Published in final edited form as:

Genetica. 2008 February ; 132(2): 187–198. doi:10.1007/s10709-007-9163-4.

Testing the rare-alleles model of quantitative variation by artificial selection

John K. Kelly

Department of Ecology and Evolutionary Biology, University of Kansas, 1200 Sunnyside Ave,
Lawrence, KS 66045-7534, USA

Abstract

The rare-alleles model of quantitative variation posits that a common allele (the 'wild-type') and one or more rare alleles segregate at each locus affecting a quantitative trait; a scenario predicted by several distinct evolutionary hypotheses. Single locus arguments suggest that artificial selection should substantially increase the genetic variance (V_g) if the rare-alleles model is accurate. This paper tests the ' ΔV_g prediction' using a large artificial selection experiment on flower size of *Mimulus guttatus*. V_g for flower size does evolve, increasing with selection for larger flower while decreasing in the other direction. These data are consistent with a model in which flower size variation is caused by rare, partially dominant alleles. However, this explanation becomes increasingly tenuous when considered with other data (correlated responses to selection and the effects of inbreeding). A combination of modern (marker-based mapping) and classical (biometric) techniques will likely to be required to determine the distribution of allele frequencies at loci influencing quantitative traits.

Keywords

Artificial selection; *Mimulus guttatus*; Quantitative traits; QTL; Rare alleles

Introduction

The prediction that genetic variation in quantitative traits should be caused by rare alleles emerges from several distinct evolutionary models. The simplest is that trait variation is caused by the pleiotropic effects of deleterious mutations (Barton 1990; Caballero and Keightley 1994; Keightley and Hill 1990; Kondrashov and Turelli 1992). Allele frequencies may also be highly asymmetric under direct stabilizing selection (Kimura 1969; Turelli 1984; Turelli 1987). Here, rare alleles are only conditionally deleterious, displacing the phenotype of an individual from a temporally constant fitness optimum. In a structured population, a particular allele may be rare in some populations but near fixation in others, particularly if there is geographical variation in selection. The balance between selection and migration can produce the rare-alleles architecture within local populations even if the polymorphism is at intermediate frequency in the species as a whole.

A number of predictions for the rare-alleles model are simple extrapolations from single locus arguments (Charlesworth and Hughes 2000; Curtsinger and Ming 1997; Houle et al. 1996; Kelly 1999). Figure 1 illustrates how the genetic variance, V_g , contributed by a single locus depends on allele frequency. V_g is least when allele frequencies are extreme (q is close to 0 or 1). If allele frequencies are made more intermediate by imposing selection on a population, V_g

should increase (Sorensen and Hill 1982). When rare alleles are consistent in the direction of their effects, perhaps reducing phenotype, V_g is predicted to increase in only one direction of selection. In the opposite direction of selection, V_g should decline rapidly and response will plateau as the common alleles fix. If the direction of effects is variable then selection will increase the contribution of some loci to V_g while reducing that of others. However, because the potential magnitude of increases is much greater than that for declines, V_g is predicted to increase in both directions of selection. For an additive locus, a 25-fold increase in V_g is predicted as q goes from 0.01 to 0.5 and an even greater increase (proportionally) will occur if rare alleles are at least partially recessive (Fig. 1). Hereafter, I use the term ' ΔV_g prediction' to denote the expectation that V_g will increase in at least one direction of selection.

Curtsinger and Ming (1997) provide 'proof of concept' that artificial selection can identify a rare-alleles basis for quantitative trait variation. Using controlled crosses, they synthesized a population of fruit flies in which ethanol resistance alleles were segregating at low frequencies. They then imposed selection for increased ethanol resistance on replicate populations. As the favoured alleles increased to intermediate frequency, the response to selection accelerated, i.e. the change in the mean per generation increased. This acceleration was repeatable across replicates (see Fig. 2 of their paper) and implies that selection caused an increase in V_g . Eventually, the favoured alleles became predominant within populations, V_g declined, and response stabilized.

Large population sizes, approximately 1000 individuals scored and 200 selected in each generation, were a key feature of the Curtsinger and Ming (1997) experiment. Genetic drift is a dominant factor in most selection experiments (Hill 1971, 1980). Drift can obscure the predicted increase in V_g under the rare-alleles model, both through its direct effect on allele frequencies as well as by generating linkage disequilibria among Quantitative Trait Loci (QTL). A recent simulation study by Zhang and Hill (2005) has shown that drift within selected populations, combined with sampling effects prior to population initiation, and selection-induced disequilibria (Bulmer and/or Hill-Robertson effects; Bulmer 1976; Hill and Robertson 1966), can effectively cancel the ΔV_g prediction. The practical implication is that each of these factors must be accommodated within an experimental design to provide a robust test of the rare-alleles model.

In this paper, I describe a large-scale selection experiment on the trait of corolla width (flower size) in the wildflower *Mimulus guttatus*. Selection is imposed in both directions, with 1000 plants measured and 200 selected within each population over six generations. The selected populations were established from a large random sampling of the natural population. The experiment is accompanied by a computer simulation study tailored specifically to the experimental design (sample sizes, duration of selection, breeding design, etc) and study system (the genetic map of *M. guttatus*). The simulation demonstrates the experimental design to be robust: if variation in the ancestral population is caused by rare alleles, selection should produce a detectable increase in V_g .

The present design differs from Curtsinger and Ming (1997) in at least four notable ways. First, I impose selection in both directions (not simply for higher trait values) because there is no *a priori* reason to assume that rare alleles are unidirectional in their effects. Second, I use controlled crosses and comparison among relatives, i.e. a breeding design, to estimate V_g within selected and control populations. This is because accelerations (or decelerations) in rate that the mean evolves can be caused by environmental differences among generations. Controlled crosses also serve to diminish linkage disequilibria generated by selection. Third, this study starts from a random genetic sample of a single natural population and not a laboratory synthesized population. Finally, I was unable to replicate the direction of selection

owing to space constraints. The simulation study speaks to the repeatability of results from this design.

