WILEY-BLACKWELL

Cladistics (2019) 1–17

Cladistics

10.1111/cla.12373

PASOS: a method for the phylogenetic analysis of shape ontogenies

Santiago A. Catalano[a,b]*, Valentina Segura[a] and Florencia Vera Candioti[a]

[a]*Unidad Ejecutora Lillo, Consejo Nacional de Investigaciones Científicas y Técnicas – Fundación Miguel Lillo, Miguel Lillo 251, 4000 S. M. de Tucumán, Argentina;* [b]*Facultad de Ciencias Naturales e Instituto Miguel Lillo, Universidad Nacional de Tucumán, Miguel Lillo 205, 4000 S. M. de Tucumán, Argentina*

Accepted 21 November 2018

Abstract

We present a novel phylogenetic approach to infer ancestral ontogenies of shape characters described as landmark configurations. The method is rooted in previously published theoretical developments to analyse landmark data in a phylogenetic context with parsimony as the optimality criterion, in this case using the minimization of differences in landmark position to define not only ancestral shapes but also the changes in developmental timing between ancestor–descendant shape ontogenies. Evolutionary changes along the tree represent changes in relative developmental timing between ontogenetic trajectories (possible heterochronic events) and changes in shape within each stage. The method requires the user to determine the shape of the specimens between two standard events, for instance birth and onset of sexual maturity. Once the ontogenetic trajectory is discretized into a series of consecutive stages, the method enables the user to identify changes in developmental timing associated with changes in the offset and/or onset of the shape ontogenetic trajectories. The method is implemented in a C language program called SPASOS. The analysis of two empirical examples (anurans and felids) using this novel method yielded results in agreement with previous hypotheses about shape evolution in these groups based on non-phylogenetic analyses.

> The best measure of paedomorphosis is the extent to which an adult descendant resembles an ancestral juvenile. S.J. Gould (1977, p. 387)

Introduction

The analysis of morphological evolution has witnessed a revolution with the advent of geometric morphometric tools (Rohlf and Marcus, 1993). In particular, methods based on the analysis of landmark data are being used to study shape changes at different scales, ranging from the analysis of variation at individual and intraspecific scales (e.g. Savriama et al., 2012) to the study of shape evolution among species or lineages (e.g. Dosik and Stayton, 2016). Analyses of shape change at a macroevolutionary scale were prompted by the development of new methodological approaches for the phylogenetic analysis of geometric morphometric data (González-José et al., 2008; Catalano et al., 2010; Klingenberg and Gidaszewski, 2010; Catalano and Goloboff, 2012; Smith and Hendricks, 2013). Phylogenetic analyses include studies that address different questions related to functional issues, convergent evolution and covariation, among others (e.g. Figueirido et al., 2010; Vera Candioti and Altig, 2010; Houle et al., 2017; Ospina-Garcés and De Luna, 2017). Morphogeometric approaches are being increasingly used in the study of ontogeny and help to compare models of allometric and heterochronic development in several groups (e.g. Ivanović et al., 2007; Rodríguez-Mendoza et al., 2011; Ponssa and Vera Candioti, 2012; Mitteroecker et al., 2013). The study of ontogenies and the way they intertwine with patterns of diversity and evolutionary changes can be traced far back in the history of biological study

*Corresponding author:
E-mail address: sacatalano@gmail.com

(Gould, 1977). Although there is an increasing focus on proximal causes of ontogenetic variations (mediated by events at molecular and genetic levels; Keyte and Smith, 2014), as stressed by Webster and Zelditch (2005), a morphological, specimen-based approach is still useful to explore and interpret the relationships between organisms and the surrounding environment, and to interpret modifications of morphological development in terms of shared changes in life history. Accordingly, while it is clear that the most appropriate context to analyse ontogenetic changes at the interspecific level is phylogenetics (Fink, 1982; Boughton et al., 1991), with a few exceptions (Foth et al., 2016; Bardin et al., 2017) the evolution of ontogenetic trajectories of shape characters is analysed without considering the phylogenetic relationships. Evolutionary changes that have accumulated from the most recent common ancestor (MRCA) of two species are inferred by making pairwise comparisons between species. This approach is limited because it is not possible to either (1) know which species changes (perhaps both species) or (2) analyse evolutionary changes that occurred among groups of species (i.e. clades).

Given that the shape of an organism changes during its lifespan, the evolution of shapes in a phylogenetic context is analysed (Rüber and Adams, 2001; Rohlf, 2002; Catalano et al., 2010; Klingenberg and Gidaszewski, 2010) by recording the shape of the different species at some standard stage, generally the adult stage. However, this standardization is only valid under the assumption of complete synchronization between the timing of shape development and that of sexual maturation. If a heterochronic event has occurred, such synchronization is lost, and sexual maturity cannot be used as a guide to define standard stages and ancestral shapes are not properly inferred. Let us consider a real case: the evolution of the skull shape in *Ambystoma*, a genus that includes the axolotl (*A. mexicanum*). The axolotl is one of the most famous cases of neoteny in vertebrates, with adults being sexually mature individuals with larval morphology. This example has been previously considered in the literature to address conceptual issues associated with heterochrony (Alberch et al., 1979; Reilly et al., 1997). Proper inference of the adult cranial shape of the MRCA of the axolotl and its sister species would require considering the shape of the fully developed axolotl, a shape that is missing given the truncation of its ontogenetic trajectory (Fig. 1). Using the shape of the sexually mature axolotl to infer the ancestral shape of the adult of the MRCA would be incorrect: it is the equivalent to, in a group with no heterochronic events, inferring the ancestral adult shape based on specimens of different ages in different species (e.g. juveniles in some species, adults in others). In fact, the shape of the sexually mature axolotl should be considered to infer the ancestral shape of the premetamorphic stages of the ancestor, but not of fully developed individuals.

The axolotl example presented above stresses the importance of including the ontogenetic perspective when inferring ancestral shapes. However, unlike in the axolotl case, in most real cases it will not be possible to determine *a priori* the existence of changes in developmental timing. Hence, it is necessary to have a method that allows inference of ancestral shapes and changes in developmental timing simultaneously. Extending the approaches developed for the inference of ancestral shapes in a single ontogenetic stage (e.g. Catalano et al., 2010; Klingenberg and Gidaszewski, 2010) to the complete organism's ontogeny would involve recording the shape at different moments, or "stages" of the ontogeny and then determining the ancestral shapes in each of these stages. This approach requires defining *a priori* the correspondence among stages in the different species analysed. However, as indicated in the axolotl example, changes in developmental timing (i.e. changes in the offset, onset or rate of development) produce truncations, extensions and/or lateral transposition of the ontogenies that alter the original frame of comparison. Hence, a method to infer ancestral shape ontogenies in the presence of changes in developmental timing should determine the correct matching among shape ontogenies. If the only change between two shape ontogenies is in timing, the optimal matching (i.e. alignment) can be determined, among all possible matchings, as the one that implies no change in shape along the ontogeny (i.e. maximizing the similarity in shapes). This same reasoning can be applied to cases in which, in addition to changes in developmental timing, other modes of ontogenetic changes have occurred (see Webster and Zelditch, 2005). In that case, among all possible matchings, the one that implies the highest similarity in shape through the ontogeny would be optimal.

To infer changes in shape along a tree in the presence of changes in developmental timing, the same criterion should be used to determine the ancestral shapes and the matching among shape ontogenies. In phylogenetic morphometrics (PM; Catalano et al., 2010), the ancestral landmark configurations are those that minimize the sum of linear (Euclidean) distances between homologous landmarks. This metric ensures extension of the parsimony criterion to characters that change in more than one dimension and is not affected by the shape orientation (i.e. the results are rotation-invariant; Catalano and Goloboff, 2012). The method presented here (PASOS: phylogenetic analysis of shape ontogenies) extends this framework by choosing not only the ancestral configurations but also the pairings between shape ontogenies that minimize the sum of linear distances between homologous landmarks. In doing so, the similarity in shape that

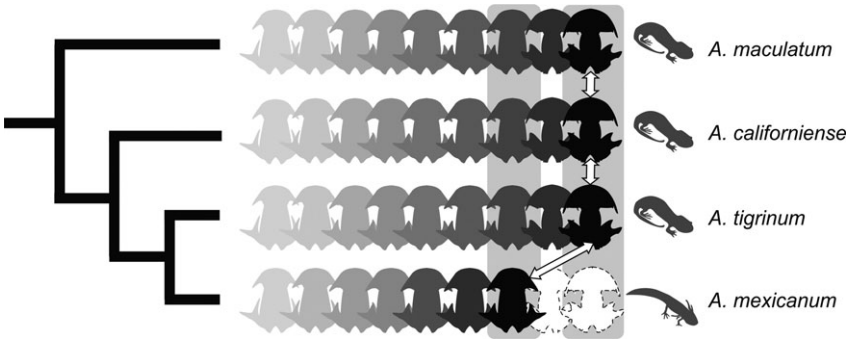

Fig. 1. A simplified real example showing the importance of including the ontogenetic dimension when inferring ancestral shapes. The example shows a representation of the skull shape ontogenies in species of *Ambystoma*, a genus that includes the neotenic species *A. mexicanum* (axolotl). Standard inference of ancestral shapes (arrows) imply considering adults (i.e. sexually mature specimens, black skulls) of all the species. However, given the change in developmental timing occurring in the branch leading to the axolotl, the adult specimens of this species should not be matched with the sexually mature specimens of the metamorphosing species but with premetamorphic specimens. As the fully developed stage of the axolotl is missing given the truncation of the ontogeny, its shape cannot be used to infer the adult shape of the ancestors. Grey rectangles indicate proper matching among shape ontogenies. See text for further explanation. Phylogenetic relationships follow Williams et al. (2013).

can be accounted for by common ancestry is maximized. An alternative, but not mutually exclusive, justification for the method is that it chooses the simplest explanation of the observed pattern: in the case of Fig. 2, it would be preferable to explain the differences in shape along the branch that leads to species A and B by a single change in developmental timing rather than considering changes in shape occurring at every stage. Despite the simplicity of the logic behind PASOS, there are several methodological challenges that complicate its implementation in a phylogenetic context. In the following sections, we describe the method, present two empirical examples and discuss the advantages and limitations of the approach.

