**Angiosperm\_Calculations.py**

This Python script facilitates statistical analyses on aligned FASTA file datasets containing trios of annual, perennial, and outgroup plant species. It offers functionalities for sequence alignment, Tajima's D test execution, and PHYLIP format conversion.

**Functionalities:**

1. **Alignment:** Aligns FASTA files in batch mode (entire directory) or compiled mode (single file).

-masce\_v2.07.jar must be in the same directory as the script

2. **Tajima's D1 test:** Performs a part Tajima's D1 test on FASTA files in batch mode or compiled mode. Outputs results (filename, M1, M2, D1, D2, chi-squared) to a CSV file. 3. **PHYLIP conversion:** Converts FASTA files to PHYLIP format (batch mode or compiled mode) for compatibility with PAML servers.

a. **Note:** This functionality currently supports FASTA files where the entire sequence for each species is on the same line directly after the '>' tag.

**Command-line Arguments:**

● -a, --annual: Fasta tag specifying the annual species.

● -p, --perennial: Fasta tag specifying the perennial species.

● -r, --outgroup: Fasta tag specifying the outgroup species.

● -t, --type: Type of sequences being analyzed (compiled or batch). ● -o, --output: Output directory or file (depends on the function).

● -i, --input: Input directory or file (depends on the function).

● -f, --function: Function to perform (align, tajima, phylip).

**Command-line Examples:**

1. **Batch Alignment:**

Bash

python3 Angiosperm\_Calculations.py -f align -i input\_dir -o output\_dir -t batch

This command aligns all FASTA files within the input\_dir directory and saves the aligned outputs in the output\_dir directory (batch mode).

2. **Compiled Alignment:**

Bash

python3 Angiosperm\_Calculations.py -f align -i input\_file.fasta -o path/to/output/filename(**no .phy**) -t compiled

This command aligns a single FASTA file named input\_file.fasta and saves the aligned output in the output\_dir directory (compiled mode).

3. **Batch Tajima's D test:**

Bash

python3 Angiosperm\_Calculations.py -f tajima -i input\_dir -o output.csv -t batch -a annual\_tag -p perennial\_tag -r outgroup\_tag

This command performs Tajima's D test on all FASTA files within the input\_dir directory. It outputs the results (M1, M2, D1, D2, chi-squared values) to a CSV file named output.csv. Replace annual\_tag, perennial\_tag, and outgroup\_tag with the corresponding Fasta tags for your species.

4. **Compiled Tajima's D test:**

Bash

python3 Angiosperm\_Calculations.py -f tajima -i input\_file.fasta -o output.csv -t compiled -a annual\_tag -p perennial\_tag -r outgroup\_tag

This command performs Tajima's D test on a single FASTA file named input\_file.fasta. It outputs the results (M1, M2, D1, D2, chi-squared values) to a CSV file named output.csv. Replace annual\_tag, perennial\_tag, and outgroup\_tag with the corresponding Fasta tags for your species.

5. **Batch PHYLIP conversion:**

Bash

python3 Angiosperm\_Calculations.py -f phylip -i input\_dir -o output\_dir -t batch

This command converts all FASTA files within the input\_dir directory to PHYLIP format and saves them in the output\_dir directory (batch mode).

6. **Compiled PHYLIP conversion:**

Bash

python3 Angiosperm\_Calculations.py -f phylip -i input\_file.fasta -o output\_dir -t compiled

This command converts a single FASTA file named input\_file.fasta to PHYLIP format and saves the output file in the output\_dir directory (compiled mode).