

The Genetic Signature of Conditional Expression

J. David Van Dyken¹ and Michael J. Wade

Department of Biology, Indiana University, Bloomington, Indiana 47405

Manuscript received September 23, 2009

Accepted for publication December 1, 2009

ABSTRACT

Conditionally expressed genes have the property that every individual in a population carries and transmits the gene, but only a fraction, ϕ , expresses the gene and exposes it to natural selection. We show that a consequence of this pattern of inheritance and expression is a weakening of the strength of natural selection, allowing deleterious mutations to accumulate within and between species and inhibiting the spread of beneficial mutations. We extend previous theory to show that conditional expression in space and time have approximately equivalent effects on relaxing the strength of selection and that the effect holds in a spatially heterogeneous environment even with low migration rates among patches. We support our analytical approximations with computer simulations and delineate the parameter range under which the approximations fail. We model the effects of conditional expression on sequence polymorphism at mutation–selection–drift equilibrium, allowing for neutral sites, and show that sequence variation within and between species is inflated by conditional expression, with the effect being strongest in populations with large effective size. As ϕ decreases, more sites are recruited into neutrality, leading to pseudogenization and increased drift load. Mutation accumulation diminishes the degree of adaptation of conditionally expressed genes to rare environments, and the mutational cost of phenotypic plasticity, which we quantify as the *plasticity load*, is greater for more rarely expressed genes. Our theory connects gene-level relative polymorphism and divergence with the spatial and temporal frequency of environments inducing gene expression. Our theory suggests that null hypotheses for levels of standing genetic variation and sequence divergence must be corrected to account for the frequency of expression of the genes under study.

IN genetically and ecologically subdivided populations, some individuals will experience a local environment very different from others, making it difficult to evolve a single adaptation adequate for all local conditions. Phenotypic plasticity allows organisms to respond adaptively to spatially and temporally varying environments by developing alternative phenotypes that enhance fitness under local conditions (SCHEINER 1993; VIA *et al.* 1995). Examples of alternative phenotypes, *i.e.*, polyphenisms, include the defensive morphologies in *Daphnia* and algae induced by the presence of predators (*e.g.*, LIVELY 1986; DEWITT 1998; HARVELL 1998; HAZEL *et al.* 2004); the winged and wingless morphs of bean beetles responding to resource variation (*e.g.*, ABOUHEIF and WRAY 2002; ROFF and GELINAS 2003; LOMMEN *et al.* 2005); and bacterial genes involved in traits such as quorum sensing, antibiotic production, biofilm formation, and virulence (FUQUA *et al.* 1996). The developmental basis of such alternative phenotypes often lies in the inducible expression of some genes in some individuals by environmental variables. That is, all individuals carry and transmit the condi-

tionally expressed genes but only a fraction of individuals, ϕ , express them when environmental conditions are appropriate.

The genes underlying plastic traits should experience relaxed selection due to conditional expression. Wade and co-workers have shown that genes hidden from natural selection in a fraction of individuals in the population by X-linked (WHITLOCK and WADE 1995; LINKSVAYER and WADE 2009) or sex-limited expression (WADE 1998; DEMUTH and WADE 2007) experience relaxed selective constraint. In *Drosophila* spp., sequence data for genes with maternally limited expression quantitatively support the theoretical predictions both for within-species polymorphism (BARKER *et al.* 2005; CRUICKSHANK and WADE 2008) and for between-species divergence (BARKER ET AL 2005; DEMUTH and WADE 2007; CRUICKSHANK and WADE 2008). Furthermore, male-specific genes in the facultatively sexual pea aphid have been shown to have elevated levels of sequence variation due to relaxed selection (BRISSEON and NUZHIDIN 2008). Genes with spatially restricted expression in a heterogeneous environment should likewise experience relaxed selection. Adaptation to the most common environment in an ecologically subdivided population (ROSENZWEIG 1987; HOLT and GAINES 1992; HOLT 1996) allows deleterious mutations to accumulate in

¹Corresponding author: Department of Biology, 1001 E. Third Street, Bloomington, IN 47405. E-mail: jdvandyk@indiana.edu

traits expressed in rare environments (KAWECKI 1994; WHITLOCK 1996).

Here we extend these results by quantifying the consequences of relaxed selection on conditionally expressed genes. Specifically, we show that, with weak selection, spatial and temporal fluctuations in selection intensity generate approximately equivalent effects on mean trait fitness, even with low rates of migration between habitats, resulting in a great simplification of analytical results. Our analytical approximations are supported with deterministic and stochastic simulations, and we note the conditions under which the approximations fail. We then derive general expressions for (1) the expected level of sequence polymorphism within populations under mutation, migration, drift, and purifying selection with conditional gene expression; (2) the rate of sequence divergence among populations, for dominant and recessive mutations; and (3) the reduction in mean population fitness due to accumulation of deleterious mutations at conditionally expressed loci. We find that the rate of accumulation of deleterious mutations for conditionally expressed genes is accelerated and the probability of fixation of beneficial mutations is reduced, causing a reduction in the fitness of conditional traits and an inflation in sequence variation within and between species. Our results suggest that evolutionary null hypotheses must be adjusted to account for the frequency of expression of genes under study, such that signatures of elevated within- or between-species sequence variation are not necessarily evidence of the action of diversifying natural selection. Furthermore, if conditional expression is due to spatial heterogeneity, we show that the level of genetic variation in a sample will often depend on whether or not genotypes were sampled from the selective habitat, the neutral habitat, or both. In the DISCUSSION we address the scope and limitations of our theory, as well as its implications for the maintenance of genetic variation, adaptive divergence between species, constraints on phenotypic plasticity, and evolutionary inference from sequence data.

RESULTS

The model

Conditionally expressed genes are defined as those genes in a population not expressed by every individual or in every generation. In particular, we examine three types of conditionally expressed genes: (1) those expressed by only a fraction, ϕ , of all individuals at every generation; (2) those expressed by all individuals but only for a subset of generations, g , over an interval of T generations, where $\phi = (g/T)$; and (3) those expressed by a fraction of the population over a subset of generations in an interval. In each case, the expressed fraction of all gene copies, ϕ , is subject to natural selection but

the unexpressed fraction, $(1 - \phi)$, is affected only by the forces of mutation and drift. As a result, the strength of natural selection acting on conditionally expressed genes is weaker than it is on constitutively expressed genes. Examples of conditionally expressed genes include recessives (expressed only in homozygotes), genes with sex-limited effects (expressed only in one but not in the other sex), and inducible genes (expressed only or predominantly by individuals experiencing the inducing environments). Another type of conditionally expressed gene, “caste genes” expressed only in the sterile workers of the eusocial insects, has been modeled by LINKSVAYER and WADE (2009).

Sequence polymorphism within populations

Expression in a fraction of individuals, ϕ , at every generation: First, we consider the effects of deleterious mutations at a conditionally expressed locus. Throughout, we compare two loci with identical selective and demographic parameters, where one locus is conditionally expressed while the other is constitutively expressed. We begin by considering a haploid population for ease of demonstration, although all methods can be readily extended to diploids. Let there be two alleles, A and a , where A is wild type and a is deleterious. When not expressed, the fitness of the deleterious allele at a conditionally expressed locus, w_a , is equal to 1; *i.e.*, it is neutral because it is not expressed. When the mutation is expressed, its fitness becomes $w_a = 1 - s$, where s is the strength of selection against the deleterious mutation manifest by the decreased viability of those carrying the gene. The average fitness of a conditionally expressed deleterious allele within a population consisting of ϕ expressers and $(1 - \phi)$ nonexpressers is $\bar{w}_a = \phi(1) + (1 - \phi)(1 - s) = 1 - \phi s$. We denote the average selection against the allele as $s_c = \phi s$, where the subscript, c , indicates the selection coefficient associated with the conditionally expressed allele.

Assuming large population size and small s , population genetic theory (*e.g.*, CROW and KIMURA 1970, pp. 58–62) has shown that, at mutation–selection balance, a constitutive deleterious mutant allele expressed in every individual in every generation will have an equilibrium frequency of $\hat{q} = u/s$, where u is the per-locus rate of mutation to the deleterious allele. In contrast, if conditionally expressed, a deleterious allele would have an equilibrium frequency of $\hat{q}_c = u/\phi s$. Hence, the relative equilibrium allele frequency of a conditionally expressed gene to a constitutively expressed paralog is

$$R_p = \frac{(u/s_c)}{(u/s)} = 1/\phi. \quad (1)$$

Polymorphism at mutation–selection balance is $\hat{q}_c(1 - \hat{q}_c)$, which is also approximately $(1/\phi)$ if we ignore terms of order $(\mu/s)^2$ (see below). By a similar derivation, it is

easily shown that Equation 1 also holds for diploids with arbitrary dominance.

Expression by all individuals but only in a fraction, ϕ , of all generations: Let ϕ be the fraction of generations, g , over an interval of T generations when the conditional gene is expressed ($\phi = g/T$). With discrete generations, the average fitness of such an allele, w_G , equals its geometric mean fitness (CROW and KIMURA 1970; FRANK and SLATKIN 1990), which can also be seen from the multiplicative gene frequency recursion. Thus, we find that $w_G = (1 - s)^\phi (1)^{1-\phi} = (1 - s)^\phi$. This is equivalent to an average, per-generation selection coefficient, s_c , equal to $1 - (1 - s)^\phi$.

