

Introduction to Quantitative Genetics

Genetics

- ✓ The study of heredity
- ✓ The study of how differences between individuals are transmitted from one generation to the next
- ✓ The study of how information in the genes is used in the development and functioning of the adult organism

Major Subdisciplines of Genetics

■ Transmission Genetics

It focuses on the transmission of genes and chromosomes in individuals from generation to generation.

■ Molecular Genetics

It focuses on the structure and function of genes at the molecular level.

■ Population Genetics

It focuses on heredity in groups of individuals for traits determined by one or only a few genes.

■ Quantitative Genetics

It focuses on heredity in groups of individuals for traits determined by many genes simultaneously (Quantitative characteristics).

It focuses on the relationship between offspring and parents

Parent-offspring relationship

- ✓ Both parents and offspring may experience the same environmental conditions

Such a situation could arise if there are strong preferences of the offspring for their natal environment. Offspring may return to the same general area for nesting in which they were raised

- ✓ The phenotype of the offspring may be determined by the phenotype but not the genotype of the parent

the ability of parents to provide food to their offspring depends on their condition which is determined by environmental factors

- ✓ The correspondence between the phenotype of the offspring and the parents is a consequence, in part, of the genes shared by the individuals

Quantitative Characteristics

- ✓ Many traits in humans and other organisms are genetically influenced, but do not show single-gene (Mendelian) patterns of inheritance.
- ✓ They are influenced by the combined action of many genes and are characterized by continuous variation. These are called **polygenic** traits.
- ✓ Continuously variable characteristics that are both polygenic and influenced by environmental factors are called **multifactorial** traits. Examples of quantitative characteristics are height, intelligence & hair color.

Types of Quantitative Traits

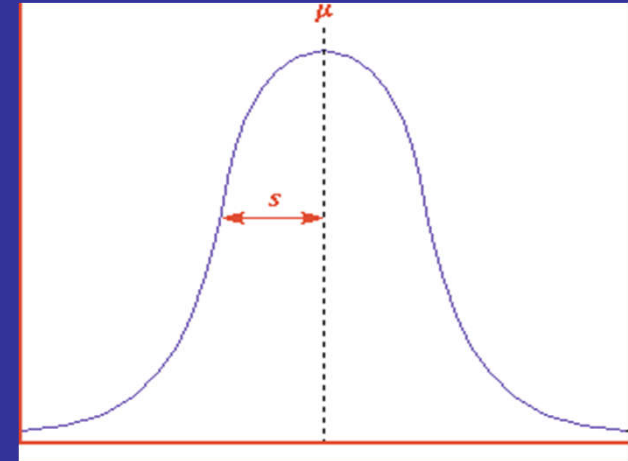
Quantitative characteristics are usually described by a measurement (quantity).

Example: How tall are you?

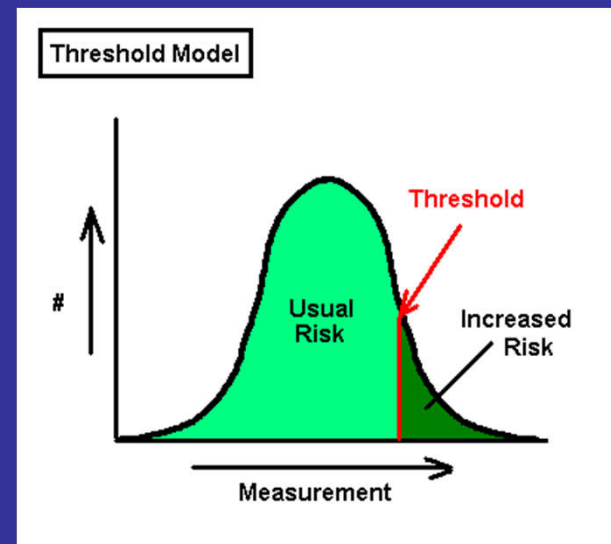
- ✓ Some quantitative traits are meristic (measured in whole numbers). For example, litter size.
- ✓ Another type of quantitative trait is a threshold characteristic which is either present or absent depending on the cumulative effect of a number of additive factors (diseases are often this type).

Types of Quantitative Trait

- In general, the distribution of quantitative traits values in a population follows the normal distribution (also known as Gaussian distribution or bell curve). These curves are characterized by the mean (mid-point) and by the variance (width). Often standard deviation, the square root of variance, is used as a measure of the curve's width.

