NATURAL SELECTION FOR WITHIN-GENERATION VARIANCE IN OFFSPRING NUMBER

JOHN H. GILLESPIE

Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19174

ABSTRACT

In this paper it is shown that natural selection can act on the within-generation variance in offspring number. The fitness of a genotype will increase as its variance in offspring number decreases. The intensity of selection on the variance component is inversely proportional to population size, although the fixation probability of a gene which differs from its allele only in the variance in its offspring number is independent of population size. The concept of effective population size is shown to be of limited use when there is genetic variation in the variance in offspring number.

THE probability that the nest of an ant bird in tropical America escapes predation is about .1. If, due to some biological constraint, ant birds can lay no more than M eggs in a breeding season, it is reasonable to ask: should all M eggs be put in one basket, or should they be evenly distributed into k baskets? The mean number of successful fledglings will be the same in both cases, but the variance in the numbers of fledglings will be $M^2(.09)$ in the first case, $M^2(.09)/k$ in the second. The second strategy seems preferable, yet the existing mathematical theory of natural selection does not seem to have treated this situation. In this paper the consequences of selection on the variance in offspring number will be explored mathematically. At the outset it is important to note that the variance in the numbers of offspring of a genotype has two components, the within-generation component resulting from different individuals of the same genotype having different numbers of offspring, and a between-generation component due to the effects of a changing environment. This latter component has been recently investigated (GILLESPIE 1973); the former will be explored in this paper.

1. *Haploid selection:* Consider a haploid population made up of two genotypes, A_1 and A_2, whose numbers in the tth generation, $t = 1, 2, 3, \ldots$, are $x_1(t)$ and $x_2(t)$, respectively. Let $1 + Y_{ij}(t)$ be a random variable representing the number of offspring of the jth individual of the ith genotype in the tth generation. If the collection of random variables $\{Y_{ij}(t), i = 1,2; j = 1,2, \ldots x_i(t)\}$ are independent, the process $(x_1(t), x_2(t))$ is the two-dimensional branching process introduced as a genetic model by FELLER (1951). As pointed out by FELLER, this process may be approximated by a diffusion process whose forward equation may be written

$$\frac{\partial \phi}{\partial t} = \sum_{i=1}^{2} \{-\mu_i \frac{\partial}{\partial x_i}[x_i\phi] + \frac{1}{2}\sigma_i^2 \frac{\partial^2}{\partial x_i^2}[x_i\phi]\} \tag{1}$$

where $1 + \mu_i$ is the mean number of offspring of genotype A_i and σ_i^2 is the variance in the number of offspring of genotype A_i. The density of $(x_1(t), x_2(t))$ is given by $\phi(x_1, x_2, t)$. This

equation may be transformed into an equation in p, the frequency of allele A_i, and n, the total population size by

$$p = \frac{x_1}{x_1 + x_2}, \qquad n = x_1 + x_2 \tag{2}$$

yielding a new forward equation

$$\frac{\partial \phi}{\partial t} = -\frac{\partial}{\partial p} \left[p(1-p)(\mu_1 - \mu_2 + \frac{1}{n}(\sigma_2^2 - \sigma_1^2)) \phi \right] + \frac{1}{2} \frac{\partial^2}{\partial p^2} \left[p \frac{(1-p)}{n} (p\sigma_2^2 \right.$$

$$+ (1-p)\sigma_1^2 \phi] - \frac{\partial}{\partial n} [n(p\mu_1 + (1-p)\mu_2) \phi] + \frac{1}{2} \frac{\partial^2}{\partial n^2} [n(p\sigma_1^2 \tag{3}$$

$$\left. + (1-p)\sigma_2^2) \phi \right] + \frac{\partial^2}{\partial p \partial n} \left[p(1-p)(\sigma_1^2 - \sigma_2^2) \phi \right].$$

If we regard this equation as approximating the behavior of a density-regulated population, n may be viewed as a fixed constant instead of a Markovian variable, and we will be left with a one-dimension process in p with drift coefficient

$$M(p) = p(1-p) \left([\mu_1 - \mu_2 + \frac{1}{n}(\sigma_2^2 - \sigma_1^2)'] \right) \tag{4}$$

and diffusion coefficient

$$V(p) = \frac{p(1-p)}{n} (\sigma_2^2 p + \sigma_1^2 (1-p)). \tag{5}$$

Unlike the equivalent coefficients used in the stochastic models pioneered by WRIGHT (1945) and KIMURA (1964), these coefficients explicitly incorporate the variances in offspring numbers of the two genotypes. Their exclusion from the WRIGHT-KIMURA theory is due to the postulated "infinite gametic pool" (CROW and KIMURA 1970, p. 327) which, through the law of large numbers, removes the possibility of variance effects being important. In the present model, one which is certainly closer to reality for many species, the "gametic pool" will be essentially the same size as the final population.

