

Genetics in support of fisheries and aquaculture management

17-19 September
Faro, Portugal

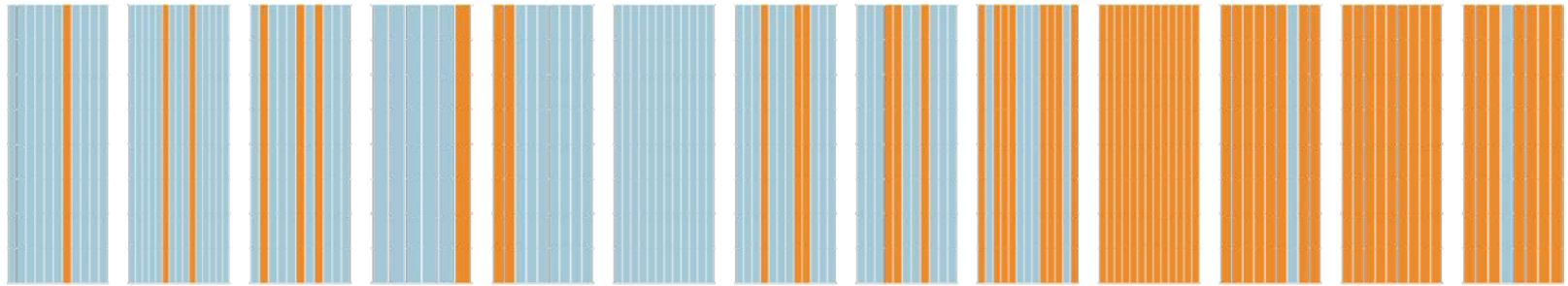


ICES
CIEM



UAAlg

CCMAR
Centro de Ciências do Mar



Quantitative genetics and application to aquaculture

17-19 September
Faro, Portugal



Animal breeding and productivity

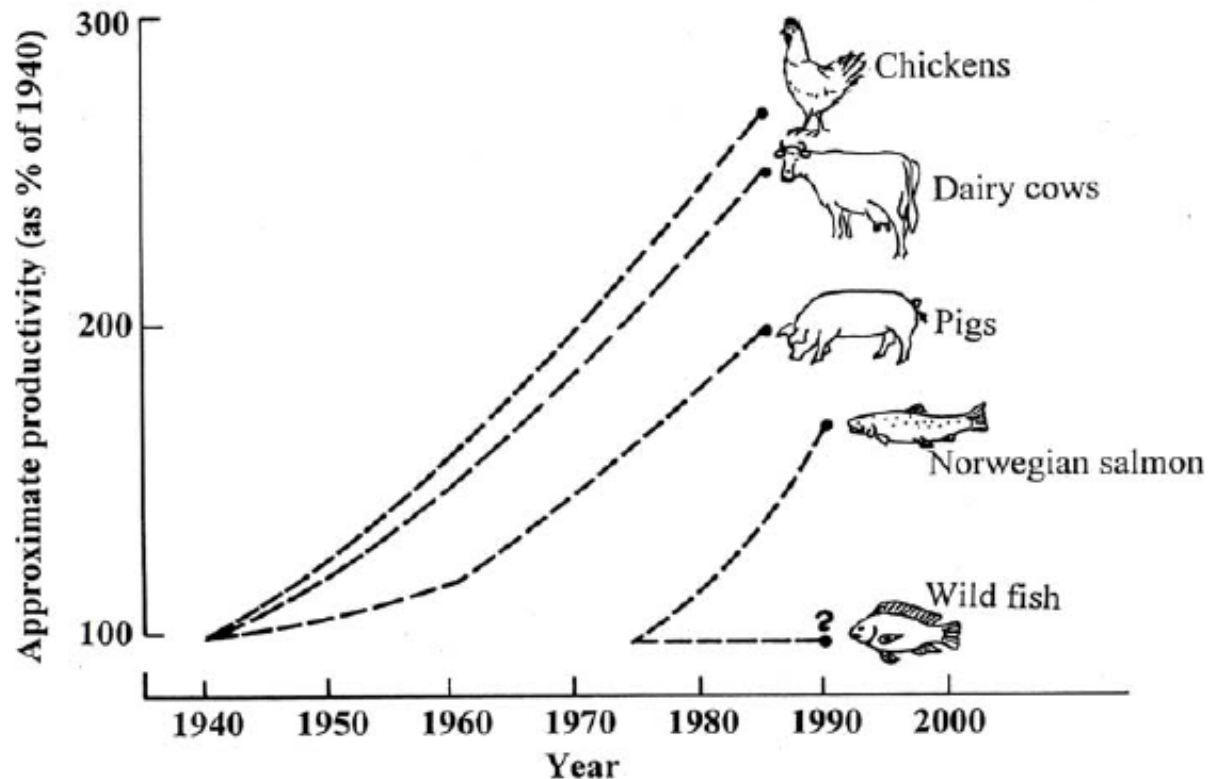
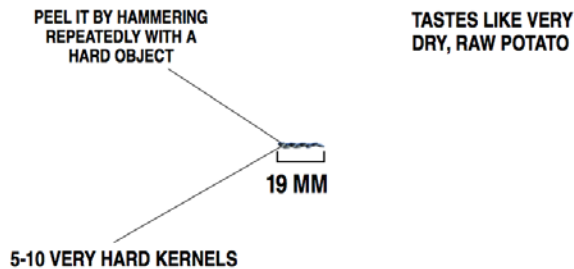


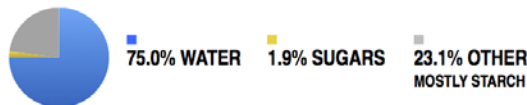
Figure 1.3. Genetic research initiated in the 1940's has resulted in remarkable developments in increasing the productivity of domestic mammals and birds. Reproduced from Eknath et al. (1991) by permission of World Fish Center.

Plant breeding and productivity

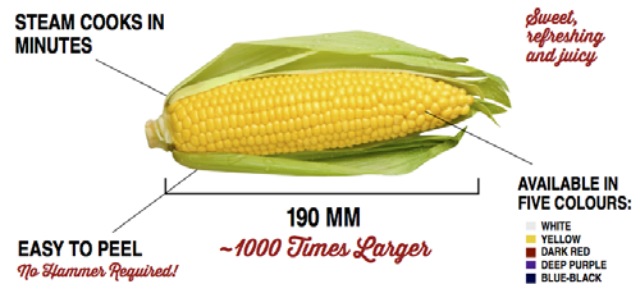
NATURAL "CORN", 7000 B.C.



8 KNOWN VARIETIES



ARTIFICIAL CORN, 2014



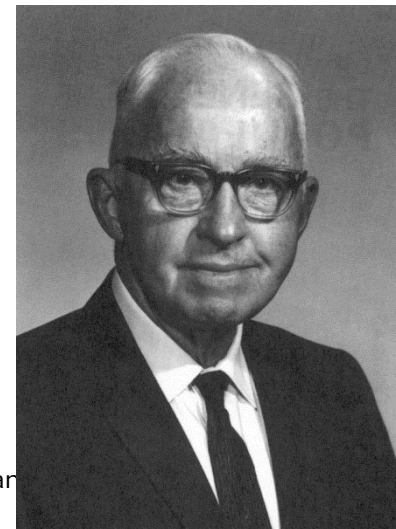
~200 VARIETIES
25-Fold Increase



JAMESKENNEDYMONASH.WORDPRESS.COM

History of breeding

- Development of breeds from 1700
- The development of the theory of animal breeding was pioneered by Sewall Wright and Jay L Lush early in last century
- In his book 'Animal breeding plans', first published in 1937, Jay L. Lush discussed the principles and elements of breeding plans for animals



Breeding

- Selective breeding exploits the underlying genetic variation in a species to change traits in the direction desired by the breeder
- Selection can change traits very fast
- Genetic changes are accumulative

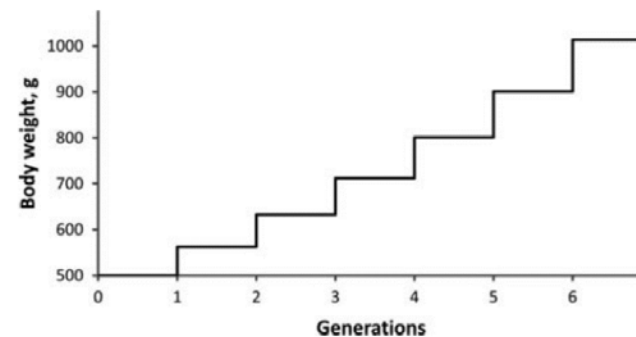


Fig. 2.1 Result of selection for growth rate over six generations when genetic improvement is 12.3% each generation (a figure that has been obtained in several breeding programs). The generation interval is illustrated by the run of the stair (distance between 1st and 2nd) and the rise of the stair is the improvement in body weight as a result of selection

Domestication

- Domestication: “The process by which a population of animals become adapted to man and to the captive environment by some combination of genetic changes occurring over generations and environmentally induced developmental events recurring during each generation” (Price 1984)
- Domestication occur primarily through selection, i.e. animals that are best adapted to a particular environment produce more progeny that survive compared to those that are less adapted
- Domestication is a slow process that takes place over a relatively long time period

Traits with a genetic background –single genes

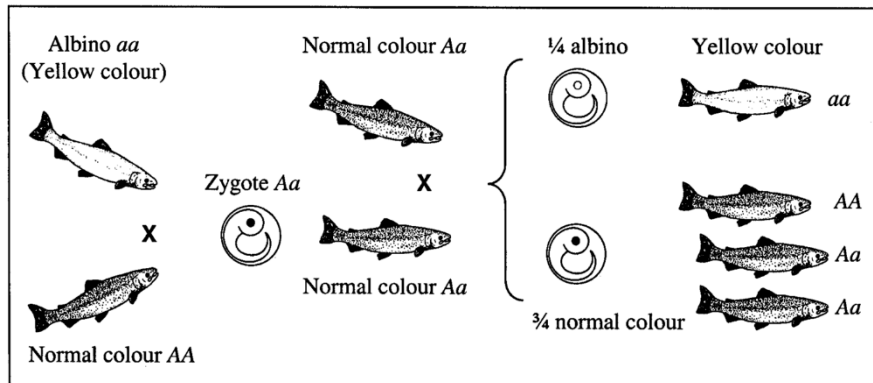
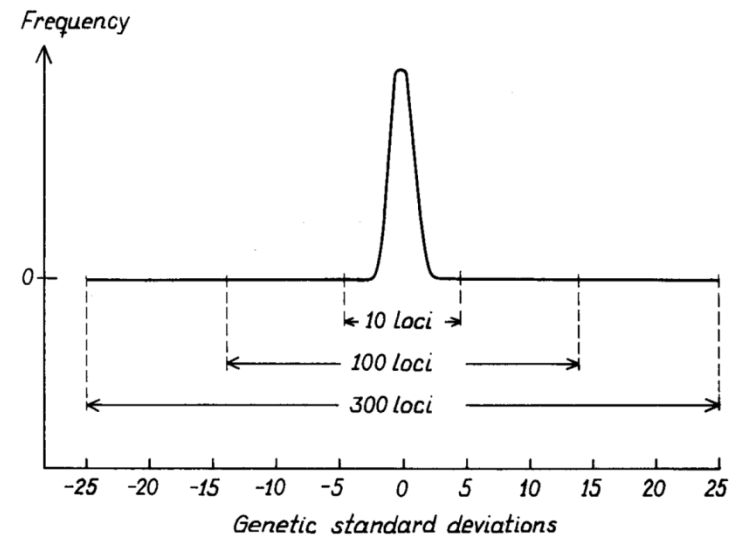
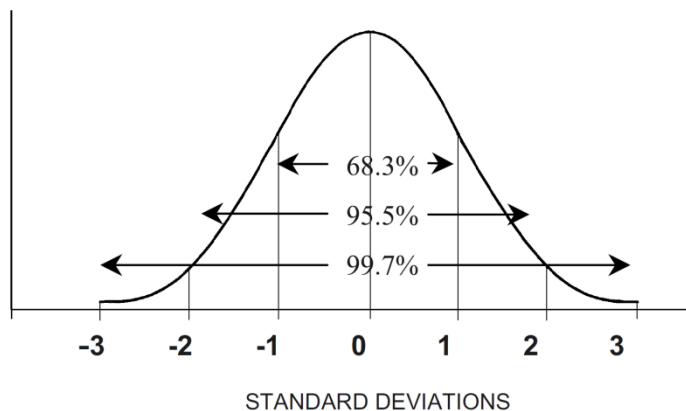
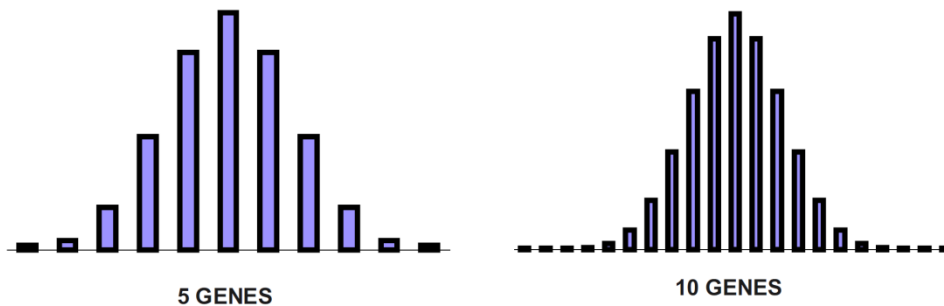


