

EVOLUTION OF QUANTITATIVE TRAITS

Introduction

Let's stop and review quickly where we've come and where we're going. We started our survey of quantitative genetics by pointing out that our objective was to develop a way to describe the patterns of phenotypic resemblance among relatives. The challenge was that we wanted to do this for phenotypic traits that whose expression is influenced both by many genes and by the environment in which those genes are expressed. Beyond the technical, algebraic challenges associated with many genes, we have the problem that we can't directly associate particular genotypes with particular phenotypes. We have to rely on patterns of *phenotypic* resemblance to tell us something about how *genetic* variation is transmitted. Surprisingly, we've managed to do that. We now know that it's possible to:

- Estimate the additive effect of an allele.¹
- Partition the phenotypic variance into genotypic and environmental components and to partition the genotypic variance into additive and dominance components.²
- Estimate all of the variance components from a combination of appropriate crossing designs and appropriate statistical analyses.

Now we're ready for the next step: applying all of these ideas to the evolution of a quantitative trait.

¹Actually, we don't know this. You'll have to take my word for it that in certain breeding designs its possible to estimate not only the additive genetic variance and the dominance genetic variance, but also the actual additive effect of "alleles" that we haven't even identified. We'll see a more direct approach soon, when we get to genome-wide associations studies.

²I should point out that this is an oversimplification. I've mentioned that we typically assume that we can simply add the effects of alleles across loci, but if you think about how genes actually work in organisms, you realize that such additivity across loci isn't likely to be very common. Strictly speaking there are epistatic components to the genetic variance too, i.e., components of the genetic variance that have to do not with the interaction among alleles at a single locus (the dominance variance that we've already encountered), but with the interaction of alleles at different loci.

Evolution of the mean phenotype

We're going to focus on how the mean phenotype in a population changes in response to natural selection, specifically in response to viability selection. Before we can do this, however, we need to think a bit more carefully about the relationship between genotype, phenotype, and fitness. Let $F_{ij}(x)$ be the probability that genotype A_iA_j has a phenotype smaller than x .³ Then x_{ij} , the genotypic value of A_iA_j is

$$x_{ij} = \int_{-\infty}^{\infty} x dF_{ij}(x)$$

and the population mean phenotype is $p^2x_{11} + 2pqx_{12} + q^2x_{22}$. If an individual with phenotype x has fitness $w(x)$, then the fitness of an individual with genotype A_iA_j is

$$w_{ij} = \int_{-\infty}^{\infty} w(x) dF_{ij}(x)$$

and the mean fitness in the population is $\bar{w} = p^2w_{11} + 2pqw_{12} + q^2w_{22}$.

Now, there's a well known theorem from calculus known as Taylor's theorem. It says that for any function⁴ $f(x)$

$$f(x) = f(a) + \sum_{k=1}^{\infty} \left(\frac{(x-a)^k}{k!} \right) f^{(k)}(a) \quad .$$

Using this theorem we can produce an approximate expression describing how the mean phenotype in a population will change in response to selection. Remember that the mean phenotype, \bar{x} , depends both on the underlying genotypic values and on the allele frequency. So I'm going to write the mean phenotype as $\bar{x}(p)$ to remind us of that dependency. The phenotype changes from one generation to the next as a result of changes in the frequency of alleles that influence the phenotype, assuming that the environmental effects on phenotypes don't change.

$$\bar{x}(p') = \bar{x}(p) + (p' - p) \left(\frac{d\bar{x}}{dp} \right) + o(p^2)$$

$$\bar{x}(p) = p^2x_{11} + 2pqx_{12} + q^2x_{22}$$

³For those of you who have had probability theory, $F_{ij}(x)$ is the cumulative distribution for the probability density for phenotype associated with A_iA_j .

