The probability of parallel genetic evolution from standing genetic variation

Journal:	*Journal of Evolutionary Biology*
Manuscript ID	JEB-2016-00344.R1
Manuscript Type:	Research Papers
Keywords:	Theory, Ecological genetics, Adaptation, Natural selection

Abstract

Parallel evolution is often assumed to result from repeated adaptation to novel, yet ecologically similar, environments. Here we develop and analyze a mathematical model that predicts the probability of parallel genetic evolution from standing genetic variation as a function of the strength of phenotypic selection and constraints imposed by genetic architecture. Our results show that the probability of parallel genetic evolution increases with the strength of natural selection and effective population size, and is particularly likely to occur for genes with large phenotypic effects. Building on these results, we develop a Bayesian framework for estimating the strength of parallel phenotypic selection from genetic data. Using extensive individual based simulations, we show that our estimator is robust across a wide range of genetic and evolutionary scenarios and provides a useful tool for rigorously testing the hypothesis that parallel genetic evolution is the result of adaptive evolution. An important result that emerges from our analyses of simulated data is that existing studies of parallel genetic evolution frequently rely on data that is insufficient for accurately estimating the strength of parallel natural selection or even distinguishing between adaptive evolution and random genetic drift. Overcoming this challenge will require sampling more populations and the inclusion of larger numbers of loci.

Key Words: Bayesian, QTL, ecological selection, genetic architecture, ecological genetics, adaptation

Introduction

As the availability of genome sequences has increased, interest in understanding how genomic architecture shapes adaptation at both the genetic and phenotypic level has grown substantially (Stapley et al., 2010). How and which genes respond to selection is a complex result of many aspects of the genotype to phenotype map, including allelic effect sizes, epistatic interactions, linkage disequilibrium, and pleiotropy. Significant work using natural populations (Nadeau & Jiggins, 2010), experimental evolution (Qi et al., 2016, Wichman et al., 1999), and evolutionary theory (Orr, 2005, Chevin et al., 2010) has been devoted to elucidating how these many factors interact to shape adaptation. Particularly useful natural systems for addressing such questions are those exhibiting parallel evolution. Many striking examples of repeated phenotypic and genetic change exist (Conte et al., 2012, Stern, 2013, Martin & Orgogozo, 2013), putatively as a consequence of adaptation to similar selective environments (Schluter, 2009). These systems can be viewed as natural experimental replicates for understanding the interplay of selection and genetic architecture in shaping patterns of adaptation (Hohenlohe et al., 2010).

There are many definitions for "parallel evolution", a phenomenon which may or may not be distinguished from "convergent evolution". These many definitions all share the common theme of repeated evolution in two or more populations, but differ in two major ways. First, in terms of whether or not these populations originated from a recent common ancestral population or are only distantly related. Second, definitions differ in the biological level at which repeated evolution occurs, ranging from the genetic to the phenotypic level (Lenormand et al., *In press.*). Here we focus on parallel genetic evolution defined as the repeated fixation of identical alleles in multiple descendent populations. We assume these alleles were initially segregating within the ancestral population at low frequency such that our focus is on parallel evolution from standing genetic variation. Our interest in this definition is

45 motivated by a number of biological systems where adaptation to a novel environment is thought to
46 result from standing genetic variation present in the ancestral population (Hoekstra et al., 2006, Steiner
47 et al., 2007, Colosimo et al., 2005). In contrast to *de novo* mutation, adaptation from standing genetic
48 variation is likely rapid (Barrett & Schluter, 2008) and may lead to distinct genomic signatures of parallel
49 adaptation (Roesti et al., 2014)

50 Understanding the conditions that promote parallel genetic evolution has been facilitated by
51 theoretical studies. For instance, Orr (2005) calculated the probability that 1 of k *de novo* beneficial
52 mutations arises and fixes repeatedly and found that parallel evolution becomes more likely as the
53 strength of selection increases and the number of possible alleles, k, decreases. These results are
54 supported by experimental adaptation of the bacteriophage $\phi X174$ to high temperatures (Wichman et
55 al., 1999) and adaptation of antifungal drug resistance in *Saccharomyces cerevisiae* (Anderson et al.,
56 2003). Taking a different approach, Chevin et al. (2010) calculated the probability that a beneficial *de*
57 *novo* mutation fixes at the same genetic locus in independent populations. By allowing mutations to
58 influence multiple phenotypic traits simultaneously, this work demonstrated that the probability of
59 parallel evolution is greatest when pleiotropy is weak. In addition, this work demonstrated that when
60 mutations have pleiotropic effects, the probability of parallel evolution is greater when populations are
61 relatively close to their adaptive optima (i.e., not too maladapted). Together, these previous theoretical
62 studies provide a solid framework for understanding the likelihood of parallel evolution arising from the
63 fixation of novel mutations.

64 Although understanding the contribution of new mutations to parallel evolution is inarguably
65 important, in some systems it may be more relevant to understand the likelihood of parallel evolution
66 from standing genetic variation. For instance, in the stickleback, *Gasterosteus aculeatus*, repeated
67 adaptation to freshwater is thought to involve genes already segregating at low frequencies within the
68 marine populations (Colosimo et al., 2005). In cases like these, the presence of adaptive alleles in the

ancestral population can have a significant effect on the probability of parallel evolution, influencing both the long term probability of parallel adaptation as well as the rate at which adaptation occurs (Ralph & Coop, 2015). Our focus here is to enhance our understanding of parallel evolution by developing a genetically explicit multi-locus framework for predicting the probability of parallel evolution from standing genetic variation. We have two specific goals: first we will develop a multi-locus theory of parallel genetic adaptation under weak selection and rapid recombination that allows us to predict the probability of parallel evolution in terms of quantities that are regularly measured in natural populations. Second, we will develop a statistical framework that uses genetic data in the form often collected in natural populations to estimate the average strength of parallel selection over time.

The Model

Biological Scenario

We envision a scenario where haploid individuals from an ancestral population colonize two or more novel environments and establish new populations (see Figure 1A). After this initial colonization we assume gene flow between the ancestral and descendent populations is negligible and that individuals within populations mate at random. The descendent populations then experience identical patterns of phenotypic selection causing population mean phenotypes at the focal trait to diverge in parallel from the ancestral population; for example, repeated selection for reduced body armor in multiple freshwater stickleback populations from their common marine ancestral phenotype (Colosimo et al., 2004).

We next envision that the genetic basis of the trait undergoing parallel phenotypic evolution is studied using one of two commonly used experimental designs (Conte et al., 2012). Figure 1B illustrates the first experimental design where parallel genetic evolution is assessed at a set of candidate genes. To identify possible candidate genes, individuals from at least one descendent population (descendent population 1 in Figure 1B) are crossed with ancestral individuals and the resulting offspring are scanned

for divergent QTL's influencing the focal trait. The remaining populations (descendent population 2 in Figure 1B) are then tested for the candidate genes using a variety of approaches such as genetic complementation tests (Hartl, 2005). This method, which we will call the "*candidate gene method*", has been used in human populations to identify the genetic basis of the multiple independent origins of lactose tolerance (Ingram et al., 2009, Enattah et al., 2008, Tishkoff et al., 2007). Alternatively, in the second design (shown in Figure 1C), descendent populations are searched independently for the genes responsible for the repeated phenotypic divergence from the ancestral population. This is done by performing independent QTL scans in each descendent population. This *"QTL method"* was used to identify separate genes responsible for a change in developmental rate in two populations of *Oncorhyncus mykiss,* (Robison et al., 2001, Nichols et al., 2007, Sundin et al., 2005).

Analytical Model

Our model assumes the trait experiencing parallel selection is controlled by n additive loci. Each locus, denoted with the index i, has two possible alleles A_i and a_i and a phenotypic effect equal to b_i associated with the A_i allele, such that the phenotype of an individual is described by

$$z = \bar{z} + \sum_{i=1}^{n} b_i(X_i - p_i), \quad (1)$$

where X_i is an indicator variable taking the value 1 if the individual carries the A allele at locus i and the value 0 if the individual carries the a allele at locus i. We assume b_i is positive for all i implying that the A_i allele always increases the value of the trait z. The frequency of the A_i allele is given by p_i and $\bar{z}$ denotes the average phenotype of the population. We assume the phenotype of the ancestral population is small, meaning that the frequency of the A_i allele is low at all loci, and initially equal to p_{0_i}. Within the new environments, individuals experience selection for large phenotypes, favoring an increase in frequency of the A alleles.

The biggest challenge to modeling evolution across multiple loci is that epistasis and linkage disequilibrium make it extremely difficult to formulate analytical predictions for the probability of

fixation at individual loci. Two key assumptions, however, make calculating the probability of fixation tractable. First, we assume the relationship between an individual's phenotype, z, and its fitness, $W(z)$, is linear:

$$W(z) = \beta z + \alpha. \quad (2)$$

Second, we assume the strength of linear directional selection, β, is weak, $O(\epsilon)$ where ϵ is a small number, and the rate of recombination between loci relatively high. Under these conditions, recombination breaks apart linkage disequilibrium more quickly than it can be built up by selection and a quasi-linkage-equilibrium (QLE) is reached where linkage disequilibrium is also small, $O(\epsilon)$ (Nagylaki, 1993, Nagylaki et al., 1999). Using the expression for the phenotypic trait z, given in equation (1), as well as the expression for fitness, given by equation (2), we can use the multi-locus methods developed by Barton and Turelli (1991) and expanded by Kirkpatrick et al. (2002) to derive the change in the frequency of the A_i allele at QLE over a single generation (*See supplementary material*)

$$\Delta p_i \approx \frac{\beta}{\alpha} b_i p_i (1 - p_i). \quad (3)$$

As expected under QLE, equation (3) does not depend on linkage disequilibrium or epistasis; instead, loci evolve independently. Later, using individual based simulations, we will relax these key assumptions and evaluate the robustness of this analytical approximation.

The independent evolution of loci enables us to utilize a classic result of the Wright-Fisher model describing the probability of fixation for an allele with initial frequency p_0 in a population of constant size N. This probability is given by

$$P_{fix} \approx \frac{\left(1 - e^{2Nsp_0}\right)}{1 - e^{2Ns}} \quad (4)$$

(Kimura, 1957, Karlin & Taylor, 1981) where s is the strength of selection acting on the allele and p_0 is its initial frequency. Under our assumption of linear directional selection, at locus i the strength of selection is $s = \frac{\beta}{\alpha} b_i$, and equation (4) can be re-written as

$$P_{fix}(i) = \frac{\left(1 - e^{2N\frac{\beta}{\alpha} b_i p_{0_i}}\right)}{1 - e^{2N\frac{\beta}{\alpha} b_i}} \quad . \qquad (5)$$

Equation (5) reveals that the probability of fixation depends on initial allele frequency, local population size, the strength of phenotypic selection, and the phenotypic effect of the locus. In the next section we will use this result to explore how these important parameters influence the extent of parallel evolution.

The probability of parallel genetic evolution at a single locus

We begin by analyzing the simplest possible scenario: a single genetic locus. For this case, parallel evolution entails the repeated fixation of the same allele in multiple descendent populations. The probability of this occurring can be calculated by using equation (5) to find the probability that at the locus of interest, i, the A_i allele fixes independently in each of m populations:

$$P_{\parallel} = \left(p_{fix}(i)\right)^m \quad . \qquad (6)$$

Requiring repeated fixation in all m populations represents a very restrictive definition of parallel genetic evolution and in some cases a less restrictive definition may be preferable. In such cases, it is straightforward to develop expressions for the probability of repeated fixation in any subset of m populations using (5). An example of the calculations for a less restrictive definition is provided in the online supplemental material.

Equations (5,6) highlight three important factors that will influence the probability of observing parallel genetic evolution. First, the probability of repeated fixation of an allele increases with its initial frequency, p_{0_i}. Second, large effect alleles, those with large b_i, are more likely to fix in parallel under

directional selection. This relationship between effect size and parallel evolution is shown in Figure 2. Third, parallel genetic evolution is more likely to occur when evolution is driven primarily by the deterministic force of natural selection rather than the stochastic force of random genetic drift. Specifically, the probability of parallel evolution increases with the product of population size and the phenotypic selection gradient in derived populations, $N\frac{\beta}{\alpha}$. This product captures the balance between drift and selection and shows that parallel evolution is more likely in large populations experiencing strong natural selection as shown by the three curves in Figure 2. As this term arises repeatedly in the derivation to follow we will denote it with the composite parameter η

$$\eta = N\frac{\beta}{\alpha}. \quad (7)$$

The probability of parallel genetic evolution at multiple loci

Although the single locus results of the previous section are insightful, they fall short of capturing the genetic richness of real populations where the extent of parallel evolution must be assessed across multiple loci. Fortunately, calculating the probability of parallel evolution across multiple loci is straightforward, and yields the following formula:

$$P_{\parallel} = \prod_{i=1}^{n}\left(p_{fix}(i)\right)^{m}, \quad (8)$$

where the product is carried over the number of loci. Not surprisingly, equation (8) shows that the factors enhancing the probability of parallel evolution at a single locus (e.g., large population size, strong selection, etc.) also increase the probability of parallel evolution across multiple loci. What distinguishes one locus from the next is the initial allele frequency and the allelic effect size. Therefore, the probability of parallel evolution across multiple loci will depend on the distribution of allelic effects. Equation (8) clarifies the connection between the effect size distribution and parallel evolution yielding several novel insights that emerge only when multiple loci are considered.