Following the same logic as above, we find that the relative equilibrium allele frequency of a conditionally expressed gene to a constitutively expressed paralog is

$$R_p = \frac{(u/s_c)}{(u/s)} = \frac{s}{1 - (1 - s)^\phi}. \quad (2)$$

Numerical investigation shows that R_p is much more sensitive to ϕ than to s . That is, varying s by many orders of magnitude ($1 \times 10^{-11} < s < 1 \times 10^{-1}$) results in a change in R_p of only 5% for any of a wide range of ϕ values (Figure 1A). Analytically, we can assume that s is small, so that $(1 - s)^\phi \sim 1 - s\phi$. With this assumption, we see that $s_c \sim s\phi$, as above, allowing us to reduce Equation 2 to

$$R_p \sim \frac{1}{\phi}. \quad (3)$$

Our result conforms to the intuitive expectation that a conditional trait expressed half the time over an interval of T generations ($\phi = 0.50$), or in half of the individuals all of the time, will accumulate twice as much mutational variation ($R_p \sim 1/0.50 = 2$) as a constitutively expressed trait under approximately the same strength of selection and mutation pressure (Figure 1A).

Expression in some individuals, some of the time: We can combine the two cases above into a single expression if we distinguish the two separate contributions to ϕ , as ϕ_I (the fraction of individuals within a generation) and ϕ_G (the fraction of generations). In this combined case, the average, per-generation selection coefficient, s_c , equals $1 - (1 - s\phi_I)^{\phi_G}$. When s is assumed to be small, $(1 - s\phi_I)^{\phi_G} \sim 1 - s\phi_I\phi_G$.

Here, the relative equilibrium allele frequency of a conditionally expressed gene to a constitutively expressed paralog is

$$R_p \sim \frac{1}{\phi_I\phi_G}. \quad (4)$$

For example, the standing polymorphism of genes with male-limited expression in facultatively sexual species, such as *Caenorhabditis elegans* (CHASNOV and CHOW 2002, SIEDEL *et al.* 2008) and the pea aphid, *Acrythosi-*

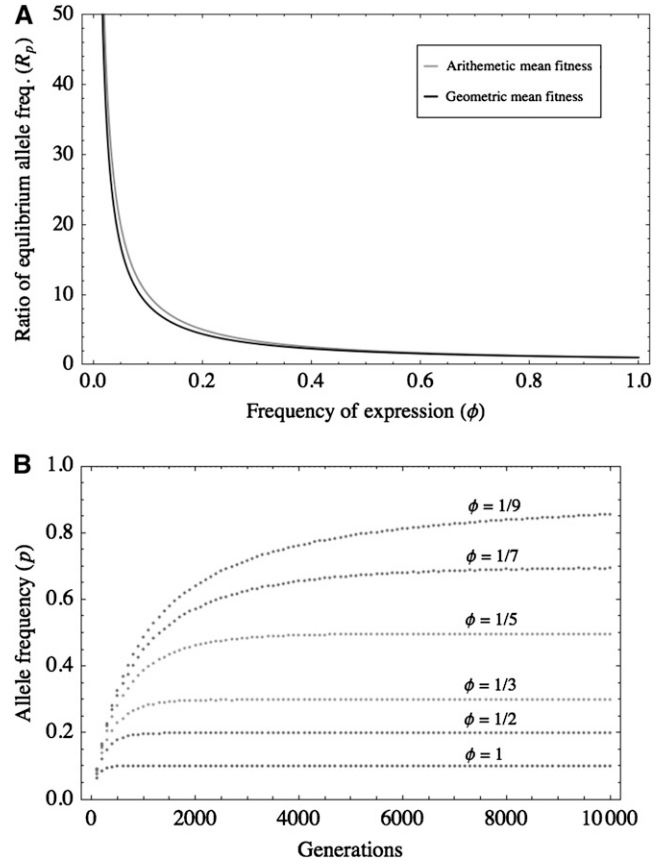


FIGURE 1.—(A) The relative increase in equilibrium frequency of deleterious mutations at mutation–selection balance that results from conditional expression (ϕ). The gray line represents Equation 1, which depends on the arithmetic mean fitness of an allele between expressed and nonexpressed states. The curve for Equation 1 is independent of selection intensity or mutation rate. The black line is Equation 2, the equilibrium allele frequency ratio based on the geometric mean fitness of the conditional locus. Although the analytical approximation that makes these two curves approximately equal required the assumption of small s , the two curves are approximately equivalent even for much larger s , as is shown in the figure with $s = 0.3$. (B) Simulation results varying s between $s = 0$ and $s = -0.01$ at frequency ϕ with recurrent one-way deleterious mutation rate $u = 0.001$. A population with a rare deleterious mutation eventually reaches mutation–selection balance at the equilibrium allele frequency approximately equal to $u/\phi s$, as predicted by theory (see text).

phon pisum (BRISSON and NUZHIDIN 2008), might well be represented by Equation 4. CHASNOV and CHOW (2002) discuss the poor mating efficiency of males in the hermaphroditic species, *C. elegans*, relative to dioecious species. They provide theoretical support for the hypothesis that deleterious mutations have accumulated in those genes with male-limited expression and this is responsible for poor male-mating efficiency. They argue further that even a very low frequency of functional males “can support male-specific genes against mutational degeneration” even in a species like *C. elegans* that is 99.9% hermaphroditic. Our Equation 4 shows that the relative degree of mutational degradation of genes with

sex-limited or environment-limited expression can be quantified and equals a simple function of ϕ . For example, imagine that males appear only once every five generations ($\phi_G = 0.20$) and, when males do appear, they represent only 5% of the population ($\phi_I = 0.05$). Thus, relative to a constitutively expressed gene that is expressed in each individual, male and female, at every generation, we would expect a gene with male-limited expression in *C. elegans* to be at least 100 times more polymorphic at equilibrium ($1/\phi_I\phi_G = [1/.20][1/0.05] = 100$). No adaptive explanation for high diversity in terms of frequency-dependent sexual selection or sex-ratio selection is necessary in this case. Or, differently put, polymorphism alone is not convincing evidence of frequency-dependent sexual selection; only polymorphism levels that exceed ($1/\phi_I\phi_G$) (*i.e.*, 100 times normal in this example) are evidence for such selection. In fact, it may well be the case that species of *Caenorhabditis* with infrequently expressed male genes are at risk of losing male function to mutation and relaxed selection as CHASNOV and CHOW (2002) suggested. Expression (4) may be useful in identifying a quantitative threshold of gene expression sufficient to engender the loss of males.

In all the cases above, the generations of neutrality experienced by conditionally expressed genes necessarily slow the removal of deleterious mutations by natural selection. Computer simulations show that periodic neutrality leads to equilibrium allele frequencies predicted by our analytical approximation (*i.e.*, Equations 1 and 3), supporting our analytical results (Figure 1B; see METHODS). Significantly, Equation 2 is equivalent to the results obtained investigating maternal-effect genes, using an arithmetic mean selection coefficient of ($s/2$), as is appropriate for a trait expressed in half of the individuals within a single generation (*i.e.*, the females) (WADE 1998; DEMUTH and WADE 2007). Thus, Equation 2 brings conditional gene expression between individuals within a single generation and conditional gene expression between generations into a single, general framework. However, an explicitly spatial model of selection on a conditionally expressed trait is required to validate this result.

Migration in a spatially heterogeneous environment:

When the cause of conditional expression is spatial heterogeneity in the distribution of trait-inducing cues or of selection pressures across the environment, a spatially explicit model is required to accurately specify the dynamics of natural selection. Two types (or “grains”) of spatial heterogeneity exist, fine grained and coarse grained (LEVINS 1968). In a fine-grained environment, individuals experience multiple environments throughout their lifetime, while in a coarse-grained environment an individual completes its entire life cycle in a single habitat (LEVINS 1968). LEVINS (1968) showed that selection in a fine-grained environment is governed by the arithmetic mean selection coefficient over habitat

types, while in a coarse-grained environment selection is governed by the geometric mean selection coefficient. Alternatively, NAGYLAKI (1980) showed that, in a (coarse-grained) spatially heterogeneous environment with migration between habitat patches, the outcome of natural selection is determined by the arithmetic mean selection coefficient over individuals in the population as long as migration between habitats is sufficiently high (*e.g.*, the “high migration limit”).

As we showed above, the geometric and arithmetic mean fitnesses are approximately equivalent for conditionally expressed traits, meaning that we do not necessarily need to make explicit assumptions about the pattern of environmental heterogeneity. The question, then, becomes at what value for migration is the high-migration limit achieved for conditionally expressed genes?

To answer this, we consider for simplicity a haploid population divided into two subpopulations, S and N, representing selective and nonselective habitats, respectively. The subpopulations are composed of different numbers of individuals, N_S and N_N , such that ϕ is defined as the fraction of the total population occupying the selective habitat, $\phi = N_S/(N_S + N_N)$. The life cycle of an individual consists of migration \rightarrow selection \rightarrow reproduction/mutation, with selection occurring only in habitat S. We assume irreversible, recurrent mutation from the wild-type to the deleterious allele in accordance with the infinite sites model (WATTERSON 1975). For simplicity, we also ignore the effects of drift.