⁴Actually there are restrictions on the functions to which it applies, but we can ignore those restrictions for our purposes.

$$\begin{aligned}
\frac{d\bar{x}(p)}{dp} &= 2px_{11} + 2qx_{12} - 2px_{12} - 2qx_{22} \\
&= 2\{(px_{11} + qx_{12} - \bar{x}/2) - (px_{12} + qx_{22} - \bar{x}/2)\} \\
&= 2(\alpha_1 - \alpha_2)
\end{aligned}$$

$$\bar{x}(p') \approx \bar{x}(p) + (p' - p)(2(\alpha_1 - \alpha_2))$$

$$\Delta\bar{x} = (\Delta p)(2(\alpha_1 - \alpha_2))$$

Now you need to remember (from lo those many weeks ago) that

$$p' = \frac{p^2w_{11} + pqw_{12}}{\bar{w}} \quad .$$

Thus,

$$\begin{aligned}
\Delta p &= p' - p \\
&= \frac{p^2w_{11} + pqw_{12}}{\bar{w}} - p \\
&= \frac{p^2w_{11} + pqw_{12} - p\bar{w}}{\bar{w}} \\
&= p \left(\frac{pw_{11} + qw_{12} - \bar{w}}{\bar{w}} \right) \quad .
\end{aligned}$$

Now,⁵ let's do a linear regression of fitness on phenotype. After all, to make any further progress, we need to relate phenotype to fitness, so that we can use the relationship between phenotype and genotype to infer the change in allele frequencies, from which we will infer the change in mean phenotype.⁶ From our vast statistical knowledge, we know that the slope of this regression line is

$$\beta_1 = \frac{\text{Cov}(w, x)}{\text{Var}(x)}$$

and its intercept is

$$\beta_0 = \bar{w} - \beta_1\bar{x} \quad .$$

⁵Since we're having so much fun with mathematics why should we stop here?

⁶Whew! That was a mouthful.

Let's use this regression equation to determine the fitness of each genotype. This is only an approximation to the true fitness,⁷ but it is adequate for many purposes.

$$\begin{aligned}
 w_{ij} &= \int_{-\infty}^{\infty} w(x) dF_{ij}(x) \\
 &\approx \int_{-\infty}^{\infty} (\beta_0 + \beta_1 x) dF_{ij}(x) \\
 &= \beta_0 + \beta_1 x_{ij} \\
 \bar{w} &= \beta_0 + \beta_1 \bar{x} \quad .
 \end{aligned}$$

If we substitute this into our expression for Δp above, we get

$$\begin{aligned}
 \Delta p &= p \left(\frac{pw_{11} + qw_{12} - \bar{w}}{\bar{w}} \right) \\
 &= p \left(\frac{p(\beta_0 + \beta_1 x_{11}) + q(\beta_0 + \beta_1 x_{12}) - (\beta_0 + \beta_1 \bar{x})}{\bar{w}} \right) \\
 &= p\beta_1 \left(\frac{px_{11} + qx_{12} - \bar{x}}{\bar{w}} \right) \\
 &= p\beta_1 \left(\frac{\alpha_1 - \bar{x}/2}{\bar{w}} \right) \\
 &= p\beta_1 \left(\frac{\alpha_1 - (p\alpha_1 + q\alpha_2)}{\bar{w}} \right) \\
 &= \frac{pq\beta_1(\alpha_1 - \alpha_2)}{\bar{w}} \quad .
 \end{aligned}$$

So where are we now?⁸ Let's substitute this result back into the equation for $\Delta \bar{x}$. When we do we get

$$\begin{aligned}
 \Delta \bar{x} &= (\Delta p) (2(\alpha_1 - \alpha_2)) \\
 &= \left(\frac{pq\beta_1(\alpha_1 - \alpha_2)}{\bar{w}} \right) (2(\alpha_1 - \alpha_2)) \\
 &= 2pq\alpha^2 \left(\frac{\beta_1}{\bar{w}} \right) \\
 &= V_a \left(\frac{\beta_1}{\bar{w}} \right) \quad .
 \end{aligned}$$

⁷Specifically, we are implicitly assuming that the fitnesses are adequately approximated by a linear function of our phenotypic measure.

⁸You don't have to tell me where you *wish* you were. I can reliably guess that it's not here.

