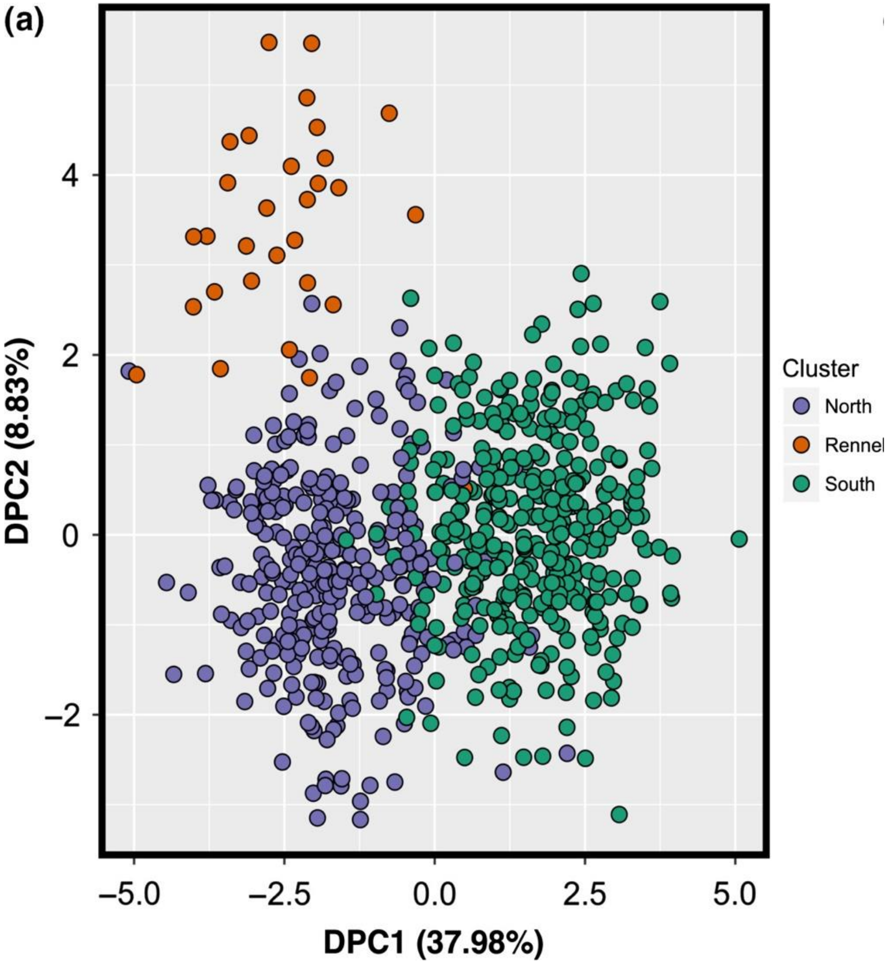
Genetic basis of local adaptation

Marine Genomics Week 9

From PCA



These populations look ”genetically different” - but why?