Figure 2.9 Inheritance of albinism in rainbow trout.



Quantitative traits

- A **quantitative trait** is a measurable phenotype that depends on the cumulative actions of many genes and the environment
- As the number of genes (each with two alleles) affecting a trait increases the number of genotypes increases as well as the variation



Quantitative traits are also influenced by the environment

- Genetically identical individuals can have different phenotypes
- Quantitative traits are determined by the combined influence of the genotype at several loci and the environment
- **genotype + environment + random variation = phenotype**

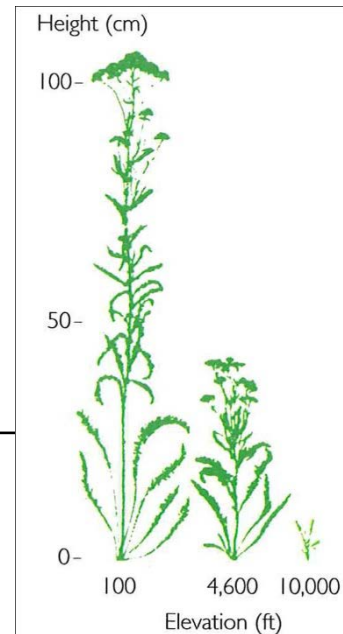
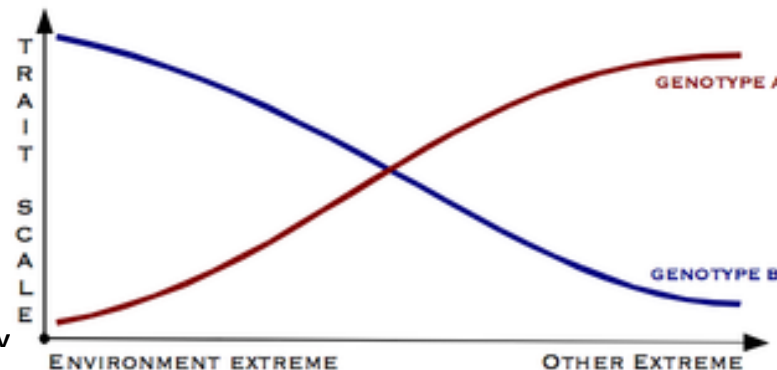


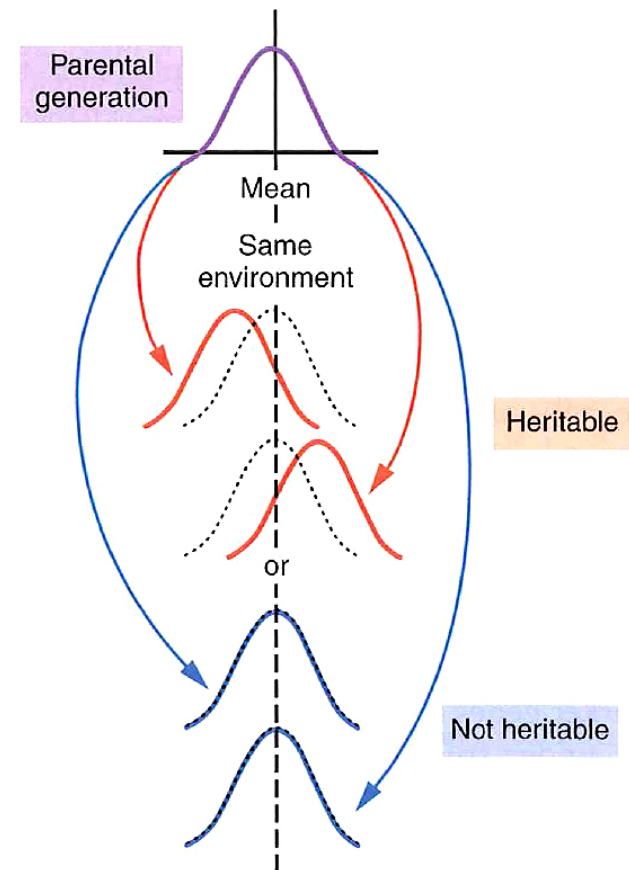
Figure 9.4 Quantitative traits are influenced by the environment as well as genotype

Reaction norm



Measuring heritability

- The fraction of the *total phenotypic variance* that is due to variation in genes is called the *heritability* of the trait
- We need to partition the total phenotypic variation (V_P) into a component due to *genetic variation* (V_G) and to a component due to *environmental variation* (V_E)
- If the phenotypic distribution can be moved by selection
 - then the trait has a *genetic component*
 - then the trait is *heritable*



Partitioning of phenotypic variation

- Phenotypic value of an individual:

$$P = G + E$$

- Phenotypic variance within a population:

$$V_P = V_G + V_E + 2Cov_{GE}$$

V_P = phenotypic variance

V_G = environmental variance

Cov_{GE} = interaction between genetic & environmental effects

Partitioning of genetic variation

- There are two broad categories of gene effects:
- **Additive gene effects**
 - occur when the combined effects of alleles at different loci are equal to the sum of their individual effects
- **Non-additive gene effects**
 - *dominance* effects, the effect of a given allele depends on the interaction with the other allele present at the same locus.
 - *epistasis*, interaction between alleles at different loci
- Only **additive genetic** effects are fully transferred to the next generation in a strict and predictable way
- selection acts upon additive effects
 - Thus we want to estimate heritability of a trait based on additive effects only
 - This will allow us to predict how a population or strain will respond to selection

Heritability

Breeders need to know how much of the phenotypic variability of a trait is due to genetic variance, and how much is due to non-genetic environmental factors

This is the 'broad-sense heritability': $H^2 = V_G/V_P$

More useful to know is the proportion of the phenotypic variation is due to **additive gene** effects

The heritability (narrow-sense) of a trait is defined as the proportion of the total phenotypic variation that is due to heritable (additive genetic) effects that can be passed on from parent to offspring :

$$V_A/V_P = h^2$$

Heritability estimates

- Offspring can resemble their parents for reasons other than their genetic relationships
- **Heritability** estimates are only reliable if there is no correlation between the environment of offspring and parents
- Heritability estimates only apply for specific environments and may change if moving the population to a new environment
- Economically important traits in aquaculture
 - $h^2 = [0.1-0.4]$
 - Estimates are trait specific but also population specific

The breeders equation

h^2 allows prediction of response to artificial (or natural) selection

The relationship between h^2 and the response to selection, is given by:

$$R = h^2 * S$$

- R is the response to selection, given by the difference between the population mean before selection and the mean of the offspring of selected parents after one generation of selection.
- S is the selection coefficient, given by the difference between the unselected population mean, and the mean of the selected parents.
- If we know the heritability of a trait and the strength of artificial selection applied to it, we can predict the response to selection >>> so...

Selection response

Response to selection

$$R = h^2 S$$

Selection differential

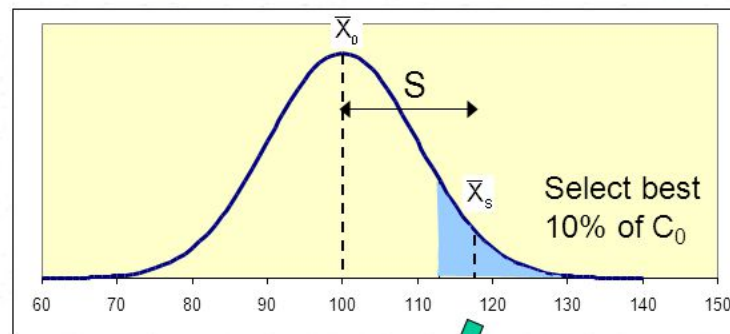
$$S = \bar{X}_S - \bar{X}_0$$

Response to selection

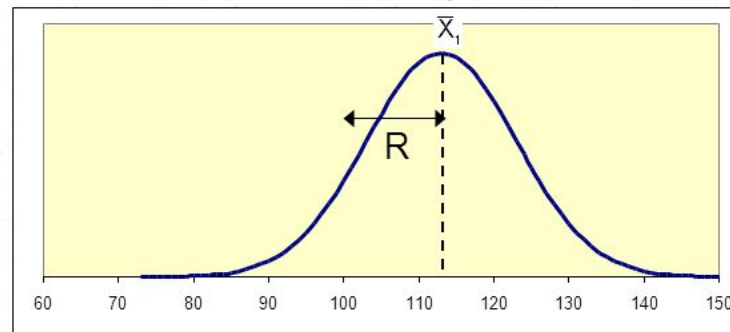
$$R = \bar{X}_1 - \bar{X}_0$$

Realized heritability

$$h^2 = \frac{R}{S} = \frac{113 - 100}{117 - 100} \approx 0.75$$



Recombine to form C_1



One more generation:

We know: $h^2 = 0.75$

We choose $X_S = 133$

What is R ?

