

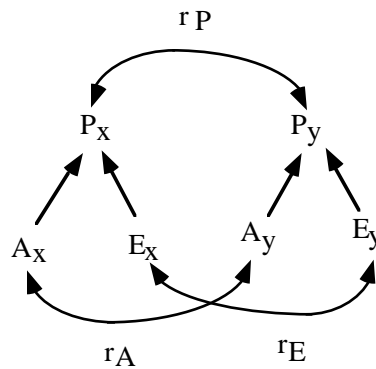
Lecture 7

Correlated Characters

Genetic and Environmental Correlations

Many characters are positively or negatively correlated at the level of phenotype (e.g., height and weight, forearm length and digit span) and we can directly measure the phenotypic correlation, r_P , between two traits X and Y .

As the following (path) diagram indicates, the phenotypic correlation r_P between two traits is generated by correlations between the genetic (r_A) and/or environmental (r_E) values of X and Y . In the figure, double-headed arrows imply possible correlations between variables.



r_A = correlation of breeding values arises from two sources

- pleiotropic effects of loci on both traits; correlation from pleiotropy indicates the extent to which the character is controlled by the same genes
- linkage disequilibrium, which will decay over time

r_E = correlation of environmental deviations

- includes non-additive genetic effects
- arises from exposure of the two traits to the same individual environment

Recall (Lecture 3) that the correlation between X and Y equals $r = cov(X, Y) / \sigma_X \sigma_Y$. Rearranging, we can express the covariance as

$$cov(X, Y) = r \sigma_X \sigma_Y$$

Therefore

$$cov_P = r_P \sigma_{XP} \sigma_{YP}, \quad cov_A = r_A \sigma_{XA} \sigma_{YA}, \quad cov_E = r_E \sigma_{XE} \sigma_{YE}$$

The covariance of phenotypic values = the sum of the genetic and environmental covariances:

$$cov_P = cov_A + cov_E$$

so that

$$r_P \sigma_{XP} \sigma_{YP} = r_A \sigma_{XA} \sigma_{YA} + r_E \sigma_{XE} \sigma_{YE} \tag{1}$$

describes the relationship between phenotypic, genetic and environmental correlations.

This expression can be simplified somewhat. First, define h_X and h_Y as the square roots of the heritabilities of characters X and Y , with

$$h = \frac{\sigma_A}{\sigma_P}, \quad \sigma_A = h\sigma_P$$

Likewise define $e^2 = 1 - h^2$

$$e^2 = \frac{\sigma_E^2}{\sigma_P^2}, \quad e = \sqrt{1 - h^2} = \frac{\sigma_E}{\sigma_P}, \quad \sigma_E = e\sigma_P$$

Substituting $\sigma_A = h\sigma_P$ and $\sigma_E = e\sigma_P$ into equation (1) gives

$$\begin{aligned} r_P &= h_X h_Y r_A + e_X e_Y r_E \\ &= r_A h_X h_Y + r_E \sqrt{(1 - h_X^2)(1 - h_Y^2)} \end{aligned}$$

Hence, the phenotypic correlation is a function of the heritabilities of the traits and the genetic and environmental correlations. If heritabilities are high, the genetic correlation is more important; if heritabilities are low, the environmental correlation is more important. In practice, phenotypic and genetic correlations often have the same sign and are of similar magnitude, but this is not always the case.

Estimating the Genetic Correlation

Methods for estimating r_A and r_E are analogous to estimating heritabilities from resemblance between relatives. The difference is that with a single trait, the covariance between the trait value in two relatives provides an estimate of the additive genetic variance of that trait, while with two traits, the covariance of trait X in one relative and trait Y in the other provides an estimate of the additive genetic covariance between the two traits.

Half sibs

The covariance of traits X and Y between sires = $cov_{XY} = (1/4)cov_A$

The variance between sires of trait $X = \sigma_{SX}^2 = (1/4)\sigma_{AX}^2$

The variance between sire of trait $Y = \sigma_{SY}^2 = (1/4)\sigma_{AY}^2$

Therefore, $r_A = cov_{XY} / (\sigma_{SX}\sigma_{SY})$

Offspring-parent

The covariance of trait X in the offspring with trait Y of the parents

$$cov(X_O, Y_P) = (1/2)cov_{A_X, A_Y}$$

The covariance of trait X in the offspring with trait X of the parents

$$cov(X_O, X_P) = (1/2)\sigma_{A_X}^2$$

The covariance of trait Y in the offspring with trait Y of the parents

$$cov(Y_O, Y_P) = (1/2)\sigma_{A_Y}^2$$

Hence,

$$r_A = \frac{cov(X_O, Y_P)}{\sqrt{cov(X_O, X_P) \cdot cov(Y_O, Y_P)}}$$

Estimates of genetic correlation have very large sampling errors, typically requiring extremely large experiments for precise estimates.

