

Marine evolutionary genomics

Week 1:
Reid Brennan

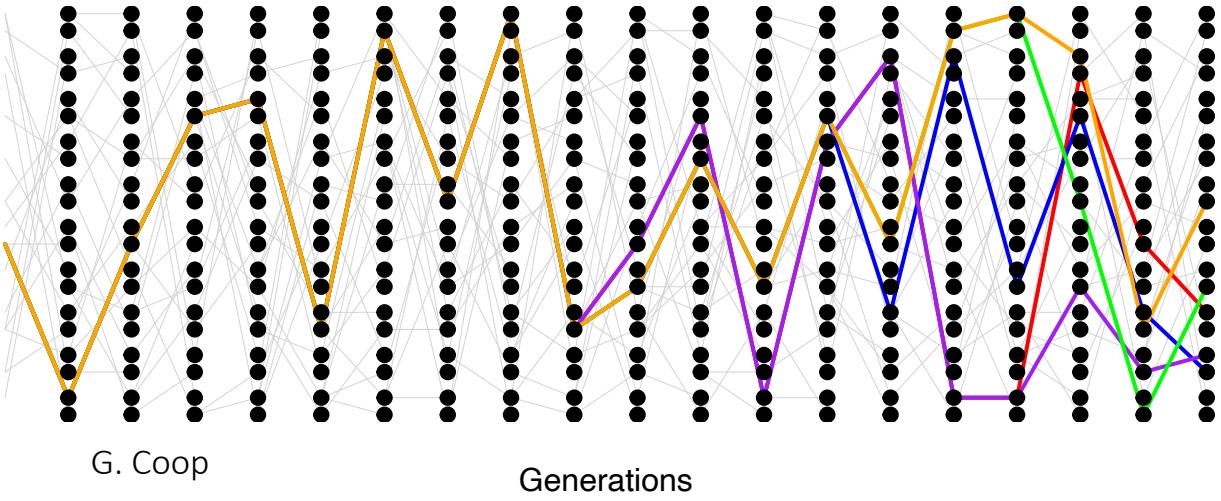
Plan for today:

- Course intro
 - Objectives
 - General information
 - Format
 - Evaluation
 - Resources
 - Content
- Intro to pop. gen:
 - Allele frequencies, genotype frequencies, and HW equil

Course objectives

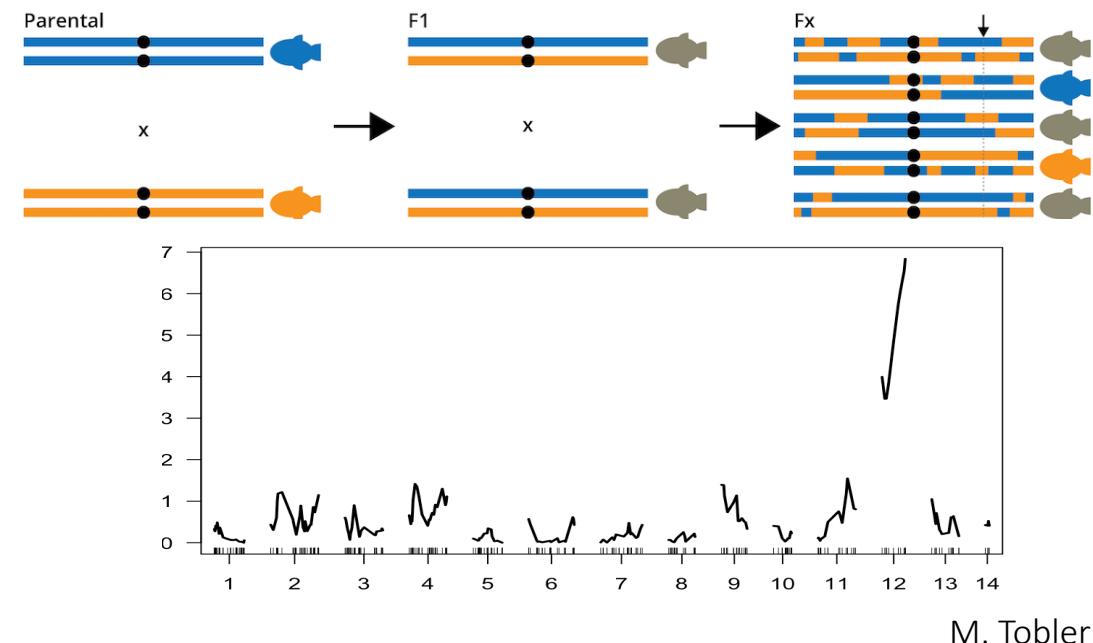
Theory

- Population genetics
- Quantitative genetics
- Evolution
- Transcriptomics and epigenetics



Practice

- What do data look like?
- How are analyses used in real systems?
- What types of questions can we address?