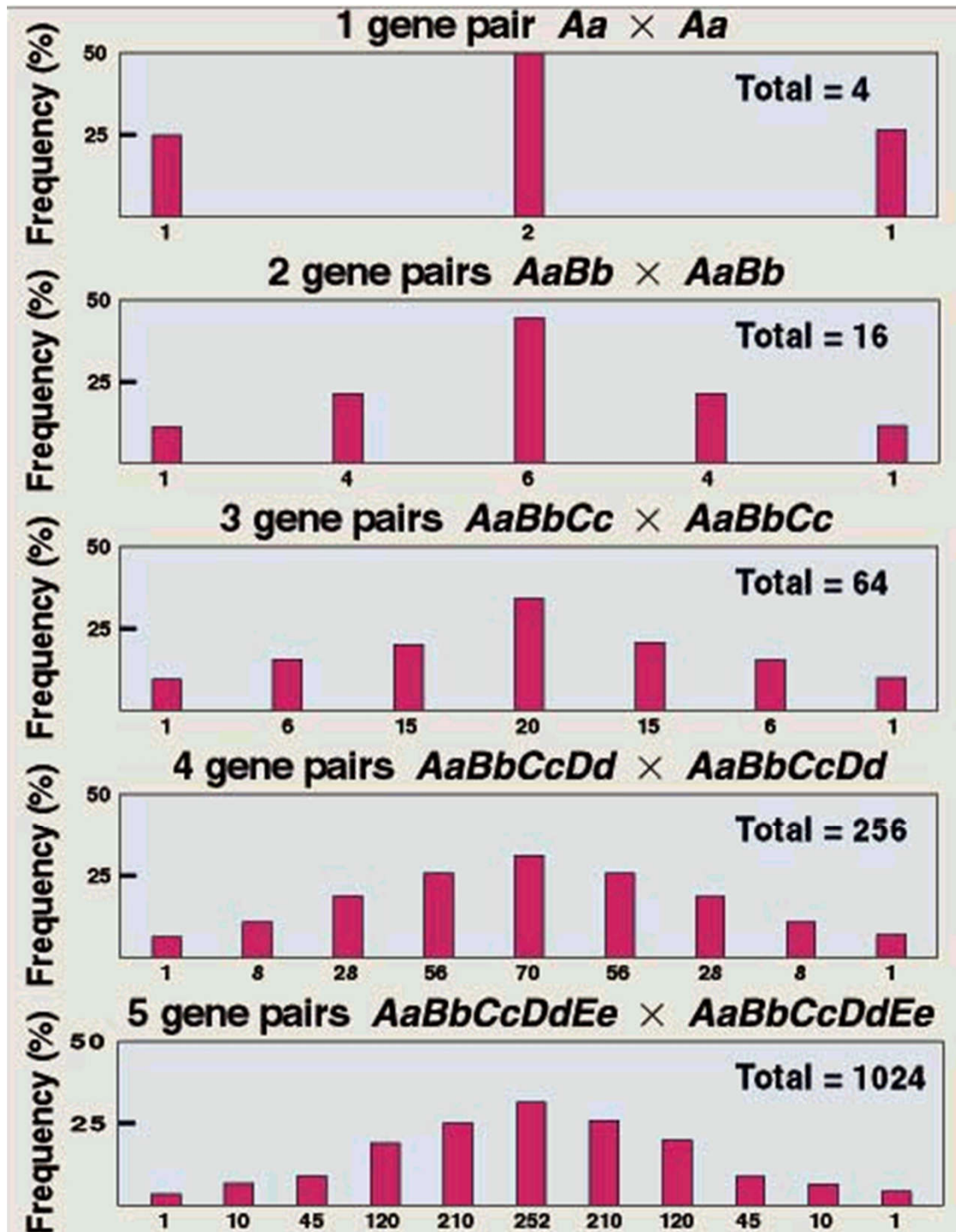


1. continuous trait: can take on any value: height, for example.
2. countable (meristic) can take on integer values only: number of bristles, for example.
3. threshold trait: has an underlying quantitative distribution, but the trait only appears only if a threshold is crossed.



Principles of Quantitative Inheritance

- Quantitative traits are influenced by the combined effects of numerous genes. These are called **polygenic** or **multifactorial** traits.
- The genes follow Mendelian laws of inheritance; however, multifactorial traits have numerous possible phenotypic categories.
- Environmental influences blur the phenotypic differences between adjacent genotypes.