Materials and methods

Study species and source population

Mimulus guttatus ($2n = 28$; Phrymaceae), the yellow monkeyflower, is a North American wildflower increasingly used as a model organism for genetic studies of floral variation and the evolution of plant mating systems (Macnair and Cumbes 1989; Ritland 1989; Fenster and Ritland 1994; Carr and Fenster 1994; Robertson et al. 1994). It ranges from Mexico to Alaska, generally west of the Rocky Mountains. The species is mixed-mating and the estimated selfing rate varies from 0 to 0.75 among populations (Awadalla and Ritland 1997; Ritland and Ganders 1987; Sweigart et al. 1999; Willis 1993b). Life-history also varies geographically with some populations annual or winter annual and others perennial.

Three experimental populations of *M. guttatus* (Low, Control, and High) were established from a common source, the Zia-1 base population (see Arathi and Kelly 2004). This population consists of 596 outbred full-sib families derived from a single natural population located on Iron Mountain in Oregon. The parents of these families, 1192 unrelated plants, were grown from seed that was collected from field samples over four distinct years (each parent from a distinct maternal plant). Iron Mountain is a large population containing hundreds of thousands of reproductive adults continuously distributed over an area of approximately 400 square meters. The plants are annual (or winter annual) and predominantly outcrossing (Willis 1993b, 1996, 1999a, b). Progeny from the Zia-1 families were dispersed uniformly among the three experimental populations to minimize their initial differences. This initial population is indexed as Generation 0.

I measured flower size as the maximal width of the corolla (see diagram in Fishman et al 2002). For most analyses, width was square root transformed to minimize scale effects, a critical concern given that the study objective is to compare variance estimates. From a genetic perspective, the criteria for the most appropriate scale are (1) a linear decline in the mean phenotype with inbreeding coefficient (Kempthorne 1957; Wright 1951), and (2) homoscedasticity across and normality within families (Wright 1952). Regarding (1), a previous study estimated the mean flower size of Iron Mountain plants with inbreeding coefficients of 0, 0.5, 0.75, and 1 (Kelly 2005). There was no significant deviation from linearity for either corolla width or square-root corolla width, although the estimated deviation was slightly smaller for the latter. Regarding (2), the square-root minimized heteroscedasticity across families within a breeding design involving inbred lines and their F1 hybrids (Kelly and Arathi 2003). In addition, Generations 0-6 of the present study are essentially a large collection of full-sib families, each with an estimated intra-family variance and substantial variation in means. Applying the generalized Box-Cox procedure to the 1640 families with four or more members yields an optimal power series transformation. The estimated exponent is 0.53, which is very close to the value for square-root (0.5). This procedure, detailed in the Appendix, finds the optimal compromise between normality within families and homoscedasticity among families (see Sokal and Rohlf 2000, p. 419).

Experimental design

In each of the first six generations, seeds were grown to flowering plants following our standard greenhouse procedures (Arathi and Kelly 2004). One thousand individuals were measured for corolla width and number of days to flower in both the Low and High populations. Between two and three hundred individuals were measured for the Control population. After flowering was complete, usually within 60 days of seeding pots, individuals were ranked according to

corolla width within populations. The highest 200 were selected in the High population, the lowest 200 in the Low population, and 200 random individuals in the Control. Selected individuals were randomly paired within populations and crossed to generate 100 outbred families in each population.

The next generation was produced by growing a nearly equal number of plants per family, 10-12 in Low and High, 2-3 in the Control. Counts for some families were slightly greater than the target number (10 for Low and High, 2 for Control) to compensate for occasional failures in germination or maturation within other families. However, the realized variance among families when censused as adults was very small. All else equal, this should nearly double the 'effective population size' relative to the census size (200), thus reducing the effect of genetic drift (Hartl and Clark 1989, p. 89). Of course, selection generates differences among families when censused at the reproductive adult stage. It also directly effects the population genealogy causing correlations of family sizes across generations (Caballero and Santiago 1995; Robertson 1961). Both factors will amplify genetic drift. Even so, the importance of drift relative to selection is much reduced with this design relative to most selection experiments. An inevitable but still unfortunate consequence of maintaining these large population sizes is that I could not replicate treatments.

Selection was not imposed on the adults of Generation six and approximately 900 plants from each population were grown to maturity and randomly paired. Each pair produced a single outbred plant for Generation seven. This plant was grown to maturity and randomly assigned as either a sire (one third) or a dam (two thirds). There were 100 sires within the Low population, 141 sires in the Control population, and 121 sires in the High population. Two dams were randomly allotted to each sire within each population and crossed to produce two (related) outbred families. Each sire and dam was also self-fertilized. The two generations between the cessation of selection and measurement of progeny allowed recombination among QTL. This reduces linkage disequilibria generated by both selection and drift (see Avery and Hill 1977), and as a consequence, the genetic variance estimates from Generation eight are more directly informative about QTL allele frequencies.

Generation eight consisted of the progeny from three distinct half-sib/full-sib breeding designs, one for each population. All Generation eight plants were measured for flower size and days to flower, and most were also measured for above-ground dry biomass at flowering, pollen number per flower, and fraction of pollen viable. The last was estimated using the Pollen Size Index (Kelly et al. 2002). The full pedigree of measured plants from High, Low, and Control contains over 21,000 plants.

The simulation program

The stochastic simulation program is tailored to the experiment described above. I assume that flower size is influenced by L loci, uniformly distributed over a genetic map of 2000cM (see Fishman et al. 2002). Each locus is polymorphic for two alleles with the rare allele at frequency q in the natural (ancestral) population. At locus i , the rare allele has effect a_i on the phenotype. I assume that effects are additive within and among loci (no dominance or epistasis) and that loci are in linkage equilibria in the natural population. While neither assumption is true, the generation of random mating between field and Generation 0 seed should at least reduce any disequilibria among loci.

For each particular 'genetic model', the a_i are drawn from a normal distribution with a specified mean (m_a) and standard deviation (s_a). The a_i values are constant throughout a simulation run. The first step of a simulation run is to sample 1000 diploid genotypes and determine the genotypic value of each individual. Once the genotypic value is established, a normally distributed environmental error (with mean zero) is added to the genotypic value of each

individual to produce the phenotype. The environmental variance is set such that the heritability of the base population is 0.33 (see Results section). This founding population (analogous to Generation 0) is then subjected to six generations of selection, with the top (or bottom) 20% selected in each generation. Selected adults are randomly paired and each couple produces 10 progeny for the next generation. Selection was imposed in both directions in distinct runs.