The first and most obvious insight to emerge from (8) is that perfectly parallel genetic evolution, where all loci are fixed for the selectively favored A_i alleles in all descendent populations, becomes less and less likely as the number of loci increases. This is a simple result of the product rule of probabilities, and arises because the overall probability of parallel evolution decreases as each additional locus is required to fix in parallel in the m descendent populations. The second insight that emerges from equation (8) is that when selection is relatively weak, population sizes are relatively small, and adaptive alleles initially infrequent, it is unlikely to observe parallel evolution at more than a single locus with large phenotypic effect (Figure 3). As selection becomes stronger, population sizes larger, or adaptive alleles initially more frequent, however, it becomes increasingly likely that parallel evolution will occur at multiple loci, including loci with moderate phenotypic effects (Figure 3). These results are, for the most part, relatively insensitive to the particular distribution of effect sizes across loci. Only in cases of strong selection and high initial allele frequency (when evolution becomes more deterministic) does the effect size distribution contribute significantly (bottom right panel of figure 3). In such cases, the probability of parallel evolution at a large number of loci increases with the mode of the effect size distribution. In other words, parallel evolution at a large number of loci is most likely when the effect size distribution is not skewed toward small effect loci (Figure 3). Together, these results suggest that the likelihood of observing parallel genetic evolution at any particular number of loci depends heavily on the value of the parameter η.

Bayesian inferences of parallel phenotypic selection

The results derived in the previous section demonstrate a strong connection between the parameter η and the probability of observing parallel genetic evolution. In this section we develop a method for estimating the value of this key parameter using a Bayesian framework that capitalizes on equation (8). Our goal is to provide a methodology that allows support for a hypothesis of adaptive parallel evolution to be assessed using data collected in empirical studies of parallel genetic evolution.

Specifically, by estimating η it becomes possible to distinguish between parallel genetic evolution caused by random genetic drift, $\eta = 0$, and parallel genetic evolution caused by natural selection.

Our Bayesian approach will rely on genetic data described by a matrix, $\mathcal{D}$, where rows represent descendent populations and columns loci. Each element of $\mathcal{D}$ takes a value of 0 or 1 depending on which allele has fixed at a particular locus in a given population (Figure 1A). Using equation (8) we can develop a likelihood function specifying the probability of observing the data, $\mathcal{D}$, given a particular value for the parameter η, empirical estimates of the effect sizes b_i and initial allele frequencies p_0. The effect sizes can be (and frequently are) estimated using QTL scans (Conte et al., 2015, Broman & Sen, 2009, Lynch & Walsh, 1998) whereas initial allele frequencies can be estimated by measuring the allele frequencies in the ancestral population. Accurate assessment of the genetic data $\mathcal{D}$, the allelic effect sizes, and the initial allele frequencies will require sufficient sample sizes from each descendent population and the ancestral population. We will later address the consequences of uncertainty in the estimation of the initial allele frequencies and allelic effect sizes using individual based simulations. Given estimates for these parameters, the likelihood expression consists of a product of terms, one for each locus of the focal trait in each population. If the A allele has fixed at a locus it contributes a term P_{fix}, as defined by equation (5). Alternatively, if the A allele is lost, it contributes a term $\left(1 - P_{fix}\right)$. Thus, for m populations and n loci, the likelihood of observing the data, $\mathcal{D}$, is given by the following product

$$\mathcal{L}(\mathcal{D}) = \prod_{j=1}^{m} \prod_{i=1}^{n} P_{fix}(\eta, i)^{\mathcal{D}_{ij}} \left(1 - P_{fix}(\eta, i)\right)^{1-\mathcal{D}_{i,j}}, \quad (9)$$

where i is an index over loci and j an index over populations. The likelihood for η as a function of the genetic data $\mathcal{D}$ and the genetic architecture of the trait under selection is based on similar principles to the method of estimating the strength of selection developed by Rice and Townsend (2012). A key

difference, however, is that here the likelihood of parallel genetic evolution is based on only the loci influencing a single phenotypic trait, rather than the entire mutational effect distribution.

The likelihood, (9), can be used in a Bayesian setting to estimate a posterior distribution for the key parameter η. Specifically, Bayes' theorem enables us to formulate estimates for η in the form of the posterior distribution $p(\eta|\mathcal{D})$ that is biologically meaningful for all possible genetic outcomes, $\mathcal{D}$,

$$p(\eta|\mathcal{D}) = \frac{\mathcal{L}(\mathcal{D}|\eta)\pi(\eta)}{\int_{\eta}\mathcal{L}(\mathcal{D}|\eta)\pi(\eta)}, \quad (10)$$

where $\pi(\eta)$ is our prior distribution for the parameter η. The denominator of this expression is the integral over the likelihood surface, and cannot be easily evaluated. For this reason, we use a Markov Chain Monte Carlo algorithm to sample from the posterior distribution and generate an estimate of the most probable value of η for the given genetic data $\mathcal{D}$. We label this estimate $\hat{\eta}$. We take two approaches to evaluating the performance of this estimator. First we analyze its performance under the assumptions of the analytical model by generating the genetic data $\mathcal{D}$ using the Wright-Fisher model. Next, we test the robustness of the estimator to violations of the assumptions of our analytical model by generating the genetic data $\mathcal{D}$ using multi-locus individual based simulations.

Wright-Fisher Simulation

We simulated the data $\mathcal{D}$ for two populations under the Wright-Fisher model by drawing a random number between 0 and 1 for each locus and population and setting $\mathcal{D}_{i,j}$ to 1 if the random number was less than p_{fix} from equation (5) and to 0 otherwise. The value of p_{fix} depends on the initial allele frequency at each locus,p_{0_i}, the allelic effect sizes of each locus, b_i, as well as the parameter η. For each simulation we drew the values of these parameters independently and at random. Initial allele frequencies were drawn independently at each locus from a uniform distribution between 0 and 0.1. Because our model envisions divergence of descendent populations from a common ancestor we assumed that the initial frequency at any one locus was the same in both populations. Allelic effect sizes

were drawn independently for each locus from a uniform distribution between 0 and 1. The value of η for each run was drawn from a uniform distribution ranging between 0 to 50. The genetic outcome $\mathcal{D}$ simulated in this manner may not, however, resemble what would be measured using experimental methods. For example, using current genomic techniques it is not possible to identify loci that have not diverged from the ancestral state. To address how experimental methodologies affect our Bayesian estimates we considered two modified forms of $\mathcal{D}$ that resemble sampling under the two experimental methods described previously (see Figure 1). The first of these methods, the candidate gene method (Figure 1B), assesses parallel genetic evolution at candidate genes which are known to have generated the phenotypic divergence in the first descendent population. This is often done by performing a cross between individuals from one of the divergent populations with the ancestral population and assessing the genetic variation in the F1s. Since the second divergent population is not independently assessed for divergent QTLs, under this method we only consider the columns of $\mathcal{D}$ (i.e., loci) where the A allele has fixed in the first population. The second experimental method, the QTL method, (Figure 1C) independently assesses divergent loci in all descendent populations. Under this method, $\mathcal{D}$ therefore contains all columns (loci) which have fixed in at least one population. Hence the effective number of loci identified using this method will always be greater than or equal to the number found by the less thorough candidate gene method.

For each simulated $\mathcal{D}$, as well as for $\mathcal{D}$ modified by the two experimental methods, we estimated η using a Metropolis-Hastings algorithm as described in the supplementary material. For the prior $\pi(\eta)$ we used a uniform distribution on the interval $\eta = \pm 80$. To analyze the performance of the estimator we ran a regression of the estimated values of $\hat{\eta}$ on the true values η, using 200 data points. Overall, this analysis revealed that the estimator was quite accurate, explaining between 80% and 85% of the variation (see Table S1). In addition, our analysis showed that the accuracy of the estimates increases with the number of loci. This trend holds regardless of the experimental method used.

However, the effective number of loci under the QTL method is always greater than when candidate genes are first identified in one population and then subsequently searched for in the other. The results of these simulations suggest our estimator performs quite well when the data meet the assumptions of our analytical model; however, this may not be the case for real data. In the next section, we explore the performance of our estimator using individual based simulations that allow us to violate key assumptions of our analytical model such as weak selection and frequent recombination.

Individual Based Simulation

Our individual based simulations consider two allopatric populations, each of which has a constant size of $N = 1000$ individuals. Initial allele frequencies and effect sizes at each locus, as well as the value of η, were drawn randomly as described above under the Wright-Fisher model. Individuals within each population undergo a two stage life cycle. During the first stage, "selection", the probability that an individual survives is given by its fitness, with fitness computed using either equation (2) which describes linear selection or an expression for stabilizing selection described below. Surviving individuals then enter the second life cycle stage, "reproduction", which consists of generating an offspring population from the surviving parental population. This is done by drawing a pair of parents at random from the pool of surviving individuals and producing an offspring from these parents by recombining the parental genomes at a specified rate r and allowing mutation between the two allelic states at a per locus mutation rate of $\mu = 10^{-6}$. This process is continued with replacement of parents until the offspring population reaches the pre-selection size of N. This life cycle is repeated until all loci approach fixation or loss (allele frequencies > 0.99 or < 0.01) at which point the simulations were terminated and the matrix of genetic data $\mathcal{D}$ filled by rounding the allele frequency to 0 or 1. As in the previous section, we formulate modified versions of $\mathcal{D}$ that resemble sampling under the two experimental methods. Then, using the Metropolis-Hastings algorithm, we compute estimates for the

296 value of η using the original outcome $\mathcal{D}$ as well as the two modified forms of $\mathcal{D}$ (*see Supplementary*
297 *Material*).

298 We used the simulations to test the robustness of the estimator when selection is strong and/or
299 non-linear. To test the effect of non-linear selection, simulations were run where an individual's fitness
300 was determined by one of two alternative forms of selection: linear directional selection described by
301 (2), or stabilizing selection toward a phenotypic optimum:

302 $$W(z) = e^{-\gamma(z-\theta)^2}, \quad (11)$$

303 where θ is the phenotypic optimum and γ is the strength of stabilizing selection. Including simulations
304 where selection is stabilizing is important because it relaxes our previous assumption that loci evolve
305 independently. Stabilizing selection is particularly useful in testing this assumption because the extent of
306 interdependence between loci can be manipulated by changing the value of the phenotypic optimum.
307 Specifically, the extent of interdependence between loci will depend on the value of the optimum
308 relative to the largest possible phenotype $z_{max} = \sum_i b_i$. When θ is greater than the largest possible
309 phenotype, z_{max}, loci remain relatively independent as directional selection predominates over
310 epistatic selection. However when $\theta < z_{max}$ this is no longer true as epistatic selection now dominates.
311 Therefore, when $\theta > z_{max}$ evolution is much more likely to resemble linear selection as our analytical
312 model assumed. We simulated these two forms of stabilizing selection respectively, by either requiring
313 that θ be larger than z_{max} or slightly smaller than z_{max} (*see Supplementary material*). Under stabilizing
314 selection η changes as the population adapts, decreasing as the population approaches the optimum
315 (Chevin & Hospital, 2008, Matuszewski et al., 2015). Therefore, we computed a "realized" strength of
316 linear selection by averaging the selection gradient, $\frac{Cov(z,w)}{Var(z)}$, over all time points for which $Var(z) \neq 0$.

