JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.13006

The probability of parallel genetic evolution from standing genetic variation

A. MACPHERSON*† & S. L. NUISMER*‡

*Program of Bioinformatics and Computational Biology, University of Idaho, Moscow, ID, USA
†Department of Zoology, University of British Columbia, Vancouver, BC, Canada
‡Department of Biological Sciences, University of Idaho, Moscow, ID, USA

Keywords:

adaptation;
Bayesian;
ecological genetics;
ecological selection;
genetic architecture;
QTL.

Abstract

Parallel evolution is often assumed to result from repeated adaptation to novel, yet ecologically similar, environments. Here, we develop and analyse 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 is that existing studies of parallel genetic evolution frequently rely on data that is insufficient for distinguishing between adaptive evolution and neutral evolution driven by random genetic drift. Overcoming this challenge will require sampling more populations and the inclusion of larger numbers of loci.

Introduction

As the availability of genome sequences has increased, interest in understanding how genomic architecture shapes adaptation at both the genetic and phenotypic levels 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 (Wichman *et al.*, 1999; Qi *et al.*, 2016) 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; Martin & Orgogozo, 2013; Stern, 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 following 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.*, 2016). Here, we focus on parallel

Correspondence: Ailene MacPherson, Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.
Tel.: 208 301 7975; e-mail: amacp@zoology.ubc.ca

genetic evolution defined as the repeated fixation of identical alleles in multiple descendent populations. We further restrict our study to cases where parallel genetic evolution results from standing genetic variation rather than from new mutations. Our interest in this specific scenario is motivated by a number of biological systems where adaptation to a novel environment is thought to result from standing genetic variation present in the ancestral population (Colosimo *et al.*, 2005; Hoekstra *et al.*, 2006; Steiner *et al.*, 2007). In contrast to *de novo* mutation, adaptation from standing genetic variation is likely rapid (Barrett & Schluter, 2008) and may lead to distinct genomic signatures of parallel adaptation (Roesti *et al.*, 2014).

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

Although understanding the contribution of new mutations to parallel evolution is inarguably important, in some systems it may be more relevant to understand the likelihood of parallel evolution from standing genetic variation. For instance, in the stickleback, *Gasterosteus aculeatus*, repeated adaptation to freshwater is thought to involve genes already segregating at low frequencies within the 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 and 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 multilocus framework for predicting the probability of parallel evolution from standing genetic variation. We have two specific goals: first, we will predict the probability of parallel evolution in terms of quantities that are regularly measured in natural populations using a multilocus model of parallel genetic adaptation that assumes weak selection and rapid recombination. Second, we will develop a statistical framework for estimating the historical average strength of parallel selection by coupling our multilocus model to routinely collected genetic data.

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 Fig. 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 resulting in reduced body armour in multiple freshwater stickleback populations relative to 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 Fig. 1B) are crossed with ancestral individuals and the resulting offspring are scanned for divergent QTLs influencing the focal trait. The remaining populations (descendent population 2 in Fig. 1b) are then tested for the candidate genes using a variety of approaches such as genetic complementation tests (Hartl & Jones, 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 (Tishkoff *et al.*, 2007; Enattah *et al.*, 2008; Ingram *et al.*, 2009). Alternatively, in the second design (shown in Fig. 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; Sundin *et al.*, 2005; Nichols *et al.*, 2007).

