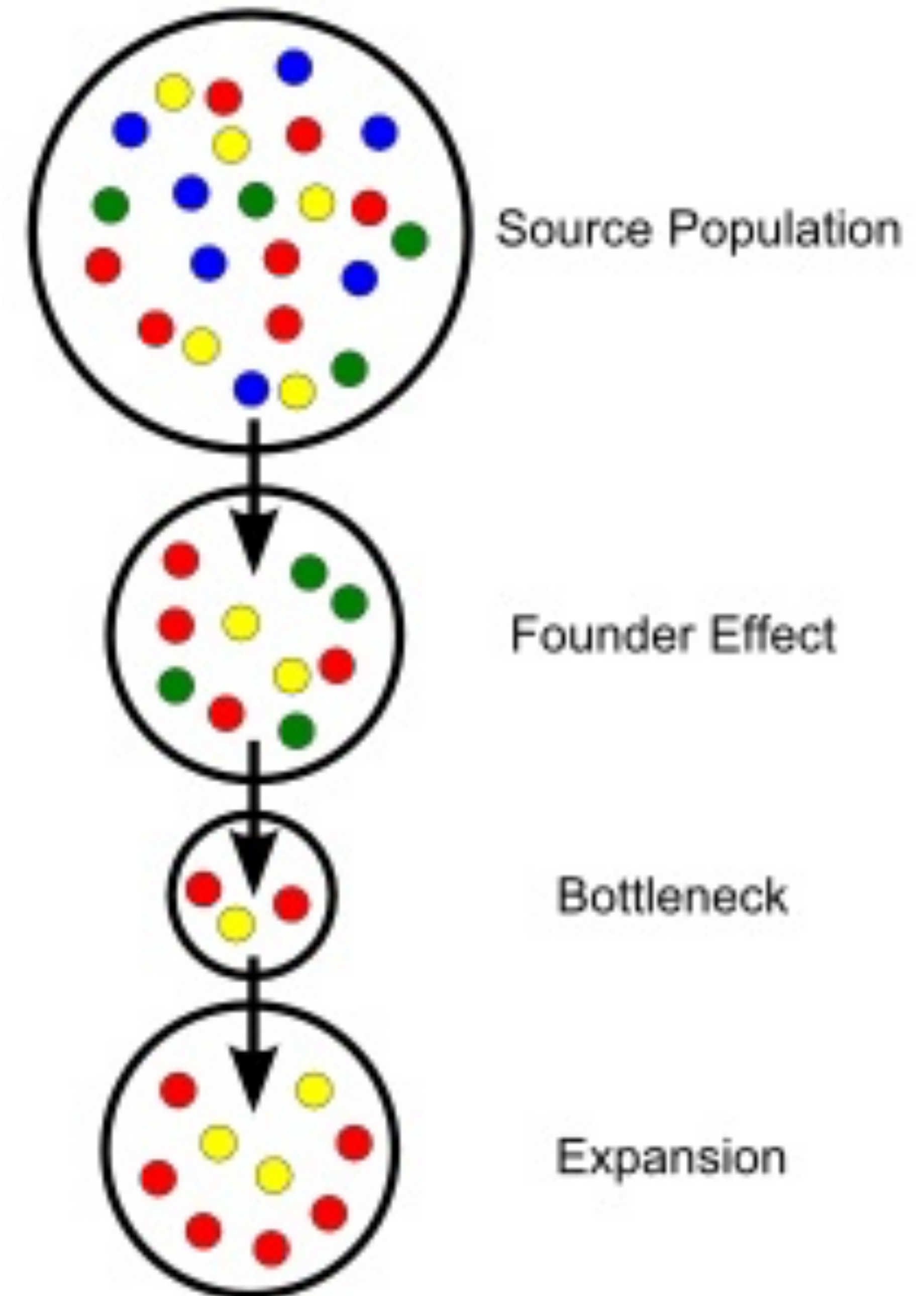# Basic population genomics

Dr Johanna Rhodes

Learning outcomes

- **Basic population genetics**

  - theory

  - using SNP data

Population genetics
## Using genomic data

- Study of genetic variation within populations

  - changes in frequencies of genes and alleles over space and time

  - many genes in a population will be polymorphic - they will occur in a number of different forms

