Selection on multiple characters

Introduction

So far we've studied only the evolution of a single trait, e.g., height or weight. But organisms have many traits, and they evolve at the same time. How can we understand their simultaneous evolution? The basic framework of the quantitative genetic approach was first outlined by Russ Lande and Steve Arnold [2].

Let z_1, z_2, ..., z_n be the phenotype of each character that we are studying. We'll use $\bar{\mathbf{z}}$ to denote the vector of these characters before selection and $\bar{\mathbf{z}}^*$ to denote the vector after selection. The selection differential, $\mathbf{s}$, is also a vector given by

$$\mathbf{s} = \bar{\mathbf{z}}^* - \bar{\mathbf{z}} \quad .$$

Suppose $p(\mathbf{z})$ is the probability that any individual has phenotype $\mathbf{z}$, and let $W(\mathbf{z})$ be the fitness (absolute viability) of an individual with phenotype $\mathbf{z}$. Then the mean absolute fitness is

$$\bar{W} = \int W(\mathbf{z})p(\mathbf{z})d\mathbf{z} \quad .$$

The relative fitness of phenotype $\mathbf{z}$ can be written as

$$w(\mathbf{z}) = \frac{W(\mathbf{z})}{\bar{W}} \quad .$$

Using relative fitnesses the mean relative fitness, $\bar{w}$, is 1. Now

$$\bar{\mathbf{z}}^* = \int \mathbf{z}w(\mathbf{z})p(\mathbf{z})d\mathbf{z} \quad .$$

Recall that $Cov(X,Y) = E(X - \mu_x)(Y - \mu_y) = E(XY) - \mu_x\mu_y$. Consider

$$
\begin{aligned}
\mathbf{s} &= \bar{\mathbf{z}}^* - \bar{\mathbf{z}} \\
&= \int \mathbf{z}w(\mathbf{z})p(\mathbf{z})d\mathbf{z} - \bar{\mathbf{z}} \\
&= E(w, z) - \bar{w}\bar{\mathbf{z}} \quad ,
\end{aligned}
$$

where the last step follows since $\bar{w} = 1$ meaning that $\bar{w}\bar{\mathbf{z}} = \bar{\mathbf{z}}$. In short,

$$\mathbf{s} = Cov(w, z) \quad .$$

That should look familiar from our analysis of the evolution of a single phenotype.

If we assume that all genetic effects are additive, then the phenotype of an individual can be written as

$$\mathbf{z} = \mathbf{x} + \mathbf{e} \quad ,$$

where $\mathbf{x}$ is the additive genotype and $\mathbf{e}$ is the environmental effect. We'll denote by $\mathbf{G}$ the matrix of genetic variances and covariances and by $\mathbf{E}$ the matrix of environmental variances and covariances. The matrix of phenotype variances and covariances, $\mathbf{P}$, is then given by[1]

$$\mathbf{P} = \mathbf{G} + \mathbf{E} \quad .$$

Now, if we're willing to assume that the regression of additive genetic effects on phenotype is linear[2] and that the environmental variance is the same for every genotype, then we can predict how phenotypes will change from one generation to the next

$$
\begin{aligned}
\bar{\mathbf{x}}^* - \bar{\mathbf{x}} &= \mathbf{G}\mathbf{P}^{-1}(\bar{\mathbf{z}}^* - \bar{\mathbf{z}}) \\
\bar{\mathbf{z}}' - \bar{\mathbf{z}} &= \mathbf{G}\mathbf{P}^{-1}(\bar{\mathbf{z}}^* - \bar{\mathbf{z}}) \\
\Delta\bar{\mathbf{z}} &= \mathbf{G}\mathbf{P}^{-1}\mathbf{s}
\end{aligned}
$$

$\mathbf{G}\mathbf{P}^{-1}$ is the multivariate version of h_N^2. This equation is also the multivariate version of the breeders equation.

But we have already seen that $\mathbf{s} = Cov(w, z)$. Thus,

$$\beta = \mathbf{P}^{-1}\mathbf{s}$$

is a set of partial regression coefficients of relative fitness on the characters, i.e., the dependence of relative fitness on that character alone holding all others constant.

