Genomic tools for conservation & management

Part II

25334 Genomic methods in breeding and management of aquatic living resources

What to conserve?

The zoo directors, curators, geneticists and population biologists who attempt to pursue the elusive goal of preservation of adaptive genetic variation are now considering the question of which gene pools they should strive to preserve.

Oliver A. Ryder (1986)

What do **YOU** think it should be conserved?

What to conserve?

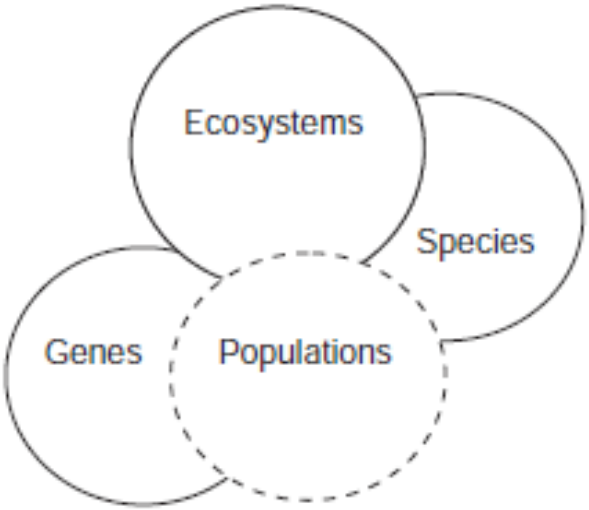


Figure 16.1 Primary levels of biodiversity recognized by the IUCN (solid circles), and a fourth level – populations – recognized as perhaps most crucial for species' long-term persistence (Hughes et al. 1997; Luck et al. 2003). In reality, biodiversity exists across a continuum of many hierarchical levels of organization including genes, genomes (i.e., multilocus genotypes), local populations, communities, ecosystems, and biomes. Additional levels of diversity include metapopulations, subspecies, genera, families, and so on.

What about fish?
Is conservation of populations of marine fish relevant?

- Local extinction in marine fish
- Dulvy et al. 2003 reported more than 60 local (population) extinctions of marine fish
- Most extinctions were due to exploitation
- There was generally a time lack of 53 years between last sighting and reported date of extinction

Which metrics to use to help us decide?

1. **Statistical estimators that** infer a population parameter using data that are related to that parameter.
   - For example, e**stimators** that measure rate of loss of genetic diversity (effective population size, N_e)
   - Measuring the degree of differentiation between populations (F_{ST})

2. **Multivariate** exploratory techniques: these allow us answering questions such as:
   - "are there major patterns in the data?", or
   - "can we assign individuals to groups (based on multilocus genotypes)?", or
   - "which variable (e.g., locus) is most useful (i.e., explains most the variance) when assigning individuals to groups?"

   For example, Principal Component Analysis or Loading plots

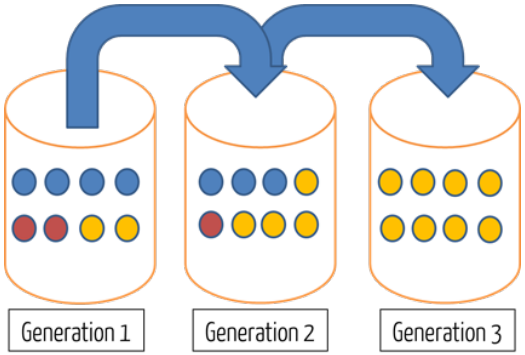
1. Estimators to help us decide what to conserve/manage:

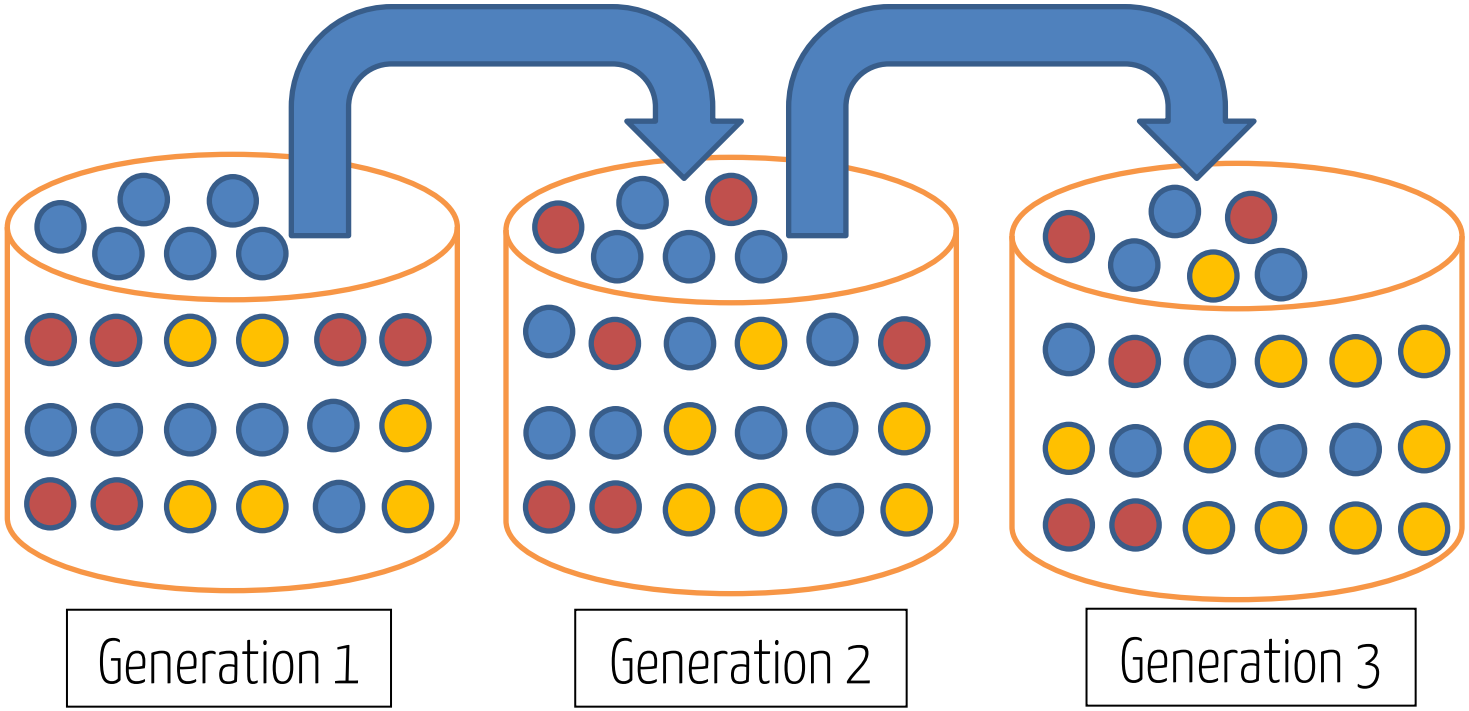
- The effective population size (N_e)
- F_{ST} or degree of differentiation

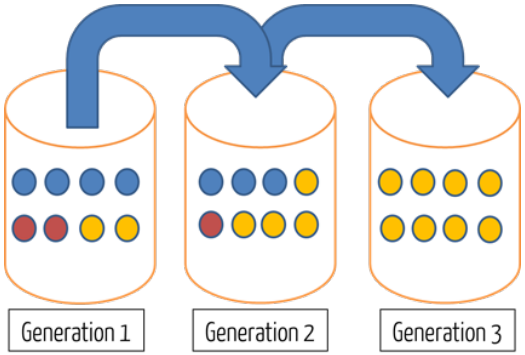
1. Estimators to help us decide what to conserve/manage:

- The effective population size (N_e)
- F_{ST} or degree of differentiation

GENETIC DRIFT







Male 1

Male 2

Male 3

Pride of females

Photos from Wikipedia, Andrea Cavallini and Tambako the Jaguar

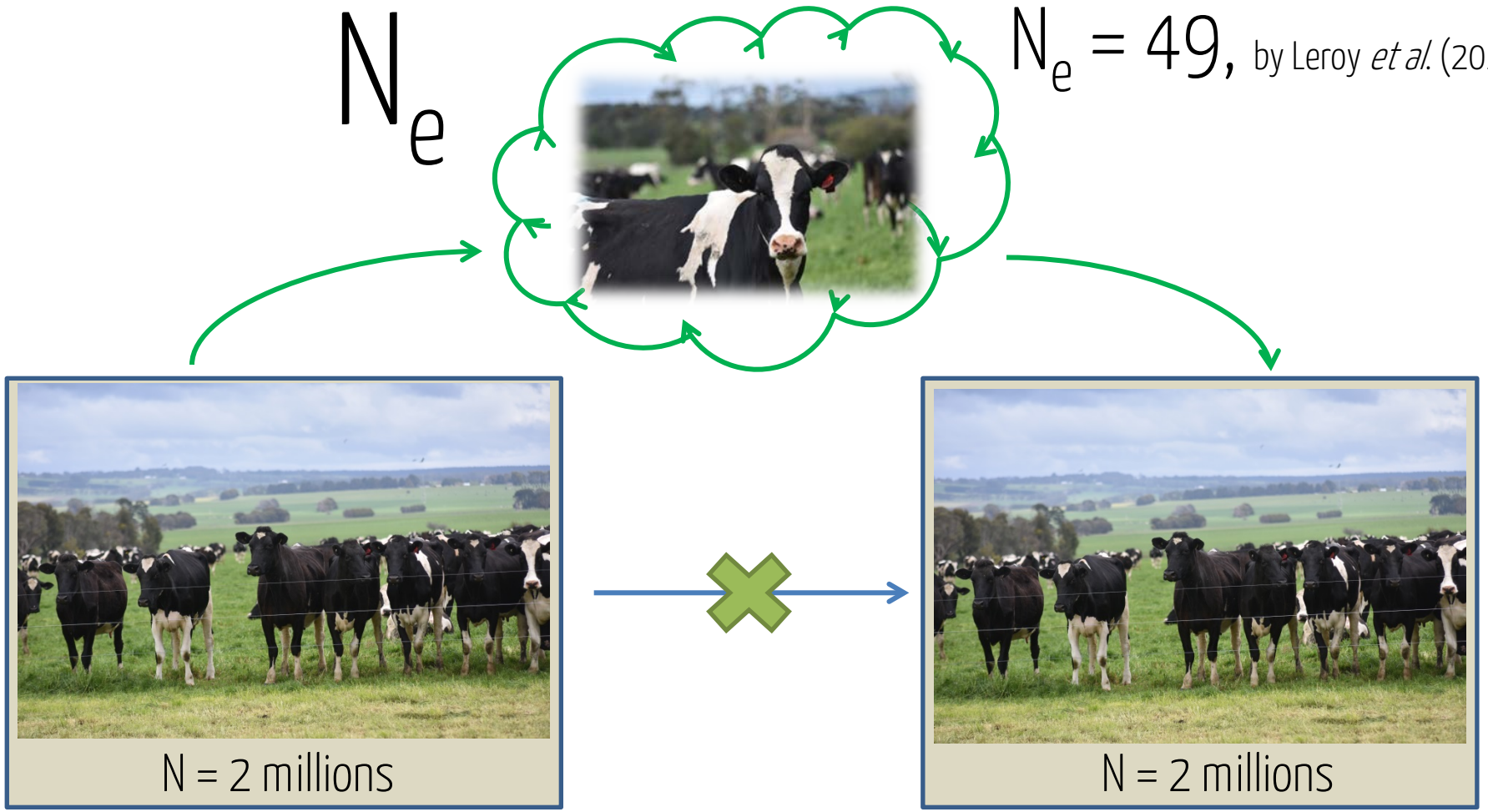
Effective population size or "N_e" =

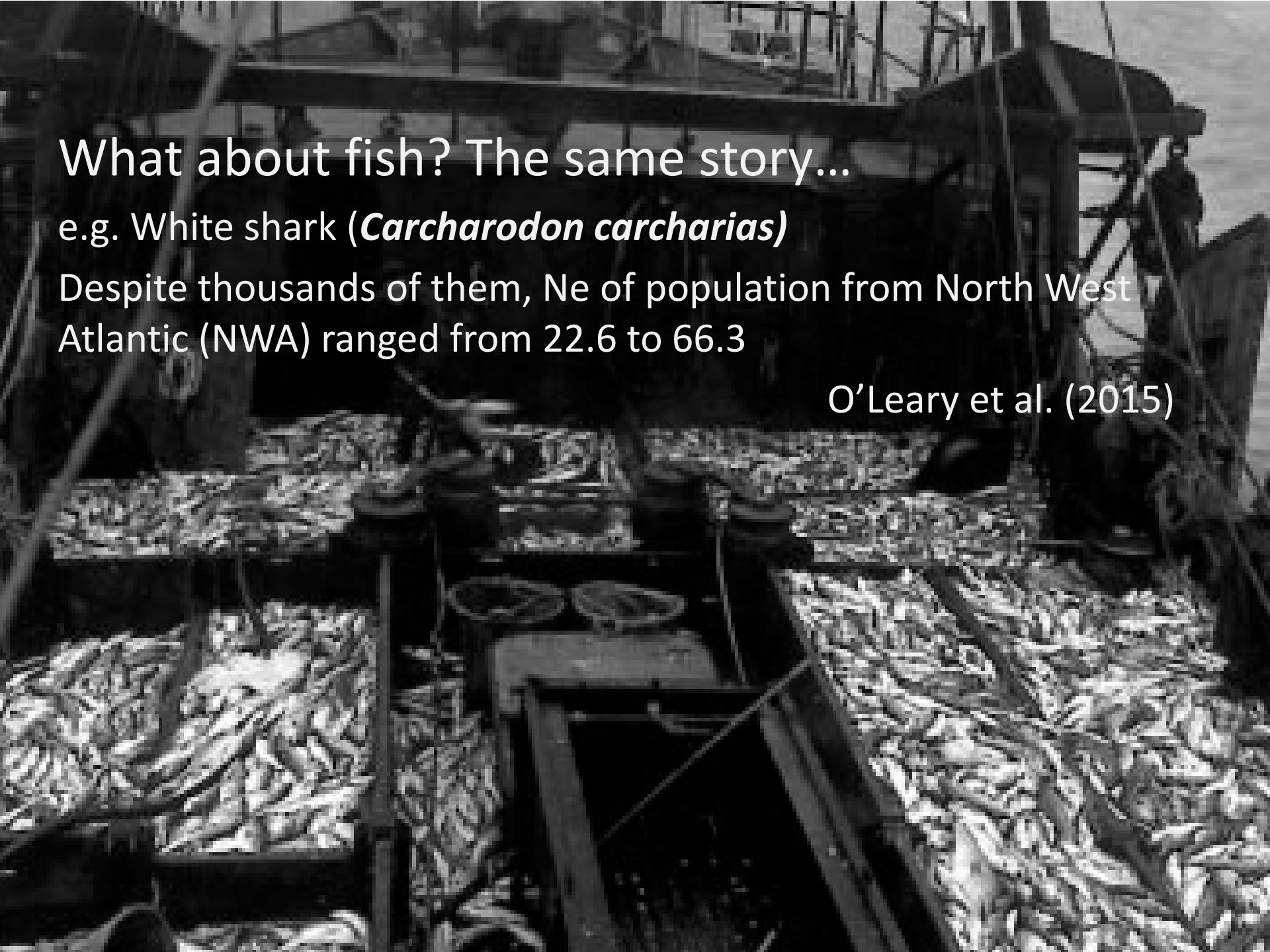
'The number of breeding individuals in an idealized population that would show the same rate of **genetic drift** as the population under consideration.'

(Wright, 1931)

Estimators to help us decide what to conserve/manage: Ne

Estimators to help us decide what to conserve/manage: Ne





What about fish? The same story...

e.g. White shark (***Carcharodon carcharias)***

Despite thousands of them, Ne of population from North West Atlantic (NWA) ranged from 22.6 to 66.3

