Quantitative genetics

BIOL 434/509

Quantitative genetics

Quantitative genetics is the study of the evolution of phenotypic traits which do not have a simple genetic basis.

Most traits are affected by more than a few loci and also by the environment, and for most traits we do not know the full details of their genetic basis.

Quantitative genetics is a statistical approach.

Quantitative genetics allows prediction of evolutionary change in traits even without a locus-by-locus understanding of the traits' genetics.

Regression of offspring on parents

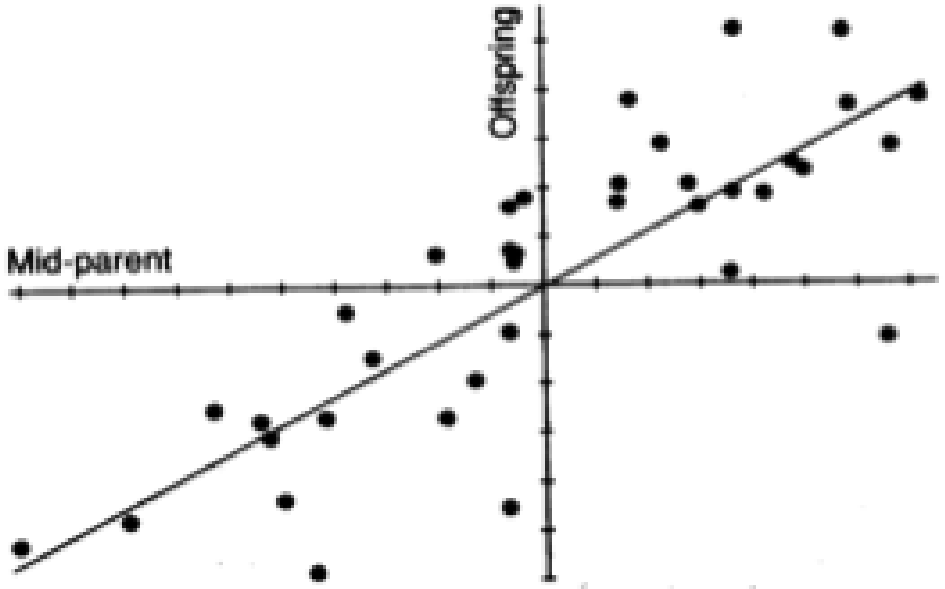
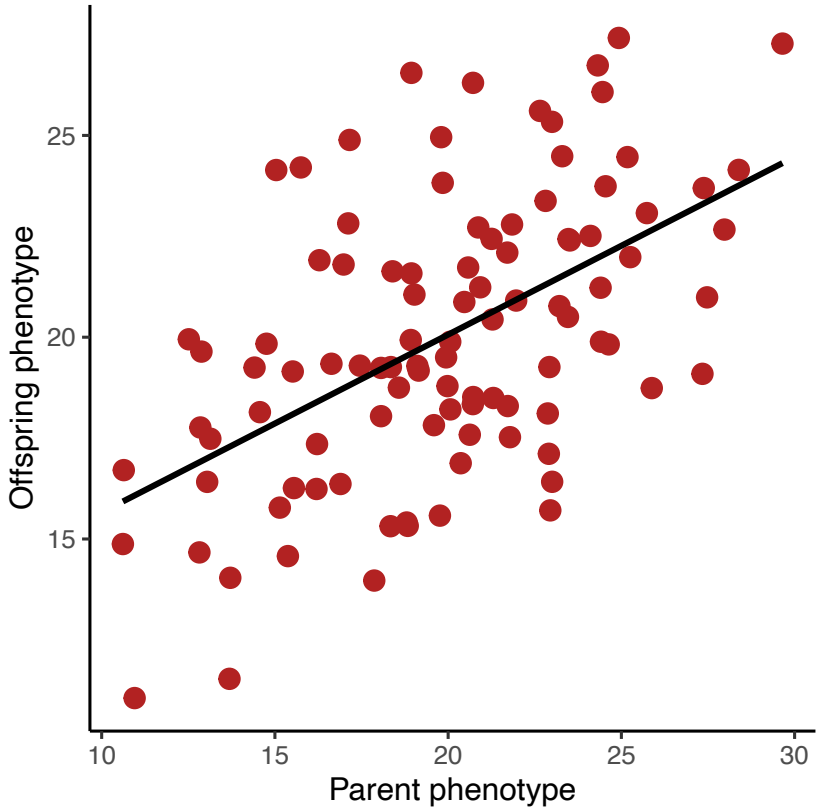


Fig. 10.1. Regression of offspring on mid-parent for wing length in *Drosophila*, as explained in Example 10.2. Mid-parent values are shown along the horizontal axis, and mean value of offspring along the vertical axis. (*Drawn from data kindly supplied by Dr E. C. R. Reeve.*)

Regression of offspring on parents



Regression

Regression predicts mean value of *Y* from value of *X*.

$$\beta = \frac{COV(X,Y)}{Var(X)}$$

Predicting the next generation

Regression is therefore exactly what we need to predict the next generation based on who reproduces.