The six generations of selection are followed by two generations of random mating without selection (analogous to Generations six and seven of the experimental design). I then measure the mean and variance of genotypic values among progeny. The genetic variance is affected by any residual linkage disequilibria, as would occur in the experiment. There are two levels of replication in these simulations: For each genetic model, specified a distinct array of a_i , selection was replicated 50 times in both directions. Variation among replicates at this level is due to sampling of founding Generation 0 and genetic drift in subsequent generations. The second level is across genetic models: For each parameter set, specified by q , L , m_a , and s_a , I sample 100 distinct a_i arrays.

Results

Simulation

Figure 2 describes simulation results for a particular genetic model in which each individual in the ancestral population carries an average of two rare alleles ($L = 100$, $q = 0.01$). Genetic effects are symmetric (equally likely to increase or decrease the trait) in this case. The response to selection over the first six generations is predicted to be fairly consistent across replicates with an apparent, although fairly modest, acceleration in later generations (Fig. 2A). The average genetic variance is greatly elevated in Generation eight, about six times that of the ancestral population (Fig. 2B). Perhaps more importantly, V_g reliably increases across replicates, i.e. the standard deviation is small relative to the mean.

Results for a range of parameter values are summarized in Table 1. A large number of loci ($L = 50$, 100, or 200) must contribute to variation under a rare-alleles model to generate substantial variation with a normal distribution of genotypic values. As expected, the increase in V_g is most pronounced with fewer loci. However, substantial changes are observed even with $L = 200$ implying that the infinitesimal limit, where V_g is constant, is approached very slowly with extreme allele frequencies. Also as expected, the increase in V_g is greatest when initial allele frequencies are most extreme (q is lowest). When rare alleles mainly reduce the trait value (lower part of Table 1), V_g increases dramatically in the Low population. However, V_g declines in the High population and response rapidly plateaus. Across parameter sets, the standard deviation among replicates for V_g in Generation eight is typically small relative to the mean.

Experiment

The divergence between High and Low populations in (transformed) mean corolla width was remarkably linear over the six generations of selection (Fig. 3A) with a cumulative difference of approximately four and a half phenotypic standard deviations of the Control population. The response was slightly asymmetric with a greater difference between High and Control than Low and Control (Fig. 3B). Each of the additional traits measured in Generation eight differed significantly among populations indicating correlated responses to selection (Table 2; ANOVAs with population as the factor, $P < 0.001$ in each case). Selection for smaller flowers also reduced development time, size at harvest, the number of pollen grains per flower, and the fraction of grains that are viable. The High population evolved delayed development and thus flowered at a larger size. Their large flowers produce more pollen grains, but a reduced fraction is viable (relative to the Control).

Two estimates for h^2 , the narrow sense heritability of transformed corolla width, can be obtained from the data of Generations 0-6. Figure 4 summarizes estimates for the linear regression of mean offspring phenotype onto the mid-parent value within the Control population. If these seven regression estimates are averaged using Generalized Least Squares to account for their differing standard errors, we obtain an overall h^2 estimate of 0.34 (SE of 0.03). The 'realized heritability' is estimated by noting that the predicted cumulative selection differential with bi-directional 20% selection is predicted to be 16.8 phenotypic SD. Averaging the SD of the Control populations in generations 0-6, this yields $h^2 = 0.28$. It is more difficult to place a SE on this estimate because the variance in response is determined more by the accumulated effects of genetic drift than the accuracy of the estimated final response (Falconer and Mackay 1996, p. 210). Both estimates are similar to the Control population estimate of $h^2 = 0.37$ obtained from the breeding design of Generations 7-8 (from Model III described next).

Maximum likelihood estimates for the variance components associated with flower size were obtained by applying a standard mixed-model to the plants grown in Generations seven and eight. There are four fixed effects in the model, the trait mean for each population and the generation effect (the difference in means between generations due to an average environmental effect). Variance components are estimated from the three contrasts among relatives, i.e. parent-offspring, full-sibs, and half-sibs. The parent-offspring covariance is $V_a/2$, the full-sib covariance is $V_a/2 + V_d/4$, and half-sib covariance is $V_a/4$. With these comparisons, maternal effects (if they exist) would likely to be absorbed into V_d estimates. For this reason, and because the contrasts provide limited information about this component, V_d estimates should thus be treated with caution. The maximum likelihood was found using the Powell algorithm (Press et al. 1992, ch. 10). Convergence was checked manually by starting replicated searches from different points in the parameter space. The programs, simple modifications of those developed previously (e.g. Kelly 2003; Kelly and Arathi 2003), are available upon request.

The number of variance components depends on the variant of the model. The simplest case (Model I) specifies a single value for V_E , V_a , and V_d across all three populations. The most elaborate case (Model IV) estimates distinct V_E , V_a , and V_d within each population. I consider two intermediate scenarios, Models II and III. Model II assumes that the genetic variance remained constant (a single V_a and V_d is estimated for all three populations), but allows V_E to differ among populations. Model III allows the genetic variances to evolve (population specific estimates for V_a and V_d), but not the environmental variance. The maximum likelihood solutions for both untransformed corolla width and square-root transformed corolla width under each model are given in Table 3. An analysis of the raw data is included here because the transformation reduces variability at the high end of the scale relative to that at the low end. Because V_g apparently increased in the High population, the effect is more pronounced when considering untransformed measurements.

I compare maximum likelihood values from each model using both the Akaike Information Criterion (AIC) and by likelihood ratios. The AIC is an Information Theoretic approach devised to find the optimal compromise between model simplicity and model fit. The AIC equals twice the difference between the number of parameters and the loglikelihood (Burnham and Anderson 2002, p. 61). The 'selected model' has the lowest AIC. Applying this criterion to results for both transformed and untransformed trait values, Model III is best followed by Models IV and II (Table 3). Model I is a distant fourth, indicating that variances have evolved due to selection. The Model III estimates are given in Fig. 5 where it is apparent that the primary difference is an elevation of V_g in the High population.