Genotype-environment interaction and the cross-environment genetic correlation

If a character is measured in two environments, it can be considered as two characters (e.g., X in environment one, Y in environment two). By rearing families in both environments, it is possible to estimate the genetic correlation between these "characters". The magnitude of this cross-environment correlation reflects the extent to which the same genes control the character in each environment.

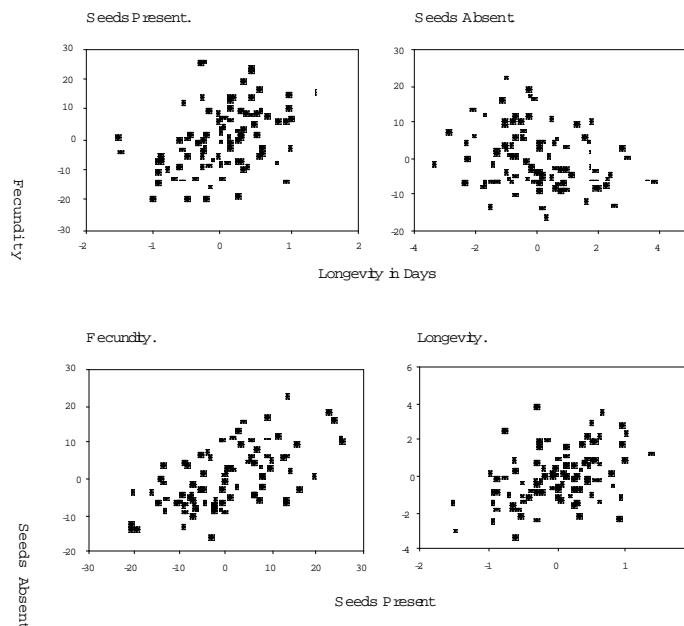
Example: Genetic correlations within and between environments in a seed beetle (data courtesy of F. Messina)

Females from each of 94 half-sib families of seed beetles (*Callosobruchus maculatus*) were placed in petri dishes which either contained or lacked seeds. The total number of eggs laid and longevity in days were recorded for each beetle ($N = 4,408$). From the data, four genetic correlations can be estimated: the correlations between fecundity and longevity within each treatment, and the correlations of longevity and fecundity across treatments.

Trait 1	Trait 2	Genetic correlation	S.E.
Fecundity, seeds present	Longevity, seeds present	0.35*	(0.16)
Fecundity, seeds absent	Longevity, seeds absent	-0.44*	(0.19)
Fecundity, seeds present	Fecundity, seeds absent	0.76***	(0.10)
Longevity, seeds present	Longevity, seeds absent	0.44*	(0.18)

(* $P < 0.05$; *** $P < 0.001$)

The scatter plots below show estimated breeding values for each of the 94 sires. These and the above genetic correlation estimates were obtained by REML using SAS (Messina and Fry, unpublished).



Fecundity and longevity showed a significant positive genetic correlation when seeds were present, but the correlation was negative when seeds were absent. Thus, these genetic correlations changed over environments.

Correlated Response to Selection

When characters are genetically correlated, selection solely on one will result in a correlated change in the second. Such a change in the unselected character is called a **correlated response**.

We compute the expected correlated response in Y given selection on X as follows. The response to selection of character X (the mean value of offspring of selected parents) is (by definition) the mean breeding value of the selected group. Thus the change in character Y in response to selection on X is the regression of the breeding value of Y on the breeding value of X . The slope of this regression is given by

$$b_{A_Y|A_X} = \frac{cov_A}{\sigma_{A_X}^2} = \frac{r_A \sigma_{A_X} \sigma_{A_Y}}{\sigma_{A_X}^2} = r_A \frac{\sigma_{A_Y}}{\sigma_{A_X}}$$

Recalling first that a regression passes through the mean of both variables (with $y - \mu_y = b_{y|x} [x - \mu_x]$) and second that the breeding values have mean zero ($\mu_{A_X} = \mu_{A_Y} = 0$), the regression of the breeding values of Y on the breeding values of X is just

$$Y = b_{A_Y|A_X} X = r_A \frac{\sigma_{A_Y}}{\sigma_{A_X}} X$$

The response of the directly selected character X is

$$R_X = i_X h_X^2 \sigma_{P_X} = i (\sigma_{A_X}^2 / \sigma_{P_X}^2) \sigma_{P_X} = i_X \sigma_{A_X}^2 / \sigma_{P_X} = i_X h_X \sigma_{A_X}$$

where $i_X = S_x / \sigma_X$. Recalling our comment above that R_X is the change in the breeding value of X , the correlated response of character Y is

$$\begin{aligned} CR_Y &= b_{A_Y|A_X} R_X \\ &= (r_A \sigma_{A_Y} / \sigma_{A_X}) (i_X h_X \sigma_{A_X}) \\ &= r_A \sigma_{A_Y} i_X h_X \end{aligned}$$