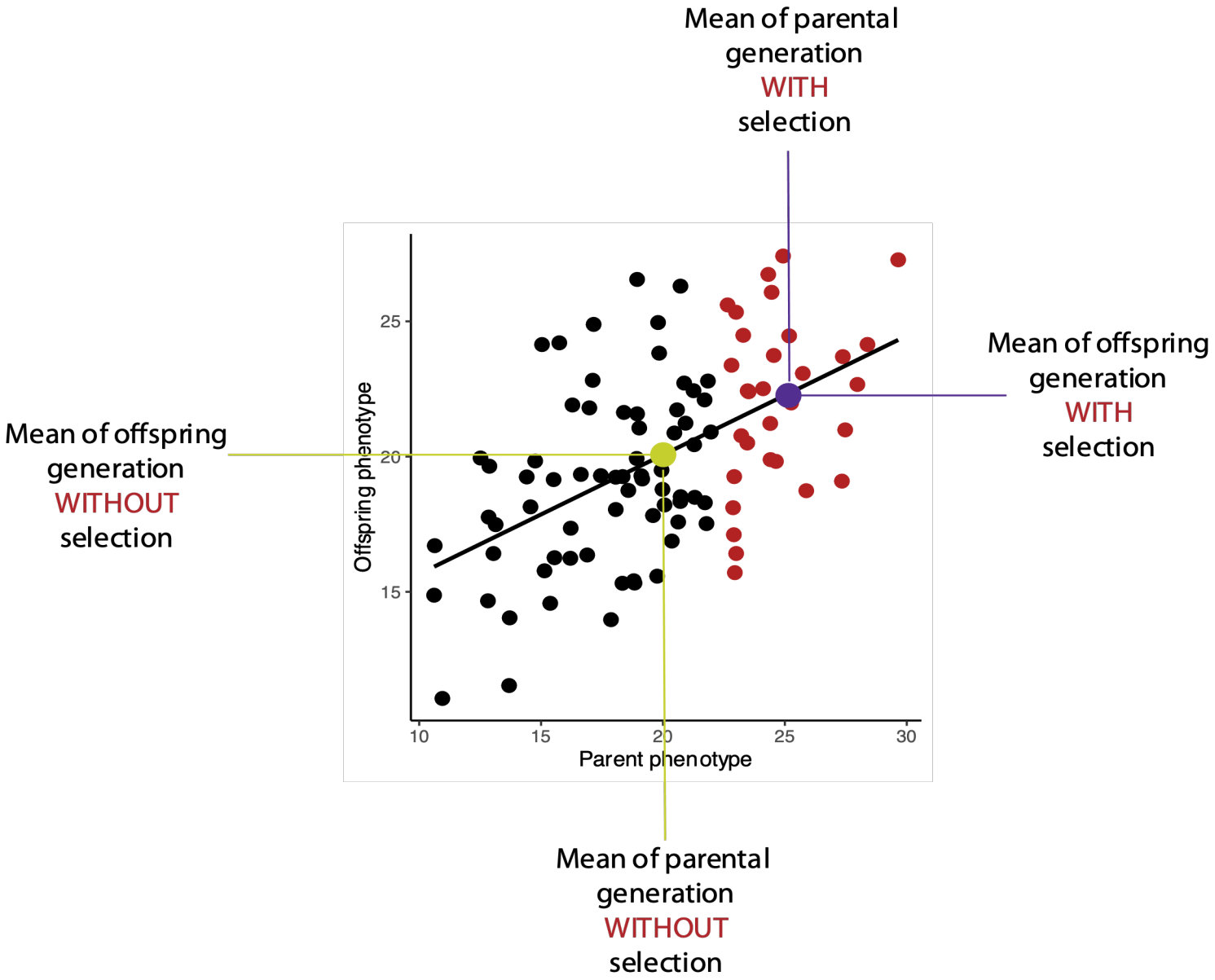
Terms

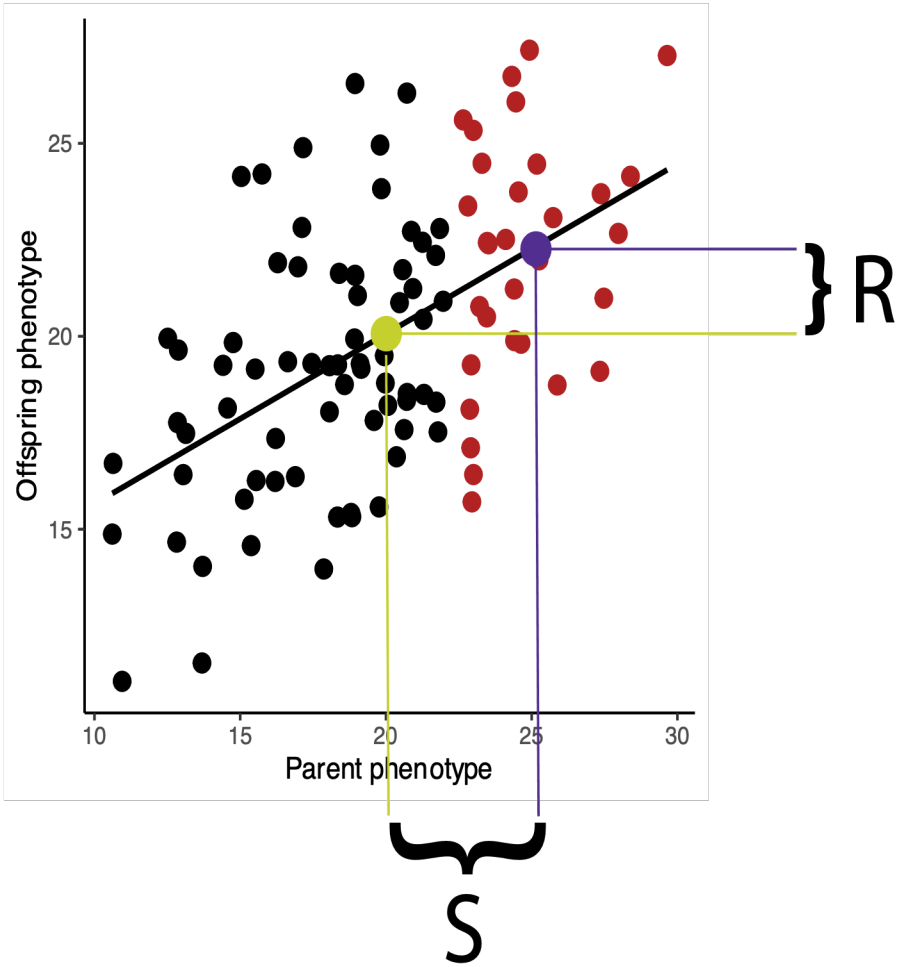
Response to selection is the change in the mean phenotype of a population caused by natural or artificial selection.

