

A population-genomic approach for estimating selection on polygenic traits in heterogeneous environments

Zachariah Gompert^{1,2} 

¹Department of Biology, Utah State University, Logan, UT, USA

²Ecology Center, Utah State University, Logan, UT, USA

Correspondence

Zachariah Gompert, Department of Biology, 5305 Old Main Hill, Utah State University, Logan, UT 84322-5305, USA.
Email: zach.gompert@usu.edu

Funding information

Division of Environmental Biology, Grant/Award Number: DEB-1844941 and DEB-1638768

Abstract

Strong selection can cause rapid evolutionary change, but temporal fluctuations in the form, direction and intensity of selection can limit net evolutionary change over longer time periods. Fluctuating selection could affect molecular diversity levels and the evolution of plasticity and ecological specialization. Nonetheless, this phenomenon remains understudied, in part because of analytical limitations and the general difficulty of detecting selection that does not occur in a consistent manner. Herein, I fill this analytical gap by presenting an approximate Bayesian computation (ABC) method to detect and quantify fluctuating selection on polygenic traits from population genomic time-series data. I propose a model for environment-dependent phenotypic selection. The evolutionary genetic consequences of selection are then modelled based on a genotype–phenotype map. Using simulations, I show that the proposed method generates accurate and precise estimates of selection when the generative model for the data is similar to the model assumed by the method. The performance of the method when applied to an evolve-and-resequence study of host adaptation in the cowpea seed beetle (*Callosobruchus maculatus*) was more idiosyncratic and depended on specific analytical choices. Despite some limitations, these results suggest the proposed method provides a powerful approach to connect the causes of (variable) selection to traits and genome-wide patterns of evolution. Documentation and open-source computer software (fsABC) implementing this method are available from GITHUB (<https://github.com/zgompert/fsabc.git>).

KEYWORDS

approximate Bayesian computation, *Callosobruchus maculatus*, computational statistics, ecological genetics, fluctuating selection, polygenic traits

1 | INTRODUCTION

Selection can cause rapid evolutionary change on ecological time scales (Bell, 2010; Ford, 1977; Messer et al., 2016; Mueller et al., 2020; Reznick et al., 1997; Thompson, 2013). The rates of evolution are lower when measured on longer time scales (Gingerich, 2019; Hendry & Kinnison, 1999; Kinnison & Hendry, 2001), in part because the form, direction and intensity of selection often vary through time

(e.g. from generation to generation; Bell, 2010; Siepielski et al., 2009). Various factors can cause selection to fluctuate in time, including climatic variability (Bergland et al., 2014; Grant & Grant, 2002, 2014; Siepielski et al., 2017), host–parasite or host–pathogen coevolution (Ford et al., 2017; Gómez & Buckling, 2011; Hall et al., 2011), and mating or survival advantages for rare phenotypes (Endler, 1988; Hughes et al., 2013; Nosil et al., 2018; Takahashi et al., 2010). Similar factors can cause selection to vary in space (Kingsolver et al., 2001; Siepielski

et al., 2013), resulting in local adaptation (Endler, 1977). Fluctuating selection may help explain fundamental biological phenomena, such as the maintenance of polymorphism (Ford, 1977; Hedrick, 2006; Turelli et al., 2001; Wittmann et al., 2017), molecular diversity levels (Coop, 2016; Gillespie, 1991; Hahn, 2008; Leffler et al., 2012), ecological specialization (Agrawal, 2020; Agrawal et al., 2010; Anderson et al., 2013; Gompert & Messina, 2016) and the evolution of plasticity (Hallsson & Björklund, 2012; King & Hadfield, 2019; Tufto, 2015). Quantitative studies of the population genetic consequences of fluctuating selection in the wild are necessary to better evaluate this possibility, but remain relatively rare (Messer et al., 2016) (but see, e.g., Bergland et al., 2014; Ford, 1977; Mueller et al., 1985).

Such studies have been hampered, in part, by limited development of appropriate statistical methods (Messer et al., 2016). Many popular statistical methods infer selection on genetic variants from static population genomic patterns (i.e. patterns lacking a temporal component), such as haplotype structure and diversity (Sabeti et al., 2002; Szpiech & Hernandez, 2014), genetic differentiation between populations (Beaumont & Nichols, 1996; Foll & Gaggiotti, 2008; Nosil et al., 2008), the inferred distribution of coalescence events (Edge & Coop, 2019; Speidel et al., 2019), genotype–environment associations (Coop et al., 2010; De Villemereuil & Gaggiotti, 2015; Rellstab et al., 2015) or population genetic patterns for trait-associated markers (Berg & Coop, 2014; Field et al., 2016; Racimo et al., 2018; Robinson et al., 2015). These methods can detect selection when it occurs in a consistent manner over time, but were not designed to identify temporally fluctuating selection. Methods (e.g. Buffalo & Coop, 2019, 2020; Feder et al., 2014; Gompert, 2016; Illingworth & Mustonen, 2011; Kelly & Hughes, 2019; Mathieson & McVean, 2013), and associated data sets (e.g. Bi et al., 2013, 2019; Rêgo et al., 2019; Reich, 2018), with temporal sampling in natural or experimental populations, are better suited for detecting fluctuating selection and are becoming more common.

Mueller et al. (1985) proposed one of the first explicit tests for fluctuating selection, which was based on the proposition that genetic loci evolving in response to the same (unknown) environmental variations should exhibit correlated patterns of allele frequency change over time. Using this method, Mueller et al. (1985) found ample evidence of fluctuating selection on allozyme loci in natural populations of two *Drosophila* species. More recently, Bergland et al. (2014) found evidence of genomically widespread fluctuating selection in *D. melanogaster* based on repeated seasonal oscillations in the frequency of multiple SNP markers. Gompert (2016) introduced a method to quantify fluctuating selection on genetic loci from correlations between patterns of allele frequency change and the state of the environment. This requires an explicit hypothesis about the environmental factor causing selection to fluctuate. Buffalo and Coop (2019) derived a method to quantify the extent and consequences of (linked) selection from population genomic time series from autocovariance in allele frequency changes across generations. This method can be used to detect fluctuating selection, provided shifts in the direction of selection do not occur frequently (Buffalo & Coop, 2019).

With the exception of Buffalo and Coop (2019), methods designed to detect fluctuating selection perform best when individual genetic

loci experience strong selection. However, many traits are polygenic (e.g. Gompert et al., 2019; Lucas et al., 2018; Pritchard et al., 2010; Shi et al., 2016; Yang et al., 2010), and thus, selection on genes can be weak even when selection on a trait is intense (Walsh & Lynch, 2018). A general approach to overcome this limitation was suggested by Berg and Coop (2014) (also see Edge & Coop, 2019; Josephs et al., 2019). Specifically, polygenic adaptation can be inferred by incorporating genotype–phenotype associations from genome-wide association studies (GWAS) in population genomic tests for selection. This makes it possible to accumulate evidence of selection across trait-associated loci, and thus detect selection even when none of the individual loci experience strong selection. Thus far, this analytical framework has mostly been applied to static population genomic data sets, and not to detect fluctuating selection from temporal data.

Herein, I fill this analytical gap by presenting an approximate Bayesian computation (ABC) method to detect and quantify fluctuating selection on polygenic traits from time-series data. With this method, phenotypic selection is modelled as an explicit function of the state of the environment (similar to Gompert, 2016). The population genomic consequences of selection are then modelled based on estimated genotype–phenotype associations, as summarized by estimated polygenic scores (similar to Berg & Coop, 2014). This allows inferences to be informed by patterns of change across multiple genetic loci, populations and generations. Pooling information in this way increases statistical power, but precludes generic genome scans for selection. Although this could be viewed as a limitation, by making the genotype–phenotype–environment–fitness hypothesis explicit, the proposed approach places an emphasis on the ecological interactions that cause selection to fluctuate and thereby provides more meaningful inferences about the evolutionary process (as advocated for in a general sense by Endler, 1986).

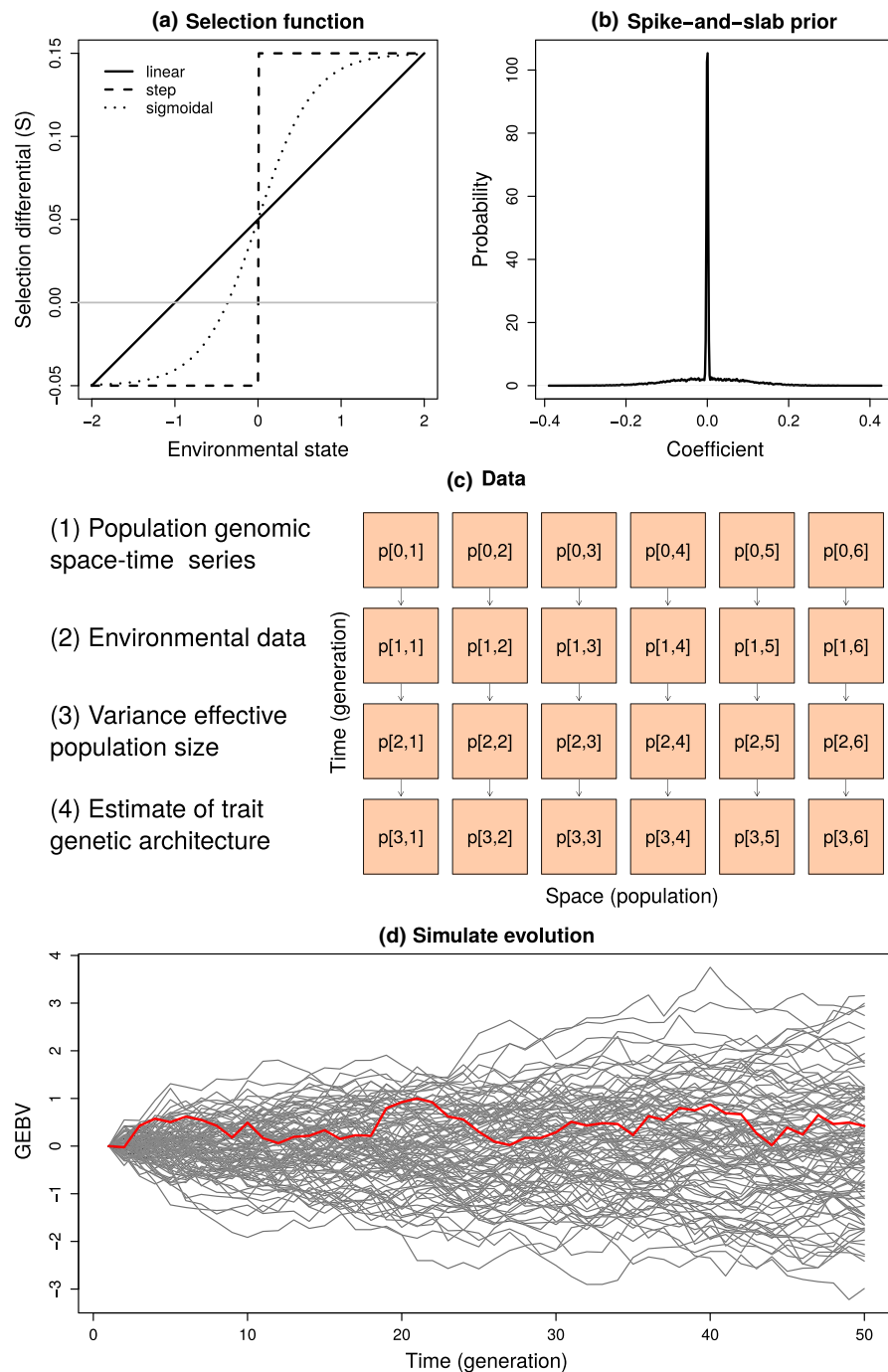
In this study, I first describe the theoretical basis for the approach. This is followed by a description of the proposed ABC method, including the data required and a framework for model evaluation and comparison. The efficacy and limitations of the proposed method are then evaluated by applying it to thousands of simulated data sets, comparing it with an existing method and applying it to a recent experimental evolution study of host adaptation in seed beetles (*Callosobruchus maculatus*). I show that the proposed method produces precise estimates of selection, especially when the generative model for the data coincides with the model assumed by the method, and detects selection on weight in *C. maculatus* despite the low heritability of this trait. I conclude by discussing the utility and limitations of the proposed method, along with potential extensions of the proposed method.