Falconer and Mackay, Chapt. 11

Selection response

Response to selection

$$R = h^2 S$$

Selection differential

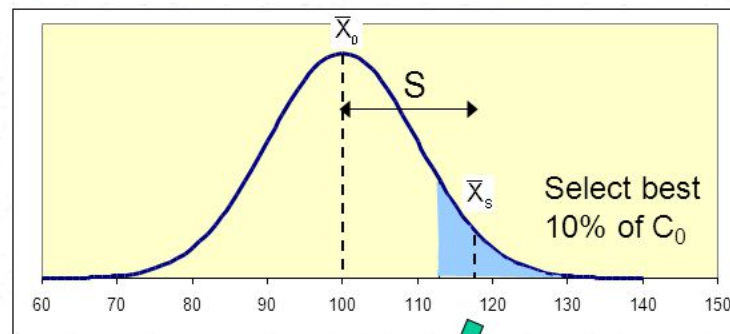
$$S = \bar{X}_S - \bar{X}_0$$

Response to selection

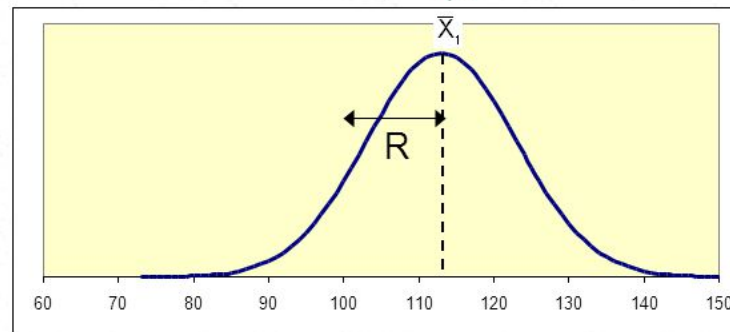
$$R = \bar{X}_1 - \bar{X}_0$$

Realized heritability

$$h^2 = \frac{R}{S} = \frac{113 - 100}{117 - 100} \approx 0.75$$



Recombine to form C_1



One more generation:

We know: $h^2 = 0.75$

We choose $X_S = 133$

What is R ?

$$S = 133 - 113 = 20$$

$$R = 0.75 * 20 = 15$$

$$X_2 = 128$$

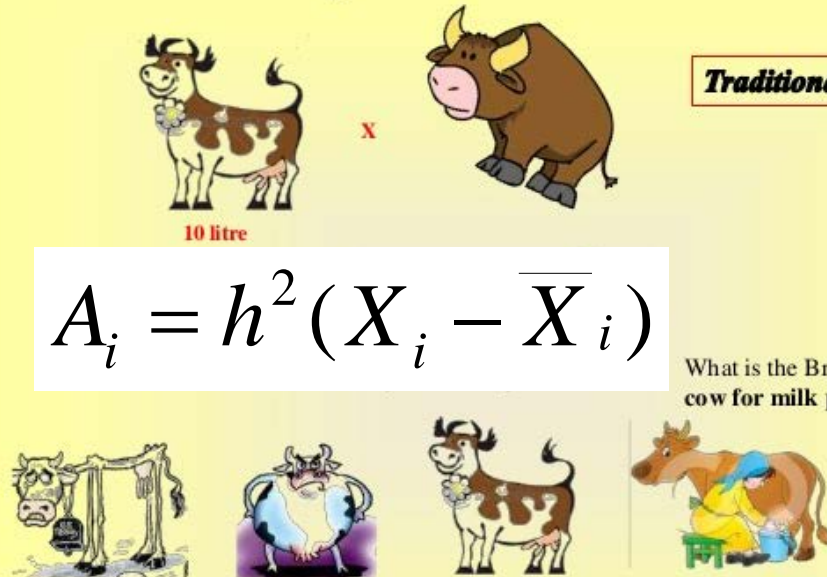
Falconer and Mackay, Chapt. 11

Selection limits

- There may be several reasons for reaching a plateau in genetic gain in a breeding program
 - Narrow genetic variation in the base population
 - Small effective population size resulting in inbreeding and increased homozygosity
 - Few loci controlling the traits selected for
 - Artificial selection may be outweighed by natural selection
- But for quantitative traits controlled by a large number of genes, selection limits will rarely be reached if inbreeding is kept low
 - Thus selection is a powerful tool to change animal populations in the desired direction

Breeding value

How to estimate Breeding value?




Traditional selection

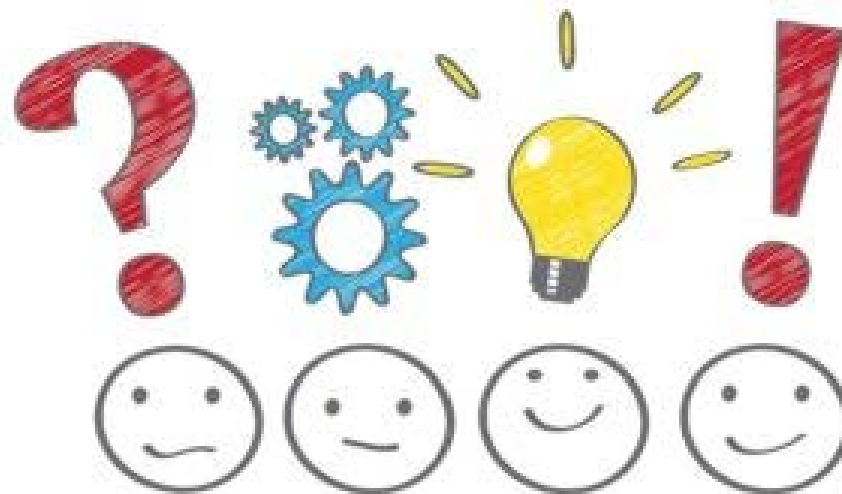
$$A_i = h^2 (X_i - \bar{X}_i)$$

What is the Breeding value of this cow for milk production?

0.5 litre 8 litre 10 litre 12 litre

Breeding value = $h^2(\text{milk production} - \text{average})$
 = $(12 - 7.625) * h^2 = 4.35 \text{ litres}$





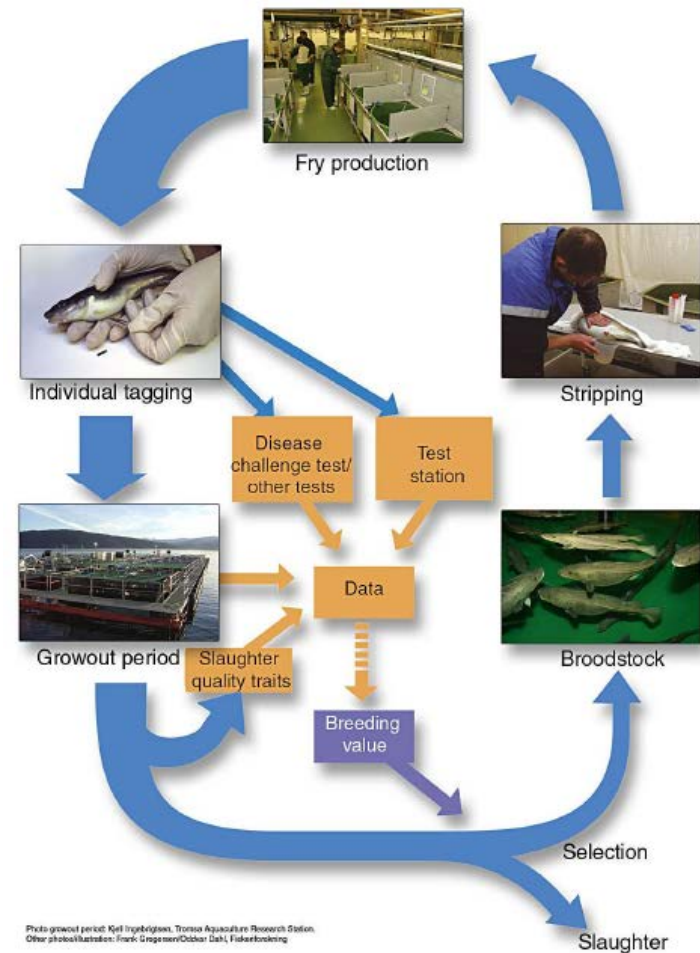
Use of genetic/genomic tools in breeding programs



"He's being bred for Hawaiian pizza!"

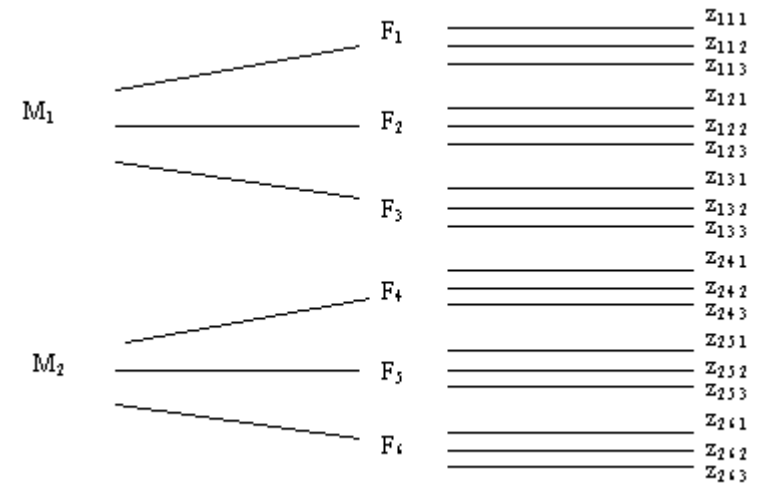
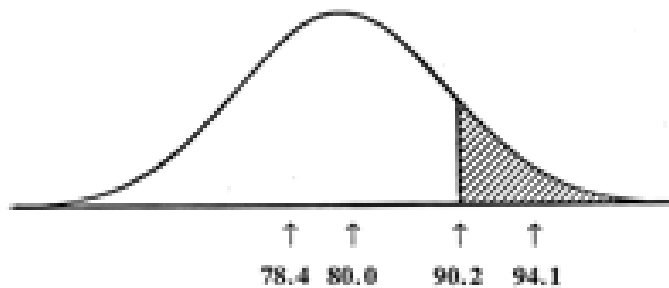
Breeding programs

- A **breeding program** is the planned breeding of a group of animals or plants, usually involving at least several individuals and extending over several generations
- Simple individual based selection – largest animals/plants used for next generation
- Family based breeding programs selecting for multiple traits



Difference between breeding programs

- Mass selection
- Family based breeding



From Lynch and Walsh, 1998

What's the advantage/disadvantages of the two methods

Methods of artificial selection

- Many economically important traits cannot be recorded on live individuals
 - Disease resistance and product quality
 - Lipid content, feed conversion, fillet yield, flesh color, etc.
- *family based selection* is a key strategy
 - Records of close relatives allow selection for traits like disease resistance and product quality

- Breeding value of individual I

$$A_i = h^2 (X_i - \bar{X}_i)$$

- Breeding value of full sib families

$$- r_G = 0.5$$

$$A_j = \frac{nh^2 (X_{jn} - \bar{X}_{jn})}{2 + (n-1)(h^2 + 2c_{FS}^2)}$$

- Breeding value of half sib families

$$- r_G = 0.25$$

$$A_j = \frac{nh^2 (X_{jn} - \bar{X}_{jn})}{4 + (n-1)(h^2 + 4c_{HS}^2)} \quad 1/2019$$