Note:

$$
\begin{aligned}
s_i &= \sum_{j=1}^{n} \beta_j P_{ij} \\
&= \beta_1 P_{i1} + \cdots + \beta_i P_{ii} + \cdots + \beta_n P_{in}
\end{aligned}
$$

is the total selective differential in character i, including the indirect effects of selection on other characters.

---

[1] Assuming that there are no genotype $\times$ environment interactions.

[2] And we were willing to do this when we were studying the evolution of only one trait, so why not do it now?

Character	Mean before selection	standard deviation
head	0.880	0.034
thorax	2.038	0.049
scutellum	1.526	0.057
wing	2.337	0.043

	head	thorax	scutellum	wing
head	1.00	0.72	0.50	0.60
thorax		1.00	0.59	0.71
scutellum			1.00	0.62
wing				1.00

Character	s	s'	β	β'
head	-0.004	-0.11	-0.7 ± 4.9	-0.03 ± 0.17
thorax	-0.003	-0.06	$11.6 \pm 3.9^{**}$	$0.58 \pm 0.19^{**}$
scutellum	-0.16^{*}	-0.28^{*}	-2.8 ± 2.7	-0.17 ± 0.15
wing	-0.019^{**}	-0.43^{**}	$-16.6 \pm 4.0^{**}$	$-0.74 \pm 0.18^{**}$

Table 1: Selection analysis of pentastomid bugs on the shores of Lake Michigan.

An example: selection in a pentastomid bug

94 individuals were collected along shoreline of Lake Michigan in Parker County, Indiana after a storm. 39 were alive, 55 dead. The means of several characters before selection, the trait correlations, and the selection analysis are presented in Table 1.

The column labeled s is the selective differential for each character. The column labeled s' is the *standardized* selective differential, i.e., the change measured in units of standard deviation rather than on the original scale.[3] A multiple regression analysis of fitness versus phenotype on the original scale gives estimates of β, the direct effect of selection on that trait. A multiple regression analysis of fitness versus phenotype on the transformed scale gives the standardized direct effect of selection, β', on that trait.

Notice that the selective differential[4] for the thorax measurement is negative, i.e., individuals that survived had smaller thoraces than those that died. But the *direct* effect of selection on thorax is strongly positive, i.e., all other things being equal, an individual with a

---

[3]To measure on this scale the data is simply transformed by setting $y_i = (x_i - \bar{x})/s$, where x_i is the raw score for the ith individual, $\bar{x}$ is the sample mean for the trait, and s is its standard deviation.

[4]The cumulative effect of selection on the change in mean phenotype.

	body	tail
body	35.4606	11.3530
tail	11.3530	37.2973

Table 2: Genetic variance-covariance matrix for vertebral number in central Californian garter snakes.

large was more likely to survive than one with a small thorax. Why the apparent contradiction? Because the thorax measurement is positively correlated with the wing measurement, and there's strong selection for decreased values of the wing measurement.

Cumulative selection gradients

Arnold [1] suggested an extension of this approach to longer evolutionary time scales. Specifically, he studied variation in the number of body vertebrae and the number of tail vertebrae in populations of *Thamnophis elegans* from two regions of central California. He found relatively little vertebral variation within populations, but there were considerable differences in vertebral number between populations on the coast side of the Coast Ranges and populations on the Central Valley side of the Coast Ranges. The consistent difference suggested that selection might have produced these differences, and Arnold attempted to determine the amount of selection necessary to produce these differences.

The data

Arnold collected pregnant females from two local populations in each of two sites in northern California 282 km apart from one another. Females were collected over a ten-year period and returned to the University of Chicago. Dam-offspring regressions were used to estimate additive genetic variances and covariances of vertebral number.[5] Mark-release-recapture experiments in the California populations showed that females with intermediate numbers of vertebrae grow at the fastest rate, at least at the inland site, although no such relationship was found in males. The genetic variance-covariance matrix he obtained is shown in Table 2.

---

[5]1000 progeny from 100 dams.