317 As expected, analysis of simulated data shows that the accuracy of our estimates depends on
318 the form of selection. Specifically, estimates for η are most accurate under linear selection, somewhat

less accurate under stabilizing selection toward a distant optimum. $\theta > z_{max}$, and least accurate under stabilizing selection toward a close optimum, $\theta < z_{max}$ (See Figure 4 and Table S2). In addition to assuming that selection is linear, we also assumed that selection is weak. By computing the variance about the regression line as η increased we were able to confirm that, for the data shown in figure 4, the accuracy of our estimates decreases with increasing selection. Next, we used our simulations to explore the sensitivity of our estimator to infrequent recombination among candidate loci (See Table S3). Not surprisingly, these simulations revealed that our estimator performs better when recombination is frequent ($r = 0.5$) than when recombination is rare ($r = 0.05$). The effect of infrequent recombination is more drastic for stabilizing selection than linear selection, and is particularly pronounced when $\theta < z_{max}$. This is expected since this latter scenario generates the strongest epistatic selection and thus has the greatest potential to cause linkage disequilibrium to accumulate.

Finally, we used the individual based simulations to test the accuracy of our estimator when several assumptions of the biological scenario envisioned above are violated. These violations included recurrent gene flow from the ancestral to the descendent populations, gene flow among descendent populations, selection that differs in strength across the two descendent populations, and estimates of the parameters p_0 and b_i that are imprecise (*see Supplementary Material and Table S4-S6 and Figure S1*). These simulations reveal that our estimates of η are robust to violations in many of these assumptions. Specifically, the estimator performs well when migration occurs between descendent populations or from the ancestral population as long as rates remain below 2 migrants per generation. In addition, estimates for the average selection gradient in the descendent populations remain robust even when the true selection gradients differ among descendent populations by up to 20%. Lastly, the estimator is robust to error in estimated values of the parameters p_0 and b_i, as long as the estimated values differ from their true values by 10% or less.

Up to this point we have focused on using the Bayesian estimator to provide single point estimates for η. Having access to the full posterior distribution allows us to calculate a 95% credible interval for the parameter η and determine whether or not it overlaps with zero. From an empirical standpoint, being able to rule out $\eta = 0$ allows us to reject the hypothesis that observed levels of parallel genetic evolution can be explained by random genetic drift alone. The open versus closed circles in Figure 4 represent data points for which credible intervals drawn from the posterior distribution do, or do not, overlap zero respectively. Figure 5 shows how the probability of rejecting 0 (filled bars) increases with η.

Discussion

It has long been understood that natural selection, parallel genetic evolution, and genomic architecture are inherently linked (Orr, 2005, Schluter, 2009, Chevin et al., 2010). Previous theoretical work has focused on repeated genetic evolution from new mutation and found that key components of genetic architecture, such as the number of possible beneficial mutations (Orr, 2005) and the distribution of mutational effects (Chevin et al., 2010) influence the probability of parallel evolution. Here we have used a multi-locus model of parallel evolution from standing genetic variation to continue to formalize these connections. We began our investigation by calculating the probability of parallel evolution at a single locus and showed that parallel evolution is most likely when phenotypic selection is strong, standing genetic variation for adaptive alleles is appreciable, adaptive alleles have large phenotypic effects, and population sizes are large. Next, we extended our analyses to multiple loci, demonstrating that the number of loci that evolve in parallel depends on the product of phenotypic selection and local population size (η). If selection is relatively weak, or population sizes are small, we expect parallel evolution at no more than a single locus. In contrast, when selection is relatively strong, or population sizes very large, parallel evolution may occur across multiple loci. These results demonstrate that without information on the strength of phenotypic selection and population size, we

have no way to assess whether the amount of parallel genetic evolution we observe in an empirical study is beyond what would be expected under neutrality. To remedy this problem, and better connect studies of parallel genetic evolution to the evolutionary processes they imply, we developed a Bayesian approach that capitalizes on available genetic data to estimate the product of phenotypic selection and local population size (η). In the following paragraphs we explore several of the key results in more detail and discuss their implications for past, present, and future studies of parallel genetic evolution.

The first important result that emerges from all of our models is that parallel evolution is most likely to be observed at loci with large phenotypic effects on traits experiencing strong phenotypic selection in novel environments. This result receives at least some support from empirical studies of parallel genetic evolution. For example the large effect gene Eda has been found in eight fresh-water descendent populations of three-spine stickleback, *Gasterosteus aculeatus,* and is largely responsible for the parallel reduction in lateral plate number in these populations. In contrast, the small effect locus LG7 has been confirmed in only two of the 8 descendent populations (Schluter et al., 2004, Colosimo et al., 2004, Conte et al., 2012). It seems likely that this example — where a stark ecological shift from salt to fresh water has occurred — corresponds to a case where natural selection is quite strong. Another important, albeit unsurprising, result of equation (6)is that the probability of parallel evolution at a single locus also depends on effect size and initial allele frequency. This may help explain why, contrary to the results described above, a recent comprehensive survey of allelic effects involved in parallel adaptation in two stickleback populations has found no correlation between effect size and probability of repeated gene use (Conte et al., 2015). Our results suggest that the lack of correlation may be the result of highly variable initial allele frequencies among loci.

By integrating multi-locus genetics into a model of adaptation, we were also able to derive expressions for the probability of observing parallel genetic evolution at various numbers of loci over the course of adaptation. The most important result to emerge from this analysis is that in the absence of

information about the likely strength of phenotypic selection in derived populations and the number of individuals composing these derived populations, there is no way to assess the significance of observing parallel evolution at any particular number of loci. Put differently, if natural selection in novel environments is quite strong or population sizes in novel environments quite large, observing parallel evolution at multiple genetic loci is not too surprising. If, however, natural selection is weak or population sizes very small, observing this same level of genetic parallelism would be rather unexpected. This suggests that if we are to more rigorously interpret the results of empirical studies of parallel genetic evolution, we must do better than simply counting up the number of parallel genetic changes observed. Our Bayesian tool accomplishes this goal by providing a methodology for tying information on the extent of parallel genetic information to underlying evolutionary processes.

Such an estimator is only useful if it produces accurate predictions across a range of parameter space. Indeed, our individual based simulations reveal that our estimator can be quite accurate and robust although there are limitations. For example, accurate estimation requires data from at least 8 total loci, be that 4 loci in two populations, 2 loci in four populations, or some intermediate combination. Whether data is gathered at fewer loci in many populations or many loci in few populations should, in principle, have no effect on the accuracy or efficiency of the estimator. Many of the studies discussed above however, have far fewer than this. For example, studies of parallel pigmentation changes in a variety of species, from beach mice (Hoekstra et al., 2006) to cave fish (Protas et al., 2006, Gross et al., 2009), focus primarily on one or two loci in somewhere between 2 and 6 populations. Therefore if future studies hope to understand the role of natural selection in driving parallel evolution, it is important that they focus on acquiring data from as many loci and as many populations as possible. Assessing parallel genetic evolution at more loci will likely require studying loci with only small phenotypic effects on the focal trait, but no such trade-off is required by utilizing additional populations. We explored two alternative experimental methods (see Figure 1 B and C), that

differ predictably in the number of loci that are detected. Because the QTL method always detects parallel evolution at an equal or larger number of loci, we recommend its use over the candidate gene method. The accuracy of the Bayesian estimate is influenced not only by the amount of available data but the accuracy of the estimated parameter values b_i and p_{0_i}. We have tested the robustness of the estimate to error in these parameters (*See the supplementary Material*). Because we may often be uncertain about the exact value of p_{0_i} due to sampling error or stochastic variation in small populations (Hermisson & Pennings, 2005), in some case it may be more appropriate to run the Bayesian estimator with a prior distribution for the parameter p_{0_i} rather than a single point estimate.

In addition to requiring information on parallel genetic evolution drawn from a reasonably large number of loci, our estimator relies on several assumptions that may affect its accuracy. For example, our approach assumes that population size is constant across time, and thus does not allow for extreme bottlenecks or extensive founder effects. This assumption may prove particularly important in cases of repeated evolution of reduced skin pigmentation in European and Asian human populations for which there is evidence for extensive bottlenecks (Schmegner et al., 2005, Amos & Hoffman, 2010). Finally, our approach assumes that selection/population size is identical in each population and that recurrent gene flow does not occur. Although these assumptions may ultimately prove important in some cases, our individual based simulations show that they have only a limited impact on the accuracy of estimates in most cases (*See Supplementary Material*).

Combined, our analyses of single and multi-locus models show that it is difficult to draw conclusions about the biological significance of parallel genetic evolution without information on the strength of parallel phenotypic selection and local population size. Our estimator provides a robust statistical methodology for translating observed levels of genetic parallelism into an estimate of the product of phenotypic selection and local population size. As a consequence, our approach provides a

much needed tool for distinguishing between adaptive and non-adaptive hypotheses for observed levels of parallel genetic evolution. Applying this method to existing and emerging data from multiple populations with common ancestry may thus provide novel insights into the importance of adaptive evolution in natural populations.

References

Amos, W. & Hoffman, J. I. 2010. Evidence that two main bottleneck events shaped modern human genetic diversity. *Proceedings of the Royal Society B-Biological Sciences* **277**: 131-137.

Anderson, J. B., Sirjusingh, C., Parsons, A. B., Boone, C., Wickens, C., Cowen, L. E. & Kohn, L. M. 2003. Mode of selection and experimental evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. *Genetics* **163**: 1287-1298.

Barrett, R. D. & Schluter, D. 2008. Adaptation from standing genetic variation. *Trends Ecol Evol* **23**: 38-44.

Broman, K. W. & Sen, S. 2009. *A guide to QTL mapping with R/qtl*. Springer, Dordrecht.

Chevin, L. M. & Hospital, F. 2008. Selective sweep at a quantitative trait locus in the presence of background genetic variation. *Genetics* **180**: 1645-1660.

Chevin, L. M., Martin, G. & Lenormand, T. 2010. Fisher's model and the genomics of adaptation: Restricted pleiotropy, heterogenous mutation, and parallel evolution. *Evolution* **64**: 3213-3231.

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Schluter, D. & Kingsley, D. M. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**: 1928-1933.

Colosimo, P. F., Peichel, C. L., Nereng, K., Blackman, B. K., Shapiro, M. D., Schluter, D. & Kingsley, D. M. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology* **2**: 635-641.

Conte, G. L., Arnegard, M. E., Best, J., Chan, Y. F., Jones, F. C., Kingsley, D. M., Schluter, D. & Peichel, C. L. 2015. Extent of QTL reuse during repeated phenotypic divergence of sympatric threespine stickleback. *Genetics*.

Conte, G. L., Arnegard, M. E., Peichel, C. L. & Schluter, D. 2012. The probability of genetic parallelism and convergence in natural populations. *Proceedings of the Royal Society B-Biological Sciences* **279**: 5039-5047.

Enattah, N. S., Jensen, T. G. K., Nielsen, M., Lewinski, R., Kuokkanen, M., Rasinpera, H., El-Shanti, H., Seo, J. K., Alifrangis, M., Khalil, I. F., Natah, A., Ali, A., Natah, S., Comas, D., Mehdi, S. Q., Groop, L., Vestergaard, E. M., Imtiaz, F., Rashed, M. S., Meyer, B., Troelsen, J. & Peltonen, L. 2008. Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. *American Journal of Human Genetics* **82**: 57-72.

Gross, J. B., Borowsky, R. & Tabin, C. J. 2009. A novel role for Mc1r in the parallel evolution of depigmentation in independent populations of the cavefish *Astyanax mexicanus*. *PLoS Genetics* **5**.

Hartl, D. J., E.W. 2005. *Genetics: Analysis of genes and genomes*. Jones & Bartlett Learning.

Hermisson, J. & Pennings, P. S. 2005. Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics* **169**: 2335-2352.

Hoekstra, H. E., Hirschmann, R. J., Bundey, R. A., Insel, P. A. & Crossland, J. P. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**: 101-104.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A. & Cresko, W. A. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6**.

Ingram, C. J. E., Mulcare, C. A., Itan, Y., Thomas, M. G. & Swallow, D. M. 2009. Lactose digestion and the evolutionary genetics of lactase persistence. *Human Genetics* **124**: 579-591.

Karlin, S. & Taylor, H. M. 1981. *A second course in stochastic processes*. Academic Press, New York.

Kimura, M. 1957. Some problems of stochastic-processes in genetics. *Annals of Mathematical Statistics* **28**: 882-901.

Lenormand, T., Chevin, L. M. & Bataillon, T. M. (*In press.*) Parallel evolution: what does it (not) tell us and why is it (still) interesting? In: *Chance in Evolution*, (Ramsey, G. P., C.H., ed.). pp. Univ. Chicago Press.

Lynch, M. & Walsh, B. 1998. *Genetics and analysis of quantitative traits*. Sinauer, Sunderland, Mass.