Because there is a nested structure to the models, classical hypothesis tests (likelihood ratios) can be applied to particular pairs of models. Model I is a special case of Models II and III, while II and III are subsets of Model IV. Twice the difference between log-likelihoods (LR) is compared to the chi-square distribution with degrees of freedom equal to the difference in number of parameters between models. For untransformed corolla width, Model I is rejected in a comparison to Model II (LR = 298.68, df = 2, $P < 0.001$) or to Model III (LR = 325.40, df = 4, $P < 0.001$). Model II is rejected in a comparison to Model IV (LR = 27.26, df = 4, $P < 0.001$), but Model III is not (LR = 0.54, df = 2, $P > 0.5$). For transformed corolla width, Model I is rejected in a comparison to Model II (LR = 51.76, df = 2, $P < 0.001$) or to Model III (LR = 61.64, df = 4, $P < 0.001$). Model II is rejected in a comparison to Model IV (LR = 10.06, df = 4, $P = 0.04$), but Model III is not (LR = 0.18, df = 2, $P > 0.5$). Models II and III cannot be compared in this way because neither is nested within the other. On the whole however, the series of Likelihood Ratio tests is thus fully consistent with the inference from AIC values.

Finally, selection also altered the allometry of plant growth. Corolla width and plant size (at the time of harvest) are positively correlated (the joint distribution in Generation eight is depicted in Fig. 6). Comparing regressions, a test for slope heterogeneity is marginally non-significant ($F_{2, 1081} = 2.58$, $P = 0.076$). I thus applied an ANCOVA with population as the factor and Log-transformed Biomass as the covariate predicting corolla width. Both factor and covariate are highly significant ($P < 0.001$), but interestingly, the difference between intercept estimates for Low and Control populations is minimal. Thus, the large difference in mean corolla width is due mainly to the fact that Low plants flowered at a smaller size than Control plants (Table 2). In contrast, High population plants flower after accumulating substantially greater vegetative tissue, but they are also produce larger flowers for a given biomass than Low or Control plants.

Discussion

The ' ΔV_g prediction' is a simple extrapolation from Fig. 1, variants of which reside in most introductory genetics texts. Two key features of an evolutionary prediction are robustness and exclusivity. Robustness implies that deviations from model assumptions do not invalidate the prediction. Exclusivity concerns the extent to which the prediction is unique to the hypothesized model. If the predicted outcome obtains, does it conclusively implicate the hypothesized cause? Here, I argue that an appropriate experimental design can provide a robust test of the ΔV_g prediction. However, because the prediction has weak exclusivity, a positive result does not prove that variation is caused by rare alleles.

Theory

The simulation study (Fig. 2, Table 1) was designed to determine whether an increase in V_g would *reliably* occur in this experiment *if* variation in flower size is caused by rare alleles. Reliability is essential because the experiment involved a single High, Low, and Control population. Genetic drift is typically a dominant factor in selection experiments (Hill 1971 1980) and can obscure the predicted increase in V_g by altering allele frequencies and generating linkage disequilibria among QTL. The adequacy of the experiment depends both on aspects of the design (founding of experimental populations, population sizes, protocol for mating and family establishment, duration of selection, the controlled crosses for V_g estimation, etc) and on features of the organism (the number of loci, the frequency of alleles at these loci, and their distribution over the genetic map). The latter are unknown, components of the hypothesis under consideration. However, the rare alleles model encompasses a range of possibilities for q (the frequency of rare alleles) and L (the number of QTL) and this compels consideration of different cases (Table 1). Generally, the simulations support the adequacy of the experiment for testing the ΔV_g prediction.

The positive results from the simulation study provide an informative contrast to those recently obtained by Zhang and Hill (2005). These authors consider response to artificial selection when variation in the ancestral population is maintained by mutation-selection balance according to their 'joint-effect model' (Zhang and Hill 2002; Zhang et al. 2002; Zhang et al. 2004). In their simulations, V_g typically declined over the first 5-6 generations of selection (compare Fig. 2 with Figure 2c of Zhang and Hill 2005). There are a number of explanations for the differing results. First, Zhang and Hill (2005) designed their simulation to the 'typical' selection experiment, based primarily on previous studies of *Drosophila*. Experimental populations were founded from laboratory stock population and not from a large random sampling of the natural population. This period of captivity tends to eliminate the rare alleles at many loci that are polymorphic in the ancestral population. A relatively small number of individuals (40) were selected in each generation and recombination was limited by the small map length of *Drosophila*. Finally, the joint-effect model actually does permit intermediate frequency polymorphisms (see Fig. 1 of Zhang and Hill 2005). Thus, even though the great majority of loci harbour rare allele(s), a limited number of intermediate frequency polymorphisms can greatly diminish the ΔV_g prediction. In the present study, the rare-alleles model is directly enforced by specifying a fixed allele frequency (q) in the ancestral population.

The contrast between the two simulation studies is directly relevant to the interpretation of previous selection experiments. Johnson and Barton (2005) reviewed artificial selection studies and failed to find a single example of an accelerating response to selection on an outbred base population (but also see Heath et al. 1995; Sorensen and Hill 1982). Response is expected to accelerate if the genetic variance increases (Curtis and Ming 1997). However, because most selection experiments were not designed to detect changes in V_g and resemble the design simulated by Zhang and Hill (2005), the extent to which their results reject the rare-alleles model is subject to question.

In contrast to the great majority of studies, the experiment described here did show an increase in V_g within the high-selected population. The response was asymmetric with V_g apparently declining in the Low population (Fig. 5, Table 3). While these results are consistent with one variant of the rare-alleles model (see next section), they do not conclusively demonstrate that variation is caused by rare alleles. An increase in V_g , particularly if it occurs in only one direction of selection, can be caused by changes at intermediate frequency polymorphisms. Over a large range of intermediate allele frequencies, the genetic variance contributed by a QTL may increase (an allele that increases from an initial frequency of 10-40% under selection, Fig. 1). Moreover, epistasis among QTL with intermediate frequencies can produce situations where V_g is still increasing even as allele frequencies substantially surpass 0.5.