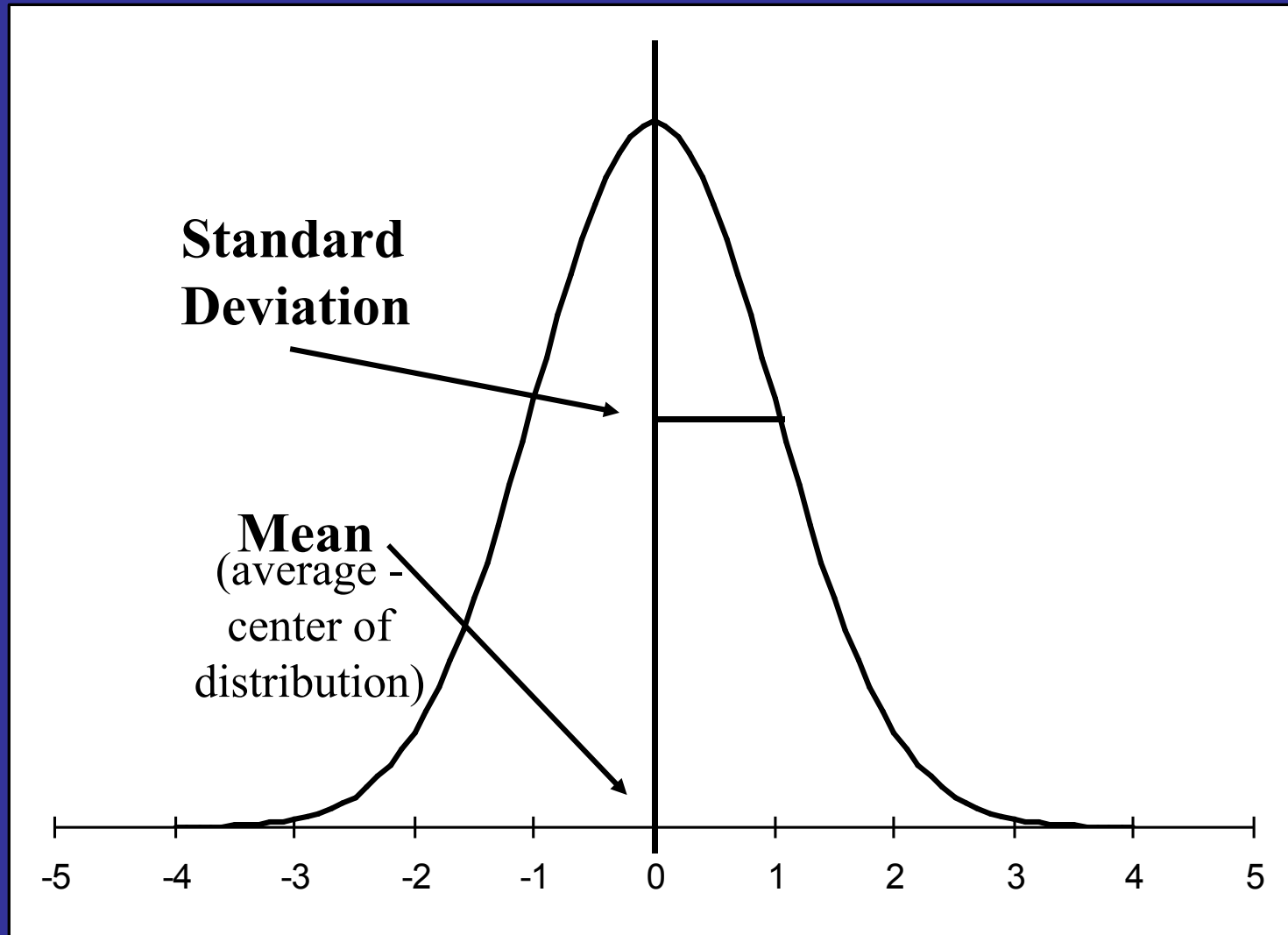
As the number of loci affecting the trait increases, the # phenotypic categories increases.

Number of phenotypic categories =

$$(\# \text{ gene pairs} \times 2) + 1$$

Connecting the points of a frequency distribution creates a bell-shaped curve called a normal distribution.

Normal Distribution



Mean $\pm 1s$ = 66% of values; $\pm 2s$ = over 95% of values

Quantitative Genetics

→ not discrete

- Continuous phenotypic variation within populations
 - Whole organism level
- Causes of variation
 - Genes vs. environment
 - Interactions between genes and environment
 - Components of genetic variation
 - Components of environmental variation

Why is quantitative genetics important?