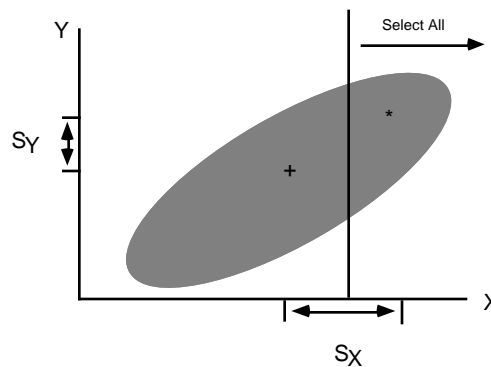
Substituting $\sigma_{A_Y} = h_Y \sigma_{P_Y}$ gives

$$CR_Y = i_X h_X h_Y r_A \sigma_{P_Y}$$

Noting that the direct response on X is $R_X = i_X h_X^2 \sigma_{P_X}$, we see that h_X^2 and $h_X h_Y r_A$ play similar roles, resulting in the later being called the **co-heritability**.

Correlated Selection Differentials

When selection is applied to character X , character Y will show a correlated selection differential when Y is phenotypically correlated with X . The figure below shows the change in the bivariate mean before (+) and after (*) selection. In this case, there is truncation selection only on trait X , but there is a correlated within-generation change in Y , with the mean of Y in the selected parents differing from the mean before selection.



Note that the correlated selection differential simply measures the *within-generation* change. Whether this translates into a *between-generation* change (i.e., a response to selection) depends on whether there is any genetic correlation between the characters. With only phenotypic correlations, none of the correlated within-generation change is passed on to the offspring.

The reason for this can be illustrated by a hypothetical example, with truncation selection only on X . Suppose that $r_A = 0$ and $r_E > 0$; therefore r_P and S_Y will be > 0 . The selected individuals, however, will not have higher than average breeding values for Y , in spite of their higher than average phenotypic values. Therefore the correlated response will be zero, regardless of the magnitude of h_Y^2 and S_Y .

Estimating the Genetic Correlation from Selection Response

Another method for estimating the genetic correlation is analogous to the realized heritability. Recall in the latter case, heritability is estimated by $h^2 = R/S$.

One procedure to obtain a realized estimate of the additive genetic correlation is as follows: From the same base population, in different lines:

- select individuals on the basis of character X . From these lines you can measure the direct response to character $X (R_X)$ and the correlated response of character $Y (CR_Y)$.
- select individuals on the basis of character Y . From these lines you can measure the direct response to character $Y (R_Y)$ and the correlated response of character $X (CR_X)$.

From the identities above, note that

$$r_A^2 = \frac{CR_X}{R_X} \frac{CR_Y}{R_Y}$$

Similarly, to obtain a realized estimate of the genetic covariance between a trait in two different environments,

- select individuals for character X in environment 1, and measure response in environment 1 (R_1) and environment 2 (CR_2)
- select individuals for character X in environment 2, and measure response in environment 2 (R_2) and environment 1 (CR_1)

The again provides an estimate of the genetic covariance of the trait in the two environments as

$$r_{GE}^2 = \frac{CR_1}{R_1} \frac{CR_2}{R_2}$$

Example: Computing the genetic correlation from a double selection experiment

The experiment is as follows: Select for increased and decreased abdominal bristle number, and increased and decreased sternopleural bristle number, from the same base population of *Drosophila melanogaster*.

- Large base population (the Raleigh population of the realized h^2 example); select for high and low values of each trait; 25 selected out of 100 scored per sex; selection continued for 25 generations.