$$R = Mean(offspring\ generation) - Mean(parent's\ generation)$$

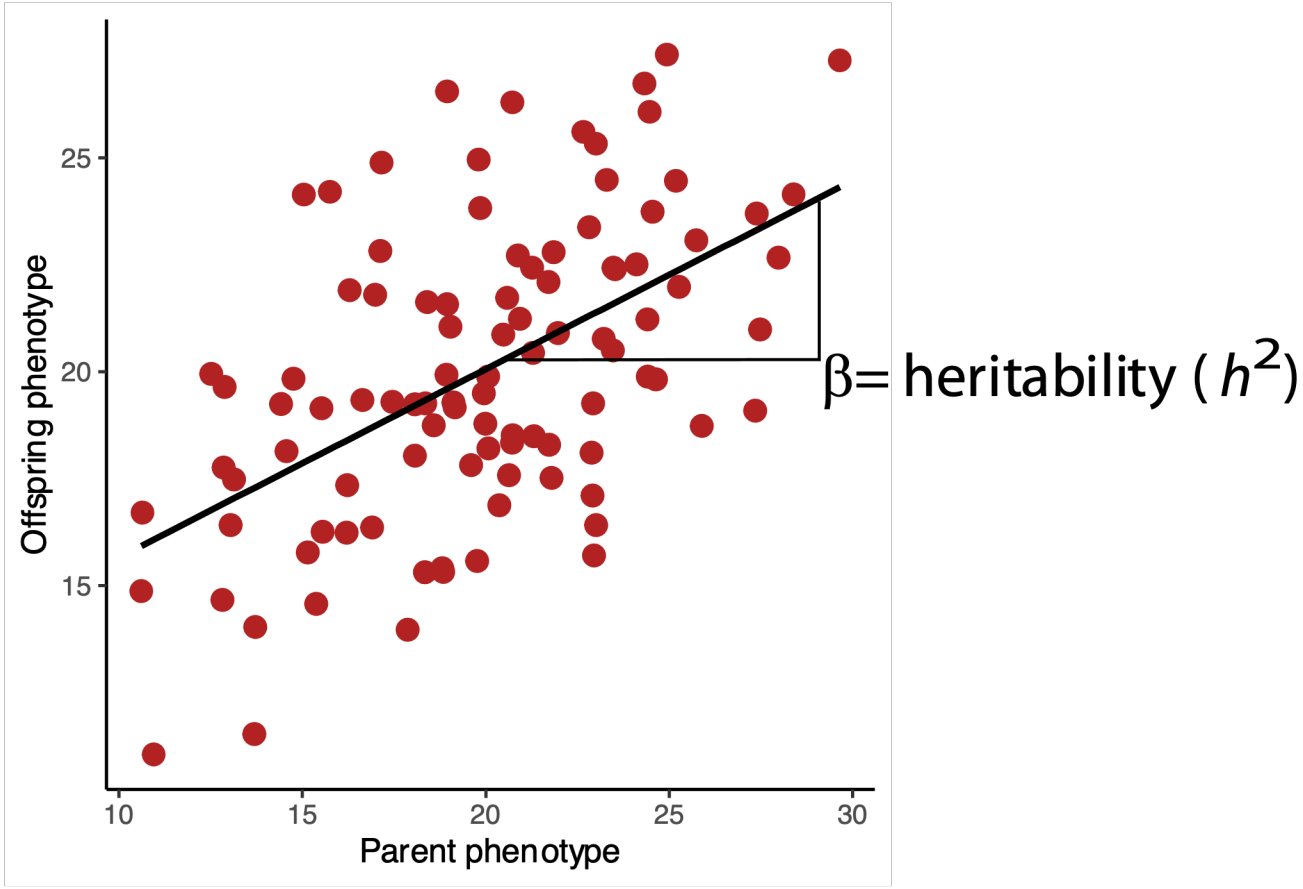
The selection differential is the difference between the mean phenotype of the successful parents and the mean phenotype of all members of the parents' generation:

$$S = Mean(successful\ parents' phenotype) - Mean(all\ phenotypes\ in\ that\ generation)$$


Mean of parental generation WITH selection
Mean of offspring generation WITH selection
Mean of offspring generation WITHOUT selection
Mean of parental generation WITHOUT selection
Offspring phenotype
Parent phenotype
25
20
15
10
15
20
25
30


25
20
15
Offspring phenotype
10
15
20
25
30
Parent phenotype
R
S

Predicting response: heritability



heritability

$$\beta = h^2$$
$$= \frac{Cov(Offspring\ phenotype, Midparent\ phenotype)}{Var(Midparent\ phenotype)}$$

The “Midparent” is the mean phenotype of the two parents of the offspring.

Why do relatives resemble each other?

- Relatives share some alleles (and so if those alleles on average cause a change in the phenotype, some of those effects are shared)
- Relatives may share patterns of alleles
  - All relatives share alleles at multiple loci, so they share the interactions of those loci.
  - Some relatives (e.g. full siblings) have some chance of sharing both alleles at a locus (and so share the effects of the dominance interactions of those alleles).
- Relatives may share environments. (Environments affect phenotypes, so these common environments affect resemblance of relatives)
- Relative sometimes share parents (and therefore share maternal effects and possibly epigenetic modifications of their DNA).

Average effects

The average effect of an allele is the difference between the phenotype of an individual who has that allele ($\bar{z}_{A?}$) and the mean phenotype of the population (μ).

$$\alpha_A = \bar{z}_{A?} - \mu$$

$$\alpha_a = \bar{z}_{a?} - \mu$$

Average effects

$$\bar{z}_{A?} = p\bar{z}_{AA} + q\bar{z}_{Aa}$$

because there is a *p* chance the other allele is an *A* and a *q* chance the other alleles is an *a*.

Average effects

With Hardy-Weinberg assumptions, the mean phenotype in this population is:

$$\mu = p^2 \bar{z}_{AA} + 2pq\bar{z}_{Aa} + q^2 \bar{z}_{aa}$$

where *p* is the frequency of the *A* allele.

Average effects

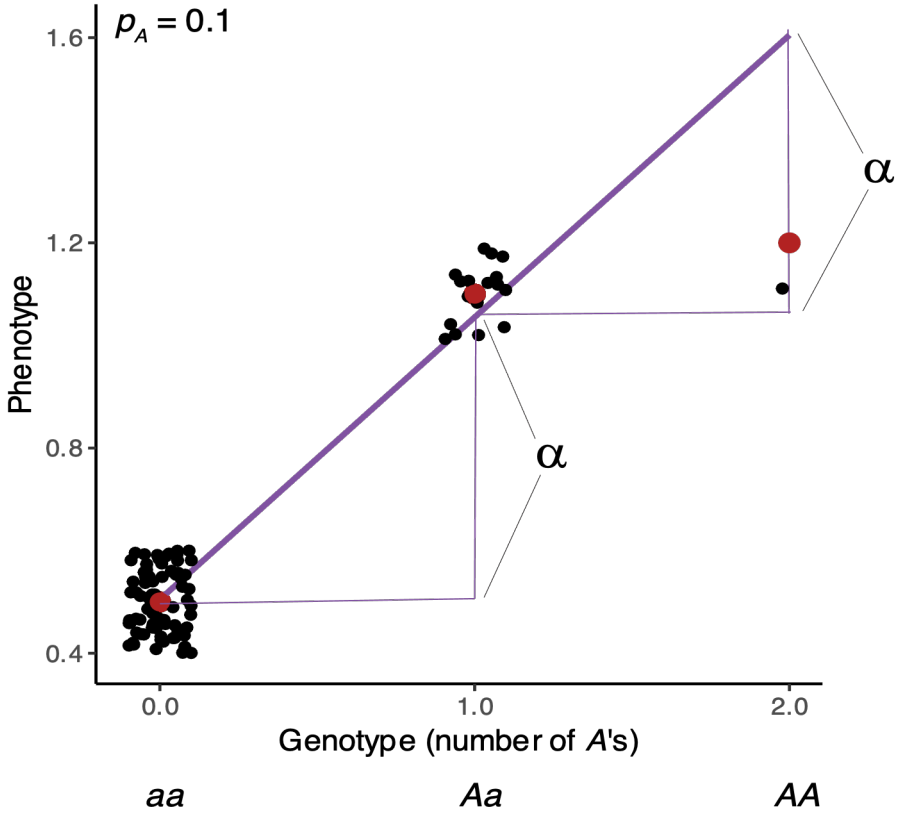
The average effect of a substitution (α) is the average change in phenotype caused by replacing one allele for another (i.e. replacing an *a* allele with an *A* allele).

$$\alpha = \alpha_A - \alpha_a \ \ (= \bar{z}_{A?} - \bar{z}_{a?})$$

The average effects capture the effect that an allele has on the phenotype, averaged over the distribution of possible genetic backgrounds.