Examination of $M(p)$ shows immediately the two main properties of the action of selection on the within-generation component of variance in offspring number: (1) Lowering the variance in offspring number will increase the fitness of a genotype. (2) The strength of selection for the variance component is inversely proportional to population size.

Selection on variance in offspring number can be studied in isolation from mean effects if we set $\mu_1 = \mu_2$. In this situation the probability of fixation of allele A_1, given an initial frequency p, is

$$u(p) = \frac{(\alpha + 1)p}{\alpha + p} \tag{6}$$

where

$$\alpha = \frac{\sigma_1^2}{\sigma_2^2 - \sigma_1^2}. \tag{7}$$

(See CROW and KIMURA 1970 for the theory which allows the calculation of $u(p)$). This probability is *independent of population size*, contrary to all other forms of natural selection (KIMURA 1962). The reason for this is that both the intensity of selection and the magnitude of the stochastic effects are inversely proportional to population size. It is also interesting to note that the probability, $u(p)$, depends on the ratio of variances rather than their absolute values, this, again, being unlike other forms of selection. Examination of $u(p)$ shows that the allele with the smaller variance in offspring number always has a better chance of fixation than a neutral allele with the same initial frequency. This further strengthens the statement that selection favors smaller variances in offspring number.

A seemingly paradoxical situation arises if we consider the limiting deterministic model which is obtained by letting $n \to \infty$. At the limit the rate of change of p is zero, yet, since $u(p)$ is independent of n, there remains a non-zero probability of the allele with the smaller variance

becoming fixed. To resolve this, examine the conditional mean time to fixation for allele, A_1, given that it does, in fact, become fixed. Using the theory of EWENS (1972), this mean time may be shown to equal

$$t_1(p) = \frac{2n}{\sigma_1{}^2} \left[-ln\, p - \frac{\alpha(1-p)}{1+\alpha} ln\,(1-p) + \frac{\alpha^2(1-p)}{1+\alpha} ln \left(\frac{\alpha}{\alpha+p} \right) \right] \qquad (8)$$

or, apprroximately if p/α is small:

$$t_1(p) \cong \frac{2n}{\sigma_1{}^2} (-ln\, p). \qquad (9)$$

The mean time is, therefore, proportional to the population size. In large populations, selection for variance in offspring numbers would proceed very slowly; at the limit it would take, on the average, an infinite time to reach fixation. Thus, the paradox is resolved.

It is possible to make a somewhat stronger statement about the behavior of the conditional waiting time to fixation than simply noting that its mean time approaches infinity. For the conditional process the drift coefficient is

$$M^*(p) = M(p) + V(p) \frac{u'(p)}{u(p)}. \qquad (10)$$

The diffusion coefficient is the same as in the unconditional process (EWENS 1972). The Laplace transform of the waiting time to fixation density $g(\lambda;p,n)$, satisfies an equation of the form

$$\frac{V(p)}{2} g'' + M^*(p)\, g' - \lambda\, g = 0. \qquad (11)$$

Since both $V(p)$ and $M(p)$ have factors of $1/n$, multiplying through this equation by n causes n to appear only as a factor of λ. The Laplace transform for populations with successively larger and larger population sizes will, therefore, be simple compressions of the smaller ones toward zero. This implies that the

$$\lim_{n \to \infty} g(\lambda; p,n) = 0, \quad \lambda > 0: \qquad (12)$$

Thus, the *entire* probability mass of the waiting time density approaches infinity.

Turning now to the more complicated case where the two alleles differ in both their mean and variance in offspring numbers, we arrive at the probability of fixation of A_1 as

$$U(p) = \frac{1 - (1 + p/\alpha)^{\frac{-2ns}{\sigma_2{}^2 - \sigma_1{}^2} + 1}}{1 - (1 + 1/\alpha)^{\frac{-2ns}{\sigma_2{}^2 - \sigma_1{}^2} + 1}} \qquad (13)$$

where

$$s = \mu_1 - \mu_2 + \frac{1}{n}(\sigma_2{}^2 - \sigma_1{}^2). \qquad (14)$$

This formula illustrates the way that increasing population size decreases the role of variance selection. In fact, as $n \to \infty$

$$\mu(p) \to \frac{1 - e^{\frac{-2n(\mu_1 - \mu_2)}{\sigma_1{}^2} p}}{1 - e^{\frac{-2n(\mu_1 - \mu_2)}{\sigma_1{}^2}}} \qquad \cdot \qquad (15)$$

which is of the same form as the fixation probability applicable to the WRIGHT-KIMURA model (KIMURA 1962). KIMURA's formula may also be obtained from (15) by allowing the variances in offspring number of the two genotypes to approach each other.

