

Day 2b – EMBO PopGen

We wish to study the evolution of spotted salamanders (*Ambystoma maculatum*) in Northern and Southern China. As such, we collected DNA samples from individuals from these two regions and obtained their genome. However, for practical reasons, we assess their genetic diversity only at three *loci*.

From Northern China, haplotypes for two diploid individuals over the three *loci* are

AAA

AAT

AGT

CAT

while for Southern China, two diploid individuals have haplotypes

AGT

AGT

AGT

AAT

.

We also have access to the putative ancestral haplotype over these three *loci*. The ancestral state of these *loci* is

AGT

(i)

Calculate the derived allele frequencies for the three *loci* for both populations.

(ii)

Calculate the unfolded site frequency spectrum for both populations.

(iii)

Estimate the population genetic parameter $\theta = 4N_e\mu$ using either Tajima's or Watterson's estimator for both populations.

(iv)

Assuming that both populations have been evolving under a scenario of constant population size and that their mutation rate is equal, discuss which population (Northern or Southern China) has a greater effective population size.