2 | METHODS

2.1 | Model

I first describe a mathematical model that connects phenotypic selection on a quantitative trait to allele frequency change at causal or linked loci (i.e. loci affecting the trait or those in linkage

FIGURE 1 Graphical overview of the model, data and proposed method. Panel (a) shows hypothetical linear (Equation 1), step (Equation 2) and sigmoidal (Equation 3) functions for the effect of the environment on the selection differential. Panel (b) depicts the general form of the spike-and-slab priors used for the coefficients in the selection functions. Most of the prior probability is near zero, but large values are not improbable. Panel (c) summarizes the data for the approximate Bayesian computation (ABC) method. The four key data inputs are listed along with a diagram summarizing the population genomic time series, which includes allele frequency data (denoted P) for multiple loci measured across populations (space) and generations (time). The ABC approach simulates evolutionary trajectories for genetic loci and genomic-estimated breeding values (GEBVs) for each population based on the input data from (c). Panel (d) shown a hypothetical observed GEBV trajectory (red line) and additional simulated trajectories (grey lines) with different values for selection. The ABC approach retains the set of simulations that produce outputs most similar to the observed data (see the main text for details) [Colour figure can be viewed at wileyonlinelibrary.com]



disequilibrium with such loci). Selection is allowed to fluctuate in space and time because of variation in the environment. Selection on the trait is connected to expected allele frequency change based on an understanding of the selected trait's genetic architecture, as might be obtained from a genome-wide association mapping study. These associations are captured by a polygenic score, similar to Berg and Coop (2014). If causal loci and their phenotypic effects are known, the selection estimated is the direct selection on the causal variants, otherwise it is the indirect selection arising from linkage disequilibrium (Gompert et al., 2017; Lande & Arnold, 1983). The goal of the model is to estimate the intensity of selection on the trait and how this varies by environment, as well as the

extent to which this translates into selection on different genetic loci. Key assumptions of the model are that selection is always directional, that each locus explains a small proportion of the phenotypic variance (either because the trait is highly polygenic or has a large environmental variance), that the phenotypic effects of loci do not depend on the environment and that the causal variants (or linked loci) exhibit minimal linkage disequilibrium with each other such that each experiences genetic drift independently (I reiterate and expand on these assumptions below). I also assume that, at least over the duration of the population genomic time series, mutation and gene flow are negligible and can be ignored. After introducing the model in this section, I describe the proposed

approximate Bayesian computation method for fitting the model in the next section.

Let $S_{jk} = \mu^* - \mu$ denote the selection differential for a quantitative trait in population j and generation k . Here, μ and μ^* denote the trait mean before and after selection, respectively. I assume phenotypic selection is a function of the environment. I consider three alternative models for the relationship between that state of the environment, denoted x_{jk} , and the selection differential, S_{jk} , specifically:

$$S_{jk} = a + bx_{jk} \quad (1)$$

$$S_{jk} = \begin{cases} a & \text{if } x_{jk} < c \\ a+b & \text{if } x_{jk} \geq c \end{cases} \quad (2)$$

$$S_{jk} = a + \frac{b-a}{e^{-cx_{jk}}} \quad (3)$$

Here, a , b and c are coefficients that determine the effect of the environment on the selection differential; Equations (1–3) specify a linear, step and sigmoidal relationship between the environment and selection differential, respectively (Figure 1a). The model thus assumes that phenotypic selection is always directional and independent of the current mean trait value, but allows flexibility in how the environment affects the direction and intensity of selection. Furthermore, I focus on the simplest case where selection is a function of a single environmental variable, but the extension to multiple environmental variables is trivial.

I then specify a model that relates phenotypic selection to expected allele frequency change. Let w_i denote the marginal, relative fitness of the non-reference allele at locus i (bi-allelic loci are assumed), that is $w_{ijk} = \int_z w_{jk}(z) f_i(z) dz$. Here, $w_{jk}(z)$ is the expected fitness of an individual with trait value z and $f_i(z)$ is the phenotypic distribution (density function) for individual's carrying a copy of the non-reference allele at locus i (this includes heterozygotes and homozygotes for the non-reference allele) (note, at present I assume that $f_i(z)$ is fixed, that is I assume no plasticity). Following Kimura and Crow (1978) and Walsh and Lynch (2018), the intensity of selection on the non-reference allele at locus i can be approximated as

$$s_{ijk} = w_{ijk} - 1 \simeq \beta_i \left(\frac{S_{jk}}{\sigma_z^2} \right) \quad (4)$$

Here, β_i denotes the average phenotypic excess of the non-reference allele at locus i , which is defined in terms of mean trait values and frequencies of the three genotypes, $\frac{Pz_{AA} + Qz_{Aa}}{P+Q} - \frac{Qz_{Aa} + Rz_{aa}}{Q+R}$, where P , $2Q$ and R are the frequencies of homozygous non-reference, heterozygous and homozygous reference genotypes and z_{AA} , z_{Aa} and z_{aa} are their respective mean trait values (Fisher, 1941). If the genotypes at each causal locus are in their Hardy-Weinberg expected proportions and in linkage equilibrium across loci, this is equivalent to the average effect of an allele, as estimated with standard genome-wide association mapping methods (e.g. Zhou et al., 2013; Zhou &

Stephens, 2012). The term σ_z^2 in Equation (4) denotes the phenotypic variation, and thus, $\frac{S_{jk}}{\sigma_z^2}$ is the standardized selection differential (i.e. the selection differential in units of standard deviations). Thus, the selection on an individual allele is approximately equal to the product of the intensity of phenotypic selection and average phenotypic excess of the allele (Walsh & Lynch, 2018). This approximation is true or approximately true for a variety of fitness functions, but requires that phenotypic variation caused by any one allele is small relative to the phenotypic variance, and thus, that selection on each individual allele is weak. This assumption can hold even when phenotypic selection is strong, if the trait is polygenic and lacks common, major effect loci.

Importantly, s_{ijk} measures selection as the average excess in relative fitness for the non-reference allele at locus i , not the more commonly used selection coefficient (which defines selection relative to a specific genotype, e.g. Gillespie, 2004). The average excess in relative fitness directly measures the effect of selection in bringing about allele frequency change; as such, this metric is not static but depends on the allele frequencies (Kimura & Crow, 1978; Walsh & Lynch, 2018). The expected change for the non-reference allele frequency at locus i due to selection is given by $\Delta p_{ijk} = p_{ijk} s_{ijk}$. Genetic drift is assumed to also contribute to evolutionary change, such that,

$$p_{ijk+1} \sim \text{binomial}(p_{ijk} + \Delta p_{ijk}, 2N_e) / 2N_e. \quad (5)$$

Equation (5) is the Wright-Fisher model with selection; N_e denotes the variance effective population size for population j and generation k (I omit subscripts on N_e for readability) (Ewens, 2004).

I assume that, conditional on the s_{ijk} , evolutionary change is independent across genetic loci (i.e. linkage disequilibrium among the genetic loci is minimal). The validity of this approximation depends on the trait genetic architecture and set of genetic loci included in the analysis. For example, an assumption of linkage equilibrium is more reasonable when causal variants are dispersed across the genome rather than clustered in one or a few genomic regions. The approximation will also be better when each causal genetic variant is represented by a single genetic marker locus, rather than a set of linked genetic markers statistically associated with the same causal variant (see the 'ABC approach' below for further discussion).

2.2 | ABC approach

The proposed method for estimating selection requires four data sources. First, estimates of allele frequencies in one or more populations for multiple time points (ideally consecutive generations) are required (these are assumed to be known with little or no error, but an option exists to relax this assumption in the software implementing this method). Whole-genome sequences are not needed (or warranted), but rather the method requires a set of genetic markers (e.g. SNPs) known to affect the trait of interest, or as will be more commonly the case, a set of markers statistically associated with the trait (i.e. in LD with the unknown causal variants). Second, information

on the effects (associations) of the genetic markers with the trait is needed. Here, I assume that each marker has a probability of affecting or being associated with the trait, denoted γ_i (e.g. a posterior inclusion probability or PIP), and an effect conditional on a true association, denoted β_i (which I equate with β_i from Equation (4) assuming the genotypes in the populations are approximately equal to their expected, Hardy–Weinberg proportions; see Fisher, 1941). Standard models for polygenic GWA mapping and genomic prediction, such as the Bayesian sparse linear mixed model approach from GEMMA, output this information (Zhou et al., 2013). Such models provide probabilistic estimates of direct phenotypic effects (and thus of direct selection with the model) when the causal variants are sequenced, but only indirectly approximate direct selection when this is not the case (Gompert et al., 2017). Alternatively, γ (the probability of association) can be set to 1 for a set of genetic variants to indicate they are known to directly affect the trait; such confidence might be appropriate if genotype–phenotype associations have been validated by genetic manipulations (e.g. Barrett et al., 2019). Lastly, model-averaged effect estimates can be used, that is $\bar{\beta}_i = \gamma_i \beta_i$, with the probability of association and then set to 1. I explore this final possibility with an empirical data set.

Third, estimates of the variance effective population size for each population are needed. These can be estimated using either LD-based methods or change over time for a set of genetic markers not associated with the (putatively) selected trait (e.g. Do et al., 2014; Gompert & Messina, 2016; Jorde & Ryman, 2007). The proposed method accounts for uncertainty in estimates of N_e by integrating over a posterior distribution.

Fourth, the proposed method requires the measurement of an environmental covariate hypothesized to determine the phenotype–fitness relationship (i.e. the selection differential). This covariate should be measured in each population and generation (or time step). I use the term environmental covariate in a broad sense, as this could include a variety of factors: temperature, precipitation, composite climatic variables from an ordination (e.g. principal components), predator abundance, resource availability, or even population density or the frequency of a trait in the population. Moreover, the relevant covariate value at time t could represent the measurement at time t or a cumulative measure over some past amount of time.

Given these data, I propose an approximate Bayesian computation (ABC) method for estimating the parameters of the selection function, a , b and c in Equations (1–3), as well as derived parameters related to selection on the genetic loci, which I describe below. In other words, inference of selection function parameters is based on how polygenic scores vary through time, but this provides a means to summarize the posterior in terms of additional derived parameters about selection on SNPs at no analytical cost. This method is implemented in the C++ program FSABC (version 0.1; available from GITHUB), which is used in combination with the R package ABC (version 2.1; Csillery et al., 2012) for inference. As with other ABC methods, the proposed method involves first sampling parameter values from prior probability distributions, then simulating evolution based on those parameter values and finally calculating summary statistics

for the simulated and observed data (all done with FSABC; Figure 1). Posterior distributions for the parameters are then obtained based on the subset of sampled parameter values and simulations that most closely match the observed data in terms of the summary statistics (e.g. Sisson et al., 2018).

I assume the form of the selection function (linear, step or sigmoidal) is known and that the coefficients describing the function (a , b and c) are the primary, unknown parameters (but see below for model selection and the possibility of model averaging) (Figure 1a). I specify spike-and-slab priors for each of these parameters, such that $\Pr(a) \sim \pi \delta_0 + (1 - \pi)U(l, u)$ (Figure 1b). Here, π is the probability of a zero coefficient and $U(l, u)$ denotes a uniform probability density function with lower bound l and upper bound u . δ_0 denotes a point mass at zero. This prior is conservative as it places a substantial proportion of the prior probability on zero (the exact value is given by π), but still allows for large coefficients (based on l and u). Values of N_e and the specific set of genetic markers associated with the trait (if either are expressed with uncertainty) are then sampled according to their probabilities.