Basic conditions for a breeding program

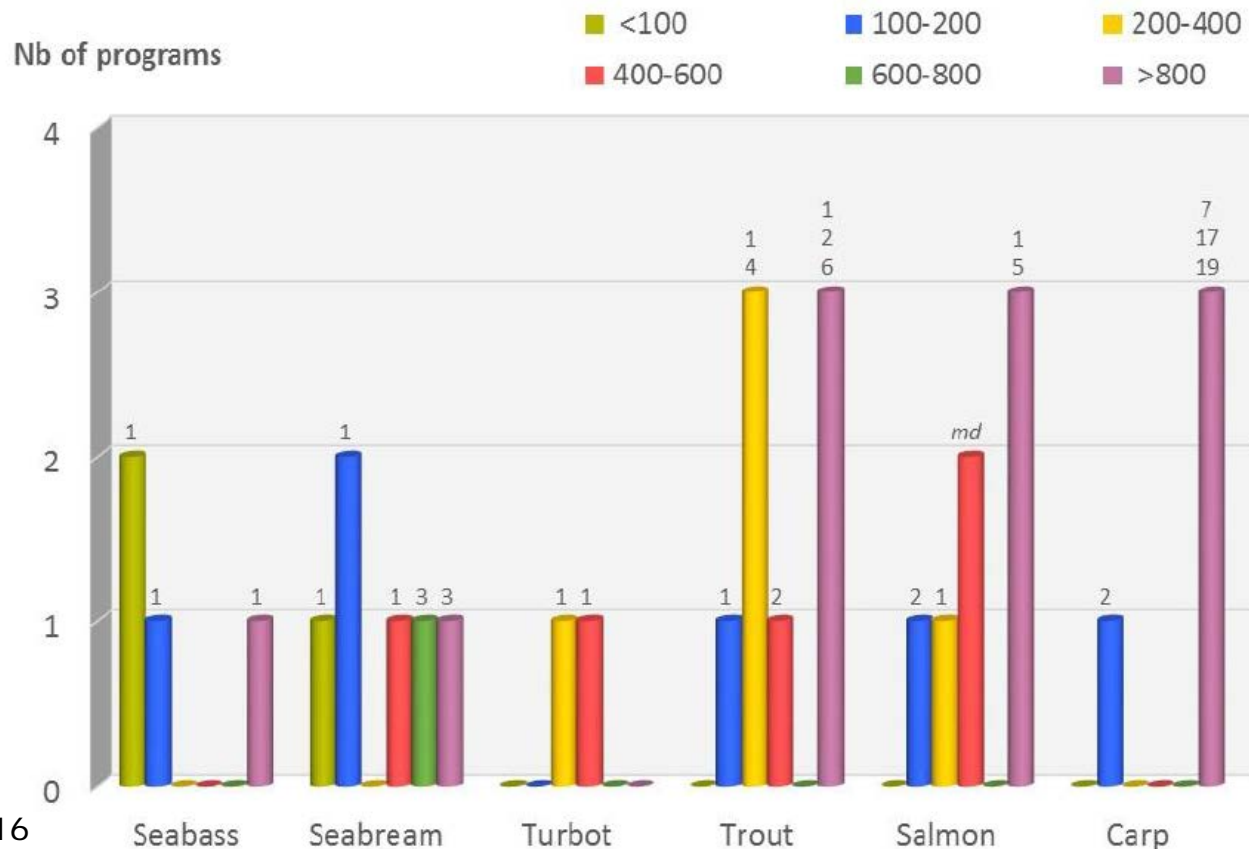
- There must be variation between individuals for the traits of interest
- A part of the variation has to be under genetic control as this is the only part which will be transferred to the next generation
- The life-cycle of the organism must be known and possible to control
- Individuals must be identifiable (tagging!) in order to keep track of their pedigree

How to start a breeding program

- Establishment of a baseline population:
- General idea to use several wild and or farmed populations to assure sufficient genetic variation
-
- However, depend on geographical and environmental coverage (local adapted populations?)
- Level of inbreeding in farmed populations should be assessed
- A minimum of 100 males and 100 females should be used – but more is better (more to choose from, less likelihood of inbreeding)

2014 Survey of number of broodfish used

Figure 3: Number of programs according to the number of broodfish involved at each generation, grouped by species.



Chavanne et al. 2016

The values above series represent the number of lines used in each program. *md*: missing data.

Breeding goals

- Reduced cost of production (ensuring price competitiveness)
- High product quality (ensuring competitiveness in the market)
- Improved welfare and reduced stress
-

Traits:

- The trait must be of economic or ethical importance
- It must show variation and part of the variance must be heritable
- It should be possible to measure the trait accurately at a reasonable cost

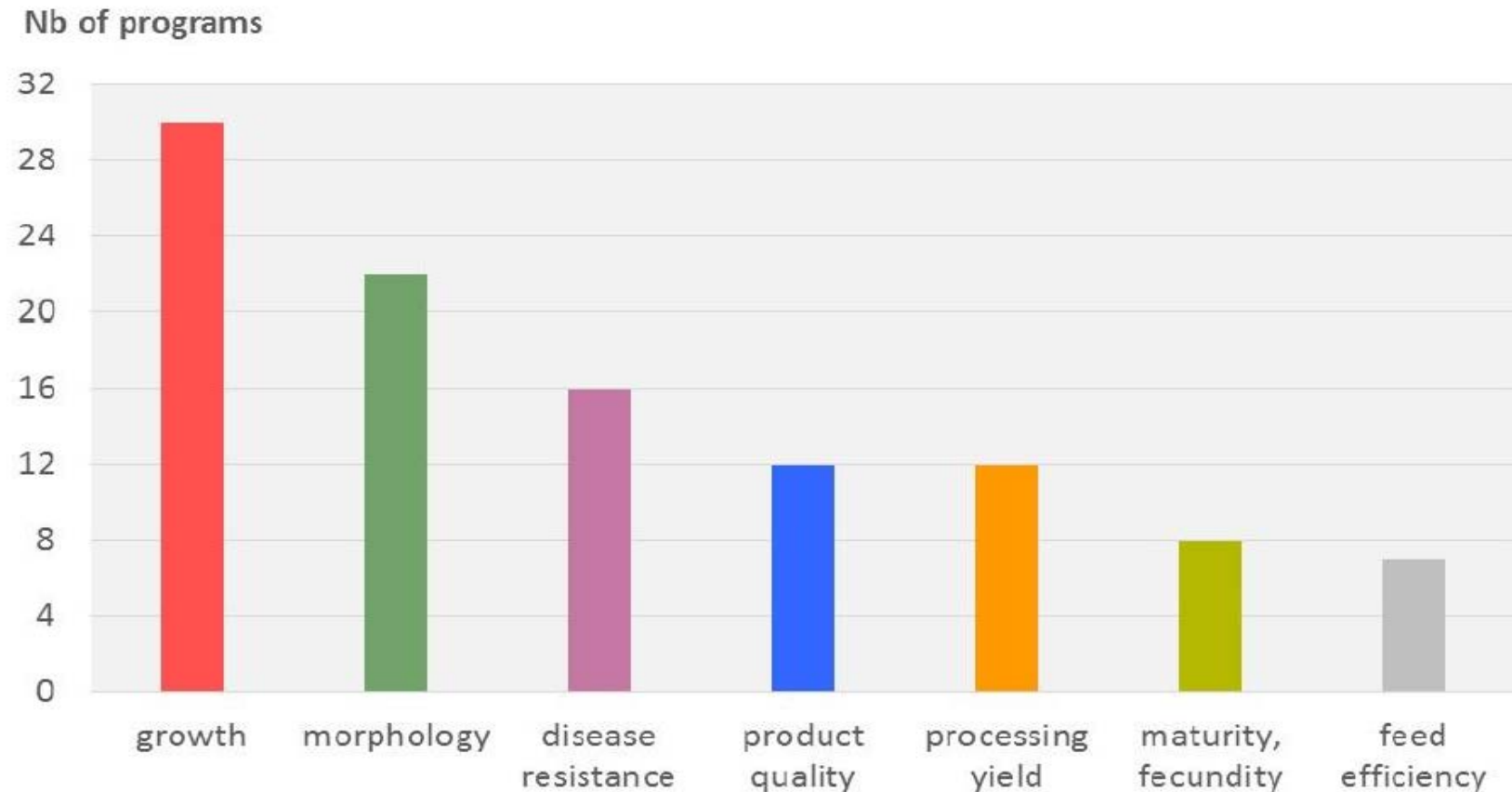
Breeding goals

- Provide as many potential breeding goals with associated traits as possible in 5 minutes.



Primary selected traits in European aquaculture

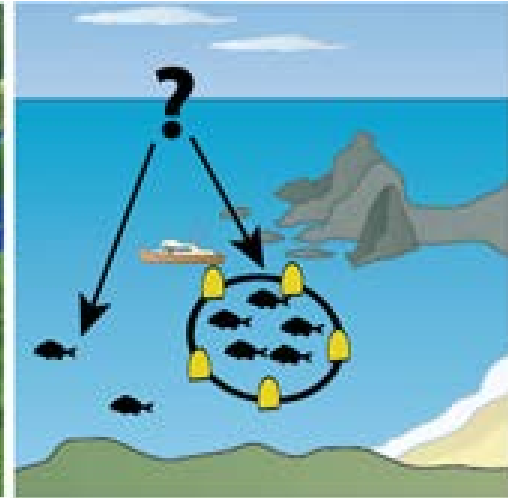
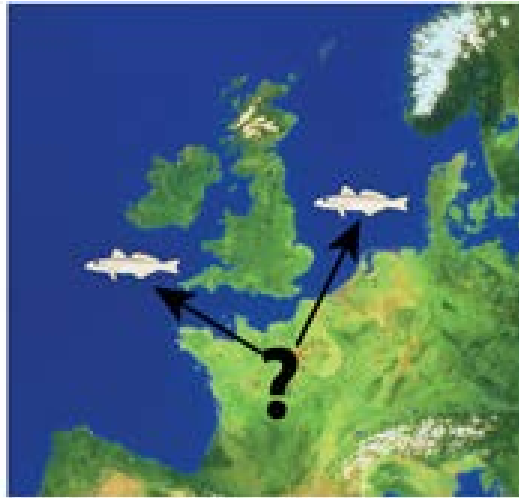
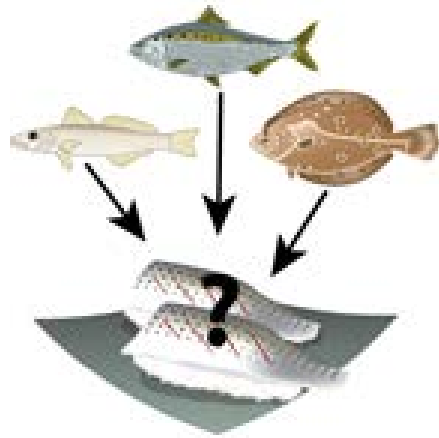
Figure 6: Number of programs according to the selected traits, all species included.



Application of genetic/genomic tools

- Genetic traceability
- Assessment of diversity/inbreeding
- Pedegreeing
- Mapping of traits (QTL)
- Marker assisted selection
- Genomic selection

Genetic traceability



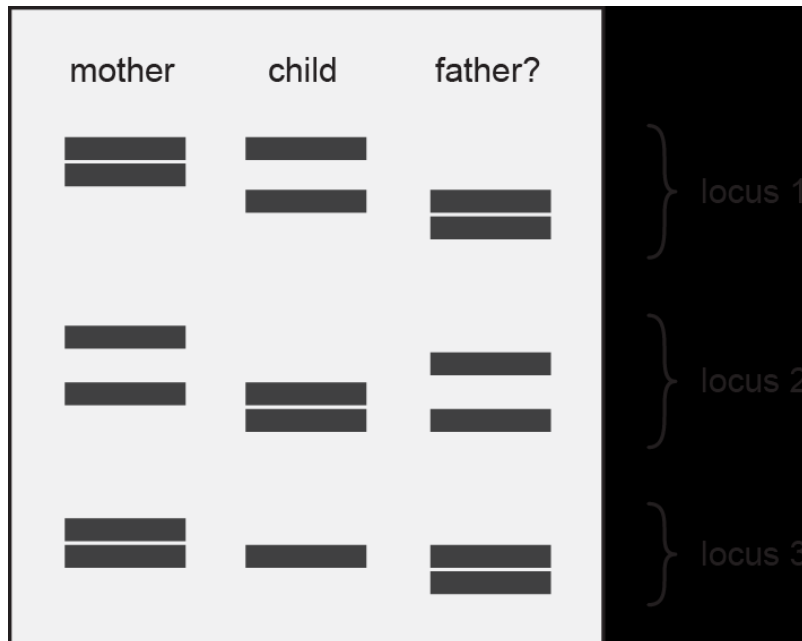
Protection of Intellectual property rights

