Genetics: Published Articles Ahead of Print, published on March 4, 2007 as 10.1534/genetics.106.069658

NEUTRAL EVOLUTION OF MULTIPLE QUANTITATIVE CHARACTERS: A GENEALOGICAL APPROACH

Cortland K. Griswold[*], Benjamin Logsdon[†,1] and Richard Gomulkiewicz[*,††]

[*]School of Biological Sciences, Washington State University Pullman, WA 99164

[†]School of Molecular Biosciences, Washington State University Pullman, WA 99164

[††]Department of Mathematics, Washington State University Pullman, WA 99164

1 – Current address: Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14583; (607) 351-6874; bal47@cornell.edu

Running Head: Neutral theory of **G** evolution

Key words: neutral, coalescent, ancestral recombination graph, eigenvalues, G matrix

Corresponding author: Richard Gomulkiewicz, School of Biological Sciences, Washington State University, Pullman, WA 99164; gomulki@wsu.edu

ABSTRACT

The **G** matrix measures the components of phenotypic variation that are genetically heritable. The structure of **G**, that is its principal components and their associated variances, determines, in part, the direction and speed of multivariate trait evolution. In this paper we present a framework and results that give the structure of **G** under the assumption of neutrality. We suggest that a neutral expectation of the structure of **G** is important because it gives a null expectation for the structure of **G** from which the unique consequences of selection can be determined. We demonstrate how the processes of mutation, recombination and drift shape the structure of **G**. Furthermore, we demonstrate how shared common ancestry between segregating alleles shapes the structure of **G**. Our results show that shared common ancestry, which manifests itself in the form of a gene genealogy, causes the structure of **G** to be non-uniform in that the variances associated with the principal components of **G** decline at an approximately exponential rate. Furthermore we show that the extent of the non-uniformity in the structure of **G** is enhanced with declines in mutation rates, recombination rates, numbers of loci and is dependent on the pattern and modality of mutation.

Phenotypic evolution is channeled through patterns of heritable variation. This is exemplified by the breeder's equation, $R = h^2 S$, which for a single quantitative trait, states that the evolutionary response (R) to selection (S) of the trait is proportional to its heritability, h^2, (see, e.g., FALCONER and MACKAY 1996; LYNCH and WALSH 1998). Likewise, the simultaneous evolutionary responses of several potentially correlated quantitative characters are predicted by the multivariate extension of the breeder's equation, $\Delta\bar{\mathbf{z}} = \mathbf{G}\boldsymbol{\beta}$, where $\Delta\bar{\mathbf{z}}$ is a vector describing the responses of each character to selection, $\boldsymbol{\beta}$—the selection gradient vector—describes the pattern of directional selection applied to the traits, and **G** is the additive-genetic covariance matrix describing the heritable patterns of variation and covariation of the traits (LANDE 1979).

Quantitative geneticists and evolutionary biologists alike are particularly interested in the principal components of the **G** matrix, since these indicate independent paths along which evolution can occur, and the variance associated with each principal component, since these modulate the speed of evolution along the direction of each principal component. For instance, the "evolutionary line of least resistance" for adaptive radiation (SCHLUTER 1996) is defined by the principal component of **G** with the largest variance, i.e., with the greatest amount of heritable variation. At the other extreme, principal components of **G** with the smallest variances represent evolutionary avenues with the greatest degrees of genetic resistance (HINE and BLOWS 2006; KIRKPATRICK and LOFSVOLD 1992).

Since the **G** matrix itself can evolve, it is natural to consider the influences of evolutionary forces (such as selection, mutation, random genetic drift, and recombination) on the patterns of variation associated with the principal components of **G**. For example, principal components with the smallest eigenvalues might reflect the strongest patterns of multivariate selection (e.g., BLOWS *et al.* 2004; HINE *et al.* 2004). Indeed, this is expected on general theoretical grounds

(LANDE 1980; LANDE and ARNOLD 1983; PHILLIPS and ARNOLD 1989) and has been verified by computer simulations (JONES *et al.* 2003).

A second force that causes the **G** matrix to evolve is random genetic drift. PHILLIPS *et al*. (2001) showed that severe inbreeding led to an average reduction in total variation of the **G** matrix that was proportional to the theoretical expectation $1 - F$, where F is the inbreeding coefficient. Furthermore, the principal components of the **G** matrices differed (in an apparently random way) among replicated inbred lines, although, on average, the **G** matrices across replicated inbred lines did not differ from the outbred control population. Subsequent work by WHITLOCK *et al*. (2002) on the same set of inbred lines showed that the **G** matrix in some lines continued to change over 20 generations of random mating, but there was no evidence that the **G** matrices were returning to the original state in the outbred control line. Based on work by AVERY and HILL (1977) and ZENG and COCKERHAM (1991), WHITLOCK *et al*. (2002) hypothesized that the subsequent changes in **G** matrices in the generations following inbreeding were caused, in part, by the decay of linkage disequilibria that were generated during the inbreeding episode. Besides the force of recombination, theoretical work by LANDE (1980) shows that mutational pressure can also cause the **G** matrix to evolve.

At equilibrium the proximate causes of the structure of a **G** matrix are linkage disequilibria and the pleiotropic and non-pleiotropic effects of segregating mutations. The ultimate forces that contribute to the structure of the **G** matrix are recombination, mutation, genetic drift and selection (e.g., PHILLIPS and MCGUIGAN 2006). To date, there is no theoretical work that presents a neutral expectation for the structure of the **G** matrix when the only forces shaping its evolution are mutation, recombination and drift, although LYNCH and HILL (1986) presented a neutral model of the evolution of genetic variance for a single character. In the case of multiple

characters LYNCH (1994) and PHILLIPS and MCGUIGAN (2006) postulate that **G** will evolve to be proportional to **M**, the mutational variance-covariance matrix.

The focus of this paper is the question of how mutation, recombination, and random genetic drift determine the eigenstructure of the **G** matrix at equilibrium, where the eigenstructure of the **G** matrix is measured according to its eigenvectors and eigenvalues. The eigenvectors of **G** are its principal components and the eigenvalues of **G** are the variances associated with each eigenvector. Naïvely, for the case of free recombination among loci, one might expect that the eigenvalues of **G** would be proportional to those of **M,** with drift induced linkage disequilibrium and stochastic sampling from **M** causing non-directional variation around this expectation. Accordingly, if the eigenvalues of **M** were all similar in size, then on average so would the eigenvalues of **G**. We explore the validity of this intuitively reasonable expectation.

One of the main insights of this paper is that this intuitively reasonable expectation is not valid. A primary reason is that segregating alleles in a population are related to one-another due to shared common ancestry or, to put it another way, alleles are related genealogically through shared common descent via reproduction. Patterns of shared common ancestry can be described for a sample from a population using Kingman's (1982) coalescent and multilocus extensions such as Griffiths and Marjoram's (1997) ancestral-recombination graph. In this paper we use the coalescent and the ancestral-recombination graph as a framework to investigate neutral evolution of **G**.

The neutral expectations developed here aid the interpretation and analysis of recent empirical work that showed significant additive genetic variance for 20 wing-shape characters in *D. melanogaster* and an approximately exponential decline in the ordered eigenvalues of the **G** matrix for these characters (MEZEY and HOULE 2005). The exponential decline in ordered

eigenvalues may be interpreted as signatures of either variable selection across characters and/or variable mutational heritabilities across characters. Our analyses show that, under the neutral model, the ordered eigenvalues of **G** will evolve toward a regular pattern. Surprisingly, even if the ordered eigenvalues of **M** are perfectly uniform and selection is absent, the expected eigenvalues of **G** in a sample from a population are predicted to decline at an approximately exponential rate.

METHODS

The Model: We consider the evolution of a multivariate quantitative trait affected by multiple loci in a diploid population of size N with Wright-Fisher random mating. (For non-ideal populations, $N = N_e$, the variance-effective population size.) Evolution is neutral as we assume the trait experiences no selection. Alleles affecting the trait act additively within and across loci (i.e., no dominance, epistasis, or genotype-environment interactions). For example, if $\mathbf{x}_i$ and $\mathbf{x}_i'$ are multivariate effects of maternally and paternally inherited alleles at locus i (e.g. the effects of each allele on m traits), then the genotypic value $\mathbf{g}$ of a typical individual is $\mathbf{g} = \sum_{i=1}^{\ell} (\mathbf{x}_i + \mathbf{x}_i')$, where ℓ is the number of loci. The phenotype is $\mathbf{z} = \mathbf{g} + \mathbf{e}$ where $\mathbf{e}$ is an independent, mean-zero random variable. The **G**-matrix is the $m \times m$ covariance matrix of additive values, $\mathbf{g}$. We hereafter ignore $\mathbf{e}$ because, in the absence of selection, it plays no role in the evolution of **G**. It is important to note, however, that $\mathbf{e}$ can affect estimates of **G**, which may affect predictions of the eigenstructure of any particular **G**, especially for small samples from a population. We do not address this issue in this paper and leave it for further study.

Under neutrality and additivity, three main processes influence **G** matrix evolution: spontaneous mutation, recombination, and random genetic drift via stochastic reproduction. In

this paper we focus on the recombination-mutation-drift equilibrium characteristics of **G**. For spontaneous mutation we examined two infinite-allele mutation models: the continuum-of-alleles model (CROW and KIMURA 1964; KIMURA 1965) and the house-of-cards model (KINGMAN 1977; KINGMAN 1978). The continuum-of-alleles model assumes that each mutation generates a new multivariate effect that is added to the current effect of an allele at a locus, where a new effect is sampled from a mutational distribution. In our simulations we used a zero-mean multivariate normal distribution with covariance matrix **A** for the distribution of mutational effects. The diagonal elements of **A** are the variances in mutational effects and the off-diagonals are the covariances of mutational effects on different traits. The house-of-cards model (KINGMAN 1978) assumes that each new mutation completely replaces the current allelic effect at a mutated locus.