Among the many different evolutionary changes that affect ontogenetic trajectories as a whole, PASOS determines only changes in developmental timing. This is because the modifications in developmental timing are the only changes that affect the matching among shape ontogenies and, consequently, affect the inference of

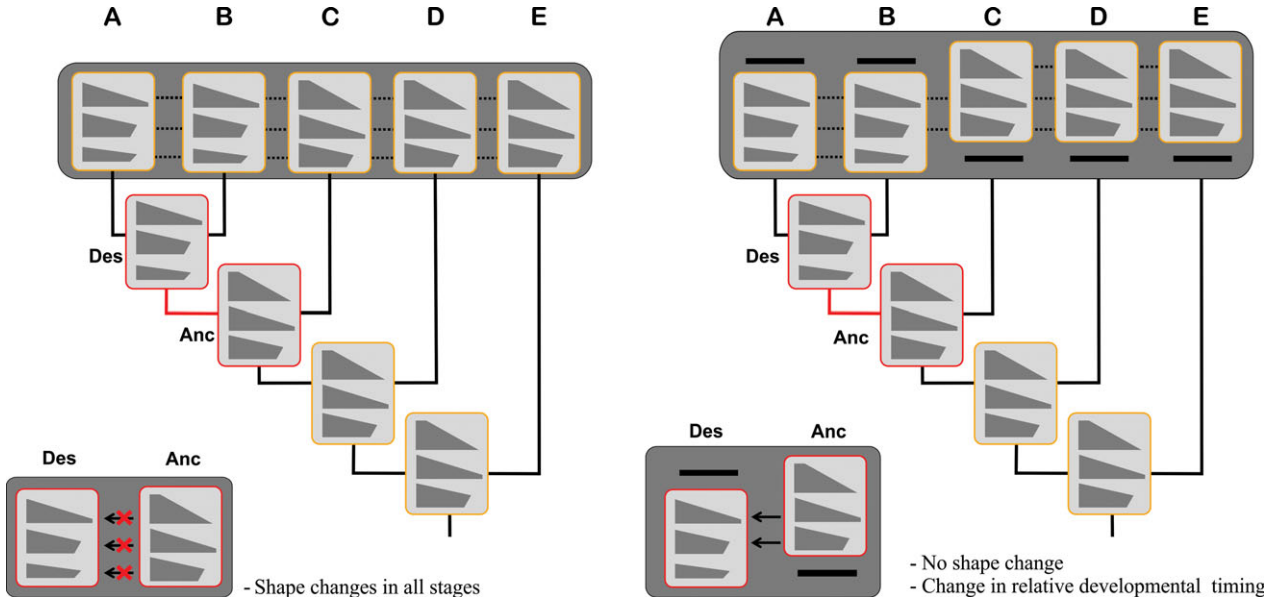

Fig. 2. Schematic representation of the rationale behind PASOS to infer ancestral shape ontogenies. In this example, the shape in each species is recorded at three different moments of the ontogeny. Left: inference of shape change considering the original frame of comparison (i.e. no change in developmental timing); changes in shape are inferred in all stages (Crosses). Right: inference of shape change considering a shift in developmental timing in the branch leading to the clade (A B); a single change in developmental timing can explain all the differences in shape between ancestor–descendant nodes. The hypothesis on the right is preferred because it is the most parsimonious explanation of the observations. Anc, ancestor; Des, descendant.

ancestral shapes. Other transformations, such as those that produce changes in the direction of the trajectories in shape space (Mitteroecker et al., 2005), do not modify the matching (e.g. the adult shape inferred at the internal nodes of the tree should be determined by comparing the adult shapes of all species analysed) and hence these changes are not inferred during the procedure. Once matching among ontogenies is established along the tree, and the optimization of shape along the trajectory is determined, the changes in the direction of the ontogenetic trajectory, as well as other changes in shape at any part of the trajectory, can be inferred by analysing the ancestral shape ontogenies using pairwise approaches (e.g. Sheets and Zelditch, 2013).

Algorithmic approach

The method starts with the description of the ontogeny of the shape under study as a set of landmark configurations recorded at different moments. The ontogeny is delimited using start and end points that are common to all species under analysis (double stage standardization; Alba, 2002), for instance, birth and onset of sexual maturity, respectively. Within that span, the method can handle data classified in *a priori* defined/tabulated stages (stages as proxy of age, such as dental ages) or with age measured on a continuous scale (e.g. hours, days or with size as a proxy of age). Specimens from all species included in the analysis should be superimposed, for instance by using a generalized Procrustes analysis (GPA) procedure (Gower, 1975; Rohlf and Slice, 1990). When the configurations are recorded from tabulated stages, the shape at each stage is established as the consensus of all configurations belonging to that stage. When age is represented in a continuous scale, the ontogeny is first discretized into a series of consecutive stages or bins, with the shape of each stage being represented by the consensus configuration of all specimens at that stage. In this case, the term "stage" is used for each of the time/age/size bins on which the ontogenetic trajectory is discretized.

The number of stages is set in PASOS as the highest possible without missing data for any stage/species. The stages can be considered as *a priori* comparable (homologous under some definitions) shape stages, and such comparability will be tested in the analysis. Note that when dealing with time recorded on a continuous scale the definition of stages is only operational—it is only an arbitrary division of a continuum. Once the ontogeny of each species is described as indicated above, and given a tree showing relationships among the species under analysis, the method selects the ancestral shapes and the pairing (matching) between ontogenetic trajectories that minimize both the difference in landmark positions and the events of change in developmental timing. Hence, the changes in shape inferred on a branch can be explained by (1) changes in developmental timing that involve a new matching between the ancestral/descendant trajectories, and/or (2) changes in one or more stages. In order to describe how the method determines the existence of changes in developmental timing, we present the simplest case concerning two terminals. In the following section, we then describe how this approach is extended to a phylogenetic context.

Quantifying changes in ontogenetic trajectories of shape characters

Starting from two discretized ontogenetic trajectories, changes in developmental timing are determined by comparing the amount of shape change implied by different pairings between the trajectories (Fig. 3).

Shape change is quantified as the sum of linear differences between the positions of all landmarks for each comparable stage (i.e. using in each stage the same metric as in PM). The pairing that produces the smallest shape change is considered the optimal matching between shape ontogenies. No change in developmental timing is inferred when the lowest score is obtained

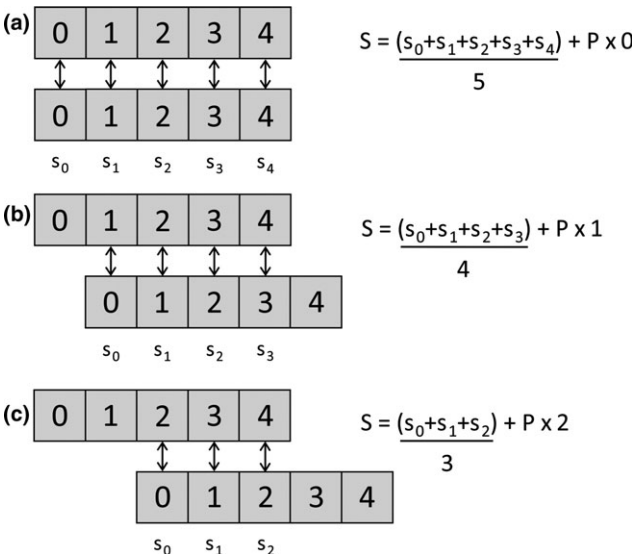

Fig. 3. Determining the optimal pairing between ontogenetic trajectories in a two-species case. Each ontogenetic trajectory is described as a series of consecutive stages, with the shape at each stage being represented by the consensus landmark configuration of all individuals corresponding to that stage. The score (S) is determined as the sum of linear distances between corresponding landmarks for each stage involved in the comparison (s_n) divided by the number of stages being compared, plus a penalty that is a function of the unitary penalty cost (P) and the number of stages shifted. The optimal pairing is that with the lowest score. The example shows three of all possible matchings: (a) no change in developmental timing, (b) shift of one stage and (c) shift of two stages. Numbers 0–4 indicate *a priori* comparable stages.

considering the original frame of comparison. Because the number of comparable stages varies in the different matchings, it is not possible to use the sum of changes through the different stages to choose the optimal matching, as using this procedure in most cases would lead to the selection of the matching that implies the lowest number of comparable stages.

Hence, the score is divided by the number of stages compared (Fig. 3), with the optimal matching being the one that produces the lowest average difference in shape between the trajectories. The number of stages into which the trajectory is divided depends on the number and distribution of the individuals sampled along the trajectory. Hence, having a good representation of specimens at different moments of the ontogeny is critical to obtain reliable results. Other approaches to analysing ontogenetic trajectories are also affected by individual sampling mainly at the extremes of the trajectory (Klingenberg, 2016).

The method evaluates different changes in developmental timing between ontogenetic trajectories. One of these changes involves a complete displacement of one of the ontogenetic trajectories (shifts, as the examples in Figs 2 and 3). This ontogenetic change is analogous to what is generally referred to as lateral transposition and pre-/post-formation (Klingenberg, 1998; Alba, 2002). When younger stages of the descendant are matched with older stages of the ancestors, by convention the shift is considered in PASOS as positive. The changes in developmental timing are in terms of stages (shifts or extensions of multiples of one stage). Hence, the higher the number of stages into which the ontogenetic trajectory can be divided (given a denser sampling of specimens), the subtler the changes that can be inferred.

In addition to the shifts between ontogenetic trajectories, the method infers changes in the offset and/or offset that produce extensions or contractions of the ontogenetic trajectories (Fig. 4). This is the sort of change that occurred in the case of the axolotl (Fig. 1), where the offset was modified but the onset was not. These changes are referred to in PASOS as stretches and can involve a change in one end-point (stretch-start or stretch-end) or both end-points of the ontogeny (double stretch). From a biological point of view, the stretches imply a modification in the rate of shape change in terms of the lifespan determined by both end-points. Whether this change represents a change in absolute (external) time can be determined *a posteriori* (see Discussion).

In the case of shifts, because the matching of each stage is modified by the same amount through the ontogenetic trajectory, the score is calculated considering the consensus configurations that represent the shapes of the original stages. In the case of stretches, because the extension of one of the ontogenetic

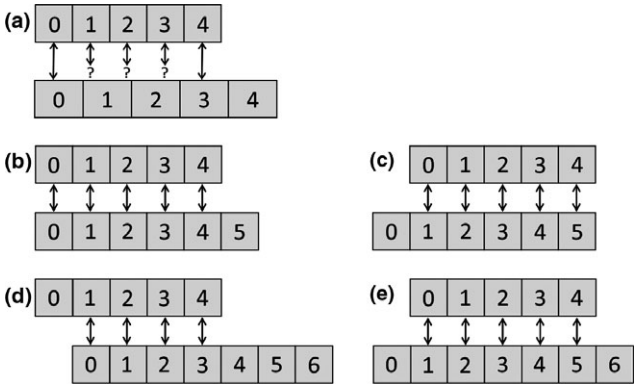

Fig. 4. Calculating the cost of changes in ontogenetic trajectories that represent extensions and contractions (stretches). These changes are product of changes in offset, in onset or in both. (a) Differing from the case of complete displacements of ontogenetic trajectories (i.e. shifts), in the case of stretches the scores cannot be directly calculated considering the original discretization stages because the one-to-one matching between stages is lost. To calculate the score it is necessary to sort the individuals into newly defined stages. Consequently, stretches can only be calculated in the case of age defined on a continuous scale. (b) The score of an extension of one stage (stretch-end +1) at the end of the descendant trajectory is calculated by dividing the trajectory into six stages and comparing the shape of only the first five. (c) Extension at the beginning of the descendant trajectory (stretch-start −1). Modifications at both extremes of the ontogenetic trajectories, (d) double-stretch +1 + 3, and (e) Double-stretch −1 + 1.

trajectories is modified, the shape at each stage should be redefined (Fig. 4). This requires grouping the individuals into new stages and re-calculating the consensus configurations for each stage. For instance, given an ontogenetic trajectory divided into five stages, to calculate the score of an extension that represents one stage, the trajectory that is modified (descendant) is divided into six stages, and only the first five stages of both trajectories are compared. This re-grouping will only be possible when the age of the individuals is measured on a continuous scale and the number of sampled individuals is enough to include at least one individual per newly defined stage. When the data are originally sorted into discrete stages, it is not possible to rearrange the configurations into new stages. In that case, only shifts can be inferred. The sign convention for stretches is the same as in shifts.

