

OntoPhylo Tutorial: Application 2 - Morphospace Dynamics

Diego S. Porto and Sergei Tarasov

20 June, 2023

Load packages.

If you are starting a new R session (again), then reload *ontophylo*.

```
library(ontophylo)
```

And all the other packages. If you have not them installed, please do so by running `install.packages()`.

```
library(tidyverse)
library(gganimate)
```

Load data.

First, load the data from the tutorial 2.

```
load("RData/step2_paramo.RData")
```

Organize data.

Let's split the lists to facilitate downstream analyses.

```
stm_amalg_head <- stm_amalg_anato$head
stm_amalg_meso <- stm_amalg_anato$mesosoma
stm_amalg_meta <- stm_amalg_anato$metasoma
```

And merge the identical adjacent state bins across all branches and trees.

```
# Merge state categories across branches.
stm_merg_head <- lapply(stm_amalg_head, function(x) merge_tree_cat(x) )
stm_merg_head <- do.call(c, stm_merg_head)

stm_merg_meso <- lapply(stm_amalg_meso, function(x) merge_tree_cat(x) )
stm_merg_meso <- do.call(c, stm_merg_meso)

stm_merg_meta <- lapply(stm_amalg_meta, function(x) merge_tree_cat(x) )
stm_merg_meta <- do.call(c, stm_merg_meta)
```

```
stm_merg_pheno <- lapply(stm_amalg_pheno, function(x) merge_tree_cat(x) )
stm_merg_pheno <- do.call(c, stm_merg_pheno)
```

Then, let's get a tree sample from each anatomical region and the entire phenome.

```
tree_hd <- stm_merg_head[[1]]
tree_ms <- stm_merg_meso[[1]]
tree_mt <- stm_merg_meta[[1]]
tree_ph <- stm_merg_pheno[[1]]
```

Overview on morphospace reconstruction with *OntoPhylo*.

OntoPhylo is able to reconstruct the morphospace dynamics through time by applying some technique of dimension reduction to multidimensional phenotypes producing a two-dimensional morphospace and then stacking morphospaces from different time slices to produce an animation. *OntoPhylo* uses the information from the amalgamated states mapped onto tree branches present in a given time slice to calculate Hamming distances among all states available and then apply Multidimensional Scaling (MDS) to get the morphospace coordinates. As for now, only MDS is available, but we are going to incorporate other dimension reduction methods in the future. There is also the option to add some noise to improve visualization of points. The number of temporal slices will depend on the resolution parameter used to discretize tree branches. Higher resolution values produce more temporal slices, and thus require much more time to process. As for now, morphospace dynamics can be reconstructed for a single tree each time, but we are planning to incorporate topological uncertainty in the future.

STEP 1. Multidimensional scaling trees across time slices.

First, let's calculate the MDS for all temporal slices of the sample trees of each anatomical region and the entire phenome.

```
# HEAD.
cat(paste0("\n", "Working on MDS: ", Sys.time(), "\n"))
MD_hd <- suppressWarnings(MultiScale.simap(tree_hd))
cat(paste0("\n", "Finished: ", Sys.time(), "\n"))

# MESOSOMA.
cat(paste0("\n", "Working on MDS: ", Sys.time(), "\n"))
MD_ms <- suppressWarnings(MultiScale.simap(tree_ms))
cat(paste0("\n", "Finished: ", Sys.time(), "\n"))

# METASOMA
cat(paste0("\n", "Working on MDS: ", Sys.time(), "\n"))
MD_mt <- suppressWarnings(MultiScale.simap(tree_mt))
cat(paste0("\n", "Finished: ", Sys.time(), "\n"))

# PHENOME
cat(paste0("\n", "Working on MDS: ", Sys.time(), "\n"))
MD_ph <- suppressWarnings(MultiScale.simap(tree_ph))
cat(paste0("\n", "Finished: ", Sys.time(), "\n"))
```

STEP 2. Plotting morphospaces.