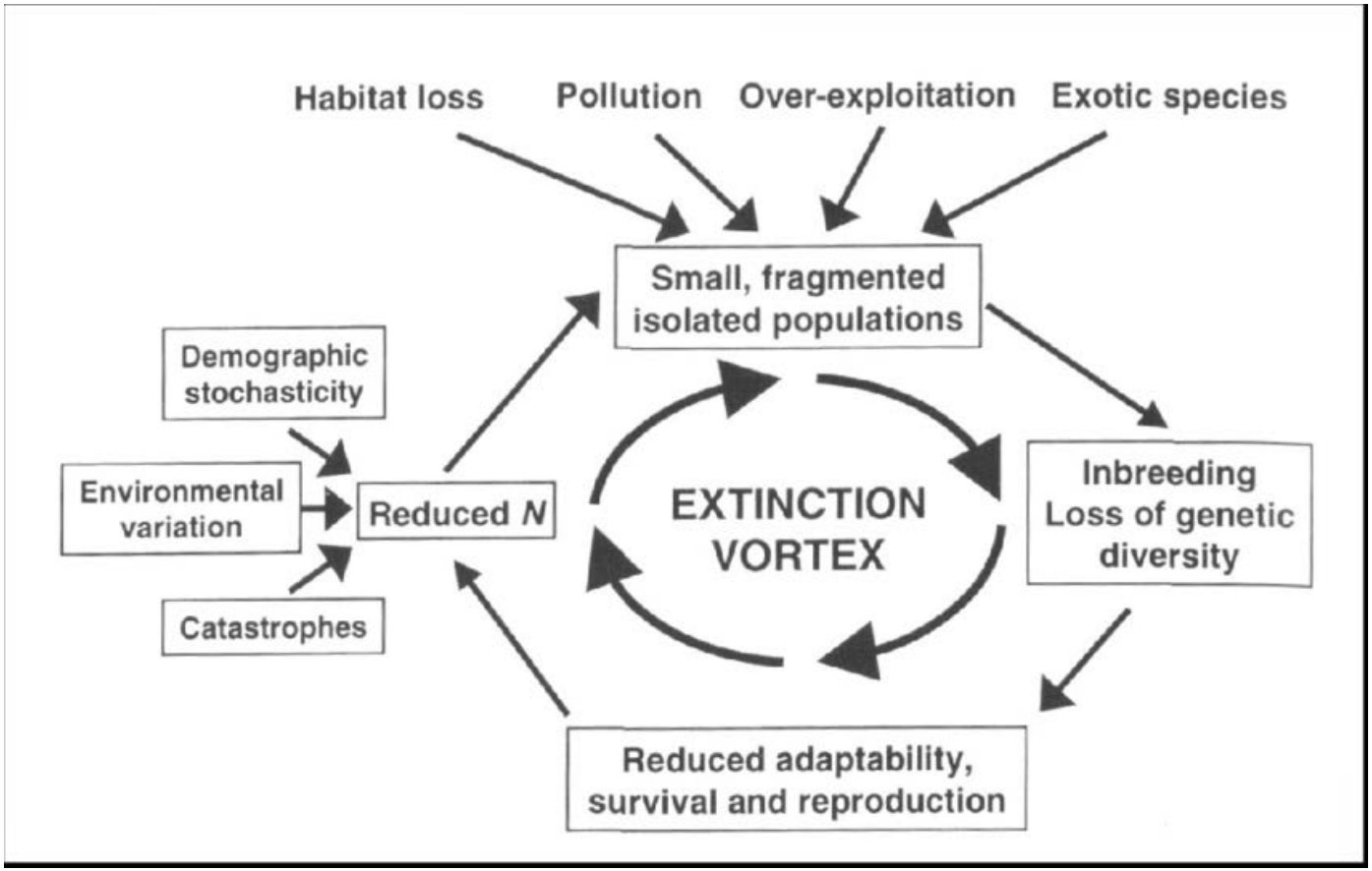
O'Leary et al. (2015)

Estimators to help us decide what to conserve/manage: Ne

- N_e is used in
  - animal breeding → aquaculture
  - molecular evolution
  - conservation
- In general, N_e can be used:
  - To assess the threat status of a population
  - To set up priority levels for conservation

An accurate estimation of N_e is thus essential!

Extinction vortex



The so called " Extinction vortex " according to FRANKHAM et al. (2002). Small, fragmented and isolated populations are fragile and sensitive to various threats which may cause either further decline or spontaneous extinction.

