Module 3: Population and Landscape Genomics

* Assessing levels of diversity across the genome (DNA)
* How are populations structured?
  + Mating between pops
  + Topographical barriers
  + Pre/post zygotic barriers
  + Effective population size
  + Local adaptation/natural selection
* Strategies for identifying signatures of selection
  + Nucleotide diversity
* Looking at the relative difference between populations
  + Fst
    - A measure of population genetic structure
    - Can look at this across the genome and identify regions associated with selection
* Will be looking at a hybrid zone in poplar

Day 2

* Fastq files-> Trimmomatic (remove adapters, quality filter) -> Clean fastq files -> BWA (program to map to reference) -> bam files -> GATK -> VCF file -> VCFtools (filtering potential sequencing errors using minor allele frequency cut-off (0.002)) -> Filtered VCF file -> Beagle (imputing missing values) -> final VCF file
* The header of a VCF file contains lots of good metadata