We assume that migration is conservative, meaning that the number of migrants sent out from each subpopulation in each generation is equal, *e.g.*, when organisms assort in space according to the ideal free distribution (FRETWELL 1972; RICE 2004). In the current model, this would be the case if the subpopulation experiencing selection is occupying a marginal habitat. Letting m_1 equal the fraction of subpopulation S composed of immigrants from subpopulation N in each generation (*e.g.*, the “backward migration rate”), we have, $N_S m_1 = N_N m_2$. Defining $\psi = N_N/N_S$ gives $m_1 = \psi m_2$. The parameter ψ is related to ϕ by $\phi = 1/(1 + \psi)$ or, equivalently, $\psi = (1/\phi) - 1$. Noting that the mean fitness in the selective habitat is $\bar{w}_S = 1 - sq_S$, at migration–mutation–selection equilibrium, the frequency of deleterious mutant alleles in the selective and neutral subpopulations, respectively, is exactly

$$\hat{q}_S = \frac{u(m_1 + m_2) + u^2}{s(m_2 + u(m_1 + m_2) + u + u^2)} \quad (5a)$$

$$\hat{q}_N = \frac{m_2 \hat{q}_S + u}{m_2 + u} \quad (5b)$$

Ignoring terms of second order or higher in s and u , substituting ψm_2 for m_1 , and using the relation $\phi = 1/(1 + \psi)$, we find that

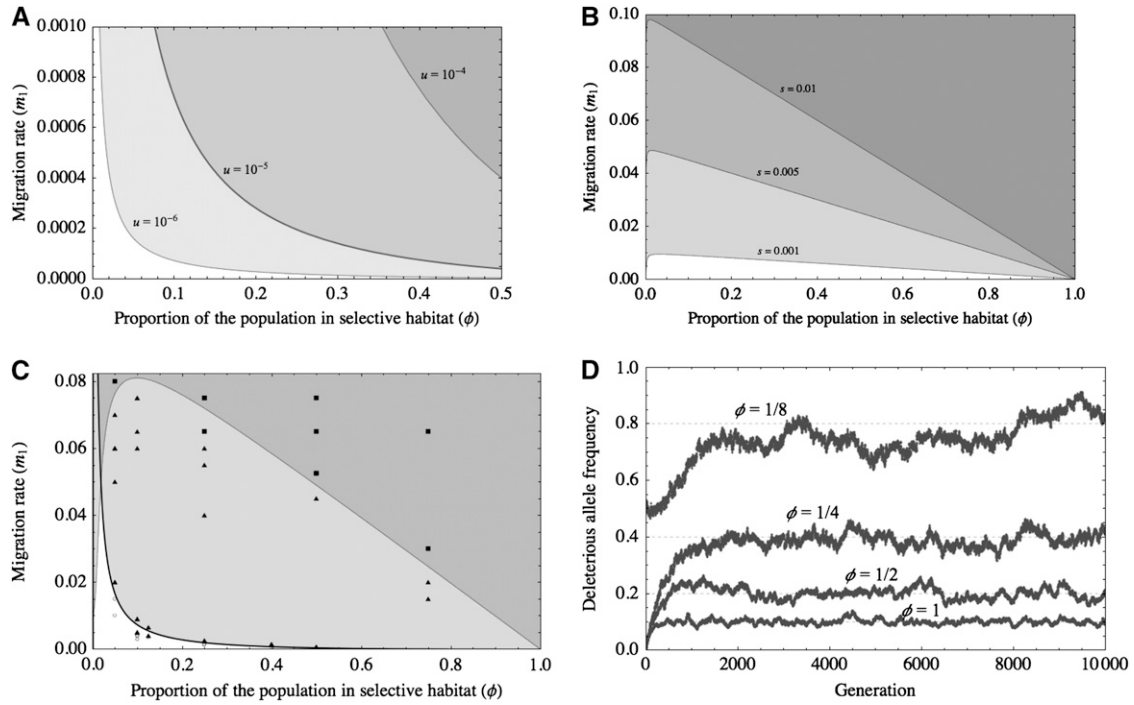


FIGURE 2.—(A) The parameter space where the approximation leading to Equation 6a is valid within 10% of the exact result. Each curve is independent of the selection coefficient and depends only on the mutation rate, u . Equation 6a is accurate for parameters lying in the space above each curve. (B) Equation 6b is accurate to within 10% of the exact result for all parameters lying in the area above each curve. For all curves, $u = 10^{-6}$. (C) The graph shows three regions: in the region above the top curve, both Equations 6a and 6b are valid within 10% of exact results; in the area between the two curves, Equation 6a is valid but Equation 6b is not; and in the area below the bottom curve, neither Equation 6a nor 6b is valid. The squares show parameter values (m_1 and ϕ) where simulation results found that Equation 6a and 6b were valid; the triangles denote parameter coordinates where simulations found Equation 6a to be accurate, but 6b to be inaccurate; and the open circles show parameter values where simulations found both Equation 6a and 6b to be inaccurate. (D) Sample outputs from individual simulations with parameter values $s = 0.01$ and $u = 0.001$. The horizontal dashed lines indicate the theoretical expectation (Equation 6a) of the respective value of ϕ , while the jagged curves are the deleterious allele frequency trajectories for the simulation (see text for details).

$$\hat{q}_S \approx \frac{u}{\phi s}. \tag{6a}$$

Assuming that $m_2 q_S \gg u$, Equation 5b gives,

$$\hat{q}_N \approx \hat{q}_S. \tag{6b}$$

Thus, with conservative migration between selective and nonselective habitats, the equilibrium frequency of deleterious mutants in the total population is approximately equal to the value we find for conditionally expressed genes in our unstructured models.

Equations 5a and 5b can be used to determine the conditions under which the analytical approximation leading to Equations 6a and 6b hold. This is done by choosing an arbitrary accuracy threshold, say 10%, and then setting the ratios (Equation 5a/Equation 6a and Equation 5b/Equation 6b) equal to 0.9 or 1.1 (depending on whether the approximate result under- or overestimates the exact result, respectively) and solving for m_1 . The result is a curve that dissects the parameter space into a region where the approximation is valid to within our chosen threshold and a region where it is not. These curves are independent of the selection coefficient

because s cancels when dividing Equation 5a by Equation 6a. Figure 2A plots the range of parameter values under which Equation 6a holds to within 10% of the exact result of Equation 5a for $u = 10^{-4}, 10^{-5}, 10^{-6}$. The approximation is valid over a wider parameter range when the mutation rate is low, and in general, the approximation of Equation 6a is valid even when migration rates are quite small (Figure 2A). While the accuracy of Equation 6a is independent of the selection coefficient, this is not true for Equation 6b. Approximation (6b), however, is less sensitive to mutation rate and more sensitive to the value of the selection coefficient. Figure 2B shows how the range of parameter values under which Equation 6b holds is greater for smaller selection coefficients. Numerical investigation shows that Equation 6b will hold to within $\eta\%$ as long as $m_1 > s/\eta$, although the exact value depends on ϕ (see Figure 2B).

Figure 2C shows the range of parameter values under which both Equations 6a and 6b hold to within 10% of the exact result of Equations 5a and 5b for $s = 0.01$ and $u = 10^{-4}$. This figure is divided into three regions: in the region below the lower curve both approximations fail; between the curves, Equation 6a holds but 6b does not;

and above the upper curve both hold to within 10% of the exact results. The parameter range where Equation 6a holds is much more extensive than that where Equation 6b holds (Figure 2C), suggesting that in nature there will often be substantial differences in allele frequencies measured between habitat types. Thus, the outcome of genetic analysis of samples from natural populations will be sensitive to the sample locations across a selectively variable environment.

It can be seen from Figures 2, A–C, that, as ϕ gets smaller, the above approximation requires higher migration rates to be valid. In addition, Figures 2A and 2C show that the parameter range for Equation 6a to hold is smaller for higher mutation rates. With realistic per-gene rates of deleterious mutation ($u = 10^{-6}$), the approximation holds for a large range of ϕ values even when migration is quite limited (clearly seen in Figure 2B). Significantly, KAWECKI (1994) derived Equation 6a by a different route, but under the assumption of complete random migration in every generation ($m = 0.5$); our results validate Equation 6a for a less restrictive parameter range with conservative migration.

Stochastic simulations testing the validity of our analytical results

To test the validity of the approximations leading to Equations 6a and 6b, we conducted extensive individual-based stochastic simulations of a population subdivided into selective and neutral subpopulations with migration between them (Figure 2B; see METHODS). Sample outputs of individual simulation runs are given in Figure 2D. Figure 2D plots the outcome of individual runs for $\phi = 1, \frac{1}{2}, \frac{1}{4}$, and $\frac{1}{8}$ with parameter values $u = 0.001$, $s = 0.01$, $m_1 = 0.1$, and $N_T = 50,000$.

Simulation results are given in Table 1, and plotted in Figure 2C (see below). Three simulations were run for each set of parameter values, and the mean and standard error of these runs are reported in Table 1. We note that the standard deviations of allele frequency are larger for lower ϕ values (and that the fluctuations of allele frequencies in Figure 2D are greater for lower values of ϕ). This is a consequence of the smaller census size of the selected subpopulation when ϕ is small, causing drift with respect to natural selection to become stronger. We chose parameter values that specifically test the validity of the curves in Figure 2C: (1) open circles are parameter values (m_1 and ϕ) where Equations 6a and 6b did not hold; (2) solid triangles indicate where Equation 6a but not Equation 6b holds; and, (3) solid squares indicate where both hold. The simulation results match the analytical predictions, giving robust support for our approximations.