Derived parameters S_{jk} (the selection differential) and s_{ijk} (selection on each locus) for population j and generation k are then calculated from Equations (1, 2 or 3), and Equation (4) based on the sampled parameter values. The simulation for each population is initiated with the known, initial allele frequencies in generation $k = 0$, and the allele frequencies in the following generation are then determined based on the stochastic Wright–Fisher model (Equation 5) (this is then the starting point for the next generation). This process is then repeated for each population and generation to simulate the complete population genomic time series.

I compute two summary statistics for inference of the selection function parameters (a , b and c): (i) the total change in the mean polygenic score (i.e. mean genomic-estimated breeding value or GEBV), and (ii) the covariance between change in the mean polygenic score and the value of the environmental covariate. The total change in the mean polygenic score captures the average selection differential (e.g. a in the linear model, Equation (1)) and is defined as $\Delta_{BV} = \sum_{ijk} \beta_i (p_{ijk+1} - p_{ijk})$. The covariance between change in the breeding value and the environmental state, that is $\text{COV}(\Delta_{BV_k}, x_{jk})$, captures the manner in which the environment affects the selection differential (e.g. b in the linear model, Equation (1)).

Primary interest is in inference of the selection function parameters, that is the parameters that describe how the environment affects the selection differential on the trait and thus change in the mean polygenic score. However, the proposed method can also estimate derived parameters describing selection on each genetic locus, that is s_{ijk} . I focus on two derived parameters that summarize generation, population and locus selection coefficients, specifically, the average absolute intensity of selection on the genetic loci (i.e. $\text{mean}(|s_{ijk}|)$) and the standard deviation of the absolute intensity of selection (i.e. $\text{SD}(|s_{ijk}|)$). These provide high-level summaries of the intensity and variability of selection on loci associated with or affecting the selected trait (i.e. indirect or direct selection depending on the nature of the genotype–phenotype map).

Various approaches have been developed to generate samples from an approximate posterior distribution in an ABC framework from summary statistics (Beaumont, 2010; Sisson et al., 2018). Here, I used the rejection with local linear regression correction algorithm proposed by Beaumont et al. (2002) and implemented in the *R* ABC package (Csillery et al., 2012). This approach performs quite well (see Results), but other approaches, such as ridge regression or non-linear regression (Blum & Francois, 2010), could be used instead for this portion of the analysis.

Beyond parameter estimation, the method also provides a convenient way to compute a posterior predictive distribution, which can be used for model assessment and comparison. For this, I simulate additional evolutionary trajectories by sampling parameter values (a , b and c) from their posterior distributions based on an initial model-fitting analysis (as described above). I then compute a new set of summary statistics for each simulated data set. Specifically, percentiles (10th to 90th percentiles in steps of 10%) of the distribution of change in the mean polygenic score across populations and generations are computed. I chose this novel set of summary statistics so that different information could be used for model fitting and model validation or comparison. By simulating multiple new data sets, posterior predictive distributions for each of the percentile summary statistics can be obtained and compared to the same summary statistics calculated from the observed data, with higher correspondence being indicative of better model performance.

2.3 | Analysis of simulated data

I first evaluated the performance of the method with simulated data. I began by testing how the method performed under a variety of conditions, but where the data were simulated under the same model used to analyse the data. This allowed me to assess the ability of the approach to provide precise and accurate parameter estimates when the model was the correct model for the data. I then conducted additional simulations to assess how the method performed for data simulated under alternative models.

The first set of simulations was conducted using the model described above and with the same software used for inference (i.e. *FSABC*). Standard (baseline) conditions involved 10 populations sampled in 10 successive generations. I further assumed 100 genetic loci affected the focal trait. Initial allele frequencies for these loci were sampled from a beta distribution with shape parameters equal to 0.5. The phenotypic effects of these 100 loci were drawn from a centred normal distribution. The standard deviation was set to ensure a heritability of ~ 0.3 (with a phenotypic variance of 1.0). That is, I assumed the trait was polygenic, lacked major effect loci and had a modest heritability (and thus a substantial non-genetic variance component). I assumed the variance effective population size was estimated with uncertainty, such that knowledge and uncertainty in N_e were characterized by a re-scaled beta distribution, $N_e \sim \beta(\alpha = 10, \beta = 10, \min = 400, \max = 600)$ (here min and max

are the lower and upper bounds). This distribution has a mean of $N_e = 500$ and standard deviation of 21.8 (i.e. a modest effective population size known with some, albeit limited, uncertainty). The environmental state for each population and generation was sampled from a standard normal distribution. I assumed a linear relationship between the environmental covariate and the selection differential, such that $S_{jk} = a + bx_{jk}$ (i.e. Equation 1), and sampled values from the spike-and-slab priors for a and b for the simulated data sets using the same prior parameter values used for the ABC inference (lower and upper bounds of the slab component of the priors set to -0.1 and 0.1 and $\pi = 0.5$; see below).

I analysed 2000 replicate simulations under these conditions (replicates differed in the values for selection, not in the trait genetic architecture or initial allele frequencies). Specifically, I generated 1.5 million simulations of evolution using the ABC simulation method and treated the first 2000 as observed data. I applied the ABC inference procedure to each using the *ABC* function from the *R* ABC package (version 2.1) (Csillery et al., 2012). I used the local linear regression method with a tolerance of 0.005, which retained 0.5% of the 1.5 million simulations (7500 samples) to form the posterior after regression adjustment. I summarized parameter estimates based on the posterior median (point estimate) and 95% equal-tail probability intervals (ETPIs). The performance of the method was then assessed by computing the mean absolute error (MAE), defined as the average absolute deviation between the true parameter value and its point estimate (posterior median), and the 95% interval coverage, that is the proportion of cases where the 95% ETPIs contained the true parameter value.

I further assessed the limitations of the method by analysing four additional sets of simulations, each deviating in some way or ways from the standard conditions described above. One set considered the effect of enhanced genetic drift by replacing the re-scaled beta distribution above with one bounded by 100 and 200, and thus with an expected variance effective population size of $N_e = 150$. A second set assumed 1000 genetic loci might affect the focal trait, but postulated uncertainty in whether or not each had a causal effect. Specifically, each genetic marker was assigned a probability of association (i.e. γ_i or posterior inclusion probability) of 0.1. This results in an expectation of 100 causal variants but with uncertain identities. Third, I considered limited sampling, specifically five populations sampled across five generations. Fourth, I assumed a step selection function (Equation 3) rather than the linear function, with $\pi = 0.5$ and slab bounds of -0.05 and 0.05 for c . Each simulated data set (2000 per set of conditions) was analysed as described above, with the exception that simulated genetic data with a step function for selection were analysed with both the linear and step function models.

Lastly, I simulated additional data sets that deviated to greater extent from the assumed model. I specified the same genetic architecture, allele frequencies, effective population size and environmental data as for the standard conditions, but assumed that the environment defined an optimum for a Gaussian fitness function. Thus, instead of a simple selection differential (which only

captures directional selection), selection would be a mixture of stabilizing and directional with specifics determined by the current composition of the population. The width of the Gaussian selection function (i.e. its standard deviation) was set to two or five (smaller standard deviations coincide with stronger stabilizing selection). The mean of the Gaussian function (i.e. that optimum phenotype) was specified by a linear model, $\mu_{jk} = \alpha + \beta x_{jk}$. I considered two values for α , -0.762 or -0.562 , and two values for β , 0 or 0.4 . The first value for α (-0.762) coincides with the expected initial mean phenotype given the trait genetic architecture and allele frequencies, and thus with no (initial) directional selection in the average environment, whereas $\alpha = -0.562$ should result in more instances of selection for larger trait values. Values for β were chosen to correspond with environment-independent ($\beta = 0$) and environment-dependent ($\beta = 0.4$) selection. I generated 20 simulated data sets for each combination of α and β (80 data sets total). This was done with an individual-based simulation written and implemented in R (version 3.5.1; code available from Dryad, DOI pending).

In these simulations, each generation genotypes for $N = 1000$ individuals were sampled from binomial probability distributions parameterized by the population allele frequencies. I then determined the phenotype of each individual by computing its polygenic score and adding environmental noise (to ensure the heritability of ~ 0.3) from a normal distribution. The relative fitness of each individual was then calculated from the Gaussian fitness function conditional on the environmental state. I then sampled N_e individuals, based on their relative fitness values, as survivors that determined the allele frequencies in the next generation. I repeated this procedure across 10 generation, 10 populations and 80 total simulations (20 replicates for each combination of selection function parameters). Because the values α and β for the selection function are for the phenotypic optimum (μ_{jk}) rather than the selection differential (S_{jk}), the method is not expected to estimate these values accurately. I thus instead compute approximate values for the true a and b from these simulations by fitting a linear model to the observed selection differentials from the simulations (i.e. from the change in the expected trait value from generation to generation), $S_{jk} = \hat{a} + \hat{b} * x_{jk}$ (the hat symbols denote that these are estimates of a and b , not parameter values defined by the model). I analysed these 80 simulated data sets as described above and measured performance by the ability to estimate \hat{a} and \hat{b} .

I conclude my analysis of these simulated data sets by illustrating how the proposed method can be used for model comparison. I chose (at random) one data set simulated based on the model with the linear selection function (and more generally, under the standard conditions). For this data set, I fit the linear selection function model (the true model) and the step selection function model (an alternative model), both as described above. I then generated 5000 samples from the posterior predictive distribution for the 10th to 90th percentiles of the distribution (over time and space) of change in polygenic scores based on each model. I compared these distributions graphically to the equivalent summary

statistics computed from the observed data (in this case, the original simulated data set).

2.4 | Comparison with an alternative statistical method

Other methods to detect temporal variation in selection on polygenic traits driven by specific environmental factors are not available at present. However, methods exist to test for spatial variation in polygenic selection by leveraging genotype–phenotype associations (Berg & Coop, 2014; Field et al., 2016; Turchin et al., 2012). In particular, Berg and Coop (2014) proposed a method to identify the phenotypic targets and environmental drives of polygenic selection from exceptionally strong correlations between polygenic scores and environmental variables across a set of contemporary populations. Whereas the method was designed to test for selection that varies across space, the general structure of the model makes it possible to apply to spatiotemporal data. Thus, I compared the performance and information obtained from my proposed ABC method to Berg and Coop's (2014) polygenic score–environmental covariate correlation method by using each to analyse additional simulated data sets.

For this, 20 new simulated data sets were generated using the individual-based model with stabilizing selection described above. Aside from the explicit exceptions that follow, simulation conditions and parameter values were as described above. For these simulations, the trait heritability was set to either 0.3 or 0.7 (ten simulations each; 100 causal variants). Parameter values (α and β) specifying the phenotypic optimum (μ_{jk}) were set to the expected trait value (α) and 0.4 (β). Thus, I simulated stabilizing selection with an environment-dependent optimum.

I analysed each simulated data set with my proposed ABC method and the polygenic score–environmental covariate correlation method (Berg & Coop, 2014). ABC analyses were based on 1.5 million simulations and used the local linear regression method with a tolerance of 0.005 to approximate the posterior distribution. For the Berg and Coop (2014) method, each sample (generation and location) was treated as a unique population. The genetic composition of the population (i.e. the mean polygenic score) was assumed to reflect the environmental state from the previous round of selection (thus the initial population samples were not included in the analysis). Null expectations were determined using a set of 10,000 simulated control SNPs, which were matched to the trait-associated SNPs based on minor allele frequencies (in bins of 0.05). For all 20 data sets and both methods, analyses were run with both the true environmental covariate and a randomized environmental covariate.