And then, let's just plot the final temporal slice of each morphospace.

```
# Set some parameters for the plots.
# Tree height.
Tmax = max(MD_hd$Points$time)
# Temporal slice (past to present).
Tslice = max(MD_hd$Points$time)

# HEAD.
# Get the MDS plot.
mds_plot_hd <- mds_plot(MD_hd, Tslice = Tslice) +
  labs(title = paste0("HEAD - ", "Time from present: ", floor(Tmax - Tslice), " Myr"))

# Save a png of the final slice of the morphospace.
ggsave(paste0("figures/", "mds_head_slice.png"), units = "in", width = 7, height = 7)

# MESOSOMA.
# Get the MDS plot.
mds_plot_ms <- mds_plot(MD_ms, Tslice = Tslice) +
  labs(title = paste0("MESOSOMA - ", "Time from present: ", floor(Tmax - Tslice), " Myr"))

# Save a png of the final slice of the morphospace.
ggsave(paste0("figures/", "mds_meso_slice.png"), units = "in", width = 7, height = 7)

# METASOMA.
# Get the MDS plot.
mds_plot_mt <- mds_plot(MD_mt, Tslice = Tslice) +
  labs(title = paste0("METASOMA - ", "Time from present: ", floor(Tmax - Tslice), " Myr"))

# Save a png of the final slice of the morphospace.
ggsave(paste0("figures/", "mds_meta_slice.png"), units = "in", width = 7, height = 7)

# PHENOME.
# Get the MDS plot.
mds_plot_ph <- mds_plot(MD_ph, Tslice = Tslice) +
  labs(title = paste0("PHENOME - ", "Time from present: ", floor(Tmax - Tslice), " Myr"))

# Save a png of the final slice of the morphospace.
ggsave(paste0("figures/", "mds_pheno_slice.png"), units = "in", width = 7, height = 7)
```

As you can see, the final morphospaces for individual anatomical regions do not look that much informative. This is because, for only 10 amalgamated character per anatomical region, there is some repetition of amalgamated states on internal branches of the tree. The morphospace looks better for the entire phenome (30 characters) but still not much informative. However, as more and more characters are amalgamated, each discrete bin on tree branches is basically a unique combination of states from the individual characters and thus, morphospace occupation through time looks much more a diffusion-like process.

Let's plot an example from the amalgamation of the entire modified data set from Sharkey et al. (2012) with 239 characters.

Finally, we can reconstruct the morphospace dynamics through time for the entire phenome by saving a GIF animation. The animation can be saved to a GIF file using the package *gganimate*.