Genetic diversity indices

- Heterozygosity
- Number of alleles
- Inbreeding
- Effective number of breeders



DNA fingerprinting (pedigree establishment)

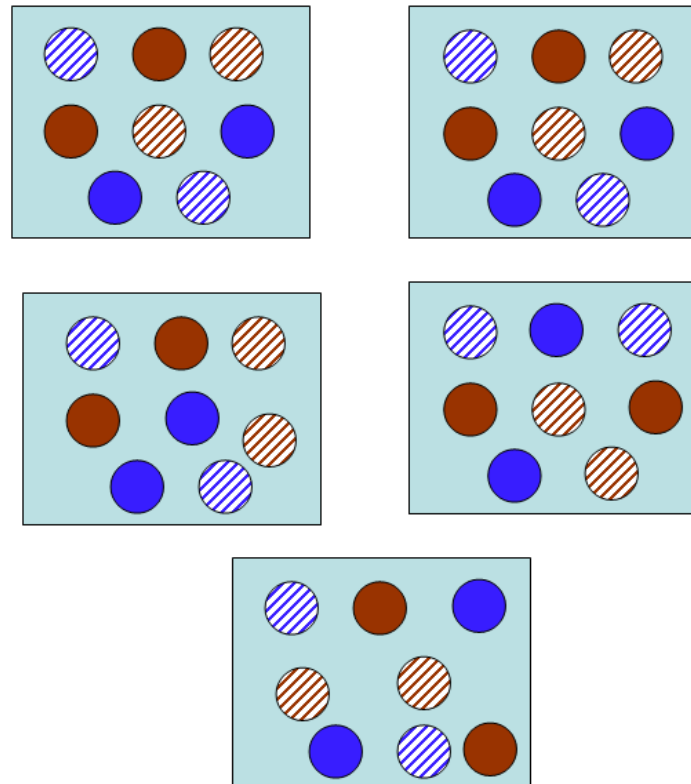


Common Garden experiments




$$P = G + E$$

No need for tagging of individuals

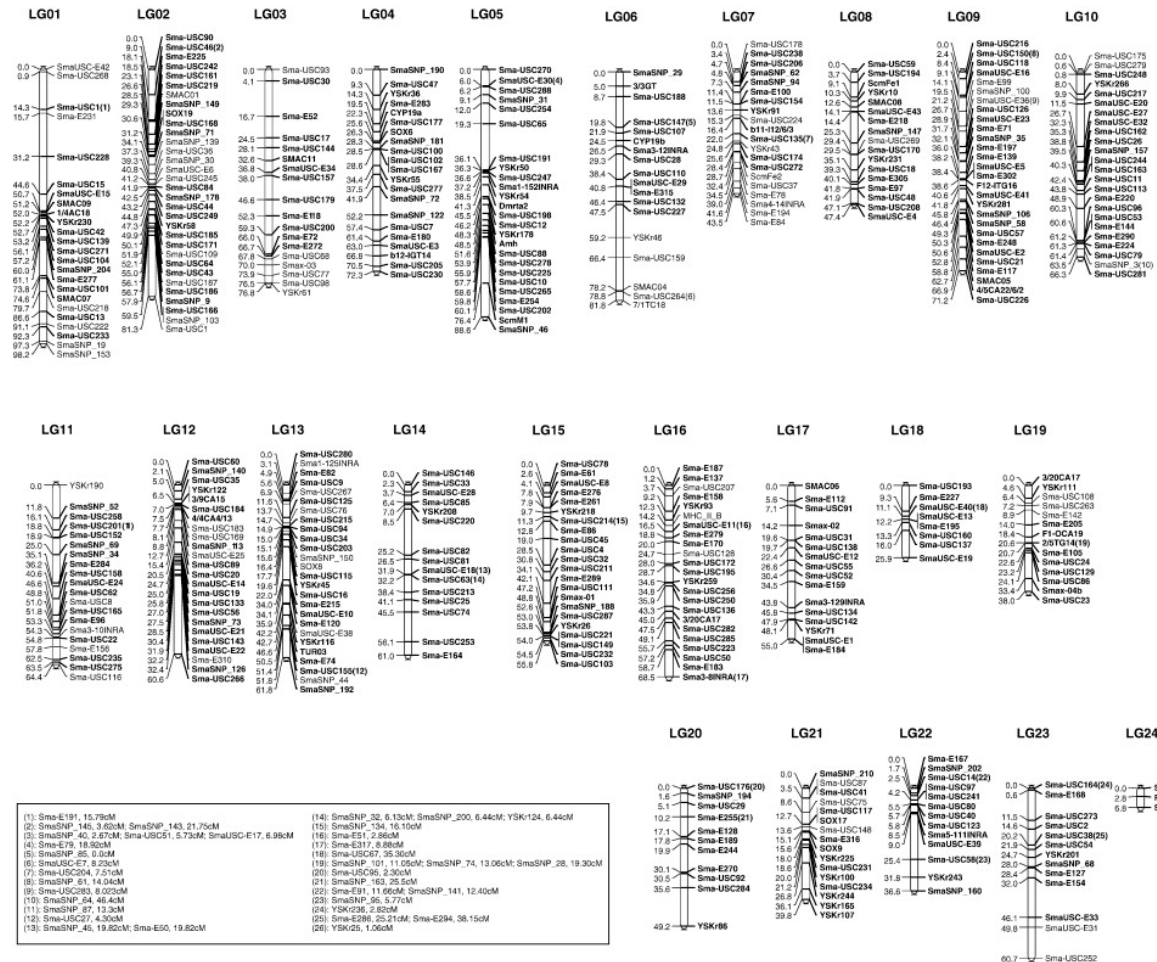
Overall Design



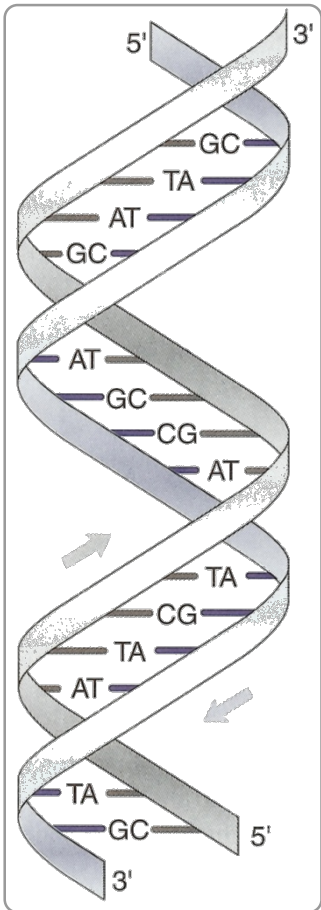
Five randomized blocks with eight pools each. Each of the 4 water and nutrient treatment combinations occurs in two pools within each block. Two pools are required since six plants are the maximum that can be placed in a single pool. Position of pools is randomized within block.

-  Low Nutrient/High H₂O
-  Low Nutrient/Low H₂O
-  High Nutrient/High H₂O
-  High Nutrient/Low H₂O

Genetic mapping

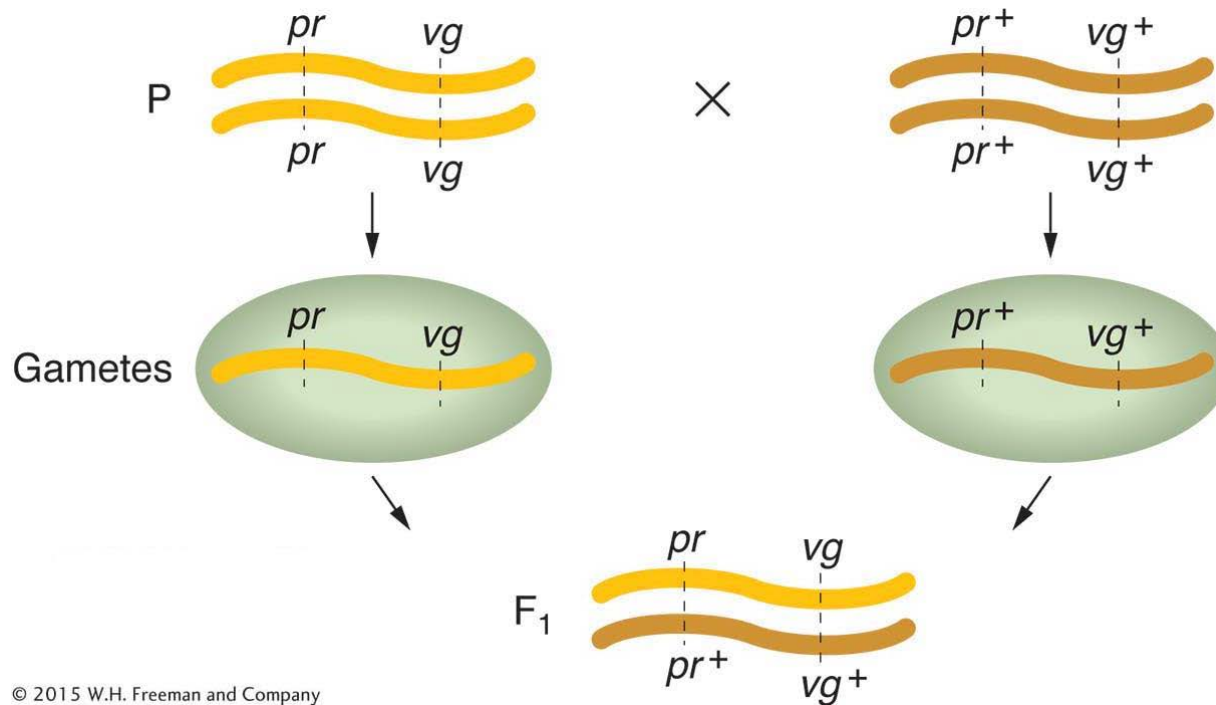


Genetic mapping - markers



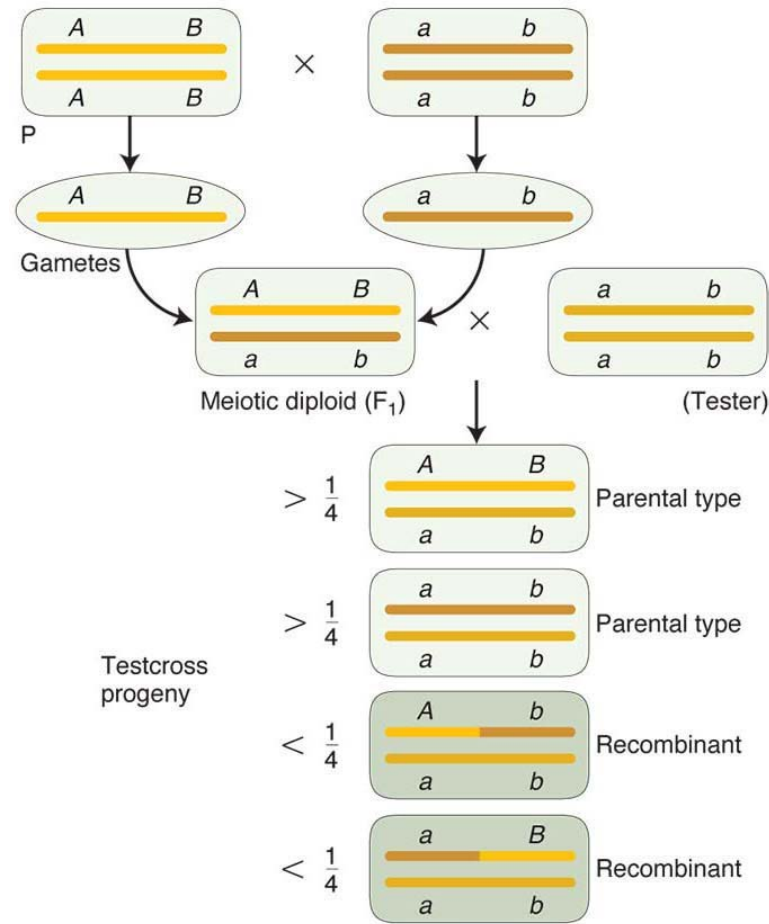
- E.g., 3 billion base pairs in the human genome
- Two genomes differ by 0.1%
 - 3 million differences between two individuals
- > 9 million known variable sites in the genome (**single nucleotide polymorphisms, SNPs**)
- Objective is to find those genetic variants that are of importance
 - directly important → genes influencing trait of interest
 - indirectly important → genetic markers