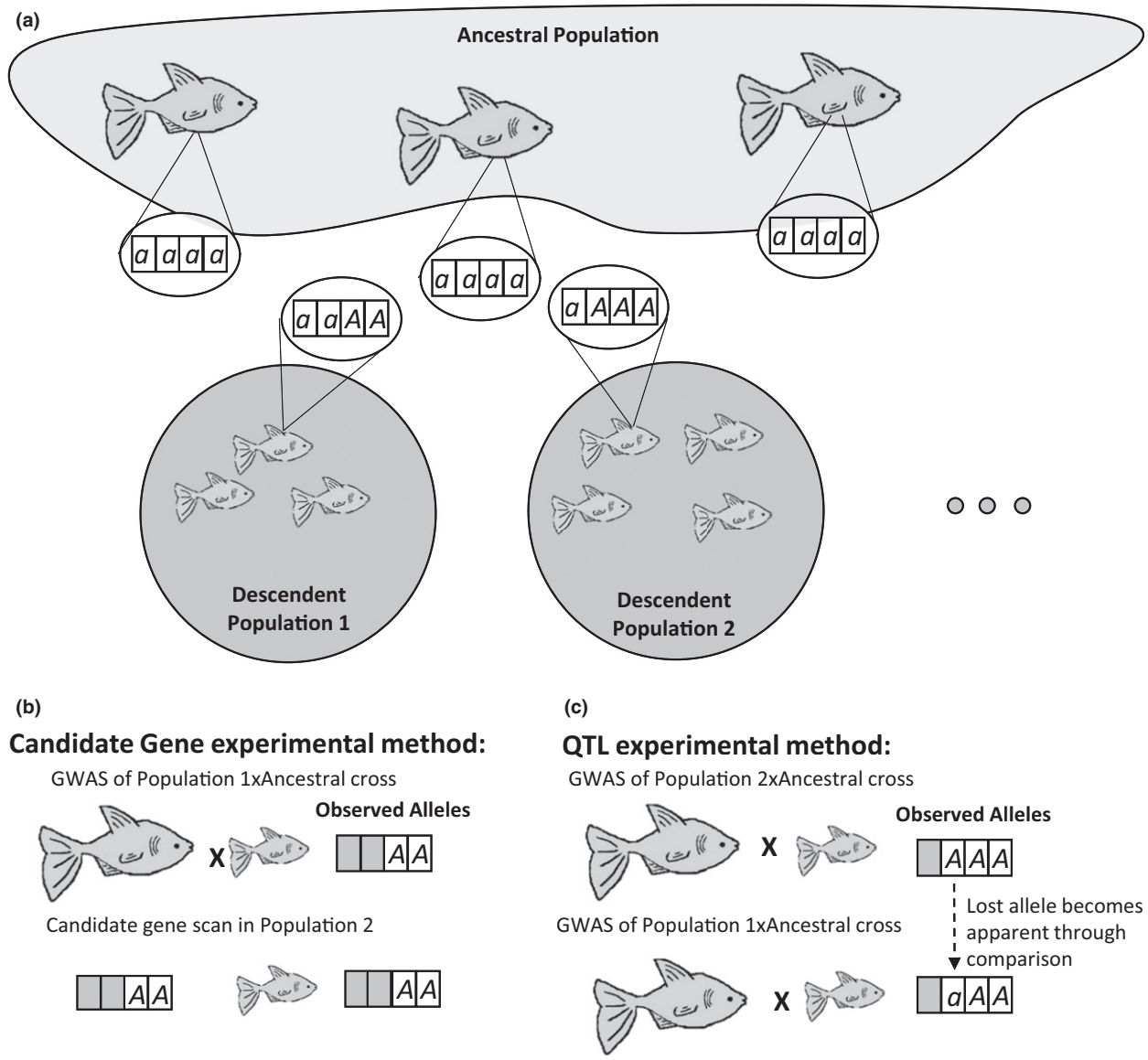

Fig. 1 Schematic of biological scenario. Panel (a) depicts two descendent populations diverging in parallel from a common ancestral population. The *a* allele predominates at all four loci in the ancestral population, whereas the *A* allele fixes at various loci in the two descendent populations. Panels (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 genomewide 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 genomewide scans, one in each population. Compared with the candidate gene method, the QTL method uncovers an additional locus driving divergence in population 2.

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 average 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, favouring an increase in frequency of the A alleles.

The biggest challenge to modelling 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:

$$W(z) = \beta z + \alpha. \tag{2}$$

and thus defined only by its intercept (α) and slope (β). Second, we assume the strength of linear directional selection, β, is weak, and that 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, and of the same order as β (Nagylaki, 1993, Nagylaki *et al.*, 1999). Using the expression for the phenotypic trait z, given in eqn (1), as well as the expression for fitness, given by eqn (2), we can use the multilocus 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

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

(see Data S1 for a full derivation). Because we have assumed linear selection and that the population is at quasi-linkage equilibrium (QLE), eqn (3) does not depend on linkage disequilibrium or the frequencies of alleles at other loci; instead, each locus evolves 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 can be approximated as

$$P_{\text{fix}} = \frac{(1 - e^{2Nsp_0})}{1 - e^{2Ns}} \tag{4}$$

(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, strength of selection acting on locus i is $s = \frac{\beta}{\alpha} b_i$, and eqn (4) can be rewritten as

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

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 analysing 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 using eqn (5) to find the probability that at the locus of interest, i, the A_i allele fixes independently in each of m populations:

$$P_{\parallel} = (p_{\text{fix}}(i))^m. \tag{6}$$

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 Data S1.