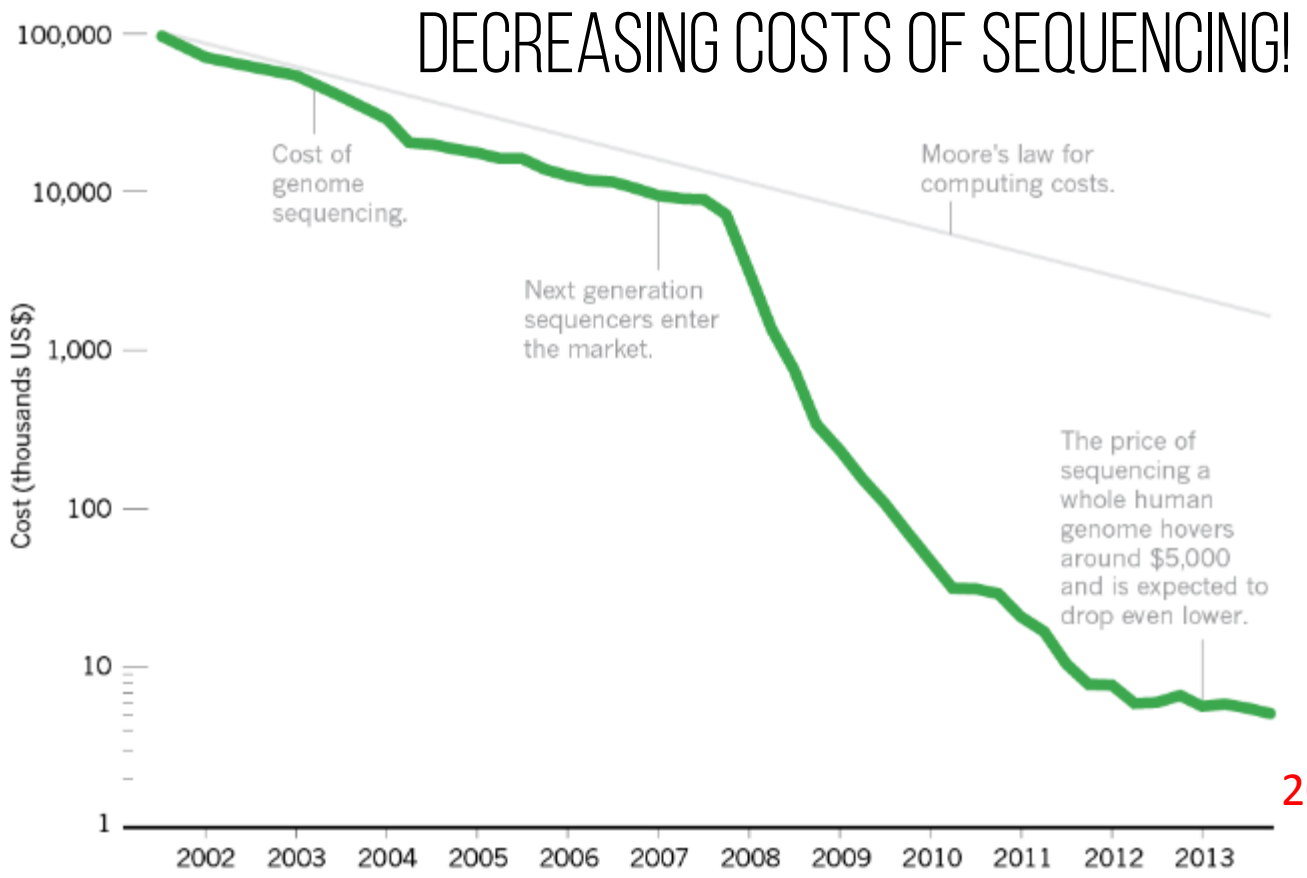
Estimators to help us decide what to conserve/manage: Ne

Different ways to estimate Ne, depending on the data we have available:

- Pedigree-based methods
- Demographic methods
- Molecular methods

  - Coalescence methods
  - Genetic Diversity methods
  - Temporal methods
  - Linkage disequilibrium methods
  - Heterozygosity methods
  - Sibship Assignment methods

Estimators to help us decide what to conserve/manage: Ne



2018: ~$100!

Macmillan Publishers Ltd: Nature News. E. Hayden. Technology: The $1,000 genome, copyright 2014

1. Estimators to help us decide what to conserve/manage:

- The effective population size (N_e)
- F_{ST} or degree of differentiation

Differences in between populations: why do we care?

Differences in between populations: why do we care?

Variation is the raw material for evolution

We can measure the degree of differenciation between two populations by the F_{ST}

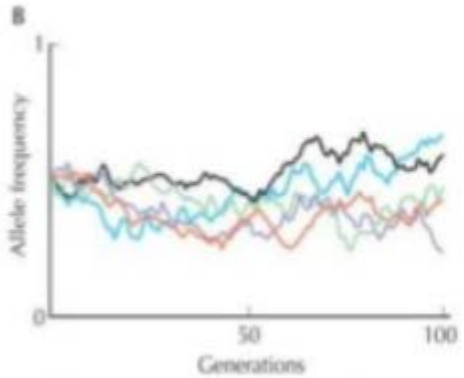
F_{ST} = the proportion of genetic diversity due to allele frequency differences among populations

F_{ST} = 0 : no differentiation
F_{ST} =1 : completely different

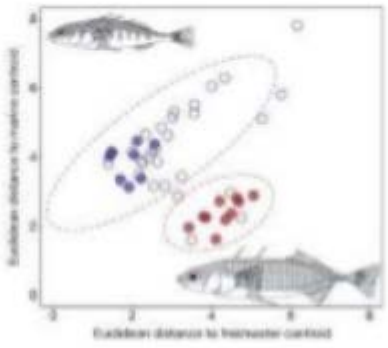
Holsinger & Weir (2009)

What causes population differentiation?