Examination of the second derivative of $u(p)$ shows it to be convex or concave depending on the sign of s. If

$$\mu_1 - \frac{1}{n}\sigma_1{}^2 > \mu_2 - \frac{1}{n}\sigma_2{}^2 \qquad (16)$$

then A_1 has a better chance of becoming fixed than a neutral allele with the same initial frequency. For this reason the best measure of the fitness of genotype A_i is

$$\mu_i - \frac{1}{n}\,\sigma_i{}^2$$

a quantity which depends on the population size. With this definition comical situations will sometimes arise. For example, if $\mu_1 > \mu_2$ and $\sigma_1{}^2 < \sigma_2{}^2$, there exists a population size for which the two alleles are neutral ($M(p)=0$).

The accuracy of these formulae may be checked by various methods. In Table 1 the results of computations of fixation probabilities when the numbers of offspring per individual are binomially distributed and the population process is generated by the direct-product branching process model of KARLIN and McGREGOR (1964) are compared to the diffusion approximations. The agreement is quite good, even for small population sizes. This justifies, in part, the use of n as a constant rather than a Markovian variable.

TABLE 1

A comparison of the exact values for the fixation probabilities with formula (6)

p	$N=10$	$a=2$ 20	$u(p)$	$N=10$	$a=1.33$ 20	$u(p)$
.1	.1367	.1374	.1429	.1510	.1428	.1629
.5	.5981	.5987	.6000	.6330	.6263	.6366
.9	.9335	.9336	.9310	.9443	.9434	.9404

2. *Diploid selection:* The diploid model is treated in an analagous fashion. If there are three genotypes, A_1A_1, A_1A_2, and A_2A_2, with means and variances in offspring numbers given by $1 + m_i$ and $1 + \sigma_i{}^2$, $i = 1, 2, 3$ respectively, it can be shown (see APPENDIX) that the drift coefficient for the diffusion approximation is

$$M(p) = p(1-p)\,[p(m_1-m_2)-(1-p)(m_3-m_2)$$
$$+ \frac{1}{n}\,((1-p)(\sigma_3{}^2 - \sigma_2{}^2)-p(\sigma_1{}^2 - \sigma_2{}^2))] \qquad (17)$$

and the diffusion coefficient is

$$V(p) = \frac{p(1-p)}{2n}\,[1+m_2 + \sigma_2{}^2 + 2p(1-p)(\sigma_1{}^2+\sigma_3{}^2 - 2\sigma_2{}^2)] . \qquad (18)$$

The drift coefficient is an obvious extension of the haploid case. The diffusion coefficient is somewhat more interesting. It should be viewed as the sum of two components, the first,

$$\frac{p(1-p)}{2n}\,(1 + m_2) , \qquad (19)$$

is a consequence of the randomness introduced by mendelian segregation. The second,

$$\frac{p(1-p)}{2n}\,[\sigma_2{}^2 + 2p(1-p)(\sigma_1{}^2 + \sigma_2{}^2)] , \qquad (20)$$

comes from the randomness due to variable offspring number. If the variances of all three genotypes are equal, and if $m_2 = 0$, we obtain the coefficients of the WRIGHT-KIMURA theory.

Examination of the drift coefficient shows that selection again has the ability to work on variances in offspring numbers. Of particular interest is the possibility of heterozygote advantage resulting from a lower variance in offspring number in the heterozygote. The strength of this selection is, again, inversely proportional to population size, and for very large populations this would be a negligible force for maintaining variations. The probability of fixation of allele A_1 may be written down for the diploid case, although the expression is rather cumbersome and yields little additional insight into the process.

The most remarkable aspects of selection for within-generation variance in offspring number are (1) the necessity of including the population size in the definition of fitness and (2) the appearance of p in the coefficient of $p(1-p)$ in $V(p)$. This latter property requires some comments since, by convention, the effective population size is precisely this coefficient, and is independent of p in the WRIGHT-KIMURA model. In the haploid model the effective size is

$$n_e(p) = \frac{n}{p\sigma_2^2 + (1-p)\,\sigma_1^2}$$

and in the diploid model is

$$n_e(p) = \frac{2n}{1 + m_2 + \sigma_2^2 + 2p(1-p)\,(\sigma_1^2 + \sigma_3^2 - 2\sigma_2^2)}\;.$$

In both cases the most common allele is the least important in determining the effective population size. In the diploid model, when either allele is rare, the heterozygote will always be the genotype of importance in determining the effective population size. By continuity, if one allele is eliminated from the population, the effective size should be determined by this now-absent allele. This rather uncomfortable conclusion suggests that the concept of effective population size loses a good deal of its value in the context of the present model. It remains useful, however, as a description of the gross reproductive pattern of the population.

