

# Lab #9

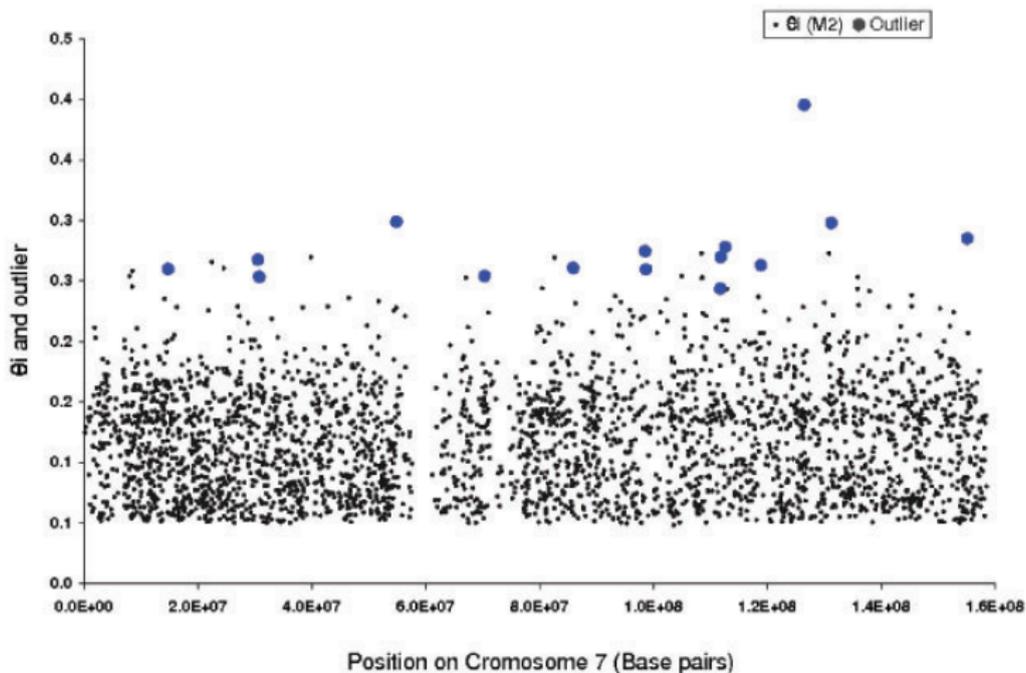


Nora Mitchell  
March 21, 2017

# Molecular Evolution

Can we detect a role for selection (or demographic processes) in molecular sequence data?

# Fst outliers



# Tajima's D

Test statistic for determining whether sequence data are consistent with population being at neutral mutation-drift equilibrium

- ▶  $\hat{D} = 0$  : no evidence for change in pop size or selection
- ▶  $\hat{D} < 0$  : pop size increasing, or purifying selection
- ▶  $\hat{D} > 0$  : pop bottleneck, or diversifying selection

## Back to *Protea repens*

Let's use *Protea repens* as an example, again.

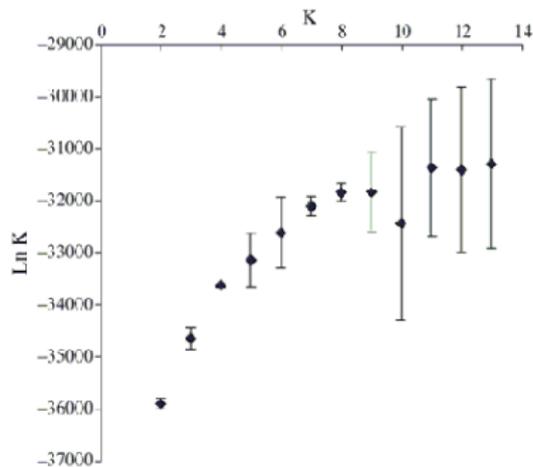
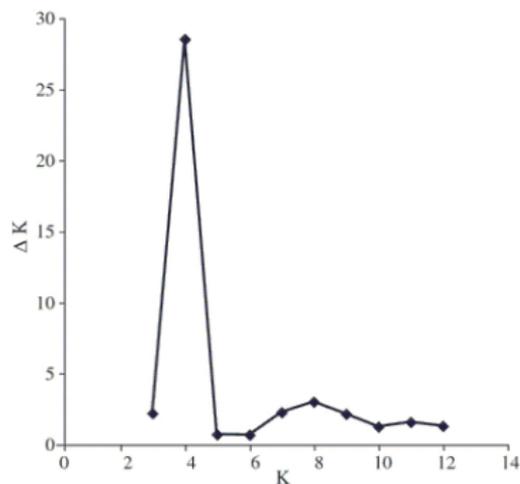
In Project #2, we used the Fst outlier loci and ran Structure on it, calculated Fst, etc.