M. Tobler

Course design

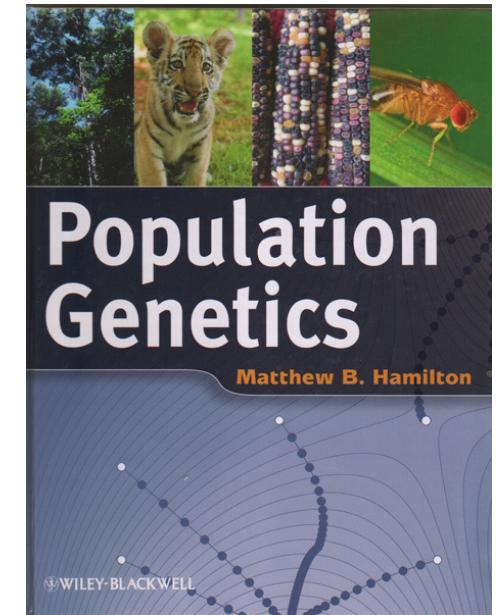
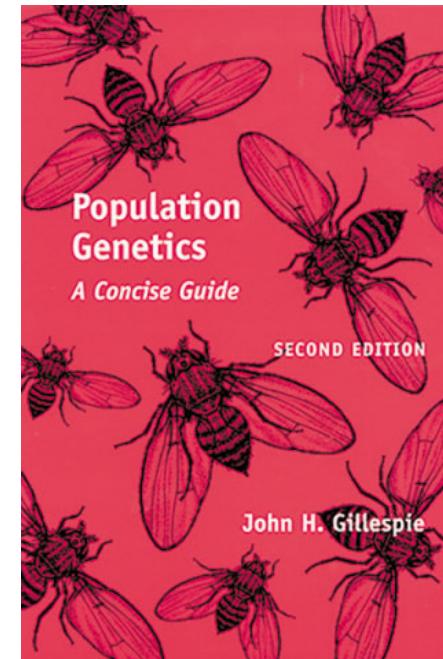
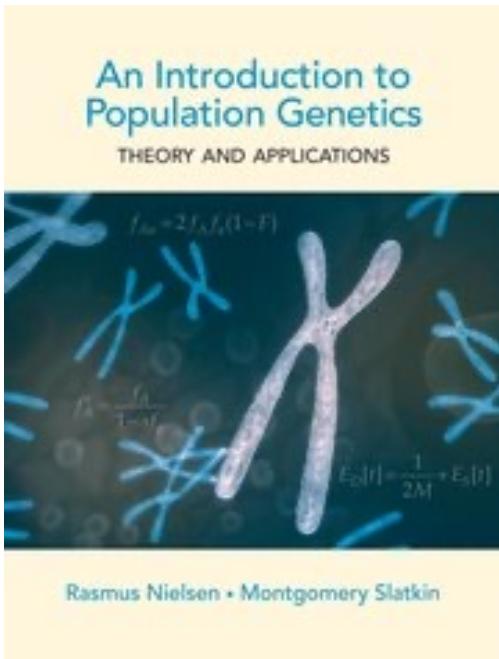
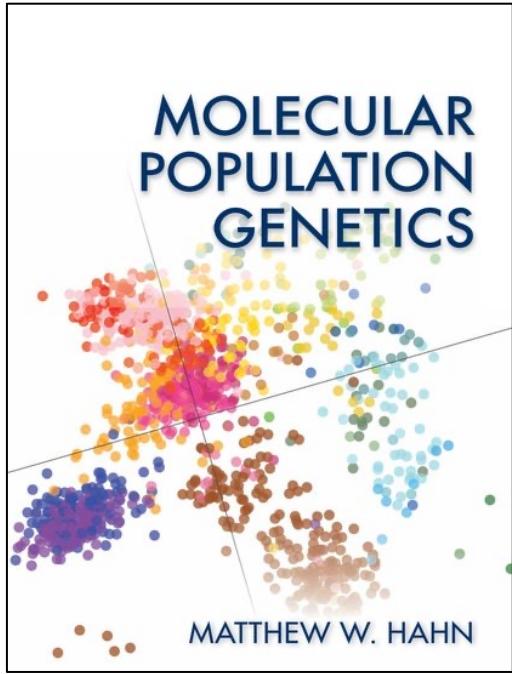
- Course website:
 - rsbrennan.github.io/EvolutionaryGenomics
- 3 parts:
 1. Lecture
 2. Discussion
 3. Exercises
- You are responsible for:
 - Reading the papers and participating in discussion
 - You will lead one week!
 - Participating in and submitting exercises
 - Writing a review and presenting
 - Any topic, approach, taxonomic group, etc. with an evolutionary genomics focus
 - Get my approval!
 - See syllabus for details of the review

Grades

- 3 parts:
 1. Review: 45%
 2. Discussion participation and lead: 25%
 3. Exercises: 30%

Resources

- Graham Coop's Population and Quantitative Genetics notes
 - <https://github.com/cooplаб/popgen-notes/releases>



And many others....

Course content

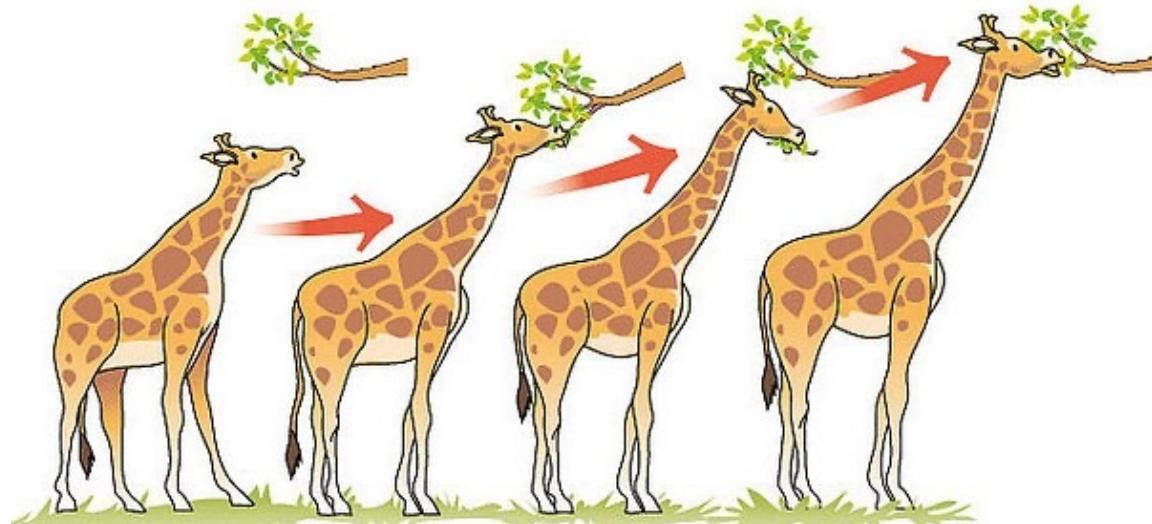
- Population genetics basics
 - Allele and genotype frequencies, HW equilibrium
 - Drift, mutation, the coalescent
 - Population structure
 - Linkage
 - Selection
- Quantitative genetics
- Transcriptomics
- Epigenetics

Each topic: theory + applied

What is evolution?