Equations (5 and 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 Fig. 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 Fig. 2. As this term arises repeatedly in the derivation to follow, we will denote it with the composite parameter η

$$\eta = N\frac{\beta}{\alpha}. \tag{7}$$

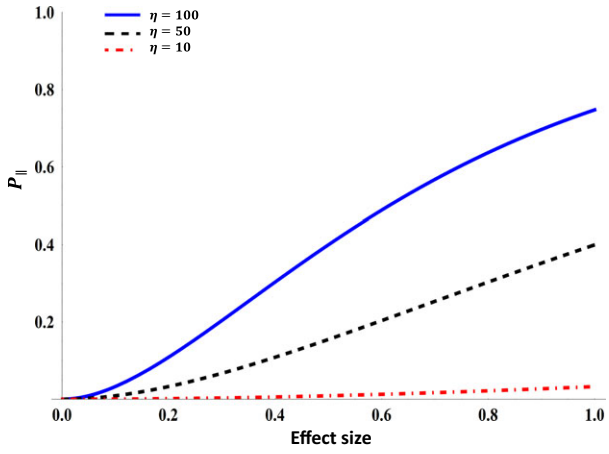

Fig. 2 The probability of parallel evolution as a function of allelic effect size, b. For a given strength of selection the probability of fixation, and hence parallel evolution, increases with allelic effect size. The rate of increase is nonlinear and depends on the strength of selection s and the population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

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:

$$P_{\|} = \prod_{i=1}^{n} (p_{\mathrm{fix}}(i))^{m}, \tag{8}$$

where the product is carried over the number of loci. Not surprisingly, eqn (8) shows that the factors enhancing the probability of parallel evolution at a single locus (e.g. large population size, strong selection) 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 favoured 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 eqn (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 (Fig. 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 (Fig. 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 Fig. 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 towards small effect loci (Fig. 3). Together, these results suggest that the likelihood of observing parallel genetic evolution at any particular number of loci depends heavily on 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 eqn (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 (Fig. 1a). Using eqn (8), we can develop a likelihood function specifying the probability of observing the data, $\mathcal{D}$, given a particular value of 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 (Lynch & Walsh, 1998; Broman & Sen, 2009; Conte *et al.*, 2015), whereas initial allele frequencies can be estimated by measuring the allele frequencies in the ancestral population. Accurate

