

Convert a Plink PED file to TreeMix:

1. Convert Genepop or whatever file to PED and MAP files for use in plink. These should have the same name, but different extensions (i.e. .ped and .map)
2. Convert this ped file to a binary ped file in Plink:  
`./plink --file InputFile --make-bed --out binaryOutput`
3. Now make a frequency cluster file sorting data into clusters (probably your sampling sites):  
`./plink --noweb --bfile binaryOutput --freq --within clusters.clust --out SNPSforTreemix`

The clusters.clust file should be a modified version of the .fam file that is created when you make the binary ped file. Open this .fam file in a text editor, remove all columns except 1 and 2, and add a third column with the cluster name you want to assign for every individual. In the end, column 1 is the 'Family' column, 2 is the individual ID, and 3 is the cluster. There cannot be any trailing spaces, and these numbers/strings should be separated by spaces only, not tabs. Also, not sure if it needs to be .clust when you're putting it into Plink, .txt should be fine too.

4. Now gzip this created output file.
5. Run the python script that came with TreeMix. Put the Plink output gzipped file in the same directory as the python script, and in the terminal type:  
`plink2treemix.py SNPSforTreemix.frq.gz treemixInput.frq.gz`
6. This output gzipped file can be run directly in TreeMix now!

\*\*\*TreeMix is fairly easy to use, the manual is helpful. Everything is run from the terminal, and then use R or RStudio to plot the resulting tree and residuals for migration events if you wish. Plotting functions for R are included in TreeMix and are also listed in the manual.