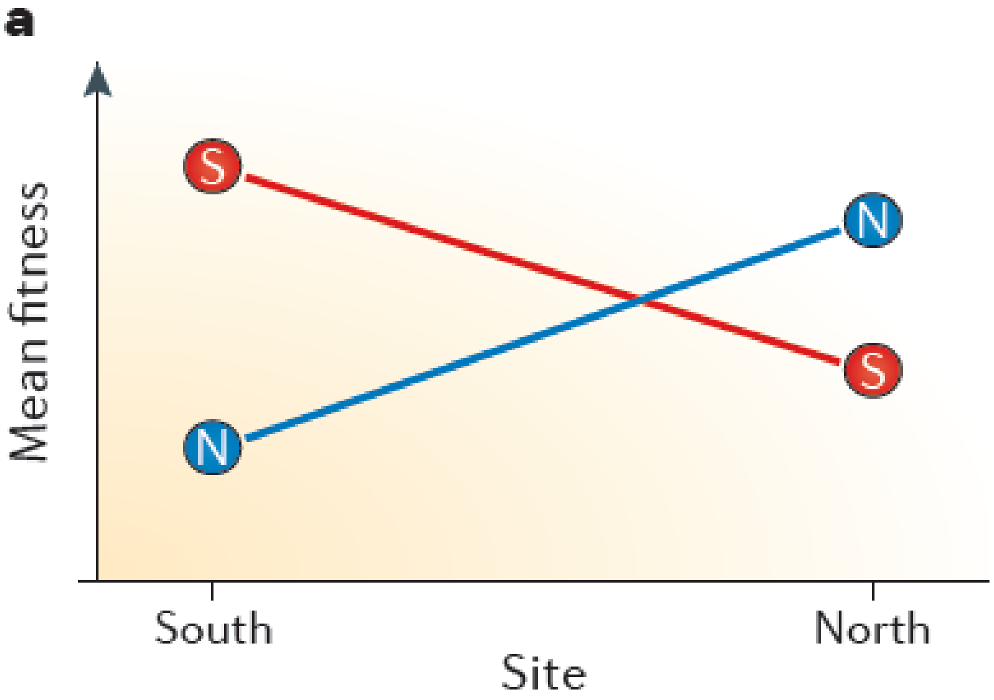
- Neutral differences
- Selection

Learning objectives

- Define local adaptation and how it is inferred from genomic data
- Describe basic measures of genetic diversity and population differentiation
- Use standard analyses to identify ‘outlier’ SNPs as candidates for local adaptation

What is local adaptation?