Medicine

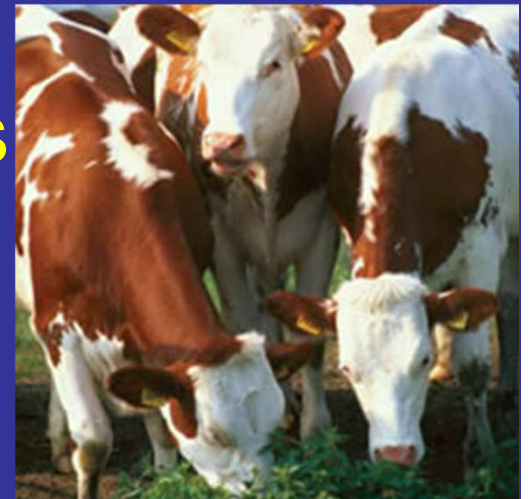
- Disease = variation
- Complex disorders
 - caused by multiple genetic and environmental factors
- Understanding genetic vs. environmental causes
 - prevention
 - genetic counselling
 - genetically-tailored treatments



Why is quantitative genetics important?

Agriculture

- Economically important traits = quantitative traits
- Quantitative genetics theory -> basis for breeding programs
- Environmental variation reduces efficiency of selection



Why is quantitative genetics important?

Consequences of inbreeding and outcrossing

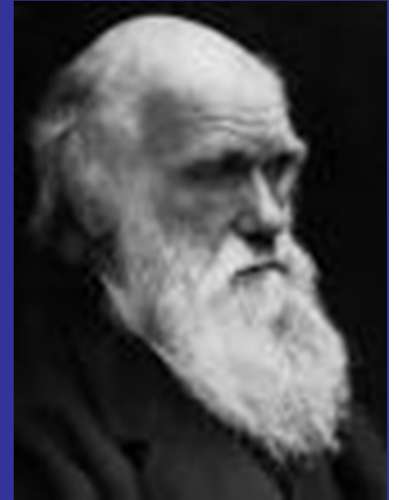
- Agriculture – inbred lines, hybrids, F1s
- Conservation – endangered species, captive breeding programs



Why is quantitative genetics important?

Evolution

- Natural selection requires heritable variation for traits
- What are the forces that maintain variation within populations?
 - Balance between selection, drift and mutation
 - Balancing selection?



Why is quantitative genetics important?

Evolution

- Does evolution proceed by small steps or big jumps?
- What is the relative importance of pre-existing variation vs. new mutations?

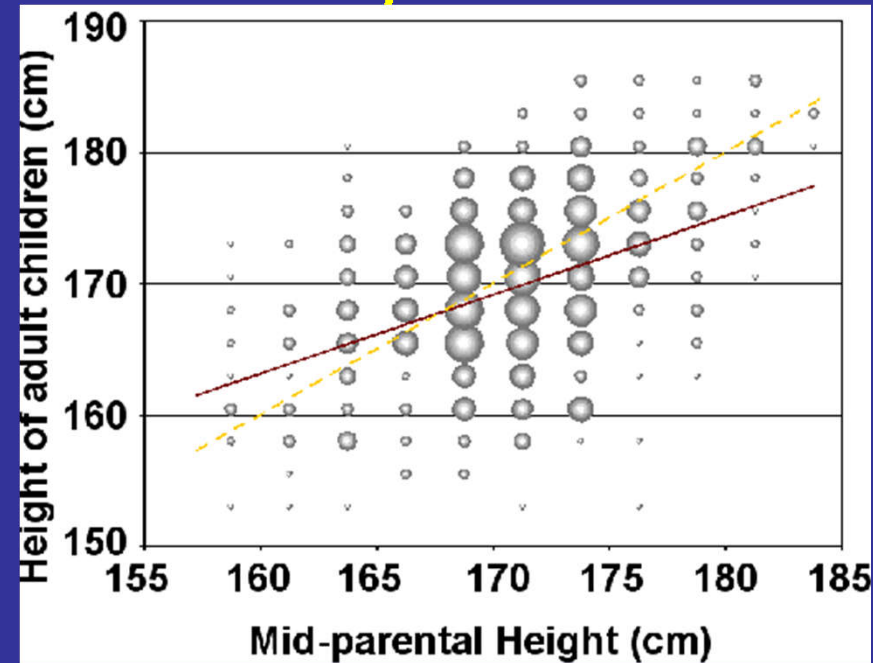


- Do genetic correlations between traits pose constraints on evolution?