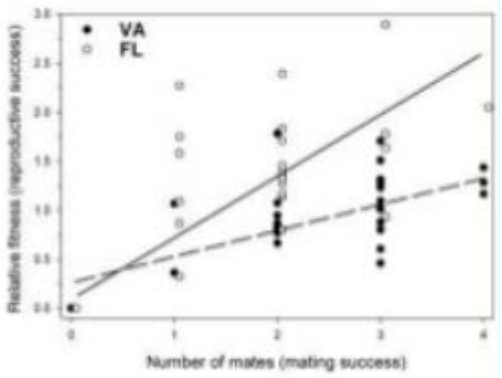
Neutral processes: genetic drift	Natural selection: local adaptation	Geographic variation in selection



Evolution © 2007 Cold Spring Harbor Laboratory Press



F Jones et al. (2012), Nature



Mobley & AG Jones(2012), Mol Ecol

F_{ST} can be measured over the genome, or by individual locus

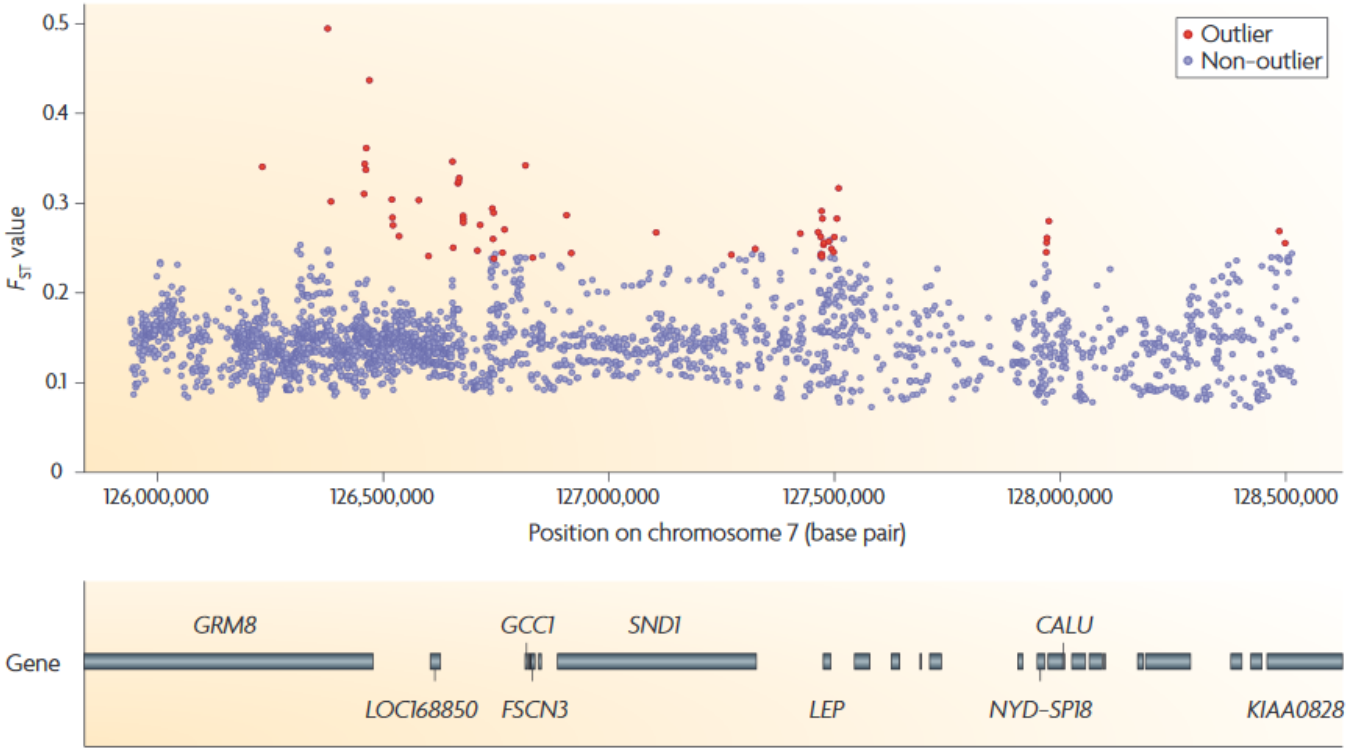


Figure 1 | **Locus-specific estimates of F_{ST} on human chromosome 7.** Estimates are as inferred from the phase II HapMap data set[95]. Horizontal bars indicate the locations of known genes. The red circles are posterior means for SNPs with estimates that are detectably different from the genomic background (purple circles). All 'outliers' show significantly more differentiation among the four populations in the sample than is consistent with the level of differentiation seen in the genomic background. The excess differentiation suggests that these SNPs are associated with genomic regions in which loci have been subject to diversifying selection among populations. *CALU*, calumenin; *FSCN3*, ascin homolog 3; *GCC1*, GRIP and coiled-coil domain containing 1; *GRM8*, glutamate receptor, metabotropic 8; *LEP*, leptin; *SND1*, staphylococcal nuclease and tudor domain containing 1. Figure is modified, with permission, from REF. 8 © (2009) American Statistical Association.