Results: At generation 25, the mean abdominal (AB) and sternopleural (ST) bristle numbers in the four selection lines were:

Selection Line	Mean Bristle Number	
	AB	ST
High AB	33.4	26.4
Low AB	2.4	12.8
High ST	22.2	45.0
Low ST	11.1	9.5

Hence, selection to increase AB gives a direct response in AB of 33.4 and a correlated response in ST of 26.4. Expressing these responses in terms of divergence selection, we have

- Response in AB = $R_{AB} = 33.4 - 2.4 = 31.0$
- Response in ST = $R_{ST} = 45.0 - 9.5 = 35.5$
- Correlated response in ST = $CR_{ST} = 26.4 - 12.8 = 13.6$
- Correlated response in AB is $CR_{AB} = 22.2 - 11.1 = 11.1$
- The estimated genetic correlation of abdominal and sternopleural bristle number is

$$r_A = \sqrt{\frac{CR_{AB}}{R_{AB}} \frac{CR_{ST}}{R_{ST}}} = \sqrt{\frac{11.1}{31} \frac{13.6}{35.3}} = 0.37$$

The positive genetic correlation between the two bristle traits may be due to linkage disequilibrium or pleiotropy. Not all loci affecting the trait necessarily have the same pleiotropic effects; one could conceive of a situation in which all loci were pleiotropic but r_A is zero, if pleiotropic effects are not directional across loci.

Indirect Selection

There are two ways the mean of a character X can change by selection:

- as a direct response to selection for trait X (R_X)
- as a correlated, or indirect, response to direct selection for trait Y (CR_X)

The relative magnitudes of the change in mean is given by the ratio

$$\frac{CR_X}{R_X} = \frac{i_Y r_A \sigma_{AX} h_Y}{i_X h_X \sigma_{AX}} = \frac{i_Y r_A h_Y}{i_X h_X}$$

Therefore the correlated response of X to selection for Y will be greater than direct response to selection for X when $i_Y r_A h_Y > i_X h_X$, or when

- character Y has a greater heritability than X , and the genetic correlation between X and Y is high. This could occur if X is difficult to measure with precision but Y is not.
- the selection intensity is much greater for Y than X . This would be true if Y were measurable in both sexes but X measurable in only one sex.

Multitrait Selection Response in Matrix Form

The response to selection of several traits is best handled using matrix notation, which we briefly introduce here. A **matrix** is an array of elements, e.g.

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} e & f \\ g & h \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} i \\ j \end{pmatrix}$$

When the matrix consists of only a single row or a single column, it is called a **vector**, whereas **A** and **B** are **square matrices**. **C** above is a column vector (consisting of a single column of entries).

Matrix multiplication is defined in such a fashion that the order of multiplication of the matrices is critical. For the above matrices,

$$\mathbf{AB} = \begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}, \quad \mathbf{BA} = \begin{pmatrix} ae + cf & eb + df \\ ga + ch & gd + dh \end{pmatrix}$$

and

$$\mathbf{AC} = \begin{pmatrix} ai + bj \\ ci + dj \end{pmatrix}, \quad \mathbf{BC} = \begin{pmatrix} ei + fj \\ gi + hj \end{pmatrix}$$

while the matrix products **CA** and **CB** are not defined. The **identity matrix**, **I**, which serves a role similar to 1 in scalar multiplication/division, is given by (for the 2 x 2 cases),

$$\mathbf{I} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Note that $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$. Finally, the **Inverse** \mathbf{A}^{-1} of a square matrix **A** is defined as satisfying

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{AA}^{-1} = \mathbf{I}$$

For a 2 x 2 matrix,

$$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

The role of the inverse and identity matrix is in solving systems of equations. Suppose we are trying to solve for **x** in $\mathbf{AX} = \mathbf{C}$. Premultiplying both sides by the inverse of **A** gives

$$\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{A}^{-1}\mathbf{C}$$

noting that $\mathbf{A}^{-1}\mathbf{Ax} = \mathbf{Ix} = \mathbf{x}$ gives $\mathbf{x} = \mathbf{A}^{-1}\mathbf{C}$

We now have all the pieces in place to express the response in matrix form. Suppose there are *n* traits under selection and we place the *n* selection differentials for each trait into a column vector **S**,

$$\mathbf{S} = \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{pmatrix}$$

Likewise, define the **phenotypic** and **additive genetic covariance matrices**, **P** and **G**, respectively, as matrices whose element in the *i*th row and *j* column is the covariance (phenotype or additive genetic) between traits *i* and *j*. Note that the diagonal elements are the variances. For example, for two characters

$$\mathbf{P} = \begin{pmatrix} \sigma^2(P_1) & \sigma(P_1, P_2) \\ \sigma(P_1, P_2) & \sigma^2(P_2) \end{pmatrix}, \quad \text{and} \quad \mathbf{G} = \begin{pmatrix} \sigma^2(A_1) & \sigma(A_1, A_2) \\ \sigma(A_1, A_2) & \sigma^2(A_2) \end{pmatrix}$$

Let **R** denote the column vector of selection responses, so that the *i*th element in the list is R_i , the change in the mean of character *i* following one generation of selection. The response to selection becomes

$$\mathbf{R} = \mathbf{GP}^{-1}\mathbf{S}$$

This equation is often referred to as the **multidimensional breeders' equation**. Recall that the response for a single character under selection is $R = h^2S = \sigma_A^2(\sigma_P^2)^{-1}S$. In the multidimensional case, the genetic and phenotypic variance are replaced by variance-covariance matrices and we use matrix inversion and multiplication.