Phylogenetic method

The approach to infer changes in development timing between shape ontogenies previously described is the basis of the phylogenetic method. As implemented, PASOS has three steps. First, changes in developmental timing and preliminary ancestral shapes are inferred on the tree. Second, considering the changes in developmental timing inferred in the first step, a multiple alignment among the ontogenetic trajectories

is generated. Finally, the assignment of ancestral shapes is improved by means of a PM optimization stage-by-stage. This method is implemented in SPASOS (software for the PASOS).

Step 1: determining changes in developmental timing and preliminary ancestral shapes. The approach uses a fixed states (FS) optimization to infer changes in developmental timing and preliminary ancestral shapes for all nodes of the reference tree. FS is an approach for the phylogenetic analysis of unaligned DNA sequences developed by Wheeler (1999). FS considers each observed DNA sequence as a possible state. A cost matrix is built with the editing cost between sequences. Once the costs are calculated, the Sankoff algorithm (Sankoff, 1975) is used to optimize characters on the tree. In PASOS, this algorithm is modified to work with ontogenetic trajectories. The observed trajectories (terminals) are the possible states of the Sankoff character. The transformation costs between states are defined as the minimum cost for all possible matchings between the two ontogenies, as indicated in the previous section (Figs 3 and 4). In addition, the optimal matching for each pair of ontogenetic trajectories is kept in memory (optimal-match table). Once the cost is calculated for every pair of terminals (i.e. the cost matrix is complete) the next step is to assign the optimal state(s) to internal nodes using the Sankoff algorithm. Given the optimal states defined in the Sankoff optimization, changes in developmental timing are determined by looking up the optimal-match table (Fig. 5).

Step 2: determining the multiple alignment among ontogenetic trajectories. The changes in developmental timing inferred along the tree in the first step are considered to determine an implied alignment (Wheeler, 2003a) among the ontogenetic trajectories of ancestral and terminal nodes (Fig. 5). The procedure to establish the implied alignment is first described for the case where only shifts were inferred in the first step. Below we show how the method proceeds when changes that modify the extension of the trajectories are inferred.

When only shifts were inferred on the tree, the procedure visits the nodes of the tree from the root to the tips (i.e. preorder traversal). When the procedure reaches a branch where a change in developmental timing was inferred in the first step of PASOS, the limit(s) of the trajectory of that node and all descending nodes, both internal and terminal, are modified accordingly. For instance, if there was a shift of one stage, the upper and lower limits of all descending nodes are modified by one stage. Once all nodes are visited, the lower and upper limits of each ontogenetic trajectory assigned to the terminals have been updated; thus, a new alignment is defined among the ontogenetic trajectories (Fig. 5).

Dealing with extensions/contractions of the ontogenetic trajectories while building the implied alignment has some differences from the case when only shifts are inferred. In PASOS, all changes in developmental timing are in terms of stages (shifts or extensions of multiples of one stage). However, this is not necessarily the case when building the implied alignment. Consider the case of a trajectory originally discretized into five stages. When starting the procedure, the root will have five stages. During tree traversal, if the procedure reaches a branch where there is a change of one stage in the extension of the trajectory (i.e. 20% of the total span of the trajectory), all descendant nodes will have a span of six stages. If the procedure continues and finds a descending branch where there is a second extension of one stage, this again represents an extension of 20% of the original trajectory. Therefore, the

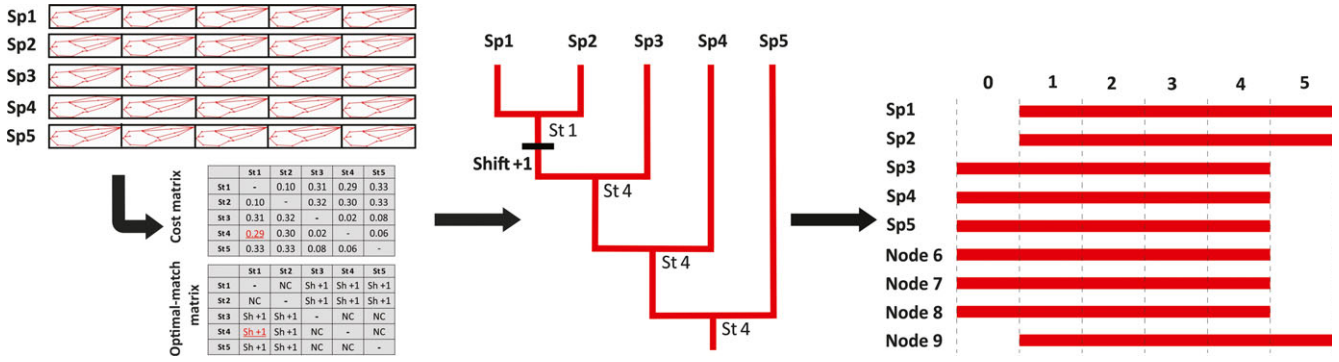

Fig. 5. Flowchart showing the first two steps of PASOS. First, the optimal matching between each pair of observed trajectories is calculated as in Figs 3 and 4. A Sankoff matrix is built with the best score for each pair of observed trajectories. In addition, a lookup table is built considering the optimal matching implied in each comparison. The cost matrix is used to optimize the character, determining the preliminary shapes for internal nodes. Changes in developmental timing are established by extracting the information from the look-up table that stores the optimal matchings. Once the changes in developmental timing along the tree are inferred, the tree is traversed from the root to the tips to generate the implied alignment (see text for details). St, state; Sh, shift; NC, no change in developmental timing.

trajectory is extended in 6 × 0.2 = 1.2 stages and not in one stage as in the case of the first change.

Once the procedure is finished, the relative position of all the trajectories is defined and a new alignment among trajectories is established. In the case that only shifts were inferred by FS (step 1), the trajectories in the terminal nodes are not modified (i.e. each node maintains the number of stages and corresponding shape, with only its relative position being changed). However, when the inferred change affects the extension of the trajectories, the number of stages of some terminal nodes may differ from the original number, which makes it necessary to regroup the individual in the new stages. Given that, the changes may imply fractions of stages (see above), and the limits of the trajectories are rounded to the nearest whole number.

Step 3: improving ancestral shape inference. After step 2, the ancestral shapes are those determined in the first step of the procedure and represent shapes of terminal taxa (the possible states in FS optimization). Once the alignment among trajectories is established, it is possible to improve the ancestral shape assignments (Fig. 6; improvement in terms of the optimality criterion considered in this method) using PM on each of the aligned stages and superimposing the configurations using the tree as a guide (Catalano and Goloboff, 2012). Note that the number of stages of the implied alignment may in general be higher than the span of each individual trajectory.

FS optimization may determine that there is ambiguity in the assignment of optimal states for one or more nodes of the tree. Given that the ancestral shapes will be updated in the following steps of PASOS, the only ambiguity that is important to take into account in the first step of the method is the one that produces different inference of changes in developmental timing. This will concomitantly affect the inference of ancestral shapes determined from the implied alignment. Given that the costs are expressed as floating point numbers, it is very unlikely that PASOS will produce ambiguous assignments. However, if there is ambiguity, SPASOS will indicate the node(s) where it occurs. A common cause of ambiguity occurs when there is a change in developmental timing at the base of the tree. This ambiguity precludes identifying the immediate descendant from the root where the change occurred. SPASOS always resolves this ambiguity by assigning the change to the branch that subtends a higher number of terminals (assuming that the other branch is the outgroup).

As currently implemented, PASOS does not have a global optimality criterion. The FS step (step 1) has an optimality criterion, but the possible states are restricted to the observed shape ontogenies. A globally optimal solution would only be obtained by this procedure in the hypothetical situation of including all possible ontogenetic trajectories. If only shifts are considered, the score can be improved by visiting each node and modifying the alignment between the target node and the neighbouring nodes [as is done in the approach proposed by Catalano and Goloboff (2012) to superimpose landmark configurations]. However, when stretches are inferred this procedure cannot be followed. This is because, at internal nodes, there are no observed specimens to regroup and redefine the stages.

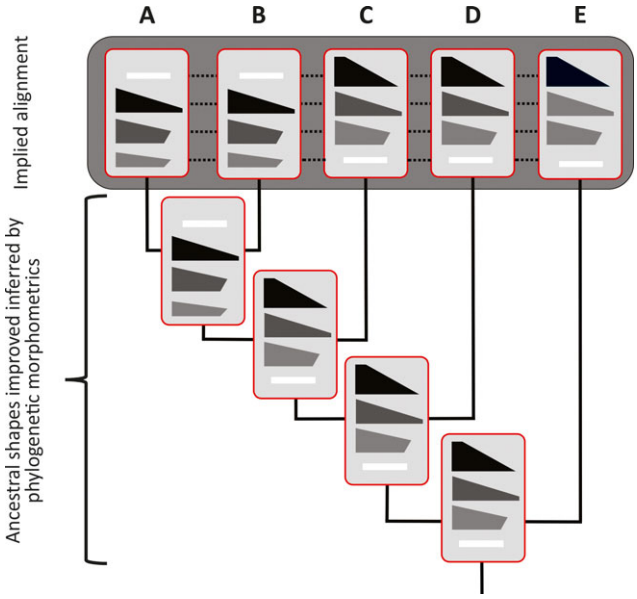

Fig. 6. Improvement of ancestral shape ontogenies by optimizing landmark configurations stage by stage departing from the implied alignment (Fig. 5). The implied alignment defined in SPASOS is entered into TNT to infer the ancestral shapes stage by stage using the phylogenetic morphometrics approach (Catalano et al., 2010; Catalano and Goloboff, 2012). White bars at internal nodes indicate shapes inferred by TNT but that are not part of the ancestral trajectory. Colour in shapes represents stages that were *a priori* considered as comparable.

Penalty for changes in developmental timing

Algorithms for the alignment of molecular sequences (e.g. Needleman and Wunsch, 1970; Smith and Waterman, 1981) depend on defining a penalty cost for the occurrence of insertions and deletions. PASOS differs from those approaches in that it does not depend on assigning a penalty value. This is because the optimal matching is defined as the one that produces the smallest averaged shape change considering all the stages compared and not the absolute amount of change, which in most cases will decrease when fewer stages are compared. However, our approach allows the user to add a penalty cost for changes in developmental timing. This penalty is included as a way to make the inference of changes in developmental timing more

conservative (only matches that clearly improve the score are considered as changes in developmental timing). In addition, it helps to consider changes in developmental timing as synapomorphies on the tree.

Given that there is no objective way to determine the gap penalty value, molecular-alignment programs usually set this value empirically. In the case of shape ontogenies there is an extra issue: the scores differ concomitantly with the scale of the configurations. Therefore, this penalty should be related to the scale of the configurations. Although there is no exact mode to determine the "correct" penalty cost (but see Felsenstein, 2004; De Laet, 2005), there are some logical limits to the values that can be given to the penalty. The maximum value can be defined considering that the penalty should never be as high as to prohibit a change in developmental timing that produces a perfect match between two trajectories (i.e. no change in shape in aligned stages). For a given dataset, this value can be approximated by calculating the cost of comparing each ontogenetic trajectory with the same trajectory but shifted one stage. This is an overly restrictive penalty, because the only ontogenetic changes that would be inferred are those that imply no modification in shape. Consequently, in SPASOS, the penalty is calculated as a fraction of this cost. Details of how this penalty is calculated are included in the Appendix.

