# Suplementary S1

# Diversification process between monogenoids (Dactylogyridae) and their marine

# catfish (Siluriformes: Ariidae) from the Atlantic coast of South America

# G.B. Soares, E.A. Adriano, M.V. Domingues, J.A. Balbuena

# Address correspondence to j.a.balbuena@uv.es

# R script to Compute Cophylospace R2 (Blasco-Costa et al. 2021)

# Input files available on Zenodo https://doi.org/10.5281/zenodo.7074210

#

##### Script starts here #######################################################

# Load libraries required

library**(**paco**)**

library**(**phytools**)**

# 1. COMPUTE PROCRUSTES R2 IN EACH COPHYLOSPACE DIMENSION

# 1.1. COMPARISON PHYLOGENY H - PHYLOGENY P

# Read trees

Htree **<-** read.tree**(**"host\_tree.txt"**)**

Ptree **<-** read.tree**(**"parasite\_tree.txt"**)**

# Read HP assoc matrix

HP **<-** read.table**(**"Dataset\_4\_matrix\_Parasite\_Host\_association.txt",

 header **=** **TRUE**, row.names **=** 1**)**

HP **<-** as.matrix**(**HP**)**

# Compute patristic distances

DH **<-** cophenetic**(**Htree**)**

DP **<-** cophenetic**(**Ptree**)**

D **<-** prepare\_paco\_data**(**DH, DP, HP**)**

D **<-** add\_pcoord**(**D**)**

D **<-** PACo**(**D, nperm **=** 10000, method**=**"r0", symmetric **=** **TRUE**,

 proc.warnings **=** **TRUE**, shuffled **=** **TRUE)**

print**(**D**$**gof**)**

R2a **<-** 1 **-** D**$**gof**$**ss

print**(**R2a**)**

# 1.2. COMPARISON PHYLOGENY P - MORPHOLOGY H

HMtree **<-** read.tree**(**"Dataset\_3\_matrixdistance\_of\_host\_morfo\_Marceniuk2012.txt"**)**

Ptree **<-** read.tree**(**"parasite\_tree.txt"**)**

DH **<-** cophenetic**(**HMtree**)**

DP **<-** cophenetic**(**Ptree**)**

D **<-** prepare\_paco\_data**(**DH, DP, HP**)**

D **<-** add\_pcoord**(**D**)**

D **<-** PACo**(**D, nperm **=** 10000, method**=** "r0", symmetric **=** **TRUE**,

 proc.warnings **=** **TRUE**, shuffled **=** **TRUE)**

print**(**D**$**gof**)**

R2HM **<-** 1 **-** D**$**gof**$**ss

print**(**R2HM**)**

# 1.3. COMPARISON PHYLOGENY H - MORPHOLOGY P

# 1.3.1 VENTRAL ANCHORS

Htree **<-** read.tree**(**"host\_tree.txt"**)**

ProAVtree **<-** read.tree**(**"Dataset\_1\_matrix\_shape\_coordinates\_Anchor\_Ventral.txt"**)**

DH **<-** cophenetic**(**Htree**)**

DP **<-** cophenetic**(**ProAVtree**)**

D **<-** prepare\_paco\_data**(**DH, DP, HP**)**

D **<-** add\_pcoord**(**D**)**

D **<-** PACo**(**D, nperm **=** 10000, method **=** "r0", symmetric **=** **TRUE**,

 proc.warnings **=** **TRUE**, shuffled **=** **TRUE)**

print**(**D**$**gof**)**

R2ProAV **<-** 1 **-** D**$**gof**$**ss

print**(**R2ProAV**)**

# 1.3.2 DORSAL ANCHORS

Htree **<-** read.tree**(**"host\_tree.txt"**)**

ProADtree **<-** read.tree**(**"Dataset\_2\_matrix\_shape\_coordinates\_Anchor\_Dorsal.txt"**)**