Local adaptation – individuals have the highest fitness at their home site

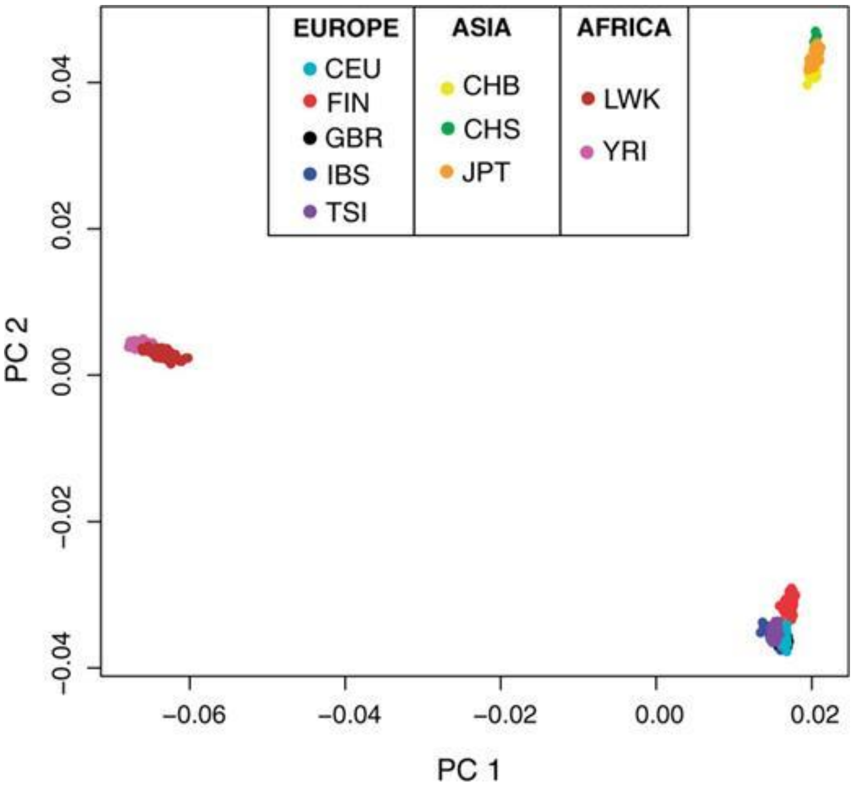


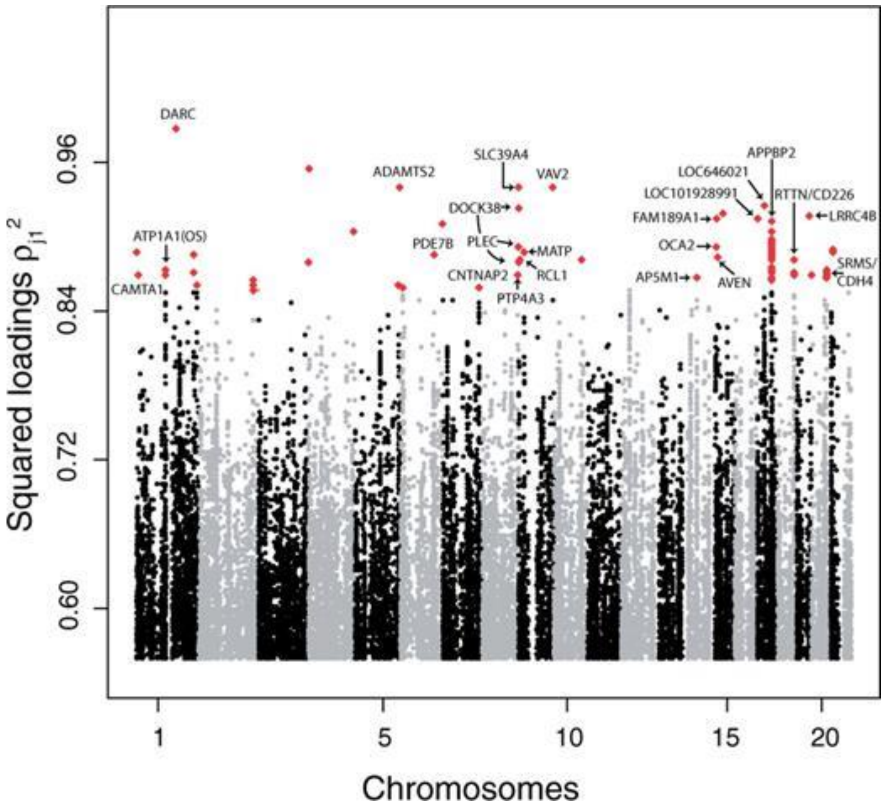
NOTE: We can’t know this from genetic data!

How do we identify the genomic basis of adaptation?

- Population genetic (genotype only)
  - Genetic differentiation outlier tests
- Phenotype mapping (genotype & phenotype)
  - Candidate genes
  - QTL mapping
  - Genome wide association studies (GWAS)
  - Bulk Segregant Analysis
  - Polygenic Scores
- Environment mapping
  - Genotype-Environment Association (GEA)

PCA-associated SNPs





Duforet-Frebourg et al. *Molecular Biology and Evolution* 2015

Quantifying genetic diversity: Allele frequencies

- ***Alleles***: B and b
    - *p*=frequency of B
        - 13/30=0.43
    - *q=frequency of b*
        - *17/30=0.57*
    - *p+q=1*

BB **bb** **Bb**

Bb **Bb** **Bb** **bb**

bb **Bb** **BB** **Bb**

Bb

Bb **bb**

Bb

Quantifying genetic diversity
Genotype frequencies and HWE

- Two alleles with frequencies *p* and *q:*
  - *What is the probability that a randomly chosen individual will have genotype BB? Bb? bb?*

Genotype	BB	Bb	bb
Frequency	p^2	$2pq$	q^2

- $p^2+2pq+q^2=1$

Quantifying genetic diversity
Genotype frequencies and HWE

Bb

BB **bb**

Bb **Bb** **Bb** **bb**

BB **Bb**

bb **Bb**

Bb

Bb **bb**

Bb

$p=0.43$, $q=0.57$

Why are predicted and observed values different?

Genotype	BB	Bb	bb
Frequency	p^2	2pq	q^2

Quantifying population differentiation

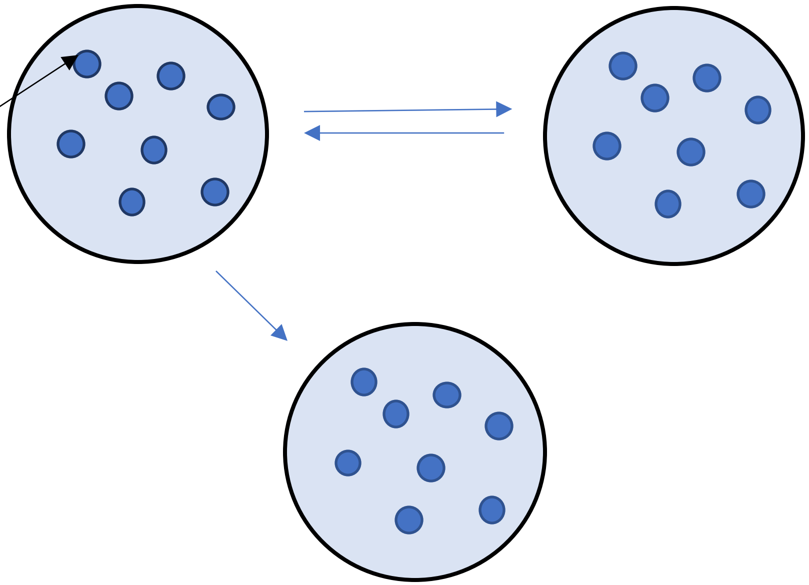
- F-statistics are based on inbreeding: $F = \frac{H_{exp} - H_{obs}}{H_{exp}}$