Martin, A. & Orgogozo, V. 2013. The loci of repeated evolution: A catalog of genetic hotspots of phenotypic variation. *Evolution* **67**: 1235-1250.

Matuszewski, S., Hermisson, J. & Kopp, M. 2015. Catch me if you can: Adaptation from standing genetic variation to a moving phenotypic optimum. *Genetics* **200**: 1255-74.

Nadeau, N. J. & Jiggins, C. D. 2010. A golden age for evolutionary genetics? Genomic studies of adaptation in natural populations. *Trends in Genetics* **26**: 484-492.

Nichols, K. M., Broman, K. W., Sundin, K., Young, J. M., Wheeler, P. A. & Thorgaard, G. H. 2007. Quantitative trait loci x maternal cytoplasmic environment interaction for development rate in *Oncorhynchus myhiss*. *Genetics* **175**: 335-347.

Orr, H. A. 2005. The probability of parallel evolution. *Evolution* **59**: 216-220.

Protas, M. E., Hersey, C., Kochanek, D., Zhou, Y., Wilkens, H., Jeffery, W. R., Zon, L. I., Borowsky, R. & Tabin, C. J. 2006. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nature Genetics* **38**: 107-111.

Qi, Q., Toll-Riera, M., Heilbron, K., Preston, G. M. & MacLean, R. C. 2016. The genomic basis of adaptation to the fitness cost of rifampicin resistance in *Pseudomonas aeruginosa*. *Proceedings of the Royal Society B-Biological Sciences* **283**.

Ralph, P. L. & Coop, G. 2015. The role of standing variation in geographic convergent adaptation. *American Naturalist* **186**: S5-S23.

Rice, D. P. & Townsend, J. P. 2012. A Test for Selection Employing Quantitative Trait Locus and Mutation Accumulation Data. *Genetics* **190**: 1533-+.

Robison, B. D., Wheeler, P. A., Sundin, K., Sikka, P. & Thorgaard, G. H. 2001. Composite interval mapping reveals a major locus influencing embryonic development rate in rainbow trout (*Oncorhynchus mykiss*). *Journal of Heredity* **92**: 16-22.

Roesti, M., Gavrilets, S., Hendry, A. P., Salzburger, W. & Berner, D. 2014. The genomic signature of parallel adaptation from shared genetic variation. *Molecular Ecology* **23**: 3944-3956.

Schluter, D. 2009. Evidence for ecological speciation and its alternative. *Science* **323**: 737-741.

Schluter, D., Clifford, E. A., Nemethy, M. & McKinnon, J. S. 2004. Parallel evolution and inheritance of quantitative traits. *American Naturalist* **163**: 809-822.

Schmegner, C., Hoegel, J., Vogel, W. & Assum, G. 2005. Genetic variability in a genomic region with long-range linkage disequilibrium reveals traces of a bottleneck in the history of the European population. *Human Genetics* **118**: 276-286.

Stapley, J., Reger, J., Feulner, P. G. D., Smadja, C., Galindo, J., Ekblom, R., Bennison, C., Ball, A. D., Beckerman, A. P. & Slate, J. 2010. Adaptation genomics: The next generation. *Trends in Ecology & Evolution* **25**: 705-712.

Steiner, C. C., Weber, J. N. & Hoekstra, H. E. 2007. Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biology* **5**: 1880-1889.

Stern, D. L. 2013. The genetic causes of convergent evolution. *Nature Reviews Genetics* **14**: 751-764.

Sundin, K., Brown, K. H., Drew, R. E., Nichols, K. M., Wheeler, P. A. & Thorgaard, G. H. 2005. Genetic analysis of a development rate QTL in backcrosses of clonal rainbow trout, *Oncorhynchus mykiss*. *Aquaculture* **247**: 75-83.

Tishkoff, S. A., Reed, F. A., Ranciaro, A., Voight, B. F., Babbitt, C. C., Silverman, J. S., Powell, K., Mortensen, H. M., Hirbo, J. B., Osman, M., Ibrahim, M., Omar, S. A., Lema, G., Nyambo, T. B., Ghori, J., Bumpstead, S., Pritchard, J. K., Wray, G. A. & Deloukas, P. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nature Genetics* **39**: 31-40.

Wichman, H. A., Badgett, M. R., Scott, L. A., Boulianne, C. M. & Bull, J. J. 1999. Different trajectories of parallel evolution during viral adaptation. *Science* **285**: 422-424.

Figure Legends

Figure 1: Schematic of biological scenario. Panel A depicts two descendent populations diverging in parallel from a common ancestral population. The a allele predominates at all four loci in the ancestral population whereas the A allele fixes at various loci in the two descendent populations. Panel B and C depict two methods for deducing the underlying genetics of reduced body size in the two descendent populations depicted in panel A. Panel B shows the candidate gene method which relies on a genome wide scan of progeny from a cross between the first descendent population and the ancestral population and subsequent candidate gene search in the second descendent population. Panel C shows the QTL method which involves two genome wide scans, one in each population. Compared to the candidate gene method, the QTL method uncovers an additional locus driving divergence in population 2.

Figure 2: The probability of parallel evolution as a function of allelic effect size, $\boldsymbol{b}$. For a given strength of selection the probability of fixation, and hence parallel evolution, increases with allelic effect size. The rate of increase is non-linear and depends on the strength of selection s and the population size N which are given by the compound parameter $\eta = Ns$. The initial alele frequency for the three curves was held constant at $p_0 = 0.01$.

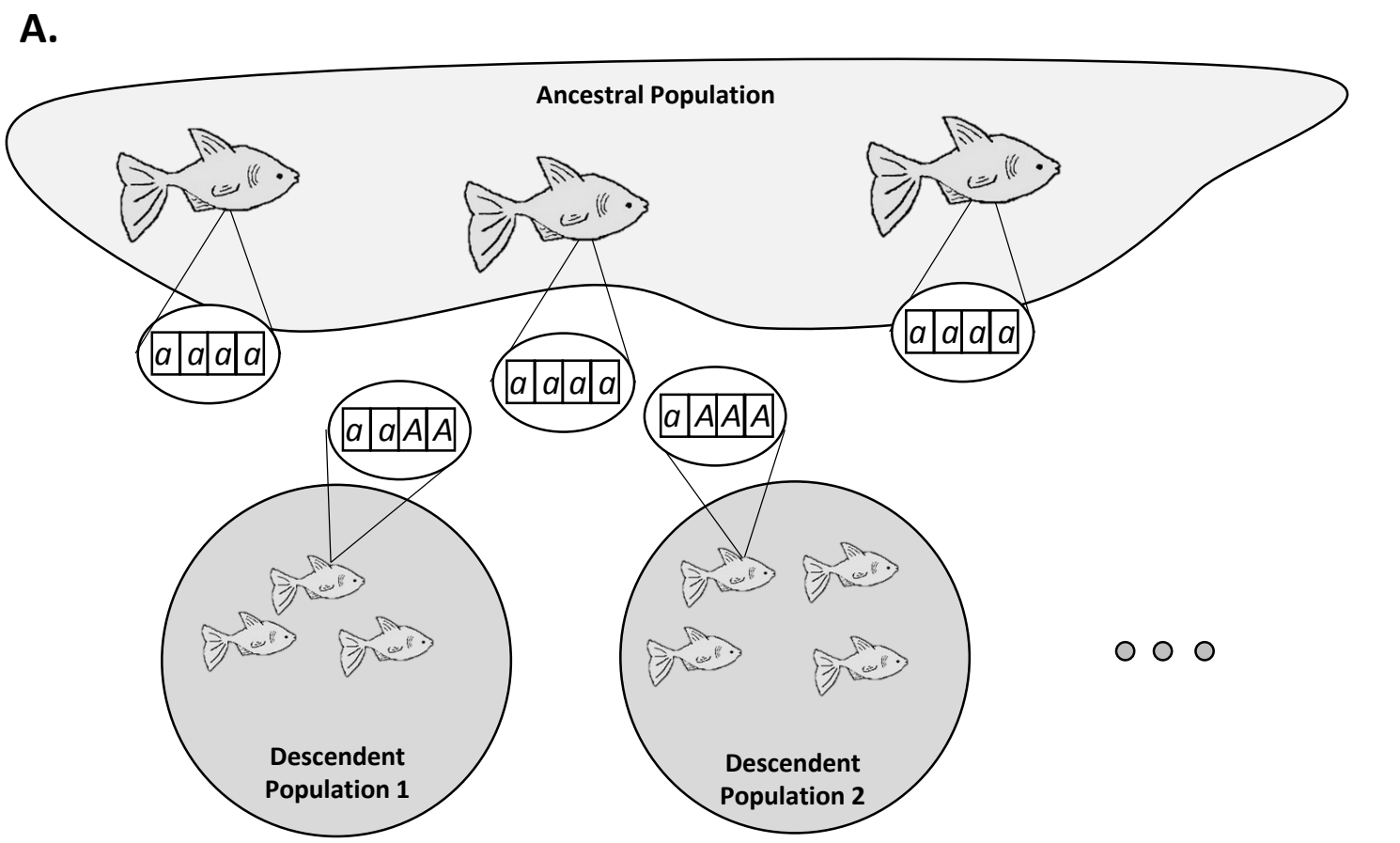
Figure 3: The probability of parallel evolution at $\boldsymbol{n}$ loci. For a trait determined by the effects of 10 total loci the probability of observing parallel evolution at exactly n loci depends on the strength of selection, which varies from near 0 to a value of $\eta = 50$, and the intial allele frequency, which is either low (0.01) or high (0.1). The probability of parallel evolution may also depend on the underlying effect size distribution depicted here (top panel) as three different gamma probability distributions with different

shape and scale parameters (Red: $k = 1, \theta = \frac{1}{2}$, Blue: $k = 2, \theta = \frac{1}{4}$, Green: $k = 5, \theta = \frac{1}{5}$,) but with the same mean effect size ($\mu = \frac{1}{2}$).

Figure 4: Regression Fit of IBS Data under the three forms of selection. Data and linear regression fit (blue), and perfect fit (red) between the time averaged values of η and the Bayesian estimate $\hat{\eta}$ for 200 replicates of the individual based simulation. Open (closed) points indicate estimates where the 95% credible interval does not (do) overlap $\eta = 0$. Sub-panels differ in the number of loci (ranging from 2 to 8) and the form of natural selection (Linear, Stabilizing with $\theta > z_{max}$, and Stabilizing with $\theta < z_{max}$). The Bayesian estimator uses genetic data filtered to resemble sampling using the QTL experimental method.

Figure 5: Significance of $\widehat{\boldsymbol{\eta}}$ estimates. The fraction of estimates that differ significantly (black) and do not differ significantly (white) from $\widehat{\boldsymbol{\eta}}$ =0 with 95% confidence across the range of time average η values. Overlapping portions are shown in grey.