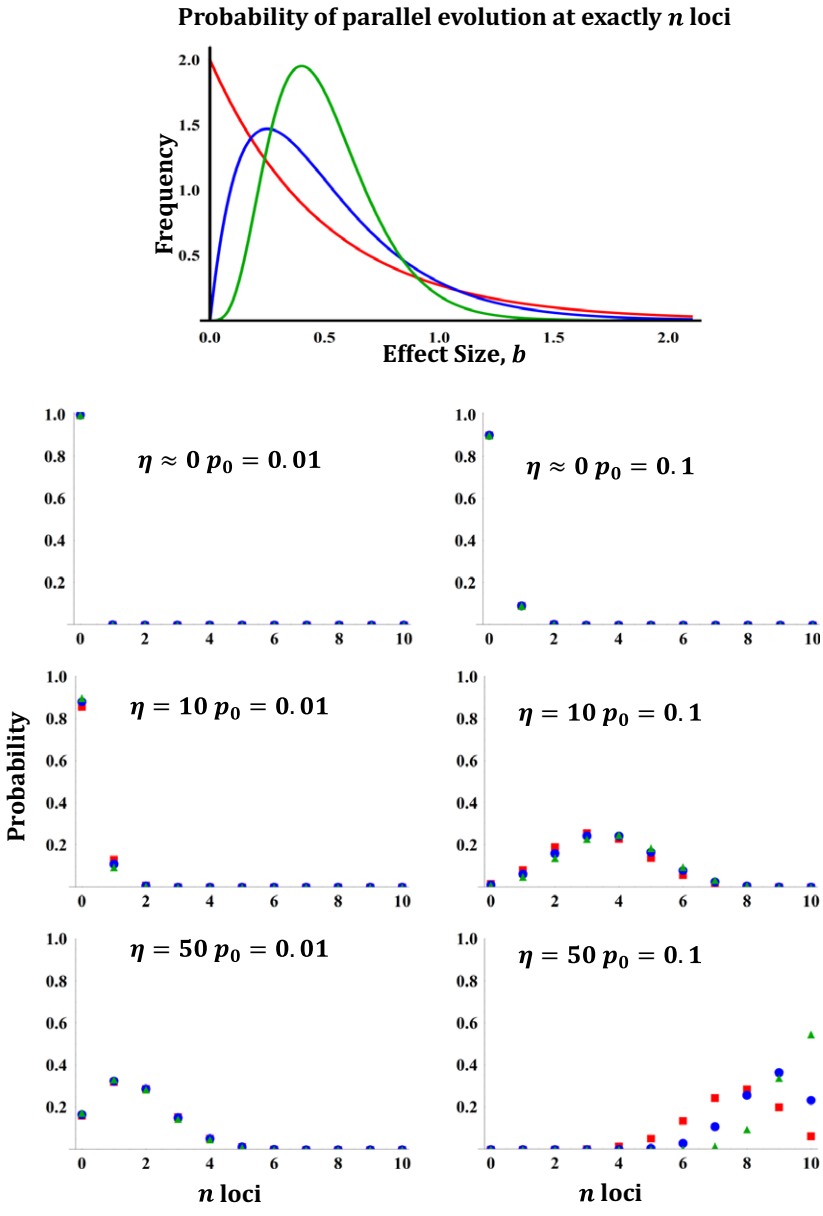

Fig. 3 The probability of parallel evolution at n loci. For a trait determined by the effects of 10 total loci, the probability of observing parallel evolution at exactly n loci depends on the strength of selection, which varies from near 0 to a value of $\eta = 50$, and the initial allele frequency, 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 ($\mu = \frac{1}{2}$).

estimation 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 eqn (5). Alternatively, if the A allele is lost, it contributes a term (1 - P_{fix}). Thus, for m populations and n loci, the likelihood of observing the data, $\mathcal{D}$, is given by the following product

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

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 principles similar to those developed by Rice & 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, $\mathcal{D}$,

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

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 analyse its performance under the assumptions of the analytical model by generating the genetic data $\mathcal{D}$ under 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 multilocus 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 eqn (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 and 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 Fig. 1). The first of these methods, the candidate gene method (Fig. 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. As 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 (Fig. 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 Data S1. For the prior $\pi(\eta)$, we used a uniform distribution on the interval $\eta = \pm 80$. To analyse 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. These simulations allow us to evaluate the consequences of violating key assumptions of our analytical model such as the weak selection and frequent recombination required for our quasi-linkage equilibrium approximation.

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 eqn (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 preselection 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 1 or 0. 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 value of η using the original outcome $\mathcal{D}$ as well as the two modified forms of $\mathcal{D}$ (see Data S1).

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

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

where θ is the phenotypic optimum and γ is the strength of stabilizing selection. Including simulations where selection is stabilizing is important because it relaxes our previous assumption that loci evolve independently. Stabilizing selection is particularly useful in testing this assumption because the relative importance of interdependence between loci can be manipulated by changing the value of the phenotypic optimum. Specifically, the relative importance of interactions between loci will depend on the value of the optimum relative to the largest possible phenotype $z_{\max} = \Sigma_i b_i$. When θ is greater than the largest possible phenotype, $z_{\max}$, loci remain relatively independent as directional selection predominates over epistatic selection. However, when $\theta < z_{\max}$, this is no longer true as epistatic selection now dominates. Therefore, when $\theta > z_{\max}$, evolution is much more likely to resemble linear selection as our analytical model assumes. We simulated stabilizing selection under these two different scenarios, by either requiring that θ be larger than $z_{\max}$ or slightly smaller than $z_{\max}$ (see Data S1). Under stabilizing selection η changes as the population adapts, decreasing as the population approaches the optimum (Chevin & Hospital, 2008; Matuszewski *et al.*, 2015). Therefore, we computed a 'realized' strength of linear selection by averaging the selection gradient, $\frac{\mathrm{COV}(z,w)}{\mathrm{var}(z)}$, over all time points for which $\mathrm{var}(z) \neq 0$.