History

Around 1900, there were two camps:

- Biometricians
 - Continuous traits
- Mendelians
 - Discrete traits

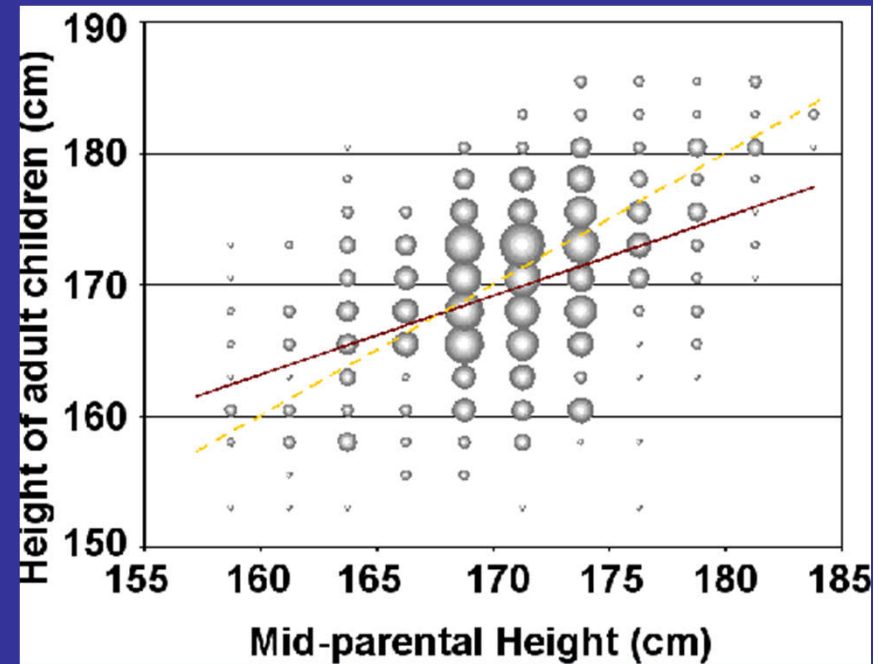


Are discrete traits inherited in the same way as quantitative traits?

History

Reconciliation:

- Multiple loci (genes) contribute to variation!



Is variation caused by a few loci of large effects or many loci with small effects?

Heterosis

- Both the parental lines and the F1's are genetically uniform. However, the parental lines are relatively small and weak, a phenomenon called "inbreeding depression": Too much homozygosity leads to small, sickly and weak organisms, at least among organisms that usually breed with others instead of self-pollinating.
- In contrast, the F1 hybrids are large, healthy and strong. This phenomenon is called "heterosis" or "hybrid vigor".
- The corn planted in the US and other developed countries is nearly all F1 hybrid seed, because it produces high yielding, healthy plants (due to heterosis) and it is genetically uniform (and thus matures at the same time with ears in the same position on every plant).

Mathematical Basis of Quantitative Genetics

- Recall the basic premise of quantitative genetics: phenotype = genetics plus environment.
- In fact we are looking at variation in the traits, which is measured by the width of the Gaussian distribution curve. This width is the variance (or its square root, the standard deviation).
- Variance is a useful property, because variances from different sources can be added together to get total variance.
- However, the units of variance are the squares of the units used to measure the trait. Thus, if length in centimeters was measured, the variances of the length are in cm^2 . This is why standard deviation is usually reported: length \pm s.d. --because standard deviation is in the same units as the original measurement. Standard deviations from different sources are not additive.
- Quantitative traits can thus be expressed as:
$$V_T = V_G + V_E$$
where V_T = total variance, V_G - variance due to genetics, and V_E = variance due to environmental (non-inherited) causes.
- This equation is often written with an additional covariance term: the degree to which genetic and environmental variance depend on each other. We are just going to assume this term equals zero in our discussions.

Heritability

- One property of interest is “heritability”, the proportion of a trait’s variation that is due to genetics (with the rest of it due to “environmental” factors). This seems like a simple concept, but it is loaded with problems.
- The broad-sense heritability, symbolized as H (sometimes H^2 to indicate that the units of variance are squared). H is a simple translation of the statement from above into mathematics:

$$H = V_G / V_T$$

- This measure, the broad-sense heritability, is fairly easy to measure, especially in human populations where identical twins are available. However, different studies show wide variations in H values for the same traits, and plant breeders have found that it doesn’t accurately reflect the results of selection experiments. Thus, H is generally only used in social science work.