Coalescent approaches: Because the alleles contributing to **g** are neutral and every mutation is novel, the alleles at individual loci should follow the distribution expected at equilibrium for the infinite alleles neutral model (see, e.g., EWENS 2004)**.** Rather than sample from this distribution directly, we used a different but equivalent approach for $n<N$, where n is the sample size — the coalescent. The coalescent is useful in that it retains information about the underlying genealogical history of alleles. Thus it can be used to model how variation in allelic values is determined by patterns of shared common ancestry.

Coalescent with independent loci: Figure 1 illustrates the approach that we used to generate the genotypic values **g** of individuals in a sample of size n based on the neutral coalescent (KINGMAN 1982). Loci were treated independently. For each locus we generated a tree with $2n$ tips and placed mutations on the tree using the standard neutral coalescent approach. (We employed the algorithm in the Appendix of HUDSON 1991.) When one unit of coalescent time is scaled to be $2N$ generations, coalescent events occur at a rate $k(k-1)/2$ per coalescent time unit

using the exponential approximation of a geometric process, where *k* is the number of lineages present at a particular time back in history. Once a gene tree is drawn, branch segments in the tree have times associated with them as a result of the coalescent process. Mutations were randomly placed along branch segments such that the number of mutations along a branch segment of the coalescent tree was Poisson distributed with mean μt_i, where μ is the haploid per locus mutation rate and t_i is the length (in coalescent time units) of branch segment *i*. With each mutation, we then associated a vector of effects drawn from a multivariate normal distribution with mean **0** and covariance **A**. A population was generated by randomly assigning two alleles from each locus (without replacement) to each of the *n* individuals. Note that this coalescent approach to generating individual genotypic values imposes the strict rule that loci are completely independent, i.e., there are no transient linkage disequilibria.

The overall process is illustrated in figure 1 for a sample of $n = 2$ individuals, each with two loci. Mutations in the coalescent tree are marked with hashes and the associated vector of effects is represented by $\vec{\delta}$. For locus A, alleles *a1* and *a2* are unmutated relative to the ancestral allele and, accordingly, have the ancestral value $\vec{0}$. Allele *a3* is differentiated from the ancestral allele by one mutation and has the value $\vec{0} + \vec{\delta}_3 = \vec{\delta}_3$. The value of allele *a4* depends on the mutation model. For the continuum-of-alleles model, the value of *a4* is $\vec{0} + \vec{\delta}_1 + \vec{\delta}_2 = \vec{\delta}_1 + \vec{\delta}_2$, but for the house-of-cards model, the value of *a4* is $\vec{0} + \vec{\delta}_2 = \vec{\delta}_2$, since the later mutational effect $\vec{\delta}_2$ completely replaces the earlier effect $\vec{\delta}_1$. The allelic effects at locus B are determined similarly.

Next, the alleles are randomly assigned two at a time, without replacement, to an individual. In figure 1, alleles *a1* and *a4* at locus A and alleles *b3* and *b4* at locus B were assigned to the first individual. Based on the alleles sampled and the mutations in the gene tree, the genotypic value

of the first individual is $\mathbf{g} = \vec{\delta}_1 + \vec{\delta}_2 + 2\vec{\delta}_4$ for the continuum-of-alleles model and $\mathbf{g} = \vec{\delta}_2 + 2\vec{\delta}_4$ assuming the house-of-cards.

Ancestral recombination graph: As stated at the outset, linkage disequilibria are also a proximate cause of additive-genetic covariances. Strictly applying independence among loci may therefore eliminate a factor promoting genetic covariances even when there is free recombination among loci. To allow for physical linkage among loci and for the possibility of linkage disequilibria we used the ancestral recombination graph coalescent approach (GRIFFITHS 1981; GRIFFITHS 1991; GRIFFITHS and MARJORAM 1996; GRIFFITHS and MARJORAM 1997; see NORDBORG 2001, pp. 195-200 for a useful introduction).

Figure 2 illustrates how a sample of two individuals each consisting of two loci is generated using the ancestral recombination graph. Instead of having two independent loci, the loci are linked. In figure 2, the uppermost rectangles each represent a haplotype in the current generation with the unshaded section corresponding to locus A and the black shaded area corresponding to locus B. In an ancestral recombination graph, two processes occur when forming a gene tree, namely recombination and coalescent events. Scaling time to be $2N$ generations, coalescent events occur as before at rate $k(k-1)/2$, where k is the number of lineages at a given time. Recombination events occur at the rate $4Nrk/2$, where r is the per haplotype recombination rate per generation. The total rate of events is $k(k-1)/2 + 4Nrk/2$, and given an event, the probability it is a coalescent is $(k-1)/(k-1+4Nr)$. Although the rate of recombination is typically much faster than the rate of coalescence, a common ancestor is guaranteed (GRIFFITHS and MARJORAM 1997).

Proceeding backward in time from the present, a recombination event is represented by the splitting of a single haplotype lineage into two lineages. For a two-locus model, one locus is

retained in one ancestral lineage, whereas the other locus is retained in the second ancestral lineage. In figure 2, gray shading corresponds to a locus that is recombined out in the descendent lineage following a recombination event. Often recombination events occur in an ancestral lineage that consists of one locus that is present in the current generation and one that is subsequently recombined out and not present in the current generation. With respect to the coalescence of alleles in the sample, these recombination events are relevant only with respect to incrementing time backward in the gene tree, but do not add additional information. Consequently, the splitting of these types of lineages is not recorded in the gene tree, although the time of the event is.

Once an ancestral recombination graph is drawn, there are times associated with each branch segment (in units of $2N$ generations) and numbers of mutations are modeled to occur along each branch segment according to a Poisson distribution with mean $\ell\mu t_i$, where t_i is the time associated with branch segment i. The per locus mutation rate is assumed to be equal across loci, such that for ℓ loci and given a mutational event, the probability the mutation occurs at a particular locus is $1/\ell$. In figure 2, hash marks represent mutations, the box next to it represents the locus at which the mutation occurred and $\vec{\delta}_i$ is the vector of effects of mutation i. To get an allele's value, simply trace its history back in the gene tree being careful to follow the correct ancestral lineage at a recombination event.

Because loci can be physically linked, individual genotypes are formed by randomly sampling haplotypes as opposed to alleles at independent loci. With a sample consisting of two individuals there are four haplotypes, each individual consisting of two haplotypes. In the example (figure 2), individual one consists of haplotypes one and four. Following the history of the allele *a1* back in time, only the mutation with effect $\vec{\delta}_7$ occurs before reaching the common

ancestor of the population so the additive effect of *a1* is $\vec{0} + \vec{\delta}_7 = \vec{\delta}_7$. Note that the mutation with effect $\vec{\delta}_3$ occurred in an ancestral allele not present in the current generation. As with the independent-loci coalescent approach described above, an allele's additive value is the sum of the effects of mutations in a lineage for the continuum-of-alleles mutation model whereas for the house-of-cards model, the additive effect is equal to the most recent mutation for the allele.

Given the relatively large sample sizes examined in this study combined with a large number of loci and moderate to fast recombination rate, there were too many possible ancestral haplotypes for computers with 1 GB of memory to simulate the ancestral recombination graph algorithm, hence there are significant computational limitations when considering recombination while utilizing the coalescent method.

Parameters: Parameters underlying the results presented in this paper are cast in terms of $4N\mu$, $2Nr$, the number of loci and **M** (see below for a definition of **M**). The expression $4N\mu/2$ gives the expected number of new mutations that enter a population each generation at each locus. In this paper, $4N\mu$ ranges from 0.1 to 10.0. The expression $2Nr/2$ gives the expected number of recombination events per generation during the meiotic phase of the life cycle for an entire population per adjacent pair of loci and it ranges from 0.0 to 1000.0. Numbers of loci vary from one to ten in the results, and it is assumed that each locus has the same value of $4N\mu$.

In this paper, mutations at a locus act either pleiotropically or non-pleiotropically. When mutations act pleiotropically, a single mutation potentially affects all characters and the covariance matrix of their effects is given by **A**. For simplicity, we concentrate on diagonal covariance matrices such that $\mathbf{A} = 0.05\mathbf{I}$, where **I** is the identity matrix. When mutations at a locus act non-pleiotropically, the mutations only affect a single character and the variance of their effects is 0.05.

Throughout the paper it is assumed that mutation rates and mutational effects are uniform across characters. Thus, the expected eigenvalues of the mutational covariance matrix, **M**, are uniform, where $\mathbf{M} = U\mathbf{A}$ and U is the zygotic mutation rate. We make these simplifying assumptions so that any non-uniformity in the eigenvalues of **G** arises via factors other than non-uniformity in the eigenvalues of the mutational covariance matrix. In reality, characters of organisms will likely have non-uniform mutational variance and this may be an important factor in determining the ordered eigenvalues of **G**. But, here we are particularly interested in demonstrating that genealogical history per se is important in shaping the ordered eigenvalues of **G**. Furthermore, we do not explore cases in which there is a non-orthogonal transformation between the underlying mutational process and the measured characters of an organism. If the transformation is not orthogonal then this will generally affect the eigenvalues of **G.**

RESULTS

All cases we examined were consistent with the prediction that, at equilibrium, the average **G** matrix is equal to $2N\mathbf{M}$ under neutrality (PHILLIPS and MCGUIGAN 2006). However, our results revealed two new, consistent features of neutral evolution of **G** matrices at recombination-mutation-drift balance. First, the eigenvalues of equilibrium **G** matrices are expected to be highly non-uniform — declining in an approximately exponential manner — even when the mutation covariance matrix **M** has perfectly uniform eigenvalues (i.e., when **M** has a single repeated eigenvalue). Second, we found that genealogical ancestry of alleles in a finite population is particularly important in determining the intensity of decay of the ranked eigenvalues of **G** matrices under neutrality. In the absence of selection, two forces, namely recombination and random genetic drift, generate genealogies of alleles, and the genotypic

values of alleles in a particular genealogy are determined by the pattern of mutation. In what follows we show in turn how mutation, drift, and recombination shape the ordered eigenvalues of **G** matrices.