Lamarck

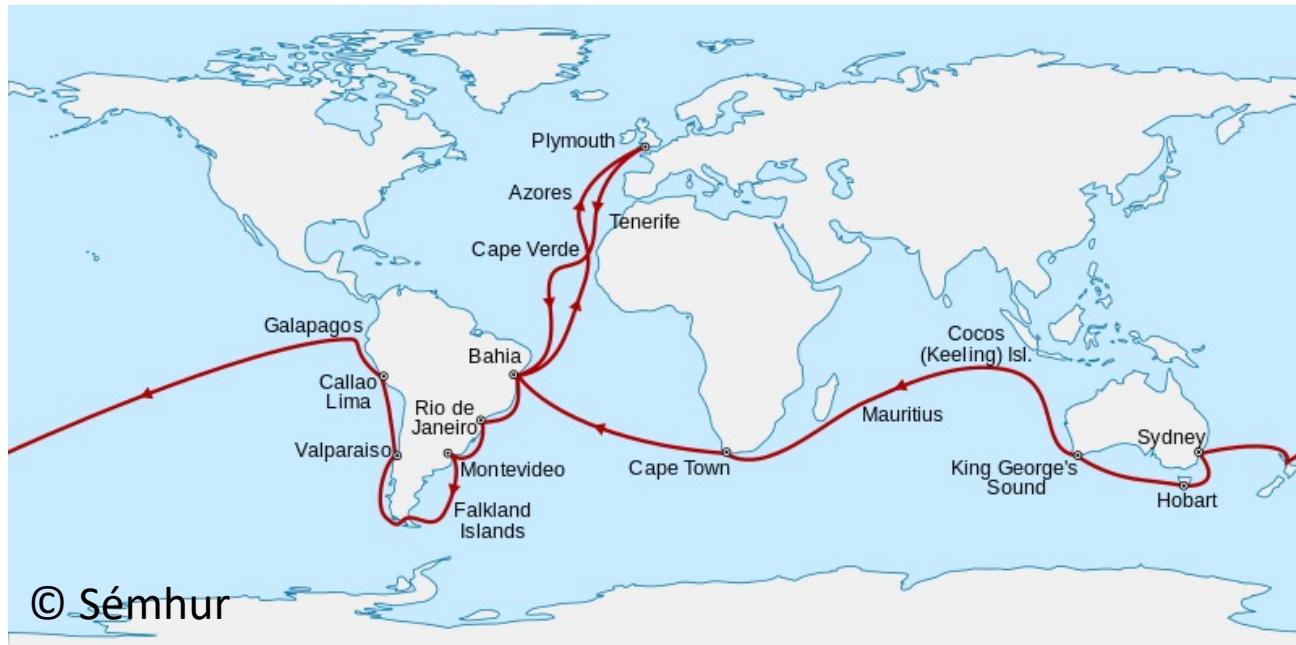
- Knew the earth was old
1. Life evolved towards complexity
 2. Inheritance of acquired characteristics



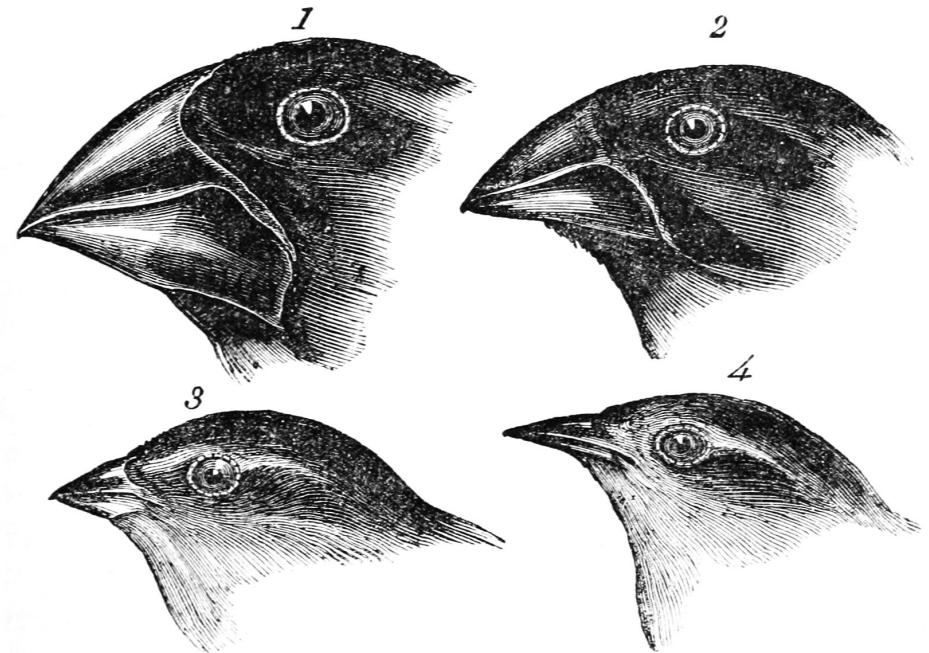
1744-1829

From Sandritaverooka - Eigenes Werk, CC BY-SA 4.0,
<https://commons.wikimedia.org/w/index.php?curid=40890964>

