation to the real (=optimal) solution. To evaluate this issue, the scores obtained

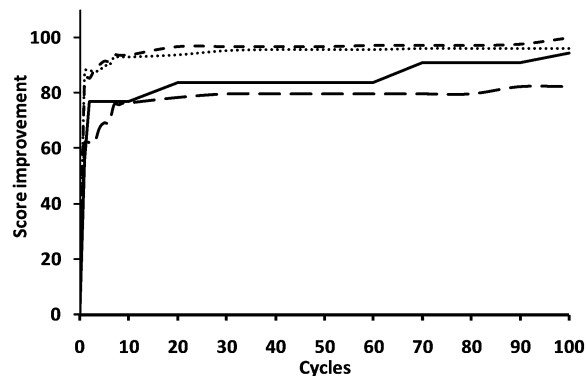


FIGURE 5. Improvement in tree score along 100 cycles of improvement/perturbation. Values expressed as percentage of the total improvement obtained after 1000 cycles. Most of the improvement is obtained in the first cycle. The runs used generalized Euclidean superimposition as a starting point. Line codes as in Figure 4.

TABLE 2. Improvement (expressed as percentage) on tree scores obtained by the dynamic approach against those derived from different fixed alignments strategies

	Generalized Euclidean	Ordinary Euclidean, mean/largest
Hominids	5.2	7.6/10.9
Didelphids	4.9	5.9/10.6
Culicids	0.7	2.6/4.6
Felids	1.2	2.2/5.4

by simultaneously mapping and aligning were compared with the scores derived from mapping configurations that were aligned a priori by (i) superimposing all configurations against a single configuration (ordinary superimposition) and (ii) superimposing all configurations against a “mean” configuration (generalized superimposition). In both the cases, the optimality criterion to align was the same as in the present method: the sum of the Euclidean distances between landmarks, so that any difference in the score between the dynamic and the static approach can be attributed solely to considering, or not, phylogenetic information in the alignment step. The results of generalized superimposition should be taken cautiously given that, unlike GPA, the generalized alignment based on Euclidean distances is not guaranteed to monotonically converge.

The analyses showed that the tree scores obtained by the dynamic approach (after 1000 cycles of improvement/perturbation) were considerably better in all the data sets, although there were differences among data sets (Table 2). The improvements were of up to 5% relative to generalized alignments and up to 11% relative to ordinary alignments.

Tree Searches

In order to evaluate the performance of the method in tree searches, we analyzed 2 data sets that have been previously used to evaluate other approaches to analyze landmark data in phylogenetics. One of them is the well-known simulated fish data set of Naylor (1996). The shapes were scanned from MacLeod (2002: p. 127, figure 7.11) and separated into 12 different configurations following MacLeod (2002). The second data set comprises landmark configurations representing the wing morphology of 11 species of *Drosophila* used in Klingenberg and Gidaszewski (2010). Given that we obtained no answer from the corresponding author to our request for this data set, we scanned the figures from the paper. In the case of Naylor’s data set, to decrease the influence on tree choice of those configurations represented by many landmarks, configurations were down-weighted in inverse proportion to the number of landmarks (Catalano et al. 2010). Landmark rescaling and optimization were performed as described by Goloboff and Catalano (2011). The search used 10 different starting points (random trees) followed by TBR. The score of each candidate tree in the TBR procedure was calculated considering the best multiple alignment obtained after 20 cycles of perturbation/improvement. The spa-

tial optimization settings (see Goloboff and Catalano (2011) for an explanation) to calculate the score were a grid of 5 × 5 cells, 3 nesting levels, and a window of 1 cell, states of the terminals included as possible states in the internal nodes, and iterative improvement of ancestral assignments. The analysis of Naylor’s data set retrieved the correct (model) tree. In the case of the *Drosophila* data set, the analysis retrieved a single tree that has the same distance to the molecular “reference” tree of Klingenberg and Gidaszewski (2010) as the tree obtained with the method of Klingenberg and Gidaszewski (2010), when tree distance was measured either with number of SPR moves (3) number of taxa in the agreement subtree (5) or by Robinson–Foulds distance (0.714).

CONCLUSIONS

The method described in this contribution is a refinement of that proposed by Catalano et al. (2010) to analyze changes in landmark configurations along a given phylogeny. Whereas in the original version of the method configurations were aligned prior to the mapping step, the extension presented here simultaneously superimposes and maps landmark configurations. Although the optimality criterion considered here is parsimony, the general approach can also be extended to alternative criteria to map landmark data on a tree (such as the minimization of the sums of squared distances). In fact, J. Felsenstein and F. Bookstein are currently working on a method which, despite some differences, can be thought as a maximum likelihood counterpart of our method (Felsenstein J., personal communication).