Genealogical effect on G: Figure 3 demonstrates how genealogical ancestry can generate non-uniformity in the eigenvalues of **G** assuming allelic variation at a single locus. In figure 3a, the upper left corner shows a genealogy of a sample of six alleles. On the genealogy are ten mutations that give rise to five allelic types out of the six sampled alleles. The bivariate effects of these ten mutations are plotted. The principal components of mutational variation are of similar magnitude suggesting that variation is evenly distributed across the two dimensions. Figure 3b shows the values of the alleles at the tips of the tree, plotted as black dots, using the continuum-of-alleles model. What is immediately apparent in this example is that allelic variation, instead of being evenly spread across two dimensions (as in fig. 3a), is primarily concentrated in one dimension. The reason for this is that allelic values are strongly dependent due to their shared common ancestry whereas the mutational effects are independent.

Figure 4 extends the principle illustrated in figure 3 by replicating the genealogical history of a single locus many times to determine if, on average, there is a greater rate of decline in the eigenvalues associated with breeding values versus mutational variation. Furthermore, a larger sample is taken from each population and there are more trait dimensions. The box-shaped points in figure 4 give the average ordered eigenvalues of mutational variation for a single locus (as was done in figure 3a, except replicated 500 times). Note that variation is non-uniform across the principal components even though mutation is sampled from a distribution with uniform variance across its principal components. The non-uniformity is the result of stochasticity, in that there are a finite number of mutations segregating at a locus and by chance

the effects of these mutations are non-uniformly distributed across characters. The triangular points in figure 4 suggest that the eigenvalues associated with the breeding values of individuals decline at a greater rate than the eigenvalues associated with mutational variation. Two processes within the genealogy appear to bring about the enhanced decline in eigenvalues. One is that genealogy builds up statistical dependence among allelic values due to the pattern of shared common ancestry (branching). The second is that along a single branch multiple mutations may occur which reduces the number of independently acting mutations. Figure 4 suggests that the point illustrated in figure 3, namely that genealogy potentially enhances the non-uniformity in the eigenvalues associated with variation in breeding values, is general because, on average, there is a faster decline in eigenvalues associated with variation in breeding values than in mutational effects. On occasion, genealogical ancestry may have no effect on eigenvalues, but on average, figure 4 suggests that it enhances the non-uniformity in their values.

Linkage: Linkage causes the genealogies of alleles at different loci to be correlated. Our results show that tighter linkage between two loci tends to enhance the non-uniformity in the expected eigenvalues of **G** (figure 5). When the loci are completely linked ($2Nr = 0$) there is greater non-uniformity in eigenvalues than with a higher rate of recombination ($2Nr = 1000$).

Other factors influencing eigenstructure: Besides finite mutation, genealogy, and linkage, our results show that several other factors contribute to the pattern of eigenvalues of **G**.

Number of loci: Alleles at each locus vary in how they contribute to patterns of additive genetic variance. As more loci are added, the relative contribution of each to the variability in **G** is reduced. Accordingly, there is more non-uniformity in the eigenvalues of **G** when there are few versus many loci (figure 6).

Mutation rate and effective population size: The rate of decline in eigenvalues is greater with smaller values of $4N\mu$ (figure 7). In fact, there is a limit to the expected number of principal components that will have nonzero variance associated with it (Table 1). In table 1, the **M** matrix had 25 dimensions with equal variances across dimensions and no covariance, yet beyond twelve principal components for a sample size of 75 and 15 principal components for a sample size of 300, principal components are expected to have zero variance associated with them. As the mutation rate or effective population size increases, the number of principal components with nonzero variance will increase (data not shown). Figure 7 illustrates another important point, namely that the decline in eigenvalues is not log-linear for high dimensional cases.

Mutational matrices with variable eigenvalues or non-zero correlations: Until now, we have only discussed cases in which the expected mutational variance-covariance matrix has perfectly uniform eigenvalues with completely uncorrelated mutational effects on different traits. This is clearly a hypothetical extreme. In reality, we expect the eigenvalues of **M** itself will vary and that mutational effects on different traits will be correlated (e.g., see AJIE *et al.* 2005; ESTES *et al.* 2005). Results suggest that non-uniformity in the eigenvalues of **M** increases non-uniformity in the eigenvalues of **G** (data not shown).

We also found that introducing non-uniformity in the eigenvalues of **M** had an important influence on the eigenvectors of **G**. Simulations for two traits show that, when **M** has uniform eigenvalues, the leading eigenvector of **G** is uniformly distributed between 0º and 360º. In contrast, when the eigenvalues of **M** differ, the leading eigenvector of **G** follows a unimodal, symmetric distribution approximately centered at the leading eigenvector of **M** (results not shown). Thus, uneven eigenvalues of **M** affect the entire expected eigenstructure (i.e., eigenvalues and eigenvectors) of **G** under neutrality.

As for non-zero mutational correlations, their impacts on the eigenstructure of **G** are channeled through their effects on the eigenstructure of **M**. For example, consider a pair of traits each with mutation variance V_m and mutational correlation ρ (as in JONES *et al.* 2003). That is, $\mathbf{M} = V_m \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, which has eigenvalues $V_m(1 \pm \rho)$. The expected eigenvalues of **G** given this **M** are exactly the same as those produced by the mutational covariance matrix $\tilde{\mathbf{M}} = V_m \begin{pmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{pmatrix}$, which has the same eigenvalues as **M** but with uncorrelated mutational effects (results not shown). This suggests that the eigenvalues of **M** —and not the specific patterns of mutational variances and covariances—determine the expected eigenvalue distribution of **G** at equilibrium. Note, however, that the distributions of expected eigenvectors of **G** would differ for the two mutational matrices since the eigenvectors of **M** and $\tilde{\mathbf{M}}$ differ.

Mutational history: We examined two models of mutation: the continuum-of-alleles model (CROW and KIMURA 1964; KIMURA 1965) and the house-of-cards model (KINGMAN 1977; KINGMAN 1978). The two models represent opposing extremes of how to model the effects of mutation. With respect to a genealogy of alleles, under the house-of-cards model the genealogy is only important in shaping allele frequencies, whereas in the continuum-of-alleles model the genealogy shapes both allele frequencies as well as the genotypic values of alleles. Thus, by comparing results from house-of-cards and continuum-of-alleles models, it is possible to disentangle, in part, the effects of variable allele frequency from the historical sequence of mutations on the eigenstructure of **G**.

Both the house-of-cards and continuum-of-alleles models yield, on average, non-uniform eigenvalues at equilibrium (figure 8). Therefore, it appears that variable allele frequencies

contribute to the non-uniformity of eigenvalues of **G** and the mutational history of alleles further increases this non-uniformity.

Pleiotropic versus non-pleiotropic mutation: All of the results discussed thus far have assumed universal pleiotropy. For example, if five characters are modeled then each random mutation could potentially affect all five characters. Under universal pleiotropy variation at each character is determined by the same set of gene trees. At the other extreme, when mutations at a locus are not pleiotropic, each character has a unique set of gene trees underlying its variation, provided that there is some recombination between loci. Figure 9 shows that pleiotropic mutation enhances non-uniformity in eigenvalues of **G** (particularly for lower mutation rates), presumably because the allelic ancestries that underlie variation of different characters overlap. If mutations act non-pleiotropically, then to the extent that loci freely recombine, the allelic ancestries underlying each character's variation will be independent.

*Whole-population properties of **G**:* The results have focused on the properties of **G** based on a sample of size n from a population of size N, where $n < N$. An important question is the extent to which the non-uniformity in **G** is the result of taking a sample from a population. In the APPENDIX, we describe results for the eigenstructure of **G** for an entire population based on individual-based simulations. The results suggest that the eigenvalues of **G** decline at an approximately exponential rate and that the decline is slightly less than the decline based on taking a subsample from the population to estimate the properties of **G**.

DISCUSSION

Our results are consistent with the equilibrium expectation of **G**, $\mathrm{E}(\mathbf{G}) = 2N\mathbf{M}$, predicted by PHILLIPS and MCGUIGAN (2006) under neutral evolution. However, our results demonstrate that

the expected eigenvalues of **G** at equilibrium and the eigenvalues of E(**G**) are highly discordant. For instance, when **M** has uniform eigenvalues associated with it, then E(**G**) has uniform eigenvalues, but the expected eigenvalues of **G** are non-uniform. The practical consequence of this is that although E(**G**) may imply that variation in breeding values is evenly distributed across characters, in fact due to stochasticity, variation in breeding values is never evenly distributed. Some characters by chance will have less genetic variation than others and furthermore, due to stochasticity and the pleiotropic nature of mutation and/or linkage, breeding values will typically covary among characters even if the expectation is independence. Our results suggest that the non-uniformity in the eigenvalues of **G** is not trivial because the non-uniformity is approximately exponential. This approximate exponential decline in eigenvalues may lead to potentially strong constraints on the freedom of a multivariate phenotype to respond to selection.

A key factor shaping the non-uniformity in the eigenvalues of **G** is genealogy. Genealogy causes the effects of alleles at a locus to be dependent. The principle that genealogy causes dependencies is not new and has been a key basis of analyses in phylogenetics (*e.g.* PIAZZA and CAVALLI-SFORZA 1974 and FELSENSTEIN 1985). That genealogy is important in shaping the eigenstructure of **G** has been until now unrecognized. Shared common descent enhances the non-uniformity in the eigenvalues of **G** due to dependencies generated by branching patterns within the genealogy. In addition, multiple mutations may occur within a single branch of a genealogy, and this too can enhance the non-uniformity in the eigenvalues of **G** because there are fewer independently acting mutations. The genealogical ancestry of alleles is intrinsic to biology. The only way to partially overcome the covariance in allelic values that is induced by genealogy is for the value of an allele to be decoupled from its ancestry. The house-of-cards mutation model is one where the value of an allele is decoupled from its ancestry and our results

show that this causes the decline in eigenvalues to be less than if the value of an allele is coupled to its ancestry (such as in the continuum-of-alleles model). Tighter linkage enhances non-uniformity in the eigenvalues of **G** because alleles at different loci do not segregate independently, thus decreasing the number of independent gene genealogies contributing to the structure of **G**.