A bi-directional increase in V_g is a more exclusive prediction. However it applies to only a subset of rare-alleles models, those in which rare alleles increase trait values at some QTL while decreasing them at other QTL (e.g. Keightley and Hill 1988; Kondrashov and Turelli 1992; Zhang and Hill 2005). When testing a robust prediction with weak exclusivity, interpretation is most straightforward if the predicted outcome (an increase in V_g) does not occur. Because V_g did increase in the High population, it is necessary to consider these results in the context of other genetic data from this population. After such a consideration, rare alleles seem at best only a partial explanation for the observed response.

Corolla size variation within the Iron Mountain population of *Mimulus guttatus*

Estimates from Generations 7-8 are consistent with a model in which rare alleles mainly *increase* flower size (Fig. 5). These alleles would also have to increase pollen per flower, days to flower, and the above-ground biomass of plants at the time of flowering (Table 2). This hypothesis further requires that rare alleles are partially dominant (on average) in their effects on corolla width, biomass, pollen viability, and pollen number per flower, but partially

recessive in their effects on days to flower. This is because inbreeding reduces the first set of traits but increases development time (Willis 1993a; unpublished results).

In previous studies, we showed that selection for smaller flowers reduces the inbreeding load of corolla width (Kelly and Willis 2001) and that inbreeding increases the genetic variance of this trait only marginally (Kelly 2003; Kelly and Arathi 2003). These results reject a model in which variation is due to rare, partially recessive alleles. However, they are actually consistent with partially dominant rare alleles increasing flower size. Of course, if such alleles had initial frequencies of say 10-30% within the High population, that could also explain the patterns depicted in Figs. 3, 5, and 6. Allele frequencies within this range are more consistent with balancing selection than the various mutation-selection or migration-selection balance models that yield a rare-alleles architecture.

Another difficulty with the 'rare-dominants model' is that we usually expect rare alleles will be at least partially recessive, either for physiological reasons (Crow 1993; Kacser and Burns 1981; Lynch et al. 1999) or because recessivity allows their persistence (Haldane 1924, 1927). It also seems unlikely that alleles reducing pollen viability and the number of grains per flower would be close to fixation. All else equal, both traits should be under strong positive selection. Relevant here is the fact that plants from the High population produce fewer flowers, while those from the Low population produce more, than Control plants over the first eight weeks of the lifespan (unpublished results). Considering pollen viability, both High and Low populations had lower average viability than the Controls (Table 2). If mutations that reduce pollen viability are bi-directional in their effects on flower size (some increasing flower size and others decreasing it), the increase in frequency of different mutations within the high and low selection populations may explain this peculiar response. Of course, an increase of rare alleles within the Low population should have inflated V_g which is not what happened (Fig. 5).

Integration of marker-based data

The complicated, conditional arguments of the preceding section illustrate the difficulties associated with using phenotype data (resemblance of relatives, response to selection or inbreeding) to infer details of inheritance (see discussion in Hill and Caballero 1992). In light of these difficulties, one might be inclined to adopt a more direct approach to estimating QTL allele frequencies. Genomic technologies have now been applied in a range of studies on *Mimulus* (Bradshaw et al. 1998; Fishman et al. 2002; Lin and Ritland 1997). If the DNA level changes that define QTL can be identified, allele frequencies can be estimated using simple population surveys. Unfortunately, the isolation of 'Quantitative Trait Nucleotides' is not a trivial endeavour, particularly for loci with small to moderate effects (Mackay 2004).

In several key regards, the rare-alleles model actually represents a worst-case scenario for current QTL mapping procedures. Artificial selection is typically used to concentrate intra-population variation into divergent lines that are subsequently used to produce mapping populations (Long et al. 1995; Valenzuela et al. 2004). Under a rarealleles model, the 'capture' of particular QTL within selection lines will be idiosyncratic. Different QTL will likely fix in different selection lines, increasing the variability of responses among lines (Frankham 1980; James 1971) and reducing the repeatability of mapping results from any specific line. Individual mapping populations provide limited information about the background population. Finally, rare-allele models typically require a great many loci to generate realistic heritabilities. As a consequence, the amount of work required to thoroughly characterize variation QTL by QTL becomes prohibitive.

QTL mapping may tell us the most about the contribution of rare alleles by exclusion. If QTL with intermediate frequencies are accurately identified, their contribution to variation can be

measured. This will establish an upper bound on the contribution of rare alleles. Of course, the unexplained variance will also likely include intermediate frequency polymorphisms for QTL with small effects. Generally speaking, we confront at least two major challenges in attempting to more accurately describe the allele frequency spectrum underlying quantitative trait variation. One is to develop of methods to quantitatively integrate of results from biometric and QTL mapping analyses. Second, we require novel experimental designs and mapping procedures that are more directly informative about population allele frequencies. The development of such methods should allow QTL mapping to more directly address evolutionary questions regarding the maintenance of genetic variability.

Acknowledgments

I thank L. Holeski, Y. W. Lee, and J. H. Willis for helpful comments on an early draft of this paper. Revisions were made in response to 2 thoughtful reviews. I gratefully acknowledge the support of NIH grant R01 GM60792-01A1 and NSF grant DEB0543052.

Appendix

The plants grown in Generations 0-6 can be clustered into full sib families. The dataset analyzed for transformation included all such families with at least four members from all three populations.

The Box-Cox procedure, as generalized to induce homogeneity of variances (Sokal and Rohlf 2000, p. 419), was applied to these 1640 families (14904 plants). The likelihood function:

$$L' = -\frac{v}{2} \ln(s_T^2) + (U - 1) \frac{v}{n} \sum_{j=1}^n \ln(Y_j) - \frac{1}{2} X^2$$

where Y_j are the individual flower size measurements, n is the number of measurements, v is the number of degrees of freedom, and s_T^2 is the Error Mean Square. The latter two quantities are equivalent to those derived from a one-way ANOVA with family as the factor and flower size as the response variable. The quantity X^2 is Bartlett's statistic:

$$X^2 = v \ln(s^2) + \sum_{i=1}^a (n_i - 1) \ln(s_i^2) \text{ and } s^2 = \frac{1}{v} \sum_{i=1}^a (n_i - 1) s_i^2$$

Here, n_i is the size of family i , s_i^2 is the variance within that family, and a is the number of families. Iteratively, I find that $U = 0.53$ maximizes L' . This transformation is essentially equivalent to the square-root transformation in regard to its effects on the variance.

