***run\_compare\_actual\_estimted\_SFS\_100reps.sh***

This code reads four input files and output the estimated and actual SFS of m population in each simulation scenario and sampling replication.

Input is

**collapse.MFA.mfa**: sequence having collapse sequences (input of BASEML)

**fltrst**: fltrst file which is the output of BASEML analysis

**sample\_seq\_m.txt**: sequences sampled from m population

**site\_used\_for\_collapse\_method.txt**: site which have < 3 states and used to make

collapse sequences

**BASEML\_input\_melpoly.MFA**: ancestral sequences of six species generated from

the simulation record.

you can set the SFS used for the weighting in BTW analysis in “***16\_make\_ancestral\_site\_probability\_mstyeo\_collapse\_only\_0%\_prob\_at\_ms\_weighted\_and\_AWP\_filter\_double\_poly.pl***” file.

@nPRF: weighting SFS of *pp* and *uu* mutation category

@puPRF: weighting SFS of *pu* mutation category

@upPRF: weighting SFS of *up* mutation category

The perl code above is only for 10 samples and,

(n, freq9, freq8, freq7, freq6, freq5, freq4, freq3, freq2, freq1, n)

**n**: no meaning

**freqX**: frequency of frequency class X among 10 samples

**EXAMPLE** folder contains examples of the five input files from a single replicate in stationary GC = 0.5 scenario.

Output file “act\_SFS\_m\_0%\_X.txt” shows the actual SFS of mutation category X.

Output file “est\_SFS\_m\_0%\_X.txt” shows the estimated SFS of mutation category X.