Differentiation between subpopulations (relative to total)

$$F_{ST} = \frac{H_T - H_S}{H_T}$$

H_T=expected heterozygosity of total population

H_S=expected subpopulation heterozygosity under HWE

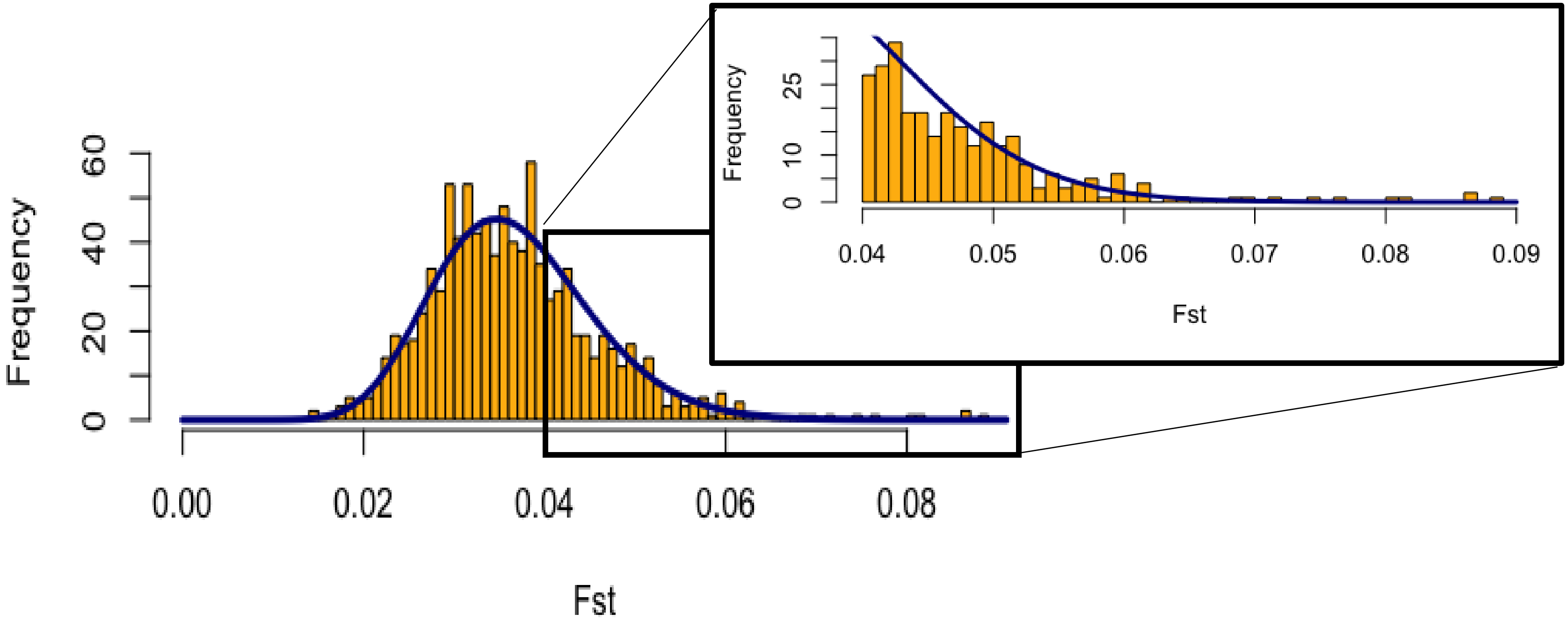


F_{ST} = 0 means no differentiation

F_{ST} = 1 means no shared alleles (complete differentiation)