Assessing support for hypotheses of changes in developmental timing

Evaluating the extent to which the results are supported by the evidence is crucial in any phylogenetic method. The extent to which the results are supported is evaluated in PASOS by two different approaches that in turn consider different facets of support. One approach evaluates the effect of the sampling of individuals within each species while the other evaluates the relative evidence that supports and contradicts the inferred pattern.

In the presence of high intraspecific variation, a limited sampling of individuals may result in the shapes included in the analysis being poor representatives of the shape of each species. In the geometric morphometrics literature, a common practice to evaluate the effect of intraspecific variation in the parameters estimated is to use a resampling procedure (Cardini et al., 2015). This procedure differs from those used in a phylogenetic framework, where characters and not individuals are resampled (Felsenstein, 1985; Farris et al., 1996; Goloboff et al., 2003). In each replicate of the resampling procedure implemented in PASOS, individuals from each species are randomly removed with a probability of 0.25. The procedure is performed with the control that at least one specimen per species remains in the first and last stage. Otherwise, the resampling may produce an artificial modification of the limits of the trajectory. This procedure is repeated 100 times. Resampling support is indicated as the percentage of replicates that obtained the same change in developmental timing (both the type and the extent of change) than that inferred considering the original matrix.

To evaluate the evidence that supports/contradicts each inference of change in developmental timing, SPASOS implements a decay approach analogous to the Bremer clade support (Bremer, 1994), by repeating the analysis considering increasing penalty values. The logic behind this approach is that changes in developmental timing that are less supported by the evidence will be lost with slight penalty-cost increases.

The software

The method presented here is implemented in SPASOS, a Windows command line program written in the C language. The executable is freely available at http://www.lillo.org.ar/phylogeny/spasos. The source code is deposited at https://github.com/sacatalano/SPASOS. The program performs all the analyses except for PM optimization, which is performed calling TNT (Goloboff et al., 2008; Goloboff and Catalano, 2016). Hence, SPASOS users need to download TNT as well, which is also freely available at http://www.lillo.org.ar/phylogeny/tnt.

Two empirical examples

Evolution of skull shape ontogeny in felids

This dataset corresponds to an ongoing analysis of the skull shape ontogeny in felids led by one of the coauthors (V. Segura, unpublished data). The analysis involves 645 individuals belonging to ten species of felids (*Leptailurus serval*, *Leopardus wiedii*, *L. pardalis*, *Lynx rufus*, *Puma concolor*, *Herpailurus yagouaroundi*, *Panthera pardus*, *P. leo*, *P. onca* and *P. tigris*). Cranial shape was described considering 66 three-dimensional (3D) landmarks (see Segura et al., 2013). The ontogenetic trajectory was delimited by dental eruption and tooth wear, from "babies" to "adult A1" stages following Segura (2015). The phylogenetic hypothesis was derived from Johnson et al. (2006). An alternative topology considering a different arrangement of *Panthera* species was also used to map ontogenetic trajectories with *P. pardus* resolved as sister to *P. leo* (Christiansen, 2008; Davis et al., 2010). Two analyses were conducted considering different proxies of age: skull size (log centroid size) and both dental eruption and tooth wear. In the latter analysis, the ontogeny

was discretized into six different *a priori* defined age classes: babies (B), juveniles (J1, J2, J3, J4) and adults (A1). Three species were not included in the former analysis because the sampling did not include individuals for all the categories (i.e. there were missing data). Specimens of all stages and species were superimposed by means of a GPA procedure (Gower, 1975; Rohlf and Slice, 1990).

The analysis based on size as proxy of age inferred six changes in developmental timing (Fig. 7), four of which were autapomorphies. Three changes represented shifts of the whole trajectory, whereas the remaining changes implied modifications at the start of the trajectories. One of the most important patterns found in the analysis was at the base of the tree. Because the pattern was at the basal node, and taking into account that the costs are symmetric, this change can be associated with the branch leading to either of the two sub-families (Pantherinae and Felinae). Because the aim of providing this example is to show how the method works, we arbitrarily assumed that the change occurred in the branch leading to Pantherinae. The change in this branch indicates that older individuals of the basal node (Node 8, Fig. 7) are more similar to younger individuals of the ancestor of Pantherinae (Node 14) than to individuals of the same *a priori* defined stage. This change is also inferred when the analysis is performed using the alternative topology (Johnson et al., 2006). However, while the resampling support is high (89%) when mapping on the topology of Johnson, the value was lower when *P. pardus* was placed as sister of *P. leo* + *P. onca* (data not shown). The change in this branch enabled us to illustrate how the inference of shape change differs when shapes are mapped in the original frame of comparison vs. the matching derived for inferring a change in developmental timing. Figure 7 shows the cranial shape of the ancestor of Pantherinae. Ancestral landmark configurations inferred by PASOS were used as the basis to generate this render. Following Muñoz et al. (2017), a lion skull (MACN-MA 26029) was digitized using a NextEngine Desktop 3-D Scanner, and a 3D surface mesh was generated. The 3D landmark coordinates were taken from the mesh using the Landmark editor software (Wiley, 2006). Visualization and graphics were made using the Morpho R package 2.5.1 (Schlager, 2017), which associates the colour pattern with shape changes. When the shape of the last stage at both nodes is compared, important changes are inferred (Fig. 7). However, when the comparison is made considering the matching derived from PASOS (older individuals of the ancestor compared with younger specimens of the descendant), much reduced shape change is observed (approximately 25% less). The change in the last stage considering the original matching between trajectories implies a narrowing of the braincase (visualized in blue), a phenomenon that is almost absent when a change in developmental timing is considered.

The analysis considering dental ages inferred three changes in developmental timing, two of them in agreement with those obtained using size as a time proxy (Fig. S1). One of the differences between analyses occurred on the branch leading to *L. wiedii* with a negative shift of one stage in the branch leading to this species. This is in agreement with a previous analysis that considered *L. wiedii* to be paedomorphic (Fagen and Wiley, 1978).

Evolution of body shape in foam frog embryos

This dataset comprises an ontogenetic series of seven species of foam frogs (*Physalaemus*; Anura: Leptodactylidae): *P. albonotatus*, *P. biligonigerus*, *P. cuqui*, *P. cuvieri*, *P. fernandezae*, *P. riograndensis* and *P. santafecinus*. This dataset is part of an ongoing project to analyse the evolution of shape ontogenies in this group led by one of the co-authors (F. Vera Candioti, unpublished data). The shapes were mapped onto a tree built following the relationships defined by Lourenço et al. (2015). The differentiation of the first gill bud was considered to be the onset of the trajectory and concealment of the right gill by the operculum (Stage 24 of Gosner, 1960) as the offset. This segment of the trajectory was selected because younger embryos have a pronounced kyphotic curvature that affects landmark superimposition, and older embryos show no evident shape change with increasing size and elapsed developmental time. The original matrix consists of photographs in lateral view of 346 embryos, in which 20 landmarks were digitized. Specimens of all stages and species were superimposed by means of a GPA procedure (Gower, 1975; Rohlf and Slice, 1990), and the centroid size was recorded for each specimen. PASOS analysis was conducted by discretizing the ontogenetic trajectories into five stages representing the shape at different relative sizes, with the default penalty value. A total of 100 replicates were performed in the resampling procedure.

The analysis recovered two changes in developmental timing (Fig. 8a). The first occurred in the basal dichotomy separating *P. fernandezae* (representative of the *P. henselii* intrageneric group) from the remaining species. Based on evidence from other studies (e.g. Lobo, 1996; Tomatis et al., 2009; Lourenço et al., 2015), we consider that the change occurred in the branch leading to *P. fernandezae*. The shift of the whole trajectory inferred at that branch implies that embryos of *P. fernandezae* are consistently paedomorphic with respect to embryos of the *P. biligonigerus* and *P. cuvieri* groups. At the end of the trajectory, concealment of the right gill in *P. fernandezae* co-occurs with body-shape features that are typical of

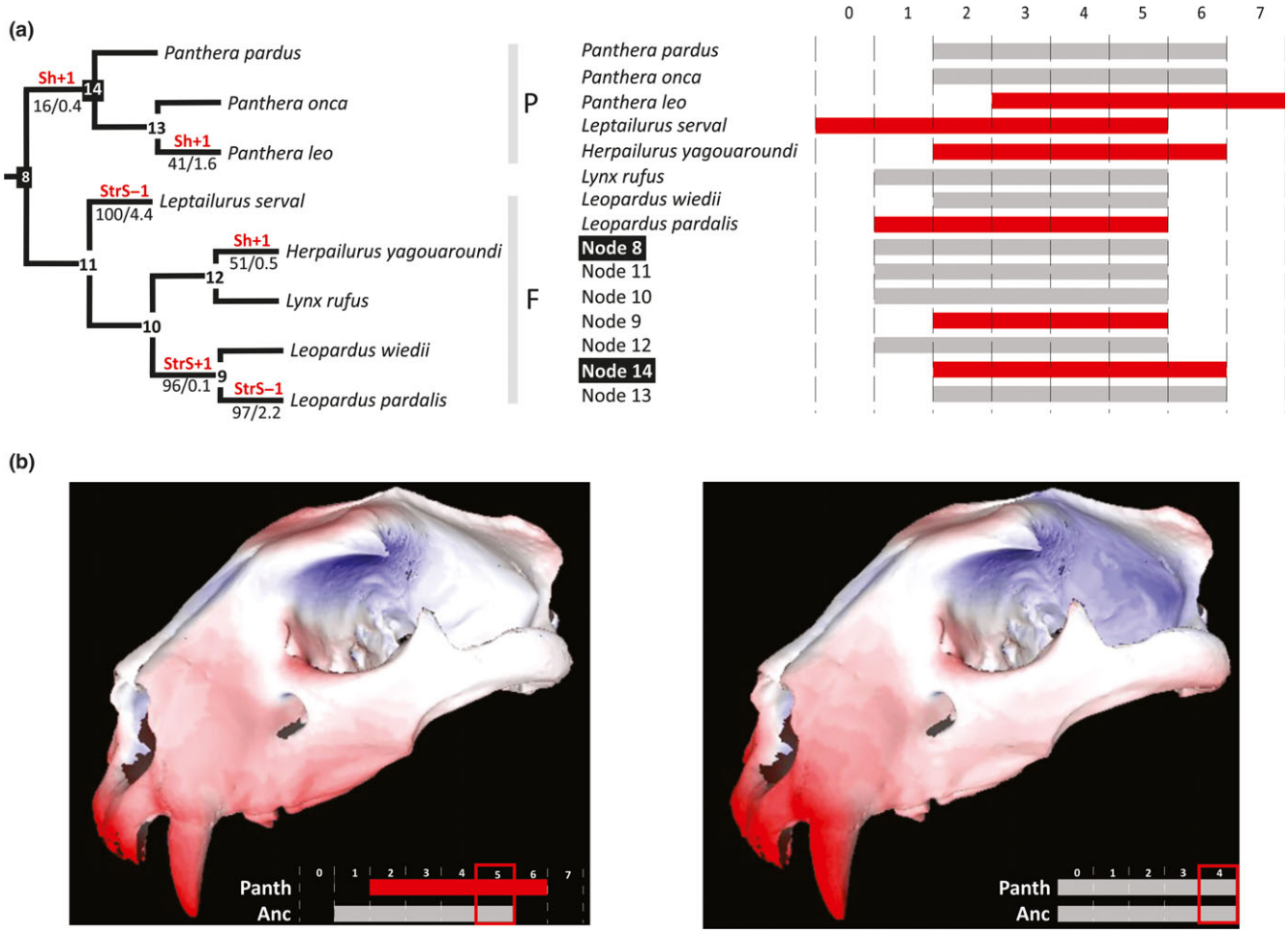

Fig. 7. Results of PASOS for the felid dataset considering logCS as proxy of age. (a) Left: changes in developmental timing inferred on the tree. Nodes below branches represent resampling/decay support values. Right: implied alignment of ontogenetic trajectories. Bars in red indicate the nodes where changes occurred. (b) Shape change in the branch leading to pantherines (Node 8–Node 14) when the optimization of shapes is performed considering the original frame of comparison (left) or when the comparison considers the change in developmental timing inferred by PASOS (right). The shape change is visualized for the last stage in the ancestor of the branch leading to pantherines (Node 8). Overall change is reduced when the pairing involves a change in developmental timing. Parts of the skull in red indicate regions that have been enlarged from the ancestor. Parts of the skull in blue indicate regions that have been reduced from the ancestor. Colour intensity is proportional to the amount of change.