Figure 1:


A.
Ancestral Population
a a a a
a a a a
a a A A
a A A A
a a a a
Descendent Population 1
Descendent Population 2

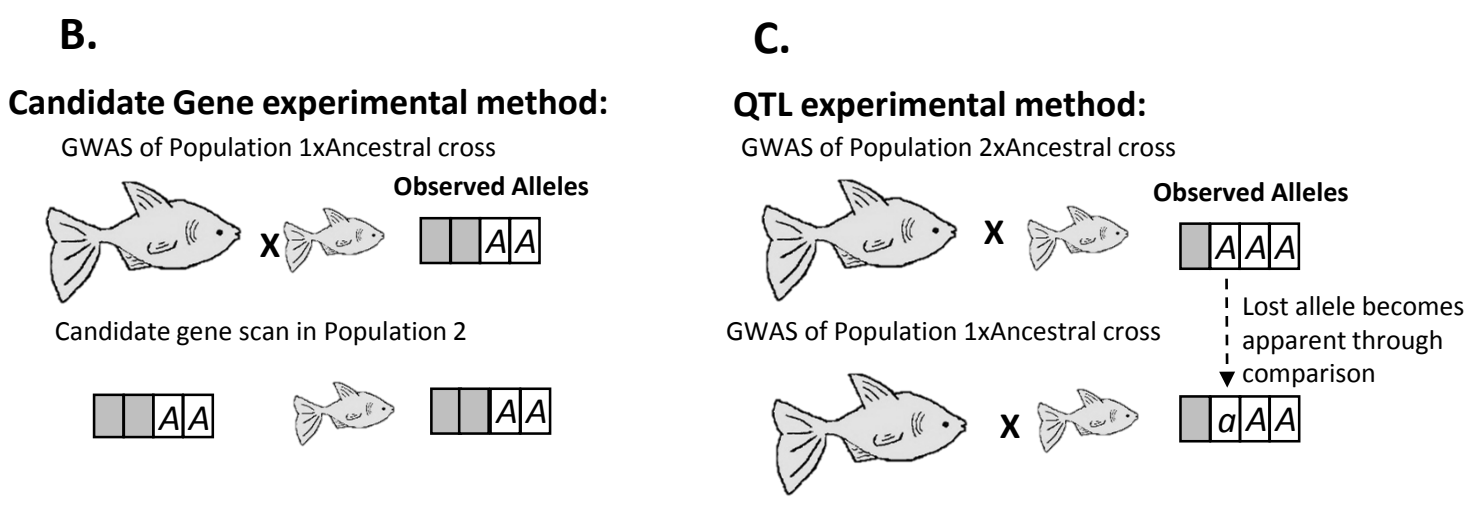

B.
Candidate Gene experimental method:
GWAS of Population 1xAncestral cross
Observed Alleles
X
A A
Candidate gene scan in Population 2
A A
A A
C.
QTL experimental method:
GWAS of Population 2xAncestral cross
Observed Alleles
X
A A A
Lost allele becomes apparent through comparison
GWAS of Population 1xAncestral cross
X
a A A

Figure 2:

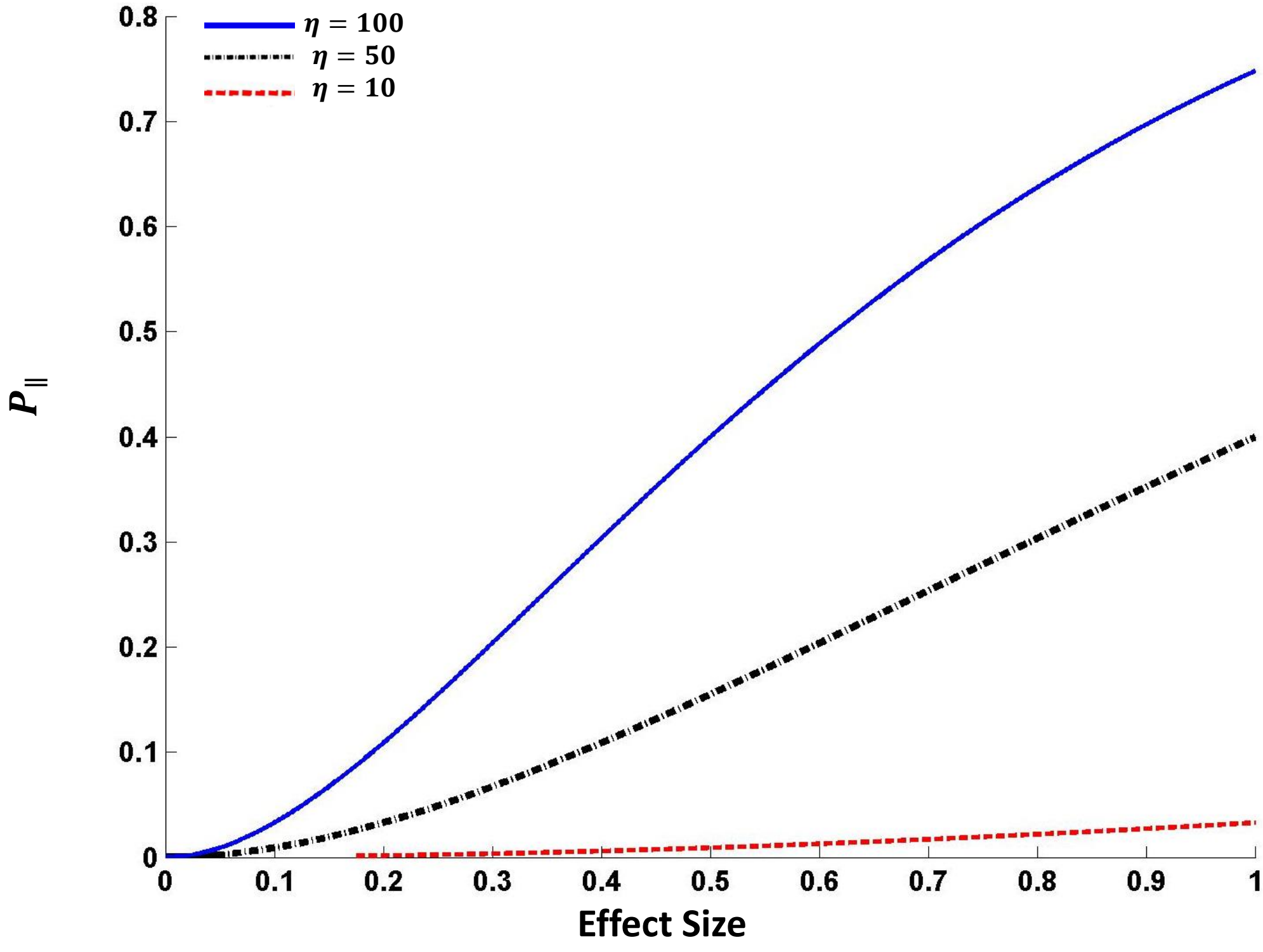

$\eta = 100$
$\eta = 50$
$\eta = 10$
$P_{\parallel}$
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Effect Size

Figure 3:

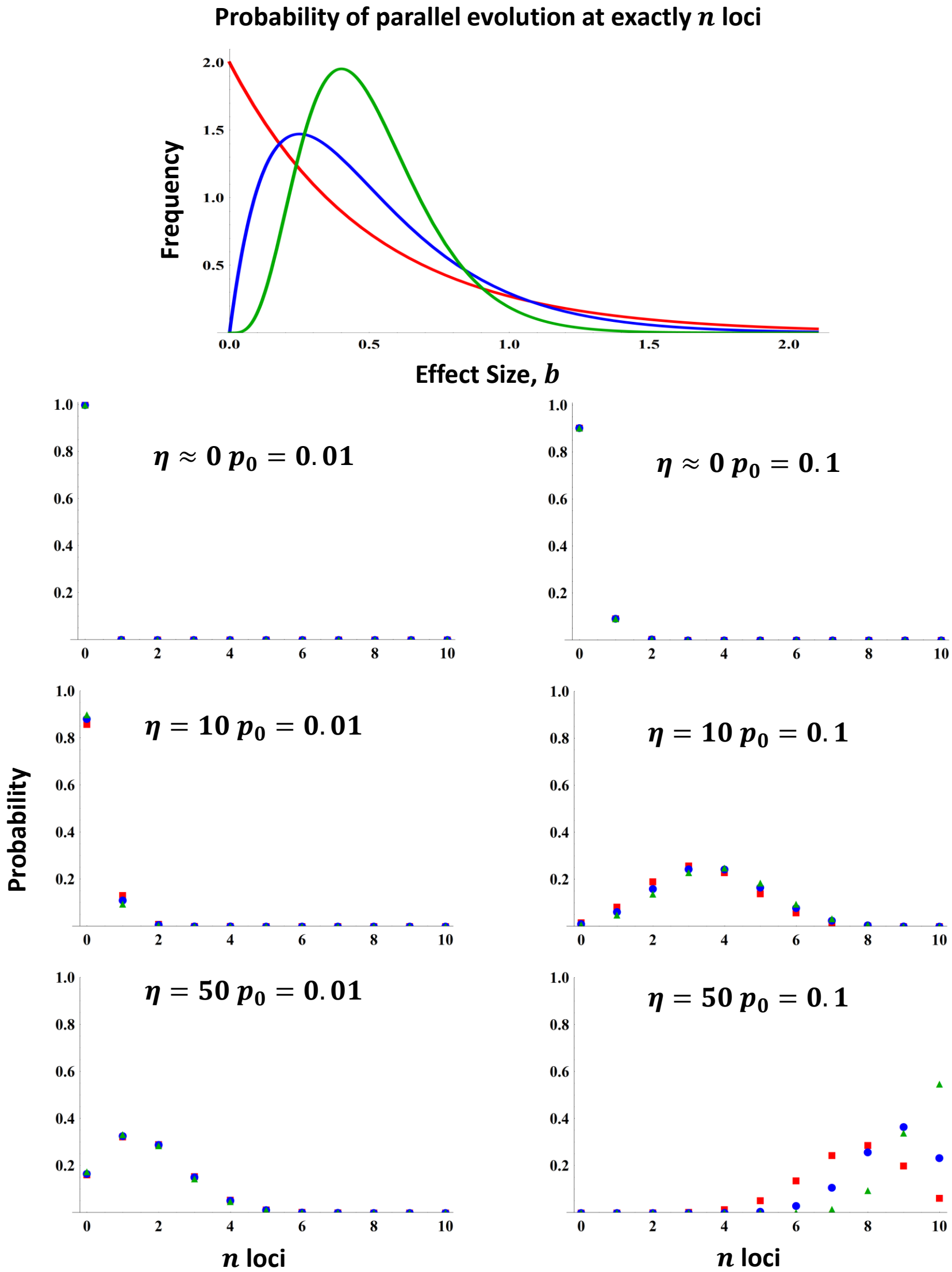

Probability of parallel evolution at exactly n loci
Frequency
2.0
1.5
1.0
0.5
0.0
0.5
1.0
1.5
2.0
Effect Size, b
$\eta \approx 0\ p_0 = 0.01$
$\eta \approx 0\ p_0 = 0.1$
$\eta = 10\ p_0 = 0.01$
$\eta = 10\ p_0 = 0.1$
$\eta = 50\ p_0 = 0.01$
$\eta = 50\ p_0 = 0.1$
Probability
1.0
0.8
0.6
0.4
0.2
0
2
4
6
8
10
n loci