Darwin: 1809-1882



5 year voyage

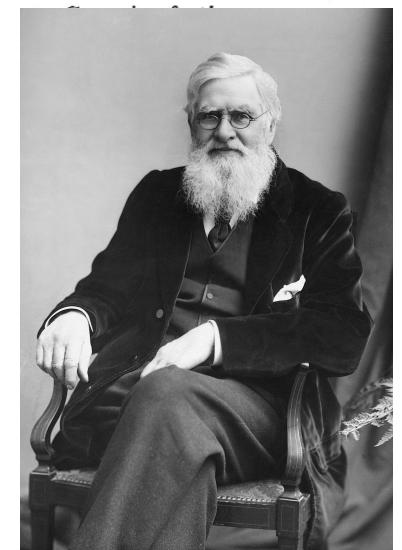
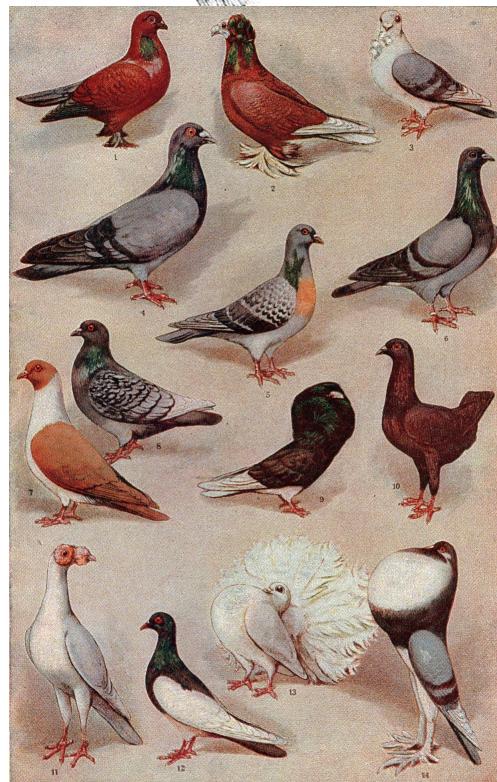
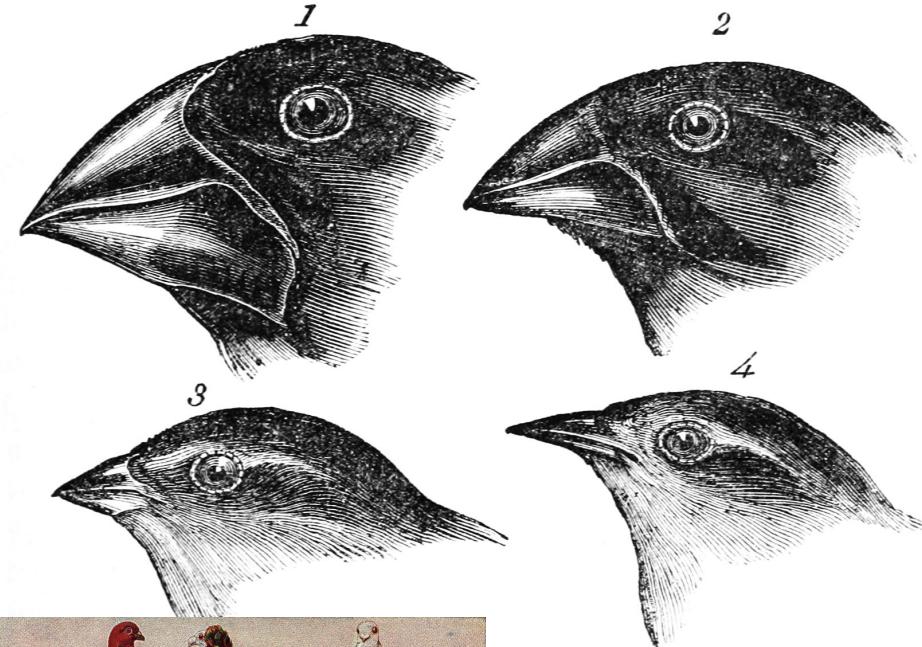


1. *Geospiza magnirostris.*
3. *Geospiza parvula.*

2. *Geospiza fortis.*
4. *Certhidea olivacea.*

Darwin's main insights

- Descent with modification
- Natural selection
 - Different reproductive success between individuals with different traits
- Not all evolution is natural selection
- What Darwin didn't know...
 - Genetics



Alfred Russel Wallace

What is evolution?

- the change over time in the genetic composition of a population.
1951, Theodosius Dobzhansky

Populations → individuals → genomes

birth – death – migration

Every individual has a unique genome and
every individual contributes to evolution



Population genetics is the study of the genetic composition of natural populations and its evolutionary causes and consequences.

Quantitative genetics is the study of the genetic basis of phenotypic variation and how phenotypic changes evolve over time.

How do the genes and phenotypes of a population change over time?

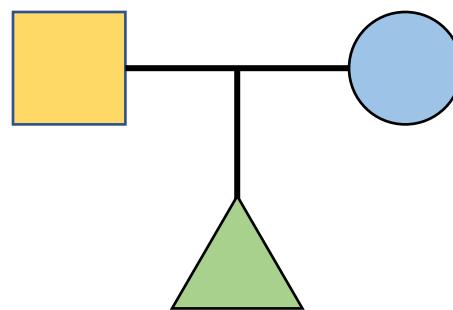
GREGOR MENDEL (1822-1884)



Experiments on Common pea (*Pisum sativum*): 1856 to 1863

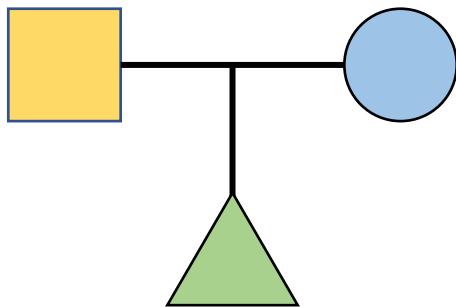


Common thought at the time was blending inheritance

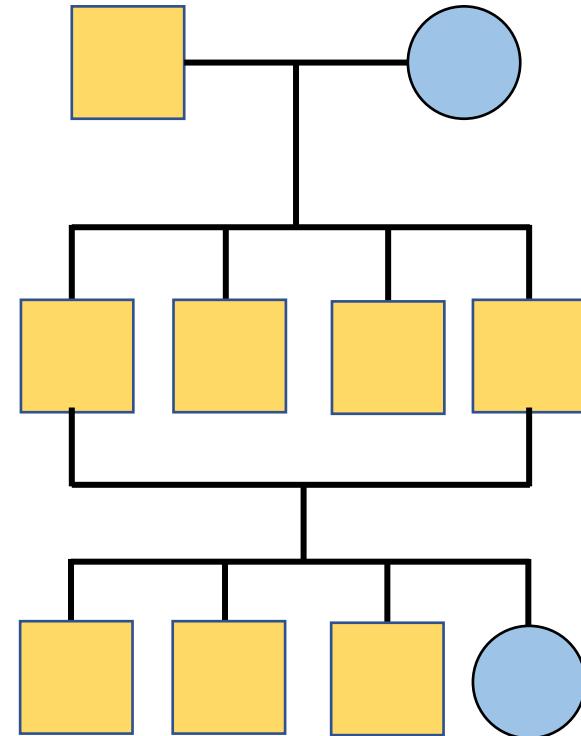


Pea coat colour (actually yellow/green...)

