

Population Genetics Homework - Week 10

Christian Polania

First we'll start with estimating Tajima's D in the complete sequences:

	D	Pval.normal	Pval.beta
cco	-0.4816287	0.6300697	0.6713393
cpk	-0.0040070	0.9968029	0.9735247
erd	-2.6511089	0.0080228	0.0001129
pp2	-2.3652443	0.0180182	0.0028890

And now only the coding sequences:

	D	Pval.normal	Pval.beta
cco	2.3553007	0.0185077	0.0200700
cpk	0.6781262	0.4976917	0.5072560
erd	-1.5036808	0.1326636	0.1218865
pp2	-2.3652443	0.0180182	0.0028890

1. Considering the complete sequences, it seems that:

- *ccoamt1* is likely to be evolving neutrally.
- *cpk3* is almost certainly evolving neutrally.
- *erd3* is almost certainly not evolving neutrally, suggesting a population expansion, bottleneck or selective sweep.
- *pp2c* is almost certainly not evolving neutrally, suggesting a population expansion, bottleneck, or selective sweep.

2. Considering the coding sequences, it seems that:

- *ccoamt1* now seems to be evolving, possibly due to a population contraction or balancing selection.
- *cpk3* is **still** probably evolving neutrally.
- *erd3* is still likely not evolving neutrally, but with less certainty than before.
- *pp2c* is **still** almost certainly not evolving neutrally.

3. Removing the possibility of population expansion or contraction, there are two patterns of selection suggested by the data:

- *ccoamt1*'s coding sequence has few rare alleles, implying strong selection for one or a few alleles (ie either directional or balancing selection).
- *erd3* and *pp2c* both have an excess of rare alleles, implying a recent selective sweep. Both rare lineages that "escaped" the sweep and recent mutation in swept genes can cause rare alleles to be more common than expected from neutral evolution.

Either kind of selection could be an indication of importance in drought response, as both suggest relatively strong selection.