**Computer Lab 10 – Population Structure**

**Conservation Genetics (BIOL 4174 / 5174)**

Due date: 4/9/2018

Part I – Structure

1. Which model (Admixture or No Admixture) seems appropriate given your background knowledge of the populations being studied? Modify the extraparams file to reflect your selection.
2. Your Lab10 folder contains the “msat-data.xlsx” file with a sheet “AGF” showing allele frequencies for each locus as histograms (1 bar per population). Study these frequencies. Based on the similarities/differences among populations, which allele frequency model would you select? (correlated or independent). Modify the extraparams file to reflect your choice.

Part II – Interpreting Structure Results

1. Provide your bar plots for *K*=2, *K*=3, *K*=4, and *K*=5 here. You can use COMMAND+CONTROL+SHIFT+4 to screenshot, then paste with COMMAND+V.
2. Briefly describe your results at each *K* value. Discuss if there are any individuals with ancestry derived from multiple populations. Are there any sampling locations that have been split among multiple populations?
3. Overall, do you see any sampling locations that have consistently clustered together across all four *K* values? Do any sampling locations cluster together in most of the plots?
4. Why is it possible for Structure to find different clusters for the same *K* value?

Part III – Structure Harvester

1. In the L(K) plot, the standard deviation among replicates are plotted as error bars. Based on these, are your replicates converging on similar answers? Are there particlar K values with very high SD?
2. Copy and paste the Delta*K* plot into your homework document. Based on this graph, how many populations are represented by the data? Does the Delta*K* plot have a secondary peak at any other *K* value? If so, provide that value as well.