Effect of conditional expression on the distribution of allele frequencies under drift, mutation, and selection: We now introduce random genetic drift into our deterministic analysis and extend the model to diploidy to explicitly account for dominance. With drift the equilib-

rium state of the population is a probability distribution of allele frequencies, rather than a single equilibrium allele frequency as above. The distribution of allele frequencies under mutation, drift, and natural selection is given by the diffusion approximation (KIMURA 1962)

$$\hat{\phi} = \frac{e^{\int (2M(x)/V(x)) dx}}{V(x)}. \quad (7)$$

The terms $M(x)$ and $V(x)$ in Equation 7 represent the mean and variance in allele frequency change per generation, respectively. Expressions for $M(x)$ and $V(x)$ can be obtained using standard population genetic formulas (e.g., CROW and KIMURA 1970). $M(x)$ incorporates directional causes of allele frequency change such as selection and mutation, while $V(x)$ incorporates nondirectional sources such as drift or temporal variation in selection intensity (KIMURA 1962; EWENS 1979). Following KIMURA (1962), we define

$$M(x) = -\phi sx(1-x)(h+x(1-2h)) - u_1 x + u_2(1-x) \quad (8a)$$

$$V(x) = V_s x^2(1-x)^2 + x(1-x)/2N_e, \quad (8b)$$

where x is mutant allele frequency, u_1 and u_2 are mutation rates from and to the deleterious allele, respectively, h is degree of dominance (for $h = 0$ the mutation is recessive, $h = 1$ dominant), V_s is the variance in the selection coefficient s , and N_e is the effective population size. The first term on the righthand side of Equation 8b accounts for variance in allele frequency change owing to variance in selection intensity, while the second term accounts for the effects of random drift. For conditionally expressed genes,

$$V_s = s^2\phi(1-\phi). \quad (9)$$

Even for relatively strong selection (e.g., $4N_e s > 1$), s^2 will be on the order of $1/N_e^2$, and the product $\phi(1-\phi) \ll 1$ for many values of ϕ . Given this, the first term of Equation 8b is negligible and is thus omitted throughout for simplicity without appreciable loss in precision. Substituting Equation 8 into Equation 7, setting $h = \frac{1}{2}$ (e.g., the mutation has additive effects on fitness), and evaluating x at q , we find,

$$\hat{\phi} = C e^{-2N_e s \phi q} (1-q)^{4N_e u_1 - 1} q^{4N_e u_2 - 1}, \quad (10)$$

where C is a normalizing constant that ensures that the area under the distribution equals one.

Equation 10 suggests that conditional expression increases the strength of genetic drift at a locus (Figure 3). When N_e is small (Figure 4A), conditional expression flattens the allele frequency distribution consistent with increased neutrality so that there is greater probability that deleterious alleles can reach high frequency or even fix in the population owing to conditional

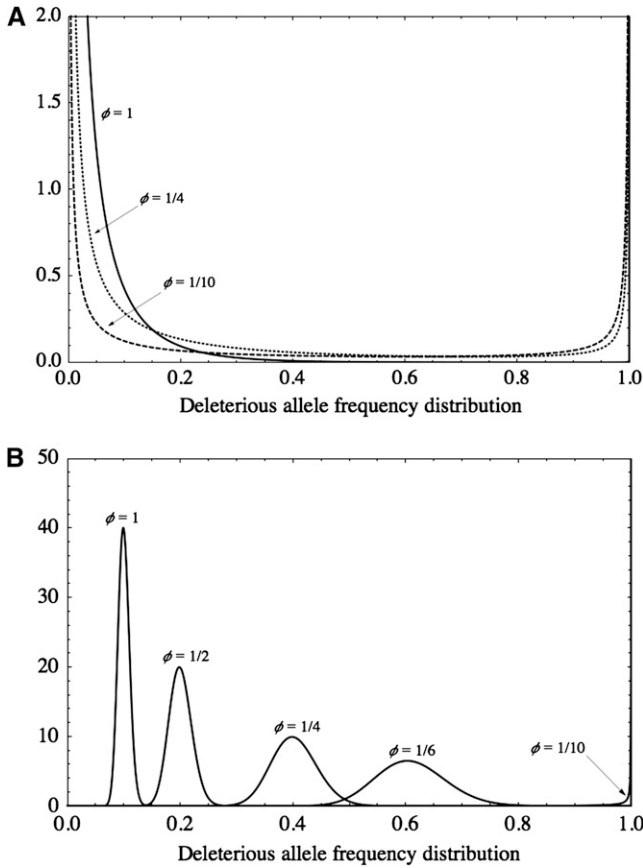


FIGURE 3.—(A) The distribution of deleterious allele frequency under one-way recurrent mutation, purifying selection, and drift when $N_e = 10^4$, $s = 10^{-3}$, and $u = 10^{-5}$. When expressed constitutively ($\phi = 1$), selection is strong enough to keep the deleterious mutation rare, but as expression becomes more infrequent, drift overcomes selection and deleterious mutations reach high frequency and fixation. (B) The distribution of deleterious allele frequency under one-way recurrent mutation, purifying selection, and drift with stronger selection ($N_e = 10^7$, $s = 10^{-4}$, and $u = 10^{-5}$), such that the distributions center about the deterministic expectation. When $\phi = 1$, the deterministic expectation is $q = 0.1$. Conditional expression shifts the distribution to the right such that the new expectation is $1/\phi$ times 0.1, as predicted by the deterministic theory (e.g., Equation 1). It can also be seen that the variance increases with lower ϕ .

expression. With large N_e (Figure 4B), the distribution shifts to the right with decreasing ϕ and is centered about the deterministic expectation, independently corroborating the deterministic results (Equation 6).

DNA sequence polymorphism under conditional expression: We need to extend the theory to use sequence data from natural populations to test its predictions. Sequence data can be used to measure the nucleotide diversity parameter, π (NEI and LI 1979; TAJIMA 1983), a commonly used measure of nucleotide polymorphism. π equals $\sum 2q_i(1 - q_i)$, where q_i is the frequency of an allele at site i , and the sum is taken over all polymorphic sites. For rare and deleterious mutations, q_i is small, and ignoring terms of order q_i^2 gives the standard approximate result, $\pi \sim \sum 2q_i$. With the approximation

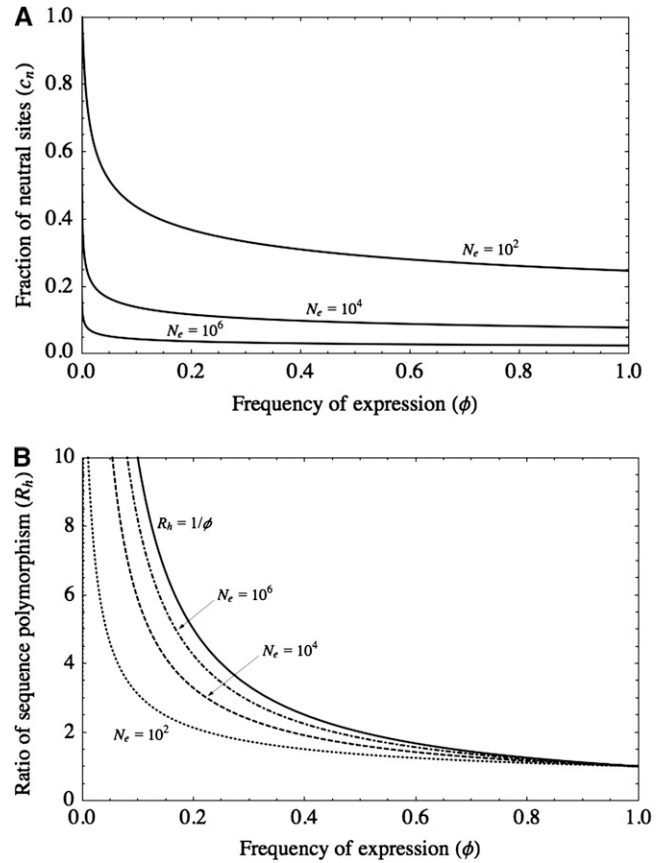


FIGURE 4.—(A) The fraction of nonsynonymous sites at conditional loci that evolve neutrally for populations with different effective population sizes. The value of c_n is determined by Equation using a Gamma DFE with shape parameter $a = 0.25$. There are more neutral nonsynonymous sites when N_e is smaller and/or when ϕ is smaller. (B) The ratio (R_h) of conditional to constitutive nonsynonymous sequence diversity with neutral sites for $\rho = 10$. R_h deviates more from the deterministic expectation (solid line) for small populations, although the approximation is quite good for large populations. R_h can be greater than ρ , the ratio when the conditional locus is evolving entirely neutrally, for small values of ϕ because selection prevents alleles from fixing, allowing variation to increase beyond that of the neutral expectation.

$s_c \sim s\phi$ and the standard assumptions of large population size and independence among sites, we can substitute the equilibrium allele frequency under mutation–selection balance (i.e., Equation 6) for q_i to find the expected replacement site diversity, π_{AC} , for a conditionally expressed trait at equilibrium: $\pi_{AC} \sim (1/\phi) \sum 2u/s$. Under the same assumptions, the expression for a constitutively expressed gene is $\pi_A \sim \sum 2u/s$. Taking the ratio of diversities, R_h , we again obtain Equation 2:

$$R_h \sim \frac{1}{\phi}. \tag{11}$$

However, Equation 11 will not hold for the ratio of polymorphism when there is drift, when some nonsynonymous sites evolve neutrally, or when ϕ becomes very small. The latter case occurs because, as ϕ becomes small,

$q \rightarrow 1$, which clearly violates the assumption of small q . Although a full treatment of the effects of ϕ on nucleotide diversity in a finite population is beyond the scope of this article, we can derive the following general results.

The expected replacement (*i.e.*, nonsynonymous) site nucleotide diversity at a locus under selection and drift is given by

$$\pi_A = c_n \theta + (1 - c_n) H_P, \quad (12)$$

where c_n is the fraction of neutral sites, $\theta = 4N_e u$ (in a diploid population), and H_P is the mean equilibrium diversity at nonneutral sites (LOEWE *et al.* 2006). Neutral sites are those sites where $-1/(4N_e) < s < 1/(4N_e)$ (OHTA and KIMURA 1971). In the case of a conditionally expressed trait, this inequality becomes $-1/(4N_e \phi) < s < 1/(4N_e \phi)$. Since ϕ is a fraction between 0 and 1, it is clear that a conditionally expressed gene will always experience random genetic drift (*i.e.*, approximate neutrality) more strongly than a constitutively expressed gene. Put differently, conditional expression increases the neutral range of mutations by reducing N_e to ϕN_e . The increased strength of drift with lower values of ϕ is seen in stochastic simulations, manifesting as an increase in the variance of mean allele frequency across replicated simulations (Table 1).