As expected, analysis of simulated data shows that the accuracy of our estimates depends on the form of selection. Specifically, estimates for η are most accurate under linear selection, somewhat less accurate under stabilizing selection towards a distant optimum, $\theta > z_{\max}$, and least accurate under stabilizing selection towards a close optimum, $\theta < z_{\max}$ (see Fig. 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 Fig. 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 as 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 descendent populations and estimates of the parameters p_0 and b_i that are imprecise (see Data S1 and Tables 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 two migrants per generation. In addition, estimates for the average η in the descendent populations remain robust even when the selection gradients differ among descendent populations by up to 20%. Lastly, estimates for η are relatively insensitive to modest levels of error (<10%) in the estimated values of the parameters p_0 and b_i. 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, however, 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 vs. closed circles in Fig. 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.

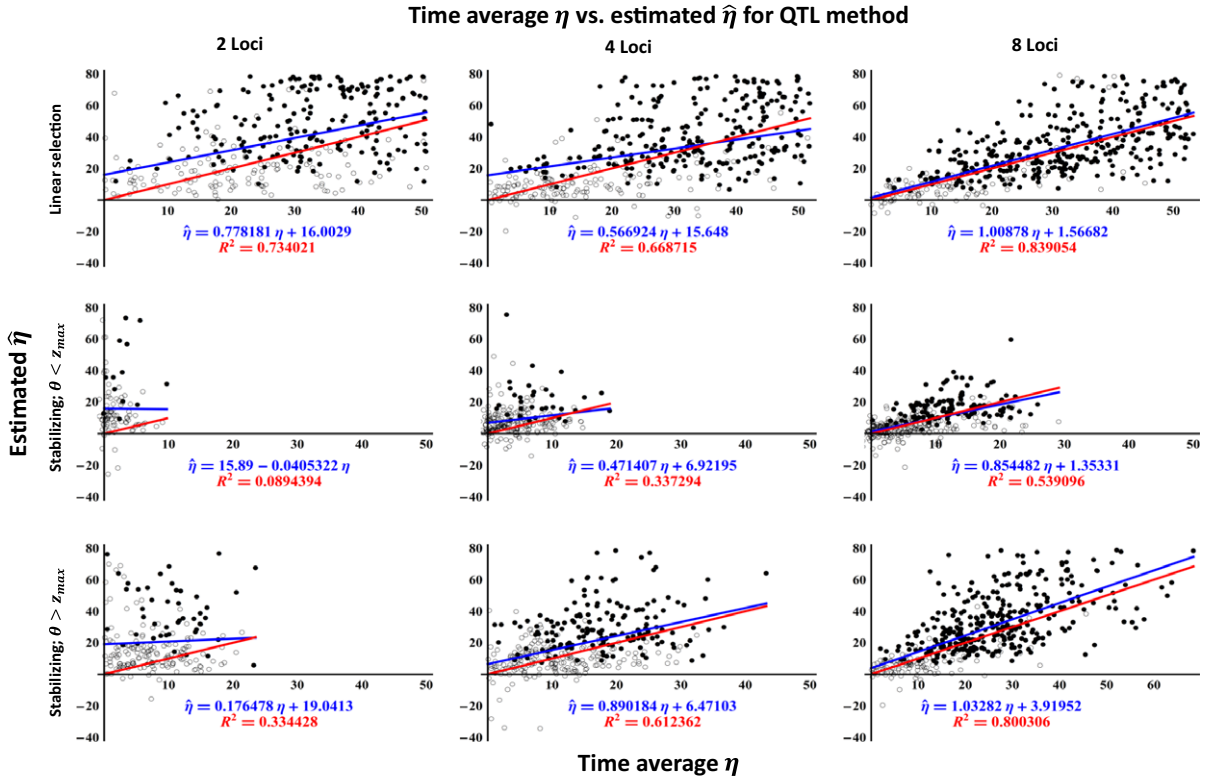

Fig. 4 Regression fit of IBS data under the three forms of selection. Data and linear regression fit (blue), and perfect fit (red) between the time averaged values of η and the Bayesian estimate $\hat{\eta}$ for 200 replicates of the individual-based simulation. Open (filled) points indicate estimates where the 95% credible interval does (not) overlap $\eta = 0$. Subpanels 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.