This is great if we've done the regression between fitness and phenotype, but what if we haven't?⁹ Let's look at $\text{Cov}(w, x)$ in a little more detail.

$$\begin{aligned}
\text{Cov}(w, x) &= p^2 \int_{-\infty}^{\infty} xw(x) dF_{11}(x) + 2pq \int_{-\infty}^{\infty} xw(x) dF_{12}(x) \\
&\quad + q^2 \int_{-\infty}^{\infty} xw(x) dF_{22}(x) - \bar{x}\bar{w} \\
&= p^2 \left(\int_{-\infty}^{\infty} xw(x) dF_{11}(x) - x_{11}\bar{w} + x_{11}\bar{w} \right) \\
&\quad + 2pq \left(\int_{-\infty}^{\infty} xw(x) dF_{11}(x) - x_{12}\bar{w} + x_{12}\bar{w} \right) \\
&\quad + q^2 \left(\int_{-\infty}^{\infty} xw(x) dF_{22}(x) - x_{22}\bar{w} + x_{22}\bar{w} \right) \\
&\quad - \bar{x}\bar{w} \\
&= p^2 \left(\int_{-\infty}^{\infty} xw(x) dF_{11}(x) - x_{11}\bar{w} \right) \\
&\quad + 2pq \left(\int_{-\infty}^{\infty} xw(x) dF_{11}(x) - x_{12}\bar{w} \right) \\
&\quad + q^2 \left(\int_{-\infty}^{\infty} xw(x) dF_{22}(x) - x_{22}\bar{w} \right) \quad .
\end{aligned}$$

Now

$$\begin{aligned}
\int_{-\infty}^{\infty} xw(x) dF_{ij}(x) - x_{ij}\bar{w} &= \bar{w} \left(\int_{-\infty}^{\infty} \frac{xw(x)}{\bar{w}} dF_{ij}(x) - x_{ij} \right) \\
&= \bar{w}(x_{ij}^* - x_{ij}) \quad ,
\end{aligned}$$

where x_{ij}^* refers to the mean phenotype of A_iA_j after selection. So

$$\begin{aligned}
\text{Cov}(w, x) &= p^2\bar{w}(x_{11}^* - x_{11}) + 2pq\bar{w}(x_{12}^* - x_{12})q^2\bar{w}(x_{22}^* - x_{22}) \\
&= \bar{w}(\bar{x}^* - \bar{x}) \quad ,
\end{aligned}$$

where \bar{x}^* is the population mean phenotype after selection. In short,¹⁰ combining our equations for the change in mean phenotype and for the covariance of fitness and phenotype and remembering that $\beta_1 = \text{Cov}(w, x)/\text{Var}(x)$ ¹¹

$$\Delta\bar{x} = V_a \left(\frac{\bar{w}(\bar{x}^* - \bar{x})}{\bar{w}} \right)$$

⁹Hang on just a little while longer. We're almost there.

¹⁰We finally made it.

¹¹You also need to remember that $\text{Var}(x) = V_p$, since they're the same thing, the phenotypic variance.

Genotype	A_1A_1	A_1A_2	A_2A_2
Phenotype	1.303	1.249	0.948

Table 1: A simple example to illustrate response to selection in a quantitative trait.

$$= h_N^2(\bar{x}^* - \bar{x})$$

$\Delta\bar{x} = \bar{x}' - \bar{x}$ is referred to as the response to selection and is often given the symbol R . It is the change in population mean between the parental generation (before selection) and the offspring generation (before selection). $\bar{x}^* - \bar{x}$ is referred to as the selection differential and is often given the symbol S . It is the difference between the mean phenotype in the parental generation before selection and the mean phenotype in the parental generation after selection. Thus, we can rewrite our final equation as

$$R = h_N^2 S \quad .$$

This equation is often referred to as the *breeders equation*.

A Numerical Example

To illustrate how this works, let's examine the simple example in Table 1.

Given these phenotypes, $p = 0.25$, and $V_p = 0.16$, it follows that $\bar{x} = 1.08$ and $h_N^2 = 0.1342$. Suppose the mean phenotype after selection is 1.544. What will the phenotype be among the newly born progeny?

$$\begin{aligned}
 S &= \bar{x}^* - \bar{x} \\
 &= 1.544 - 1.08 \\
 &= 0.464 \\
 \Delta\bar{x} &= h_N^2 S \\
 &= (0.1342)(0.464) \\
 &= 0.06 \\
 \bar{x}' &= \bar{x} + \Delta\bar{x} \\
 &= 1.08 + 0.06 \\
 &= 1.14
 \end{aligned}$$

Genotype	A_1A_1	A_1A_2	A_2A_2
Frequency	p^2	$2pq$	q^2
Fitness	w_{11}	w_{12}	w_{22}
Additive fitness value	$2\alpha_1$	$\alpha_1 + \alpha_2$	$2\alpha_2$

Table 2: Fitnesses and additive fitness values used in deriving Fisher's Fundamental Theorem of Natural Selection.