2.5 | Analysis of a seed beetle evolve-and-resequence experiment

I further assessed the efficacy of the proposed method by using it to test for selection on female weight during an evolve-and-resequence

experiment with the cowpea seed beetle, *Callosobruchus maculatus*. These insects infest stores of grain legumes (Tuda et al., 2014). Female beetles attach eggs to the surface of legume seeds. Hatching larvae burrow into the seed and must complete development in the single, natal seed.

My colleagues and I previously conducted an evolve-and-resequence experiment to analyse the evolutionary dynamics of host adaptation in *C. maculatus* when experimentally shifted to a stressful legume host, lentil (*Lens culinaris*), from an ancestral host, mung bean (*Vigna radiata*) Rêgo et al., (2019). Low survival (~1%) at the onset of the experiment caused a population bottleneck, but adaptive evolution quickly rescued the population from extinction with survival rates increasing to 90% by the F10 generation (Rêgo et al., 2019). Previous analyses of generation-scale allele frequency change from this experiment showed strong selection caused rapid adaptation and substantial evolutionary change at many genetic loci (Rêgo et al., 2019). Genome-wide association mapping in a back-cross (BC) mapping population derived from this line documented modest heritability for female weight at eclosion, with 17% of the trait variance explained by genetic markers (Rêgo et al., 2020). Female weight is one of several traits that often evolves during adaptation to lentil by *C. maculatus* (Messina & Durham, 2015; Messina & Jones, 2011; Messina et al., 2009). However, there was considerable uncertainty in individual genotype–weight associations, and we failed to detect excess overlap between SNPs associated with weight and those evolving most rapidly during the evolve-and-resequence experiment (Rêgo et al., 2020).

Here, I reanalyse these data with the proposed method and specifically ask whether the method can detect selection on female weight and whether selection depends on population density (the hypothesized environmental covariate). Explicit data on population density are lacking, but qualitatively the population size had rebounded by the F4 generation, and thus, I encode density as a binary covariate with a value of 0 for the first four generations and 1 thereafter. For this analysis, the data set included 10,409 SNPs sequenced in each of 10 samples (48 beetles per sample)—the founders (P), F1–F8 generations and the F16 generation—and 251 sequenced female beetles that comprise the BC mapping population. Additional details regarding these data and experiments are provided by Rêgo et al., (2019) and Rêgo et al., (2020). SNP–genotype associations for the mapping population were estimated with the Bayesian sparse linear mixed model approach implemented in GEMMA (version 0.98) (Zhou et al., 2013). Consequently, for each SNP, there is an estimated posterior probability of association (i.e. posterior inclusion probability or PIP) and an estimated phenotypic effect conditional on the association. I assume the phenotypic variance in the experiment was equal to that in the mapping population ($\sigma^2 = 0.43$ mg; see the Discussion for limitations of this assumption). The Bayesian allele frequency estimates were taken from Rêgo et al., (2020). I estimated the variance effective population size over the course of the experiment based on patterns of allele frequency change between the P and F16

generation using the Bayesian approach implemented in VARNE (version 0.9) (Gompert & Messina, 2016).

I used the proposed method to estimate selection on female weight. The model was fit in two ways, using either (i) model-averaged effect estimates (i.e. taking the product of the posterior inclusion probability and effect estimate as a certain, model-averaged effect estimate) or (ii) integrating over uncertainty in genotype–phenotype associations. My main focus was on the analysis with population density (a qualitative, binary indicator variable) as an environmental covariate, but I also considered a null model with directional selection only and a model with a random environmental covariate (with environmental states drawn from a standard normal distribution). I bounded the priors on the selection model parameters (a and b) by -2.0 and 2.0 (here S is in mg). In each case, I based the analysis on one million simulations. In the models with uncertainty in genotype–phenotype associations, observed summary statistics were obtained for 100 independent samples of genotype–phenotype associations (based on their PIPs). Inferences were based on 1000 samples from the posterior after applying the local linear regression adjustment in ABC (Beaumont et al., 2002; Csillery et al., 2012).

3 | RESULTS

3.1 | Results from stimulated data

The core simulations (i.e. those based on the model) resulted in weak-to-modest phenotypic selection, such that the average, absolute intensity of selection was less than 0.01 for 55% of the simulations and had a maximum value of ~0.1 (Figure S1). Variability in selection across space and time was of the same order of magnitude as the average, absolute intensity of selection, consistent with Siepielski et al. (2009) and Siepielski et al. (2013) (mean standard deviation in $|S_{ij}| = 0.025$, maximum = 0.10) (Figure S1). The average selection intensity on individual genetic loci (measured by $|S_{ijk}|$) was about an order of magnitude lower and rarely exceeded 0.01 (Figure S2).

Under the standard conditions (i.e. linear selection function, 10 populations, 10 generations, 100 causal loci, $\bar{N}_e = 500$, $h^2 = 0.3$), I obtained precise and accurate estimates of the mean selection differential (a) and environmental effect (b) (Figure 2). For example, the MAE (i.e. the average absolute deviation between the true and estimated parameter value) was 0.0028 for a and 0.0025 for b . Moreover, the 95% interval coverage was 0.95 and 0.96 for a and b , respectively (Figure S3). High accuracy was mostly independent of the strength of these effects (Figure S4). Estimates of the average selection on genetic loci were similarly accurate with a MAE of 0.00031 and 95% interval coverage of 0.96 for the standard conditions (Figure 3). There was a slight decrease in the accuracy of the estimates for simulations with reduced N_e (i.e. cases where drift was more important), uncertainty in genetic architecture, reduced spatial and temporal sampling or with the step selection function (Figure 4). For example, MAE increased, approximately doubling, with reduced

FIGURE 2 Plots summarize the performance of the method under the standard simulation conditions. Panels (a) and (b) show the true (blue points) and estimated (black points with vertical lines for 95% equal-tail probability intervals [ETPIs]) parameter values for the selection coefficients a and b , respectively. Results are shown for 300 representative simulations. Panels (c) and (d) provide scatter plots depicting the relationship between the true parameter values (x-axis) and the point estimate (posterior median, y-axis) for a (c) and b (d). A one-to-one line is shown. Parameter estimates were highly correlated with their true values (Pearson's $r > .99$) [Colour figure can be viewed at wileyonlinelibrary.com]

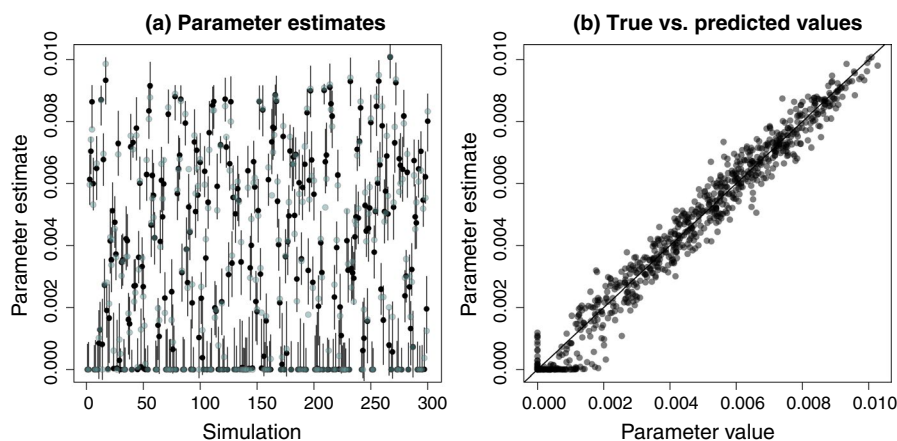
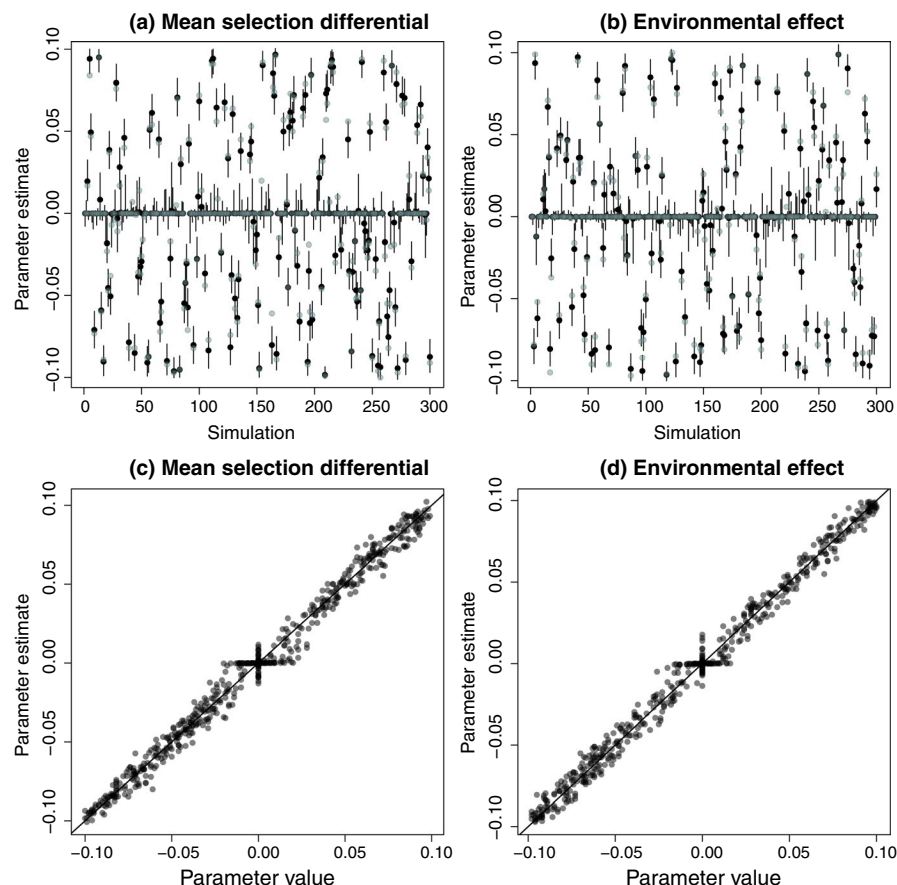


FIGURE 3 Plots summarize the performance of the method in terms of selection on genetic loci under the standard simulation conditions. Panel (a) shows the true (blue points) and estimated (black points with vertical lines for 95% equal-tail probability intervals [ETPIs]) mean absolute intensity of selection on the genetic loci. Results are shown for 300 representative simulations. Panel (b) provides a scatter plot depicting the relationship between the true mean intensity of selection on genetic loci (x-axis) and the point estimate (posterior median, y-axis). A one-to-one line is shown. Parameter estimates were highly correlated with their true values (Pearson's $r = .99$) [Colour figure can be viewed at wileyonlinelibrary.com]

sampling (MAE = 0.0064 for a and 0.0061 for b), though 95% interval coverage remained high (~95%).