The Directional Selection Gradient

The multivariate breeder's equation can also be written as

$$\mathbf{R} = \mathbf{G}\boldsymbol{\beta}$$

where

$$\boldsymbol{\beta} = \mathbf{P}^{-1}\mathbf{S}$$

is called the **directional selection gradient**. The i th element of $\boldsymbol{\beta}$, β_i , measures the amount of direct selection on trait X_i (i.e., the effects of correlated selection differentials are removed). From the rules of matrix multiplication, the response in trait j can be written as

$$R_j = \sigma^2(A_j) \beta_j + \sum_{i \neq j} \sigma(A_j, A_i) \beta_i$$

where the first term is the change due to direct selection on trait j and the sum is the indirect contribution to the response due to the correlated response of selection on other traits. Likewise note that $\mathbf{P}\boldsymbol{\beta} = \mathbf{S}$, so that the selection differential on trait j can be written as

$$S_j = \sigma^2(P_j) \beta_j + \sum_{i \neq j} \sigma(P_j, P_i) \beta_i$$

where the first term represents the contribution from direct selection on trait j and the sum term the contribution to the within-generation change due to direct selection on phenotypically correlated traits.

Effect of Selection on Genetic Correlation

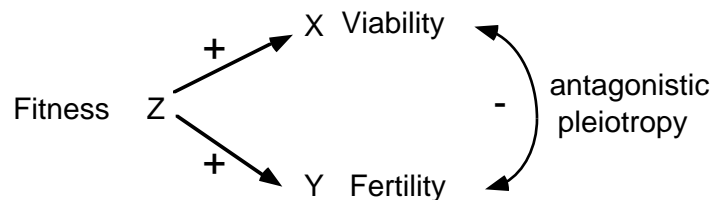
Sometimes two or more correlated characters are selected jointly — e.g. selection for increased body length and weight of mice. This is also true of natural selection, for which all traits contributing positively to the composite trait, fitness, are jointly selected.

One consequence of simultaneous selection for correlated characters is that the genetic correlation between them can become negative.

Consider a simple example of a composite trait Z controlled by multiple genes, which have different pleiotropic effects on two component traits, X and Y :

Locus	Effects of allele 1 relative to allele 2:	
	on X	on Y
A	+	+
B	+	-
C	-	+
D	-	-

The effect of selection on Z will be to fix allele 1 at locus A and allele 2 at locus D. However alleles at loci B and C will remain segregating at intermediate frequencies. The genetic correlation of traits X and Y tend to become negative because the only genes that remain affect X and Y in opposite directions. The heritability of trait Z will tend to 0 and it will exhibit no further selection response. The heritabilities of traits X and Y will be greater than 0, and there would be a response to selection for either trait separately. In this case the mean value of the correlated trait would decline.



Correlated Characters Problems

1. Consider selection acting on two traits (1 and 2). The phenotypic variances and covariances are $\sigma^2(P_1) = \sigma^2(P_2) = 10$, $\sigma(P_1, P_2) = -5$, while the additive genetic variances and covariances are $\sigma^2(A_1) = 4$, $\sigma^2(A_2) = 9$, $\sigma(A_1, A_2) = 3$. Compute the response on both characters when:

- a: We select directly on trait 2, with $S_2 = 10$
- b: We select directly on trait 1, with $S_1 = 10$
- c: $S_1 = 5, S_2 = 5$ (use the multivariate breeders equation)

2. The following covariances of performance in randomly chosen dam and daughter pairs of dairy cattle were obtained from an analysis. Estimate the heritabilities of milk yield and fat % and the genetic, phenotypic, and environmental correlations between them.

	Dam's yield	Dam's fat %
Dam's yield (in 100 kg units)	68	-0.55
Dam's fat %	-0.55	0.11
Daughter's yield	7.8	-0.20
Daughter's fat	-0.18	0.035

3. The heritability of growth rate in pine seedlings is 0.5 when they are grown in the greenhouse, and 0.2 when grown in the field. The genetic correlation between growth rate in the two environments is 0.8. Suppose you wish to select for increased growth in the field. In which environment should you do the selection? How would your conclusions change if the genetic correlations between growth rate in the two environments was 0.5?