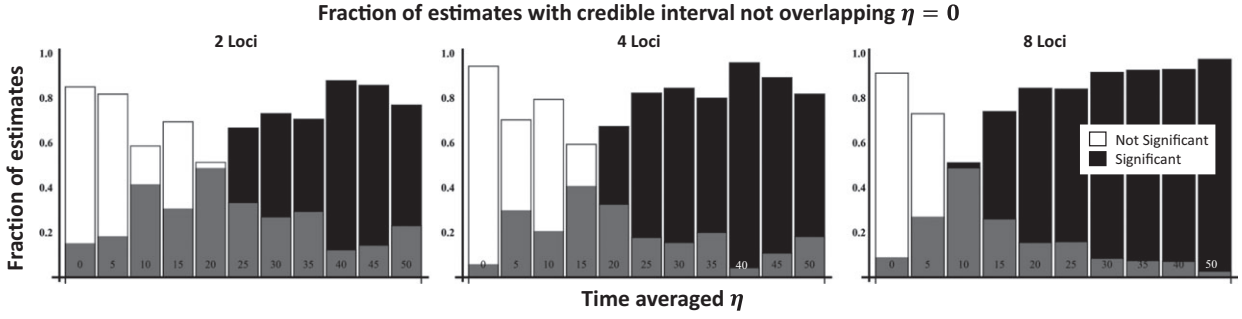

Fig. 5 Significance of $\hat{\eta}$ estimates. The fraction of estimates that differ significantly (black) and do not differ significantly (white) from $\hat{\eta} = 0$ with 95% confidence across the range of time average η values. Overlapping portions are shown in grey.

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 multilocus model of parallel evolution from standing genetic variation to further 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 used a quasi-linkage equilibrium approximation to extend 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 if 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 are 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 freshwater descendent populations of threespine 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 eight descendent populations (Colosimo *et al.*, 2004; Schluter *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 eqn (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 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 multilocus 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.

For our Bayesian approach to be broadly useful, it must produce reliable estimates across a broad range of parameter space. For this reason, we used individual-based simulations to assess the accuracy and robustness of our approach when key assumptions of the underlying model are violated. These simulations demonstrate that our estimator is indeed both accurate and robust, although there are limitations. For example, accurate estimation requires data from at least eight total loci, be that four loci in two populations, two loci in four populations, or some intermediate combination. Whether data are 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 two and six 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. Because increasing the number of loci at which parallel genetic evolution is assessed will almost certainly require studying loci with smaller phenotypic effects on the focal trait, there may be hard limits to the number of loci that can be usefully included. Fortunately, no such limitation exists with respect to the number of populations that can be included.

The demonstration that increasing the number of loci at which parallel evolution is assessed is important

suggests that some experimental approaches may perform better than others. For instance, we explored two alternative experimental methods (see Fig. 1b,c) that differ predictably in the number of loci that are detected: the QTL method and the candidate gene method. 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. Finally, our results show that the accuracy of the Bayesian estimate is influenced not only by the amount of available data but by the accuracy of the estimated parameter values b_i and p_{0_i}. Fortunately, however, as long as error in these parameters remains modest, our Bayesian approach continues to produce reliable estimates (see Table S1). Alternatively, because the approach we have developed is Bayesian, it is also possible to input *a prior* distribution for the parameter p_{0_i} rather than a single-point estimate in cases where the exact value of p_{0_i} is in doubt due to sampling error or stochastic variation in small populations (Hermisson & Pennings, 2005).

In addition to requiring information on parallel genetic evolution drawn from a reasonably large number of loci or populations, our estimator relies on several key assumptions that affect its accuracy. The most important of these assumptions is that selection is weak and linear. When combined with the assumption that recombination is frequent, these key assumptions allowed us to utilize a quasi-linkage equilibrium approximation. Another potentially important assumption of our approach is that population size is constant across time. As a result, we implicitly ignore potentially important impacts of sporadic population bottlenecks or 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 Table S1).

Combined, our analyses of single- and multilocus 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. We have overcome this hurdle by developing a robust statistical methodology for translating observed levels of genetic parallelism into an estimate of the product of phenotypic selection and local population size. This statistical approach provides a much needed tool for distinguishing between adaptive and nonadaptive hypotheses for observed levels of parallel genetic evolution. Applying this method to existing and emerging data from multiple populations with common ancestry may thus offers novel insights into the importance of adaptive evolution in natural populations.

Acknowledgments

We thank Paul Joyce, Richard Gomulkiewicz, Lyudmila Barannyk and Paul Hohenlohe for helpful suggestions and discussions on this work. Funding was provided by NSF grants DEB 1118947 and DEB 1450653 to SLN and BCB fellowships to AM through the University of Idaho's IBEST program.

