README file: Running a STRUCTURE analysis on the *Medicago polymorpha* full GBS population

Emily Helliwell, 9/11/2018

The purpose of this experiment is to use STRUCTURE (Pritchard et al., 2000) to find genotypic subpopulations of *M. polymorpha* based upon allelic frequencies. The input file (Mpol\_wUSDA\_structure\_input.txt) consists of 446 individuals and 9658 polymorphic loci. The organism is diploid, so there are two rows per individual, or one row for each allele. 1 = A, 2 = C, 3 = G, and 4 = T. -999 designates missing data.

The first column gives the unique identification of each individual, designated W0XXX with X being a numerical value.

It is important to note that for this experiment, STRUCTURE does not require a header row.

The file ‘structure\_mainparams.txt’ designates the specifications for each run.

The command is:

structure -m /mnt/home/f0001839/Structure/mainparams -e /mnt/home/f0001839/Structure/extraparams

To visualize our structure results, we constructed separate maps for the native and invaded ranges, with pie charts over each population. The code for each of these is in ‘Mpol native range jittered map.R’ and ‘Mpol invaded range jittered map.R’. There are no input files for either of these, everything stems from the structure output file pulled from the previous step and the GPS coordinates from each population, jittered by view.

Citation: Pritchard J. K., Stephens M., Donnelly P. (2000). Inference of population structure using multilocus genotype data. Genetics 155, 945–959