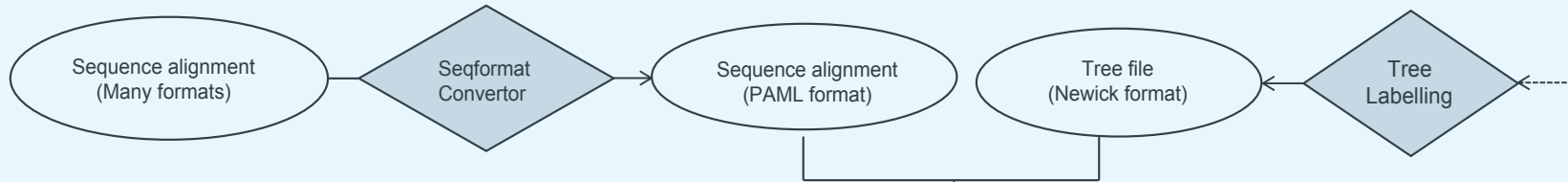EasyCodeML: A visual tool for analysis of selection using CodeML

A quick guide to EasyCodeML
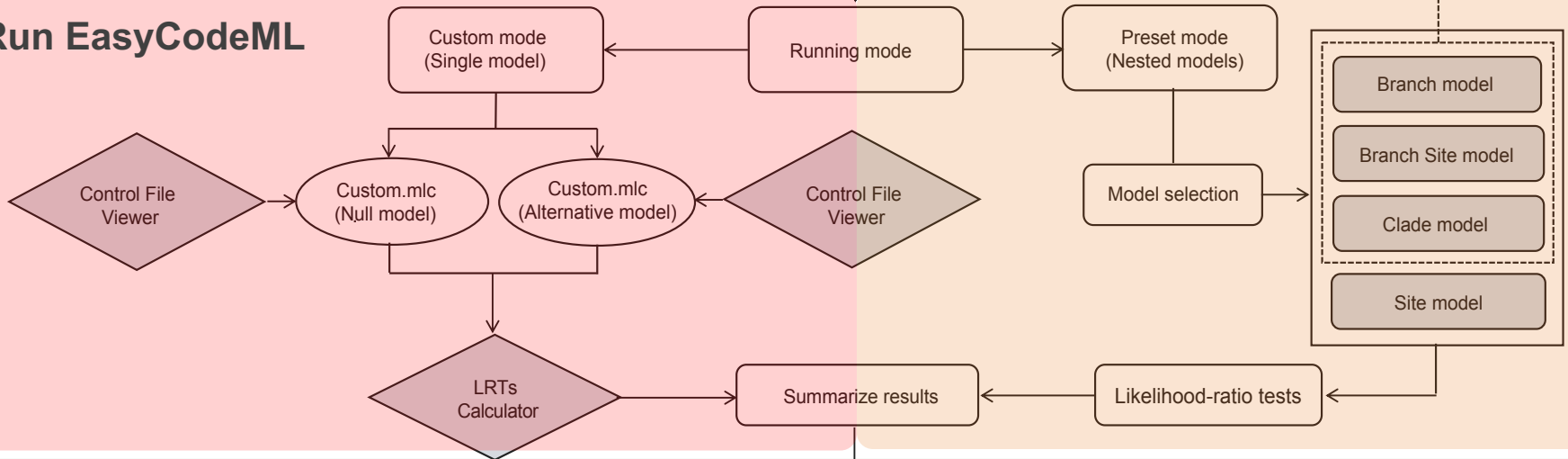
16 Jan, 2019
Fangluan Gao et. al

EasyCodeML workflow
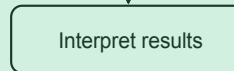
Prepare data

```
Sequence alignment
(Many formats)
```
→ Seqformat Convertor →
```
Sequence alignment
(PAML format)
```

```
Tree file
(Newick format)
```
← Tree Labelling

Run EasyCodeML

Running mode

Custom mode (Single model) ←

→ Preset mode (Nested models)

Custom.mlc (Null model)

Custom.mlc (Alternative model)

