

SPASOs v.1.1: User Guide

1-. Overview

PASOS (Phylogenetic Analysis of Shape OntogenieS) is a method to infer ancestral shape ontogenies and changes in developmental timing along a phylogeny departing from shapes described as landmark configurations (Catalano et al., 2019). The approach is rooted in previous developments for the analysis of landmark data in a phylogenetic context with parsimony as optimality criterion (Catalano et al., 2010, Goloboff & Catalano 2011; Catalano & Goloboff 2012). The ontogenetic trajectory is delimited by starting and end points common to all species under analysis (double stage standardization Alba, 2002). The method allows analyzing ontogenetic trajectories with shapes recorded at different moments of the trajectories with age defined as either *a priori* defined stages or continuous time. This method is implemented in SPASOS (Software for Phylogenetic Analysis of Shape OntogenieS), a windows command line program. SPASOS is written in C language being open source (Catalano et al submitted). The program reads the data from a tps file including all the specimens under analysis. Alternatively, the program also accepts data in a format similar to that called "txt" in MorphoJ (Klingenberg 2011). The reference tree should be given in parenthetical notation in a different file. While the inference of changes in developmental timing is performed in SPASOS, the optimization of landmark data on the tree is performed in TNT (SPASOS launches TNT). The program produced a series of output files that allow different a posteriori processing and visualization of the results.

NOTE: Users are urged to read the paper describing the method (Catalano *et al.* 2019; Catalano et al submitted) before running the analyses.

2-. Input Data

Landmark and timing data. SPASOS can read landmark coordinates from either a tps format file or a file with a format similar to that called "txt" in MorphoJ (Klingenberg 2011). All the specimens should be included in the same file. The ID of each specimen should include the species name (a name without white spaces), optionally followed by a specimen number identifier, with a slash symbol ("/") separating the species and specimen identifier. The place of each specimen in the trajectory (age, size, or developmental stage) is indicated in a line preceded by the string "AGE=". If landmarks are in three dimensions, the string indicating the number of landmarks should be "LM3=" instead of "LM=". In the case of "txt" format files, the program assumes that the first row includes the column headings, and hence it is skipped. The first column includes the name of each specimen following the format previously described. Next column includes the age information, while subsequent columns include landmark coordinates. Data in different columns are separated by space or tab characters. Period is used as decimal separator.

Example tps format file

```
LM= 5
-0.173 -0.054
-0.152 0.012
-0.128 -0.093
-0.093 -0.075
-0.098 0.069
ID=Panthera_leo/45
AGE= 5.7
LM= 5
-0.093 -0.058
-0.152 0.011
-0.138 -0.023
-0.043 -0.074
-0.102 0.069
ID=Panthera_tigris
AGE= 4.3
```

Example "txt" MorphoJ format file

```
Panthera_leo/45 5.7 -0.173 -0.054 -0.152 0.012 -0.128 -0.093 -0.093 -0.075 -0.098 0.069
Panthera_tigris 4.3 -0.093 -0.058 -0.152 0.011 -0.138 -0.023 -0.043 -0.074 -0.102 0.069
```

Phylogenetic Tree. The topology considered to map the landmark data is read from a text file with the tree in TNT parenthetical notation and preceded by the string "tread". The tree should be completely resolved, no polytomic nodes are accepted. Species names should match those included in the landmark data. Missing species are not allowed either in the landmark file or in the tree file.

Example

```
tread
(one (two (three four ) ) ) ;
```

3-. *Running the analysis*

In order to run the analysis, executables of SPASOS and TNT should be in the same folder. TNT can be downloaded [here](#):

Mandatory arguments

-i "myinputfile" Input data file

Input files can be in "tps" or "txt" format, with the corresponding extension. If the data is in "txt" format, the number of landmarks and dimensions (2-3) should follow the file name. If data is in tps format no specification about dimensions or number of landmarks should be indicated. If the file is not in the same folder than SPASOS, give the full path.