Figure 4

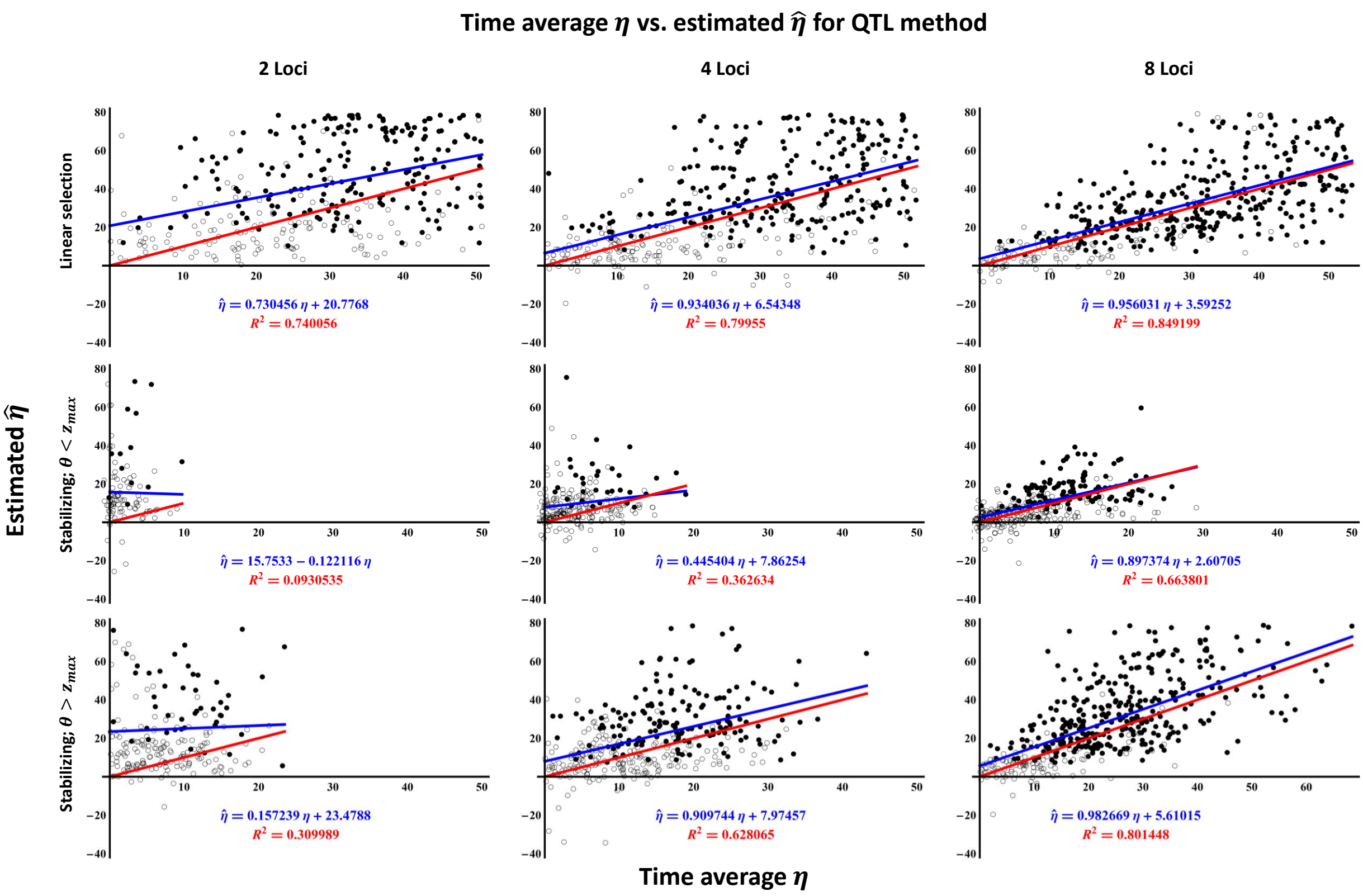

Time average η vs. estimated $\hat{\eta}$ for QTL method
2 Loci
4 Loci
8 Loci
Linear selection
Stabilizing; $\theta < z_{max}$
Stabilizing; $\theta > z_{max}$
Estimated $\hat{\eta}$
Time average η
$\hat{\eta} = 0.730456\,\eta + 20.7768$
$R^2 = 0.740056$
$\hat{\eta} = 0.934036\,\eta + 6.54348$
$R^2 = 0.79955$
$\hat{\eta} = 0.956031\,\eta + 3.59252$
$R^2 = 0.849199$
$\hat{\eta} = 15.7533 - 0.122116\,\eta$
$R^2 = 0.0930535$
$\hat{\eta} = 0.445404\,\eta + 7.86254$
$R^2 = 0.362634$
$\hat{\eta} = 0.897374\,\eta + 2.60705$
$R^2 = 0.663801$
$\hat{\eta} = 0.157239\,\eta + 23.4788$
$R^2 = 0.309989$
$\hat{\eta} = 0.909744\,\eta + 7.97457$
$R^2 = 0.628065$
$\hat{\eta} = 0.982669\,\eta + 5.61015$
$R^2 = 0.801448$

Figure 5:

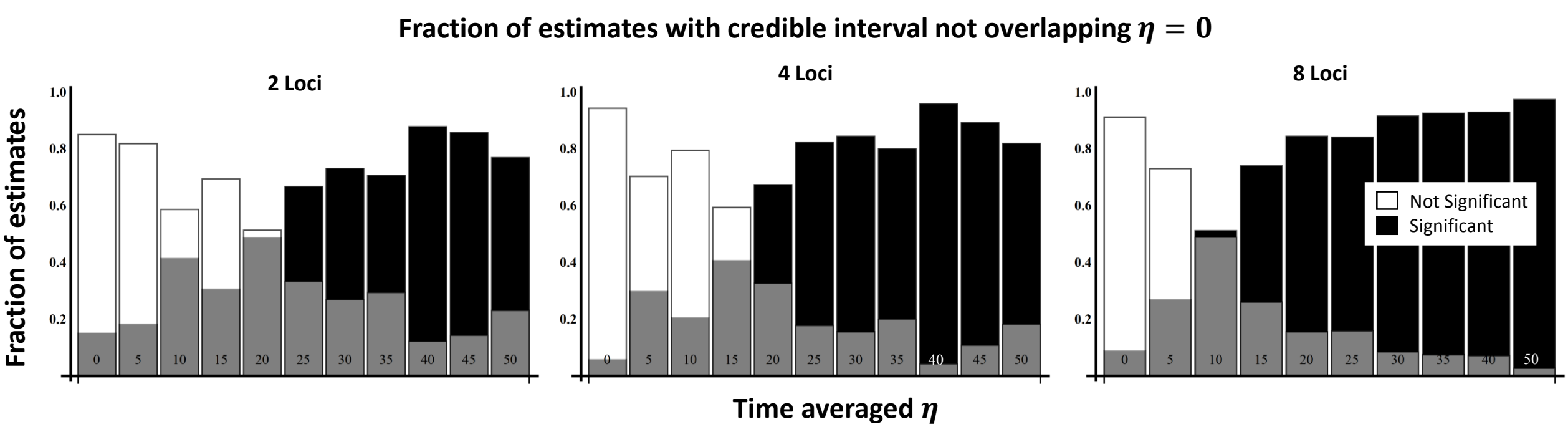
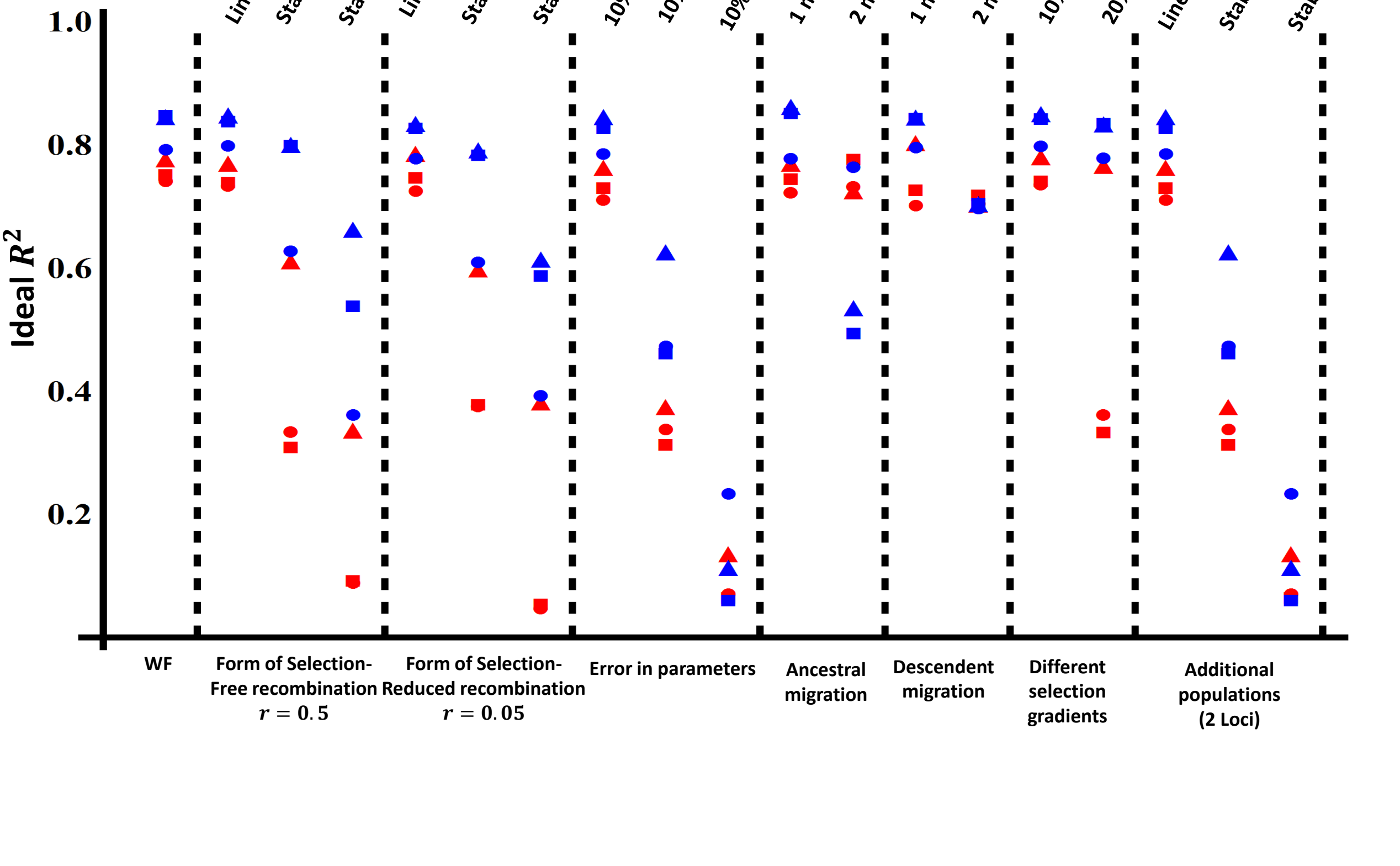

Fraction of estimates with credible interval not overlapping $\eta = 0$
2 Loci
4 Loci
8 Loci
Fraction of estimates
1.0
0.8
0.6
0.4
0.2
0
5
10
15
20
25
30
35
40
45
50
Not Significant
Significant
Time averaged η

Figure S1:


Linear selection
Stabilizing $\theta > z_{max}$
Stabilizing $\theta < z_{max}$
Linear Selection
Stabilizing $\theta > z_{max}$
Stabilizing $\theta < z_{max}$
10% error in p_0 only
10% error in b only
10% error in p_0 and b
1 migrant/generation
2 migrants/ generation
1 migrant/generation
2 migrants/generation
10% different
20% different
Linear selection
Stabilizing $\theta > z_{max}$
Stabilizing $\theta < z_{max}$
Ideal R^2
1.0
0.8
0.6
0.4
0.2
WF
Form of Selection- Free recombination $r = 0.5$
Form of Selection- Reduced recombination $r = 0.05$
Error in parameters
Ancestral migration
Descendent migration
Different selection gradients
Additional populations (2 Loci)

Supplementary Material:

A: *Evolution of Allele Frequencies and Linkage Disequilibrium under directional selection*

We study parallel phenotypic evolution in two or more descendent populations which were colonized by a single common ancestral population (Figure 1). Within the ancestral population, selection is assumed to favor small values of the phenotype, z; larger values of the phenotype z are favored in the descendent populations. We further assume the phenotype z is determined by the additive action of n diallelic loci with alleles A and a such that

$$z = \bar{z} + \sum_{i=1}^{n} b_i \zeta_i + e_z, \quad \text{(S1)}$$

where $\bar{z}$ is the average phenotype of the population, b_i is the effect of the A allele relative to the a allele, $\zeta_i = (X_i - p_i)$ where X_i is an indicator variable which takes on the value 1 if the individual carries the A_i allele and a value of 0 if it carries the a_i allele, and p_i is the allele frequency of the A allele. The variable e_z describes the random environmental component of the phenotype. For simplicity we assume there is no environment effect and hence $e_z = 0$. Given this simplification, equation (S1) reduces to equation (1) of the main text. Because we assume selection favors a small value of the phenotype, z, in the ancestral population and ignore mutation, allele frequencies, p, within the ancestral population will be near zero. We assume that the approximate values of these allele frequencies are known.

We begin our analysis by focusing on the simplest possible selective scenario capable of generating parallel phenotypic evolution: directional selection of identical strength within each of the descendent populations. Specifically, we assume that selection is linear such that absolute fitness as a function of phenotype is given by

$$W(z) = \beta z + \alpha, \quad \text{(S2)}$$

where β and α describe the slope and intercept of the selection surface respectively. Equation (S2) is the same as equation (2) of the main text. Averaging (S2) over individuals gives us the following expression for the average fitness of a population,

$$\bar{W} = \beta\bar{z} + \alpha.$$

Here $\bar{z}$ is the average phenotype. The relative fitness of an individual with phenotype z is given by the ratio $w(z) = \frac{W(z)}{\bar{W}}$. The resulting expression for relative fitness is simplified by assuming that selection is weak and Taylor expanding about $\beta = 0$

$$w(z) = \frac{\beta z+\alpha}{\beta\bar{z}+\alpha} \approx 1 + \frac{\beta}{\alpha}(z - \bar{z}) + \mathcal{O}(\beta^2). \quad \text{(S3)}$$

Substituting equation (S1) into equation (S3) results in an expression for relative fitness as a function of the individual effect of each locus

$$w(z) \approx 1 + \sum_{i=1}^{n} \frac{\beta}{\alpha} b_i \zeta_i. \quad \text{(S4)}$$

Expression (S4) is very useful as it allows us to determine the strength and direction of selection acting on each locus. To do so, we begin with equation (7) from Kirkpatrick et al. (2002), which gives a general expression for relative fitness

$$w(z) = 1 + \sum_i^n a_i(\zeta_i) + \sum_i^n \sum_{j<i}^n a_{i,j}\left(\zeta_i\zeta_j - D_{ij}\right) + \cdots, \quad \text{(S5)}$$

where D_{ij} is the linkage disequilibrium between the i^{th} and j^{th} locus, a_i is the selection coefficient on the A_i allele, and $a_{i,j}$ is a selection coefficient describing epistatic selection acting on the combination of the A_i and A_j alleles. We can solve for these selection coefficients (the $a's$) by comparing like terms between (S4) and (S5). This reveals that

$$a_i = \frac{\beta}{\alpha} b_i \tag{S6a}$$

$$a_{i,j} = 0. \tag{S6b}$$

With these selection coefficients in hand we can describe the change in the allele frequency of the A_i allele over a single generation using equation (10) from Kirkpatrick et al. (2002),

$$\Delta p_i = a_i p_i (1 - p_i) + \sum_{j \neq i}^{n} a_j D_{ij}, \tag{S7}$$

where D_{ij} is the linkage disequilibrium between loci i and j. To further simplify this expression we assume that recombination is frequent relative to the strength of selection, an assumption which allows the populations to reach a state known as quasi-linkage-equilibrium (QLE). At QLE, the D_{ij} terms are small and of the same order as the selection coefficient, $\mathcal{O}(a_i)$ (Nagylaki, 1993, Nagylaki et al., 1999). This allows us to simplify the change in allele frequency in (S7) to

$$\Delta p_i \approx a_i p_i (1 - p_i) + \mathcal{O}\left(a_i^2\right), \tag{S8}$$

which further simplifies to equation (3) of the main text after substituting the value of a_i from equation (S6a).