Control File Viewer →

← Control File Viewer

Model selection →

Branch model

Branch Site model

Clade model

Site model

LRTs Calculator →

Summarize results ←

Likelihood-ratio tests ←

Interpret results

Interpret results

I Prepare data

● **Sequence alignment in PAML format**

EasyCodeML fully supports PAML format (e.g., Examples/Example1.pml). We have incorporated a file-format convertor (Fig. 1) that can convert diverse types of data (e.g., Clustal, FASTA, MEGA, Nexus, and Phylip formats) into PAML format.
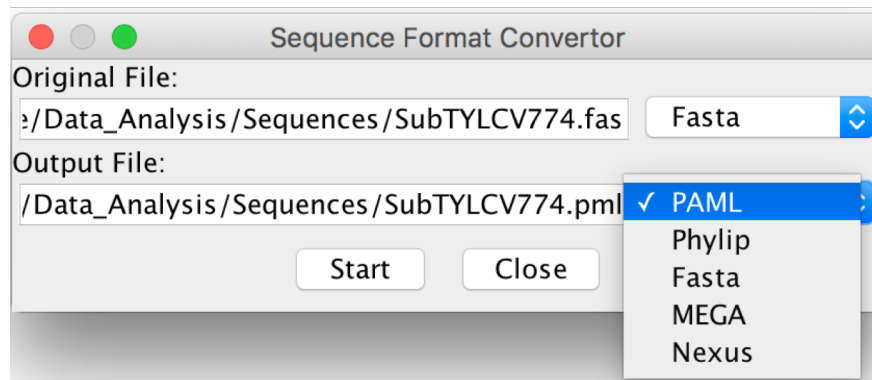

Fig. 1 Format conversion in Sequence Format Convertor

Ⅰ Prepare data

● **Tree file in Newick format**

The tree file must be in plain Newick format (e.g., Examples/Example1.tree; Fig. 2).

Branch lengths can be included and each node can have a single label. Taxon names

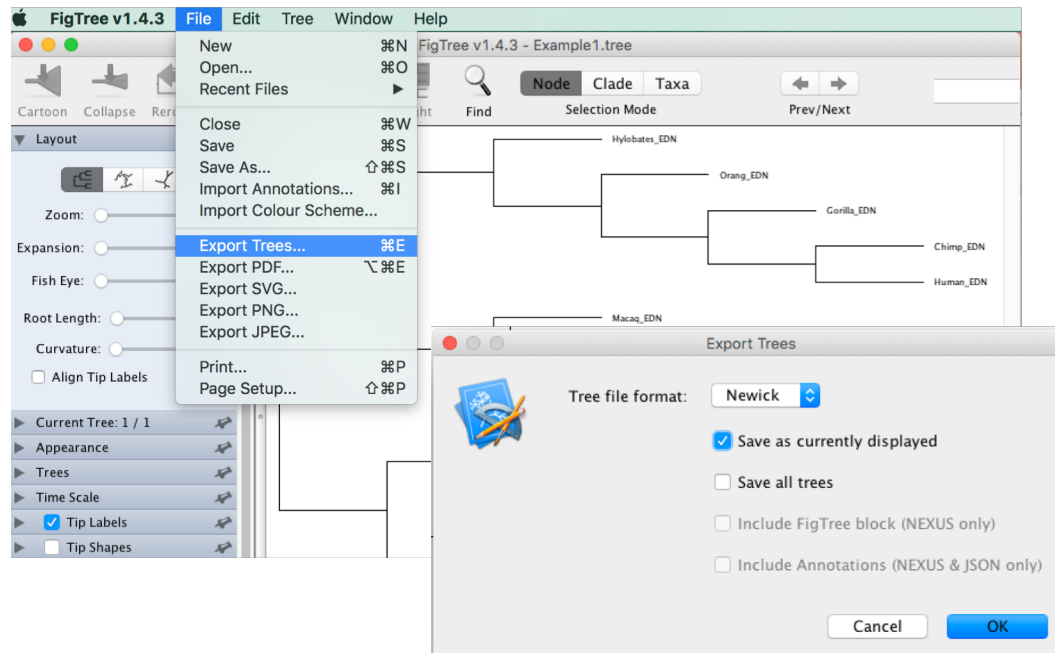cannot have illegal characters (such as spaces and semicolons).

Fig. 2 Exporting a tree in Newick format

Ⅰ Prepare data

● Tree labelling

Tree labelling is necessary when using the branch-related models (branch models, branch-site models, and clade models; Fig. 3), but not with the site models.

Fig. 3 Labelling branches in a tree for the (A) branch and branch-site and (B) clade models