Source Population

Founder Effect

Bottleneck

Expansion

Population genetics
Using genomic data

- Principles in fungal populations include:

  - mutation

  - gene flow

    - transfer of genes among populations

  - genetic drift

    - natural selection

Source Population

Founder Effect

Bottleneck

Expansion
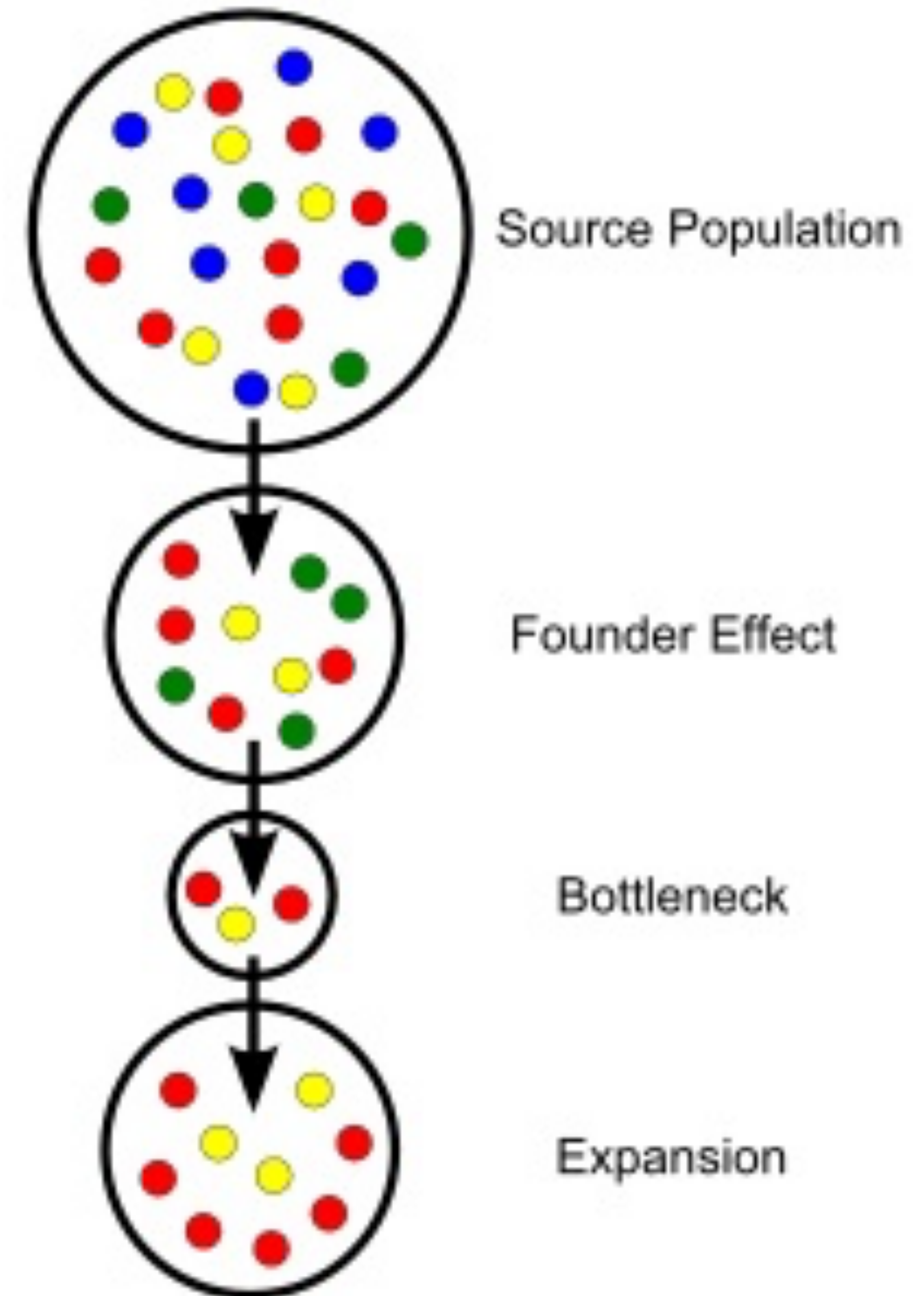
What is a population?
Using wgSNPs

- Principal component analysis (PCA)