There was a greater reduction in performance for simulations based on the Gaussian fitness function, that is where the analysis model differed from the generative evolutionary process used to simulate the data (Figures 4 and 5). In particular, MAE increased

to 0.0075 for a and 0.0105 for b , and the 95% interval coverage dropped to ~0.49 for both. Despite this drop in performance, the parameter estimates were still indicative of the 'true' parameter values (recall that \hat{a} and \hat{b} are not true parameter values in the same sense for these data sets, but rather best approximations when forcing a linear model for the selection differential). For example, Pearson's

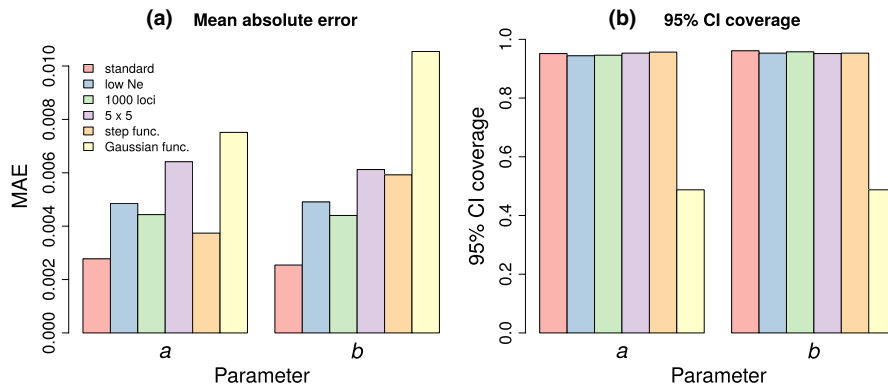


FIGURE 4 Bar plots depict how the performance of the method depends on simulation conditions. Results are shown for mean absolute error (MAE) (a) and 95% interval coverage (b) [Colour figure can be viewed at wileyonlinelibrary.com]

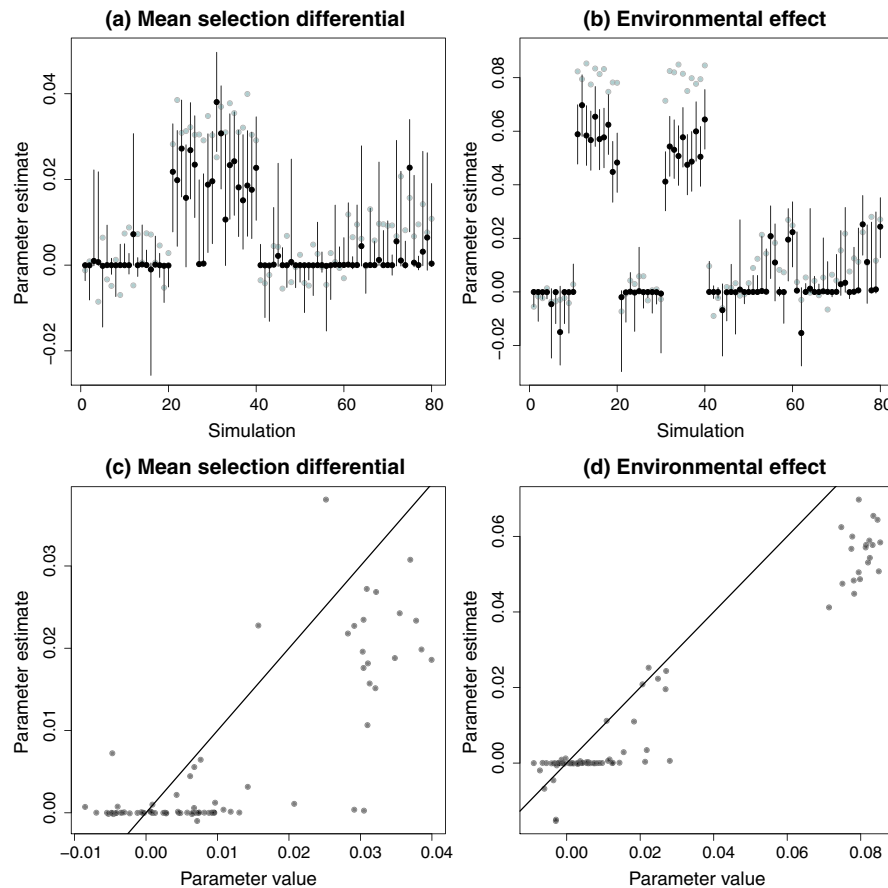


FIGURE 5 Plots summarize the performance of the method for data simulated assuming a Gaussian fitness function with an optimum determined by the environment. Panels (a) and (b) show the true (blue points) and estimated (black points with vertical lines for 95% equal-tail probability intervals [ETPIs]) parameter values for the selection coefficients *a* and *b*, respectively. Panels (c) and (d) provide scatter plots depicting the relationship between the true parameter values (x-axis) and the point estimate (posterior median, y-axis) for *a* (c) and *b* (d). A one-to-one line is shown [Colour figure can be viewed at wileyonlinelibrary.com]

correlations between true and estimated values (posterior medians) for *a* and *b* across the 80 data sets simulated under this model were 0.80 (95% confidence interval = 0.71–0.87) and 0.97 (95% confidence interval = 0.96–0.98), respectively.

A comparison of the posterior predictive distribution for models with either linear or step selection functions fit to a simulated data set where the true function was linear illustrates how the proposed method can be used for model comparison (see Figure S5 for the model fits). As expected, the true summary statistics fell closer to the centre of the posterior predictive distribution for the true linear model than the step function model (Figure 6). This held across the various percentiles considered for the summary statistics.

Both the proposed ABC method and the environmental correlation approach introduced by Berg and Coop (2014) were able to detect fluctuating selection on a polygenic trait when it

occurred (Figure 7). For example, with the ABC approach, posterior probabilities for environment-dependent selection ($b \neq 0$) were all >0.97 . Likewise, with the Berg and Coop (2014) method, polygenic scores were more strongly correlated with the previous environment at a location than expected by chance in all 20 simulated data sets (Spearman's ρ , all $p < .05$). Moreover, neither approach incorrectly detected environment-dependent selection when the analysis was based on a randomized environmental covariate (Figure 7). Specifically, the posterior probability of an environmental effect on selection was greater than 0.9 in only two of 20 cases and exceeded 0.95 in only one case for the ABC method, and only one spurious significant correlation was detected by the Berg and Coop (2014) approach ($p = .012$, all others $p > .07$). Whereas the methods performed similarly in terms of detecting selection, only the proposed ABC method simultaneously tests

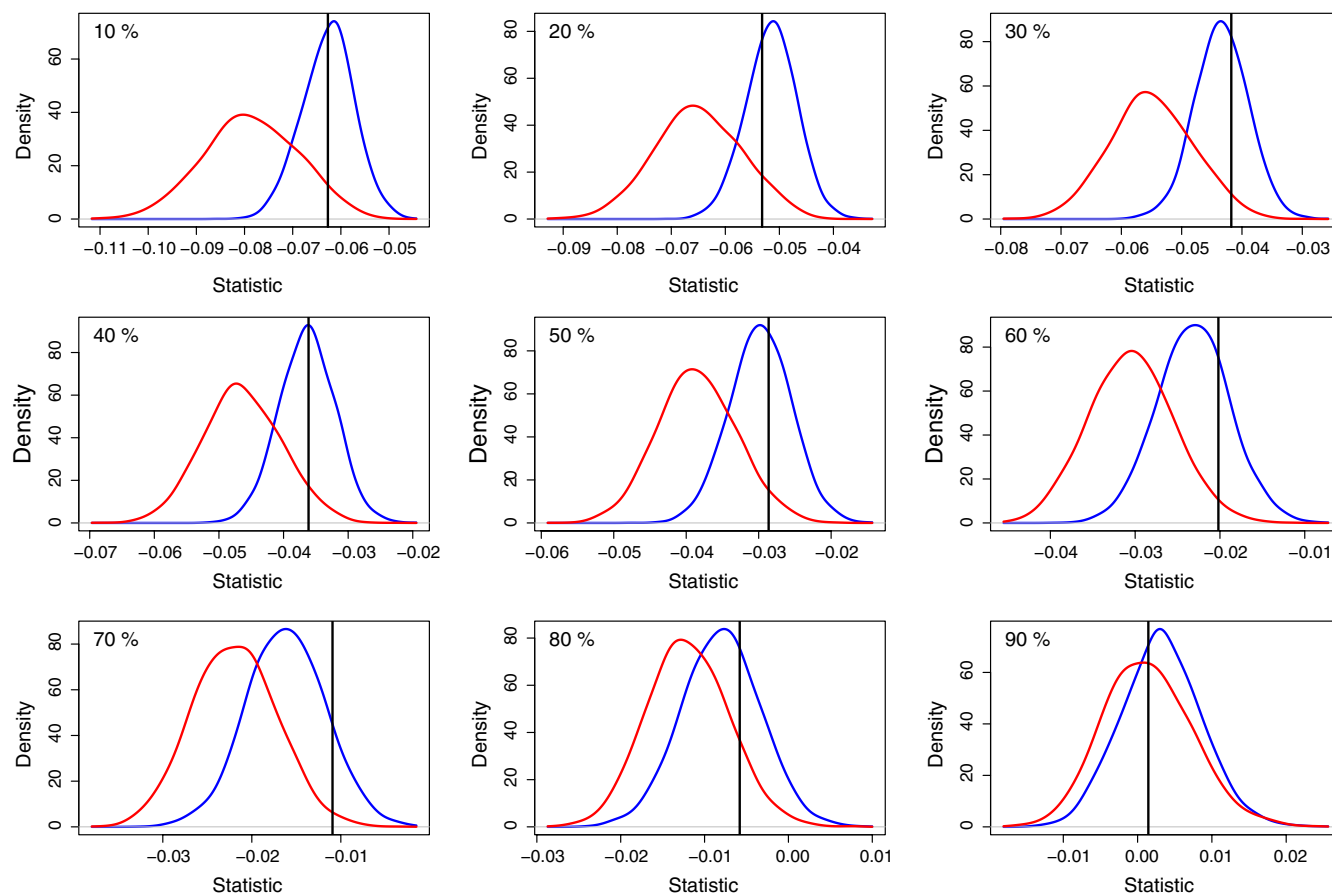


FIGURE 6 Graphical summary of a posterior predictive check. In each panel, the posterior predictive distribution is shown for the linear selection model (blue) and the step selection model (red). The former is the true model. The summary statistic from the data set used to fit the models (i.e. the observed data) is shown as a vertical black line. Each plot shows results for a different summary statistic, specifically a different percentile of the distribution of change in the mean polygenic score across populations and generations (the percentile is given in the upper left corner of each plot) [Colour figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

for environment-independent selection and provides estimates of the strength of selection and how strongly selection intensity is affected by the environment.

3.2 | Results from the seed beetle evolve-and-resequence experiment

I found compelling evidence of strong selection for increased female weight when using model-averaged genotype-phenotype effects (Figure 8a). Estimates of the average selection on weight (a) were mostly insensitive to my choice of an environmental covariate, with estimate of a equal to 1.95 (95% ETPI = 1.70–2.01), 1.99 (1.85–2.03) or 2.00 (1.88–2.06) when using population density, a random covariate or no covariate, respectively. In contrast, fully accounting for uncertainty in genotype-phenotype associations dramatically reduced the evidence for selection (Figure 8A). When fully accounting for uncertainty in the genotype-phenotype map in simulations and the observed data, the estimate of a was ~ 0 (posterior mean = 0.156, 95% ETPI = -0.69 to 1.57, posterior prob. $a > 0 = 0.62$). Using the mean of summary statistics across replicate genotype-phenotype

maps for the observed data decreased uncertainty in a some and increased the evidence for $a > 0$ slightly, but had little effect on the point estimate, 0.0003 (posterior mean = 0.020, 95% ETPI = -0.00 to 0.28, posterior prob. $a > 0 = 0.83$). This discrepancy between results based on model-averaged phenotypic effects and fully accounting for uncertainty in the genotype-phenotype associations is consistent with calculations of mean breeding values or polygenic scores over the course of the experiment and likely reflects the great uncertainty in genotype-phenotype associations in this data set (the highest SNP posterior inclusion probability was 0.03, and only 13 SNPs had posterior inclusion probabilities > 0.01) (Figure S6).