Crosses between 'pure' breeding lines



Parents



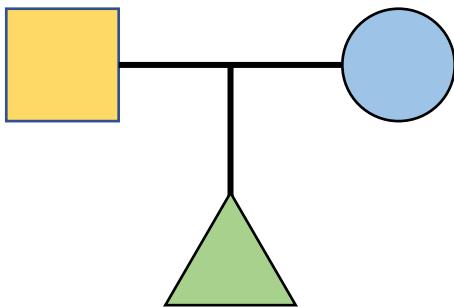
F1

F2

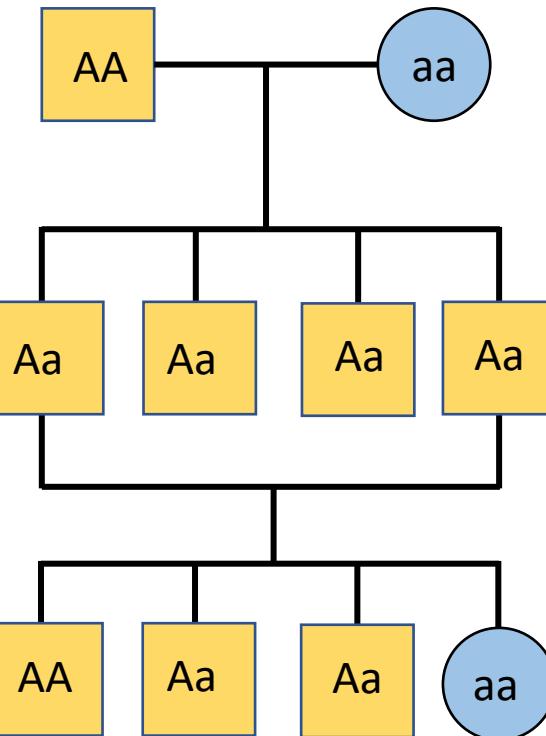
~30,000 plants!

Pea coat colour (actually yellow/green...)

Crosses between 'pure' breeding lines



Parents



F1

F2

Concept of alleles and dominance!

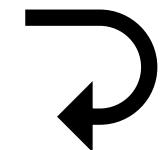
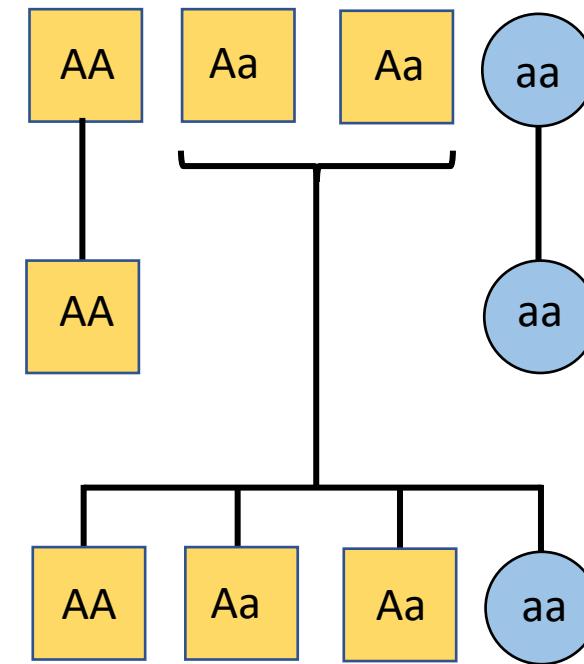
	A	a
A	AA	Aa
a	Aa	aa

Punnet squares (1905)

3:1 phenotypes

F2

F3



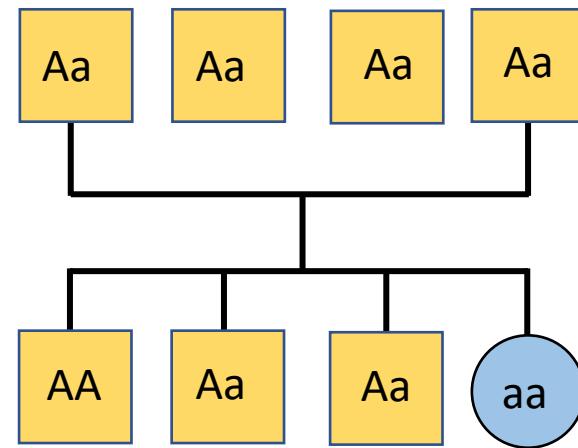
~30,000 plants!