In general, then, the non-uniformity in the eigenvalues of **G** is exaggerated with more dependency among allelic effects. We found these dependencies can be affected by a variety of other factors as well. For instance, when there are fewer loci or the per locus mutation rate is lower the observed non-uniformity in the eigenvalues of **G** is increased. The degree of pleiotropy of new mutations affects the dependency of allelic effects. Under universal pleiotropy all characters are influenced by the same set of gene trees and this overlap increases the observed non-uniformity in the eigenvalues of **G**. In contrast, in the absence of pleiotropy, each character has a unique set of gene trees (as long as there is some recombination), and this independence decreases correlations between genotypic values and reduces non-uniformity in the expected eigenvalues of **G**.

What leads to the discrepancy between the expected eigenvalues of **G** and the eigenvalues of E(**G**)? Mathematically, the discordance traces to the fact that eigenvalues are nonlinear functions of the elements of **G**, and hence the expectations of the eigenvalues of **G** will not generally be equal to the eigenvalues of the expectation of **G**. This can be seen more directly in the following two-trait case. Suppose the expected **G** matrix is diagonal with the repeated eigenvalue $\hat{\lambda}$: $E(\mathbf{G}) = \begin{pmatrix} \hat{\lambda} & 0 \\ 0 & \hat{\lambda} \end{pmatrix}$, where $\hat{\lambda} > 0$. A realized **G** under neutral evolution can be viewed as a perturbation of E(**G**), say, $\mathbf{G} = \begin{pmatrix} \hat{\lambda} + \varepsilon_1 & \varepsilon_2 \\ \varepsilon_2 & \hat{\lambda} + \varepsilon_3 \end{pmatrix}$, where the ε_i's are random variables

with zero means and variances much smaller than $\hat{\lambda}$. If we imagine that the perturbations are drawn, say, from a zero-mean normal distribution with variance $\sigma^2 << \hat{\lambda}$ then it is not hard to show that the expected eigenvalues of **G** are not uniform (i.e., not repeated). Indeed, the expected leading eigenvalue of **G** exceeds $\hat{\lambda}$ by at least $\sigma\sqrt{2/\pi}$ whereas the smaller eigenvalue of **G** is expected to fall below $\hat{\lambda}$ by at least the same amount. This suggests that non-uniformity among the expected eigenvalues of **G** is driven by the variability of perturbations from E(**G**). This variability arises from the assorted sources of stochasticity inherent to evolution in finite populations.

That the eigenvalues of random covariance matrices (such as a **G** matrix) are non-uniform is not a new result (*e.g.*, see a review by EDELMAN and RAO [2005] and WAGNER'S [1984] random pleiotropy model). Additionally, similar patterns found in this paper for the average of ranked eigenvalues are observed if it is assumed that the variances for characters are independently drawn from a χ^2 for a given degrees of freedom. In particular, a deviation from a purely exponential decline occurs as was observed in figure 7. Broadly, stochasticity causes the non-uniformity in eigenvalues, but there are biologically unique processes that contribute to this stochasticity. This paper suggests that genealogical ancestry is one important process.

Our neutral theoretical predictions for the expected structure of **G** at recombination-mutation-drift balance form an important set of null expectations for empirical studies of multivariate evolution. In particular, our finding that, under neutrality, the eigenvalues of **G** tend to be unequal and decay approximately exponentially from largest to smallest is consistent with empirical estimates of **G** (e.g., MEZEY and HOULE 2005). However, our results also show that the rate of decay is sensitive in predictable ways to numerous factors including population size, the number and patterns of linkage among loci, and the rate of and modality of mutation. If one

includes selection as well (as in JONES *et al.* 2003), current theoretical considerations suggest that inferring underlying process from an equilibrium estimate of **G** may be difficult because different underlying processes have similar effects on the eigenvalues associated with **G**. For instance, it may be difficult to disentangle mutation rates from numbers of loci and the pattern of recombination among loci. Single locus expectations and statistics may be developed, but this ignores the likely possibility that more than one locus contributes variation to **G**. Rather, to infer processes that impact the evolution of **G** may require direct approaches (see ESTES *et al.* 2004) or possibly the study of transient changes in **G** (PHILLIPS *et al.* 2001; WHITLOCK *et al.* 2002).

ACKNOWLEDGEMENTS

We thank P. Phillips and K. McGuigan for allowing us to cite their then unpublished work, and the National Science Foundation (grants DEB 0337582, EF 0328594, and EF 0531870 to RG) for financial support. The IBEST program at the University of Idaho kindly provided computer resources and they are supported by NSF grant EPS 0080935 and NIH grants P20 RR16454 and P20 RR16448 from the COBRE and INBRE programs of the National Center for Research Resources. Comments by Ruth Shaw, Joe Felsenstein and an anonymous reviewer improved the presentation and content of this paper. We thank the anonymous reviewer for suggesting that we examine the case in which the variance for each character is independent and χ^2 distributed.

APPENDIX

Whole-population predictions for G: To investigate neutral evolution of **G** for an entire population we used forward time, individual-based simulations. Our approach is similar to that of JONES *et al*. (2003) except for the absence of selection and our use of Wright-Fisher mating (i.e., sampling parents with replacement). The individual-based simulations were founded on the same model assumptions as the two coalescent-based methods. Events in the individual-based simulations were ordered mutation $\rightarrow$ recombination $\rightarrow$ random mating within generations. At the beginning of a replicate simulation run, each individual was initialized with a genotypic value of zero by assigning allelic values of zero to all loci. Population sizes of 140 and 1366 were simulated. An inspection of simulations suggests that populations reach quasi-recombination-mutation-drift balance by 12,000 generations for populations consisting of 140 and 1366 individuals given our mutation rate and rates of recombination. The basis for this inference was that the average **G**-matrix from simulations was not significantly different from $2N\mathbf{M}$ the expected **G**-matrix. For each replicate run, we calculated the **G** matrix at generation 12,000 using the genotypic values of all individuals in the population. We ran 20 replicates for a particular parameter set when using 50 loci, and 200 replicates when using 2 and 5 loci.

Some of the parameters used in both the individual-based simulations were similar to those utilized by JONES *et al*. (2003). For example, the per haploid locus mutation rate utilized was 0.0002 and all traits were uncorrelated with an expected per character variance in the effects of mutations of 0.05.

Results: Figure A1 illustrates that when the whole population is modeled, the eigenvalues of **G** decline at an approximately exponential rate and that the rate of decline is greater for smaller populations. Figure A2 suggests that non-independent segregation of alleles among loci

enhances the non-uniformity in the eigenvalues of **G** when the whole population is modeled. Overall these whole-population simulations combined with the coalescent simulations suggest that the non-uniformity in the eigenvalues of **G** is a property of both an entire population and subsamples from the population. Furthermore, a comparison of individual-based and coalescent results in figures A1 and A2 suggests that estimating the eigenstructure of **G** based on a subsample from a population (as is done with the coalescent) results in slightly enhanced non-uniformity in the eigenvalues of **G** relative to the true value of the whole population.

REFERENCES

AJIE, B. C., S. ESTES, M. LYNCH and P. C. PHILLIPS, 2005 Behavioral degradation under mutation accumulation in Caenorhabditis elegans. Genetics **170:** 655-660.

AVERY, P. J., and W. G. HILL, 1977 Variability in genetic parameters among small populations. Genetical Research **29:** 193-213.

BLOWS, M. W., S. F. CHENOWETH and E. HINE, 2004 Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. American Naturalist **163:** E329-E340.

CROW, J. F., and M. KIMURA, 1964 The theory of genetic loads, pp. 495-505 in *Proceedings of the XI international congress of genetics*. Pergamon Press, Oxford, UK.

EDELMAN, A., and N. R. RAO, 2005 Random matrix theory. Acta Numerica **14:** 233-297.

ESTES, S., B. C. AJIE, M. LYNCH and P. C. PHILLIPS, 2005 Spontaneous mutational correlations for life-history, morphological and behavioral characters in Caenorhabditis elegans. Genetics **170:** 645-653.

ESTES, S., P. C. PHILLIPS, D. R. DENVER, W. K. THOMAS and M. LYNCH, 2004 Mutation accumulation in populations of varying size: The distribution of mutational effects for fitness correlates in Caenorhabditis elegans. Genetics **166:** 1269-1279.

EWENS, W. J., 2004 *Mathematical population genetics. I. Theoretical introduction*. Springer, New York.

FALCONER, D. S., and T. F. C. MACKAY, 1996 *Introduction to Quantitative Genetics*. Longman, Essex, England.

GRIFFITHS, R. C., 1981 Neutral two-locus multiple allele models with recombination. Theoretical Population Biology **22:** 169-186.

GRIFFITHS, R. C., 1991 The two-locus ancestor graph in *Selected proceedings of the symposium of applied probability*, edited by I. V. BASAWA and R. L. TAYLOR.

GRIFFITHS, R. C., and P. MARJORAM, 1996 Ancestral inference from samples of DNA sequences with recombination. Journal of Computational Biology **3:** 479-502.

GRIFFITHS, R. C., and P. MARJORAM, 1997 An ancestral recombination graph, pp. 257-270 in *Progress in population genetics and human evolution* edited by P. DONNELLY and S. TAVARÉ. Springer Verlag, New York, NY.

HINE, E., and M. W. BLOWS, 2006 Determining the effective dimensionality of the genetic variance-covariance matrix. Genetics **173**.