Similar to the results for a , I found evidence that selection varied by population density when using model-averaged effects, but not when accounting for uncertainty in genotype-phenotype associations (Figure 8b). Even when using model-averaged effects, I found no evidence that selection on weight varied as a function of the random covariate ($b = -0.19$, 95% ETPI = -1.33 to 1.31, posterior prob. $b < 0 = 0.69$). Regardless of the specific model, estimates of the average and variability of selection on individual loci were always low (< 0.001). Lastly, a posterior predictive check based on the best performing model, that is the population density

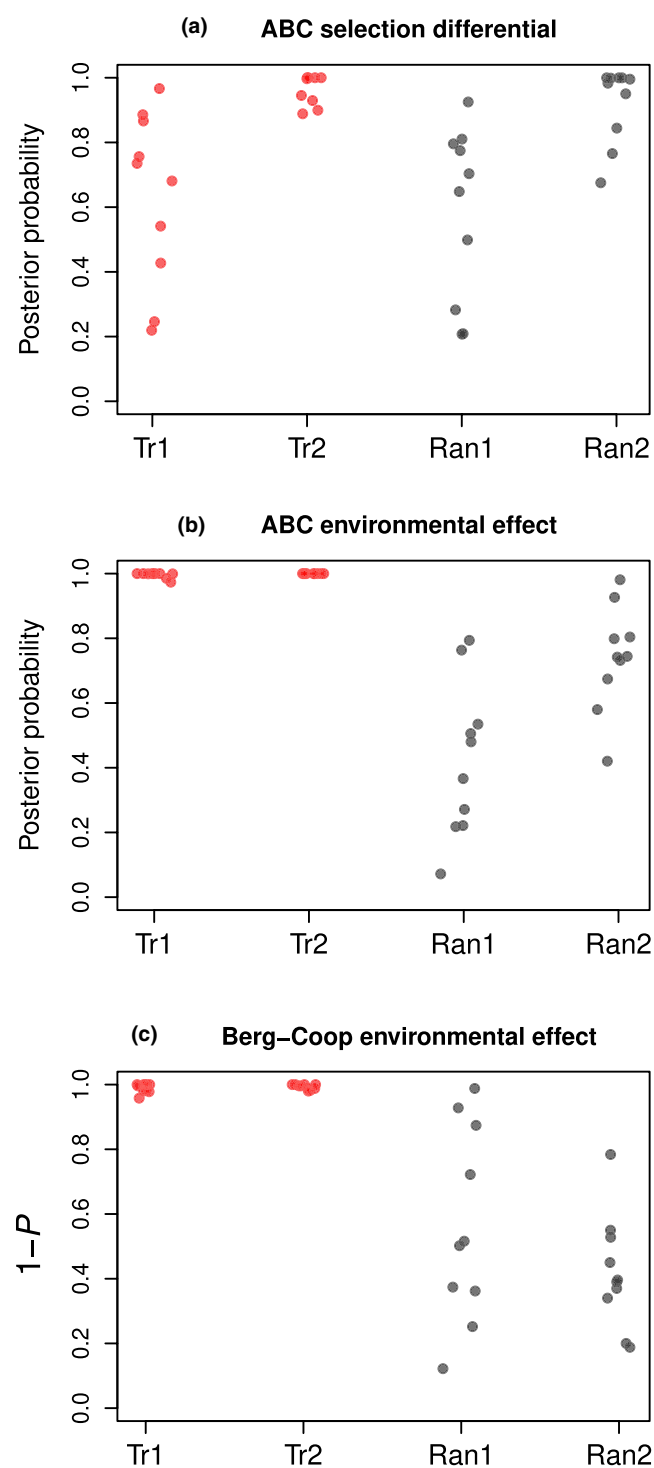


FIGURE 7 Comparison of the performance of the proposed ABC method versus the polygenic selection method introduced by Berg and Coop (2014). Panels (a) and (b) show the inferred posterior probability of selection, that is of *a* (a) or *b* (b) non-zero and in the correct direction. Points denote results from each of 20 simulations with $h^2 = 0.3$ (Tr1 or Ran1) or 0.7 (Tr2 or Ran2). Tr1 and Tr2 indicate analyses based on the actual environmental covariate, whereas Ran1 and Ran 2 use a randomized environmental covariate. Panel (c) shows analogous results from the Spearman correlation test from Berg and Coop (2014). Here, points denote $1 - P$ [Colour figure can be viewed at wileyonlinelibrary.com]

model with model-averaged effects, underestimated the variability in change in the expected polygenic score across generations (Figure 8c).

4 | DISCUSSION

Rapid evolution on ecological time scales is at least somewhat common (Bergland et al., 2014; Campbell-Staton et al., 2020; Endler, 1986; Grant & Grant, 2014; Hendry, 2016; Nosil et al., 2018; Thompson, 2013), but remains understudied from a population genomic perspective (Messer et al., 2016). This is perhaps not surprising, as most existing methods were designed to detect weaker selection operating in a consistent manner over longer periods of time (reviewed in Walsh & Lynch, 2018). I proposed and evaluated a statistical method to help fill this gap. Critically, the method connects environmental drivers of selection with traits and patterns of evolutionary change in a population genomic context. Taking a population genomic approach has the advantage of estimating selection on genetic variants (not just traits) and of mostly disentangling temporal variation in traits due to evolutionary change versus plasticity. Via simulations, I showed that the proposed method can generate very accurate and precise estimates of fluctuating selection on polygenic traits. However, performance depended on simulation conditions and is generally expected to be sensitive to details of the data input. I discuss the performance at length below, along with current limitations and possible extensions of the proposed method.

4.1 | Overall performance

The proposed method was computationally efficient, and I do not foresee problems arising even when applying it to large, whole-genome data sets. As an example, I was able to generate one million ABC simulations for the *C. maculatus* data set in ~300 CPU minutes on a standard desktop computer (Intel Core i7-3930K CPU at 3.20 GHz \times 12 with 64 GB RAM and running Ubuntu 18.04.4 LTS). Moreover, each simulation is independent, and thus, massive parallelization of the ABC simulations is trivial. Computational time should scale linearly with the number of populations and generations, and with the number of genetic variants associated with or affecting the phenotype. Notably, the number of phenotype-associated genetic markers will likely be on the order of hundreds to perhaps tens of thousands even when whole-genome sequences are analysed (i.e. not every genetic variant affects a trait).

The proposed ABC method and a related method designed to detect spatial variation in selection on polygenic traits (Berg & Coop, 2014) had similar power to detect polygenic selection that varied in time. Although this will not always be true, this finding shows that methods designed for spatial analyses can, in some cases, be productively applied to temporal data sets simply by treating samples in time as distinct populations. Nonetheless, methods developed specifically

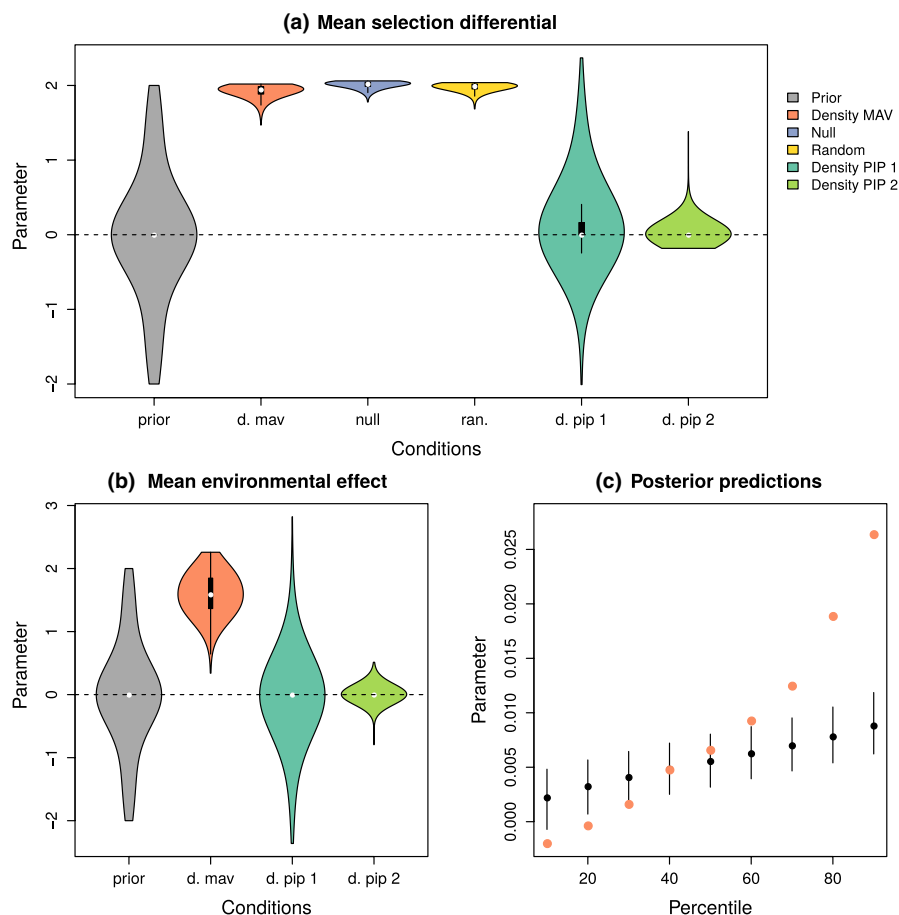


FIGURE 8 Plots summarize the application of the proposed method to the *Callosobruchus maculatus* data set. Violin plots in panel (a) depict prior (grey) or posterior (other colours) distributions for the mean selection differential parameter, and *a.* posteriors are shown with density as a covariate and with model averaging (density MAV) or uncertainty with the distribution (density PIP 1) or mean (density PIP 2) observed summary statistics, and for no covariate (null) or a random covariate (random). White dots denote the median, black boxes denote the 1st and 3rd quartiles, vertical bars extend to the minimum and maximum value or $1.5 \times$ the interquartile range, and kernel densities show the full prior or posterior distribution. Similar plots in panel (b) show the prior and posteriors for the environmental effect, and *b.* panel (c) summarizes the posterior predictive check. Coloured points shown percentiles of change in the mean polygenic score for the observed data (using model averaging), and black dots and lines show the median and 95% intervals from posterior predictive simulations [Colour figure can be viewed at wileyonlinelibrary.com]

for temporal data sets, such as the one proposed here, retain some advantages. In particular, by modelling the temporal evolutionary process, it is possible to move beyond null-hypothesis significance testing (i.e. tests for selection) to parameter estimation with uncertainty, that is to estimating the intensity of selection and how it varies by the environment (see, e.g., Kruschke & Liddell, 2018). The benefits of this approach will depend on the extent to which the model captures key aspects of reality, as I discuss in more detail below.

In terms of accuracy of inference for the proposed method, the performance was best under the standard simulation conditions (i.e. 10×10 populations \times generations, moderately large N_e and 100 known causal variants). I documented minor degradation of performance with more limited space-time sampling (5×5 populations \times generations) or reduced effective population size (i.e. increased genetic drift). Deviations between the assumed and true selection function and uncertainty in genetic architecture had greater effects, and I discuss each in turn next.

4.2 | Performance and the selection function

The performance of the method declined substantially when applied to simulations of environment-dependent stabilizing selection described by a Gaussian fitness function. For example, the true parameter value (or at least the linear approximation to the true parameter value) was not within the 95% ETPIs of the estimate about half of the time. Nonetheless, the estimates tended to be correct qualitatively, that is the direction and approximate magnitude of the mean selection differential and environmental effect were correctly estimated. Future work could allow for additional selection functions, but even then, quantitative estimates of selection on traits and genetic loci will only be strictly interpretable when the function assumed matches (or very closely approximates) reality. Estimates of selection function parameters cannot be combined across functions as these parameters do not have a consistent meaning, but Bayesian model averaging for derived parameters is possible.

Perhaps even more important from a practical perspective, the estimates of selection, or at least the environment-dependent component of selection, depend on correctly identifying the environmental covariate responsible for fluctuations in selection. In some cases, such as negative or positive frequency-dependent selection, this might be relatively easy. And a recent meta-analysis suggests that some aspects of the environment, most notably precipitation, might generally be important drivers of variation in selection (Siepielski et al., 2017). Nonetheless, it will often be difficult to designate a single environmental covariate a priori. Alternative models with different covariates could be fit and evaluated based on their posterior predictive distributions. Indeed, the relatively poor fit of the population density model to the *C. maculatus* data set was uncovered based on the posterior predictive distribution (unfortunately, I lack a clear, alternative covariate to explain variability in change in the mean polygenic score over time in this data set). However, caution is warranted as considering many possible environmental covariates may uncover spurious relationships. When many covariates are tested, it would be prudent to divide the data into training and validation sets for cross-validation. Lastly, in some cases variation in selection likely results from the interaction of multiple environmental factors (e.g. Benkman et al., 2010; Benkman & Siepielski, 2004; Nosil et al., 2018). I intend to extend the method and FSABC software to allow for multiple covariates and their interactions in the future.