If we define $\Psi(s)$ as the distribution of fitness effects (DFE) of new mutations (EYRE-WALKER and KEIGHTLEY 2007), then the fraction of neutral sites for a conditionally expressed trait is given by the equation

$$c_n = \int_{-1/(4N_e \phi)}^{1/(4N_e \phi)} \Psi(s) ds. \quad (13)$$

In the limit as $\phi \rightarrow 0$, $1/(4N_e \phi) \rightarrow \infty$, and $-1/(4N_e \phi) \rightarrow -\infty$, giving $c_n = 1$ in the limit by the definition of a probability distribution. Thus, as ϕ becomes small, the nonsynonymous site polymorphism at a locus approaches the expectation for complete neutrality, $\pi_A = \theta = 4N_e u$. As expected, conditional expression increases the neutral range of mutations, and in the limit, genes that are never expressed should become pseudogenized by mutation.

Unlike the deterministic case, the upper limit on polymorphism set by the neutral expectation ($\pi_A = \theta$) prevents the ratio of polymorphisms of conditional to constitutive genes (R_h) from increasing without bound as ϕ becomes small. We now derive an equation for R_h that takes this into account. We consider the case where there are two classes of nonsynonymous sites: neutral and selected (this is the most extreme case, because it results in the highest variance in s among sites, and thus places an upper bound on the possible effect of V_s). As selection is relaxed by conditional expression, previously selected sites will become neutral (Equation 13) and will have a level of polymorphism equal to the neutral expectation, θ . We define π_s as the polymorphism at sites under selection due to mutation-

selection balance, c_{n1} and c_{n2} are the proportion of neutral nonsynonymous sites at conditional and constitutive loci, respectively, and $\rho = \theta/\pi_s$. With the assumption of two discrete types of nonsynonymous sites, $H_P = \sum u_i/\phi s_i = (1/\phi) \pi_s$. Applying these assumptions and definitions to Equation 12, taking the ratio of conditional π_A to constitutive π_A , and rearranging gives the relative polymorphism of a conditional gene,

$$R_h = \frac{1/\phi + c_{n1}(\rho - 1/\phi)}{1 + c_{n2}(\rho - 1)}. \quad (14)$$

When $c_{n1} = 1$ (*e.g.*, all nonsynonymous sites at the conditional locus evolve neutrally), Equation 11 reduces to $R_h = \theta/\pi_A$ as expected (where π_A is given by Equation 12).

The effect of ϕ on Equation 14 depends on the DFE (through the parameters c_{n1} and c_{n2}) and the relative magnitude of π_s and θ . To investigate Equation 14, we define the DFE as a Gamma-distribution with shape parameter $\alpha = 0.25$, which accords with empirical results from a number of species (EYRE-WALKER *et al.* 2006; LOEWE *et al.* 2006; EYRE-WALKER and KEIGHTLEY 2007). Figure 4A shows that the fraction of neutral nonsynonymous sites at a conditionally expressed locus (c_{n1}) depends on N_e , as expected (see Equation 10), and that it is sensitive to ϕ only when ϕ is small. Figure 4B shows that the deterministic value of R_h (Equation 11) is a good approximation of Equation 14 when N_e is large, but that it consistently overestimates R_h for smaller N_e . The degree of overestimation depends on N_e because with small N_e drift is stronger at selected sites causing π_A to approach θ , such that $R_h \rightarrow 1$ in the limit when drift overwhelms selection at all nonsynonymous sites. For realistic population sizes ($N_e > 10^3$), however, Equation 11 is a reasonable approximation for the effect of conditional expression on sequence polymorphism.

Sequence divergence between populations or species

We extend our results for polymorphism within species to sequence divergence among species, which is determined by the fixation process. The probability of fixation of an allele in a population experiencing selection and drift is given by the diffusion approximation (KIMURA 1962)

$$u(p) = \frac{\int_0^p G(x) dx}{\int_0^1 G(x) dx}, \quad (15)$$

where

$$G(x) = e^{-2 \int M(x)/V(x) dx}. \quad (16)$$

The terms $M(x)$ and $V(x)$ are given above by Equation 8. As above, we assume that V_s is negligible because it is typically many orders of magnitude smaller than drift variance.

TABLE 1

Results for stochastic, individual-based computer simulations of a population subdivided into selective and nonselective habitats

ϕ	m_1	q_S	q_N	% difference between q_S and q_N	Observed q_S within %10 of expectation	$q_S \sim q_N$
0.05	0.01	0.159 (0.018)	0.282 (0.026)	77.2		
	0.015	0.154 (0.017)	0.239 (0.025)	54.8		
	0.02	0.209 (0.013)	0.283 (0.017)	35.3	X	
	0.05	0.180 (0.005)	0.205 (0.005)	14.1	X	
	0.06	0.194 (0.008)	0.215 (0.008)	10.8	X	
	0.07	0.189 (0.001)	0.209 (0.001)	10.6	X	
	0.08	0.198 (0.005)	0.216 (0.005)	9.0	X	X
	0.09	0.183 (0.006)	0.1970 (0.006)	7.7	X	X
0.1	0.003	0.072 (0.005)	0.255 (0.016)	256.2		
	0.004	0.069 (0.003)	0.204 (0.006)	196.6		
	0.005	0.106 (0.005)	0.123 (0.020)	158.1	X	
	0.009	0.090 (0.003)	0.174 (0.007)	94.0	X	
	0.05	0.106 (0.006)	0.123 (0.007)	16.5	X	
	0.065	0.101 (0.006)	0.113 (0.006)	11.6	X	
	0.075	0.093 (0.004)	0.103 (0.005)	10.3	X	
0.125	0.0035	0.051 (0.001)	0.172 (0.002)	238.5		
	0.004	0.073 (0.006)	0.224 (0.020)	205.9	X	
	0.0065	0.080 (0.006)	0.165 (0.011)	107.7	X	
0.25	0.0015	0.030 (0.003)	0.204 (0.006)	573.7		
	0.0025	0.043 (0.013)	0.161 (0.018)	271.3	X	
	0.04	0.043 (0.002)	0.050 (0.002)	18.0	X	
	0.055	0.040 (0.002)	0.045 (0.002)	12.6	X	
	0.06	0.040 (0.001)	0.045 (0.001)	10.5	X	
	0.065	0.039 (0.001)	0.042 (0.001)	9.0	X	X
	0.075	0.038 (0.001)	0.042 (0.001)	9.6	X	X
0.4	0.00035	0.018 (<0.001)	0.221 (0.013)	1127.9		
	0.0015	0.023 (0.002)	0.215 (0.060)	834.0	X	
0.5	0.0001	0.013 (0.001)	0.340 (0.037)	2505.2		
	0.0007	0.018 (0.001)	0.117 (0.013)	564.3	X	
	0.045	0.020 (0.001)	0.022 (0.001)	10.2	X	
	0.05	0.019 (<0.001)	0.021 (<0.001)	9.3	X	X
	0.065	0.021 (0.001)	0.022 (0.001)	7.0	X	X
	0.075	0.020 (0.001)	0.022 (0.001)	5.8	X	X
	0.75	0.015	0.014 (<0.001)	0.016 (<0.001)	14.5	X
0.02		0.013 (<0.001)	0.014 (0.001)	11.4	X	
0.03		0.013 (<0.001)	0.014 (<0.001)	6.4	X	X
0.065		0.013 (<0.001)	0.013 (<0.001)	2.3	X	X

Parameter values for the simulation are $s = 0.01$ and $u = 0.001$. ϕ is the proportion of individuals occupying the selective habitat, q_S is the frequency of deleterious mutants in the selective habitat, q_N is the frequency of deleterious mutants in the neutral habitat, and m_1 is the proportion of individuals in the selective habitat replaced by migrants from the neutral habitat every generation. The values for q_S and q_N are the mean allele frequency averaged over three replicated runs, and the standard error over these runs is given in parentheses. The “expectation” in the second column from the left is the approximate result given by Equation 6a, which is $q_S \sim u/\phi s$. See text for details.

We combine Equations 8a, 8b, 15, and 16 and ignore terms of order q^2 because $q = \frac{1}{2}N$ for a new mutation (RICE 2004). Making this approximation and setting $q = \frac{1}{2}N$, we find that the probability of fixation of a new conditionally expressed mutation is approximately

$$u_c(p) = \frac{1 - e^{-2hs\phi}}{1 - e^{-4N_ehs\phi}} \tag{17}$$

Equation 17 was also found by WHITLOCK (1996). However, this is an approximate result that fails for

completely recessive mutations (where $h = 0$). We obtain exact results by numerical integration of Equation 15 to determine the effect of dominance on divergence of conditionally expressed loci. We are interested in the *relative* probability of fixation for a conditionally expressed mutation, which is the ratio of fixation probabilities of conditional to constitutive mutations. This ratio, R_F , is plotted against ϕ for dominant ($h = 1$), additive ($h = \frac{1}{2}$), and recessive ($h = 0$) mutations for constant values of N_e and s (see Figure 5A). Figure 5 demonstrates that degree of dominance has little effect

