HyPhy on HPC

What is HyPhy?

HyPhy (**Hy**pothesis Testing using **Phy**logenies) is an open-source HyPhy package for the analysis of genetic sequences (in particular the inference of natural selection) using techniques in phylogenetics, molecular evolution, and machine learning. It features a rich scripting language for limitless customization of analyses. Additionally, HyPhy features support for parallel computing environments (via message passing interface).

[Source]

Links:

Official Website

Documentation

Versions Available:

The following versions are available on the cluster:

- Hyphy- v2.7.0

How to load HyPhy?

To load HyPhy, use the following commands:

```
#Load the HyPhy module
module load bio/hyphy/2.7.0
```

To verify if the module and dependencies are loaded correctly, use the following command.

```
#Show all the modules loaded
module list
```

This should list all the HyPhy and dependencies that are loaded. In this case, openmpi will be loaded.

How to use HyPhy?

To demonstrate the usage of this HyPhy, download the test file,

```
#Download test files
svn checkout https://github.com/veg/hyphy-
tutorials/trunk/docs/selection/data
```

To run HyPhy interactively, jump into a compute node

```
# Jump to compute node
srun -p main --qos main --ntasks=8 --ntasks-per-core=1 --mem=20g  --
pty bash


# Run HyPhy interactively

mpirun -np $SLURM_NTASKS HYPHYMPI -i
```

Follow following steps (CLI) after launching HyPhy interactively,

1. Select the appropriate analysis to run

- **[GUI]** Choose *Analysis:Standard Analyses:Positive Selection:BUSTED.bf*
- **[CLI]** When presented with the list of standard analysis options upon launch, choose *Positive Selection*, then option 4 (*Run the Branch-site Unrestricted Statistical Test for Episodic Diversification to test for evidence of episodic alignment-wide selective pressure.*)

2. **Universal** genetic code option
3. The file to process

- **[GUI]** In the file dialog, navigate to and select `HIV.nex`
- **[CLI]** Input the full path name to the file (make sure there is no trailing space), e.g. `/Users/sergei/Coding/hyphy-tutorials/selection/data/HIV.nex`

4. Confirm that the tree included in the file will be used

- **[GUI]** Type **y** into the bottom box of the console window and hit Enter
- **[CLI]** Type **y** and hit Enter

5. Choose all branches to include in the test (**All**)

- **[CLI]** You will need to type **d** and hit "Enter" after selecting the **All** option to exit the selection dialog.

The analysis will now run for a few minutes and produce the following output

```
[BUSTED] Selected 26 branches as the test (foreground) set: R20_239,R20_245,Node5,R20_240,R20_238,R20_242,Node4,R20_
[BUSTED] Obtaining initial branch lengths under the GTR model
[BUSTED] Log(L) = -2114.132335772669
[BUSTED] Fitting the unconstrained branch-site model
[BUSTED] Log(L) = -2039.992959126133. Unrestricted class omega = 104.6591567580357 (weight = 0.02032866068122922)
[BUSTED] Fitting the branch-site model that disallows omega > 1 among foreground branches
[BUSTED] Log(L) = -2076.666683221396
[BUSTED] Likelihood ratio test for episodic positive selection, p = 1.110223024625157e-16
```

For example, in this case the analysis inferred that

1. A proportion of sites (~0.02) is evolving with dN/dS > 1 (~105) along a subset of the branches (it is not known which).
2. Forcing dN/dS = 1 provides a significantly worse (p ~ 10^{-16}) fit to the data, i.e. rejects the hypothesis of no positive selection in the alignment.

In addition to this output, HyPhy will also generate a JSON file with a more detailed analysis output. The JSON will be written to same directory as the input alignment file, with `BUSTED.json` appended to the file name, e.g. `/Users/sergei/Coding/hyphy-tutorials/selection/data/HIV.nex.BUSTED.json` This JSON file can be visualized and explored with this web app, e.g., to obtain plots like those shown in the lecture slides

User can also run HyPhy non-interactively with sbatch script. Please see the tutorial on HyPhy website.

Where to find help?

If you are stuck on some part or need help at any point, please contact OIT at the following address.

https://ua-app01.ua.edu/researchComputingPortal/public/oitHelp