Mendel's first 'law': independent segregation of alleles (at a single locus)

- *Two members of a gene pair (alleles) segregate separately into gametes so that half of the gametes carry one allele and the other half carry the other allele.*



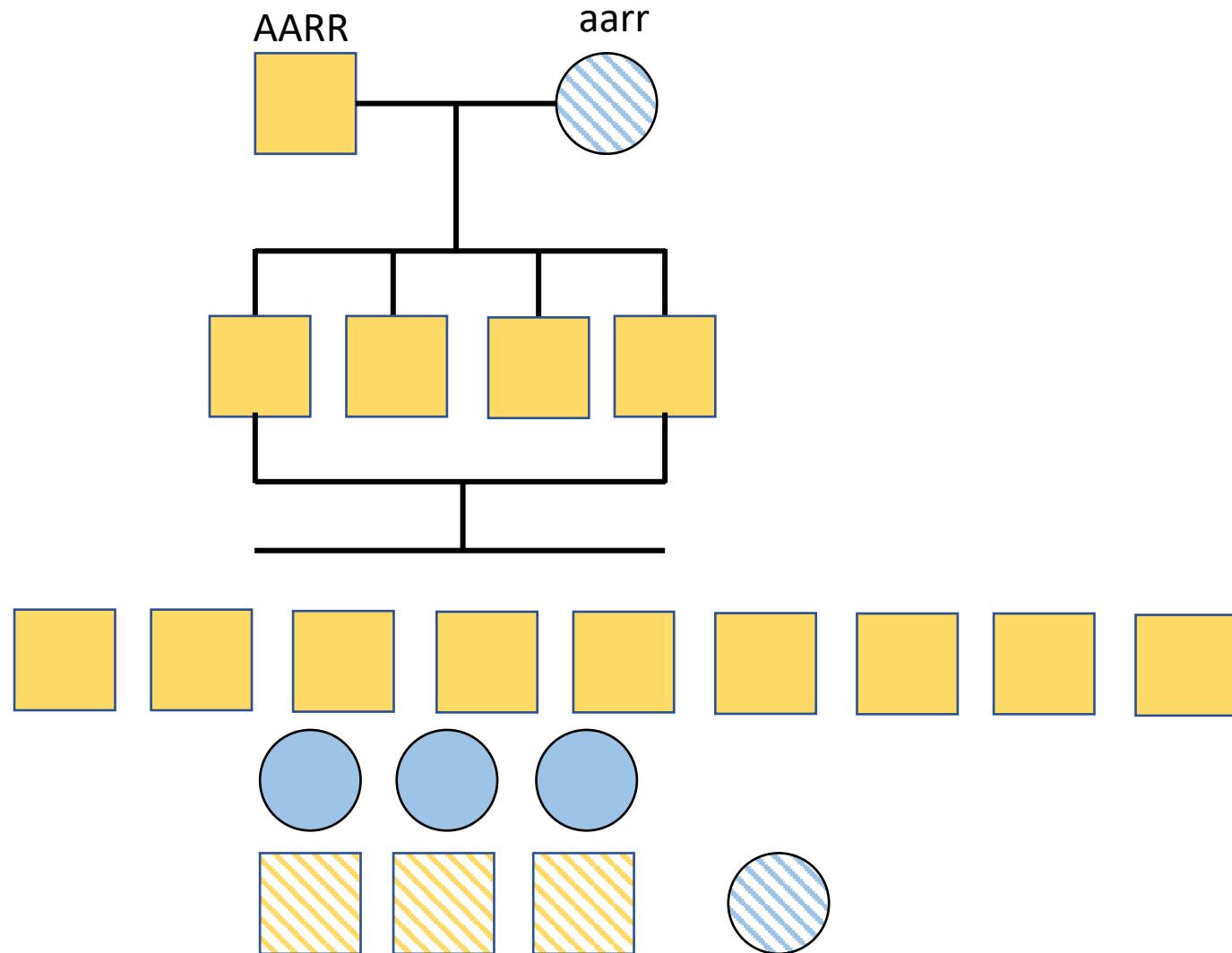
A locus is heterozygote if it carries different alleles (A and a), and homozygote if it carries the same two alleles (a and a)



	A	a
A	AA	Aa
a	Aa	aa

Mendel's second 'law': independent assortment of multiple loci

- The segregation of alleles in one gene is independent of the segregation of alleles in another gene.*



AR	aR	Ar	ar
AR			

9:3:3:1

GREGOR MENDEL



- phenotypes are determined by discrete units (genes) that are inherited intact and unchanged through generations: *particulate inheritance*
- Published in 1866 in the *Proceedings of the Natural Sciences of Brünn* (in German)
- Darwin's contemporary, but unnoticed for 35 years!

Establishes the first and most fundamental prediction of population genetics: ***expected genotypic frequencies***



Allele frequencies, genotype frequencies, and HW equilibrium

Definitions....

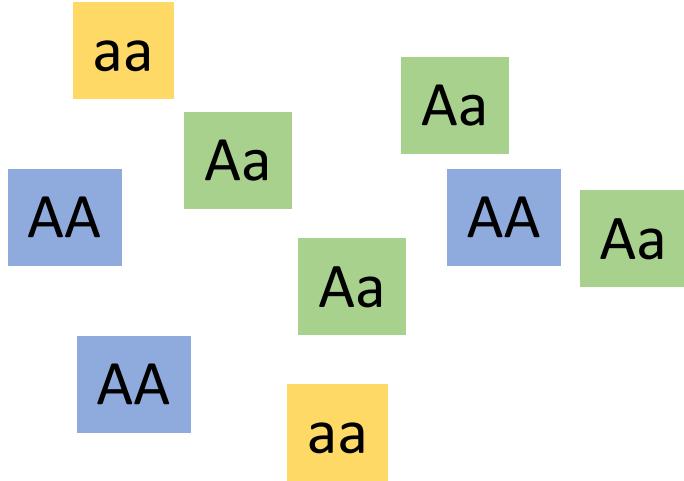
Locus: any location in the genome: gene, single basepair, microsatellite....

Allele: the variants at a locus (polymorphic when more than one)

Genotype: set of alleles possessed by an individual (at one locus or several loci)

Phenotype: an organism's observable attribute (morphological, developmental, biochemical, physiological, behavioral, ...)