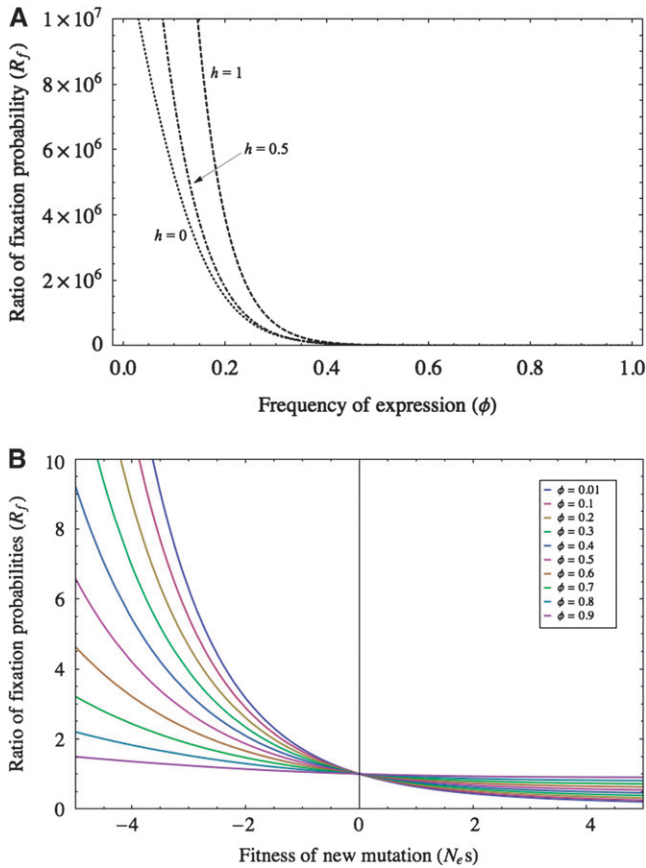


FIGURE 5.—(A) Numerical integration of Equation 15 is used to determine the effect of dominance on the ratio of fixation probabilities R_f . Conditional expression has a stronger effect on dominant than additive or recessive mutations when ϕ is small, although the effects of each are within an order of magnitude for all values of ϕ . (B) The relative probability of fixation of dominant mutations with fitness effect $N_e s$ for different frequencies of gene expression, ϕ . Mutations with deleterious fitness effects have a significantly inflated probability of fixation under conditional expression, corresponding to large values of R_f , leading to increased divergence between species at conditional loci. For beneficial mutations, however, the probability of fixation is reduced. When $N_e s$ is large and positive, the probability of fixation is $2s\phi$, giving a value of $R_f = \phi$. Thus, adaptive evolution is slowed by conditional expression, but the rate of fixation of deleterious mutations is increased.

on the relative fixation probability when $\phi > 0.5$. However, when ϕ is small, dominant mutations are disproportionately affected by conditional expression because deleterious recessive mutations are always hidden from selection when rare, whereas dominant mutations are not. Conditional expression, then, provides a selective reprieve for dominant mutants that they normally do not receive, making them more sensitive to infrequent expression.

Figure 5B plots the ratio of fixation probabilities against the strength of selection ($N_e s$) for the case of dominant mutations ($h=1$) and shows that conditional expression greatly increases the probability of fixation of deleterious mutations. On the other hand, the spread

of beneficial mutations ($s > 0$) in conditionally expressed traits is slowed (see also WHITLOCK 1996). To show this latter result analytically, we assume that $N_e = N$ as per HALDANE (1927) and take the limit of Equation 14 as $N_e \rightarrow \infty$. We find the probability of fixation of a dominant, beneficial mutation ($s > 0$) is (for non-infinitesimal ϕ)

$$u_c(p) \sim 2s\phi \quad (18)$$

analogous to HALDANE's (1927) classic result that the probability of fixation is twice the selective benefit. Taking the ratio of Equation 18 to the classic result of HALDANE (1927), who assumed constant s , we find

$$R_{f, \text{beneficial}} = \phi. \quad (19)$$

That is, conditionally expressed genes suffer a reduction in probability of fixation directly proportional to ϕ , the frequency of gene expression (Figure 5B), slowing the rate of adaptive divergence between species.

The combined effect of increased probability of fixation of deleterious mutations and a decreased probability of fixation of beneficial mutations should be a net increase in overall sequence divergence between species at conditionally expressed loci, since the former is believed to be orders of magnitude greater than the latter. The logic is as follows: the relative probability of fixation of slightly deleterious alleles increases *exponentially* with decreasing ϕ (Figure 5B). However, the probability of fixation of beneficial mutations decreases *linearly* with the frequency of trait expression, ϕ . The distribution of fitness effects of new mutations is highly skewed toward deleterious mutations (EYRE-WALKER and KEIGHTLEY 2007), so that the relaxation of selective constraint vis à vis purifying selection outweighs its effect on adaptive substitution, increasing the rate of sequence divergence among species.

The plasticity load and the mutational cost of complexity

The mutational load is the decrease in mean population fitness due to recurrent deleterious mutation (MULLER 1950). In a haploid population, with a deleterious allele, in frequency q , the mean fitness of a population is $\bar{w} = 1 - qs$. Substituting the allele frequency at mutation–selection balance, \hat{q} (Equation 6), for q gives, $\bar{w} = 1 - u$. Thus, the mutational load is equal to the deleterious mutation rate (MULLER 1950).

For a conditionally expressed locus, the mean population fitness is $\bar{w}_c = 1 - qs\phi = 1 - us\phi/s\phi = 1 - u$. Thus, conditional expression does not affect the mutational load over the *total* population. However, if we consider the mutational load only in generations or habitats where the gene is expressed, then s is not diminished by averaging over nonexpressive conditions and we have $\bar{w}_c^* = 1 - qs = 1 - u/\phi$, where the asterisk

denotes marginal fitness (*e.g.*, the fitness only in the selective environment). Defining the load, L_c , as $1 - \bar{w}_c^*$, we have

$$L_c = u/\phi. \quad (20)$$

We refer to this as the *plasticity load*. For diploids, the same methods lead to $L_c = 2u/\phi$. For haploids, the plasticity load is maximal (*i.e.*, $L_c = 1$) when $\phi = u$, which corresponds to complete mutational decay. Equation 20 demonstrates the connection between the accumulation of deleterious mutations under conditional expression (*e.g.*, Equation 6) and the resulting decline in adaptedness of the population to the conditional environment. KAWECKI (1994) and WHITLOCK (1996) derived the equivalent of Equation 20 for a spatially subdivided population with random migration between habitat patches. Our methods extend these results to any conditionally expressed locus (conditional in time and space). Furthermore, because the ϕ approximation is accurate in the selective environment over a much wider range of parameter values than it is over the total population (see Figures 2, A and B), our results for the plasticity load hold much more generally than suggested by KAWECKI (1994) and WHITLOCK (1996).

We can extend these results to account for the load experienced when more than a single locus contributes to fitness. Consider multiple nucleotide sites contributing multiplicatively to total fitness such that $W = \prod(1 - qs)$. Taking the logarithm of both sides, using the approximation $\log(1 + x) \sim x$ for small x , and taking the exponent gives $W = e^{-nu}$ and $L = 1 - e^{-nu}$, where n is the total haploid number of sites at the locus under consideration and u is the average mutation rate over these sites (for diploids, $W = e^{-2nu}$; *e.g.*, GILLESPIE 2004). The mutational load for a conditionally expressed trait when expressed is

$$L_c = 1 - e^{-nu/\phi}. \quad (21)$$

A population expressing a trait conditionally will have a lower mean fitness than a population of specialists that express a trait unconditionally by a factor of

$$R_W = e^{-nu(1/\phi-1)} \quad (22)$$

Thus, specialists will always be better adapted to the conditional environment than nonspecialists, even when antagonistic pleiotropy is not a factor, as long as specialism involves suites of genes that are conditionally expressed (KAWECKI 1994; WHITLOCK 1996). Since Equations 21 and 22 depend on the product nu , the cost to plasticity relies heavily on the number of loci coding for the plastic phenotype. More elaborate plastic phenotypes, then, are more evolutionarily constrained by mutation than simple plastic phenotypes as demonstrated in Figure 6 by the different curves corresponding to different values of nu .

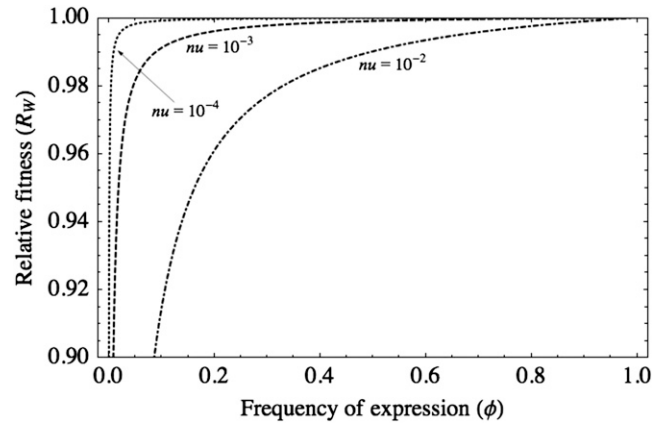


FIGURE 6.—The relative fitness of a conditionally expressed trait under selective conditions, compared to the same trait in a population that always experiences selection on the trait. Fitness is reduced by conditional expression and is more severe for larger values of nu , the number of mutations per generation at loci underlying the conditional trait. The plasticity load is greater for more elaborate traits (*e.g.*, traits with a larger number, n , of underlying genes), indicating a mutational cost to complexity of conditional traits.

METHODS

Temporal variation simulation: To simulate temporal fluctuations in gene expression, we carried out a deterministic simulation in Mathematica (Wolfram), using the standard haploid allele frequency recursion, $q = q(1 - s)/\bar{W} + u(1 - q)$ (CROW and KIMURA 1970), where the mutation rate, u , was set at 0.001. Periodicity of selection was simulated by switching the value of the selection coefficient, s , between 0 and 0.01 at a frequency determined by ϕ . For instance, for $\phi = 1/2$, s alternated between 0 and 0.01 every other generation. For $\phi = 1/10$, $s = 0$ for every generation except generations that were multiples of 10, when s would become equal to 0.01. The results of the simulation are plotted in Figure 1B for multiple values of ϕ . Simulations were run for 10,000 generations, with data points sampled and plotted every 100 generations.