HINE, E., S. F. CHENOWETH and M. W. BLOWS, 2004 Multivariate quantitative genetics and the lek paradox: Genetic variance in male sexually selected traits of Drosophila serrata under field conditions. Evolution **58:** 2754-2762.

HUDSON, R. R., 1991 Gene genealogies and the coalescent process. Oxford Surveys in Evolutionary Biology **7:** 1-44.

JONES, A. G., S. J. ARNOLD and R. BÜRGER, 2003 Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. Evolution **57:** 1747-1760.

KIMURA, M., 1965 A stochastic model concerning the maintenance of genetic variability in quantitative characters. Proceedings of the National Academy of Science USA **54:** 731-736.

KINGMAN, J. F. C., 1977 On the properties of binlinear models for the balance between genetic mutation and selection. Philosophical proceedings of the Cambridge philosophical society **81 :** 443-453.

KINGMAN, J. F. C., 1978 A simple model for the balance between selection and mutation. Journal of Applied Probability **15:** 1-12.

KINGMAN, J. F. C., 1982 The coalescent. Stochastic processes and their applications **13:** 235-248.

KIRKPATRICK, M., and D. LOFSVOLD, 1992 Measuring selection and constraint in the evolution of growth. Evolution **46:** 954-971.

LANDE, R., 1979 Quantitative genetics analysis of multivariate evolution, applied to brain:body size allometry. Evolution **33:** 402-416.

LANDE, R., 1980 The genetic covariance between characters maintained by pleiotropic mutations. Genetics **94:** 203-215.

LANDE, R., and S. J. ARNOLD, 1983 The measurement of selection on correlated characters. Evolution **37:** 1210-1226.

LYNCH, M., and W. G. HILL, 1986 Phenotypic evolution by neutral mutation. Evolution **40:** 915-935.

LYNCH, M., and B. WALSH, 1998 *Genetics and analysis of quantitative traits*. Sinauer Associates, Inc., Sunderland, MA USA.

MEZEY, J. G., and D. HOULE, 2005 The dimensionality of genetic variation for wing shape in Drosophila melanogaster. Evolution **59:** 1027-1038.

NORDBORG, M., 2001 Coalescent theory, pp. 179-212 in *Handbook of statistical genetics*, edited by D. J. BALDING, M. BISHOP and C. CANNINGS. John Wiley & Sons, Ltd., Chichester, England.

PIAZZA, A. and L.L. CAVALLI-SFORZA, 1974 Spectral analysis of patterned covariance matrices and evolutionary relationships in *Proceedings of the Eighth International Conference on Numerical Taxonomy,* edited by G.F. ESTABROOK. W.H. Freeman, San Francisco.

PHILLIPS, P. C., and S. J. ARNOLD, 1989 Visualizing multivariate selection. Evolution **43:** 1209-1222.

PHILLIPS, P. C., and K. L. MCGUIGAN, 2006 Evolution of genetic variance-covariance structure, pp. 310-325 in *Evolutionary genetics: concepts and case studies*, edited by C. W. FOX and J. B. WOLF. Oxford University Press, Oxford, England.

PHILLIPS, P. C., M. C. WHITLOCK and K. FOWLER, 2001 Inbreeding changes the shape of the genetic covariance matrix in Drosophila melanogaster. Genetics **158:** 1137-1145.

SCHLUTER, D., 1996 Adaptive radiation along genetic lines of least resistance. Evolution **50:** 1766-1774.

WAGNER, G. P., 1984 On the eigenvalue distribution of genetic and phenotypic dispersion matrices: evidence for a nonrandom organization of quantitative character variation. Journal of Mathematical Biology **21:** 77-95.

WHITLOCK, M. C., P. C. PHILLIPS and K. FOWLER, 2002 Persistence of changes in the genetic covariance matrix after a bottleneck. Evolution **56:** 1968-1975.

ZENG, Z.-B., and C. C. COCKERHAM, 1991 Variance of neutral genetic variances within and between populations for a quantitative character. Genetics **129:** 535-553.

Table 1: Limits to multivariate trait variation. When the effective population size or genomic mutation rate for a set of characters is low enough, some principal components are expected to have zero variance associated with them. For instance, given below are the eigenvalues ranked in order from largest to smallest when $4N\mu = 0.1$ per locus and the number of loci is ten. The **M** matrix had dimensionality of 25 with the same variance for each principal component.

	Sample size	
Eigenvalue rank	75	300
1	6.41e-1 +/- 3.20e-2	6.22e-1 +/- 2.80e-2
2	3.09e-1 +/- 1.86e-2	3.08e-1 +/- 1.74e-2
3	1.56e-1 +/- 1.25e-2	1.56e-1 +/-1.17e-2
4	7.37e-2 +/- 7.32e-3	7.67e-2 +/- 7.22e-3
5	3.63e-2 +/- 4.69e-3	3.87e-2 +/- 4.62e-3
6	1.72e-2 +/- 2.71e-3	1.92e-2 +/- 2.78e-3
7	7.85e-3 +/- 1.60e-3	9.48e-3 +/- 1.69e-3
8	3.31e-3 +/- 8.07e-4	4.38e-3 +/- 9.05e-4
9	1.27e-3 +/- 4.42e-4	1.91e-3 +/- 5.04e-4
10	5.23e-4 +/- 2.44e-4	8.12e-4 +/- 2.95e-4
11	2.02e-4 +/- 1.34e-4	3.33e-4 +/- 1.74e-4
12	9.08e-6 +/- 1.78e-5	1.27e-4 +/- 9.90e-5
13	0.0	6.02e-5 +/- 5.00e-5
14	0.0	1.49e-5 +/- 1.60e-5
15	0.0	1.19e-6 +/- 2.34e-6
16	0.0	0.0

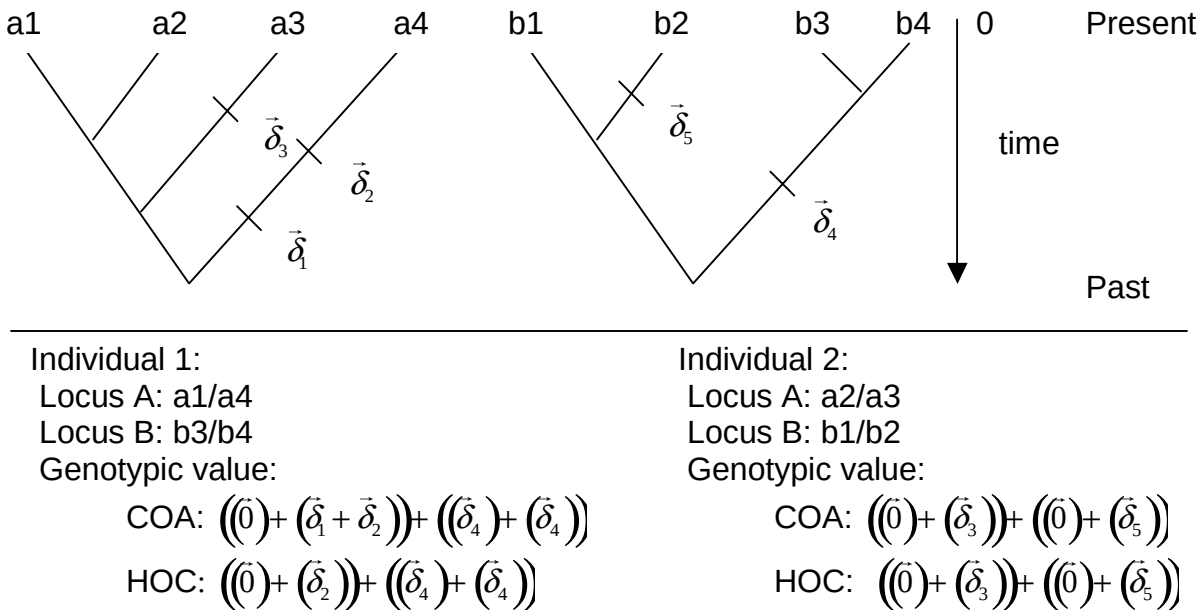
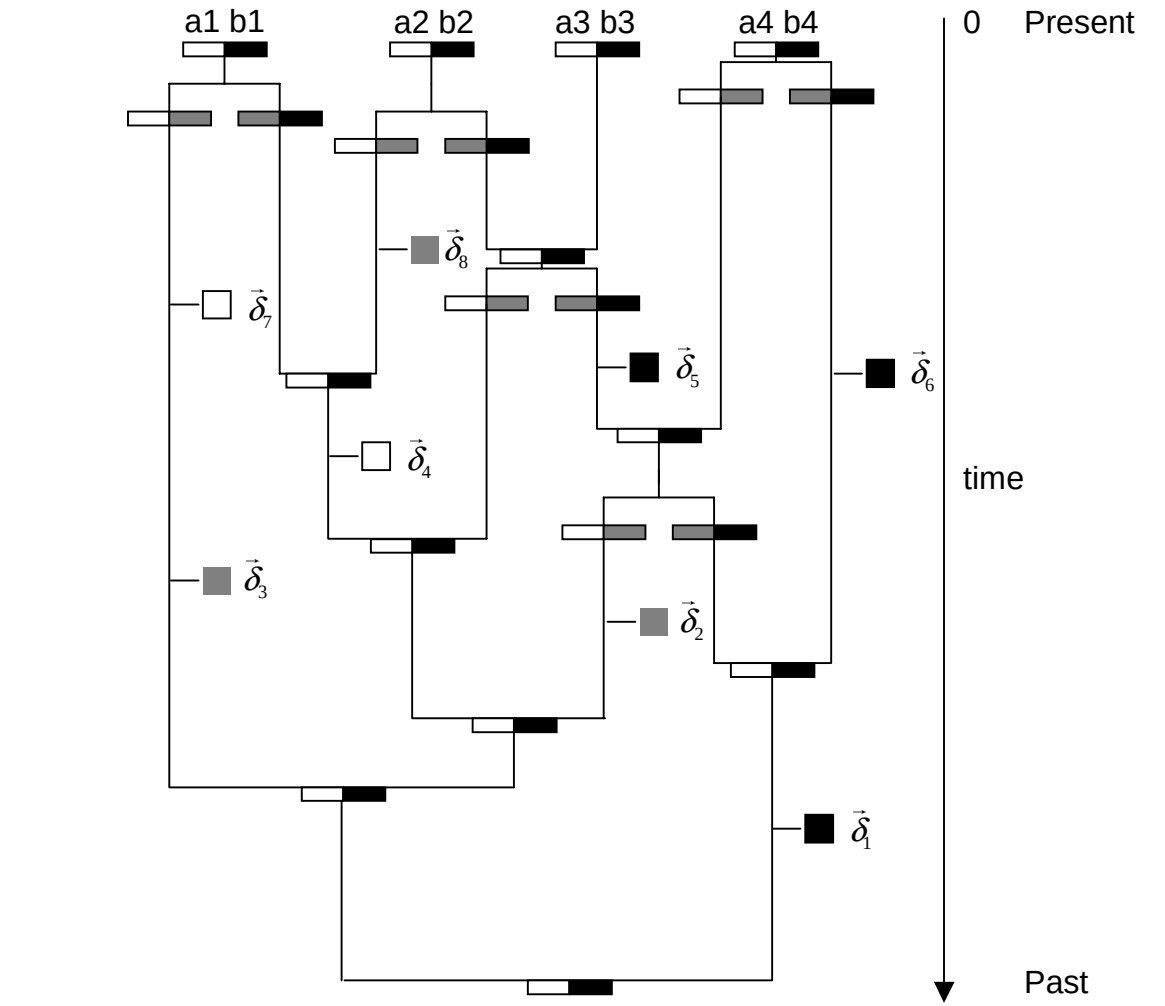