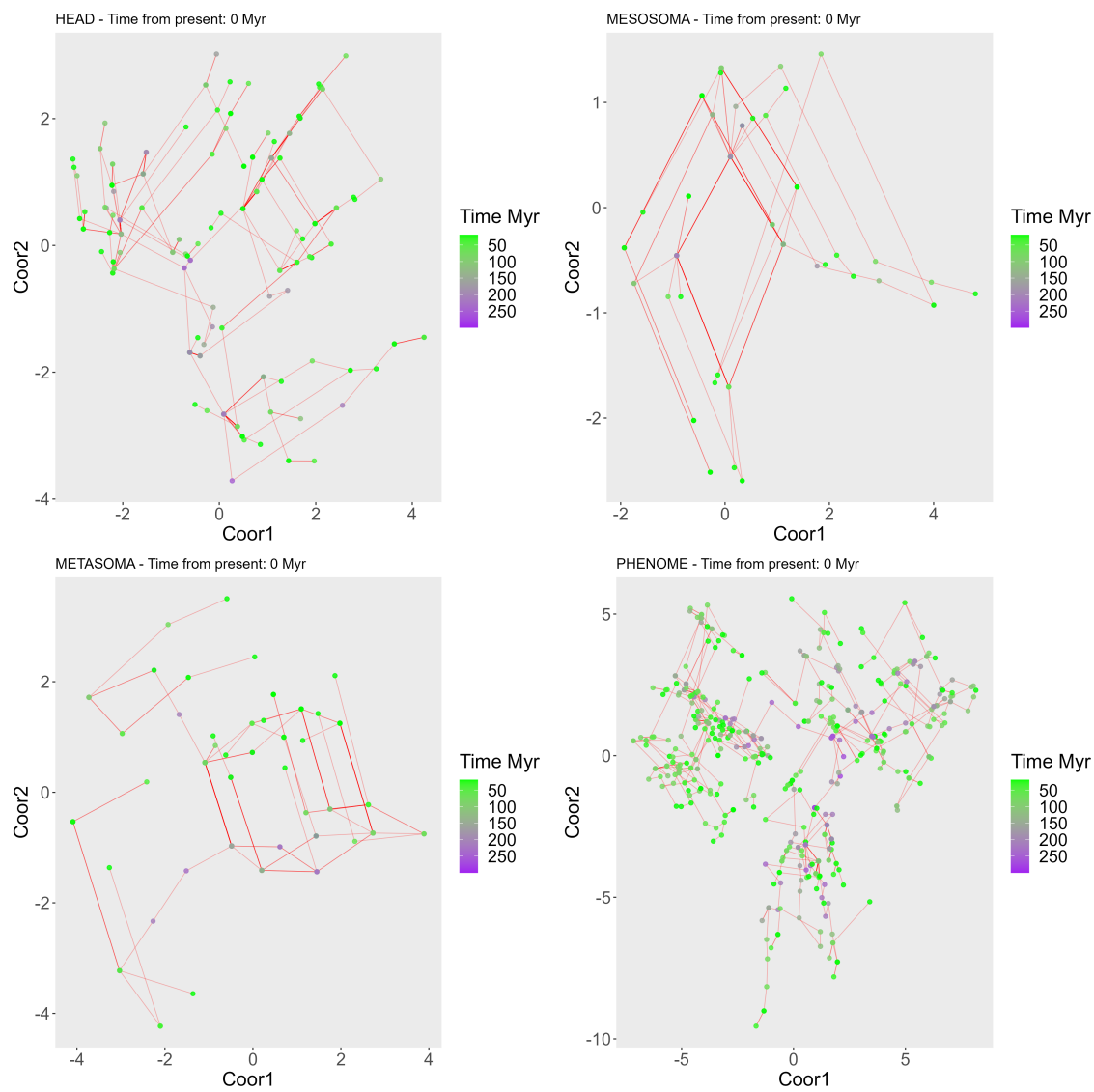


Figure 1: From left to right, top to bottom, morphospaces of head, mesosoma, metasoma, and phenome.

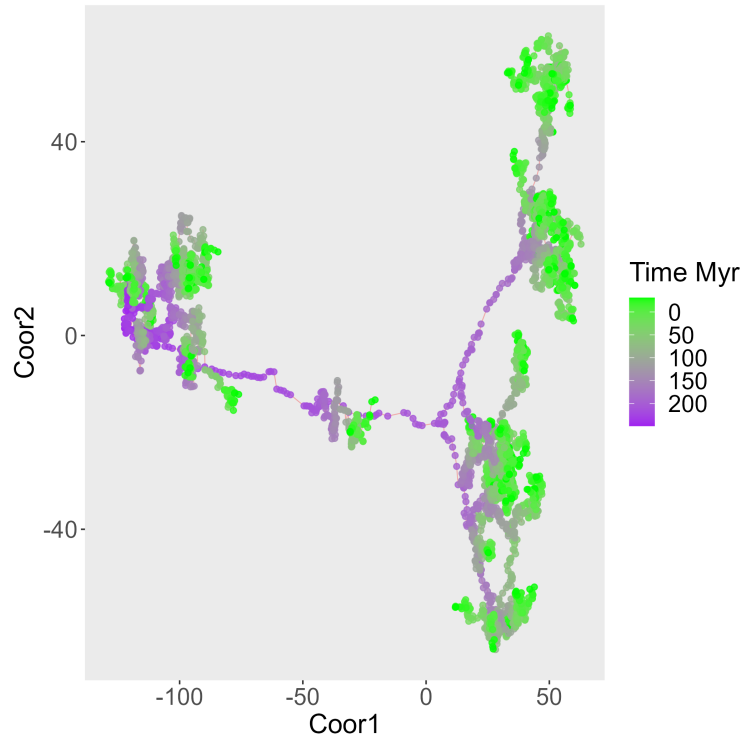


Figure 2: Morphospaces of the full phenome of Hymenoptera.

```
# Add transitions.
mds_plot_ph_anim <- mds_plot_ph + transition_reveal(time, keep_last = TRUE) +
  labs(title = "Myr: {ceiling(abs(as.integer(frame_along)-Tmax))}")

# Animate temporal slices.
animate(mds_plot_ph_anim, height = 500, width = 600, nframes = 100, res = 100)

# Save gif #
anim_save("figures/phenotype_flux.gif")
```

And finally, save all the results obtained so far.

```
save.image("RData/step4_mds.RData")
```

References

Sharkey, M. J., Carpenter, J. M., Vilhelmsen, L., Heraty, J., Liljeblad, J., Dowling, A. P., Schulmeister, S., Murray, D., Deans, A. R., Ronquist, F., et al. (2012). Phylogenetic relationships among superfamilies of hymenoptera. *Cladistics*, 28(1):80–112.