Average effects: Regression interpretation

The average effect is also the slope of the regression of phenotype on the number of copies of one of the alleles.

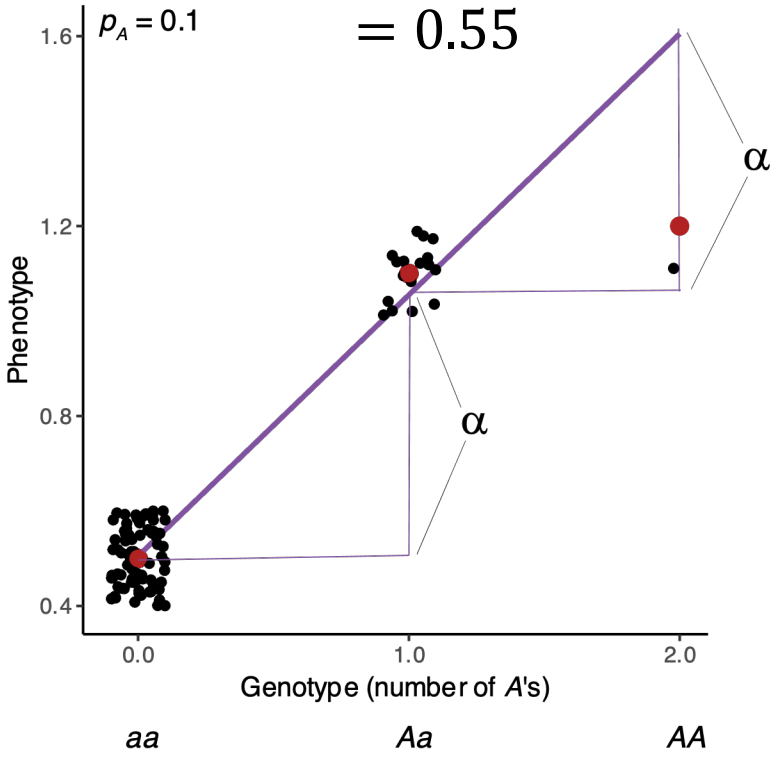


Example: Imagine the phenotypes of the three genotypes are 0.5, 1.1, and 1.2 units.

$$\bar{z}_{A?} = p_A(1.2) + p_a(1.1) \qquad \bar{z}_{a?} = p_A(1.1) + p_a(0.5)$$

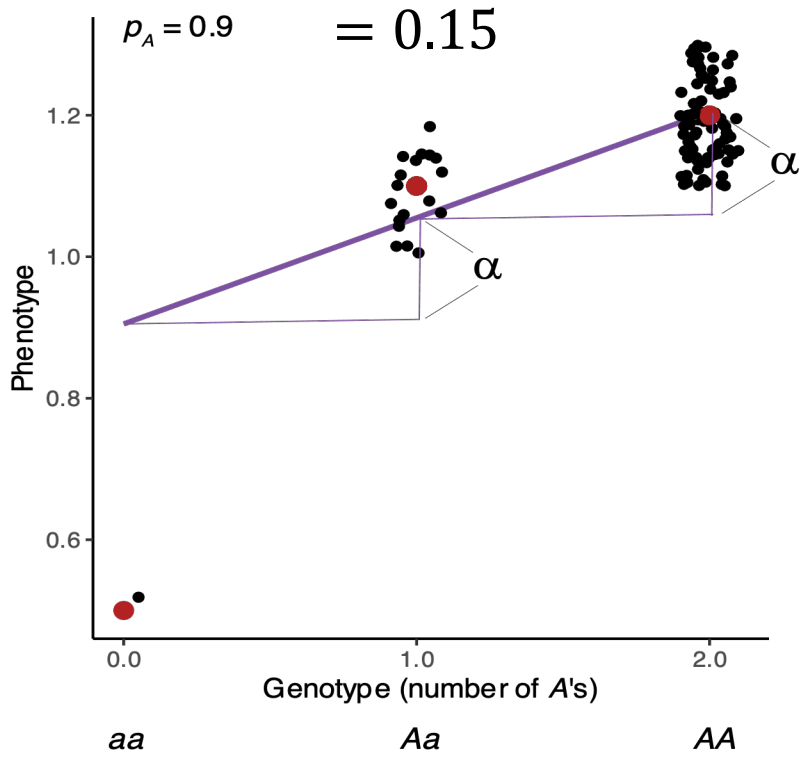
If the population allele frequency is $p_A = 0.1$, then

$$\alpha = (0.1)(1.2) + (0.9)(1.1) - \big((0.1)(1.1) + (0.9)(0.5)\big) = 0.55$$



If the population allele frequency is $p_A = 0.9$, then

$$\alpha = (0.9)(1.2) + (0.1)(1.1) - \big((0.9)(1.1) + (0.1)(0.5)\big) = 0.15$$



Breeding value

The breeding value of an individual is the sum of the average effects of all of its alleles, summed over all loci.

$$BV = \mu + \sum \left(\alpha_{i,1} + \alpha_{i,2} \right)$$

It is called the breeding value because these average effects determine the qualities of the offspring of that individual. The offspring of an individual will differ from the population mean by half of the breeding value of its parent, on average. (Because the offspring inherits half of its alleles for that parent.)

Additive genetic variance

The additive genetic variance for a trait (V_A) is the variance of the breeding values of the individuals in that population. It is a measure of the evolutionary potential of that population.

$$V_A = \sum 2p_i q_i \alpha_i^2$$

Phenotypic variance

The phenotypic variance of a trait (V_P) is simply the variance among individuals of the observed values of that trait.

Narrow sense heritability

The **narrow sense heritability** (h^2) is the ratio of the additive genetic variance over the phenotypic variance for a trait.

$$h^2 = \frac{V_A}{V_P}$$

This is the slope of the regression of offspring on the average of their parents.

Why additive effects matter: an example

Consider a plant whose height is affected by a single locus with overdominance, and the allele frequency is ½:

Genotype:	$A_1 A_1$	$A_1 A_2$	$A_2 A_2$
Frequency:	0.25	0.5	0.25
Height:	10 cm	12 cm	10 cm

Is there genetic variation?

What is the mean phenotype in the current generation?