Inspection of intra-family variance estimates indicates that this transformation effectively eliminates the positive relationship between family mean and within family variance that exists on the original scale of measurement.

References

- Arathi HS, Kelly JK. Corolla morphology facilitates both autogamy and bumblebee pollination in *Mimulus guttatus*. *Int J Plant Science* 2004;165:039–1045.
- Avery PJ, Hill WG. Variability in genetic parameters among small populations. *Genetical research* 1977;29:93–213.

- Awadalla P, Ritland K. Microsatellite variation and evolution in the *mimulus guttatus* species complex with contrasting mating systems. *Mol Biol Evol* 1997;14:1023–1034. [PubMed: 9335142]
- Barton NH. Pleiotropic models of quantitative trait variation. *Genetics* 1990;124:773–782. [PubMed: 2311921]
- Bradshaw HD Jr, Otto KG, Frewen BE, McKay JK, Schemske DW. Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics* 1998;149:367–382. [PubMed: 9584110]
- Bulmer MG. The effect of selection on genetic variability: a simulation study. *Genet Res* 1976;28:101–117. [PubMed: 1010337]
- Burnham, KP.; Anderson, DR. Model selection and multimodel inference: a practical information-theoretic approach. Springer-Verlag; New York: 2002.
- Caballero A, Keightley PD. A pleiotropic nonadditive model of variation in quantitative traits. *Genetics* 1994;138:883–900. [PubMed: 7851783]
- Caballero A, Santiago E. Response to selection from new mutations and effective size of partially inbred populations I Theoretical results. *Genet Res* 1995;66:213–225.
- Carr DE, Fenster CB. Levels of genetic variation and covariation for *mimulus* (Scrophulariaceae) floral traits. *Heredity* 1994;72:606–618.
- Charlesworth, B.; Hughes, KA. The maintenance of genetic variation in life history traits. In: Singh, RS.; Krimbas, CB., editors. *Evolutionary genetics from molecules to morphology*. Cambridge University Press; Cambridge, U.K: 2000. p. 369-392.
- Crow JF. Mutation, mean fitness, and genetic load. *Oxford Surv Evol Biol* 1993;9:3–42.
- Curtsinger JW, Ming R. Non-linear selection response in *Drosophila*: a strategy for testing the rare-alleles model of quantitative genetic variability. *Genetica* 1997;99:59–66. [PubMed: 9226437]
- Falconer, DS.; Mackay, TFC. *Introduction to quantitative genetics*. Prentice Hall; London: 1996.
- Fenster CB, Ritland K. Quantitative genetics of mating system divergence in the yellow monkeyflower species complex. *Heredity* 1994;73:422–435.
- Fishman L, Kelly AJ, Willis JH. Minor quantitative trait loci underlie floral traits associated with mating system divergence in *mimulus*. *Evolution* 2002;56:2138–2155. [PubMed: 12487345]
- Frankham, R. The founder effect and response to artificial selection in *Drosophila*. In: Robertson, A., editor. *Selection experiments in laboratory and domestic animals*. commonwealth agricultural bureaux; Slough, UK: 1980. p. 87-90.
- Haldane JBS. A mathematical theory of natural and artificial selection, Part I. *Trans Camb Philos Soc* 1924;23:19–41.
- Haldane JBS. A mathematical theory of natural and artificial selection. Part V. Selection and Mutation. *Proc Camb Phil Soc* 1927;23:838–844.
- Hartl, DL.; Clark, AG. *Principles of population genetics*. Sinauer associates; Sunderland, Massachusetts: 1989.
- Heath SC, Bulfield G, Thompson R, Keightley PD. Rates of change of genetic-parameters of body-weight in selected mouse lines. *Genet Res* 1995;66:19–25. [PubMed: 8575666]
- Hill WG. Design and efficiency of selection experiments for estimating genetic parameters. *Biometrics* 1971;28:293–311. [PubMed: 5130167]
- Hill, WG. Design of quantitative genetic selection experiments. In: Robertson, A., editor. *Selection experiments in laboratory and domestic animals*. Commonwealth Agricultural Bureaux; Slough, UK: 1980. p. 1-13.
- Hill WG, Caballero A. Artificial selection experiments. *Ann Rev Ecol Syst* 1992;23:287–310.
- Hill WG, Robertson A. The effect of linkage on limits to artificial selection. *Gen Res* 1966;8:269–294.
- Houle D, Morikawa B, Lynch M. Comparing mutational variabilities. *Genetics* 1996;143:1467–1483. [PubMed: 8807316]
- James JW. The founder effect and response to artificial selection. *Genet Res* 1971;16:241–250. [PubMed: 5512250]
- Kacser H, Burns JA. The molecular basis of dominance. *Genetics* 1981;97:639–666. [PubMed: 7297851]
- Keightley PD, Hill WG. Quantitative genetic-variability maintained by mutation-stabilizing selection balance in finite populations. *Genet Res* 1988;52:33–43. [PubMed: 3181758]

