Software Details

January 27, 2024

Here we provide commands used to run the methods in our study.

1 Maximum Parsimony with PAUP*

This command is used for both multi-state maximum parsimony analyses (MP 1-4) as well as standard parsimony on the binary encoded data. We use Version 4.0a168 of PAUP* [Swofford, 2002] with the following command:

```
begin paup;
set criterion=parsimony maxtrees=100 increase=no;
hsearch start=stepwise addseq=random nreps=25 swap=tbr;
filter best=yes;
pscores all/ ci ri rc hi scorefile=paup_out.scores replace=yes;
savetrees file=paup_out.trees replace=yes format=nexus;
quit;
end;
```

2 UPGMA and NJ (Distance Methods)

For distance methods, we again use Version 4.0a168 of PAUP* [Swofford, 2002], providing with a custom distance matrix and the following command:

```
begin paup;
UPGMA;
savetrees file=paup_out.trees replace=yes format=nexus;
quit;
end;
```

When running NJ, `UPGMA` should be replaced with `NJ`.

3 Dollo Parsimony

We provide the following command to PAUP*:

```
begin assumptions;
options deftype=dollo.up;
end;

begin paup;
set criterion=parsimony maxtrees=100 increase=no;
hsearch start=stepwise addseq=random nreps=25 swap=tbr;
filter best=yes;
```

```
pscores all/ ci ri rc hi scorefile=paup_out.scores replace=yes;
savetrees file=paup_out.trees replace=yes format=nexus;
quit;
end;
```

4   GRAY-ATKINSON WITH MRBAYES

We use Version 3.2.7a of MrBayes [Huelsenbeck and Ronquist, 2001] with the following command:

```
begin mrbayes;
set autoclose=yes nowarn=yes;
lset rates=gamma;
mcmcp ngen=100000 printfreq=10000 samplefreq=500
nruns=1 nchains=4 savebrlens=yes filename=Bayes_out;
mcmc;
set nowarnings=yes;
sumt filename=Bayes_out burnin=100;
quit;
end;
```

Note that the character data type must also be set as `datatype=restriction`.

5   STOCHASTIC DOLLO

We use TraitLab [Nicholls et al., 2013] (downloaded from Github in October 2022) in batch mode, with the following configuration:

```
% FULL PATH OF DATA FILE INCLUDE .NEX EXTENSION
Data_file_name = input.nex

% ONE OF THE FOLLOWING THREE OPTIONS MUST BE ONE, THE OTHERS 0
Start_from_rand_tree = 1
Start_from_tree_in_output = 0
Start_from_true_tree = 0

% VALUE OF THETA IGNORED UNLESS Start_from_rand_tree == 1
Theta = 0.001000

% NEXT TWO FIELDS IGNORED UNLESS Start_from_tree_in_output == 1
% FULL PATH OF OLD OUTPUT FILE INCLUDE .NEX EXTENSION
Tree_file_name =
Use_tree = 0

Omit_taxa = 0
% LIST IS IGNORED UNLESS Omit_taxa == 1 CAN USE MATLAB VECTOR NOTATION
Omit_taxa_list =

Omit_traits = 0
% LIST IS IGNORED UNLESS Omit_traits == 1 CAN USE MATLAB VECTOR NOTATION
Omit_trait_list =

% ONE OF THE FOLLOWING TWO OPTIONS MUST BE 1, THE OTHER 0
Yule_prior_on_tree = 0
Flat_prior_on_tree = 1
```

```
% ONE OF THE FOLLOWING TWO OPTIONS MUST BE 1, THE OTHER 0
Uniform_prior_on_tree_topologies = 1
Uniform_prior_on_labelled_histories = 0
% FOLLOWING IS IGNORED UNLESS Flat_prior_on_tree == 1
Max_root_age = 16000

Vary_topology = 1
Account_rare_traits = 1

Impose_clades = 1
% LIST IS IGNORED UNLESS Impose_clades == 1 CAN USE MATLAB VECTOR NOTATION
Omit_clade_list =
Omit_clade_ages_list =

Vary_loss_rate = 1
% FOLLOWING IS IGNORED WHEN Random_initial_loss_rate == 1
Initial_loss_rate = 0.637419
Random_initial_loss_rate = 1

Account_for_lateral_transfer = 0
% FOLLOWING IS IGNORED WHEN Account_for_lateral_transfer == 0
Vary_borrowing_rate = 0
Random_initial_borrowing_rate = 0
% NEXT LINE IS IGNORED WHEN Random_initial_borrowing_rate == 1
Initial_borrowing_rate = 0

Include_catastrophes = 0
% NEXT 6 LINES ARE IGNORED WHEN Include_catastrophes = 0
% FOLLOWING IS IGNORED WHEN Random_initial_cat_death_prob = 1
Initial_cat_death_prob = 0.322368
Random_initial_cat_death_prob = 1
% FOLLOWING IS IGNORED WHEN Random_initial_cat_rate = 1
Initial_cat_rate = 8.15221e-05
Random_initial_cat_rate = 1

Model_missing = 0

Run_length = 200000
Sample_interval = 50

Seed_random_numbers = 0
% FOLLOWING IS IGNORED UNLESS Seed_random_numbers == 1
With_seed = 0

% OUTPUT FILE NAME OMITTING PATH AND ANY EXTENSIONS
Output_file_name = inf_tree1_1
% FULL PATH FOR DIRECTORY TO OUTPUT FILES
Output_path_name = output/

% Gaps are treated as missing data. To change this, edit GlobalValues.m.

Coupled_markov_chains = 0
% FOLLOWING IS IGNORED WHEN Coupled_markov_chains == 0
Coupling_lag = 0
```

We allow each taxon time to vary between the present (time 0) and the max root time (16,000) by imposing the following clade constraint for each taxon in the input Nexus file:

```
begin clades;
CLADE NAME = t1
ROOTMIN = 0
ROOTMAX = 15999
TAXA = t1;

...

end clades;
```

References

J. P. Huelsenbeck and F. Ronquist. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17(8):754–755, 2001.

G. K. Nicholls, R. J. Ryder, and D. Welch. TraitLab: a MatLab Package for Fitting and Simulating Binary Trait-Like Data, 2013. available online at `https://www.stats.ox.ac.uk/~nicholls/TraitLab/`.

D. L. Swofford. *Phylogenetic analysis using parsimony (PAUP*) 4.0.* Sinauer Associates: Sunderland, MA, USA, 2002.