Imagine there is very strong selection favoring taller plants: all the 10 cm plants die, and the survivors breed to produce the next generation. What will the mean height be in the next generation?

Decomposition of variance

The phenotypic variance can be broken down into parts:

$$V_P = V_G + V_E + V_{GxE}$$

where

V_P is the phenotypic variance

V_G is the genetic variance

V_E is the environmental variance

V_{GxE} is the variance due to interactions between environment and genotype

Decomposition of genetic variance

The genetic variance can be further decomposed:

$$V_G = V_A + V_D + V_I$$

where

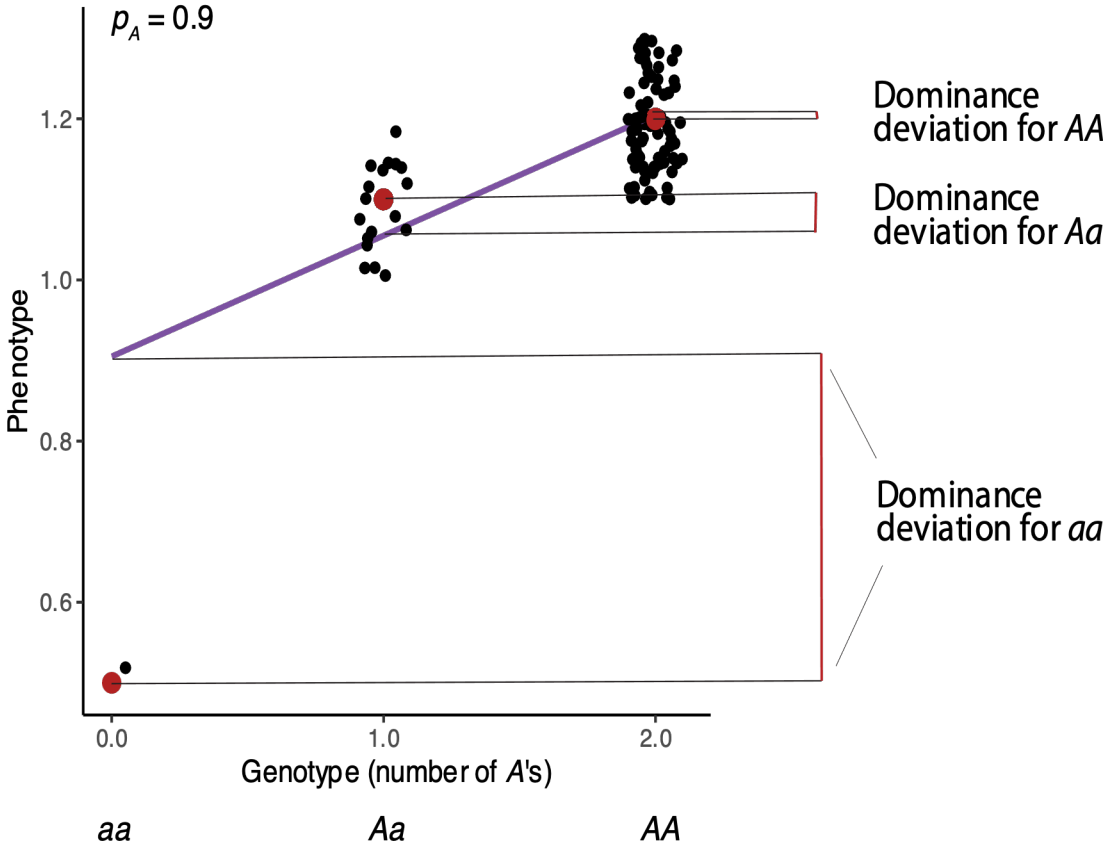
V_A is the additive genetic variance

V_D is the dominance genetic variance

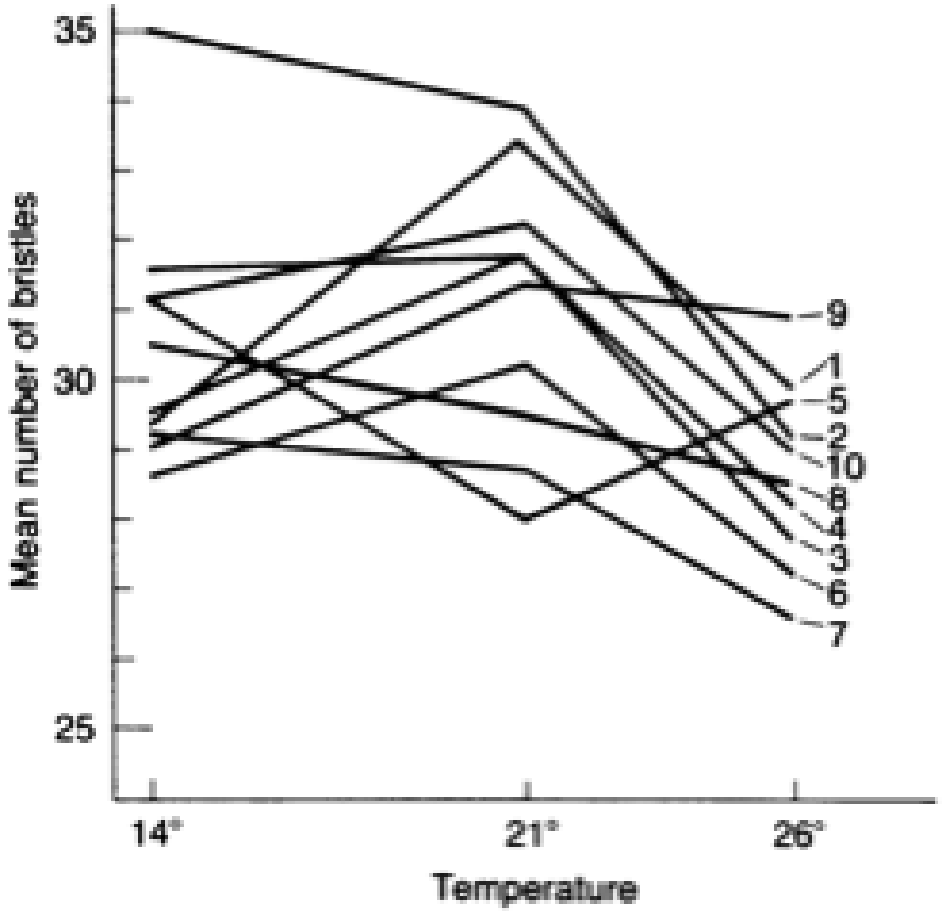
V_I is the epistatic variance (due to interactions of effects at different loci)

Dominance variance

The dominance variance describes variation in phenotypes caused by the effects of single loci, after removing the additive effects of each allele.