Outlier analysis

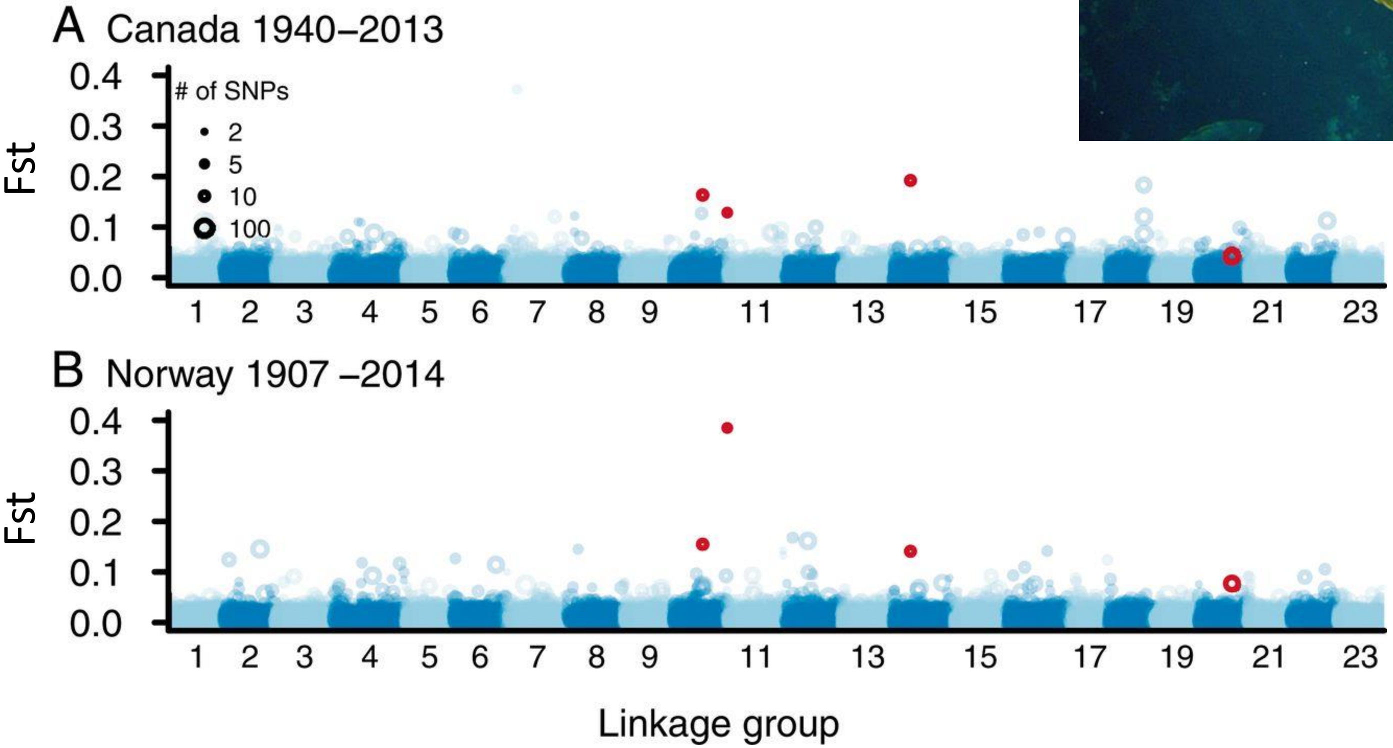
- Which loci are more differentiated than ’expected’
- Null distribution can be simulated for fit to ‘neutral’ data