-t "mytreefile.tre" Input tree file

Optional arguments

-c Colour tree branches with violet as intermediate colour (black as default)

- m Colour tree branches according to changes in Onset (changes in Offset as default)
- n Show changes as number of stages (percentage of the trajectory, total span as default)
- o Infer ancestral ontogenetic trajectories in the original frame of comparison (i.e. force no change in developmental timing in all branches)
- p N Define a penalty factor (0 no penalty, 1 lowest, 5 highest). Default= 0
- r N Perform resampling with N replicates
- u "myprefix" Give a prefix for output files
- y Calculate decay support values

Example: perform an analysis with landmark data included in the file *cats.tps*, the tree in the file *cats.tre*. Calculate decay values and perform resampling analysis using 100 replicates.

C: \spasos -i cats.tps -t cats.tre -y -r 100

Changes in developmental time inferred in SPASOS

Previous version of SPASOS classified changes in developmental timing as shifts (when offset and onset timing varies equally), and stretches (when only one of the extremes of the trajectory varies). In the new version of SPASOS, the changes in developmental timing at each limit of the ontogenetic trajectory are indicated separately as positive/negative percentages. For instance a change in -10% at the end of ontogeny implies a shortening of the descendant ontogenetic trajectory with a span that represent the 10% of the ancestral ontogenetic trajectory -a pattern compatible with paedomorphosis. Changes at the beginning of the ontogenetic trajectory are represented in the same way, with negative values representing extensions of the descendant ontogenetic trajectory towards “unmature” shapes.

Penalty

SPASOS allows assigning a penalty cost for the changes in developmental timing that is proportional to the extent of the modification. The user can choose different penalty levels, from 0 (no penalty) to 5, being 0 de default value. Penalty calculation is described in Catalano et al. (2019).

Support values

SPASOS implements two different approaches to evaluate the strength of the evidence that support the inferences made. One of the approaches evaluates the effect of the intraspecific sampling in the pattern obtained. This is performed by a resampling procedure (*alla* bootstrap, i.e. sampling with no replacement), where 33 % of the specimens are randomly excluded and the analysis is rerun. Once this is performed, the changes inferred in each replicate are compared to the original changes, counting as a congruent result when a change has the same direction irrespective of the amount of change. By default this is repeated 100 times. In addition, SPASOS implements another measure of support, a decay index, by running the analysis with increasing penalty values until no inference of change in the ontogenies is determined. The option to run the resampling analysis is -r followed by the number of replicates.

Decay index

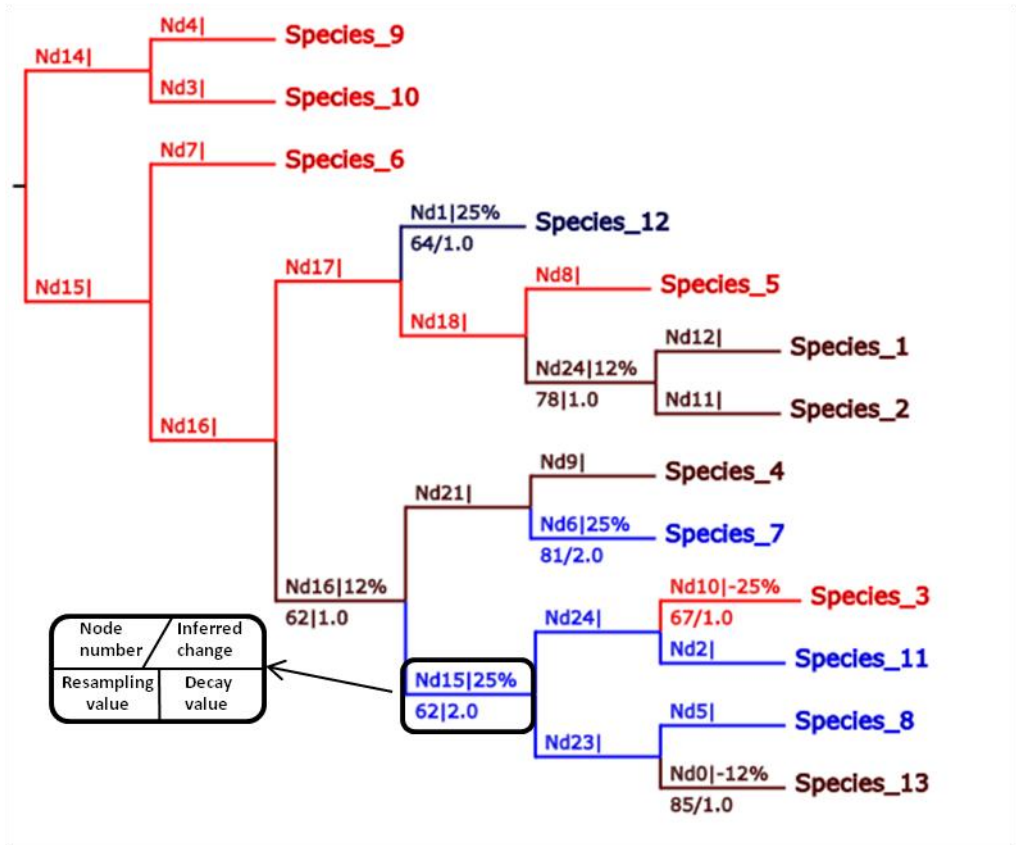
SPASOS calculates a decay index that determines how strong is the evidence that supports a certain inference of change in developmental timing. SPASOS reanalyses the data considering increasing penalty values (increase of 20% each time). Hence a value of 0.2 indicates that the transformation originally inferred is no longer retrieved when the analysis is repeated considering a penalty 20% higher. The option to run the decay analysis is -y. These values allows comparing the support among nodes of the same analysis.