Genotype by environment interactions



Response to selection: The Breeder's equation

$$R = h^2 S$$

This is sometimes expressed in the evolution literature as

$$\Delta\bar{z} = V_A \beta$$

where $\Delta\bar{z}$ is the change in mean phenotype caused by selection ($\Delta\bar{z} = R$) and β is the slope of the regression predicting fitness from the phenotype of an individual (called the **selection gradient**).

Response to selection:

$$\Delta\bar{z} = V_A\beta$$

is probably better because it is probably more likely that the functional relationship between phenotype and fitness is not changed by changes in V_P.

Pace of evolution is determined more by V_A than by h^2.

How to estimate h^2?

Resemblance between relatives

Response to selection

Estimating h^2 from response to selection

The response to selection can also be used to estimate heritability. If the selection differential and the change in mean phenotype are known, we can solve for the heritability.

$$h^2 = \frac{R}{S}$$

Resemblance of relatives

All classes of relative may covary in their phenotype because of partially shared genotypes.

(Also relatives may share environments, which complicates things.)

One parent and offspring

The covariance between the phenotype is a parent and its offspring is

$$COV(Parent, offspring) = \frac{V_A}{2}$$

assuming that they share no features of the environment.

An individual parent and one of its offspring share half of their alleles, and they never have identity by descent at both alleles at a locus. Therefore the share no dominance effects, and they share half of their additive effects. Hence the covariance between them is half the additive genetic variance.

One parent and offspring

$$COV(Parent, offspring) = \frac{V_A}{2}$$

$$\hat{V}_A = 2COV(Parent, offspring)$$

Regression

The slope of the regression of offspring phenotype on the phenotype of one of its parents is given by:

$$\beta_{offspring,parent} =$$

$$\frac{COV(Parent, offspring)}{Var(parent)} = \frac{V_A}{2V_P} = \frac{h^2}{2}$$

Half-siblings

Half-sibling share one parent (but not the other).

An allele in one half sib has a ¼ chance of being identical by descent in the other half sib, and there is no chance that they have both alleles at a locus that are IBD.

The covariance between half sibs is therefore

$$COV(half-sib, half-sib) = \frac{V_A}{4}$$

Half-siblings sharing a mother

$$COV(half-sib, half-sib) = \frac{V_A}{4} + Cov_{maternal\ effects}$$

Shared effects on phenotype due to maternal effects are much more common than paternal effects. Using maternal half-sibs can therefore bias estimates of V_A unless estimates of the maternal effects are known.

Half-siblings sharing a father

$$COV(half-sib, half-sib)$$
$$= \frac{V_A}{4} + Cov_{paternal\ effects}$$

Shared effects on phenotype due to maternal effects are much less common than maternal effects.

Full-siblings

Full siblings share both parents.

- share on average half of their alleles as identical by descent.
- share both alleles at the same locus causing them to also share some dominance effects.

$$COV(sib, sib) = \frac{V_A}{2} + \frac{V_D}{4} + Cov_{maternal\ effects}$$

Full siblings can produce biased estimates of V_A. They can be used in conjunction with half-sibs, however, to also estimate V_D.

Identical twins

Identical twins (monozygotic twins) share their genomes in their entirety. Therefore the covariance between monozygotic twins is

$$COV(monozygotic\ twins) = V_G + Cov_{maternal\ effects}$$

Identical twins usually share a great deal of their environment, causing the simple covariance between them to exaggerate the genetic effects.

Identical twins

Estimates of heritability based on identical twins is not an estimate of the **narrow sense heritability** ($h^2 = V_A/V_P$) but of what is called the broad-sense heritability ($H^2 = V_G/V_P$).

General equation

r is the relatedness between two individuals (the probability that a randomly chosen allele in one individual is present in an identical by descent copy in the other)

r_2 is the probability that both alleles at a locus are identical by descent in the two individuals

$$COV(relatives) = rV_A + r_2V_D + \cdots$$

Terms after … are due to epistatic effects.

Challenges in estimating V_A

Common environmental effects must be carefully accounted for in experimental design (i.e. by randomizing the housing containers an location of each individual.)

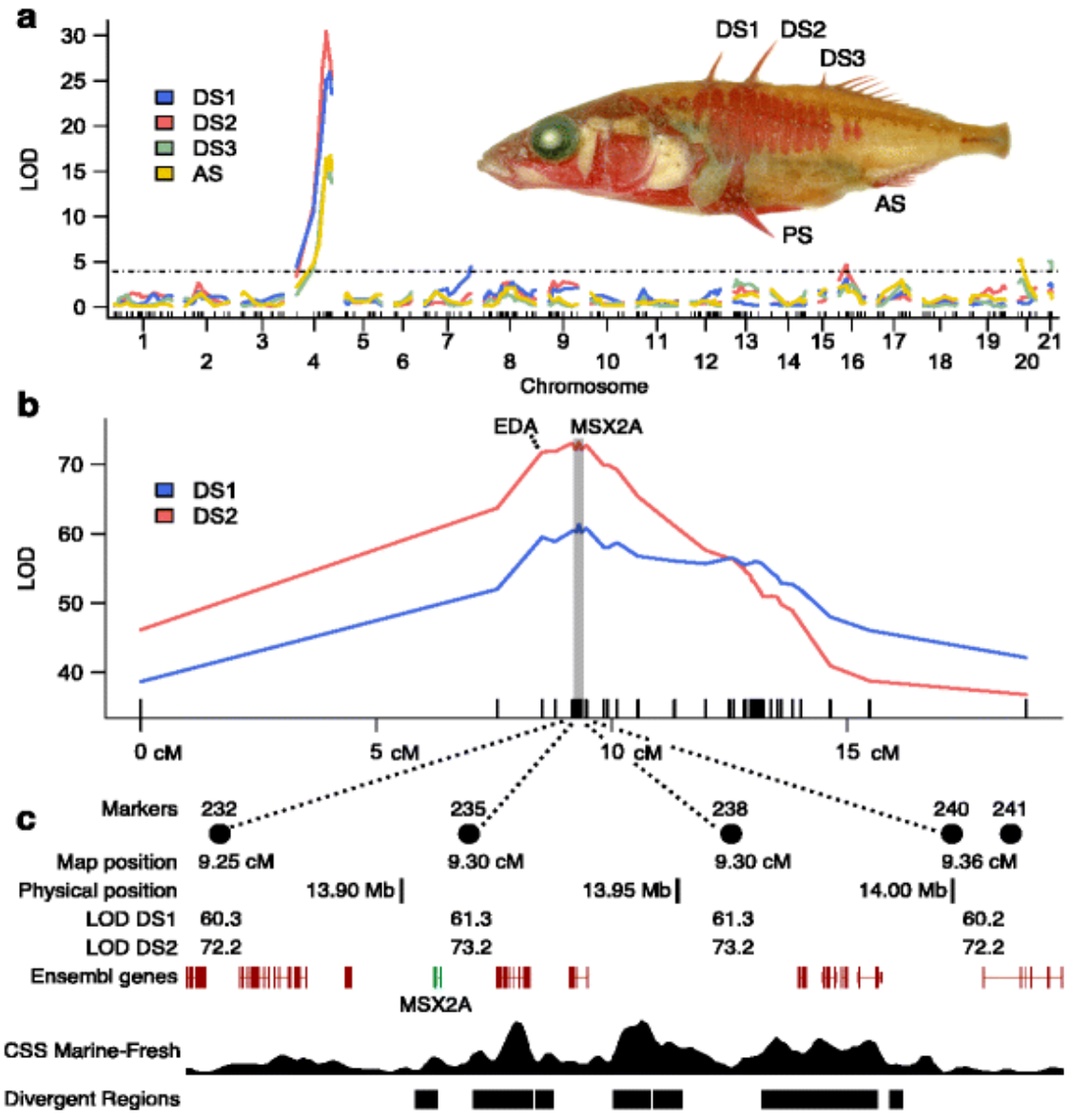
Maternal effects can be large.