Heritability

Measured using resemblance
between relatives

$$h^2 = \frac{\text{genetic variation}}{\text{phenotypic variation}}$$

Genetic + environmental + interaction



Heritability

(broad-sense)

Heritability (broad-sense) is the proportion of a population's phenotypic variance that is attributable to genetic differences

Additive vs. Dominance Genetic Variance

- The biggest problem with broad sense heritability comes from lumping all genetic phenomena into a single V_G factor. Paradoxically, not all variation due to genetic differences can be directly inherited by an offspring from the parents.
- Genetic variance can be split into 2 main components, additive genetic variance (V_A) and dominance genetic variance (V_D).
$$V_G = V_A + V_D$$
- Additive variance is the variance in a trait that is due to the effects of each individual allele being added together, without any interactions with other alleles or genes.
- Dominance variance is the variance that is due to interactions between alleles: synergy, effects due to two alleles interacting to make the trait greater (or lesser) than the sum of the two alleles acting alone. We are using dominance variance to include both interactions between alleles of the same gene and interactions between different genes, which is sometimes a separate component called epistasis variance.
- The important point: dominance variance is not directly inherited from parent to offspring. It is due to the interaction of genes from both parents within the individual, and of course only one allele is passed from each parent to the offspring.

Heritability

(narrow sense)

Heritability (narrow sense) is the proportion of a population's phenotypic variance that is attributable to **additive** genetic variance as opposed to dominance genetic variance (interaction between alleles at the same locus).

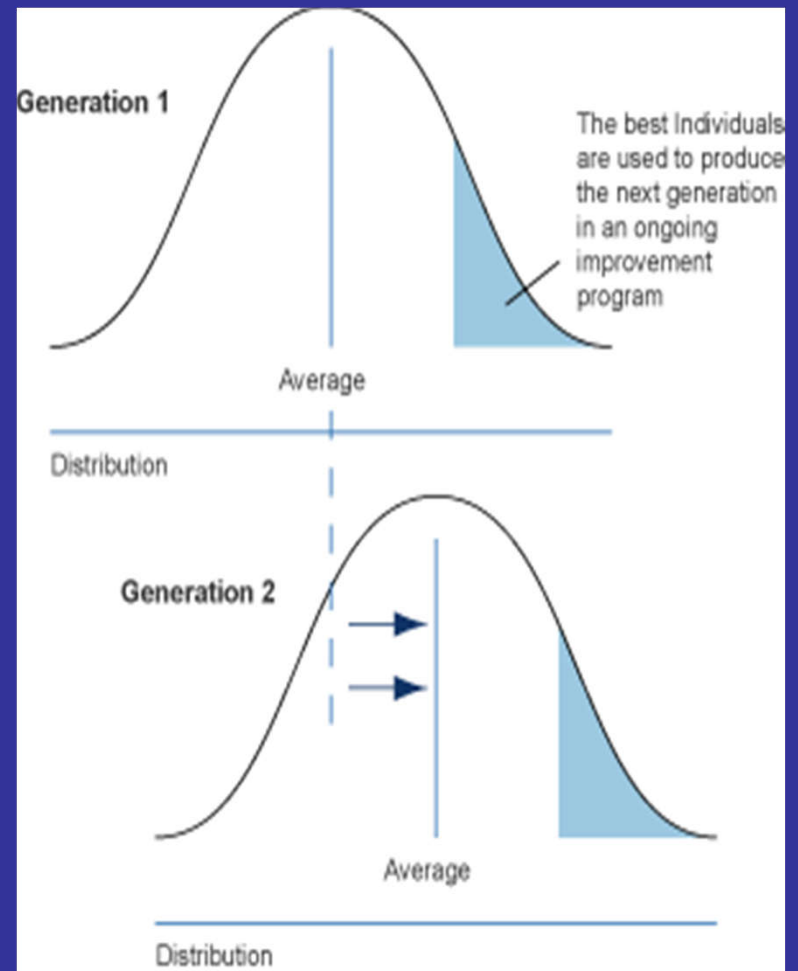
Additive genetic variance responds to selection

Narrow Sense Heritability

- For a practical breeder, dominance variance can't be predicted, and it doesn't affect the mean or variance of the offspring of a selection cross in a systematic fashion. Thus, only additive genetic variance is useful. Breeders and other scientists use "narrow sense heritability", h , as a measure of heritability.

$$h = V_A / V_T$$

- Narrow sense heritability can also be calculated directly from breeding experiments. For this reason it is also called "realized heritability".