  - reduces data to its essential features - principal components

  - principal components explain the variance of the variables

PCA coloured by countries
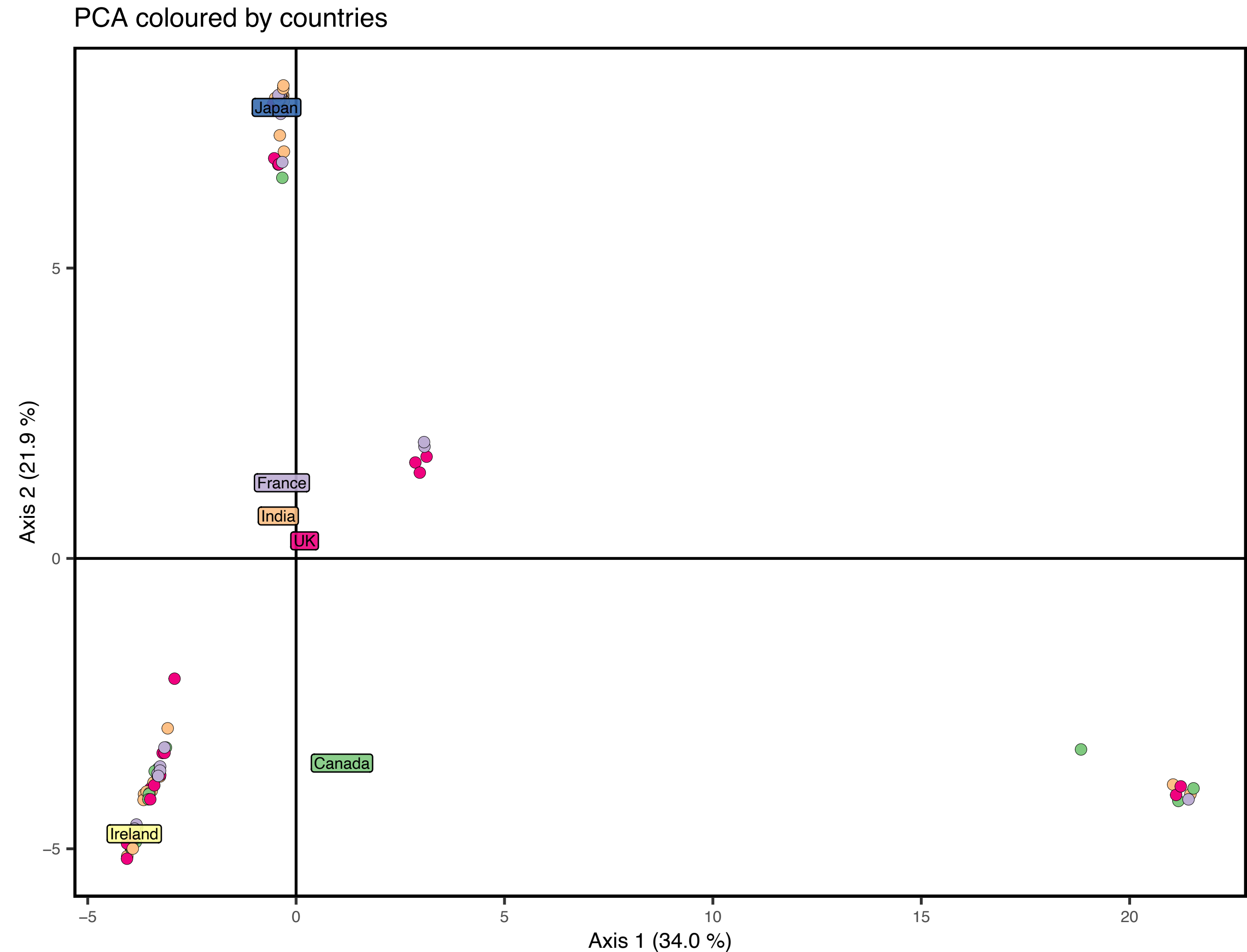
What is a population?
Using wgSNPs

- Discriminant analysis of principal components (DAPC)