Figure 1: Demonstration of how individual genotypic values were generated using the neutral coalescent in which the evolutionary histories of loci are completely independent. In the figure, it is assumed that there are two individuals in a sample and each individual is composed of two loci. The gene tree on the left is that for locus *A* and the gene tree on the right is that for locus *B*. For each individual two alleles from each locus are chosen without replacement. Genotypic values under the continuum-of-alleles (COA) and house-of-cards (HOC) mutation models are given. See main text for further details.



Individual 1:
Locus A: a1/a4
Locus B: b1/b4
Genotypic value:

COA: $\left(\left(\vec{\delta}_7\right)+\left(\vec{0}\right)\right)+\left(\left(\vec{0}\right)+\left(\vec{\delta}_6+\vec{\delta}_1\right)\right)$

HOC: $\left(\left(\vec{\delta}_7\right)+\left(\vec{0}\right)\right)+\left(\left(\vec{0}\right)+\left(\vec{\delta}_6\right)\right)$

Individual 2:
Locus A: a2/a3
Locus B: b2/b3
Genotypic value:

COA: $\left(\left(\vec{\delta}_4\right)+\left(\vec{0}\right)\right)+\left(\left(\vec{\delta}_5+\vec{\delta}_1\right)+\left(\vec{\delta}_5+\vec{\delta}_1\right)\right)$

HOC: $\left(\left(\vec{\delta}_4\right)+\left(\vec{0}\right)\right)+\left(\left(\vec{\delta}_5\right)+\left(\vec{\delta}_5\right)\right)$

Figure 2: Demonstration of how individual genotypic values were generated using the ancestral recombination graph. Now loci are linked such that individuals inherit haplotypes as opposed to independent alleles. Haplotypes are presented as rectangles and each haplotype consists of two loci: locus *A* is unshaded and locus *B* is shaded black. Points at which a lineage bifurcates into two lineages (back in time) correspond to recombination events. Where lineages join are coalescent events. In this example, each individual consists of a random sample of two haplotypes (without replacement). Genotypic values under the continuum-of-alleles (COA) and house-of-cards (HOC) mutation models are given. See main text for further details.

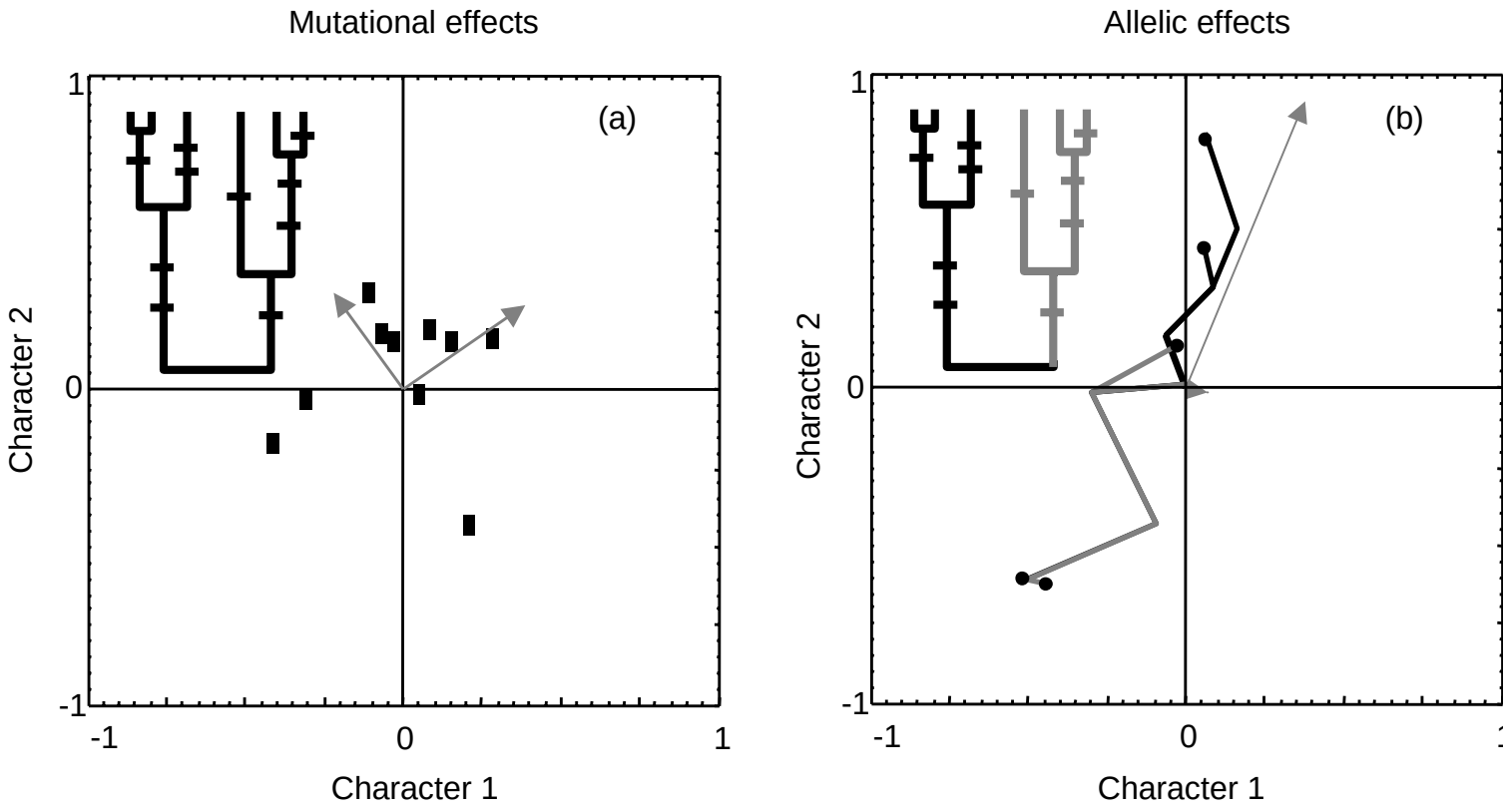


Figure 3: The effect of genealogy at a single locus on the eigenstructure of allelic variation. In the upper left corner of both panels is an illustration of a genealogy of six alleles with ten mutations. (a) The bivariate effects of ten random mutations are plotted assuming normality, where the average effect of a mutation per character is zero and effects are uncorrelated. The gray arrows represent the principal components of the variation in mutational effects with the length of each arrow proportional to the variance explained by each direction. The arrows are close in length suggesting that mutational variation is evenly distributed across the two dimensions. (b) The effects of the six alleles at the tips of the genealogy are plotted as black dots. There are only five dots because two alleles are identical in state. Lines of descent, shaded to aid comparison with the genealogy, connect allelic values. As in (a), the gray arrows indicate the principal components of allelic variation. Comparing (a) and (b) suggests that shared common ancestry leads to a reduction in the dimensionality of allelic variation relative to mutational variation. Allelic variation is concentrated along one principal component axis, whereas mutational variation is distributed relatively evenly over two principal component axes.