early embryos, such as a shorter snout, higher adhesive glands and lower tail fins. These changes can be recognized by comparing the ancestral trajectory inferred considering a developmental change with the one derived from the original alignment among trajectories, where shape change is much pronounced around those regions (Fig. 8b). The *P. henselii* group has the southernmost distribution within the genus, and these species breed during the winter (Lourenço et al., 2015). Embryos of these species are among the smallest of the dataset analysed (F. Vera Candioti, unpublished data). Given the known effect of low temperatures on anuran growth (e.g. Derakhshan and Nokhbatolfoghahai, 2015), some delayed development was expected. Interestingly, while certain individual characters previously analysed showed a similar pattern (e.g. external gills in this group occur as two short, scarcely ramified pairs; Grosso, 2017), others exhibit a rather accelerated, peramorphic development (e.g. the oral apparatuses have the most complex configuration within the genus; Vera Candioti et al., 2011).

Phylogenetic analysis of shape ontogenies also inferred an extension in the trajectory on the branch subtending the clade *P. biligonigerus* + *P. santafecinus* (Fig. 8a), such that their offset shapes are peramorphic compared to those of their ancestor. The shape ontogeny at that node exhibits a synapomorphic displacement at the origin of the dorsal fin, a change that begins at the final stages of the trajectory (Fig. S2) and that is maintained in both species in this clade. This body-shape pattern fully agrees with what is known about oral apparatus ontogeny within the clade, with

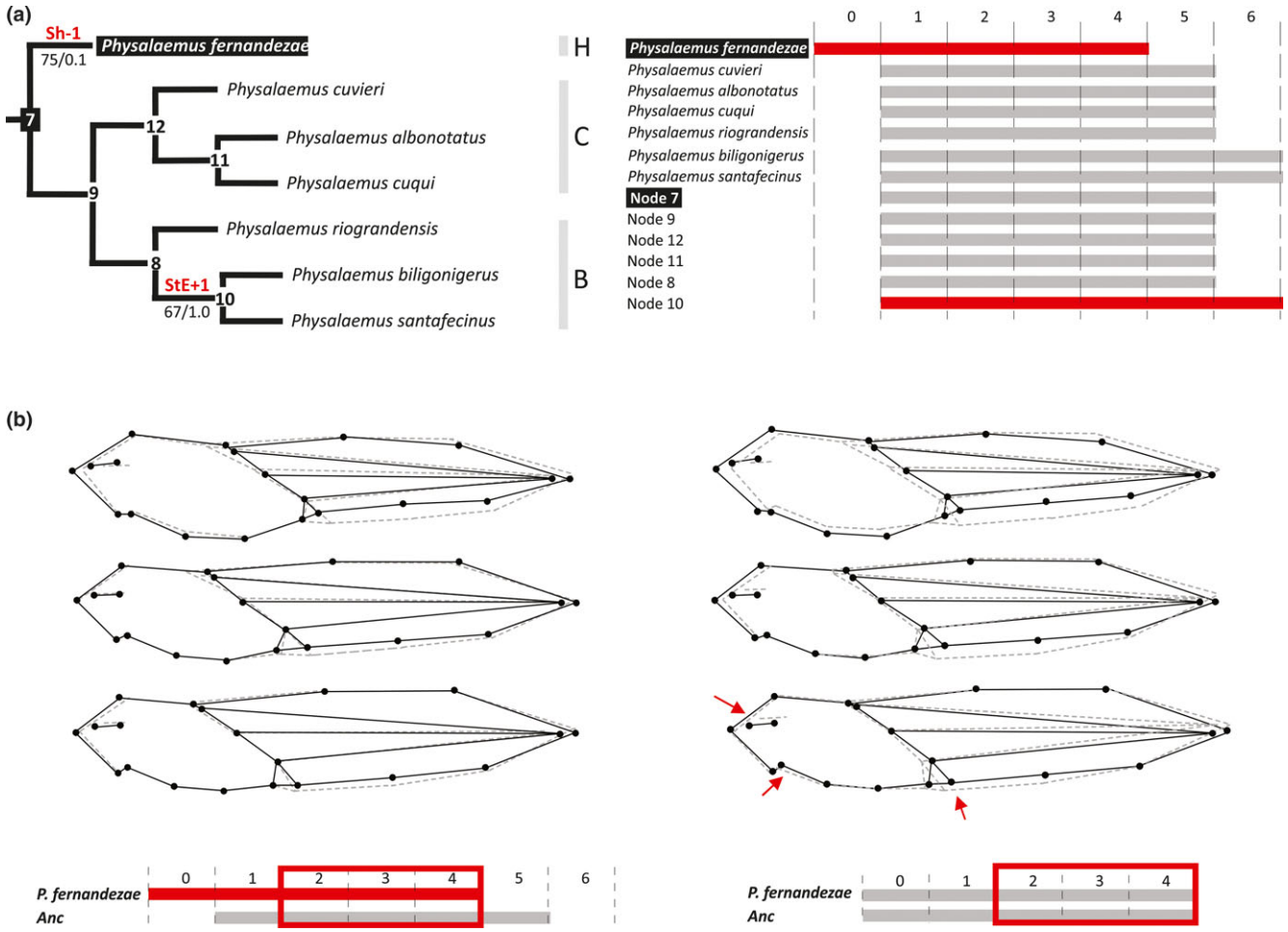

Fig. 8. Results of PASOS for the *Physalaemus* embryos dataset. (a) Left: changes in developmental timing mapped on the reference phylogeny. Numbers on branches represent resampling percentage and decay index. Two changes were inferred, a positive shift of the whole trajectory in the branch leading to *P. fernadezae* and an extension at the end of the trajectory in the branch leading to *P. biligonigerus* + *P. santafecinus*. These results are in agreement with previous evidence in the group under study. B = *P. biligonigerus* group; C = *P. cuvieri* group; H = *P. henselii* group. Right: implied alignment of ontogenetic trajectories. Nodes with changes in developmental timing are shown in red. (b) Comparison of the shape changes inferred in the branch leading to *P. fernandezae* considering the shift in the whole trajectory inferred in that branch (optimal hypothesis; left) vs. shape change inferred considering the original frame of comparison (right). Only the last three stages (where changes are more evident) are shown. Note the overall change being reduced across the whole trajectory in PASOS alignment. The arrows highlight areas with embryonic features in the offset shape in *P. fernandezae*, concentrated on the snout region, adhesive glands and ventral tail fin.

an extended oral trajectory in *P. biligonigerus* and *P. santafecinus* (Vera Candioti et al., 2011).

Discussion

The PASOS method proposed here is rooted in the theoretical framework previously developed for the analysis of landmark data in phylogenetics considering parsimony as the optimality criterion (Catalano et al., 2010; Goloboff and Catalano, 2011; Catalano and Goloboff, 2012). Other methods that analyse the evolution of ontogenies also use parsimony as their optimality criterion. Jeffery et al. (2002, 2005) proposed different approaches for the analysis of sequence heterochrony using parsimony to identify the minimum number of events required to explain all observed event-pair changes along any given branch of a phylogenetic tree. Schulmeister and Wheeler (2004) presented a parsimony method to analyse developmental sequences using a search-based character optimization (Wheeler, 2003b). More recently, Giannini (2014; see also Segura, 2014) proposed studying evolutionary allometry by mapping allometric coefficients as continuous characters with parsimony as the optimality criterion (Farris, 1970; Goloboff et al., 2006). The most recent approach that uses parsimony to analyse the evolution of ontogenies was proposed by Bardin et al. (2017), using the parameters of the fitting models as continuous characters that are subsequently optimized

on the tree. The general framework presented here can also be considered to develop an approach based on an explicitly statistical framework, for instance using squared-change parsimony (Maddison, 1991) or extending the Brownian motion model to deal not only with changes in shape but also the changes in developmental timing.

Heterochrony

In describing the PASOS method, we deliberately omitted any references to whether the changes in developmental timing inferred on the tree represent heterochronic changes or not. This omission is grounded on two factors. First, as stressed in several reviews, the concept of heterochrony has changed deeply since its original introduction by Haeckel (Klingenberg, 1998). Different definitions emphasize different aspects, such as changes in features that produce parallels between ontogeny and phylogeny, in developmental processes, and in developmental events (reviewed by Webster and Zelditch, 2005). In the context of geometric morphometrics, some current debates discuss the idea and evaluation of "shared shape ontogenetic trajectories" as a prerequisite to test the hypothesis of heterochrony (e.g. for an interesting counterpoint see Mitteroecker et al., 2005 and Lieberman et al., 2007). In this scenario, the concept is so versatile that it can be charged with either "explaining everything" or of being so restrictive that it could be impossible to identify in real data (e.g. McNamara, 1997; McNamara and McKinney, 2005; Lieberman et al., 2007). Second, while there is general consensus that approaches to the study of heterochrony should explicitly deal with time, it is also widely acknowledged that time data are hard to obtain and also hard to handle (e.g. measures of extrinsic, intrinsic, absolute, or normalized time, in captive vs. wild populations, from transversal vs. longitudinal data; e.g. Klingenberg, 1998; Jeffery et al., 2002; Lieberman et al., 2007). Accordingly, with a few exceptions, studies on developmental trajectories use size as a proxy for absolute time, and interpret changes in allometric patterns in terms of heterochronic types of changes (e.g. Piras et al., 2011; Foth et al., 2016; Esquerré et al., 2017). Irrespective of the theoretical framework followed by each user, the ancestral ontogenetic trajectories inferred by PASOS can be analysed in combination with additional information on size and/or absolute time to describe changes in developmental timing in more depth. Figure 9 shows the plot of shape stages aligned by PASOS against size (log of centroid size, logCS) for extant and ancestral nodes of Pantherinae compared to the ontogenetic trajectory of the Felinae ancestor. logCS for each internal node was established by optimizing the logCS for the stages at the extremes of the extant trajectories independently. Values of logCS for intermediate stages at internal nodes were calculated by dividing the total span of logCS for each node by the number of stages. The plot shows a clear similarity among Pantherinae trajectories in comparison with the Felinae ancestor, with an extension of the shape trajectory in Pantherinae associated with an increase in size. This is in agreement with several studies in Pantherinae (Sicuro, 2011; Sicuro and Oliveira, 2011; Sakamoto and Ruta, 2012; Segura et al., 2017) that have suggested that the evolution of skull shape would be related to variation in size.