Whitlock & Lotterhos 2015. *The American Naturalist*

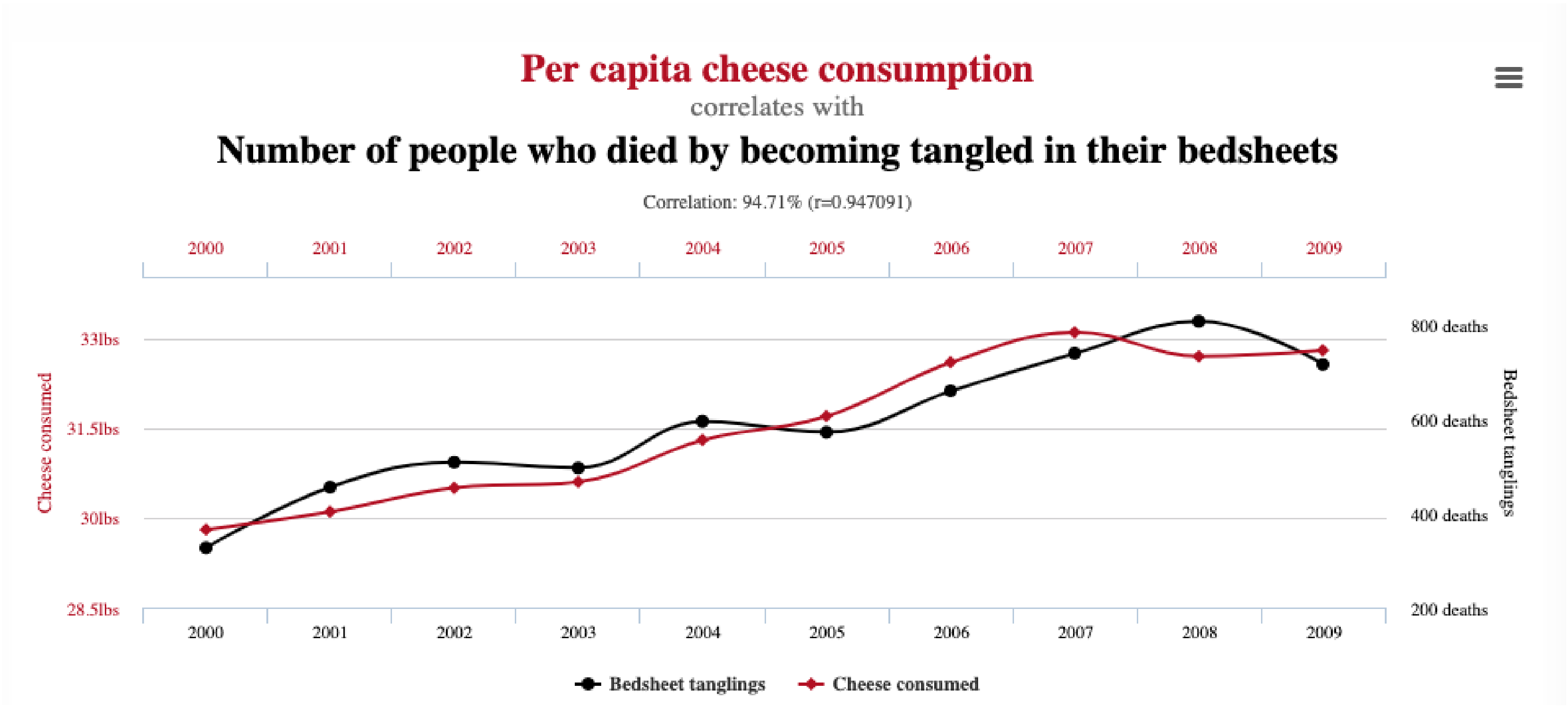
F_{ST} outlier analysis

Genome scans are often visualized with *Manhattan plots:*



Pinsky et al. *PNAS* 2021

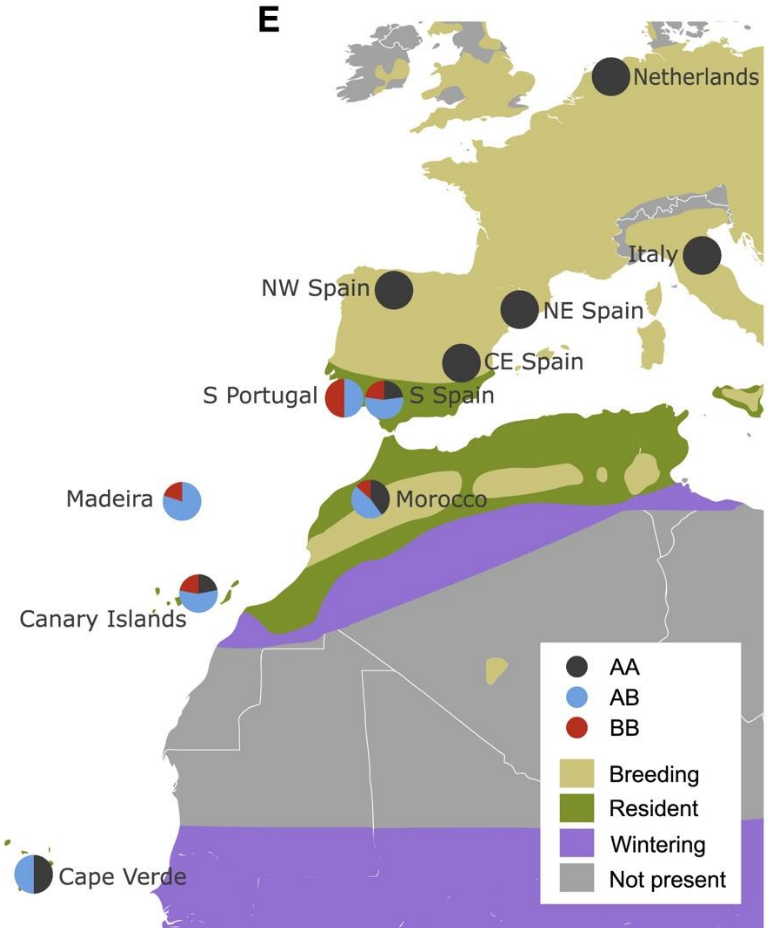
Caution: The problem with tons of data

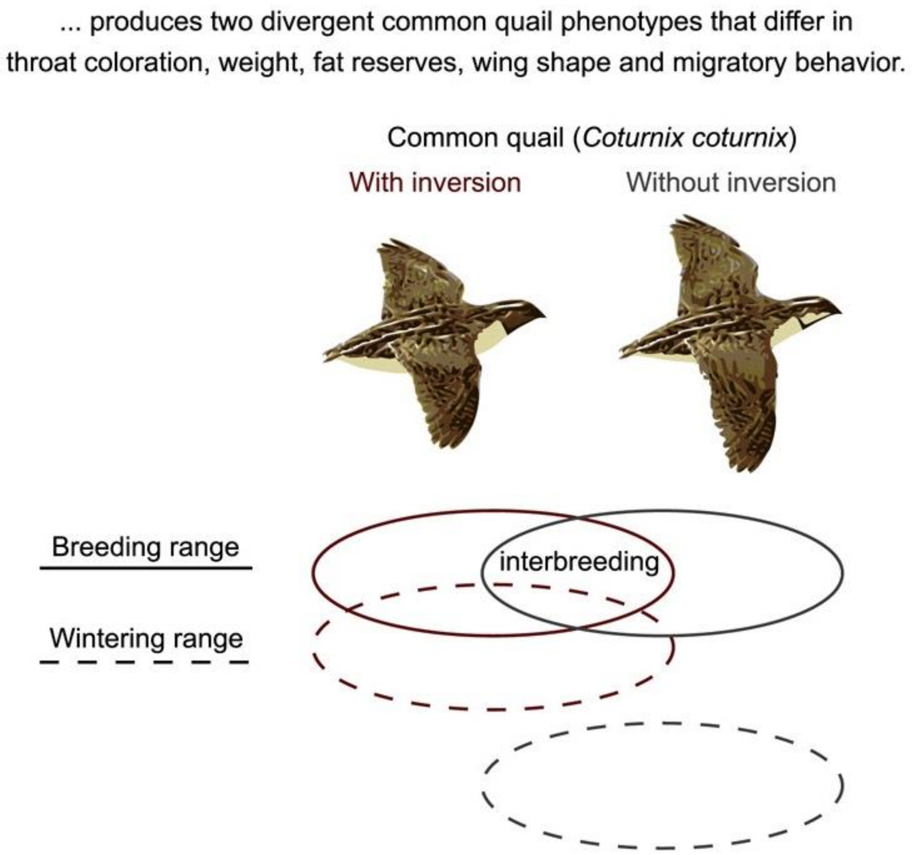


- When conducting >1000 tests, a
- Some results are likely to be ***false positives***
- We use p-value corrections to attempt to minimize this problem

Today’s exercise:

- Which SNPs drive differences among locations?





... produces two divergent common quail phenotypes that differ in throat coloration, weight, fat reserves, wing shape and migratory behavior.