- Keightley PD, Hill WG. Variation maintained in quantitative traits with mutation selection balance-pleiotropic side-effects on fitness traits. *Proc Roy Soc Lond B Biol Sci* 1990;242:95–100.
- Kelly JK. An experimental method for evaluating the contribution of deleterious mutations to quantitative trait variation. *Genet Res* 1999;73:263–273. [PubMed: 10425922]
- Kelly JK. Deleterious mutations and the genetic variance of male fitness components in *mimulus guttatus*. *Genetics* 2003;164:1071–1085. [PubMed: 12871916]
- Kelly JK. Epistasis in Monkeyflowers. *Genetics* 2005;171:1917–1931. [PubMed: 15944350]
- Kelly JK, Arathi HS. Inbreeding and the genetic variance of floral traits in *mimulus guttatus*. *Heredity* 2003;90:77–83. [PubMed: 12522429]
- Kelly JK, Willis JH. Deleterious mutations and genetic variation for flower size in *mimulus guttatus*. *Evolution* 2001;55:937–942. [PubMed: 11430654]
- Kelly JK, Rasch A, Kalisz S. A method to estimate pollen viability from pollen size variation. *Am J Bot* 2002;89:1021–1023.
- Kemphorne, O. An introduction to genetic statistics. Wiley; New York: 1957.
- Kimura M. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc Natl Acad Sci U S A* 1969;41:144–150. [PubMed: 16589632]
- Kondrashov AS, Turelli M. Deleterious mutations, apparent stabilizing selection and the maintenance of quantitative variation. *Genetics* 1992;132:603–618. [PubMed: 1427047]
- Lin JZ, Ritland K. Quantitative trait loci differentiating the outbreeding *mimulus guttatus* from the inbreeding *M-Platycalyx*. *Genetics* 1997;146:1115–1121. [PubMed: 9215912]
- Long AD, Mullaney SL, Reid LA, Fry JD, Langley CH, Mackay TFC. High-resolution mapping of genetic-factors affecting abdominal bristle number in *Drosophila-melanogaster*. *Genetics* 1995;139:1273–1291. [PubMed: 7768438]
- Lynch M, Blanchard J, Houle D, Kibota T, Schultz S, Vassilieva L, Willis JH. Spontaneous deleterious mutation. *Evolution* 1999;53:645–663.
- Mackay, TFC. Genetic dissection of quantitative traits. In: Singh, RS.; Uyenoyama, MK., editors. *The evolution of population biology*. Cambridge University Press; Cambridge: 2004. p. 51-73.
- Macnair MR, Cumbes QJ. The genetic architecture of interspecific variation in *mimulus*. *Genetics* 1989;122:211–222. [PubMed: 17246497]
- Press, WH.; Teukolsky, SA.; Vetterling, WT.; Flannery, BP. *Numerical recipes in C*. Cambridge University Press; Cambridge: 1992.
- Ritland K. Genetic differentiation, diversity, and inbreeding in the Mountain Monkeyflower (*Mimulus-Caespitatus*) of the Washington cascades. *Can J Bot-Revue Canadienne De Bota-nique* 1989;67:2017–2024.
- Ritland K, Ganders FR. Covariation of selfing rates with parental gene fixation indexes within populations of *Mimulus guttatus*. *Evolution* 1987;41:760–771.
- Robertson A. Inbreeding in artificial selection programmes. *Gen Res* 1961;2:189–194.
- Robertson AW, Diaz A, MacNair MR. The quantitative genetics of floral characters in *mimulus guttatus*. *Heredity* 1994;72:300–311.
- Sokal, RR.; Rohlf, FJ. *Biometry*. Freeman and company; New York: 2000.
- Sorensen DA, Hill WG. Effect of short-term directional selection on genetic-variability-experiments with *Drosophila Melanogaster*. *Heredity* 1982;48:27–33. [PubMed: 6804417]
- Sweigart A, Karoly K, Jones A, Willis JH. The distribution of individual inbreeding coefficients and pairwise relatedness in a population of *mimulus guttatus*. *Heredity* 1999;83:625–632. [PubMed: 10620036]
- Turelli M. Heritable genetic variation via mutation -selection balance: Lerch's Zeta meets the Abdominal Bristle. *Theor Popul Biol* 1984;25:138–193. [PubMed: 6729751]
- Turelli, M. Quantitative genetics, Ecology, and evolution. In: Weir, BS.; Eisen, EJ.; Goodman, MM.; Namkoong, G., editors. *Second international conference on quantitative genetics*. Sinauer Associates; 1987.
- Valenzuela RK, Forbes SN, Keim P, Service PM. Quantitative trait loci affecting life span in replicated populations of *Drosophila melanogaster*. *II. Response to Selection Genetics* 2004;168:313–324.

- Willis JH. Effects of Different Levels of Inbreeding on Fitness Components in *Mimulus Guttatus*. *Evolution* 1993a;47:864–876.
- Willis JH. Partial self fertilization and inbreeding depression in two populations of *Mimulus guttatus*. *Heredity* 1993b;71:145–154.
- Willis JH. Measures of phenotypic selection are biased by partial inbreeding. *Evolution* 1996;50:1501–1511.
- Willis JH. Inbreeding load, average dominance, and the mutation rate for mildly deleterious alleles in *Mimulus guttatus*. *Genetics* 1999a;153:1885–1898. [PubMed: 10581293]
- Willis JH. The role of genes of large effect on inbreeding depression in *Mimulus guttatus*. *Evolution* 1999b;53:1678–1691.
- Wright S. The genetical structure of populations. *Ann Eugen* 1951;15:323–354.
- Wright, S. The genetics of quantitative variability. In: Reeve, ECR.; Waddington, CH., editors. *Quantitative inheritance*. Her majesty's stationary office; London, UK: 1952. p. 5-41.
- Zhang X-S, Hill WG. Joint effects of pleiotropic selection and stabilizing selection on the maintenance of quantitative genetic variation at mutation-selection balance. *Genetics* 2002;162:459–471. [PubMed: 12242254]
- Zhang X-S, Hill WG. Predictions of patterns of response to artificial selection in lines derived from natural populations. *Genetics* 2005;169:411–425. [PubMed: 15677752]
- Zhang X-S, Wang J, Hill WG. Pleiotropic model of maintenance of quantitative genetic variation at mutation-selection balance. *Genetics* 2002;161:419–433. [PubMed: 12019255]
- Zhang X-S, Wang J, Hill WG. Influence of dominance, leptokurtosis and pleiotropy of deleterious mutations on quantitative genetic variation at mutation-selection balance. *Genetics* 2004;166:597–610. [PubMed: 15020447]

Fig. 1.

The genetic variance (V_g) is predicted as a function of allele frequency. The average phenotypic values of genotypes A_0A_0 , A_0A_1 , and A_1A_1 are $-a$, d , and a , respectively, and q is the frequency of A_0 (as in Falconer & Mackay, 1996, chps 7-8). For the solid line, $a = 1$ and $d = 0.8$. For the broken line, $a = 1$ and $d = 0.0$.

Fig. 2.

A summary of replicate simulations from a genetic model in which $L = 100$ and $q_i = 0.01$.