4-. Output files

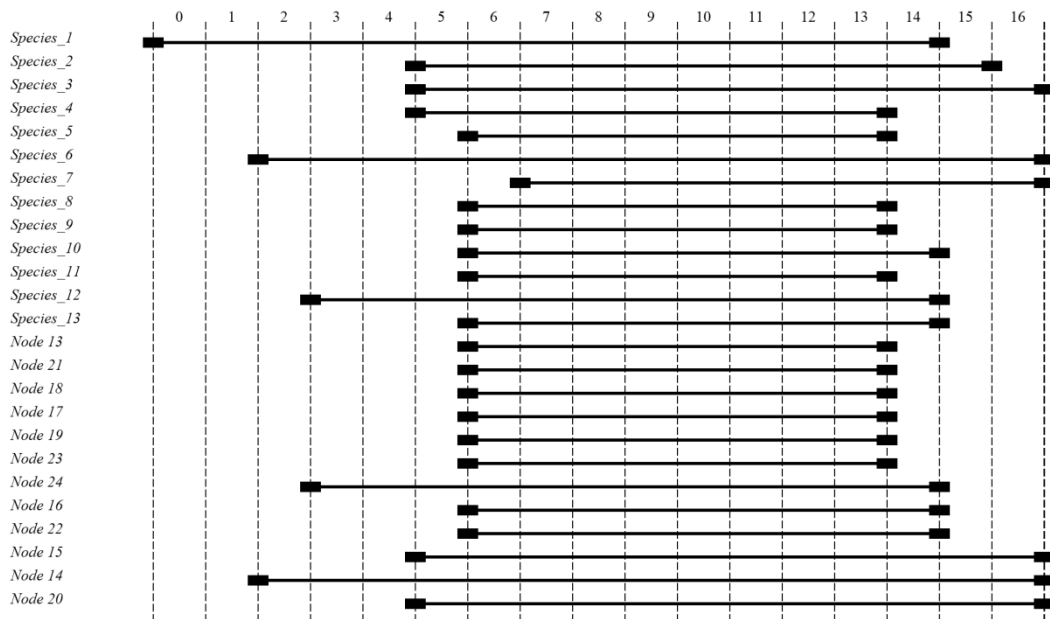
SPASOS produces a series of output files that allows visualization a post-processing of the results. All files are named including as prefix the name of tps/txt file. Alternatively, the user can give a different name using the option -u "mynewname".