  - PCA fails to discriminate groups accurately; DAPC maximises the separation between groups while minimising variation within groups
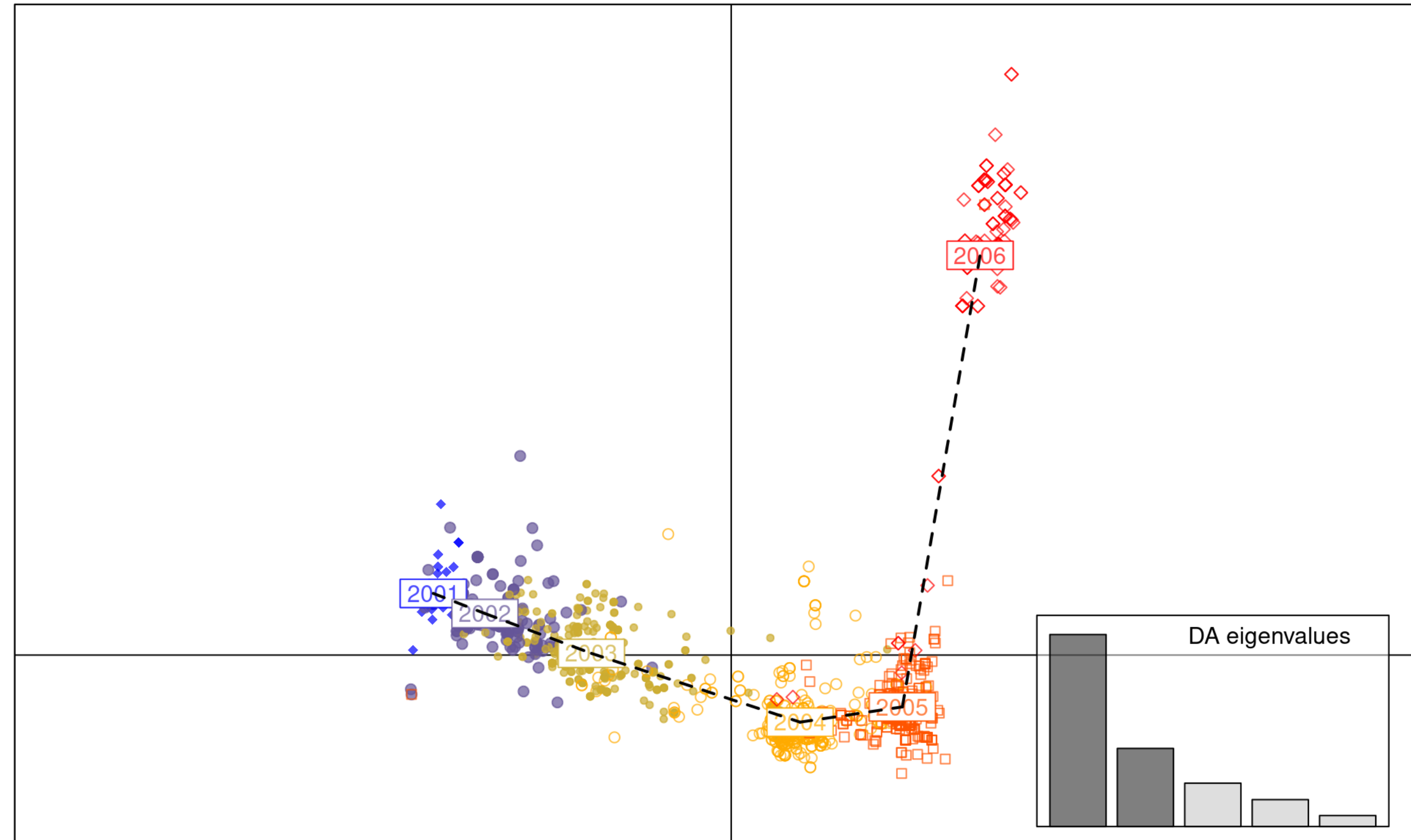
Principal component analysis (PCA)
In practice

- In R/RStudio

- Need:

  - multiFASTA file

  - installation of packages: ape, adegenet, viridis

  - txt file with all isolates plus group assignment

Discriminant Analysis of Principal Components (DAPC)
In practice

- Same inputs as for PCA. How does it look?

- Do other groupings explain the variation better?

  - finding the true 'k' value (hypothetically)

  - 'find.clusters(x)' in the dapc package

**Value of BIC
versus number of clusters**