4.3 | Performance and the genotype-phenotype map

Uncertainty in trait genetic architecture had a modest effect on the performance of the method with simulated data, but had a profound effect on the analysis of the *C. maculatus* experiment. At least two factors likely contributed to this discrepancy. First, the simulations involved 1000 genetic markers each having a modest probability of being associated with the trait (all PIPs = 0.1), whereas more than 10,000 SNPs had small but non-zero probabilities of association with female weight in *C. maculatus* (minimum = 0.001, maximum = 0.030, median = 0.004). Thus, the *C. maculatus* data exhibited much greater uncertainty in the genotype-phenotype map than the simulated data. Uncertainty in real data sets will be a function of sample size, which tends to be low except in human studies and some model systems. Second, the method assumes the posterior inclusion probabilities (PIPs) across genetic markers are independent (i.e. all information arises from the marginal probabilities of association). This assumption was valid for the simulated data, but was likely violated for *C. maculatus* data set. With the Bayesian sparse linear mixed model approach used to estimate genotype-phenotype associations for *C. maculatus* weight, sets of SNPs in the model (i.e. with non-zero effects) in any given Markov chain Monte Carlo (MCMC) step are unlikely to be associated with the same causal variant (or in high LD with each other) (Zhou et al., 2013). Instead, on average each causal variant should be represented by ~1 SNP marker in any given MCMC step. Information on sets of SNPs (not just individual SNPs) is not

encoded in the PIPs. Thus, the marginal association probabilities jet-tison information from the full, joint posterior and sampling sets of SNPs for the proposed method based on these could result in poor inference in cases where the marginal posteriors are misleading.

I see several ways to overcome the problem of uncertainty in the genotype-phenotype map. First, additional large-scale experiments could be conducted, including genetic manipulations or fine-scale genetic mapping, to refine an initial set of candidate genetic markers or QTL regions (e.g. Barrett et al., 2019; Linnen et al., 2013). Second, when sets of linked SNPs are associated with a trait, all but the most strongly associated could be dropped and the sum of the posterior inclusion probabilities across the set could be assigned to the retained SNP (this could be repeated for different sets of traits). This should work best when genetic markers fall into multiple, easy to delineate, sets of trait-associated markers (which was not the case for the *C. maculatus* data set; Rêgo et al., 2020). Third, the method could be modified to work with samples from the joint posterior for genotype-phenotype associations. Such information can be extracted from the MCMC output of common polygenic GWA models (including GEMMA; Zhou et al., 2013). I plan to implement this final approach in a future version of FSABC. Short of this, my results suggest that using model-averaged effect estimates is likely the better solution when dealing with considerable uncertainty in the genotype-phenotype map. At minimum, this approach allowed me to detect selection in *C. maculatus*.

An additional key consideration of the genotype-phenotype map concerns the choice of a mapping population. I expect the proposed method to perform best when the mapping population harbours the same genetic variants segregating at the same frequencies, and the same phenotypic variance, as the populations included in the population-genomic time series. Thus, mapping in those actual populations will often be the best choice. Indeed, differences between the genetic composition of the *C. maculatus* mapping population (a BC between the source population and the lentil-adapted line) and the focal line during the course of the experiment likely further limit quantitative interpretation of the magnitude of selection on weight during the experiment (more generally, mapping populations generated from crosses might often violate the linkage and Hardy-Weinberg equilibrium assumptions for equating average effects with average phenotypic excess). Of course, if the genetic composition of the populations changes considerably over the course of the study (perhaps because of sustained directional selection) or if the phenotype is highly plastic, a single genotype-phenotype map may be limiting regardless of the mapping populations (e.g. Coop, 2019). In such cases, it should be possible to include multiple genotype-phenotype maps (for different populations, generations or environments). Doing so would allow selection and the genotype-phenotype map to depend on the environment and would thus facilitate the analysis of plastic traits. I hope to include this in a future version of FSABC.

Accurate inference of selection requires substantial LD between genetic markers and the causal variants affecting traits and fitness. Trait-mapping experiments might often be biased towards common variants relative to the set of causal variants, and these two sets of

loci can exhibit different patterns of LD and evolutionary dynamics (e.g. Gompert et al., 2014; Mathieson & McVean, 2012). This is especially likely if mapping populations are genotyped using a SNP panel identified from a small sample of individuals. Even when direct sequencing is used, small sample sizes can limit power to detect and analyse rare variants. Moreover, rare variants, including rare structural variants, might be enriched among true causal variants for traits affecting fitness (e.g. Bomba et al., 2017; Chakraborty et al., 2019; Cirulli & Goldstein, 2010; Josephs et al., 2017; Park et al., 2011). Thus, care should be taken to capture rare variants to the extent possible in the mapping population, and even then, caution is warranted as it can be hard to determine how well the population genetic properties for a set of genetic markers and the (unknown) true causal variants match.

5 | CONCLUSIONS

Combining population genomic time-series data with data on trait genetics and the environment has the potential to advance our understanding of how selection acts in nature and drives evolution across the genome. I hope that the proposed method contributes to this aim. I expect the biggest limitation at present for the use of this (or related) method is a lack of appropriate data. Large-scale spatiotemporal population-genomic data sets are still uncommon, especially for natural populations. However, advances in the analysis of ancient or museum DNA and the expanding field of evolve-and-resequence experiments may help overcome this limitation (e.g. Barghi et al., 2017; Bi et al., 2013; Cridland et al., 2018; Rêgo et al., 2019; Sproul & Maddison, 2017). Moreover, few systems currently combine such sampling with detailed studies of trait genetics and the environment. Nonetheless, I hope that methods, like the one proposed here, help illustrate the potential value of such combined data sets, and more generally of long-term studies in the age of genomics (see, e.g., Grant & Grant, 2014).

ACKNOWLEDGMENTS

This manuscript was improved by comments from A. Bergland and C. Nice. This research was funded by the National Science Foundation (DEB-1638768 and DEB-1844941 to ZG). The support and resources from the Center for High Performance Computing at the University of Utah are also gratefully acknowledged.

AUTHOR CONTRIBUTIONS

ZG designed the study, developed the proposed method, wrote the computer software, generated and analysed the simulated data, analysed the *C. maculatus* data, and wrote and revised the manuscript.

DATA AVAILABILITY STATEMENT

The FSABC software, manual and an example data set are available from GITHUB (<https://github.com/zgompert/fsabc.git>). Additional scripts and simulated data sets analysed in this manuscript are available from Dryad (<https://doi.org/10.5061/dryad.pvmcndkh>).

ORCID

Zachariah Gompert  <https://orcid.org/0000-0003-2248-2488>

REFERENCES

- Agrawal, A. A. (2020). A scale-dependent framework for trade-offs, syndromes, and specialization in organismal biology. *Ecology*, 101, e02924.
- Agrawal, A. A., Conner, J. K., & Rasmann, S. (2010). Tradeoffs and negative correlations in evolutionary ecology. In M. A. Bell D. J. Futuyma W. F. Eanes & J. S. Levinton (Eds.), *Evolution since Darwin: The first 150 years* (pp. 243–268). Sinauer Associates Inc.
- Anderson, J. T., Lee, C. R., Rushworth, C. A., Colautti, R. I., & Mitchell-Olds, T. (2013). Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology*, 22, 699–708.
- Barghi, N., Tobler, R., Nolte, V., & Schlotterer, C. (2017). *Drosophila simulans*: A species with improved resolution in evolve and resequence studies. *G3: Genes, Genomes, Genetics*, 7, 2337–2343.
- Barrett, R. D. H., Laurent, S., Mallarino, R. et al (2019). Linking a mutation to survival in wild mice. *Science*, 363, 499–504.
- Beaumont, M. A. (2010). Approximate Bayesian Computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41, 379–406.
- Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings Of the Royal Society B-Biological Sciences*, 263, 1619–1626.
- Beaumont, M. A., Zhang, W. Y., & Balding, D. J. (2002). Approximate Bayesian computation in population genetics. *Genetics*, 162, 2025–2035.
- Bell, G. (2010). Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365, 87–97.
- Benkman, C. W., Parchman, T. L., & Mezquida, E. T. (2010). Patterns of coevolution in the adaptive radiation of crossbills. *Annals Of the New York Academy of Sciences*, 1206, 1–16.
- Benkman, C. W., & Siepielski, A. M. (2004). A keystone selective agent? pine squirrels and the frequency of serotiny in lodgepole pine. *Ecology*, 85, 2082–2087.
- Berg, J. J., & Coop, G. (2014). A population genetic signal of polygenic adaptation. *PLoS Genetics*, 10, e1004412.
- Bergland, A. O., Behrman, E. L., O'Brien, K. R., Schmidt, P. S., & Petrov, D. A. (2014). Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in *Drosophila*. *PLoS Genetics*, 10, e1004775.
- Bi, K., Linderth, T., Singhal, S. et al (2019). Temporal genomic contrasts reveal rapid evolutionary responses in an alpine mammal during recent climate change. *PLoS Genetics*, 15, e1008119.
- Bi, K., Linderth, T., Vanderpool, D., Good, J. M., Nielsen, R., & Moritz, C. (2013). Unlocking the vault: Next-generation museum population genomics. *Molecular Ecology*, 22, 6018–6032.
- Blum, M. G. B., & Francois, O. (2010). Non-linear regression models for Approximate Bayesian Computation. *Statistics and Computing*, 20, 63–73.
- Bomba, L., Walter, K., & Soranzo, N. (2017). The impact of rare and low-frequency genetic variants in common disease. *Genome Biology*, 18, 1–17.
- Buffalo, V., & Coop, G. (2019). The linked selection signature of rapid adaptation in temporal genomic data. *Genetics*, 213, 1007–1045.
- Buffalo, V., & Coop, G. (2020). Estimating the genome-wide contribution of selection to temporal allele frequency change. *Proceedings of the National Academy of Sciences*, 117, 2067220680.
- Campbell-Staton, S. C., Winchell, K. M., Rochette, N. C., Fredette, J., Maayan, I., Schweizer, R. M., & Catchen, J. (2020). Parallel selection on thermal physiology facilitates repeated adaptation of city lizards to urban heat islands. *Nature Ecology & Evolution*, 4, 652–658.