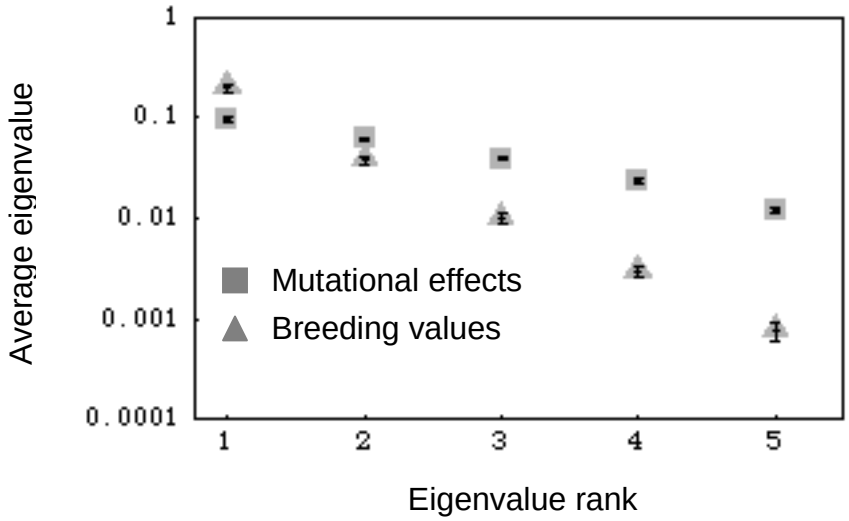


Figure 4: The effects of random mutation and genealogy on the eigenvalues of **G**. Shown are results for a single locus that pleiotropically affects five characters with $4N\mu = 1.0$ and a sample size of 300. To generate this figure 500 replicate coalescent trees were generated and random mutation placed on the tree according to a Poisson process. Mutations were drawn from a multivariate normal distribution with zero mean and uncorrelated effects. To determine the eigenvalues of mutational effects, for each tree a covariance matrix was estimated based on the effects of mutations on the tree (as in figure 3a). Allelic effects were calculated assuming the continuum-of-alleles model. Individual breeding values were calculated by sampling without replacement two alleles from a coalescent tree. In the figure, eigenvalues are on a log scale and error bars represent 95% confidence intervals of the mean eigenvalue. The internal hash mark within an error bar is the mean eigenvalue.

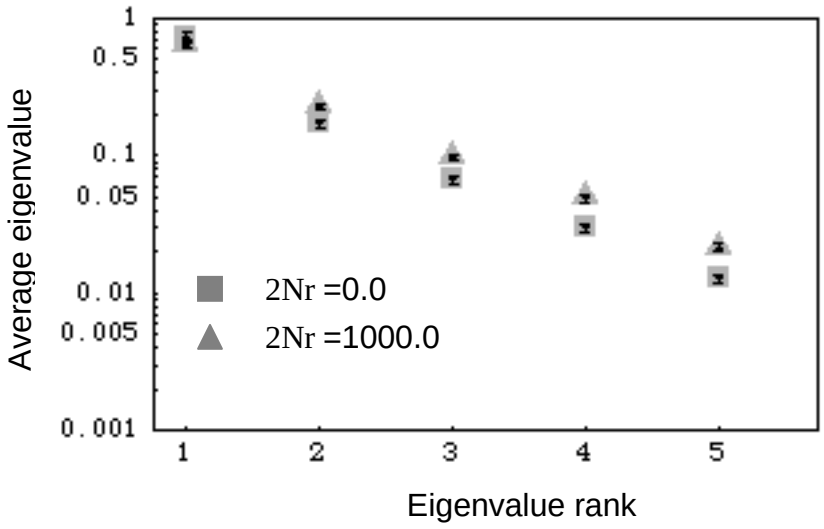


Figure 5: The effects of recombination on the eigenvalues of **G**. A sample of 300 individuals was drawn from a population in which $4N\mu = 1.0$ per locus and four loci were modeled using the ancestral recombination graph. *2Nr* is given as the rate of recombination per adjacent pair of loci, thus the total genomic recombination rate is *6Nr*. Mutations pleiotropically affected five characters. 500 replicate recombination graphs were simulated. In the figure, eigenvalues are on a log scale and error bars represent 95% confidence intervals of the mean eigenvalue. The internal hash mark within an error bar is the mean eigenvalue.

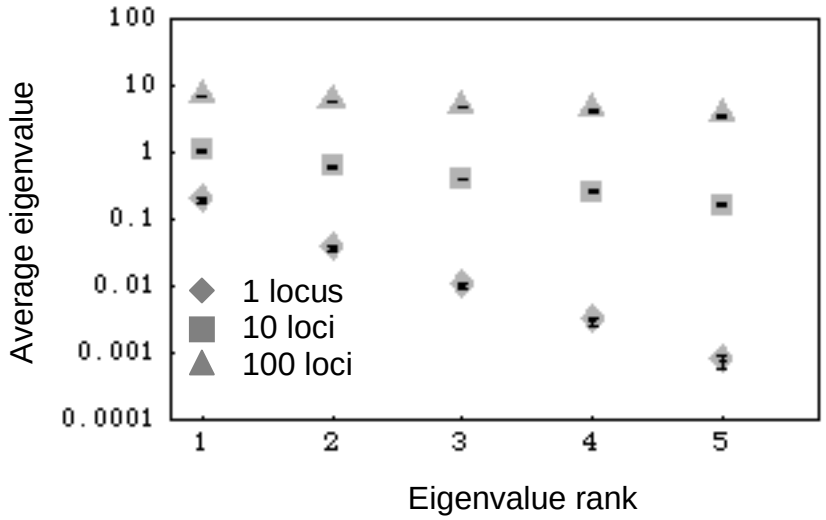


Figure 6: The effects of varying the number of loci on the eigenvalues of **G**. A sample of 300 individuals was drawn from a population in which $4N\mu = 1.0$ per locus. For the ten and 100 loci cases, independent gene trees for each locus were simulated using Hudson's (1990) algorithm. Mean eigenvalues and their confidence intervals are based on 500 replicates. In the figure, eigenvalues are on a log scale and error bars represent 95% confidence intervals of the mean eigenvalue. The internal hash mark within an error bar is the mean eigenvalue.

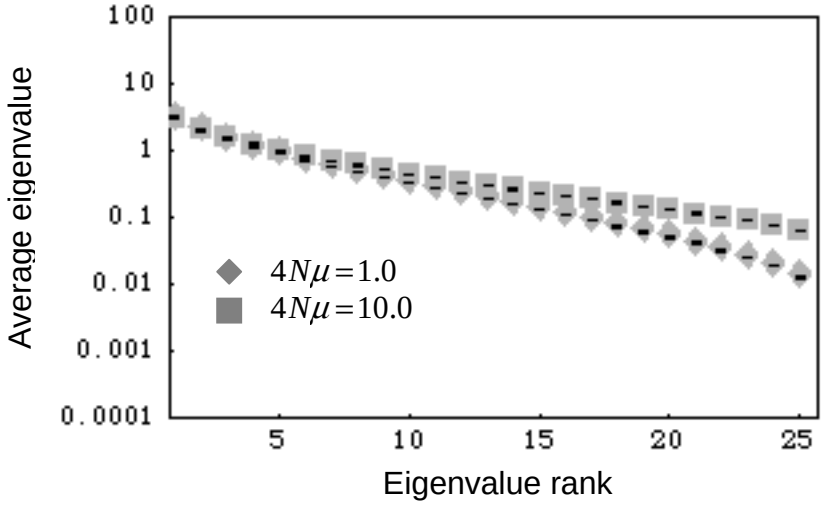


Figure 7: The consequence of varying $4N\mu$ (per locus) with a fixed number of loci (ten). To aid direct comparison between the different $4N\mu$ values with respect to how rank eigenvalues decline, the points for $4N\mu = 10.0$ were all shifted down an equal amount such that the average leading eigenvalues for $4N\mu = 1.0$ and $4N\mu = 10.0$ coincide. The figure demonstrates that different per locus mutation rates (or effective population sizes) lead to different rates of decrease in the eigenvalues, with smaller $4N\mu$ values showing a faster decline. In the figure n=300 and the independent locus coalescent approach was used. Eigenvalues are on a log scale and error bars represent 95% confidence intervals of the mean eigenvalue. The internal hash mark within an error bar is the mean eigenvalue.

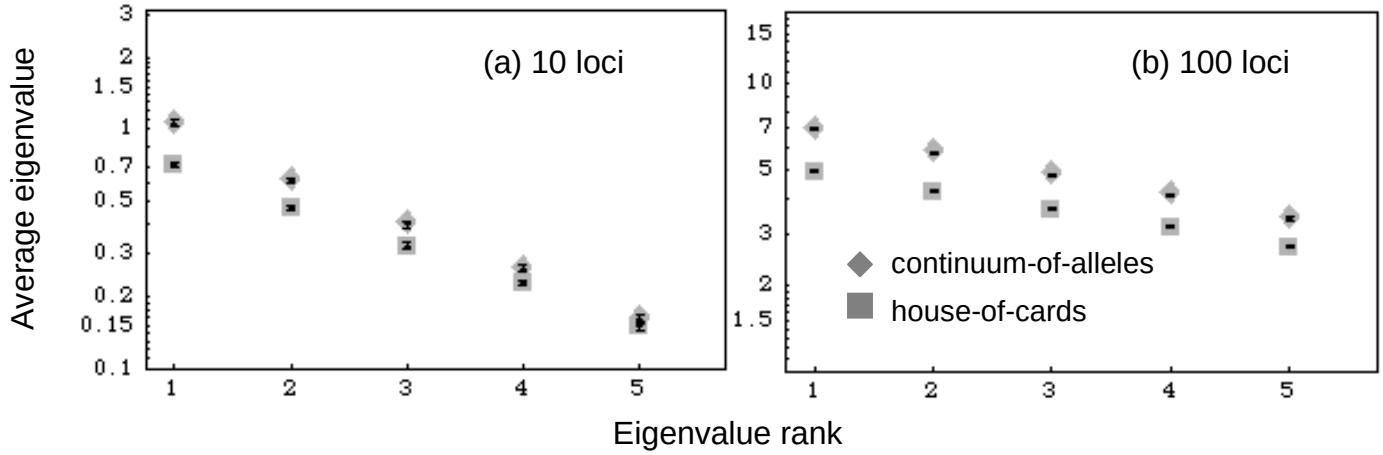


Figure 8: The effect of different mutational models: house-of-cards versus continuum-of-alleles. $4N\mu = 1.0$ for each locus and 300 individuals were sampled. The independent locus coalescent approach was used. Estimates are based on 500 replicates. In the figure, eigenvalues are on a log scale and error bars represent 95% confidence intervals of the mean eigenvalue. The internal hash mark within an error bar is the mean eigenvalue.