The analysis of the 4 real data sets showed improvements in tree score of up to 5% (up to 11% if the comparison is done against scores derived from ordinary alignments). In other terms, it means that 5–11% of the shape changes inferred on the trees when mapping a fixed alignment were not changes in shape but merely differences produced by suboptimal alignments. These marked differences give empirical support to our claim that, when mapping landmark configurations, the alignment procedure should take into account the phylogenetic information. Thus, using an approach as the one described here seems highly desirable. In fact, it seems logical to expect that similar improvements will be obtained when other metrics (such as least squares minimization) are used to map landmark configurations. Obviously, in some specific cases, there may be reasonable alternatives to define the frame of comparison other than the pure minimization of landmark displacements described, such as when relevant external biological information is available. A typical case is the evolutionary conservativeness of parts of the structure that allows using a 2-point registration approach (Bookstein 1982).

One of the reviewers indicated that a drawback of our method is that, given that it is not based on any statistical model, one really does not know how to evaluate its suitability. Our main goal was to develop a framework

that allows incorporating landmark data into a phylogenetic analysis either alone or in combination with other sources of evidence. In this context, it is completely possible to test the method: if several empirical analyses indicate that the phylogenetic results obtained with the method are in general agreement with those obtained with alternative sources of evidence, the general validity of the approach is strongly upheld.

A limitation of our approach to select a phylogenetic tree (i.e., to do tree searches) is the time required to evaluate each tree, making it possible to effect TBR searches for only a few dozen taxa. Two possible strategies may help overcome this time limitation: (i) parallelize the job, which can be easily implemented with TNT scripts, and (ii) establish meta-strategies for searches in which not all topologies are evaluated according to their best alignments (for instance, the configurations are realigned dynamically on the tree, but only every X rearrangements accepted). Alternatively, and until faster implementations of dynamic alignment are available, searching under static alignments as an approximation to the results of searching under dynamic alignments (as proposed by De Laet (2005) for sequence data) may be the only option for large data sets. The ranking of trees evaluated according to static or dynamic alignments may present differences (e.g., a tree A considered to be worse than a tree B under a static evaluation may, however, be better when a dynamic evaluation is used). The rank differences, however, are much smaller when comparing only the trees in a TBR neighborhood (i.e., all the trees produced by TBR rearrangements from a single tree), suggesting that static evaluations may produce reasonable approximations during heuristic tree searches.

The approach presented here, although sharing with Kendall's framework the same definition of shape, diverges from that approach in using an alternative metric to superimpose and evaluate differences in shape, a metric that allows extending the parsimony principle for the analysis of landmark data. Given that the specific geometric constructions Kendall (1984) proposed were "designed to suit statistical needs", working outside such framework can hardly be considered a drawback for a method like the one presented here, which is not proposed as a probabilistic approach to the problem. The quantification of shape changes as the sum of differences in landmark positions (once the configurations were superimposed) is obviously an overly simplistic approach to define the ordering of states in characters that describe shape. One of its limitations (shared with other approaches to map landmark configurations) is that under some particular changes in shape that could occur during the course of evolution, it would not be possible to establish the superimposition that allows tracing the actual events of changes (for an example, see figure 13 of Richtsmeier et al. 2002). Other limitation is that (as implemented herein) it does not take into account the dependence among landmarks. Despite these limitations, the quantification of shape differences by the analysis of differences in landmark positions is an

improvement compared with the traditional analysis of shape characters in phylogenetics, where states (and their ordering, i.e., transformation costs) were defined in a much more subjective way.

SUPPLEMENTARY MATERIAL

Supplementary material, including data files and/or online-only appendices, can be found at <http://www.sysbio.oxfordjournals.org/>.

FUNDING

This work was supported by CONICET (PIP-CONICET 6502 to M.L. de Grosso), ANPCyT (PICT 2007-01314 to P.A.G.).

ACKNOWLEDGMENTS

The authors thank the Willi Hennig Society for subsidizing the program TNT and making it freely available. We would like to thank Miriam Morales for her unpublished feline data set and Diego Astúa for his Didelphid data set. Many ideas presented in this article have been inspired by Ward Wheeler's work on direct optimization—an intellectual debt which we feel compelled to acknowledge. We greatly appreciate the invaluable comments and criticism given by J. Felsenstein, 2 anonymous reviewers that considerably helped to improve the manuscript, as well as associate editor N. MacLeod. We also wish to thank many colleagues for useful critical comments on the manuscript, in particular J. Salvador Arias, Ignacio Escapa, Rolando González-José, Iván Pérez, and Florencia Vera Candioti. The research for this paper was carried out at the Instituto Superior de Entomología "Dr. Abraham Willink" and the Instituto Argentino de Estudios Filogenéticos.