Holsinger & Weir (2009)

2. Principal Component Analysis (PCA) as a tool for exploration of population differentiation

What is PCA?

From Wikipedia:

-”A statistical procedure that converts a set of observations of possibly correlated variables (entities each of which takes on various numerical values, for example, allele frequencies) into a set of values of linearly uncorrelated variables called **principal components”**.

-”This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible)”.

In other terms, an exploratory tool to visualize potential differences/similarities between individuals/populations.

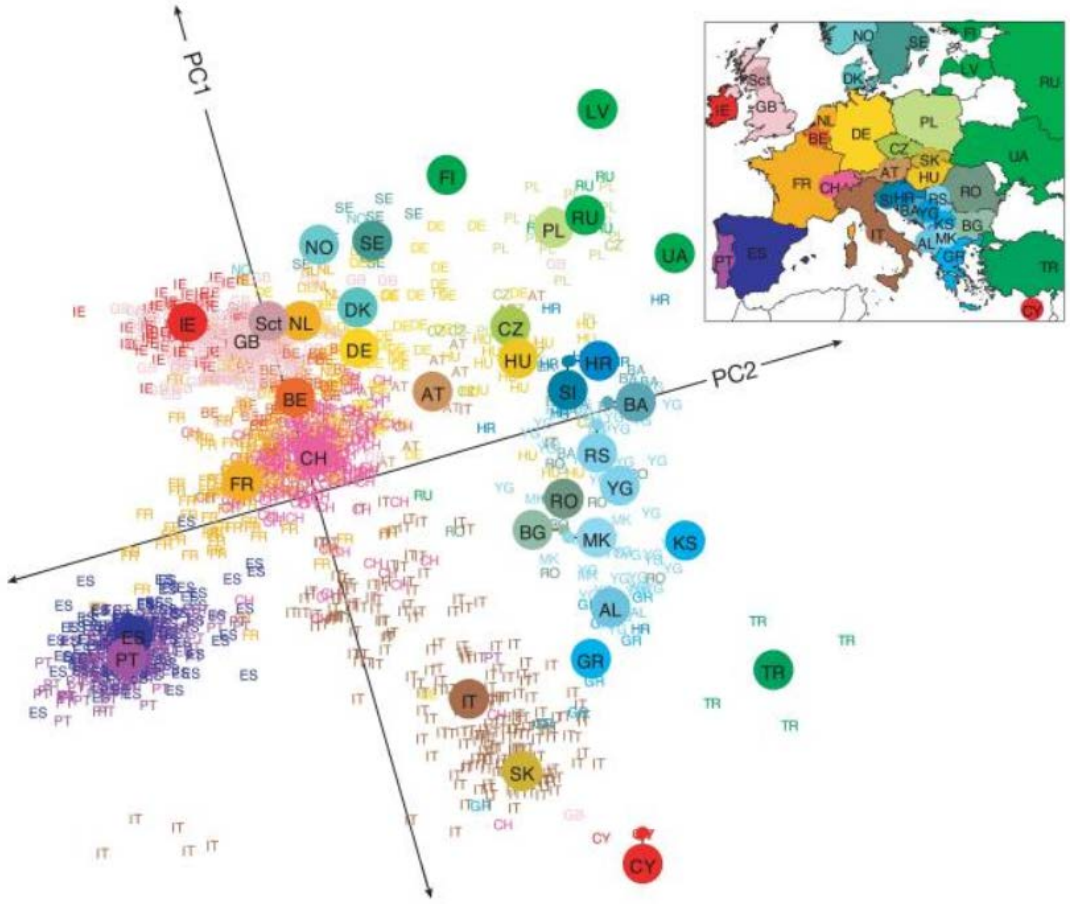


Figure 1 of Novembre et al. (2008), Genes mirror geography within Europe
PCA of modern humans in Europe

Questions?

When you are ready, open the 2nd Practical session:

https://github.com/BelenJM/Conservation-Genomics-course-DTU