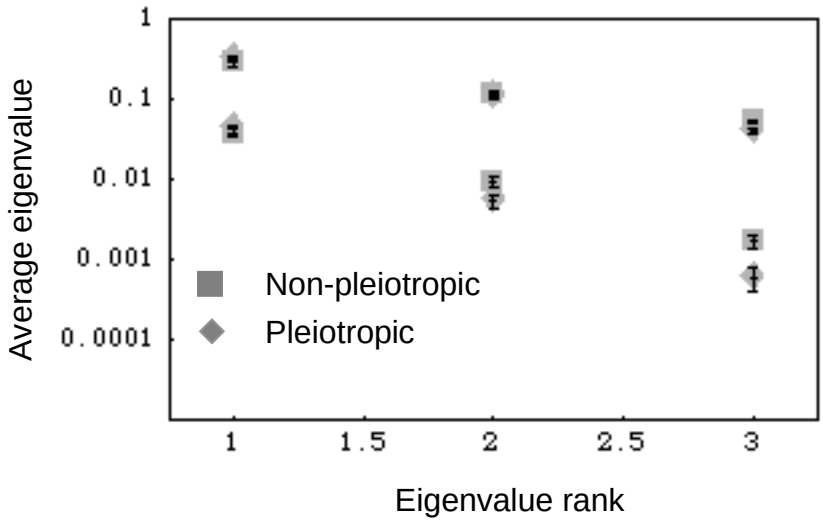


Figure 9: The effects of pleiotropy on the non-uniformity in **G**. Three characters and loci were modeled. In the pleiotropy cases, mutations at the three loci affect all characters, whereas in the no pleiotropy case a single locus only affects one of the characters. The upper set of points correspond to a *per character* $4N\mu = 3.0$ and the lower set of points correspond to a *per character* $4N\mu = 0.3$. Mutation rate was modeled on a per character basis so that a direct comparison between the pleiotropy and no pleiotropy case can be made. The ancestral-recombination graph was simulated in which the genomic recombination rate was $4Nr = 2000.0$. In the figure, eigenvalues are on a log scale and error bars represent 95% confidence intervals of the mean eigenvalue. The internal hash mark within an error bar is the mean eigenvalue.

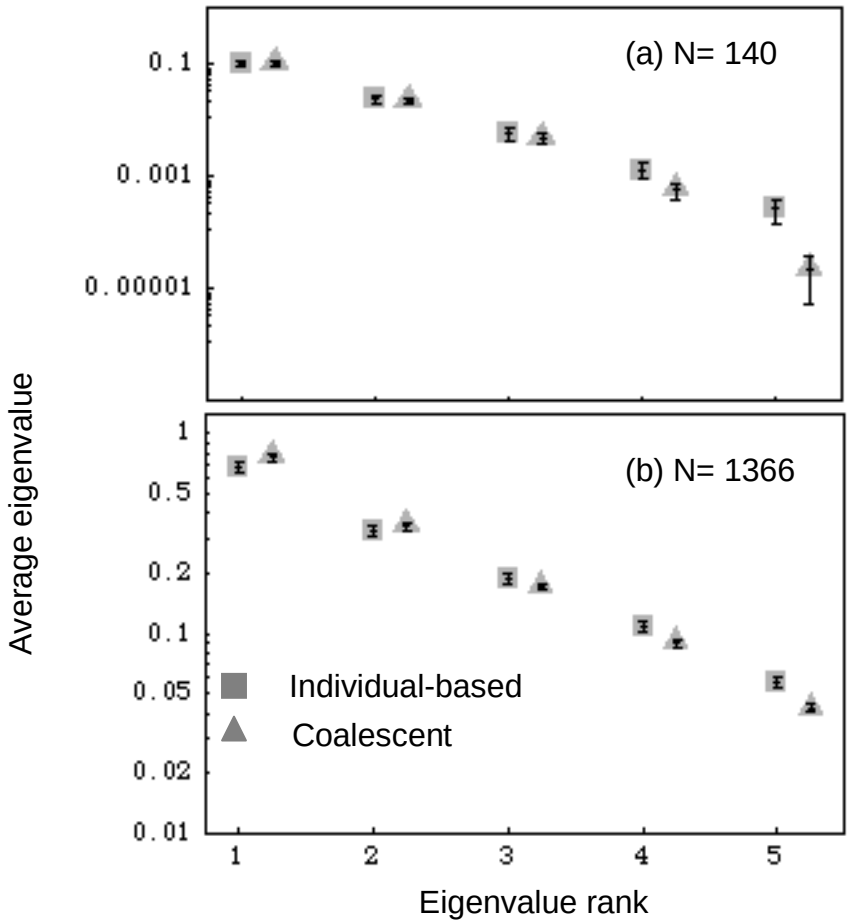


Figure A1: The eigenvalues of **G** for an entire population based on individual-based simulations (square points) for small (a) and large population sizes (b). Mutations pleiotropically affected five characters at five loci. The continuum of alleles mutation model was used. In the individual-based simulations, the recombination rate (*r*) was 0.5 per locus pair. For comparison, results based on the independent locus coalescent are provided for a sample of ten (a) and thirty (b) individuals from populations of size 140 and 1366, respectively. The per locus mutation rate and number of loci was the same in the coalescent and individual-based simulations. Results are based on 200 replicates for both the individual-based and 500 replicates for the coalescent simulations. In the figure, eigenvalues are on a log scale and error bars represent 95% confidence intervals of the mean eigenvalue. The internal hash mark within an error bar is the mean eigenvalue. Points using the coalescent approach are offset relative to points using the individual-based approach to aid comparison.

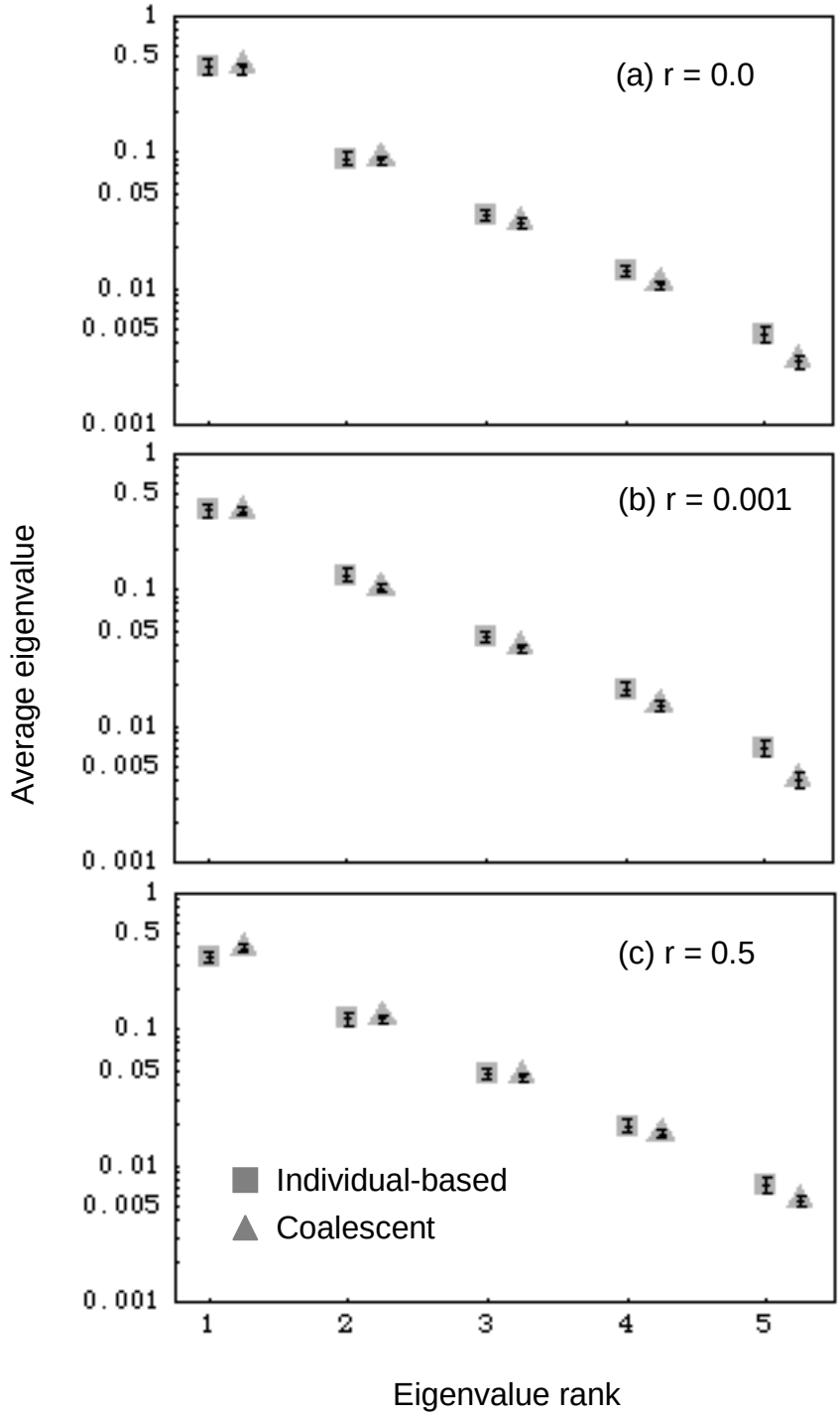


Figure A2: The eigenvalues of **G** for an entire population based on individual-based simulations and when the recombination rate varies between loci (square points). The population size in all cases was *N=1366*. Mutations pleiotropically affected five characters at two loci. The continuum-of-alleles mutation model was used. For comparison, results based on the ancestral recombination graph are provided (triangular points) in which the same number of loci, population size, per locus mutation rate and recombination rate is modeled, except now results are based on a sample of thirty individuals from the population. Results are based on 200

replicates for the individual-based and 500 replicates for the coalescent simulations. In the figure, eigenvalues are on a log scale and error bars represent 95% confidence intervals of the mean eigenvalue. The internal hash mark within an error bar is the mean eigenvalue. Points using the ancestral recombination graph approach are offset relative to points using the individual-based approach to aid comparison.