Let's use a different subset of the data ( $n = 187$  individuals, 150 SNP loci) to try and detect Fst outliers

Then, let's use sequence data on just 8 individuals (but 274,405 bp) to estimate Tajima's D.

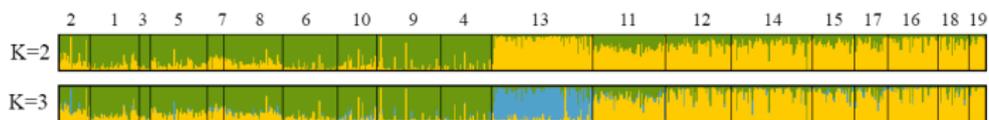
## Previous work

Remember that Prunier et al. found evidence for just 2-3 populations using non-Fst outlier loci

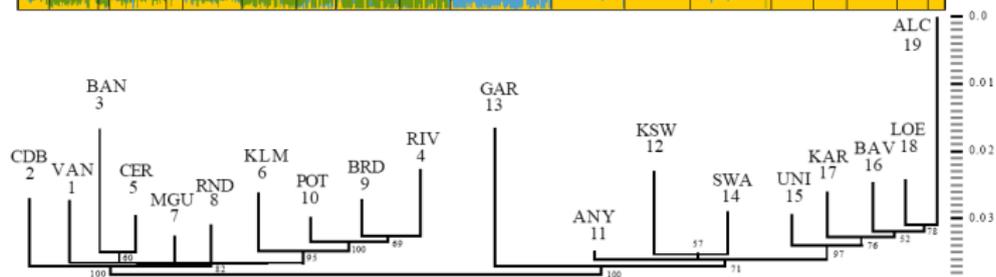


# Previous work

A

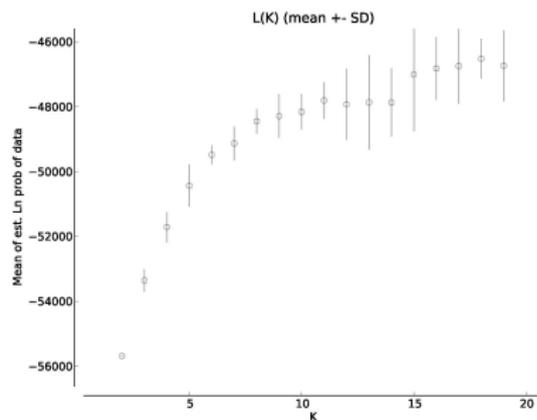
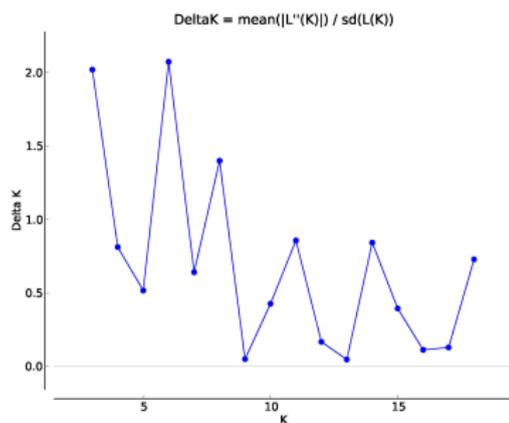


B



# Previous work

In project # 2, you found evidence for more populations



## To do:

- ▶ Use the handout as a guide
- ▶ Visually check for Fst outliers using the `repens_150.stru` data
- ▶ Run Structure on `repens_150.stru` and compare with previous findings
- ▶ Estimate the pieces of Tajima's D on `repens.fasta` data, then estimate Tajima's D using a built-in function
- ▶ Next week: project # 5 assigned!