Linked alleles tend to be inherited together



© 2015 W.H. Freeman and Company

Tester cross



© 2015 W.H. Freeman and Company

Concepts: Linkage and Linkage Mapping

Linkage map:

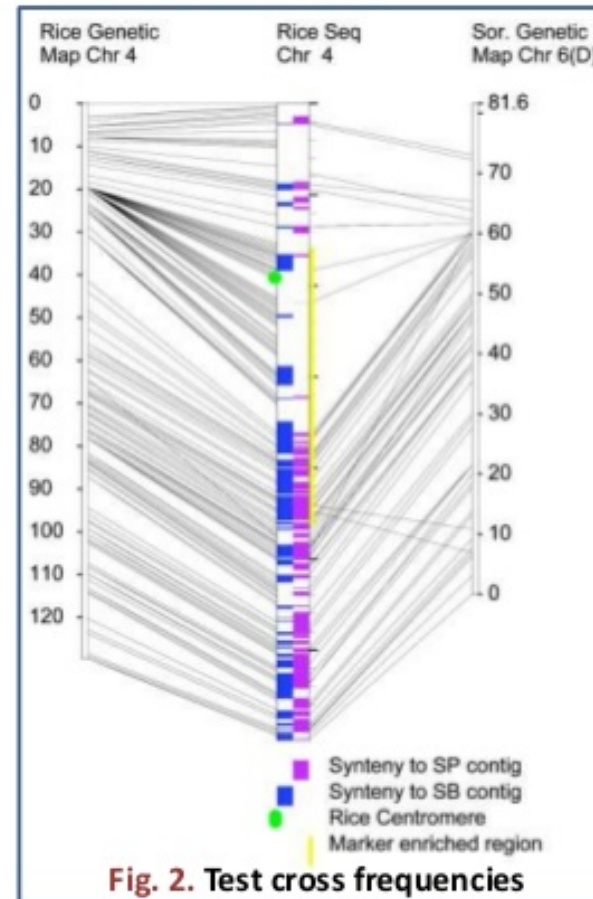
- “is a map of the frequencies of recombination that occur between markers on homologous chromosomes during meiosis.”
- distance is measured in cM.

Physical map:

- “shows the physical locations of genes and other DNA sequences of interest.
- distance measure in base pairs

Comparative map:

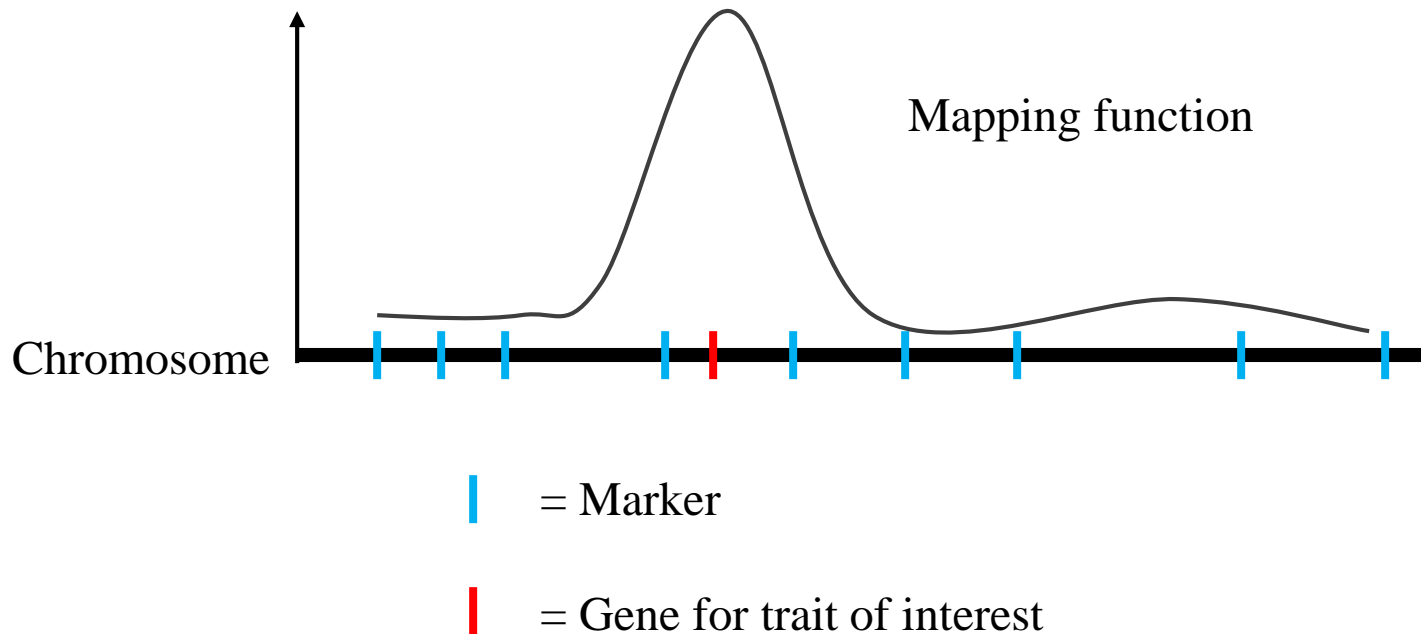
- a map that compares linkage maps or physical maps of related species based on shared markers or sequences, respectively (Fig. 2)



Source: Fig. 2 - www.pnas.org/content/102/37/13206/F3.expansion.html

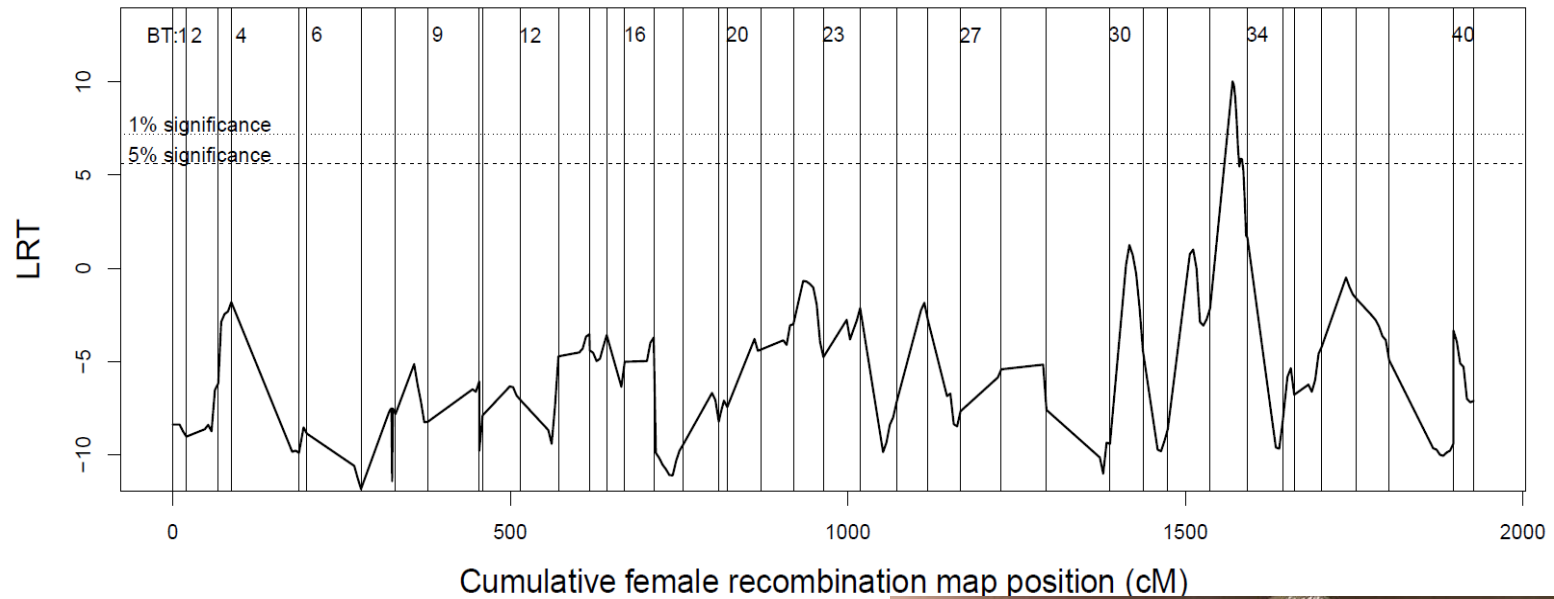
Basic principles of genetic mapping

Pr (Gene location | data)



Need many polymorphic markers and information on trait variation in strain

Length at age (1) in brown trout

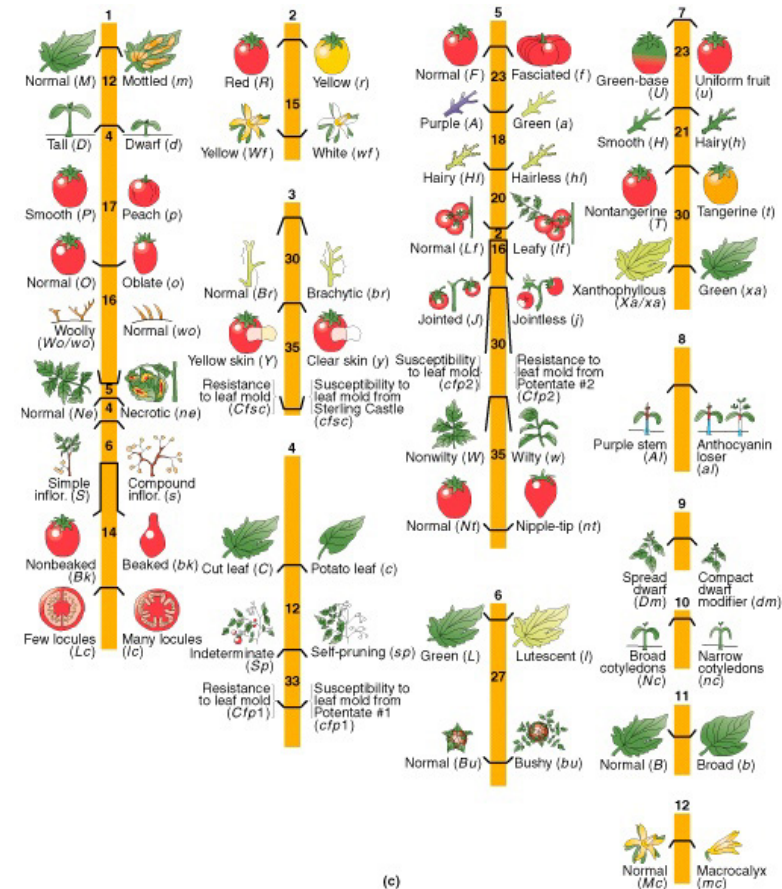


Different approaches to mapping

- **Genetic mapping** - Locating genes which affect the phenotype by examining genetic markers
- **Pedigree based methods:**
 - Linkage-mapping
 - Count recombination events between markers and putative gene
 - Goal: use markers to find genes via linkage with markers
- **Population based methods:**
 - Linkage disequilibrium or association mapping
 - Use population history to detect markers in proximity with the putative gene
 - Goal: use markers to find genes via Linkage Disequilibrium (LD)