The genetic Correlation

Traits are not inherited as independent unit, but the several traits tend to be associated with each other

This phenomenon can arise in 2 ways:

1. A subset of the genes that influence one trait may also influence another trait (pleiotropy)
2. The genes may act independently on the two traits, but due to non random mating, selection, or drift, they may be associated (linkage disequilibrium)

Basic formula:

$$r_G = \text{cov}_{XY} / (\text{var}_X \cdot \text{var}_Y)^{0.5}$$

r_G often used both for additive (r_A)
and genotypic (r_G) correlation!

Phenotypic correlation:

**A combination of genetic and environmental (incl.
nonadd gen effects) corr:**

$$r_P = h_X \cdot h_Y \cdot r_G + (1-h_X^2)^{0.5} \cdot (1-h_Y^2)^{0.5} \cdot r_E$$
$$r_P = h_X \cdot h_Y \cdot r_G + e_X \cdot e_Y \cdot r_E$$

The magnitude and even the sign of r_G cannot be
determined from r_P alone!

The use of genetic correlations

1. Trait-trait correlation

Relation between different traits.

For studies of how the improvement of one trait will affect another trait.

2. Age-age correlation

Relation between a trait at young and mature age. Gives info about when reliable estimations can be achieved.

3. Site-site correlation

Relation between genotype and environment. For deliniation of breeding and seed zones and for optimization of number of trials per zone

Another basic use of r_G is prediction of genetic gain.

Two basic estimations of r_G :

- **Burdon correlation, type A:**
Both traits are measured on the same individual (true genetic corr.). Trait-trait and age-age correlations
- **Burdon correlation, type B:**
Two traits are measured on different individuals (approximated genetic corr.). One trait expressed at two sites are considered as two different traits. Site-site correlations.

Some features of genetic correlations

$$r_G = \text{cov}_{XY} / (\text{var}_X \cdot \text{var}_Y)^{0.5}$$

- 1) The three components are hard to estimate with any precision, i.e. large materials are needed.
- 2) Strongly influenced by gene frequencies, i.e. it is valid for a certain population only. Genetic correlations are easily changed by selection.

Correlated response

If we select for character X, what will be the change of the correlated character Y?

$$CR_Y = i \cdot h_X \cdot h_Y \cdot r_G \cdot \sigma_{PY}, \text{ where}$$

CR_Y = the correlated response in trait Y,

i = the intensity of selection,

h_X and h_Y = the square root of the h^2

r_G = the genetic correlation between traits X and Y

σ_{PY} = the phenotypic standard deviation for trait Y.

The CR_Y can be expressed in percent by relating it to the phenotypic mean of variable Y.

Indirect selection

When we want to improve character X, but select for another character (Y) and achieve progress due to the correlated response.

$$CR_X / R_X = (i_Y \cdot r_A \cdot h_Y) / (i_X \cdot h_X)$$

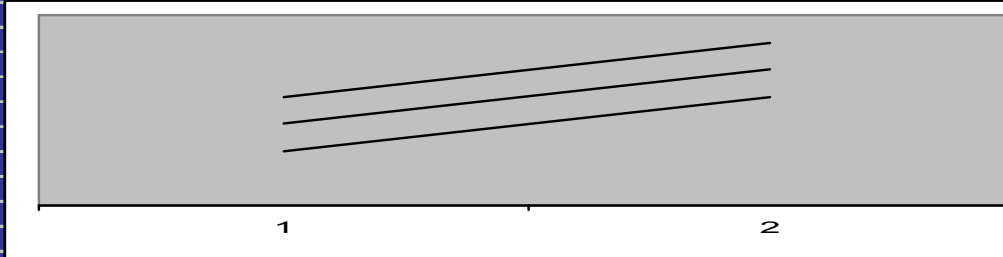
Presumptions:

$H_Y > H_X$ and strong CR or $i_Y > i_X$.

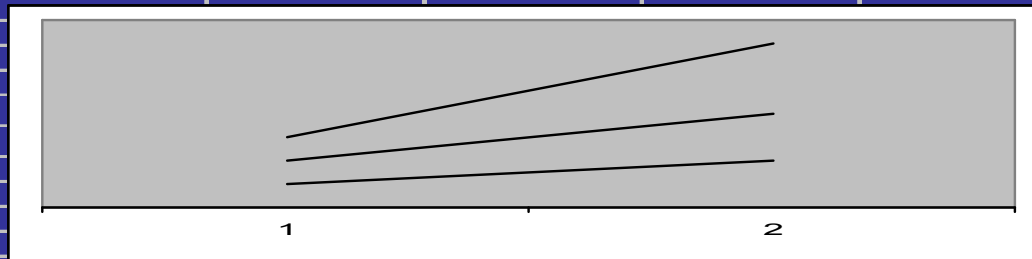
Usable when difficult to apply selection directly to the desired character:

- 1) Hard to measure with any precision, which reduces h^2
- 2) The desired trait is costly to measure. Then it would be better to select for an easily measurable, correlated trait.