Instead of plotting the aligned shape stages against size, it is also possible to work directly with the ancestral shapes inferred by PASOS. These ancestral shapes are obtained by optimizing landmark data using PM in each stage of the implied alignment (Fig. 6). Once the ancestral trajectories are obtained, it is possible to analyse the evolution of ontogenetic trajectories at any node of the tree using the approaches proposed for pairwise comparisons. For instance, changes in the ontogenetic trajectories can be visualized by plotting shape (as summarized by regression scores following the approach by Drake and Klingenberg, 2008) against size. When this approach was followed, the results in the felid example were very similar to those obtained when working with the aligned shape stages (Fig. 10). Approaches like those proposed by Sheets and Zelditch (2013) to establish evolutionary changes between pairs of ontogenetic trajectories can also be conducted considering the ancestral ontogenies inferred by PASOS.

Comparison with other approaches

To the best of our knowledge, PASOS is the first method proposed to infer ancestral shapes along the

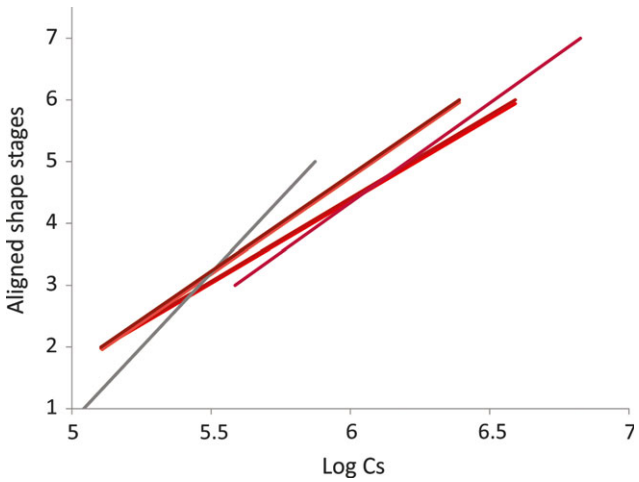

Fig. 9. Plot of aligned shape stages against size for extant and ancestral nodes of the Pantherinae clade (red shades) vs. Felinae ancestor (grey). Note the similarity among Pantherinae trajectories in comparison with Felinae ancestor.

ontogenetic trajectory that takes into account possible changes in developmental timing. Hence, it was not possible for us to make a direct comparison of the performance of the present method with other approaches. While different approaches have been adopted to compare shape ontogenies from different species (e.g. Adams and Nistri, 2010; Sheets and Zelditch, 2013), most of them do not consider phylogenetic information explicitly. Esquerré et al. (2017) analysed the evolution of head and body shape in pythons by performing a series of tests to determine the possible existence of heterochrony that implied comparison between pair of species. The limitation of that approach has already been indicated: pairwise comparisons neither enable one to properly infer the changes among clades, nor to determine which of the compared species has changed. In addition, it does not allow inference of the ancestral shapes along the ontogeny nor the changes in shape along the tree at different ontogenetic stages.

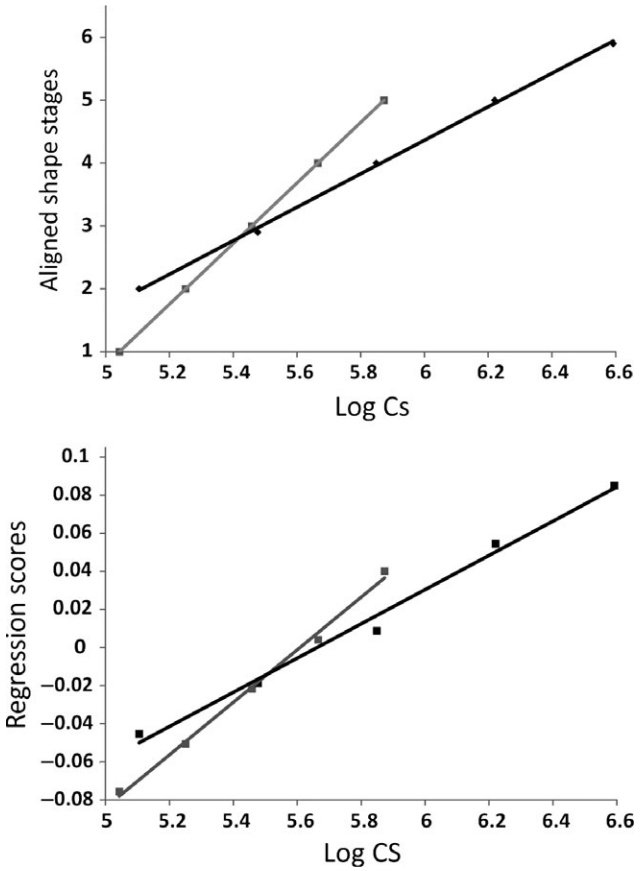

Fig. 10. Comparison of plotting aligned shape stages (above) or regression scores (below) against logCS in two different nodes in the feline example, representing the change in the ontogenetic trajectory in the branch leading to Pantherinae (Fig. 7). The evolutionary inference is the same in both approaches: a peramorphic change in the branch leading to Pantherinae with more extreme shapes associated with larger skulls. Grey = ancestor; black = descendant.

Strelin et al. (2016) and Godoy et al. (2018) attempted to establish heterochronic events including the phylogenetic dimension indirectly. The ontogenetic regression model of basal species was used to perform an allometric correction for shapes of mature specimens of all taxa. The logic of their approach was that if species with adult shapes of different size have the same residual shapes after size correction, the differences in shape can be attributed to ontogenetic scaling. One of the underlying assumptions is that basal equates with ancestral, something that has been addressed the literature for years (e.g. Krell and Cranston, 2004; Crisp and Cook, 2005). Another limitation of their approach is that given that the method is not phylogenetic, it is not possible to assess evolution along the branches of the phylogenetic tree, it is a just a comparison between two groups. Finally, as in the approach followed by Esquerré et al. (2017), the approach proposed by Strelin et al. (2016) is not useful to determine either the ancestral shapes along the ontogeny or the changes in shape along the tree, which are the main goals of PASOS.

Foth et al. (2016) proposed an explicitly phylogenetic approach to analyse the ontogenetic shape evolution on saurischian skull. In that approach ancestral shapes were inferred at different ages, the ancestral ontogenetic trajectories were reconstructed and the trajectories at different nodes were compared in order to identify possible changes in developmental timing. They recorded the shape of each species at two stages (ages): juveniles and adults. The shapes of these stages were independently mapped on a tree using squared-change parsimony (Maddison, 1991). The ancestral shapes were then used to infer heterochronic changes along the phylogeny by comparing the trajectories using different approaches (Piras et al., 2011; Klingenberg, 2016). This approach is the most similar to PASOS. However, it has the limitation that the ancestral-shape ontogenies are established assuming an *a priori* correspondence between stages. As previously indicated, this procedure affects the inference of ancestral shapes in the presence of changes in developmental timing that may in turn affect the inference of developmental timing along the tree.

Alternative approaches to analyse shape ontogenies in a phylogenetic context imply establishing ancestral values for the regression parameters of the shape variables against size/age (e.g. Giannini, 2014; Bardin et al., 2017). Those approaches can be extended to infer ancestral landmark ontogenies. PASOS has the advantage that the same criterion (minimization of changes in landmark positions) is used both to infer ancestral shapes and to determine the optimal matching among trajectories. In the case of function-based approaches, this connection is lost: minimization of the function parameters is not equivalent to

minimization of changes in landmark positions. If the aim is to infer either ancestral ontogenies (shapes) or heterochronic events, it is necessary to have a method, like PASOS, that allows us to establish ancestral shapes and changes in developmental timing simultaneously using the same criterion.

Although PASOS infers changes in developmental timing and ancestral shape ontogenies simultaneously, the changes in developmental timing are determined only by taking into account observed shape ontogenies (the possible states in the first step of PASOS). This limited number of states may limit the ability of the method to infer changes in developmental timing. Nevertheless, the empirical examples presented here demonstrate that the method is powerful enough to identify changes in developmental timing that are compatible with previous knowledge of the groups under study.

Support and discretization

A key aspect of any phylogenetic method is how to assess the evidence that supports the inferred patterns and how to determine the methodological or empirical conditions under which the method is prone to fail. However, the methods proposed to analyse ontogenies phylogenetically do not often evaluate support of the inferred patterns. While in phylogenetic searches the notion of support is associated with the evidence supporting individual groupings of the tree (Felsenstein, 1985; Bremer, 1994), in the context of PASOS what should be evaluated is the support of the inference of change in developmental timing. The strength of the evidence that supports the optimal hypothesis vs. alternative explanations is quantified in PASOS in terms of score improvement in absolute values or as a percentage of the penalty value. In addition, PASOS evaluates another aspect of the support: that related to the sample of individuals within each species. In the presence of high intraspecific variation, and under a poor sampling of specimens, the shapes included in the analysis may not be a good estimate of the average shape of the species. PASOS evaluates this potential sampling error using a resampling procedure in which the analysis is repeated considering random subsamples of the specimens in each species/stage.

The method proposed here has several simplifications that enable us to make the problem analysable in a phylogenetic framework; the most important simplification is that shape ontogenetic trajectories are discretized and described as a series of consecutive stages. This discretization of the trajectory is one of the possible sources of error. First, because the changes are determined in units of stages, the method is incapable of recognizing changes in developmental timing that are subtler. Second, because shapes of different ages are averaged to represent a single shape for each stage, descriptions of shape variation along ontogeny become less detailed. This discretization may limit the inference of changes in developmental timing. As such, the method can be considered conservative, sometimes being unable to infer certain evolutionary events of changes in developmental timing. This error is likely to be reduced with an increasing numbers of stages, which can be achieved by increasing the sampling of specimens along the whole trajectory. Alternatively, it is possible to increase the number of stages by extrapolating shapes to intervals of the ontogenetic trajectory that lack data. This extrapolation can be conducted, among other approaches, by using a weighted-moving-average approach (S. Catalano, in preparation), a non-parametric technique to interpolate values commonly used in time-series analyses (Fuller, 2009). Following this approach, the number of stages on which the ontogenetic trajectory is separated can be increased to a level where the discretization has almost no effect on the results.

Although the drawbacks of discretizing the ontogenetic trajectory are clearly evident when the shape along the complete trajectory is mapped, any inference of ancestral shapes, even in the most common case where only the shape of adult specimens is analysed, inevitably includes some sort of discretization. The very definition of "adult shape" implies a certain level of discretization: not all adults are of the same age or size, and individuals continue changing in shape even after reaching sexual maturity. This discretization is also present in many ontogenetic analyses where the ontogenetic vectors are constructed considering juvenile/small and adult/large categories as endpoints (e.g. Ivanović et al., 2007; Collyer and Adams, 2013).