DH **<-** cophenetic**(**Htree**)**

DP **<-** cophenetic**(**ProADtree**)**

D **<-** prepare\_paco\_data**(**DH, DP, HP**)**

D **<-** add\_pcoord**(**D**)**

D **<-** add\_pcoord**(**D, correction **=** **)**

D **<-** PACo**(**D, nperm **=** 10000, method **=** "r0", symmetric **=** **TRUE**,

 proc.warnings **=** **TRUE**, shuffled **=** **TRUE)**

print**(**D**$**gof**)**

R2ProAD **<-** 1 **-** D**$**gof**$**ss

print**(**R2ProAD**)**

# 2. COMPUTE PROCRUSTES R2 AND CONFIDENCE INTERVALS BASED ON

# POST-PROBABILITY TREES

## Read post-probability trees; parasites and hosts

Ptrees **<-** read.nexus**(**"parasite\_Tree\_run1000tree.t"**)**

Ptrees **<-** Ptrees**[-**c**(**1,502**)]** # remove burn in set

# Trees need to be rooted correctly

Ptrees **<-** lapply**(**Ptrees, **function(**x**)** ape**::**root**(**x, outgroup **=** c**(**"G\_bu", "G\_co"**)))**

class**(**Ptrees**)** **<-** "multiPhylo"

# Root taxa dropped

Ptrees **<-** lapply**(**Ptrees, **function(**x**)** ape**::**drop.tip**(**x, tip **=** c**(**"G\_bu", "G\_co"**)))**

class**(**Ptrees**)** **<-** "multiPhylo"

Htrees **<-** read.nexus**(**"host\_Tree\_run2000tree.t"**)**

Htrees **<-** Htrees**[-**c**(**1,502**)]** # remove burn in set

# Trees need to be rooted correctly

# We tried to root trees at c("G\_at", "G\_fe", "G\_pe", "I\_pu", "C\_sp", "A\_at")

# but some taxa were not always at the root

# Since the consensus tree shows "A\_at" at root, trees were rooted as follows

Htrees **<-** lapply**(**Htrees, **function** **(**x**)** ape**::**root**(**x, outgroup **=**"A\_at"**))**

class**(**Htrees**)** **<-** "multiPhylo"

# Root taxa dropped

Htrees **<-** lapply**(**Htrees, **function** **(**x**)** ape**::**drop.tip**(**x,

 tip **=** c**(**"G\_at", "G\_fe", "G\_pe", "I\_pu", "C\_sp", "A\_at"**)))**

class**(**Ptrees**)** **<-** "multiPhylo"

# There are 2000 host trees. Select 1000 randomly

selectT **<-** sample**(**1**:**2000, 1000**)**

Htrees **<-** Htrees**[**selectT**]**

# Compute patristic distances post-probability trees

DisH **<-** lapply**(**Htrees, cophenetic**)**

DisP **<-** lapply**(**Ptrees, cophenetic**)**

# 4.1 PHYLOGENY H - PHYLOGENY P

# Compute R2 post-probability trees

R2 **<-** rep**(NA**, length**(**DisH**))**

**for(**i **in** 1**:**length**(**DisH**))** **{**

 D **<-** prepare\_paco\_data**(**DisH**[[**i**]]**, DisP**[[**i**]]**, HP**)**

 D **<-** add\_pcoord**(**D**)**

 D **<-** PACo**(**D, nperm **=** 1, method **=** "r0", symmetric **=** **TRUE)**

 R2**[**i**]** **<-** 1 **-** D**$**gof**$**ss

**}**

# 4.2 PHYLOGENY P - MORPHOLOGY H

# Read morphology data

MH **<-** read.tree**(**"Dataset\_3\_matrixdistance\_of\_host\_morfo\_Marceniuk2012.txt"**)**