Measurement and statistical error can be very large in these studies.

QTL

Quantitative Trait Loci

(QTL) are loci identified to play a role in the genetic variation for a trait. QTL are identified by an association between particular genotypes and the phenotype in the F2 or later offspring of a cross between genetically (and phenotypically) diverged individuals.



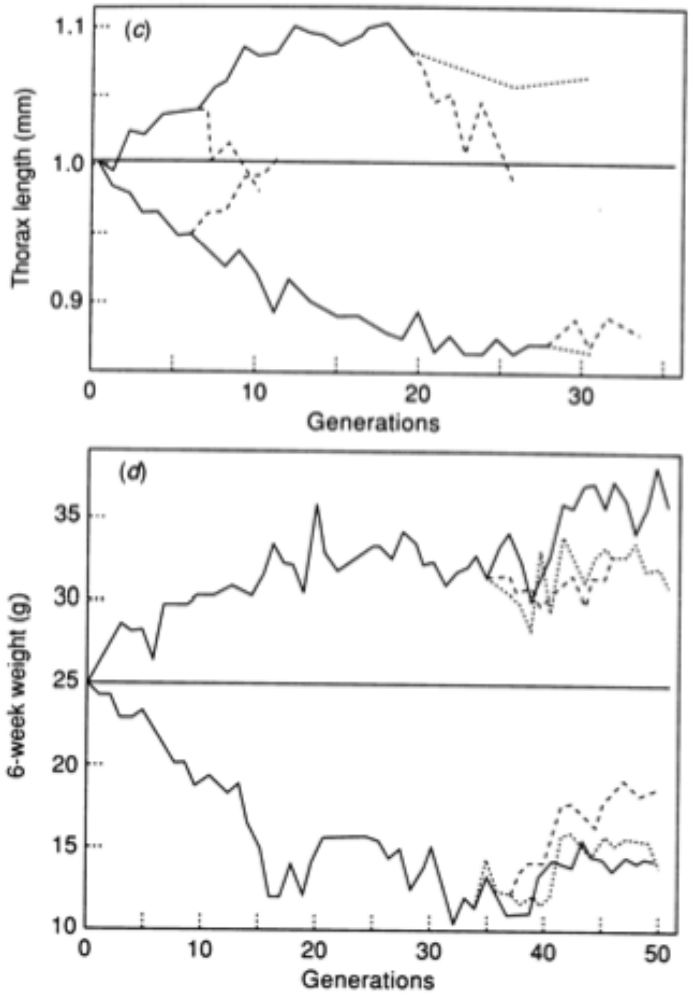
Howes et al., *BMC Biology* **volume 15**, Article number: 115 (2017)

Long-term response to selection

Quantitative genetics predicts response to selection in the short term.

Only if V_A and β are unchanged over time is it possible to make longer term predictions.

Lab experiments frequently detect limits to selection response.



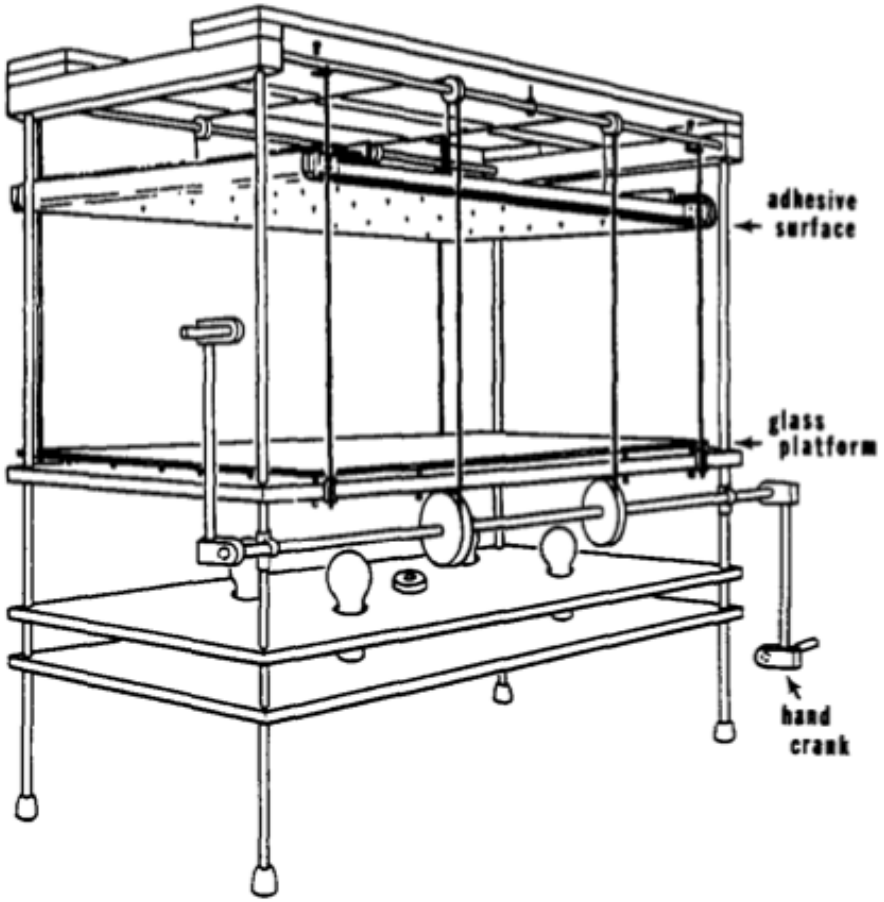
(*c*) *Drosophila melanogaster*, thorax length. (*After F. W. Robertson, 1955.*)
(*d*) Mouse, six-week body weight. (*Adapted from Roberts, 1966b.*)
Dashed lines are responses to selection in the reverse direction; dotted lines are responses t
natural selection, with artificial selection suspended.
(*All figures redrawn from the above sources with permission of the authors and publishers.*)