Because (S8) shows that evolution of allele frequencies across loci is independent at QLE, we can make use of several classical results derived from the Wright-Fisher model to study the balance between selection and drift within the colonizing populations. Specifically, this can be seen by comparing equation (S8) to the classical expression for the change in allele frequency under the Wright-Fisher model

$$\Delta p = s\, p(1 - p) \tag{S9}$$

(Hartl & Clark, 2007). Comparison of (S8) to (S9) reveals that under linear selection and at QLE $s = \frac{\beta}{\alpha} b_i$. We can use this expression for s to utilize another classical result of the Wright-Fisher model, the probability of fixation of an allele under linear directional selection (Karlin & Taylor, 1981):

$$P_{fix} = \frac{\left(1-e^{2Nsp_o}\right)}{1-e^{2Ns}} = \frac{\left(1-e^{2N\frac{\beta}{\alpha}p_o}\right)}{1-e^{2N\frac{\beta}{\alpha}}}, \quad \text{(S10)}$$

which corresponds to equation (4) and (5) of the main text.

Extension of equation (S10) to the probability of parallel evolution at a single locus (equation 6 of the main text) is straight-forward. Parallel genetic evolution at a single locus in m independent populations requires the independent fixation of the derived allele in the m populations and is therefore simply the m^{th} power of equation (S10). The extension of (S10) to the multi-locus probability of parallel evolution given by equation 7 of the main text is similarly straight-forward. The only additional complications are that the allelic effect size, b_i, and the initial allele frequency, p_0, now require a subscript to denote the locus they refer to. The resulting probability of parallel evolution across multiple loci now represents a product of the single locus result (S10) across loci, denoted by the i subscript.

As mentioned in the main text, equations (6) and (7) represent a very restrictive definition of parallel evolution, requiring repeated adaptation in *all* m descendent populations. Alternative definitions of parallel evolution can however be easily derived from equation 5 of the main text. For example, one plausible alternative definition of parallel evolution is the repeated fixation in *x* of the m descendent populations. The probability of parallel evolution defined in this way is given by

$$P_{\parallel} = \binom{m}{x}\left(\frac{\left(1-e^{2N\frac{\beta}{\alpha}b_i p_{0_i}}\right)}{1-e^{2N\frac{\beta}{\alpha}b_i}}\right)^{x}\left(1-\left(\frac{\left(1-e^{2N\frac{\beta}{\alpha}b_i p_{0_i}}\right)}{1-e^{2N\frac{\beta}{\alpha}b_i}}\right)\right)^{m-x}, \quad \text{(S11)}$$

where $\binom{m}{x}$ is the binomial coefficient and the second term is the probability of parallel evolution in x populations.

B: *Markov Chain Monte Carlo Simulation of Posterior Distributions.*

At the conclusion of the individual based simulations, adaptation in the descendent populations is summarized by the genetic data $\mathcal{D}$, a $2 \times n$ matrix of 1's and 0's indicating the fixation or loss of alleles at the n loci in the two descendent populations. Simulations assume the values for the allelic effect sizes and initial allele frequencies are known and these are used as inputs. We use the following Metropolis-Hastings algorithm to sample the posterior distribution, $p(\eta|\mathcal{D})$ where $\eta = N\frac{\beta}{\alpha}$. The basic algorithm can be described in 7 steps, the first of these steps initializes the algorithm, the 2nd through 4th steps are recursive and generate samples from the posterior (Marjoram et al., 2003), and the 5th and 6th steps address the convergence and termination of the algorithm respectively (see below). Convergence to the true posterior distribution can be computed by simulating multiple independent sequences of points and then comparing the variance between versus within these sequences; we will simulate $M = 5$ sequences to assess convergence. We treated the first half of the sequences as a burn-in period as the values at these points depend primarily on the initial conditions of each sequence and not the posterior distribution (Gelman, 2004).

The Metropolis-Hastings algorithm can be described by the following steps.

Initialization

1. Draw M estimates of η from the prior distribution. These estimates will serve as the starting points for each of the M sequences.

Recursive algorithm: (Repeat for each of the M sequences)

2. From the current estimate, η, propose a move to a new point, η^*, where $\eta^*$is drawn from the jump distribution $J(\eta, \eta^*)$, a distribution describing the probability of moving from η to η^*. The jump distribution used here was a uniform distribution ranging from $\eta \pm 100$.
3. Calculate the probability of accepting the point η^*, h.

$$h = \min\left(1, \frac{P(D|\eta^*)\pi(\eta^*)J(\eta, \eta^*)}{P(D|\eta)\pi(\eta)J(\eta^*, \eta)}\right)$$

4. Move to $\eta^*$with probability h, otherwise stay at η.

Assessing Convergence: (After simulating 2n points in each of the m sequences)

5. Denote these M sequences by $\psi_{i,j}$ where i is the index over points and ranges from 1 to $2n$ and j is the index over the independent sequences and ranges from 1 to M. Discard the first n points in each sequence for "burn in" so that the index i now ranges from 1 to n. Calculate the following expression, which is a weighted comparison of the variance between versus within each sequence:

   $R = \sqrt{\frac{\frac{n-1}{n}W + \frac{1}{n}B}{W}}$,

   where:

   $B = \frac{n}{M-1}\sum_{j=1}^{M}\left(\frac{1}{n}\sum_{i=1}^{n}\psi_{i,j} + \frac{1}{M}\sum_{j=1}^{M}\frac{1}{n}\sum_{i=1}^{n}\psi_{i,j}\right)^2$ and

   $W = \frac{1}{M}\sum_{j=1}^{M}\frac{1}{n-1}\sum_{i=1}^{n}\left(\psi_{i,j} - \frac{1}{n}\sum_{i=1}^{n}\psi_{i,j}\right)^2$

6. If $R < 1$ simulating additional points will not significantly improve the estimation of the mode of the posterior distribution (Gelman, 2004).

After enough points have been simulated to reach a ratio of $R < 1$, we use the last n points in each of the $M = 5$ sequences to generate a single histogram, consisting of 100 bins, which approximates the posterior distribution. The estimate for η, $\hat{\eta}$, was given by the histogram bin containing the most points.

C: *Sensitivity of the Bayesian Estimator to migration from the ancestral population and to migration among descendent populations.*

Our analytical model assumes the descendent populations diverge from their ancestor in perfect allopatry. Although this is likely true in many well-studied cases, in others recurrent gene flow from the ancestral population or among descendent populations may occur. We investigate this scenario by integrating recurrent gene flow into our individual based simulations and then using these simulations to test the accuracy of the estimator. We found that the estimator is fairly robust to the presence of recurrent gene flow occurring at rate ν, remaining quite accurate even when up to $N\nu = 2$ migrants move among populations per generation regardless of the source (ancestral or descendent) of these migrants (Table S4, S5). Simulations with $N\nu = 3$ and $N\nu = 4$ confirm that the estimator is very robust to migration amongst descendent populations but performs poorly under migration from the ancestral population at these high levels. As expected however, migration among descendent populations does increase the frequency of the selectively favored allele within these populations, leading to slight overestimates of the strength of selection when gene flow becomes appreciable, $N\nu = 2$ Table (S5).

D: *Sensitivity of Bayesian Estimator to differing selection gradients in descendent populations.*

An additional assumption of our analytical model and estimator is that the strength of natural selection is identical in the descendent populations. We tested the effect of differing selection gradients among descendent populations with our individual based simulations and then used these simulations to evaluate the accuracy of the estimator. Specifically, we compared our Bayesian estimate of η to the

average value of η actually experienced by the two populations when the selection gradient between the two populations differed by up to 20%. We saw no significant difference between the accuracy of the estimator under differing selection compared to when selection was identical, Table (S6).

E: *Sensitivity of Bayesian Estimator to error in parameters*

Thus far we have discussed the robustness of the estimator to violations in the assumptions of the analytical model; these tests, however, were performed with an assumption of their own. Specifically, the Bayesian estimate of η was found by assuming that initial allele frequencies and effect sizes of the alleles were known without error. Although these values can be estimated in natural systems, estimates will likely be associated with substantial error. To investigate how error in the estimates of these key parameters influenced the accuracy of our estimator we ran individual based simulations where, rather than using the true values for the initial frequencies and effect sizes, we estimated $\hat{\eta}$ using parameter estimates accompanied by a specified amount of error. Specifically, we drew estimated initial allele frequencies and effect sizes from a Gaussian distribution centered around the true value and with a variance of either 0% or 10% of the true value, the estimated initial allele frequencies were assumed to be the same in both descendent populations. The results of these simulations are shown in Table (S7), and reveal that the estimator was relatively insensitive to error in initial allele frequency and only moderately sensitive to error in effect size.

F: References:

Gelman, A. 2004. *Bayesian data analysis*. Chapman & Hall/CRC, Boca Raton, Fla. *pg 289-297.*

Hartl, D. L. and A. G. Clark. 2007. *Principles of population genetics*. Sinauer Associates, Sunderland, Mass.

Karlin, S. and H. M. Taylor. 1981. *A second course in stochastic processes*. Academic Press, New York.

Kirkpatrick, M., T. Johnson, and N. Barton. 2002. General models of multilocus evolution. Genetics 161:1727-1750.

Marjoram, P., J. Molitor, V. Plagnol, and S. Tavare. 2003. Markov chain Monte Carlo without likelihoods. P Natl Acad Sci USA 100:15324-15328.

Nagylaki, T. 1993. The Evolution of Multilocus Systems under Weak Selection. Genetics 134:627-647.

Nagylaki, T., J. Hofbauer, and P. Brunovsky. 1999. Convergence of multilocus systems under weak epistasis or weak selection. J Math Biol 38:103-133.