The method

We know from Lande and Arnold's results that the change in multivariate phenotype from one generation to the next, $\Delta\bar{\mathbf{z}}$, can be written as

$$\Delta\bar{\mathbf{z}} = \mathbf{G}\beta \quad,$$

where $\mathbf{G}$ is the genotypic variance-covariance matrix, $\beta = \mathbf{P}^{-1}\mathbf{s}$ is the set of partial regression coefficients describing the direct effect of each character on relative fitness.[6] If we are willing to assume that $\mathbf{G}$ remains constant, then the total change in a character subject to selection for n generations is

$$\sum_{k=1}^{n}\Delta\bar{\mathbf{z}} = \mathbf{G}\sum_{k=1}^{n}\beta \quad.$$

Thus, $\sum_{k=1}^{n}\beta$ can be regarded as the cumulative selection differential associated with a particular observed change, and it can be estimated as

$$\sum_{k=1}^{n}\beta = \mathbf{G}^{-1}\sum_{k=1}^{n}\Delta\bar{\mathbf{z}} \quad.$$

The results

The overall difference in vertebral number between inland and coastal populations can be summarized as:

$$\begin{aligned}
\text{body}_{\text{inland}} - \text{body}_{\text{coastal}} &= 16.21 \\
\text{tail}_{\text{inland}} - \text{tail}_{\text{coastal}} &= 9.69
\end{aligned}$$

Given the estimate of $\mathbf{G}$ already obtained, this corresponds to a cumulative selection gradient between inland and coastal populations of

$$\begin{aligned}
\beta_{\text{body}} &= 0.414 \\
\beta_{\text{tail}} &= 0.134
\end{aligned}$$

Applying the same technique to looking at the differences between populations within the inland site and within the coastal site we find cumulative selection gradients of

$$\begin{aligned}
\beta_{\text{body}} &= 0.035 \\
\beta_{\text{tail}} &= 0.038
\end{aligned}$$

---

[6]$\mathbf{P}$ is the phenotypic variance-covariance matrix and $\mathbf{s}$ is the vector of selection differentials.

for the coastal site and

$$\beta_{\text{body}} = 0.035$$
$$\beta_{\text{tail}} = -0.004$$

for the inland site.

The conclusions

"To account for divergence between inland and coastal California, we must invoke cumulative forces of selection that are 7 to 11 times stronger than the forces needed to account for differentiation of local populations."

Furthermore, recall that the selection gradients can be used to partition the overall response to selection in a character into the portion due to the direct effects of that character alone and the portion due to the indirect effects of selection on a correlated character. In this case the overall response to selection in number of body vertebrae is given by

$$\mathbf{G}_{11}\beta_1 + \mathbf{G}_{12}\beta_2 \quad ,$$

where $\mathbf{G}_{11}\beta_1$ is the direct effect of body vertebral number and $\mathbf{G}_{12}\beta_2$ is the indirect effect of tail vertebral number. Similarly, the overall response to selection in number of tail vertebrae is given by

$$\mathbf{G}_{12}\beta_1 + \mathbf{G}_{22}\beta_2 \quad ,$$

where $\mathbf{G}_{22}\beta_2$ is the direct effect of tail vertebral number and $\mathbf{G}_{12}\beta_1$ is the indirect effect of body vertebral number. Using these equations it is straightforward to calculate that 91% of the total divergence in number of body vertebrae is a result of direct selection on this character. In contrast, only 51% of the total divergence in number of tail vertebrae is a result of direct selection on this character, i.e., 49% of the difference in number of tail vertebrae is attributable to indirect selection as a result of its correlation with number of body vertebrae.

The caveats

While the approach Arnold suggests is intriguing, there are a number of caveats that must be kept in mind in trying to apply it.

- This approach assumes that the $\mathbf{G}$ matrix remains constant.

- This approach cannot distinguish strong selection that happened over a short period of time from weak selection that happened over a long period of time.

6

- This approach *assumes* that the observed differences in populations are the result of selection, but populations isolated from one another will diverge from one another even in the absence of selection simply as a result of genetic drift.

  - Small amount of differentiation between populations within sites could reflect relatively recent divergence of those populations from a common ancestral population.
  - Large amount of differentiation between populations from inland versus coastal sites could reflect a more ancient divergence from a common ancestral population.

References

[1] S J Arnold. Quantitative genetics and selection in natural populations: microevolution of vertebral numbers in the garter snake *Thamnophis elegans*. In B S Weir, E J Eisen, M M Goodman, and G Namkoong, editors, *Proceedings of the Second International Conference on Quantitative Genetics*, pages 619–636. Sinauer Associates, Sunderland, MA, 1988.

[2] R Lande and S J Arnold. The measurement of selection on correlated characters. *Evolution*, 37:1210–1226, 1983.

Creative Commons License