Long-term response to selection

Such limits could be due to:

- Genetic drift
- Lack of replication
- Changing environmental conditions
- Depletion of genetic variance due to selection
- Loss of genetic variance due to drift
- Logical or physical constraints

Test for effect of small population size



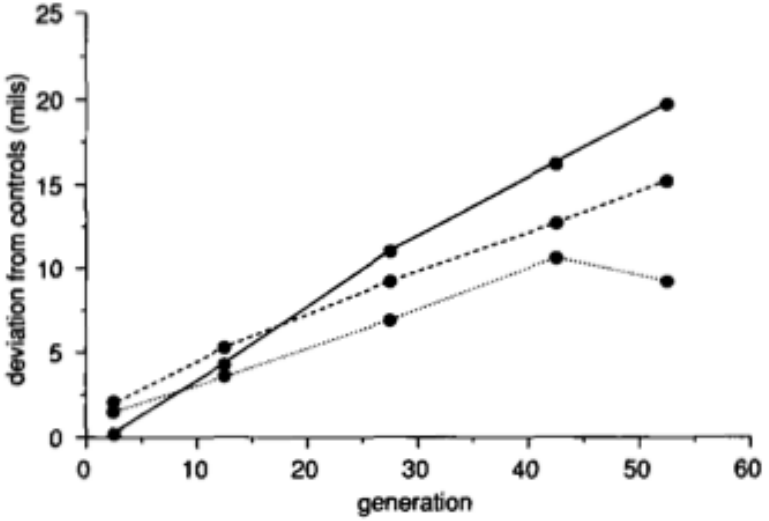


FIGURE 2.—Mean deviation from controls within treatments, averaged over successive periods of 5, 15, 15, 15 and 5 generations. Points are at midpoint of each period. Large lines (——); medium lines (- - - - -); small lines (· · · · ·).

Quantitative genetics with multiple traits

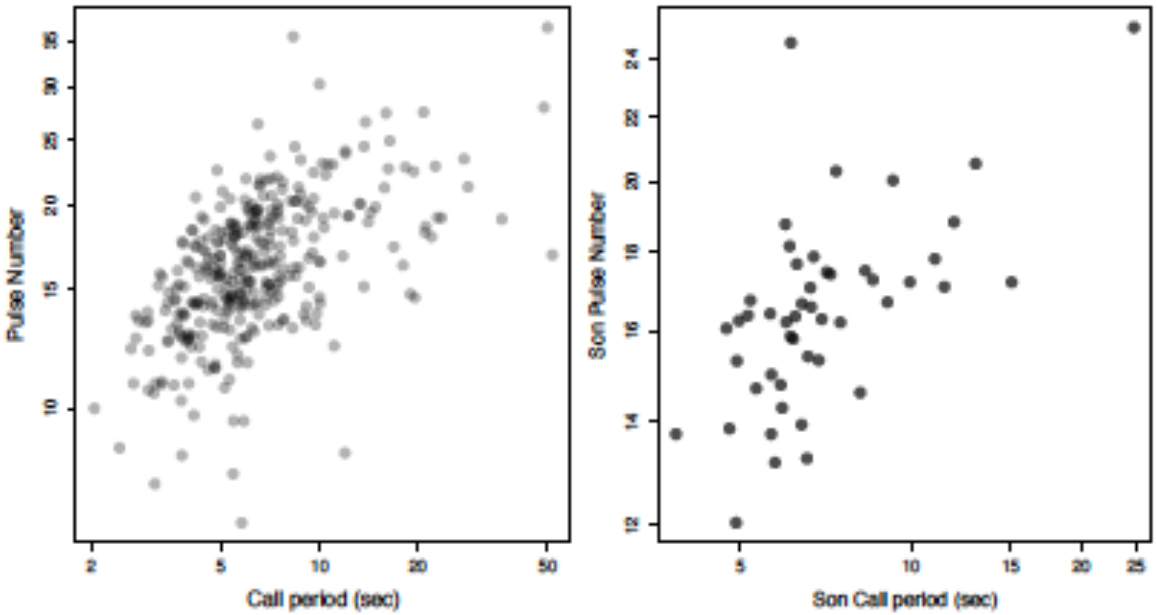


Figure 7.17: Phenotypic and Genetic correlations in male grey treeforg (*Hyla versicolor*) calls. On the left each male is shown as a dot, recording their inter-call period and the number of pulses in each call. One the right each dot corresponds to a father with the mean of sons for both phenotypes. Data from WELCH *et al.* (2014) downloaded from dryad, Code here.

Genetic covariance

With genetic covariance, individuals covary in multiple traits because genetic effects on these traits are associated.

Genetic covariance can be due to linkage disequilibrium (between an allele which affects one trait and another allele that affects the other trait) or pleiotropy (when alleles have effects on both traits).

Genetic covariance

With genetic covariance, selection on one trait can have an indirect effect on the mean of the other trait.

The additive genetic covariance is defined just like the additive genetic variance, except one trait is measured in one relative and the other trait is measured in the other.

Modeling selection on multiple traits

An indirect response to selection is a change in the mean phenotype of one trait caused by selection on a genetically correlated trait.

For example, selection for longer call period in the tree frogs would be expected to result in a higher pulse number.

Modeling selection on multiple traits

Let $COV_{A,1,2}$ be the additive genetic covariance between trait 1 and trait 2.

The indirect response to selection on trait 1 from selection on trait 2 is given by:

$$\Delta\bar{z}_1 = COV_{A,1,2}\beta_2$$

Combining the direct response to selection and the indirect response to selection we get:

$$\Delta\bar{z}_1 = V_A\beta_1 + COV_{A,1,2}\beta_2$$

Modeling selection on multiple traits

Expanded to consider more than two traits:

$$\Delta\bar{z}_1 = V_A\beta_1 + COV_{A,1,2}\beta_2 + COV_{A,1,3}\beta_3 + COV_{A,1,4}\beta_4 + \cdots$$

This is often captured in linear algebra, using G (the additive genetic variance-covariance matrix), $\boldsymbol{\Delta\bar{z}}$ (the vector of changes in mean trait for a list of all traits considered), and β (the vector of selection gradients for each trait).

$$\boldsymbol{G} = \begin{bmatrix} V_{A,1} & COV_{A,1,2} \\ COV_{A,1,2} & V_{A,2} \end{bmatrix}$$

$$\boldsymbol{\Delta\bar{z}} = \boldsymbol{G\beta}$$