1. *An image file (svg) with the phylogenetic tree indicating the changes in developmental timing inferred for the offset the trajectory at each branch.* Support values (resampling and decay index) are shown below those branches where changes in developmental timing were inferred. Branches are coloured according to the position in the multiple alignment of ontogenetic trajectories with bluish colours representing extended shape trajectories and reddish colour representing truncated trajectories (intermediate values are black). An image file showing the same information but

regarding the onset of the trajectory can optionally be generated. Files are named “mydata”_tree_offset.svg and “mydata”_tree_onset.svg. Node numbering follows that in TNT.



2. An image file (svg) showing the multiple alignment of ontogenetic trajectories. One of the results of PASOS is the determination of the correspondence among ontogenetic trajectories. The alignment is established taking into account the changes in developmental timing inferred by SPASOS (see Catalano et al. 2019). A visual representation of the correspondence among trajectories is included in a file named “mydata”_traj.svg. Node numbering follows that shown in the svg tree file (Fig 3).



3-. A *tps* file including shapes for ancestral and observed ontogenetic trajectories. Each ontogenetic trajectory is separated in different stages from the onset to the offset. The number of stages of each node varies according to the extension of each trajectory. Shapes for ancestors are represented by the shapes inferred in TNT taking into account the multiple alignment of the shape ontogenies determined in SPASOS. The label of each configuration includes the species name and the corresponding stage in the multiple alignment (represented in the *svg* file). First and last stage for each node are labelled as *speciesname_Onset_stgN* and *speciesname_Offset_stgN* respectively.

4-. A TNT format file with the alignment among ontogenetic trajectories. This allows inferring ancestral shapes taking into account this new comparison. To accomplish that, SPASOS generates a TNT file where each stage is included as a different character/block in consecutive order (the first character represents the youngest stage). Care should be taken when visualizing the results because TNT infers a shape even if it is not included in the trajectory. In order to recognize which stage is included in each node, the user may look at the *svg* file where there is a visual representation of the implied alignment (see Catalano et al. 2019). This file is named as “myfile”_aln.tnt.

(5) A *txt* file with complementary information for each configuration included in the *tps*. This includes classifiers (species, stage) and also covariates (relative position of each configuration in the trajectories and absolute values of the variable chosen to order

specimens along the ontogeny, e.g., log centroid size). For ancestral ontogenetic trajectories, these values are obtained by mapping them in TNT as a continuous character. Notice that these values are only to have a general reference since only a single value is considered for each limit of the ontogenetic trajectory, even if there is ambiguity in node reconstruction. A thorough analysis should include taken into account the potential existence of ambiguity. This file is named as “myfile”_info.txt.

Example:

```
Configuration,Species/IntNode,Aligned_Stage,Rel_Pos
P_albofrons/Stg_0, P_albofrons ,0,0.000
P_albofrons/Stg_1, P_albofrons ,1,0.250
P_albofrons/Stg_2, P_albofrons ,2,0.500
P_albofrons/Stg_3, P_albofrons ,3,0.750
P_albofrons/Stg_4, P_albofrons ,4,1.000
P_camacan/Stg_1, P_camacan,1,0.000
P_camacan/Stg_2, P_camacan,2,0.200
P_camacan/Stg_3, P_camacan,3,0.400
P_camacan/Stg_4, P_camacan,4,0.600
P_camacan/Stg_5, P_camacan ,5,0.800
P_camacan/Stg_6, P_camacan ,6,1.000
```

5-. References

Alba D.M. 2002. Shape and stage in heterochronic models. In: Minugh-Purvis N., McNamara K.J., editors. Human evolution through developmental change. Baltimore: The Johns Hopkins University Press. pp 28-50.

Catalano SA, MF Vera Candioti, V Segura. (2019). “PASOS: a method for the Phylogenetic Analysis of Shape Ontogenies”. *Cladistics*. 35: 671-687.

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

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

Goloboff P.A., Catalano S.A. 2011. Phylogenetic morphometrics (II): algorithms for landmark optimization. *Cladistics*. 27:42–51.