Ⅱ **Run EasyCodeML**

● **Preset running mode**

1. Select a local folder as working directory

2. Select a model type for selection analysis

3. Load sequence alignment in PAML format

4. Load tree file in Newick format

5. Check consistency of taxon labels between the tree and sequence files

6. Label branches or clades in the phylogenetic tree (branch-related models only)

7. Save parameter settings

8. Start CodeML analysis in normal /silent modes

9. Perform a likelihood-ratio test for the nested models

10. Export a publication-quality table that contains the results

11. Launch Microsoft Excel to view the saved results file

Step 1
Load data and
configure parameters

Step 2
CodeML analysis

Step 3
Summarize results

Ⅱ   Run EasyCodeML

- **Custom running mode**

1. Select a local folder as working directory

2. Load sequence alignment in PAML format

3. Load tree file in Newick format

4. Choose a data type

5. Check consistency of taxon labels between the tree and sequence files

6. Label branches or clades in the phylogenetic tree (branch-related models only)

7. Load a codon model from Control File Viewer (Fig. 4)

8. Save parameter settings

9. Start CodeML analysis

10. View the saved results file

11. Run a likelihood-ratio test (Fig. 5)

II Run EasyCodeML

Step 1
Choose a running mode and input data

Step 2
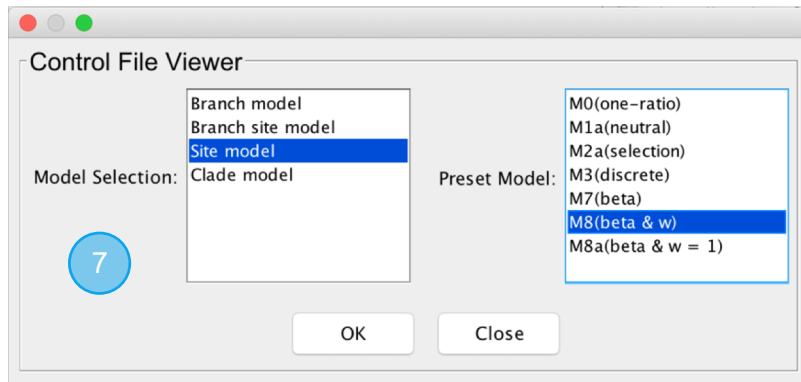Configure the running parameters

Step 3
CodeML analysis

Fig. 4 Control File Viewer that includes codon models with pre-optimized parameters
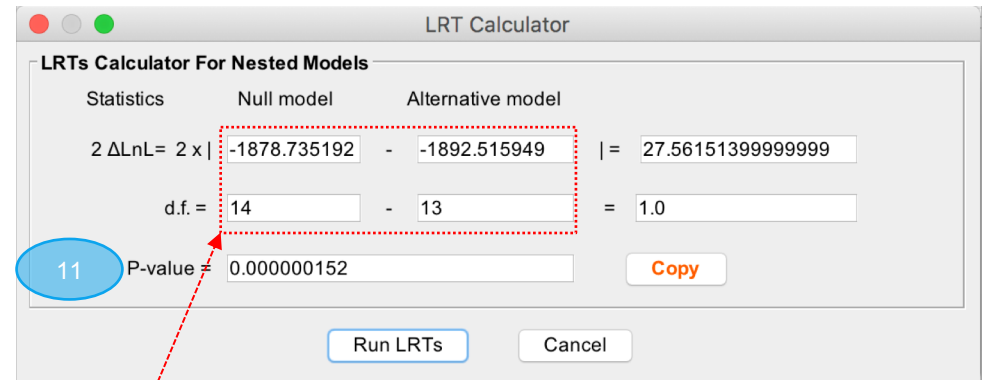
Fig. 5 LRT Calculator for running a likelihood-ratio test

M8a:
TREE # 1: ((((4, 6), 2), 5), 1, 3);   MP score: 58
lnL(ntime: 9 np: 13): -1892.515949   +0.000000
   7..8   8..9   9..10   10..4   10..6   9..2   8..5   7..1   7..3
  0.043343 0.011395 0.002754 0.005543 0.011176 0.000004 0.002428 0.040393 0.098672 1.288211 0.828982 0.005000
3.272293

M8:
TREE # 1: ((((4, 6), 2), 5), 1, 3);   MP score: 58
check convergence..
lnL(ntime: 9 np: 14): -1878.735192   +0.000000
   7..8   8..9   9..10   10..4   10..6   9..2   8..5   7..1   7..3
  0.057571 0.012156 0.002972 0.006652 0.013314 0.000004 0.003824 0.064314 0.163142 1.617384 0.966804 0.031342
0.166519 19.657673

Troubleshooting

- **Java 1.6 or higher**

  EasyCodeML requires Java Runtime Environment (JRE) 1.6 or higher.

  To check your version of Java, type java –version at a command prompt (CMD in Windows, terminal in Mac or Linux).

  The latest version of Java can be downloaded from http://www.java.com

- **Run EasyCodeML in debug mode**

  To run EasyCodeML in debug mode, type java -jar EasyCodeML.jar at the command prompt (CMD in windows or Terminal in Mac or Linux).

Contact us

- **EasyCodeML is an ongoing project**

  We welcome bug reports, feedback, and suggestions

- **https://github.com/BioEasy/EasyCodeML**

- **Contact us**

  Fangluan Gao   raindy@fafu.edu.cn
  Chengjie Chen   ccj0410@gmail.com