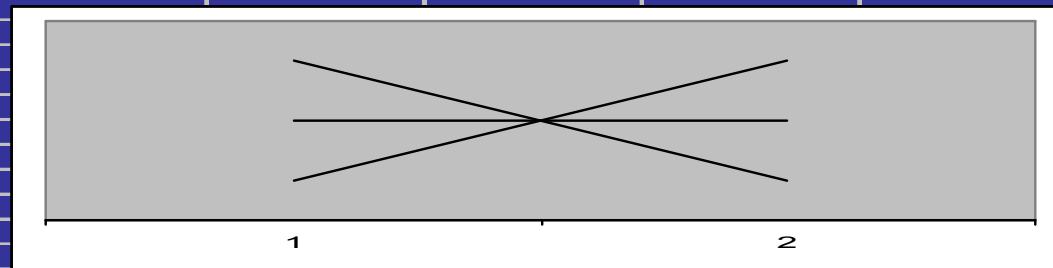
G x E interaction



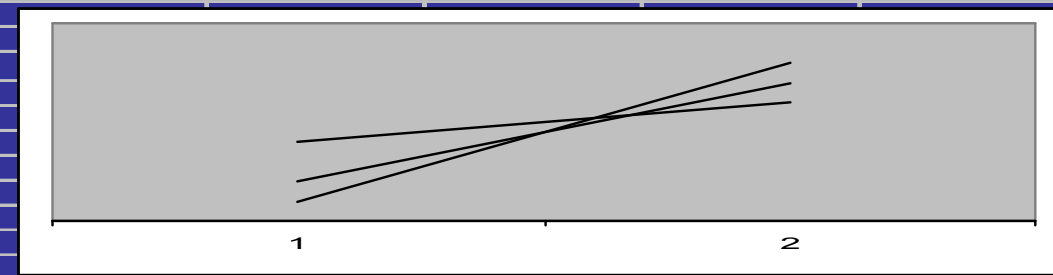
**Parallel and
no reaction norm**



Scale effects



True interaction



**Both scale and
true interaction**

- It is the "true interactions" that should affect breeding strategies.
- Scale effects can be handled by transformation prior to analysis to ensure homogeneity of among-genotype variances in environments.
- The question is whether breeding should be producing genotypes suitable for specific environments or genotypes adapted to a wide range of environments?
- $G \times E$ can be used in practice when interactions and the specific environments are well defined.
- The smaller $G \times E$, the fewer test sites are needed.

Calculations of G x E

1. **ANOVA** according to the simple model: $Y = G + E + G \cdot E$.

- ✓ The model assumes homogenous variances between sites. Scale effects (not true interactions) will generate an interaction!
- ✓ Not independent of whether environment is a fixed or random effect.

2. G x E as genetic correlations:

I. Yamada: $r_G = \text{var}_G / (\text{var}_G + \text{var}_I)$

var_G = genetic variance component from ANOVA involving data from two environments and var_I = G x E variance component from ANOVA.

II. Burdon: $r_G = r_{XY} / (h_X \cdot h_Y)$

r_{XY} = phenotypic correlation between family means in environment X and Y

$h_X \cdot h_Y$ = square roots of heritabilities of the genetic family means in environment X and Y.

III. GCA-approach: $r_G = r / (r_{ax} \cdot r_{ay})$

r = Pearson correlation between BLUP-values in environment X and Y

$r_{ax} \cdot r_{ay}$ = estimated relation between the “true” and the “predicted” breeding values calculated as $(h^2 \cdot k) / (1 + h^2(k-1))$ where k is the harmonic mean of the number of replications per family.

Type B correlations are routinely made by univariate methods

Problems:

- 1)** Correlation estimates are biased for unbalanced data and when variances across environments are heterogenous.
- 2)** The estimates are frequently out of the theoretical parameter space due to sampling errors of genetic variances and covariances ($r_G > 1.0$).
- 3)** The correlations are seldom normally distributed unless the test population is large. Std err of genetic correlations are difficult to estimate and are often approximated! Estimates of std err. should be interpreted with caution. However they indicate the relative reliability.