Possible applications and extensions

One application of the PASOS method is the analysis of morphological and developmental disparity. Disparity refers to the degree of morphological differentiation among taxa within groups (Foote, 1999; Eble, 2000; McNamara and McKinney, 2005) and can be studied along ontogenies by comparing morphological diversity at different developmental stages (Eble, 2002). Ontogenetic disparity can be ascribed to external, ecological constraints (e.g. larval damselfishes are suggested to be more similar to each other than adult stages in part because they share similar pelagic diets; Frédérich and Vandewalle, 2011) or to developmental, functional constraints (e.g. in some cave salamanders, ontogenetic convergence results in an adult foot morphology probably related to efficient climbing; Adams and Nistri, 2010). Analysing ontogenetic trajectories using PASOS could be appropriate in these kinds of

studies because in principle different alignments of shape stages would produce different measurements of disparity across stages. The results of PASOS allow also us to determine, for instance, which proportion of the variation in adult shapes can be accounted for by changes in developmental timing.

Phylogenetic analysis of shape ontogenies deals with the evolution of shape ontogenies considering actual time (or size as proxy) or stages defined in normal developmental tables. In the case of tabulated stages the method assumes that the order of the events considered to define those stages is not modified in the species analysed. If tabulated stages represent discrete developmental stages (e.g. instars in arthropods) other possible changes in the ontogeny are the insertions and deletions of stages. As currently implemented, the method cannot deal with those events. PASOS also cannot determine changes in developmental timing that are restricted to a segment of the ontogenetic trajectory. In that case two trajectories are alike until a certain point where a change in the developmental rate occurs, producing a different offset shape. These kinds of transformations can be determined by modifying the algorithms currently included in SPASOS. Hence, future versions of PASOS will deal with those changes.

The method can also be the basis of an approach to infer phylogenetic relationships. One possibility is to perform a phylogenetic search using the Sankoff matrix calculated in the first step of PASOS (Fig. 5). Alternatively, the method can be extended to calculate the score considering the best possible alignment for each tree, as in POY (Wheeler et al., 2015) for the case of DNA sequences.

Acknowledgements

We thank the Willi Hennig Society for subsidizing TNT and making it freely available. We also thank N. Giannini, P. Goloboff and F. Prevosti for helpful discussions; D. Baldo, D. Barrasso, C. Borteiro and F. Kolenc for lending us anuran embryos; and G. Cassini and N. Muñoz for producing the felid renders. Prof. Bardin and the Associate Editor provided helpful comments. Prof. Felsenstein provided helpful comments and suggestions on earlier versions of the manuscript. This work was supported by Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT, PICT 2011-1679, PICT CONICET 2014-1930 to SAC; PICT 2014-3182 to VS).

References

Adams, D.C., Nistri, A., 2010. Ontogenetic convergence and evolution of foot morphology in European cave salamanders (Family: Plethodontidae). BMC Evol. Biol. 10, 216.

Alba, D.M., 2002. Shape and stage in heterochronic models. In: Minugh-Purvis, N., McNamara, K.J. (Eds.), Human Evolution through Developmental Change. The Johns Hopkins University Press, Baltimore, MD, pp. 28–50.

Alberch, P., Gould, S.J., Oster, G.F., Wake, D.B., 1979. Size and shape in ontogeny and phylogeny. Paleobiology 5, 296–317.

Bardin, J., Rouget, I., Cecca, F., 2017. Ontogenetic data analyzed as such. Syst. Biol. 66, 23–37.

Boughton, D.A., Collette, B.B., McCune, A.R., 1991. Heterochrony in jaw morphology of needlefishes (Teleostei: Belonidae). Syst. Biol. 40, 329–354.

Bremer, K., 1994. Branch support and tree stability. Cladistics 10, 295–304.

Cardini, A., Seetah, K., Barker, G., 2015. How many specimens do I need? Sampling error in geometric morphometrics: testing the sensitivity of means and variances in simple randomized selection experiments Zoomorphology 134, 149–163.

Catalano, S.A., Goloboff, P., 2012. Simultaneously mapping and superimposing landmark configurations with parsimony as optimality criterion. Syst. Biol. 61, 392–400.

Catalano, S.A., Goloboff, P., Giannini, N., 2010. Phylogenetic morphometrics (I): the use of landmark data in a phylogenetic framework. Cladistics 26, 539–549.

Christiansen, P., 2008. Phylogeny of the great cats (Felidae: Pantherinae), and the influence of fossil taxa and missing characters. Cladistics 24, 977–992.

Collyer, M.L., Adams, D.C., 2013. Phenotypic trajectory analysis: comparison of shape change patterns in evolution and ecology. Hystrix 24, 75–83.

Crisp, M.D., Cook, L.G., 2005. Do early branching lineages signify ancestral traits? Trends Ecol. Evol. 20, 122–128.

Davis, B.W., Li, G., Murphy, W.J., 2010. Supermatrix and species tree methods resolve phylogenetic relationships within the big cats, *Panthera* (Carnivora: Felidae). Mol. Phylogenet. Evol. 56, 64–66.

De Laet, J.E., 2005. Parsimony and the problem of inapplicables in sequence data. In: Albert, V.A. (Ed.), *Parsimony Phylogeny and Genomics*. Oxford University Press, Oxford, pp. 81–116.

Derakhshan, Z., Nokhbatolfoghahai, M., 2015. Thermal tolerance limits and effects of temperature on the growth and development of the green toad, *Bufotes viridis*. Salamandra 51, 129–136.

Dosik, M., Stayton, T., 2016. Size, shape, and stress in tortoise shell evolution. Herpetologica 72, 309–317.

Drake, A.G., Klingenberg, C.P., 2008. The pace of morphological change: historical transformation of skull shape in St Bernard dogs. Proc. R. Soc. B 275, 71–76.

Eble, G.J., 2000. Contrasting evolutionary flexibility in sister groups: disparity and diversity in Mesozoic atelostomate echinoids. Paleobiology 26, 56–79.

Eble, G.J., 2002. Multivariate approaches to development and evolution. In: Minugh-Purvis, N., McNamara, K.J., (Eds.), Human Evolution through Developmental Change. Johns Hopkins University Press, Baltimore, MD, pp. 51–78.

Esquerré, D., Sherratt, E., Keogh, J.S., 2017. Evolution of extreme ontogenetic allometric diversity and heterochrony in pythons, a clade of giant and dwarf snakes. Evolution 71, 2829–2844.

Fagen, R.M., Wiley, K.S., 1978. Felid paedomorphosis, with special reference to *Leopardus*. Carnivore 1, 72–81.

Farris, J., 1970. Methods for computing Wagner trees. Syst. Zool. 19, 83–92.

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., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Felsenstein, J., 2004. *Inferring phylogenies*. Sinauer Associates, Sunderland, MA.

Figueirido, B., Serrano-Alarcón, F.J., Slater, G.J., Palmqvist, P., 2010. Shape at the cross-roads: homoplasy and history in the evolution of the carnivoran skull towards herbivory. J. Evol. Biol. 23, 2579–2594.