Spatial variation simulation: To simulate spatial variation in gene expression, we carried out a stochastic simulation in Mathematica. A population of N_T individuals was divided into a selective and a neutral subpopulation, each population composing an array of individuals denoted by 0 for wild type and 1 for mutant, where $N_S = \phi N_T$ individuals composed the selective subpopulation and $N_N = (1 - \phi)N_T$ the neutral subpopulation. Individuals were sampled at random with replacement within each subpopulation in accordance with the Wright–Fisher model, and wild-type individuals were mutated with probability u before being placed into the next generation. All randomly chosen individuals in the neutral population were placed into the next generation, whereas mutant individuals in the selective habitat were retained with probability $(1 - s)$. Generations were discrete such that a “generation” consisted

of sampling, selection (in the selective habitat) and mutation until N_S and N_N individuals occupied the next generation in the selective and neutral habitats, respectively. To simulate migration, at the end of each generation, mN_S individuals from each subpopulation were chosen at random and exchanged between subpopulations, where m is the migration rate. This constitutes “conservative migration,” as was assumed in our analytical results. All simulations began with zero mutants, except for the simulation for $\phi = \frac{1}{8}$ plotted in Figure 2C, which was initiated with $q = 0.5$, in order that it reached equilibrium within 10,000 generations. Each simulation was run for 10,000 generations. The parameter values for Figure 2C were $u = 0.001$, $s = 0.01$, $m = 0.1$, and $N_T = 50,000$.

For the mean and standard deviation of allele frequency reported in the text, each simulation was run three times, and the allele frequency from each run was computed as the mean allele frequency of the last 5000 generations.

DISCUSSION

Our results show that spatial or temporal fluctuations in the induction of gene expression generate relaxed selective constraint, increasing the level of standing polymorphism within a species and the rate of non-adaptive divergence between species. Specifically, relative to genes expressed in all individuals all the time, conditional expression increases polymorphism by a factor of approximately $(1/\phi)$, where the approximation is closer for populations of large effective size. This factor is the inverse of ϕ , the fraction of gene-expressing individuals. When gene expression is both spatially (ϕ_I) and temporally (ϕ_T) conditional, the equilibrium polymorphism at a locus is proportional to the product of the inverses, $(1/\phi_I\phi_G)$. We show that the assumption of random mixing between selective and neutral habitats every generation is not necessary and that using the mean selection coefficient ($s\phi$) for a conditionally expressed locus in a spatially variable population is valid even for small rates of migration (Figures 2, A–C). In general, our theory connects gene-level relative polymorphism and divergence with the spatial and temporal frequency of environments inducing gene expression.

Conditionally expressed genes suffer increased mutational and drift loads due to the increased maintenance and fixation of slightly deleterious mutations. As suggested previously (KAWECKI 1994, 1997; WHITLOCK 1996; KAWECKI *et al.* 1997), this poses a serious constraint on the elaboration and maintenance of phenotypically plastic traits. In contrast to other hypotheses for constraints on phenotypic plasticity such as antagonistic pleiotropy and energetic costs (reviewed in DEWITT *et al.* 1998), mutation accumulation due to relaxed selection is a *necessary* consequence of conditional

expression experienced by any conditionally expressed gene. Our theory provides quantitative predictions for how such effects can be observed in gene sequence data. Furthermore, Equation 22 demonstrates that the constraint is greater not only for more infrequently expressed genes, but for traits encoded by a larger suite of genes. If complexity is measured by the number and diversity of distinct morphs encoded by a single genotype, then mutation accumulation due to relaxed selection imposes a potentially severe cost to complexity, limiting the number and elaborateness of distinct morphs in polyphenic species.

Relaxed selection has been proposed as a cause of the observed stability of species' niches over long periods of geologic time (“niche conservatism,” HOLT and GAINES 1992; HOLT 1996). Our theory extends these previous results by demonstrating how parameter values affect the predicted plasticity load. According to Figure 2A, the plasticity load will exist even with relatively low rates of migration between selective and nonselective habitats. The level of migration necessary to eliminate the load may be low enough to facilitate speciation as suggested by KAWECKI (1997).

Importantly, the relaxed selection described here is not experienced by genes expressed conditionally within a single individual's lifetime. In many cases, genes are expressed with varying intensity across tissues or cell types or at different life stages in an individual's lifetime. If there is a relationship between the intensity of gene expression and the intensity of selection (*i.e.*, the magnitude of s), then it might well be possible to extend our theory. If the relationship between expression level and selection was given by the linear regression coefficient, β , then two genes with different expression levels could be compared for standing polymorphism and divergence using our theory by substituting β for ϕ . Recent empirical work has shown that levels of gene expression within an individual are correlated with rates of divergence (LEMONS *et al.* 2005) suggesting that such relationships may exist and permit tests of our theory. Similarly, if the tissue specificity of gene expression could be related to strength of selection, predictions regarding the relative rate and diversity of evolution of one gene to another could be made. This is an active area of current research and it is too early to determine whether such relationships will be discovered and whether, if discovered in one species, they hold in another.

Applications of the Theory and Tests of its Predictions: Our theory is supported by previous studies, which serve as a guide for future testing. One way of testing the theory is to apply the methods used in testing the predictions of maternal effects theory (WADE 1998; LINKSVAYER and WADE 2009), which is a special case of the current model with $\phi = \frac{1}{2}$. Theory predicts that, all else being equal, maternal effect genes should experience relaxed selection relative to genes expressed in

both sexes. The specific predictions are: (1) standing variation for maternal effect genes should be twice that of genes expressed in both sexes, and (2) maternal effect genes should be differentiating more rapidly across taxa. BARKER *et al.* (2005) tested these two predictions in two species of fruit flies, using the tandem duplicate gene pair *bicoid* (*bcd*, maternal expression) and *zerknult* (*zen*, zygotic expression in both sexes). For the ratios of (*bcd/zen*) polymorphism, they found it to be 1.08 for silent sites and 2.18 for nonsynonymous sites within *Drosophila melanogaster*, and 0.89 and 2.68 for silent and nonsynonymous sites, respectively, in *Drosophila simulans*. Both ratios provide quantitative support for the theory predictions. In a broader test of the theory, CRUICKSHANK and WADE (2008; Figure 3, A and B) compared sequence diversity within *D. simulans* and sequence divergence among several species for 42 genes, critical to early embryo development (9 maternal and 33 zygotic). For five maternal-zygotic homologs, like *bcd* and *zen*, they observed a nonsynonymous site diversity to average 0.55 ± 0.22 SE for the maternal genes and 0.28 ± 0.15 SE for the zygotic genes for a ratio of 1.96, very nearly identical to the theoretically predicted ratio of 2.0. Across species, CRUICKSHANK and WADE (2008, Figure 2) found that “. . .the average level of between-species sequence divergence ranges from two to four times greater for maternal than for zygotic genes.” These empirical findings also quantitatively support the theoretical predictions.

In some species, sex-limited gene expression may result in values of ϕ that differ from $\phi = \frac{1}{2}$. In parthenogenic species with facultative sexuality, male-specific genes experience a frequency of expression such that $\phi < \frac{1}{2}$. BRISSON and NUZHIDIN (2008) identified male-specific, female-specific, and sex-neutral genes in the facultatively sexual pea aphid, *A. pisum*, and measured levels of sequence divergence between species for these classes of genes. This species reproduces clonally for 10 to 20 generations before producing a sexual generation composed of males and females in the fall (MORAN 1992; BRISSON and NUZHIDIN 2008). Consistent with the predictions of relaxed selective constraint, male-specific genes had more than a twofold increase in divergence compared to female-specific and sex-neutral genes, which had approximately the same levels of divergence, and tests of selection showed that this variation accumulated neutrally (BRISSON and NUZHIDIN 2008). The pattern of expression for this species gives a frequency of expression of between $\phi = 0.05$ and $\phi = 0.1$. While our theory shows that the level of standing polymorphism at equilibrium should be inflated 10- to 20-fold, the consequence for divergence is not as straightforward. From Equation 17 and Figure 5, a ϕ value of 0.1 will produce a 2-fold increase in divergence if the mutations accumulated experience an average selection intensity of $N_e s \sim -1.3$, although this requires the assumption that deleterious mutations are dominant. Further empirical work on the effect sizes of

spontaneous mutations in this species, or estimation of N_e and comparison with mutational effect sizes from other species, will be necessary to determine if these data support our theory. This highlights the fact that less data are required to apply our theory to polymorphism data than to divergence data. While divergence data are sensitive to the selection coefficient and dominance, polymorphism data are not. Unfortunately, at present, polymorphism data are not as easily obtained as divergence data.

With polymorphism data, our theory can be used in multiple ways. Given a known value of ϕ , a null expectation for levels of sequence polymorphism can be obtained. Our theory shows that null expectations should always be adjusted to account for the effects of the frequency of expression. Alternatively, the parameter ϕ can be estimated from the data, or if ϕ is known, differences in selection pressures can be measured between paralogs or classes of genes. The two major barriers to such an analysis are the availability of polymorphism data and of expression data across individuals.

The results presented here for conditionally expressed genes can also be applied to social genes. Recent theory (LINKSVAYER and WADE 2009) hypothesizes that genes with indirect genetic effects or genes whose primary fitness effect is at the level of the social group (called “kin selection” genes) should experience relaxed selective constraint in a manner analogous to conditionally expressed genes. In the Linksvayer–Wade theory, the coefficient of relatedness, r , between the individual expressing the gene and the group experiencing its fitness effects, plays the same role as ϕ in our theory for conditionally expressed genes. The quantitative predictions developed here, then, can be applied with modification to kin selection genes.