Fisher's Fundamental Theorem of Natural Selection

Suppose the phenotype whose evolution we're interested in following is fitness itself.¹² Then we can summarize the fitnesses as illustrated in Table 2.

Although I didn't tell you this, a well-known fact about viability selection at one locus is that the change in allele frequency from one generation to the next can be written as

$$\Delta p = \left(\frac{pq}{2\bar{w}} \right) \left(\frac{d\bar{w}}{dp} \right) \quad .$$

Using our new friend, Taylor's theorem, it follows immediately that

$$\bar{w}' = \bar{w} + (\Delta p) \left(\frac{d\bar{w}}{dp} \right) + \left(\frac{(\Delta p)^2}{2} \right) \left(\frac{d^2\bar{w}}{dp^2} \right) \quad .$$

Or, equivalently

$$\Delta\bar{w} = (\Delta p) \left(\frac{d\bar{w}}{dp} \right) + \left(\frac{(\Delta p)^2}{2} \right) \left(\frac{d^2\bar{w}}{dp^2} \right) \quad .$$

Recalling that $\bar{w} = p^2w_{11} + 2p(1-p)w_{12} + (1-p)^2w_{22}$ we find that

$$\begin{aligned} \frac{d\bar{w}}{dp} &= 2pw_{11} + 2(1-p)w_{12} - 2pw_{12} - 2(1-p)w_{22} \\ &= 2[(pw_{11} + qw_{12}) - (pw_{12} + qw_{22})] \\ &= 2[(pw_{11} + qw_{12} - \bar{w}/2) - (pw_{12} + qw_{22} - \bar{w}/2)] \\ &= 2[\alpha_1 - \alpha_2] \\ &= 2\alpha \quad , \end{aligned}$$

¹²The proof of the fundamental theorem that follows is due to C. C. Li [1]

where the last two steps use the definitions for α_1 and α_2 , and we set $\alpha = \alpha_1 - \alpha_2$. Similarly,

$$\begin{aligned}\frac{d^2\bar{w}}{dp^2} &= 2w_{11} - 2w_{12} - 2w_{12} + 2w_{22} \\ &= 2(w_{11} - 2w_{12} + w_{22})\end{aligned}$$

Now we can plug these back into the equation for $\Delta\bar{w}$:

$$\begin{aligned}\Delta\bar{w} &= \left\{ \left(\frac{pq}{2\bar{w}} \right) \left(\frac{d\bar{w}}{dp} \right) \right\} \left(\frac{d\bar{w}}{dp} \right) + \frac{\left\{ \left(\frac{pq}{2\bar{w}} \right) \left(\frac{d\bar{w}}{dp} \right) \right\}^2}{2} [2(w_{11} - 2w_{12} + w_{22})] \\ &= \left\{ \left(\frac{pq}{2\bar{w}} \right) (2\alpha) \right\} (2\alpha) + \left\{ \left(\frac{pq}{2\bar{w}} \right) (2\alpha) \right\}^2 (w_{11} - 2w_{12} + w_{22}) \\ &= \frac{2pq\alpha^2}{\bar{w}} + \frac{p^2q^2\alpha^2}{\bar{w}^2} (w_{11} - 2w_{12} + w_{22}) \\ &= \frac{V_a}{\bar{w}} \left\{ 1 + \frac{pq}{2\bar{w}} (w_{11} - 2w_{12} + w_{22}) \right\} \quad ,\end{aligned}$$

where the last step follows from the observation that $V_a = 2pq\alpha^2$. The quantity $\frac{pq}{2\bar{w}}(w_{11} - 2w_{12} + w_{22})$ is usually quite small, especially if selection is not too intense.¹³ So we are left with

$$\Delta\bar{w} \approx \frac{V_a}{\bar{w}} \quad .$$

References

- [1] C C Li. *First Course in Population Genetics*. Boxwood Press, Pacific Grove, CA, 1976.

Creative Commons License

These notes are licensed under the Creative Commons Attribution 4.0 International License (CC BY 4.0). To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, 559 Nathan Abbott Way, Stanford, California 94305, USA.

¹³Notice that it's exactly equal to 0 if the fitness of the heterozygote is exactly intermediate. In that case, all of the variance in fitness is additive.