Genotype frequency



$$f_{AA} = \frac{N_{AA}}{N}$$

$$f_{Aa} = \frac{N_{Aa}}{N}$$

$$f_{aa} = \frac{N_{aa}}{N}$$

$$N_{AA} = 3$$

$$N_{Aa} = 4$$

$$N_{aa} = 2$$

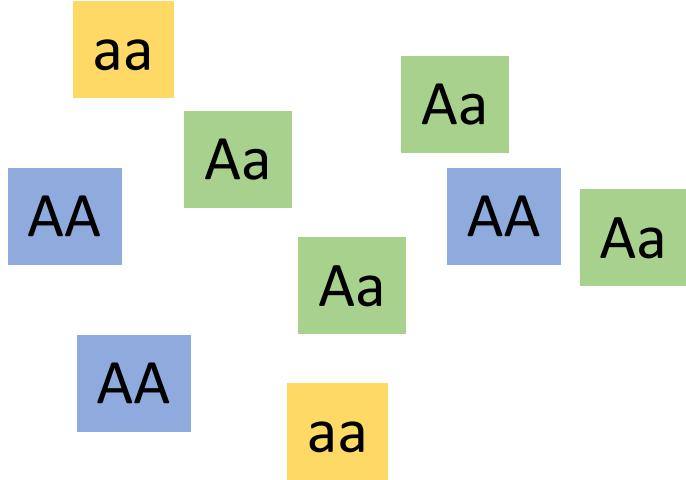
$$N = 9$$

$$f_{AA} = 3/9 = 0.33$$

$$f_{Aa} = 4/9 = 0.44$$

$$f_{aa} = 2/9 = 0.22$$

Allele frequency



$$f_A = p = \frac{2N_{AA} + N_{Aa}}{2N} = f_{AA} + f_{Aa}/2$$

$$f_a = q = 1 - p$$

$$N_{AA} = 3$$

$$N_{Aa} = 4$$

$$N_{aa} = 2$$

$$N = 9$$

$$f_A = 0.33 + 0.44/2 = 0.55$$

$$f_a = 0.22 + 0.44/2 = 0.44$$

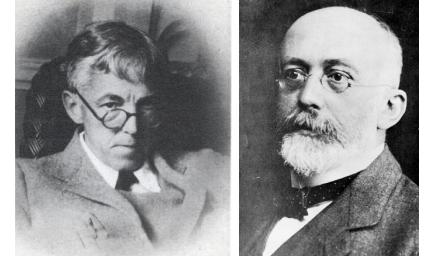
$$f_{AA} = 3/9 = 0.33$$

$$f_{Aa} = 4/9 = 0.44$$

$$f_{aa} = 2/9 = 0.22$$

Hardy-Weinberg Equilibrium

- Predicting genotype frequencies from allele frequencies
- Need to assume random mating



HARDY-WEINBERG (1908)

- Remember Mendel's first law...
- To get AA genotype →
- To get aa genotype →
- To get Aa genotype →

Godfrey H. Hardy (1877-1947)
Wilhelm Weinberg (1862-1937)

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

AA Aa/aA aa

A null expectation → Deviations = evolution

A single generation of reproduction will result in a population that meets the expected Hardy-Weinberg frequencies, i.e. is at Hardy-Weinberg (HW) equilibrium

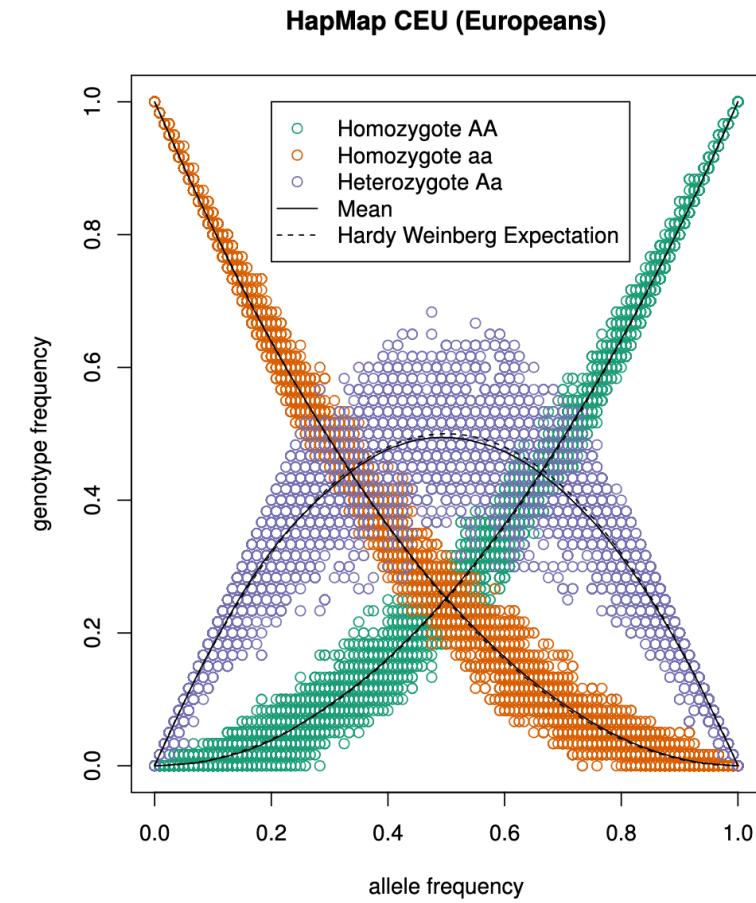
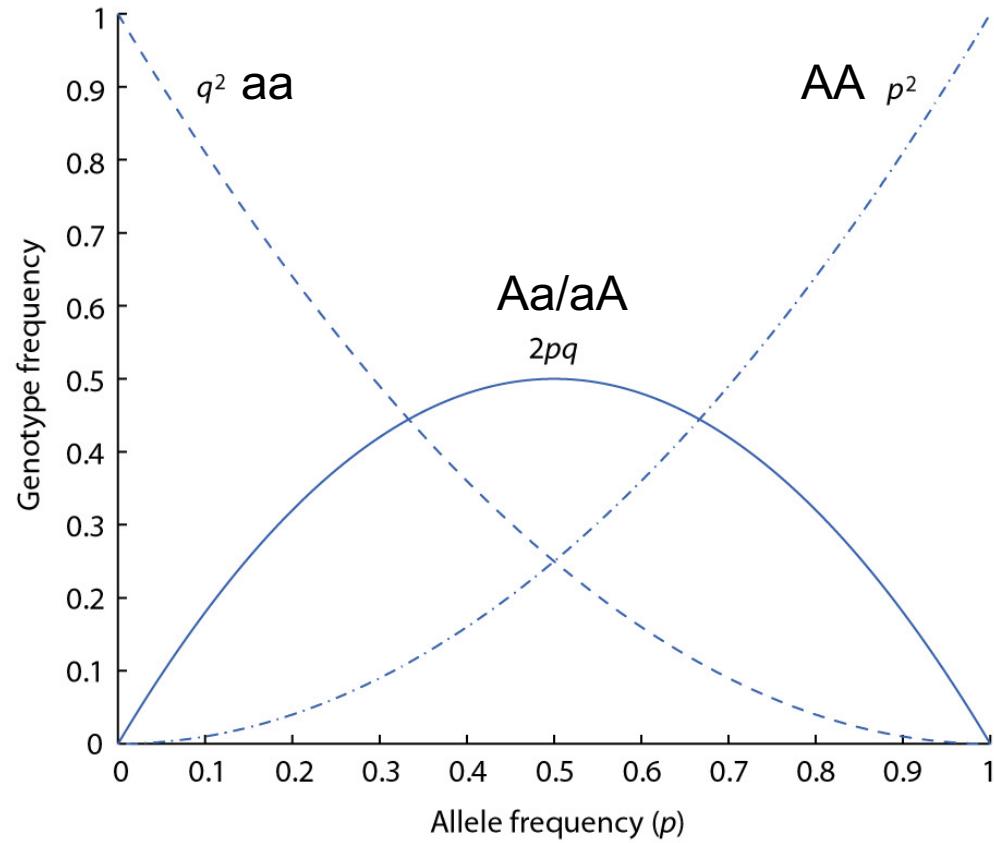
Assuming an “ideal” population, i.e. :