(A) The divergence between the means of High and Low selected populations is given as a function of generation for five replicates. Here, divergence is measured in units of phenotypic standard deviations of the base population. (B) The distribution of V_g for the Low population in Generation eight across 500 replicate simulations. Trait variation is standardized so that the base population genetic variance is 1

Fig. 3.

The response to selection is depicted as **(A)** the divergence in mean corolla width between High and Low populations over the first six generations, and **(B)** the means of plants from all three populations in Generation eight. The units are $\sqrt{mm}$ and the error bars (often obscured by the points) define 95% confidence intervals

Fig. 4. The linear regression coefficient of mean offspring corolla size onto the mid-parent value is given for the Control population in generations 0-7. The error bars denote $\pm$ one standard error on the regression estimate

Fig. 5.

Estimates from Model III are given for the genetic variance, V_g , of square-root transformed corolla width (open bars) and untransformed corolla width (solid bars) for each population in Generation eight. Error bars are $\pm$ one SE. The SE is calculated from the sampling variances of, and covariance between, V_a and V_d from the asymptotic dispersion matrix

Fig. 6. The joint distribution of corolla width at above-ground dry mass at the time of flowering is depicted for each population. The units are $\sqrt{mm}$ and Log(mg)

Table 1

For the symmetric effects cases, $m_a = 0$ and $s_a = 1/L$, while for the asymmetric cases, $m_a = -2/L$ and $s_a = 1/L$

q	L	Divergence	Low $V_g/V_g(0)$		High $V_g/V_g(0)$	
			Mean	SD	Mean	SD
Symmetric effects						
0.002	50	24.8	29.80	5.00	29.80	5.00
0.002	100	20.5	24.10	3.30	24.10	3.30
0.002	200	16.2	17.13	2.07	17.13	2.07
0.005	50	16.9	12.39	1.70	12.39	1.70
0.005	100	13.8	10.61	1.21	10.61	1.21
0.005	200	11.1	7.73	0.80	7.73	0.80
0.01	50	12.3	6.65	0.78	6.65	0.78
0.01	100	10.3	5.80	0.58	5.80	0.58
0.01	200	8.4	4.35	0.40	4.35	0.40
0.02	50	9.2	3.73	0.38	3.73	0.38
0.02	100	7.9	3.29	0.28	3.29	0.28
0.02	200	6.7	2.56	0.22	2.56	0.22
Asymmetric effects						
0.002	50	13.96	45.89	3.92	0.03	0.05
0.002	100	11.06	29.24	2.31	0.02	0.04
0.002	200	8.83	17.72	1.46	0.02	0.03
0.005	50	9.91	20.61	1.54	0.02	0.03
0.005	100	8.09	13.58	0.99	0.01	0.02
0.005	200	6.76	8.53	0.66	0.02	0.02
0.01	50	7.81	11.49	0.79	0.01	0.02
0.01	100	6.62	7.80	0.54	0.02	0.02
0.01	200	5.81	5.14	0.38	0.05	0.02
0.02	50	6.37	6.52	0.42	0.02	0.02
0.02	100	5.67	4.66	0.31	0.05	0.02
0.02	200	5.25	3.31	0.23	0.12	0.03

Divergence is the average difference between Low and High population means in units of phenotypic SD of the ancestral population. The fourth and fifth columns give the mean and SD of the Low population genetic variance in Generation eight (averaged across genetic models for each parameter set). V_g is standardized relative the genetic variance of the ancestral population. The sixth and seventh columns give the corresponding values for the High population

The population specific means and standard deviations are given for five traits measured in Generation eight. Above-ground biomass is reported in mg, while the other measurements are counts or proportions

Kelly

Table 2

Population	Low		Control		High	
	Mean	SD	Mean	SD	Mean	SD
Biomass	28.0	13.1	44.8	28.1	81.9	62.5
Day of flowering	30.4	3.2	31.0	3.9	34.4	4.1
Total pollen	6396	2900	8440	3051	11807	4123
Pollen viability	0.540	0.197	0.625	0.209	0.553	0.209
Est. viable pollen	3586	2222	5382	2722	6656	3570

Table 3

The number of estimated parameters, log-likelihood, AIC, and variance component estimates are given for each model when fit to data from Generations 7-8

Model	Par.	Log-likelihood	AIC	Variance component Estimates
<i>Corolla width</i>				
I	7	-1661.13	317.41	$V_a = 0.40, V_d = 0.15, V_E = 0.45$
II	9	-1511.79	22.73	$V_a = 0.38, V_d = 0.06$ L: $V_E = 0.33, C: V_E = 0.56, H: V_E = 1.60$
III	11	-1498.43	0	$V_E = 0.44$ L: $V_a = 0.25, V_d = 0.01$ C: $V_a = 0.38, V_d = 0.18$ H: $V_a = 0.94, V_d = 0.85$
IV	13	-1498.16	3.45	L: $V_a = 0.25, V_d = 0.01, V_E = 0.44$ C: $V_a = 0.38, V_d = 0.21, V_E = 0.41$ H: $V_a = 0.92, V_d = 0.49, V_E = 0.79$
<i>Square-root(Corolla width)</i>				
I	7	-1506.56	53.63	$V_a = 0.38, V_d = 0.14, V_E = 0.48$
II	9	-1480.68	5.87	$V_a = 0.41, V_d = 0.13$ L: $V_E = 0.42, C: V_E = 0.46, H: V_E = 0.87$
III	11	-1475.74	0	$V_E = 0.52$ L: $V_a = 0.31, V_d = 0.10$ C: $V_a = 0.37, V_d = 0.12$ H: $V_a = 0.64, V_d = 0.36$
IV	13	-1475.65	3.80	L: $V_a = 0.32, V_d = 0.05, V_E = 0.56$ C: $V_a = 0.37, V_d = 0.22, V_E = 0.41$ H: $V_a = 0.65, V_d = 0.34, V_E = 0.53$

The analysis of untransformed corolla width is given in the upper part of the table and square-root transformed values in the lower part. AIC is standardized to a minimum value of 0. Par. is the number of parameters per model. Variance estimates from each model are scaled so that the Control population has unit variance