References

Amos, W. & Hoffman, J.I. 2010. Evidence that two main bottleneck events shaped modern human genetic diversity. *Proc. R. Soc. B Biol. Sci.* **277**: 131–137.

Anderson, J.B., Sirjusingh, C., Parsons, A.B., Boone, C., Wickens, C., Cowen, L.E. *et al.* 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.

Barton, N.H. & Turelli, M. 1991. Natural and sexual selection on many loci. *Genetics* **127**: 229–255.

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., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. *et al.* 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol.* **2**: 635–641.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J. *et al.* 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science* **307**: 1928–1933.

Conte, G.L., Arnegard, M.E., Peichel, C.L. & Schluter, D. 2012. The probability of genetic parallelism and convergence in natural populations. *Proc. R. Soc. B Biol. Sci.* **279**: 5039–5047.

Conte, G.L., Arnegard, M.E., Best, J., Chan, Y.F., Jones, F.C., Kingsley, D.M. *et al.* 2015. Extent of QTL reuse during repeated phenotypic divergence of sympatric threespine stickleback. *Genetics* **201**: 1189–1200.

Enattah, N.S., Jensen, T.G.K., Nielsen, M., Lewinski, R., Kuokkanen, M., Rasinpera, H. *et al.* 2008. Independent introduction of two lactase-persistence alleles into human populations reflects different history of adaptation to milk culture. *Am. J. Hum. Genet.* **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 Genet.* **5** e1000326. doi: 10.1371/journal.pgen.1000326.

Hartl, D.J. & Jones, E.W. 2005. *Genetics: Analysis of Genes and Genomes*. Jones & Bartlett Learning, Burlington, MA.

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 Genet.* **6**: e1000862. doi: 10.1371/journal.pgen.1000862.

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. *Hum. Genet.* **124**: 579–591.

Karlin, S. & Taylor, H.M. 1981. *A Second Course in Stochastic Processes*. Academic Press, New York, NY.

Kimura, M. 1957. Some problems of stochastic-processes in genetics. *Ann. Math. Stat.* **28**: 882–901.

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

Lenormand, T., Chevin, L.M. & Bataillon, T.M. 2016. Parallel evolution: what does it (not) tell us and why is it (still) interesting? In: *Chance in Evolution*, (G.P. Ramsey & C.H. Pence, eds.), pp. 201–221. Univ. Chicago Press, Chicago, IL.

Lynch, M. & Walsh, B. 1998. *Genetics and Analysis of Quantitative Traits*. Sinauer, Sunderland, MA.

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**: 1255–1274.

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

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. *J. Math. Biol.* **38**: 103–133.

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 mykiss*. *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. *et al.* 2006. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nat. Genet.* **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*. *Proc. R. Soc. B* **283**: 20152452. doi: 10.1098/rspb.2015.2452

Ralph, P.L. & Coop, G. 2015. The role of standing variation in geographic convergent adaptation. *Am. Nat.* **186**: S5–S23.

Rice, D.P. & Townsend, J.P. 2012. A test for selection employing quantitative trait locus and mutation accumulation data. *Genetics* **190**: 1533–1545.

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*). *J. Hered.* **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. *Mol. Ecol.* **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. *Am. Nat.* **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. *Hum. Genet.* **118**: 276–286.

Stapley, J., Reger, J., Feulner, P.G.D., Smadja, C., Galindo, J., Ekblom, R. *et al.* 2010. Adaptation genomics: the next generation. *Trends Ecol. Evol.* **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 Biol.* **5**: 1880–1889.

Stern, D.L. 2013. The genetic causes of convergent evolution. *Nat. Rev. Genet.* **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. *et al.* 2007. Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* **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.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Figure S1 Ideal R^2 values across robustness tests.

Data S1 Model derivation, MCMC algorithm, supplementary tables, and figure.

Table S1 Wright-Fisher Simulations.

Table S2 Evaluating accuracy across different forms of selection.

Table S3 Variable forms of selection with constrained recombination.

Table S4 Estimation of η with recurrent gene flow from the ancestral population.

Table S5 Estimation of η with recurrent gene flow between the descendent populations.

Table S6 Estimation of η with differing selection in descendent populations.

Table S7 Estimation of η with error in p_0 and b.

Table S8 Estimation of η with 2, 4, and 8 populations.

Received 25 February 2016; revised 11 October 2016; accepted 18 October 2016