QTL linkage-mapping

- In organisms where experimental breeding is possible
- Often done by crossing two inbred lines that have been exposed to selection on the trait of interest but in opposite directions
 - Not possible in humans
 - Not possible in blue whales

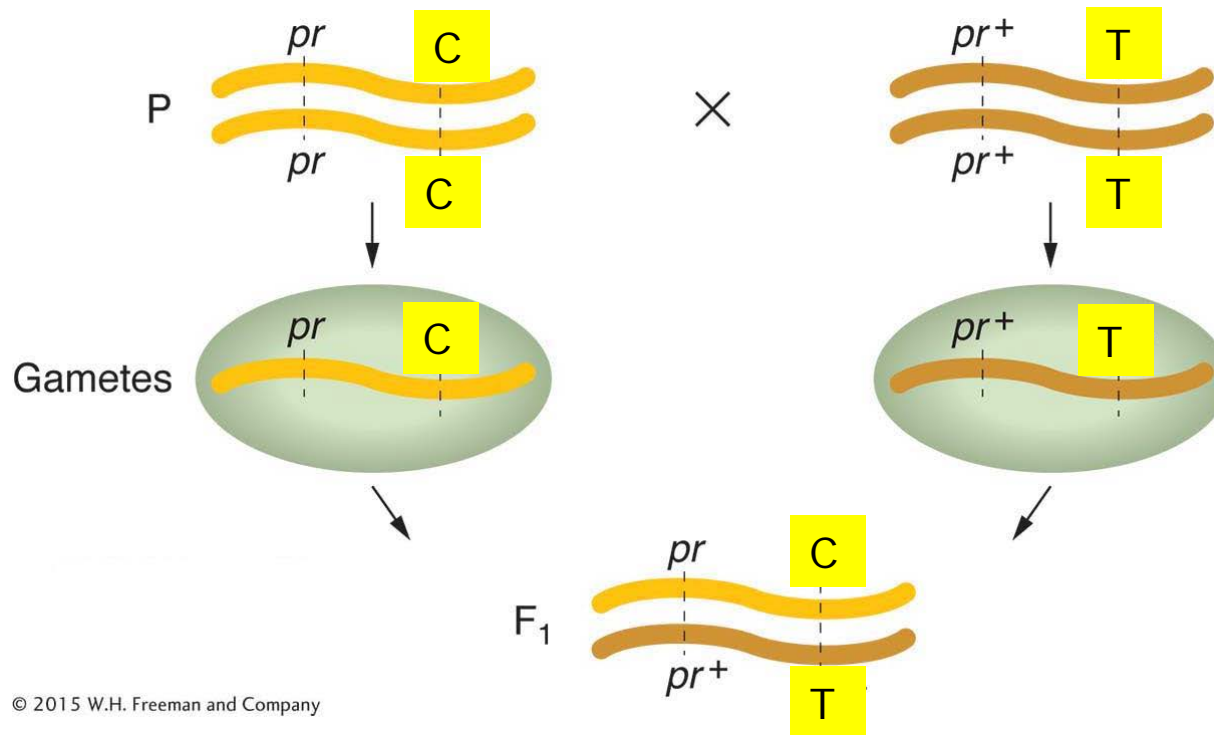


Association mapping

Goal: Use genetic markers to find functional genes via linkage disequilibrium (LD)

The essential idea is that markers close to the gene coding for the functional variation may also have allele frequency differences between phenotypes if there is linkage disequilibrium between the marker locus and the gene of interest

Linked alleles tend to be inherited together

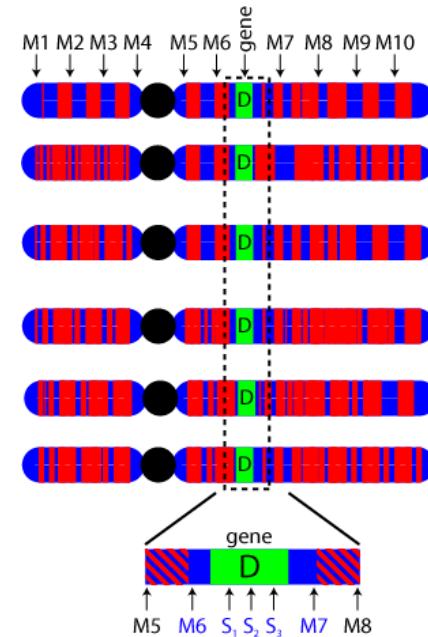


© 2015 W.H. Freeman and Company

Causes of association

A marker M can be informative of the gene D due to:

- Direct causation ($M=D$)
 - Linkage disequilibrium (M in LD with D)
 - Population substructure
 - Statistical artefacts
 - Technical artefacts
- } false-positive association



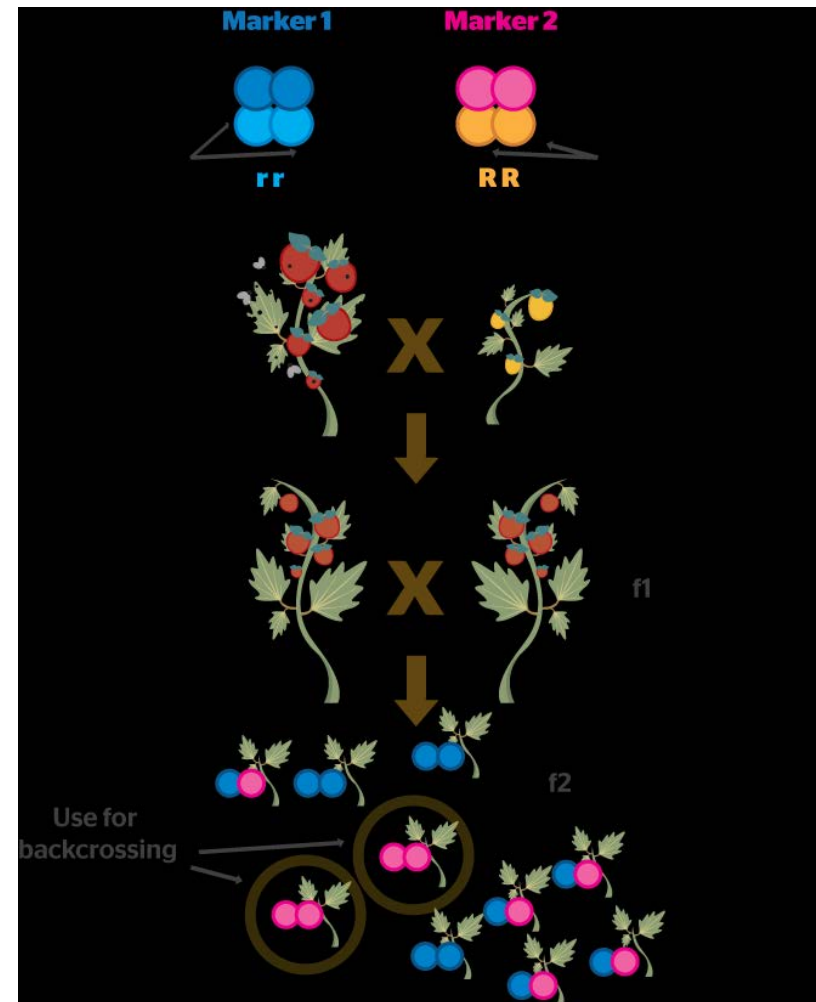
Marker Assisted Selection

- **MAS** is an indirect selection process where a trait of interest is selected based on a marker linked to a trait of interest (e.g. productivity, disease resistance, abiotic stress tolerance, and quality), rather than on the trait itself
- Can supplement/replace trait measurement, which can be tedious and expensive
- Can be used for traits not recordable on live individuals
- Is not influenced by environment
- Is not influenced by life/developmental stage of individual

Marker assisted selection

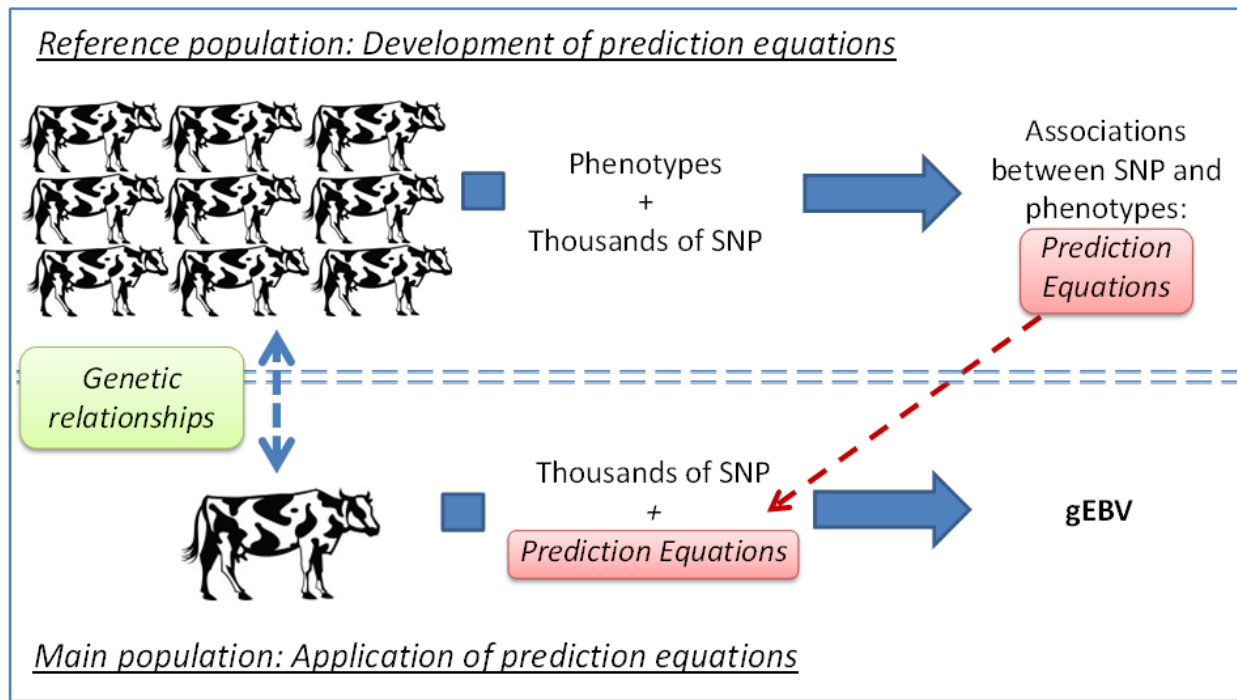
Development of resistant strains of strawberry

- Marker linked to QTL for resistance
- Crossing of elite non-resistant strain with wild conspecific
- Selection of F2 seedlings (embryos) homozygous for the resistance linked marker