Fink, W.L., 1982. The conceptual relationship between ontogeny and phylogeny. Paleobiology 8, 254–264.
Foote, M., 1999. Morphological diversity in the evolutionary radiation of Paleozoic and post-Paleozoic crinoids. Paleobiology 25, 1–116.
Foth, C., Hedrick, B.P., Ezcurra, M.D., 2016. Cranial ontogenetic variation in early saurischians and the role of heterochrony in the diversification of predatory dinosaurs. PeerJ 4, e1589.
Frédérich, B., Vandewalle, P., 2011. Bipartite life cycle of coral reef fishes promotes increasing shape disparity of the head skeleton during ontogeny: an example from damselfishes (Pomacentridae). BMC Evol. Biol. 11, 82.
Fuller, W.A., 2009. Introduction to Statistical Time Series, Vol. 428. John Wiley & Sons, New York, NY.
Giannini, N.P., 2014. Quantitative developmental data in a phylogenetic framework. J. Exp. Zool. B. Mol. Dev. Evol. 322, 558–566.
Godoy, P.L., Ferreira, G.S., Montefeltro, F.C., Vila Nova, B.C., Butler, R.J., Langer, M.C., 2018. Evidence for heterochrony in the cranial evolution of fossil crocodyliforms. Palaeontology 61, 543–558.
Goloboff, P.A., Catalano, S.A., 2011. Phylogenetic morphometrics (II): algorithms for landmark optimization. Cladistics 27, 42–51.
Goloboff, P.A., Catalano, S.A., 2016. TNT version 1.5, including a full implementation of phylogenetic morphometrics. Cladistics 32, 221–237.
Goloboff, P.A., Farris, J.S., Källersjö, M., Oxelman, B., Ramírez, M.N.J., Szumik, C.A., 2003. Improvements to resampling measures of group support. Cladistics 19, 324–332.
Goloboff, P., Mattoni, C., Quinteros, A., 2006. Continuous characters analyzed as such. Cladistics 22, 589–601.
Goloboff, P., Farris, J.S., Nixon, K., 2008. TNT, a free program for phylogenetic analysis. Cladistics 24, 774–786.
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.
Gosner, K.L., 1960. A simplified table for staging anurans embryos and larvae with notes on identification. Herpetologica 16, 183–190.
Gould, S.J., 1977. Ontogeny and Phylogeny. Harvard University Press, Cambridge, MA.
Gower, J.C., 1975. Generalized procrustes analysis. Psychometrika 40, 33–51.
Grosso, J., 2017. Patrones de Variación durante la Ontogenia Temprana en Leptodactyliformes. Un Abordaje desde la Ecomorfología y la Filogenia. Ph.D. Thesis, Universidad Nacional de Córdoba, Argentina.
Houle, D., Bolstad, G.H., van der Linde, K., Hansen, T.F., 2017. Mutation predicts 40 million years of fly wing evolution. Nature 548, 447–450.
Ivanović, A., Vukov, T., Džukic, G., Tomašević, N., Kalesić, M.L., 2007. Ontogeny of skull size and shape changes within a framework of biphasic lifestyle: a case study in six *Triturus* species (Amphibia, Salamandridae). Zoomorphology 126, 173–183.
Jeffery, J.E., Richardson, M.K., Coates, M.I., Bininda-Emonds, O.R., 2002. Analyzing developmental sequences within a phylogenetic framework. Syst. Biol. 51, 478–491.
Jeffery, J.E., Bininda-Emonds, O.R., Coates, M.I., Richardson, M.K., 2005. A new technique for identifying sequence heterochrony. Syst. Biol. 54, 230–240.
Johnson, W.E., Eizirik, E., Pecon-Slattery, J., Murphy, W.J., Antunes, A., Teeling, E., O'Brien, S.J., 2006. The late Miocene radiation of modern Felidae: a genetic assessment. Science 311, 73–77.
Keyte, A.L., Smith, K.K., 2014. Heterochrony and developmental timing mechanisms: changing ontogenies in evolution. Semin. Cell Dev. Biol. 34, 99–107.
Klingenberg, C.P., 1998. Heterochrony and allometry: the analysis of evolutionary change in ontogeny. Biol. Rev. 73, 79–123.
Klingenberg, C.P., 2016. Size, shape, and form: concepts of allometry in geometric morphometrics. Dev. Genes. Evol. 226, 113–137.
Klingenberg, C.P., Gidaszewski, N.A., 2010. Testing and quantifying phylogenetic signals and homoplasy in morphometric data. Syst. Biol. 59, 245–261.
Krell, F.T., Cranston, P.S., 2004. Which side of the tree is more basal? Syst. Entomol. 29, 279–281.
Lieberman, D.E., Carlo, J., Ponce de León, M., Zollikofer, C.P.E., 2007. A geometric morphometric analysis of heterochrony in the cranium of chimpanzees and bonobos. J. Hum. Evol. 52, 647–662.
Lobo, F., 1996. Nuevas observaciones sobre la osteología del género *Physalaemus* (Anura: Leptodactylidae). Acta Zool. Lilloana 43, 317–326.
Lourenço, L.B., Targueta, C.P., Baldo, D., Nascimento, J., Garcia, P.C.A., Andrade, G.V., Haddad, C.F.B., Recco-Pimentel, S.M., 2015. Phylogeny of frogs from the genus *Physalaemus* (Anura, Leptodactylidae) inferred from mitochondrial and nuclear gene sequences. Mol. Phylogenet. Evol. 92, 204–216.
Maddison, W.P., 1991. Squared-change parsimony reconstructions of ancestral states for continuous-valued characters on a phylogenetic tree. Syst. Biol. 40, 304–314.
McNamara, K.J., 1997. Shapes of Time: The Evolution of Growth and Development. Johns Hopkins University Press, Baltimore, MD.
McNamara, K.J., McKinney, M.L., 2005. Heterochrony, disparity, and macroevolution. Paleobiology 31, 17–26.
Mitteroecker, P., Gunz, P., Bookstein, F.L., 2005. Heterochrony and geometric morphometrics: a comparison of cranial growth in *Pan paniscus* versus *Pan troglodytes*. Evol. Dev. 7, 244–258.
Mitteroecker, P., Gunz, P., Windhager, S., Schaefer, K., 2013. A brief review of shape, form, and allometry in geometric morphometrics, with applications to human facial morphology. Hystrix 24, 59–66.
Muñoz, N.A., Cassini, G.H., Candela, A.M., Vizcaíno, S.F., 2017. Ulnar articular surface 3-D landmarks and ecomorphology of small mammals: a case study of two early Miocene typotheres (Notoungulata) from Patagonia. Earth Env. Sci. T. R. So. 106, 315–323.
Needleman, S.B., Wunsch, C.D., 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48, 443–453.
Ospina-Garcés, S.M., De Luna, E., 2017. Phylogenetic analysis of landmark data and the morphological evolution of cranial shape and diets in species of *Myotis* (Chiroptera: Vespertilionidae). Zoomorphology 136, 251–265.
Piras, P., Salvi, D., Ferrara, G., Maiorino, L., Delfino, M., Pedde, L., Kotsakis, T., 2011. The role of post-natal ontogeny in the evolution of phenotypic diversity in Podarcis lizards. J. Evol. Biol. 24, 2705–2720.
Ponssa, M.L., Vera Candioti, F., 2012. Patterns of skull development in anurans: size and shape relationship during postmetamorphic cranial ontogeny in five species of the *Leptodactylus fuscus* Group (Anura: Leptodactylidae). Zoomorphology 131, 349–362.
Reilly, S.M., Wiley, E.O., Meinhardt, D.J., 1997. An integrative approach to heterochrony: the distinction between interspecific and intraspecific phenomena. Biol. J. Linn. Soc. 60, 119–143.
Rodríguez-Mendoza, R., Muñoz, M., Saborido-Rey, F., 2011. Ontogenetic allometry of the bluemouth, *Helicolenus dactylopterus dactylopterus* (Teleostei: Scorpaenidae), in the Northeast Atlantic and Mediterranean based on geometric morphometrics. Hydrobiologia 670, 5–22.
Rohlf, F.J., 2002. Geometric morphometrics and phylogeny. In: MacLeod, N., Forey, P.L. (Eds.), Morphology, Shape, and Phylogeny. Taylor & Francis, London, pp. 175–193.
Rohlf, F.J., Marcus, L.F., 1993. A revolution in morphometrics. Trends Ecol. Evol. 8, 129–132.
Rohlf, F.J., Slice, D., 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst. Zool. 39, 40–59.
Rüber, L., Adams, D.C., 2001. Evolutionary convergence of body shape and trophic morphology in cichlids from Lake Tanganyika. J. Evol. Biol. 14, 325–332.
Sakamoto, M., Ruta, M., 2012. Convergence and divergence in the evolution of cat skulls: temporal and spatial patterns of morphological diversity. PLoS ONE 7, 1–13.

Sankoff, D., 1975. Minimal mutation trees of sequences. SIAM J. Appl. Math. 28, 35–42.
Savriama, Y., Gómez, J.M., Perfectti, F., Klingenberg, C.P., 2012. Geometric morphometrics of corolla shape: dissecting components of symmetric and asymmetric variation in *Erysimum mediohispanicum* (Brassicaceae). New Phytol. 196, 945–954.
Schlager, S., 2017. Morpho and Rvcg - shape analysis in R. In: Zheng, G., Li, S., Szekely, G. (Eds.), Statistical Shape and Deformation Analysis. Academic Press, London, pp. 217–256.
Schulmeister, S., Wheeler, W.C., 2004. Comparative and phylogenetic analysis of developmental sequences. Evol. Dev. 6, 50–57.
Segura, V., 2014. Ontogenia Craneana Postnatal en Cánidos y Félidos Neotropicales: Funcionalidad y Patrones Evolutivos. Ph.D. Thesis, Universidad Nacional de La Plata, Argentina.
Segura, V., 2015. A three-dimensional skull ontogeny in the bobcat (*Lynx rufus*; Carnivora: Felidae): a comparison with other carnivores. Can. J. Zool. 93, 225–237.
Segura, V., Prevosti, F., Cassini, G., 2013. Cranial ontogeny in the Puma lineage, *Puma concolor*, *Herpailurus yagouaroundi*, and *Acinonyx jubatus* (Carnivora: Felidae): a three-dimensional geometric morphometric approach. Zool. J. Linn. Soc. 169, 235–250.
Segura, V., Cassini, G.H., Prevosti, F.J., 2017. Three-dimensional cranial ontogeny in pantherines (*Panthera leo*, *P. onca*, *P. pardus*, *P. tigris*; Carnivora: Felidae). Biol. J. Linn. Soc. 120, 210–227.
Sheets, H.D., Zelditch, M.L., 2013. Studying ontogenetic trajectories using resampling methods and landmark data. Hystrix 24, 67–73.
Sicuro, F.L., 2011. Evolutionary trends on extant cat skull morphology (Carnivora: Felidae): a three-dimensional geometrical approach. Biol. J. Linn. Soc. 103, 176–190.
Sicuro, F.L., Oliveira, L.F.B., 2011. Skull morphology and functionality of extant Felidae (Mammalia: Carnivora): a phylogenetic and evolutionary perspective. Zool. J. Linn. Soc. 161, 414–462.
Smith, U.E., Hendricks, J.R., 2013. Geometric morphometric character suites as phylogenetic data: extracting phylogenetic signal from gastropod shells. Syst. Biol. 62, 366–385.
Smith, T.F., Waterman, M.S., 1981. Identification of common molecular subsequences. J. Mol. Biol. 147, 195–197.
Strelin, M.M., Benitez-Vieyra, S.M., Fornoni, J., Klingenberg, C.P., Cocucci, A.A., 2016. Exploring the ontogenetic scaling hypothesis during the diversification of pollination syndromes in *Caiophora* (Loasaceae, subfam. Loasoideae). Ann. Bot. 117, 937–947.
Tomatis, C.G., Baldo, D., Kolenc, F., Borteiro, C., 2009. Chromosomal variation in the species of the *Physalaemus henselii* group (Anura, Leiuperidae). J. Herpetol. 43, 555–560.
Vera Candioti, F., Altig, R., 2010. A survey of shape variation in keratinized labial teeth of anuran larvae as related to phylogeny and ecology. Biol. J. Linn. Soc. 101, 609–625.
Vera Candioti, F., Haad, B., Baldo, D., Kolenc, F., Borteiro, C., Altig, R., 2011. Different pathways are involved in the early development of the transient oral apparatus in anuran tadpoles (Anura: Leiuperidae). Biol. J. Linn. Soc. 104, 330–345.
Webster, M., Zelditch, M.L., 2005. Evolutionary modifications of ontogeny: heterochrony and beyond. Paleobiology 31, 354–372.
Wheeler, W.C., 1999. Fixed character states and the optimization of molecular sequence data. Cladistics 15, 379–385.
Wheeler, W.C., 2003a. Implied alignment: a synapomorphy-based multiple sequence alignment method and its use in cladogram search. Cladistics 19, 261–268.
Wheeler, W.C., 2003b. Search-based optimization. Cladistics 19, 348–355.
Wheeler, W.C., Lucaroni, N., Hong, L., Crowley, L.M., Varón, A., 2015. POY version 5: phylogenetic analysis using dynamic homologies under multiple optimality criteria. Cladistics 31, 189–196.
Wiley, D.F., 2006. Landmark Editor 3.0. IDAV, University of California, Davis. http://graphics.idav.ucdavis.edu/research/EvoMorph.
Williams, J.S., Niedzwiecki, J.H., Weisrock, D.W., 2013. Species tree reconstruction of a poorly resolved clade of salamanders (Ambystomatidae) using multiple nuclear loci. Mol. Phylogenet. Evol. 68, 671–682.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Results of PASOS for the cat dataset considering dental ages.

Fig. S2. Ancestral shapes inferred for the node subtending the clade *P. biligonigerus* + *P. santafecinus* (Node 10, Fig. 8). Arrows show how the placement of the dorsal fin limit is modified during development; this placement is a synapomorphic feature for this clade. Dashed lines indicate the shape of the previous node. S and E represent start and end points of the ontogenetic trajectory.

Appendix

Penalty and score for different pairings between trajectories

The general formula to calculate the cost (C) associated with a particular pairing between trajectories is

$$C = D + (P/S) \times N$$

where D is the sum of landmark displacement in each stage divided by the number of stages compared, P is the penalty factor calculated as indicated in the main text and S is the number of stages in which the ontogenetic trajectory is discretized. N depends on the particular modification between trajectories:

Case 1. For shifts of trajectories, N is the number of stages the ontogenetic trajectory is shifted.
Case 2. For stretching of trajectories that modifies only one of the limits, N is the number of stages this limit is modified.
Case 3. For stretching of trajectories that modifies both the limits in the same direction, N is calculated as the largest change of both limits (in absolute value) plus the difference between the changes at both limits in absolute values. This case reduces to case 2 when the change in one of the limits is equal to zero and reduces to case 1 if the change in both limits is the same.
Case 4. For stretching of trajectories that modifies both the limits in different directions, N is calculated as the sum of the changes in both limits in absolute value. This case reduces to case 2 when the change in one of the limits is equal to zero.