REFERENCES

- Adams D.C., Cardini A., Monteiro L.R., O'Higgins P., Rohlf F.J. 2011. Morphometrics and phylogenetics: principal components of shape from cranial modules are neither appropriate nor effective cladistic characters. *J. Hum. Evol.* 60:240–243.
- Adams D., Rosenberg M. 1998. Partial warps, phylogeny, and ontogeny: a comment on Fink and Zelditch (1995). *Syst. Biol.* 47: 168–173.
- Astúa D. 2009. Evolution of scapula size and shape in didelphid marsupials (Didelphimorphia: Didelphidae). *Evolution.* 63:2438–2456.
- Benson R.H., Chapman R.E., Siegel A.F. 1982. On the measurement of morphology and its change. *Paleobiology.* 8:328–339.
- Bookstein F.L. 1982. Foundations of morphometrics. *Annu. Rev. Ecol. Syst.* 13:451–470.
- Bookstein F.L. 1994. Can biometrical shape be a homologous character? In: Hall B.K., editor. *Homology: the hierarchical basis of comparative biology.* New York: Academic Press. p. 197–227.
- Bookstein F.L. 2002. Creases as morphometric characters. In: MacLeod N., Forey P., editors. *Morphology, shape and phylogeny.* Systematics association special volume series 64. London: Taylor & Francis. p. 139–174.
- Catalano S.A., Goloboff P.A., Gianinni N. 2010. Phylogenetic morphometrics (I): the use of landmark data in a phylogenetic framework. *Cladistics.* 26:539–549.