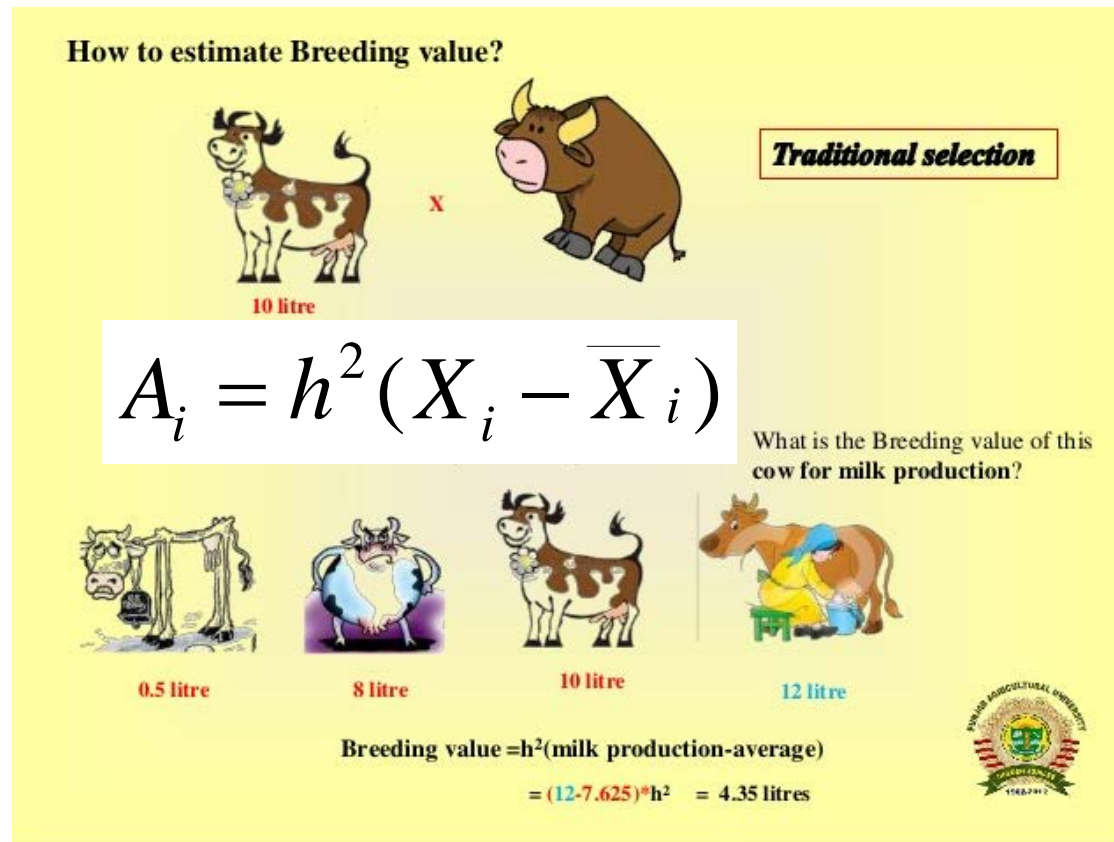
Genomic selection

- Genomic selection is a form of marker-assisted selection in which genetic markers covering the whole genome are used so that all quantitative trait loci (QTL) are in linkage disequilibrium with at least one marker



Breeding value

How to estimate Breeding value?




Traditional selection

$$A_i = h^2 (X_i - \bar{X}_i)$$

What is the Breeding value of this cow for milk production?

0.5 litre 8 litre 10 litre 12 litre

Breeding value = $h^2(\text{milk production} - \text{average})$
 = $(12 - 7.625) * h^2 = 4.35 \text{ litres}$

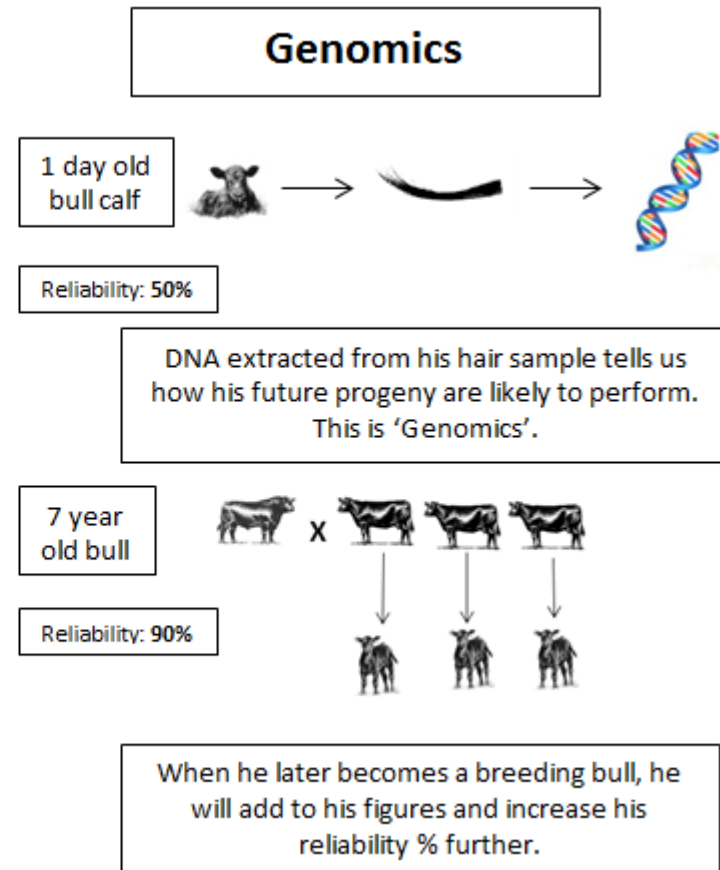
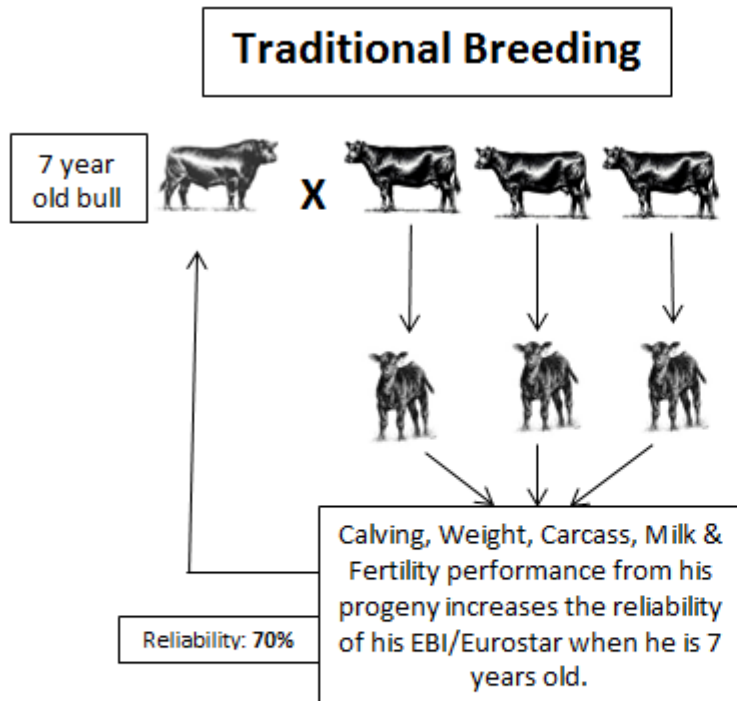


Genomic breeding values (GEBVs)

Are recorded through:

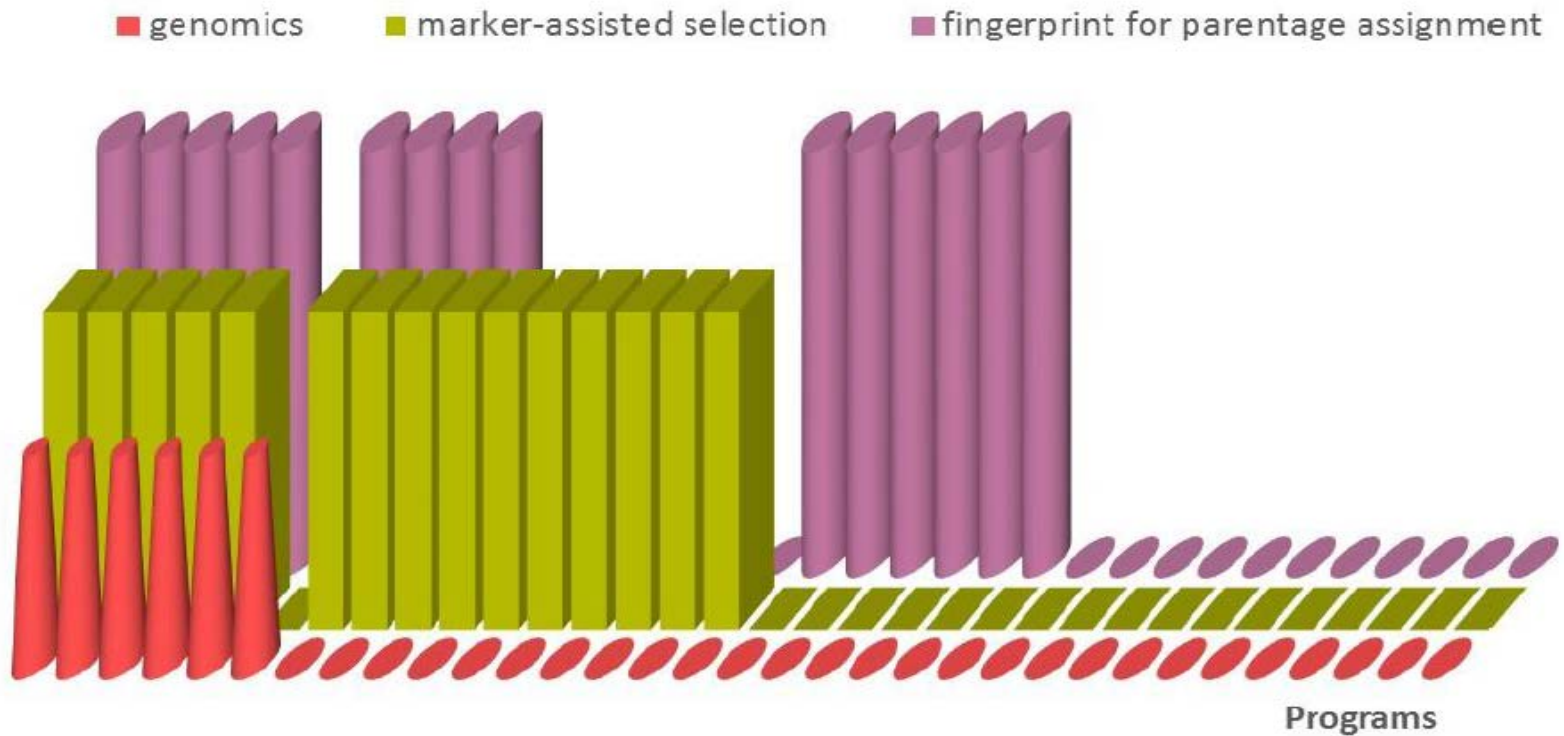
- Reference or training population
 - 1) Estimation of trait values
 - 2) Molecular marker scores (LD with QTL's)
 - 3) Pedigree information or kinship
 - 4) Establishment of prediction equations
- Validation population
 - Evaluation of precision of GEBVs

Advantages of Genomic Selection



Use of molecular tools in European fish breeding

Figure 5: Use of molecular tools in each program.



Summary

- The theory of breeding genetics was developed early 20th century
- Estimation of the heritability is key for all breeding genetics
- The breeders equation allow estimation of the expected response to selection
- The breeding value of an individual and family can be estimated
- QTL mapping is an important application in modern breeding
- Mapping allow Marker Assisted Selection and Genomic selection to be performed
- Genetic tools are also used for traceability, variability and pedigree establishment