- Diploid organisms
- Sexual reproduction (as opposed to clonal)
- Random mating (as opposed to e.g. assortative) with respect to genotype
- Random union of gametes
- Discrete, non-overlapping generations
- Very large (infinite) population
- No migration
- No population structure
- No natural selection
- Two alleles
- Identical allele frequencies in both sexes

-> Departures from HW equilibrium may indicate:

- Inbreeding
- Assortative mating
- Migration
- Natural selection
- Population structure
- ...

What is the expectation for genotype frequencies?
When do we have the highest heterozygosity?



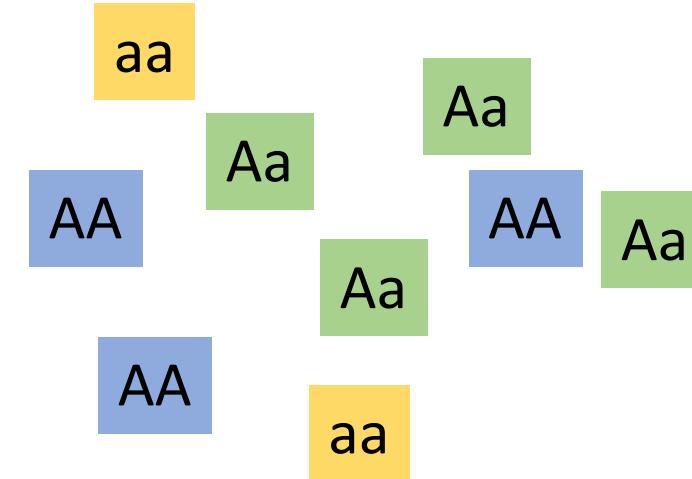
G. coop

Inbreeding coefficient: F

Fixation index (F): proportion by which heterozygosity is reduced or increased relative to the heterozygosity of a population at HW equilibrium with the same allele frequencies.

$$F = \frac{H_e - H_o}{H_e} = \frac{2fAf_a - fAa}{2fAf_a}$$

$$f_{Aa} = 2fAf_a(1 - F)$$



What is F for our population?

$$\frac{2(0.55)(0.44) - 0.44}{2(0.55)(0.44)}$$

$$f_A = 0.55$$

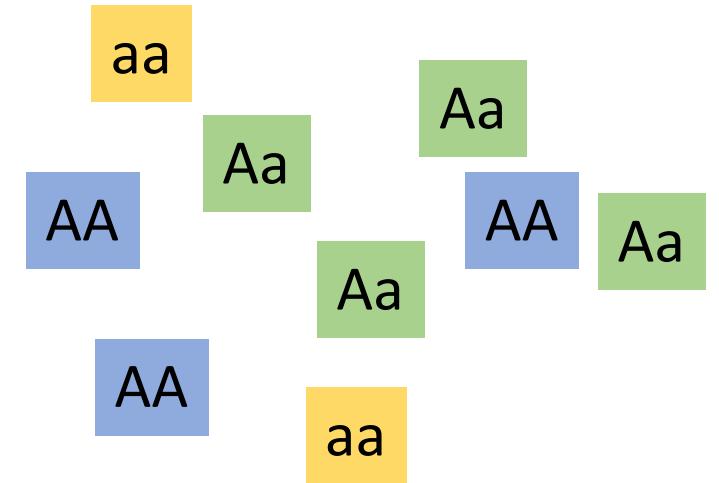
$$f_a = 0.44$$

Deviations from HWE

- Chi-Square test

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Find expected genotype counts, compare to observed genotype counts



Paper for next week

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Genomic signatures of clonality in the deep water kelp *Laminaria rodriguezii*

Lauric Reynes¹  | Thierry Thibaut¹  | Stéphane Mauger²  | Aurélie Blanfuné¹  |
Florian Holon³  | Corinne Cruaud⁴  | Arnaud Couloux⁵  | Myriam Valero²  |
Didier Aurelle^{1,6} 

H_e : expected heterozygosity

H_o : observed heterozygosity

F_{is} : fixation index

Hardy-Weinberg