DMH **<-** cophenetic**(**MH**)**

# Compute R2s

R2MH **<-** rep**(NA**, length**(**DisH**))**

**for(**i **in** 1**:**length**(**DisH**))** **{**

 D **<-** prepare\_paco\_data**(**DMH, DisP**[[**i**]]**, HP**)**

 D **<-** add\_pcoord**(**D**)**

 D **<-** PACo**(**D, nperm**=**1, method**=**"r0", symmetric **=** **TRUE)**

 R2MH**[**i**]** **<-** 1 **-** D**$**gof**$**ss

**}**

D **<-** prepare\_paco\_data**(**DMH, DP, HP**)**

D **<-** add\_pcoord**(**D**)**

D **<-** PACo**(**D, nperm**=**1, method**=**"r0", symmetric **=** **TRUE)**

R2CMH **<-** 1 **-** D**$**gof**$**ss

# 4.3 PHYLOGENY H - MORPHOLOGY P

# 4.3.1 VENTRAL ANCHORS

# Read P morphological data

MPv **<-** read.tree**(**"Dataset\_1\_matrix\_shape\_coordinates\_Anchor\_Ventral.txt"**)**

DMPv **<-** cophenetic**(**MPv**)**

#Compute R2s

R2MPv **<-** rep**(NA**, length**(**DisH**))**

**for(**i **in** 1**:**length**(**DisH**))** **{**

 D **<-** prepare\_paco\_data**(**DisH**[[**i**]]**, DMPv, HP**)**

 D **<-** add\_pcoord**(**D**)**

 D **<-** PACo**(**D, nperm**=**1, method**=**"r0", symmetric **=** **TRUE)**

 R2MPv**[**i**]** **<-** 1 **-** D**$**gof**$**ss

**}**

D **<-** prepare\_paco\_data**(**DH, DMPv, HP**)**

D **<-** add\_pcoord**(**D**)**

D **<-** PACo**(**D, nperm**=**1, method**=**"r0", symmetric **=** **TRUE)**

R2CMPv **<-** 1 **-** D**$**gof**$**ss

# 4.3.2 DORSAL ANCHORS

# Read P morphological data

MPd **<-** read.tree**(**"Dataset\_2\_matrix\_shape\_coordinates\_Anchor\_Dorsal.txt"**)**

DMPd **<-** cophenetic**(**MPd**)**

# Compute R2s

R2MPd **<-** rep**(NA**, length**(**DisH**))**

**for(**i **in** 1**:**length**(**DisH**))** **{**

 D **<-** prepare\_paco\_data**(**DisH**[[**i**]]**, DMPd, HP**)**

 D **<-** add\_pcoord**(**D**)**

 D **<-** PACo**(**D, nperm**=**1, method**=**"r0", symmetric **=** **TRUE)**

 R2MPd**[**i**]** **<-** 1 **-** D**$**gof**$**ss

**}**

D **<-** prepare\_paco\_data**(**DH, DMPd, HP**)**

D **<-** add\_pcoord**(**D**)**

D **<-** PACo**(**D, nperm**=**1, method**=**"r0", symmetric **=** **TRUE)**

R2CMPd **<-** 1 **-** D**$**gof**$**ss

# Plot results

boxplot**(**R2, R2MH, R2MPv, R2MPd, names **=** c**(**"Cophylogenetic signal",

 "Host interaction",

 "Parasite interaction (VA)",

 "Parasite interaction (DA)"**)**,

 ylab **=** "Procrustes R2"**)**

lines**(**c**(**1, 1**)**, quantile**(**R2, c**(**0.025, 0.975**))**, col **=**"red"**)**

lines**(**c**(**2, 2**)**, quantile**(**R2MH, c**(**0.025, 0.975**))**, col **=**"red"**)**

lines**(**c**(**3, 3**)**, quantile**(**R2MPv, c**(**0.025, 0.975**))**, col **=**"red"**)**

lines**(**c**(**4, 4**)**, quantile**(**R2MPd, c**(**0.025, 0.975**))**, col **=**"red"**)**

#Script ends here #############################################################

# REFERENCE

# Blasco-Costa et al. 2021. Trends in Ecology & Evolution 36: 907-918

# https://doi.org/10.1016/j.tree.2021.06.006