Tables:

Table S1:

	Candidate Gene Method				QTL Method			
# of loci	**Intercept**	**Slope**	**R^2**	**Ideal R^2**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**
2	15.16692	0.847294	0.798549	0.742167	19.8606	0.780984	0.829471	0.752435
4	4.313604	0.992298	0.789878	0.778282	7.161146	0.945582	0.813414	0.793367
8	0.880424	1.007933	0.849744	0.84871	1.318149	1.015741	0.848798	0.846376

Table S2:

		Candidate Gene Method				QTL Method			
Selection	**# of loci**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**
Linear	2	16.00289	0.778181	0.786338	0.734021	20.77678	0.730456	0.815123	0.740056
Linear	4	4.908612	0.959635	0.781458	0.770902	6.543477	0.934036	0.813979	0.79955
Linear	8	1.566819	1.008783	0.84166	0.839054	3.592519	0.956031	0.853823	0.849199
$\theta > z_{max}$	2	19.04129	0.176478	0.570297	0.334428	23.47881	0.157239	0.613157	0.309989
$\theta > z_{max}$	4	6.471029	0.890184	0.653319	0.612362	7.974566	0.909744	0.686446	0.628065
$\theta > z_{max}$	8	3.919524	1.032816	0.819669	0.800306	5.610145	0.982669	0.823882	0.801448
$\theta < z_{max}$	2	15.89003	-0.04053	0.416807	0.089439	15.75335	-0.12212	0.373713	0.093054
$\theta < z_{max}$	4	6.921948	0.471407	0.474124	0.337294	7.862538	0.445404	0.501893	0.362634
$\theta < z_{max}$	8	1.353306	0.854482	0.544687	0.539096	2.607046	0.897374	0.681847	0.663801

Table S3:

		Candidate Gene Method				QTL Method			
Selection	**# of loci**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**
Linear	2	15.87892	0.775401	0.77533	0.725998	19.81205	0.735114	0.810594	0.747602
Linear	4	6.138736	0.92206	0.798193	0.786955	9.198878	0.861285	0.796594	0.77863
Linear	8	3.654056	0.893438	0.830906	0.827941	4.976391	0.88004	0.840244	0.835328
$\theta > z_{max}$	2	13.61869	0.580798	0.554989	0.376406	14.42015	0.871419	0.587351	0.378834
$\theta > z_{max}$	4	6.962834	1.005779	0.663432	0.59926	8.065374	1.035488	0.692149	0.610299
$\theta > z_{max}$	8	6.550021	0.906787	0.80325	0.7844	7.121709	0.922785	0.816295	0.792451
$\theta < z_{max}$	2	17.98822	-0.40386	0.430633	0.047842	20.97166	0.293755	0.504912	0.05447
$\theta < z_{max}$	4	4.724858	0.706142	0.471843	0.383263	6.840519	0.609338	0.532616	0.393311
$\theta < z_{max}$	8	1.064135	0.991435	0.594402	0.58825	2.302309	0.919698	0.631262	0.614892

Table S4:

			Candidate Gene Method				QTL Method			
Selection	**# of loci**	**# migrants**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**
Linear	2	1	9.462552	1.000405	0.768803	0.72351	12.86251	0.973032	0.808819	0.745215
Linear	4	1	5.85923	0.865046	0.7767	0.771027	7.986257	0.865206	0.791305	0.77893
Linear	8	1	1.388972	1.043136	0.857805	0.852324	2.034207	1.039227	0.87068	0.863085
Linear	2	2	9.651092	0.702307	0.7416	0.733154	12.04375	0.744005	0.788795	0.777064
Linear	4	2	-2.08238	0.868612	0.76941	0.726144	-2.95146	0.941102	0.787093	0.764882
Linear	8	2	-5.43055	0.795074	0.785624	0.494597	-3.55962	0.742378	0.805221	0.535897

Table S5:

			Candidate Gene Method				QTL Method			
Selection	**# of loci**	**# migrants**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**
Linear	2	1	18.87484	0.594862	0.755336	0.702178	20.06695	0.652962	0.786618	0.727683
Linear	4	1	5.095263	1.016475	0.822661	0.804144	9.645867	0.915799	0.823622	0.796352
Linear	8	1	1.580135	1.002535	0.846162	0.843802	2.158628	1.006954	0.850158	0.845805
Linear	2	2	20.29168	0.707668	0.789092	0.704971	24.31354	0.682712	0.827893	0.718936
Linear	4	2	18.34987	1.089859	0.866779	0.704424	21.89371	0.996555	0.869351	0.697563
Linear	8	2	19.42666	1.171849	0.898938	0.705512	20.60433	1.133726	0.900089	0.705511

Table S6:

			Candidate Gene Method				QTL Method			
Selection	**# of loci**	**difference in gradient**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**	**Intercept**	**Slope**	**R^2**	**Ideal R^2**
Linear	2	10%	14.13968	0.80753	0.774469	0.736087	18.18815	0.841374	0.813448	0.741789
Linear	4	10%	5.052786	0.915427	0.787541	0.781258	5.393386	0.947227	0.808594	0.798983
Linear	8	10%	1.634081	0.989398	0.844835	0.843117	3.400554	0.928276	0.853845	0.85118
Linear	2	20%	38.23155	0.091405	0.741292	0.362411	41.03655	0.08503	0.772002	0.333705
Linear	4	20%	4.2148	0.984006	0.777632	0.767716	6.646766	0.932552	0.793882	0.779343
Linear	8	20%	0.17403	1.054042	0.838123	0.835511	1.531522	1.008696	0.837089	0.834684

Table S7:

				Candidate Gene Method				QTL Method			
Selection	# of loci	$Error\ in\ p_0$	$Error\ in\ b$	Intercept	Slope	R^2	Ideal R^2	Intercept	Slope	R^2	Ideal R^2
Linear	2	10%	10%	34.83995	0.102524	0.743151	0.220846	39.56384	0.080656	0.786806	0.183825
Linear	4	10%	10%	5.392651	0.957575	0.793362	0.78166	7.479929	0.938856	0.814904	0.795675
Linear	8	10%	10%	0.34253	1.002631	0.860966	0.860806	1.26521	1.003183	0.85683	0.855334
Linear	2	10%	0%	19.93728	0.664966	0.785625	0.723271	23.69046	0.652236	0.807125	0.724296
Linear	4	10%	0%	7.29575	0.918863	0.790391	0.774034	8.508657	0.907506	0.807808	0.787979
Linear	8	10%	0%	1.90241	0.931618	0.829667	0.828801	4.181844	0.875054	0.841528	0.838672
Linear	2	0%	10%	16.70543	0.709101	0.754913	0.713937	17.49327	0.819431	0.814171	0.750978
Linear	4	0%	10%	6.935046	0.946104	0.796294	0.777997	7.947352	1.002682	0.825405	0.79228
Linear	8	0%	10%	2.06413	0.978143	0.844339	0.842551	3.512649	0.948235	0.847899	0.84454

Table S8:

		Candidate Gene Method				QTL Method			
Selection	# of pops	Intercept	Slope	R^2	Ideal R^2	Intercept	Slope	R^2	Ideal R^2
Linear	2	20.17902	0.704162	0.778937	0.711608	24.19928	0.62949	0.813061	0.731029
Linear	4	7.330957	0.92165	0.782471	0.763815	9.809639	0.983656	0.82849	0.786108
Linear	8	3.548657	0.985801	0.835579	0.827861	4.840564	0.973264	0.857317	0.846613
$\theta > z_{max}$	2	20.66103	-0.02662	0.597459	0.338874	20.21517	0.145328	0.539055	0.313713
$\theta > z_{max}$	4	9.066178	0.446497	0.47196	0.37569	7.222451	0.998224	0.585238	0.473687
$\theta > z_{max}$	8	1.706663	0.8564	0.46868	0.462201	0.693301	1.01186	0.630704	0.627024
$\theta < z_{max}$	2	18.89695	-0.59417	0.443024	0.071341	22.44341	-1.60471	0.460389	0.065313
$\theta < z_{max}$	4	5.255587	0.617513	0.288124	0.136686	5.110795	1.482254	0.402424	0.234122
$\theta < z_{max}$	8	2.001947	0.484536	0.087474	0.060867	1.274492	0.6845	0.124804	0.114393

Table Legends:

Table S1: Wright-Fisher Simulations. The estimate $\hat{\eta}$ is calculated from data simulated using the Wright-Fisher model with $Ns = \eta$. The table records the slope, intercept and R^2 values for the regression fit for the simulated data as well as the R^2 values for the data fit to the ideal $\hat{\eta} = \eta$ line. Simulations were performed for between 2 and 8 loci and estimates for the strength of selection were made using the data from all loci, as well as data filtered to resemble the two experimental methods, QTL and candidate gene.

Table S2: Evaluating accuracy across different forms of selection. The table shows the slope, intercept, and R^2 values for the linear-regression fit of predicted values of $\hat{\eta}$ to the true value in the simulation η, and the R^2 values for the line of perfect fit $\hat{\eta} = \eta$. Estimates for the strength of selection were made using the data from all loci, as well as data filtered to resemble the two experimental methods, QTL and

candidate gene. Accuracy was evaluated for 2, 4, and 8 loci and three forms of natural selection; linear directional, stabilizing with $\theta < z_{max}$, and stabilizing with $\theta > z_{max}$. The loci underwent free recombination, $r = 0.5$.

Table S3: Variable forms of selection with constrained recombination. The table shows the slope, intercept, and R^2 values for the linear-regression fit of predicted values of $\hat{\eta}$ to the true value in the simulation η, and the R^2 values for the line of perfect fit $\hat{\eta} = \eta$. Estimates for the strength of selection were made using the data from all loci, as well as data filtered to resemble the two experimental methods, QTL and candidate gene. Accuracy was evaluated for 2, 4, and 8 loci and three forms of natural selection; linear directional, stabilizing with $\theta < z_{max}$, and stabilizing with $\theta > z_{max}$. In contrast to table S2 recombinaton was constrained at a rate of $r = 0.05$

Table S4: Estimation of η with recurrent gene flow from the ancestral population. The table shows the slope, intercept, and R^2 values for the linear-regression fit of predicted values of $\hat{\eta}$ to the true value in the simulation η, and the R^2 values for the line of perfect fit $\hat{\eta} = \eta$. Estimates for the strength of selection were made using the data from all loci, as well as data filtered to resemble the two experimental methods, QTL and candidate gene. Accuracy was evaluated for 2, 4, and 8 loci under linear selection. Each generation each descedent population received a set number of migrants (1 or 2) from the ancestral population.

Table S5: Estimation of η with recurrent gene flow between the descendent populations. The table shows the slope, intercept, and R^2 values for the linear-regression fit of predicted values of $\hat{\eta}$ to the true value in the simulation η, and the R^2 values for the line of perfect fit $\hat{\eta} = \eta$. Estimates for the strength

of selection were made using the data from all loci, as well as data filtered to resemble the two experimental methods, QTL and candidate gene. Accuracy was evaluated for 2, 4, and 8 loci under linear selection. Each generation each descedent population received a set number of migrants (1 or 2) from the other descendent population.

Table S6: Estimation of η with differing selection in descendent populations. The table shows the slope, intercept, and R^2 values for the linear-regression fit of predicted values of $\hat{\eta}$ to the true value in the simulation η, and the R^2 values for the line of perfect fit $\hat{\eta} = \eta$. Estimates for the strength of selection were made using the data from all loci, as well as data filtered to resemble the two experimental methods, QTL and candidate gene. Accuracy was evaluated for 2, 4, and 8 under linear selection. The strength of linear selection was allowed to vary up to 10% between the two populations.

Table S7: Estimation of η with error in p_0 and b. The table shows the slope, intercept, and R^2 values for the linear-regression fit of predicted values of $\hat{\eta}$ to the true value in the simulation η, and the R^2 values for the line of perfect fit $\hat{\eta} = \eta$. Estimates for the strength of selection were made using the data from all loci, as well as data filtered to resemble the two experimental methods, QTL and candidate gene. Accuracy was evaluated for 2, 4, and 8 loci under linear selection. The estemated values of $\hat{\eta}$ where calculated using initial allele requencies, p_0, and alellic effect sizes, b, that differed from their true values.

Table S8: Estimation of η with 2, 4, and 8 populations. The table shows the slope, intercept, and R^2 values for the linear-regression fit of predicted values of $\hat{\eta}$ to the true value in the simulation η, and the R^2 values for the line of perfect fit $\hat{\eta} = \eta$. All simulations where run with $n = 2$ loci and a variable

number of populations, $m = 2,4, or\ 8$. Estimates for the strength of selection were made using the data from 2 loci, as well as data filtered to resemble the two experimental methods, QTL and candidate gene.

Figure Legends:

Figure S1: Ideal R^2 values across robustness tests. Ideal R^2 values for QTL method (blue) and candidate gene method (red) estimates using data from 2 (circle), 4 (square), or 8 (triagnle) loci across different robustness tests, except for the "additional populations**"** points which are all simulated with 2 loci but with variable number of popualtions, 2 (circle), 4 (square), or 8 (triagnle) popualtions respectivley. Wright Fisher (WF) simulations represent an ideal case. Explinations of different robustness tests are given in main text as well, supplimentary sections C,D, and E, as well as table legends.

References

Gelman, A. 2004. *Bayesian data analysis,* 2nd ed. Chapman & Hall/CRC, Boca Raton, Fla.

Hartl, D. L. & Clark, A. G. 2007. *Principles of population genetics,* 4th ed. Sinauer Associates, Sunderland, Mass.

Karlin, S. & Taylor, H. M. 1981. *A second course in stochastic processes*. Academic Press, New York.

Kirkpatrick, M., Johnson, T. & Barton, N. 2002. General models of multilocus evolution. *Genetics* **161**: 1727-1750.

Marjoram, P., Molitor, J., Plagnol, V. & Tavare, S. 2003. Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 15324-15328.

Nagylaki, T. 1993. The Evolution of Multilocus Systems under Weak Selection. *Genetics* **134**: 627-647.

Nagylaki, T., Hofbauer, J. & Brunovsky, P. 1999. Convergence of multilocus systems under weak epistasis or weak selection. *Journal of Mathematical Biology* **38**: 103-133.