Selection for the within-generation component of variance in offspring number shares certain properties with selection for the between-generation component. The most striking similarity is in the respective measures of fitness. For the between-generation component, the best measure of fitness of genotype A_i is

$$\mu_i - \frac{1}{2}\,\sigma_i^2$$

(GILLESPIE 1973). The similarity of this expression to (16) suggests that a similar phenomenon is operating in both models. One expression of this phenomenon is that the advantage which a genotype gains through producing many offspring in a good year does not balance the disadvantage from producing few offspring in a bad year. Thus a lowering in the variance in the offspring number, be it the within- or between-generation components, can only raise the probability of leaving offspring behind.

LITERATURE CITED

CROW, J. F. and M. KIMURA, 1970   *An Introduction to Population Genetics Theory.* Harper and Row, New York.

EWENS, W. J., 1972   Concepts of substitutional load in finite populations. Theoret. Population Biol. **3**: 153–161.

FELLER, W., 1951   Diffusion processes in genetics. Proc. 2nd Berkeley Symp. on Math Statistics and Probability 227–246.

GILLESPIE, J. H., 1973   Natural selection with varying selection coefficients—a haploid model. Genet. Res. **21**: 115–120.

KARLIN, S. and J. MCGREGOR, 1964   Direct product branching processes and related Markov chains. Proc. Natl. Acad. Sci. U.S. **51**: 598–602.

KIMURA, M., 1954   Processes leading to quasi-fixation of genes in natural populations due to random fluctuation in selection intensities. Genetics **39**: 280–295. ——, 1962   On the probability of fixation of mutant genes in a population. Genetics **47**: 713–719. ——, 1964   Diffusion models in population genetics. J. Appl. Probability **1**: 177–232.

WRIGHT, S., 1945   The differential equation of the distribution of gene frequencies. Proc. Natl. Acad. Sci. U.S. **31**: 382–389.

APPENDIX

We will begin by deriving the drift and diffusion coefficients for the diploid forward equation corresponding to (1). If, in generation t there are $x_1(t)$ A_1 alleles and $x_2(t)$ A_2 alleles, the numbers of these three genotypes are assumed to be np^2, $n2pq$, and nq^2 where $2n = x_1 + x_2$, $p = x_1/n$, and $q = 1-p$. The mean and variance of the numbers of A_1 and A_2 alleles produced by these genotypes, under the assumption of random mating, may be derived by combining the contributions of the three genotypes considered separately. The mean number of A_1 and A_2 alleles produced will be

$$np^2 (1 + m_1) + pq (1 + m_2)$$
$$nq^2 (1 + m_3) + pq (1 + m_2)$$

so the expected changes in the number of A_1 and A_2 alleles are

$$M_1 (x_1, x_2) = n(p^2 m_1 + pq m_2)$$
$$M_2 (x_1, x_2) = n(pq m_2 + q^2 m_3).$$

These correspond to $\mu_1 x_1$ and $\mu_2 x_2$ in (1). They have been notated using n and p for convenience, although it should be kept in mind that the process being described is (x_1, x_2) and not, at this point, (n,p).

The variance in the number of A_1 alleles produced may also be calculated by considering the contributions of the three genotypes taken separately. For A_1, the contribution from $A_1 A_1$ is $np^2 \sigma_1^2$ and from $A_1 A_2$ is $npq\sigma_2^2/2 + npq (1 + m_2)/2$. The contribution from the heterozygote may be calculated from the generating function of the number of A_1 alleles produced per heterozygote, $F(1/2 + 1/2s)$, where $F(s)$ is the generating function of the number of offspring per heterozygote. This, and the analogous calculations for A_2, lead to the diffusion coefficients.

$$V_1 (x_1, x_2) = n\, p^2 \sigma_1^2 + \frac{npq}{2}\, \sigma_2^2 + npq(1 + m_2)$$

$$V_2 (x_1, x_2) = nq^2 \sigma_3^2 + \frac{npq}{2}\, \sigma_2^2 + npq(1 + m_2).$$

Finally, there is a covariance term which comes from the fact that each gamete produced by a heterozygote is either A_1 or A_2. The covariance in the number of A_1 and A_2 alleles is readily shown, again using generating functions, to be

$$C(x_1, x_2) = \frac{npq}{2} (-(1 + m_2) + \sigma_2^2).$$

M_i, V_i, and C define a diffusion process in x_1 and x_2. Using the transformation (2) this is converted into one in p and n. Holding n constant yields (17) and (18).