While we focused primarily on relaxed selection as an evolutionary constraint on conditionally expressed traits, we cannot discount the possibility that much of the accumulated mutational variation at conditional loci may provide the raw material for adaptive evolutionary change under appropriate conditions. The increased level of standing genetic variation caused by relaxed selection may facilitate the rapid emergence of evolutionary novelties when formerly rare, low ϕ environments become more common. Further, the larger neutral range experienced by conditionally expressed genes may allow populations to explore previously inaccessible regions of sequence space, allowing populations to escape local optima in favor of globally higher fitness peaks. The possibility of such adaptive peak shifts (WRIGHT 1931) is of considerable medical relevance since virulence and antibiotic resistance are both conditionally expressed traits, and since all new host species and drug therapies begin as rare, low ϕ environments. Similarly, given the broad but predictable patterns of climate change associated with global warming, it is likely that, at many latitudes, formerly rare

environments will become more common. Conditionally expressed genes, then, may contribute disproportionately to the evolution of novel form and function under appropriate ecological and population genetic conditions.

We thank M. Feldman, M. Hahn, and M. Lynch for helpful discussions and Y. Brandvain, T. Cruickshank, T. Platt, and P. Zee for helpful comments on the manuscript. This work was supported by National Institutes of Health grant R01 GM65414 to M.J.W. and a National Science Foundation Integrative Graduate Education and Research Traineeship (NSF IGERT) Fellowship to J.D.V.D.

LITERATURE CITED

- ABOUHEIF, E., and G. A. WRAY, 2002 Evolution of the gene network underlying wing polyphenism in ants. *Science* **297**: 249–252.
- BARKER, M. S., J. P. DEMUTH and M. J. WADE, 2005 Maternal expression relaxes constraint on innovation of the anterior determinant, *bicoid*. *PLoS Genet* **1**: e57.
- BRISSON, J. A., and S. V. NUZHIDIN, 2008 Rarity of males in pea aphids results in mutational decay. *Science* **319**: 58.
- CHASNOV, J. R., and K. L. CHOW, 2002 Why are there males in the hermaphroditic species *Caenorhabditis elegans*? *Genetics* **160**: 983–994.
- CROW, J. F., and M. KIMURA, 1970 *An Introduction to Population Genetics Theory*. Harper & Row, New York.
- CRUICKSHANK, T., and M. J. WADE, 2008 Microevolutionary support for a developmental hourglass: gene expression patterns shape sequence variation and divergence in *Drosophila*. *Evol. Dev.* **10**: 583–590.
- DEMUTH, J., and M. WADE, 2007 Maternal expression increases the rate of bicoid evolution by relaxing selective constraint. *Genetica* **129**: 37–43.
- DEWITT, T. J., 1998 Costs and limits of phenotypic plasticity: tests with predator-induced morphology and life history in a freshwater snail. *J. Evol. Biol.* **11**: 465–480.
- DEWITT, T. J., A. SIH and D. S. WILSON, 1998 Costs and limits of phenotypic plasticity. *Trends Ecol. Evol.* **13**: 77–81.
- EYRE-WALKER, A., and P. D. KEIGHTLEY, 2007 The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* **8**: 610–618.
- EYRE-WALKER, A., M. WOOLFIT and T. PHELPS, 2006 The distribution of fitness effects of new deleterious amino acid mutations in humans. *Genetics* **173**: 891–900.
- FRANK, S. A., and M. SLATKIN, 1990 Evolution in a Variable Environment. *Am. Nat.* **136**: 244–260.
- FRETWELL, S. D., 1972 *Populations in a Seasonal Environment*. Princeton University Press, Princeton, NJ.
- FUQUA, C., S. C. WINANS and E. P. GREENBERG, 1996 Census and consensus in bacterial ecosystems: the LuxR-LuxI family of quorum-sensing transcriptional regulators. *Annu. Rev. Microbiol.* **50**: 727–751.
- GILLESPIE, J. H., 2004 *Population Genetics: A Concise Guide*. JHU Press, Baltimore, Maryland.
- HALDANE, J. B. S., 1927 A mathematical theory of natural and artificial selection. Part V. Selection and mutation. *Proc. Camb. Philol. Soc.* **23**: 838–844.
- HARVELL, C. D., 1998 Genetic variation and polymorphism in the inducible spines of a marine Bryozoa. *Evolution* **52**: 80–86.
- HAZEL, W., R. SMOCK and C. M. LIVELY, 2004 The ecological genetics of conditional strategies. *Am. Nat.* **163**: 888–900.
- HOLT, R., and M. GAINES, 1992 Analysis of adaptation in heterogeneous landscapes: implications for the evolution of fundamental niches. *Evol. Ecol.* **6**: 433–447.
- HOLT, R. D., 1996 Demographic constraints in evolution: towards unifying the evolutionary theories of senescence and niche conservatism. *Evol. Ecol.* **10**: 1–11.
- KAWECKI, T., N. H. BARTON and J. D. FRY, 1997 Mutational collapse of fitness in marginal habitats and the evolution of ecological specialisation. *J. Evol. Biol.* **10**: 407–429.
- KAWECKI, T. J., 1994 Accumulation of deleterious mutations and the evolutionary cost of being a generalist. *Am. Nat.* **144**: 833–838.
- KAWECKI, T. J., 1997 Sympatric speciation via habitat specialization driven by deleterious mutations. *Evolution* **51**: 1751–1763.
- KIMURA, M., 1962 On the probability of fixation of mutant genes in a population. *Genetics* **47**: 713–719.
- KIMURA, M., and T. OHTA, 1971 Protein polymorphism as a phase of molecular evolution. *Nature* **229**: 467–469.
- LEMONS, B., B. R. BETTENCOURT, C. D. MEIKLEJOHN and D. L. HARTL, 2005 Evolution of proteins and gene expression levels are coupled in *Drosophila* and are independently associated with mRNA abundance, protein length, and number of protein-protein interactions. *Mol. Biol. Evol.* **22**: 1345–1354.
- LEVINS, R., 1968 *Evolution in Changing Environments: Some Theoretical Explorations*. Princeton University Press, Princeton, NJ.
- LINKSVAYER, T. A., and M. J. WADE, 2008 Genes with social effects are expected to harbor more sequence variation within and between species. *Evolution* **63**: 1685–1696.
- LIVELY, C. M., 1986 Predator-induced shell dimorphism in the acorn barnacle *Chthamalus anisopoma*. *Evolution* **40**: 232–242.
- LOEWE, L., B. CHARLESWORTH, C. BARTOLOME and V. NOEL, 2006 Estimating selection on nonsynonymous mutations. *Genetics* **172**: 1079–1092.
- LOMMEN, S. T. E., P. W. DE JONG and P. M. BRAKEFIELD, 2005 Phenotypic plasticity of elytron length in wingless two-spot ladybird beetles, *Adalia bipunctata* (Coleoptera: Coccinellidae). *Eur. J. Entomol.* **102**: 553–556.
- MORAN, N. A., 1992 The evolution of aphid life cycles. *Annu. Rev. Entomol.* **37**: 321–48.
- MULLER, H. J., 1950 Our load of mutations. *Am. J. Hum. Genet.* **2**: 111–176.
- NAGYLAKI, T., 1980 The strong-migration limit in geographically structured populations. *J. Math. Biol.* **9**: 101–114.
- NEI, M., and W. H. LI, 1979 Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* **76**: 5269–5273.
- OHTA, T., and M. KIMURA, 1971 On the constancy of the evolutionary rate of cistrons. *J. Mol. Evol.* **1**: 18–25.
- RICE, S. H. 2004 *Evolutionary Theory: Mathematical and Conceptual Foundations*. Sinauer Associates, Inc., Sunderland, MA.
- ROFF, D. A., and M. B. GÉLINAS, 2003 Phenotypic plasticity and the evolution of trade-offs: the quantitative genetics of resource allocation in the wing dimorphic cricket, *Gryllus firmus*. *J. Evol. Biol.* **16**: 55–63.
- ROSENZWEIG, M. L., 1987 Habitat selection as a source of biological diversity. *Evol. Ecol.* **1**: 315–330.
- SCHEINER, S. M., 1993 Genetics and evolution of phenotypic plasticity. *Annu. Rev. Ecol. Systematics* **24**: 35–68.
- SEIDEL, H. S., M. V. ROCKMAN and L. KRUGLYAK, 2008 Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *Science* **319**: 589–594.
- TAJIMA, F., 1983 Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**: 437–460.
- VIA, S., R. GOMULKIEWICZ, G. DE JONG, S. M. SCHEINER, C. D. SCHLICHTING *et al.*, 1995 Adaptive phenotypic plasticity: consensus and controversy. *Trends Ecol. Evol.* **10**: 212–217.
- WADE, M. J., 1998 The evolutionary genetics of maternal effects, pp. 5–21 in *Maternal Effects as Adaptations*, edited by T. A. MOUSSEAU and C. W. FOX. Oxford University Press, New York.
- WATTERSON, G. A., 1975 On the number of segregating sites in genetic models without recombination. *Theor. Popul. Biol.* **7**: 256–276.
- WHITLOCK, M. C., and M. J. WADE, 1995 Speciation: Founder Events and Their Effects on X-Linked and Autosomal Genes. *Am. Nat.* **145**: 676–685.
- WHITLOCK, M. C., 1996 The Red Queen beats the Jack-of-All-Trades: the limitations on the evolution of phenotypic plasticity. *Am. Nat.* **148** (Suppl): S65–S77.
- WRIGHT, S., 1931 Evolution in Mendelian populations. *Genetics* **16**: 97–159.