- De Laet J. 2005. Parsimony and the problem of inapplicables in sequence data. In: Albert V.A., editor. Parsimony, phylogeny, and genomics. New York: Oxford University Press. p. 81–116.
- Dryden I.L., Walker G. 1999. Highly resistant regression and object matching. *Biometrics*. 55:820–825.
- Farris J. 1983. The logical basis of phylogenetic analysis. In: Platnick N.I., Funk V.A., editors. *Advances in cladistics II*. New York: Columbia University Press. p. 7–36.
- Goloboff P.A., Farris J., Nixon K. 2003. TNT: tree analysis using new technology. Program and documentation. Available from: <http://www.zmuc.dk/public/phylogeny/TNT>.
- Goloboff P.A., Farris J.S., Nixon K.C. 2008. TNT, a free program for phylogenetic analysis. *Cladistics*. 24:774–786.
- Goloboff P.A., Catalano S.A. 2011. Phylogenetic morphometrics (II): algorithms for landmark optimization. *Cladistics*. 27:42–51.
- González-José R., Escapa I., Neves W., Cúneo R., Pucciarelli N. 2008. Cladistic analysis of continuous modularized traits provides phylogenetic signals in Homo evolution. *Nature*. 453:775–778.
- González-José R., Escapa I., Neves W., Cúneo R., Pucciarelli N. 2011. Morphometric variables can be analyzed using cladistic methods: a reply to Adams et al. *J. Hum. Evol.* 60:244–245.
- Gower J.C. 1975. Generalized Procrustes analysis. *Psychometrika*. 40:33–51.
- Harbach R.E., Kitching I.J. 1998. Phylogeny and classification of the Culicidae (Diptera). *Syst. Entomol.* 23:327–370.
- Jones N.S., Moriarty J. 2010. Evolutionary inference for functional data: using Gaussian processes on phylogenies to study shape evolution. *ArXiv:1004.4668*.
- Johnson W., Eizirik E., Pecon-Slattery J., Murphy W., Antunes A., Teeling E., O'Brian S. 2006. The late Miocene radiation of modern Felidae: a genetic assessment. *Science*. 311:73–77.
- Kendall D.G. 1977. The diffusion of shape. *Adv. Appl. Prob.* 9:428–430.
- Kendall D.G. 1984. Shape-manifolds, procrustean metrics and complex projective spaces. *Bull. Lond. Math. Soc.* 16:81–121.
- Klingenberg C.P., Gidaszewski N.A. 2010. Testing and quantifying phylogenetic signals and homoplasy in morphometric data. *Syst. Biol.* 59:245–261.
- Larsen R. 2008. L1 generalized Procrustes 2D shape alignment. *J. Math. Imaging Vis.* 31:189–194.
- MacLeod N. 2002. Phylogenetic signal in morphometric data. In: MacLeod N., Forey P., editors. *Morphology, shape and phylogeny*. London: Taylor & Francis. p. 100–138.
- Monteiro L. 2000. Why morphometrics is special: the problem with using partial warps as characters for phylogenetic inference. *Syst. Biol.* 49:796–800.
- Naylor G. 1996. Can partial warp scores be used as cladistic characters? In: Marcus L.F., Corti M., Loy A., Naylor G., Slice D.E., editors. *Advances in morphometrics*. New York: Plenum. p. 519–530.
- Richtsmeier J.T., DeLeon V.B., Lele S.R. 2002. The promise of geometric morphometrics. *Yearb. Phys. Anthropol.* 45:63–91.
- Rohlf F.J. 2008. tpsRelw, ver. 1.39. Computer program and documentation, Department of Ecology and Evolution, State University of New York at Stony Brook.
- Rohlf F.J. 1990. Rotational fit (Procrustes) methods. In: Rohlf F.J., Bookstein F.L., editors. *Proceedings of the Michigan morphometrics workshop*. Museum of Zoology, Ann Arbor (MI): University of Michigan. p. 227–236.
- Rohlf F.J., Slice D. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. *Syst. Zool.* 39:40–59.
- Rohlf F.J. 1998. On application of geometric morphometrics to studies of ontogeny and phylogeny. *Syst. Biol.* 47:147–158.
- Rohlf F.J. 2002. Geometric morphometrics and phylogeny. In: MacLeod N., Forey P., editors. *Morphology, shape and phylogeny*. London: Taylor & Francis. p. 175–193.
- Sankoff D.D., Rousseau P., Cedergren R.J. 1973. Evolution of 5S RNA and the non-randomness of base replacement. *Nature*. 245:232–233.
- Sankoff D. 1975. Minimal mutation trees of sequences. *SIAM J. Appl. Math.* 28:35–42.
- Sankoff D.D., Rousseau P. 1975. Locating the vertices of a Steiner tree in arbitrary space. *Math. Program.* 9:240–246.
- Sankoff D.D., Cedergren R.J. 1983. Simultaneous comparison of three or more sequences related by a tree. In: Sankoff D.D. and Kruskal J.B., editors. *Time warps, string edits, and macromolecules: the theory and practice of sequence comparison*. Reading, MA: Addison-Wesley. p. 253–264.
- Smith G.R. 1990. Homology in morphometrics and phylogenetics. In: Rohlf F.J., Bookstein F.L., editors. *Proceedings of the Michigan morphometrics workshop*. Museum of Zoology, University of Michigan. Special Publication 2.
- Siegel A., Benson R. 1982. A robust comparison of biological shapes. *Biometrics*. 38:341–350.
- Schwikowski B., Vingron M. 1997. The deferred path heuristic for the generalized tree alignment problem. *J. Comput. Biol.* 4:415–431.
- Slice D.E. 1996. Three-dimensional generalized resistant fitting and the comparison of least-squares and resistant-fit residuals. In: Marcus L.F., Corti M., Loy A., Naylor G., Slice D.E., editors. *Advances in morphometrics*. New York: Plenum. p. 179–199.
- Slice D.E. 2005. Modern morphometrics. In: Slice D.E., editor. *Modern morphometrics in physical anthropology*. New York: Kluwer. p. 1–45.
- Stone J.R. 2003. Mapping cladograms into morphospaces. *Acta Zool.* 84:63–68.
- Varón A., Vinh L.S., Wheeler W.C. 2010. POY version 4: phylogenetic analysis using dynamic homologies. *Cladistics*. 26:72–85.
- Wheeler W.C. 1996. Optimization alignment: the end of multiple sequence alignment in phylogenetics? *Cladistics*. 12:1–9.
- 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., Aagesen L., Arango C.P., Faivovich J., Grant T., D'Haese C., Janies D., Smith W.L., Varón A., Giribet G. 2006. Dynamic homology and phylogenetic systematics: a unified approach using POY. New York: American Museum of Natural History. p. 365.
- Zelditch M.L., Fink W.L., Swiderski D. 1995. Morphometrics, homology, and phylogenetics: quantified characters as synapomorphies. *Syst. Biol.* 44:179–189.
- Zelditch M.L., Fink W.L., Swiderski D.L., Lundrigan B.L. 1998. On applications of geometric morphometrics to studies of ontogeny and phylogeny: a reply to Rohlf. *Syst. Biol.* 47:159–167.