- Chakraborty, M., Emerson, J., Macdonald, S. J., & Long, A. D. (2019). Structural variants exhibit widespread allelic heterogeneity and shape variation in complex traits. *Nature Communications*, 10, 1–11.
- Cirulli, E. T., & Goldstein, D. B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. *Nature Reviews Genetics*, 11, 415–425.
- Coop, G. (2016). Does linked selection explain the narrow range of genetic diversity across species? *BioRxiv*, 042598.
- Coop, G. (2019). Reading tea leaves? polygenic scores and differences in traits among groups. *arXiv Preprint arXiv:1909.00892*.
- Coop, G., Witonsky, D., Di Rienzo, A., & Pritchard, J. K. (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics*, 185, 1411–1423.
- Cridland, J. M., Ramirez, S. R., Dean, C. A., Sciligo, A., & Tsutsui, N. D. (2018). Genome sequencing of museum specimens reveals rapid changes in the genetic composition of honey bees in California. *Genome Biology and Evolution*, 10, 458–472.
- Csillery, K., Francois, O., & Blum, M. G. B. (2012). abc: An R package for approximate bayesian computation (ABC). *Methods in Ecology and Evolution*, 3(3), 475–479.
- De Villemereuil, P., & Gaggiotti, O. E. (2015). A new FST-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*, 6, 1248–1258.
- Do, C., Waples, R. S., Peel, D., Macbeth, G., Tillett, B. J., & Ovenden, J. R. (2014). NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Molecular Ecology Resources*, 14, 209–214.
- Edge, M. D., & Coop, G. (2019). Reconstructing the history of polygenic scores using coalescent trees. *Genetics*, 211, 235–262.
- Endler, J. A. (1977). *Geographic variation, speciation, and clines*. Princeton University Press.
- Endler, J. A. (1986). *Natural selection in the wild*. Princeton University Press.
- Endler, J. A. (1988). Frequency-dependent predation, crypsis and aposematic coloration. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 319, 505–523.
- Ewens, W. J. (2004). *Mathematical population genetics: I. Theoretical introduction*, vol. 27. Springer Science & Business Media.
- Feder, A. F., Kryazhimskiy, S., & Plotkin, J. B. (2014). Identifying signatures of selection in genetic time series. *Genetics*, 196, 509–522.
- Field, Y., Boyle, E. A., Telis, N., Gao, Z., Gaulton, K. J., Golan, D., Yengo, L., Rocheleau, G., Froguel, P., McCarthy, M. I., & Pritchard, J. K. (2016). Detection of human adaptation during the past 2000 years. *Science*, 354, 760–764.
- Fisher, R. A. (1941). Average excess and average effect of a gene substitution. *Annals of Eugenics*, 11, 53–63.
- Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180, 977–993.
- Ford, E. B. (1977). *Ecological genetics*. Springer.
- Ford, S. A., Williams, D., Paterson, S., & King, K. C. (2017). Co-evolutionary dynamics between a defensive microbe and a pathogen driven by fluctuating selection. *Molecular Ecology*, 26, 1778–1789.
- Gillespie, J. H. (1991). *The causes of molecular evolution*. Oxford University Press.
- Gillespie, J. (2004). *Populations genetics: A concise guide* (2nd ed.). Johns Hopkins University Press.
- Gingerich, P. D. (2019). *Rates of evolution: A quantitative synthesis*. Cambridge University Press.
- Gómez, P., & Buckling, A. (2011). Bacteria-phage antagonistic coevolution in soil. *Science*, 332, 106–109.
- Gompert, Z. (2016). Bayesian inference of selection in a heterogeneous environment from genetic time-series data. *Molecular Ecology*, 25, 121–134.
- Gompert, Z., Brady, M., Chalyavi, F. et al (2019). Genomic evidence of genetic variation with pleiotropic effects on caterpillar fitness and plant traits in a model legume. *Molecular Ecology*, 28, 2967–2985.
- Gompert, Z., Egan, S. P., Barrett, R. D., Feder, J. L., & Nosil, P. (2017). Multilocus approaches for the measurement of selection on correlated genetic loci. *Molecular Ecology*, 26, 365–382.
- Gompert, Z., Lucas, L. K., Buerkle, C. A., Forister, M. L., Fordyce, J. A., & Nice, C. C. (2014). Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular Ecology*, 23, 4555–4573.
- Gompert, Z., & Messina, F. J. (2016). Genomic evidence that resource-based trade-offs limit host-range expansion in a seed beetle. *Evolution*, 70, 1249–1264.
- Grant, P. R., & Grant, B. R. (2002). Unpredictable evolution in a 30-year study of Darwin's finches. *Science*, 296, 707–711.
- Grant, P. R., & Grant, B. R. (2014). *40 years of evolution: Darwin's finches on Daphne Major Island*. Princeton University Press.
- Hahn, M. W. (2008). Toward a selection theory of molecular evolution. *Evolution*, 62, 255–265.
- Hall, A. R., Scanlan, P. D., Morgan, A. D., & Buckling, A. (2011). Host-parasite coevolutionary arms races give way to fluctuating selection. *Ecology Letters*, 14, 635–642.
- Hallsson, L. R., & Björklund, M. (2012). Selection in a fluctuating environment leads to decreased genetic variation and facilitates the evolution of phenotypic plasticity. *Journal of Evolutionary Biology*, 25, 1275–1290.
- Hedrick, P. W. (2006). Genetic polymorphism in heterogeneous environments: the age of genomics. *Annual Review in Ecology, Evolution and Systematics*, 37, 67–93.
- Hendry, A. P. (2016). *Eco-evolutionary dynamics*. Princeton University Press.
- Hendry, A. P., & Kinnison, M. T. (1999). Perspective: The pace of modern life: Measuring rates of contemporary microevolution. *Evolution*, 53(6), 1637–1653.
- Hughes, K. A., Houde, A. E., Price, A. C., & Rodd, F. H. (2013). Mating advantage for rare males in wild guppy populations. *Nature*, 503, 108–110.
- Illingworth, C. J. R., & Mustonen, V. (2011). Distinguishing driver and passenger mutations in an evolutionary history categorized by interference. *Genetics*, 189, 989–1000.
- Jorde, P. E., & Ryman, N. (2007). Unbiased estimator for genetic drift and effective population size. *Genetics*, 177, 927–935.
- Josephs, E. B., Berg, J. J., Ross-Ibarra, J., & Coop, G. (2019). Detecting adaptive differentiation in structured populations with genomic data and common gardens. *Genetics*, 211, 989–1004.
- Josephs, E. B., Stinchcombe, J. R., & Wright, S. I. (2017). What can genome-wide association studies tell us about the evolutionary forces maintaining genetic variation for quantitative traits? *New Phytologist*, 214, 21–33.
- Kelly, J. K., & Hughes, K. A. (2019). Pervasive linked selection and intermediate-frequency alleles are implicated in an evolve-and-resequencing experiment of *Drosophila simulans*. *Genetics*, 211, 943–961.
- Kimura, M., & Crow, J. F. (1978). Effect of overall phenotypic selection on genetic change at individual loci. *Proceedings of the National Academy of Sciences*, 75, 6168–6171.
- King, J. G., & Hadfield, J. D. (2019). The evolution of phenotypic plasticity when environments fluctuate in time and space. *Evolution Letters*, 3, 15–27.
- Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M. et al (2001). The strength of phenotypic selection in natural populations. *The American Naturalist*, 157, 245–261.
- Kinnison, M. T., & Hendry, A. P. (2001). The pace of modern life II: From rates of contemporary microevolution to pattern and process. In M. T. Kinnison & A.P. Hendry (Eds.), *Microevolution rate, pattern, process* (pp. 145–164). Springer.

- Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychonomic Bulletin & Review*, 25, 178–206.
- Lande, R., & Arnold, S. (1983). The measurement of selection on correlated characters. *Evolution*, 37, 1210–1226.
- Leffler, E. M., Bullaughey, K., Matute, D. R., Meyer, W. K., Ségurel, L., Venkat, A., Andolfatto, P., & Przeworski, M. (2012). Revisiting an old riddle: What determines genetic diversity levels within species? *PLoS Biology*, 10, e1001388.
- Linnen, C. R., Poh, Y. P., Peterson, B. K., Barrett, R. D. H., Larson, J. G., Jensen, J. D., & Hoekstra, H. E. (2013). Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, 339, 1312–1316.
- Lucas, L. K., Nice, C. C., & Gompert, Z. (2018). Genetic constraints on wing pattern variation in *Lycaeides* butterflies: A case study on mapping complex, multifaceted traits in structured populations. *Molecular Ecology Resources*, 18, 892–907.
- Mathieson, I., & McVean, G. (2012). Differential confounding of rare and common variants in spatially structured populations. *Nature Genetics*, 44, 243–246.
- Mathieson, I., & McVean, G. (2013). Estimating selection coefficients in spatially structured populations from time series data of allele frequencies. *Genetics*, 193, 973–984.
- Messer, P. W., Ellner, S. P., & Hairston, N. G. (2016). Can population genetics adapt to rapid evolution? *Trends in Genetics*, 32, 408–418.
- Messina, F. J., & Durham, S. L. (2015). Loss of adaptation following reversion suggests trade-offs in host use by a seed beetle. *Journal of Evolutionary Biology*, 28, 1882–1891.
- Messina, F. J., & Jones, J. C. (2011). Inheritance of traits mediating a major host shift by a seed beetle, *Callosobruchus maculatus* (Coleoptera: Chrysomelidae: Bruchinae). *Annals of the Entomological Society of America*, 104, 808–815.
- Messina, F. J., Mendenhall, M., & Jones, J. C. (2009). An experimentally induced host shift in a seed beetle. *Entomologia Experimentalis Et Applicata*, 132, 39–49.
- Mueller, J. C., Carrete, M., Boerno, S., Kuhl, H., Tella, J. L., & Kempnaers, B. (2020). Genes acting in synapses and neuron projections are early targets of selection during urban colonization. *Molecular Ecology*, 29, 3403–3412.
- Mueller, L. D., Barr, L. G., & Ayala, F. J. (1985). Natural selection vs. random drift: Evidence from temporal variation in allele frequencies in nature. *Genetics*, 111, 517–554.
- Nosil, P., Egan, S. P., & Funk, D. J. (2008). Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. *Evolution*, 62, 316–336.
- Nosil, P., Villoutreix, R., de Carvalho, C. F., Farkas, T. E., Soria-Carrasco, V., Feder, J. L., Crespi, B. J., & Gompert, Z. (2018). Natural selection and the predictability of evolution in *Timema* stick insects. *Science*, 359, 765–770.
- Park, J. H., Gail, M. H., Weinberg, C. R., Carroll, R. J., Chung, C. C., Wang, Z., Chanock, S. J., Fraumeni, J. F., & Chatterjee, N. (2011). Distribution of allele frequencies and effect sizes and their interrelationships for common genetic susceptibility variants. *Proceedings Of the National Academy of Sciences*, 108, 18026–18031.
- Pritchard, J. K., Pickrell, J. K., & Coop, G. (2010). The genetics of human adaptation: Hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, 20, R208–R215.
- Racimo, F., Berg, J. J., & Pickrell, J. K. (2018). Detecting polygenic adaptation in admixture graphs. *Genetics*, 208, 1565–1584.
- Rêgo, A., Chaturvedi, S., Springer, A., Lish, A. M., Barton, C. L., Kapheim, K. M., Messina, F. J., & Gompert, Z. (2020). Combining experimental evolution and genomics to understand how seed beetles adapt to a marginal host plant. *Genes*, 11, 400.
- Rêgo, A., Messina, F. J., & Gompert, Z. (2019). Dynamics of genomic change during evolutionary rescue in the seed beetle *Callosobruchus maculatus*. *Molecular Ecology*, 28, 2136–2154.
- Reich, D. (2018). *Who we are and how we got here: Ancient DNA and the new science of the human past*. Oxford University Press.
- Reilstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24, 43484370.
- Reznick, D. N., Shaw, F. H., Rodd, F. H., & Shaw, R. G. (1997). Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science*, 275, 1934–1937.
- Robinson, M. R., Hemani, G., Medina-Gomez, C., Mezzavilla, M., Esko, T., Shakhbazov, K., Powell, J. E., Vinkhuyzen, A., Berndt, S. I., Gustafsson, S., Justice, A. E., Kahali, B., Locke, A. E., Pers, T. H., Vedantam, S., Wood, A. R., van Rheenen, W., Andreassen, O. A., Gasparini, P., ... Visscher, P. M. (2015). Population genetic differentiation of height and body mass index across Europe. *Nature Genetics*, 47, 1357–1362.
- Sabeti, P. C., Reich, D. E., Higgins, J. M., Levine, H. Z. P., Richter, D. J., Schaffner, S. F., Gabriel, S. B., Platko, J. V., Patterson, N. J., McDonald, G. J., Ackerman, H. C., Campbell, S. J., Altshuler, D., Cooper, R., Kwiatkowski, D., Ward, R., & Lander, E. S. (2002). Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 419, 832–837.
- Shi, H., Kichaev, G., & Pasaniuc, B. (2016). Contrasting the genetic architecture of 30 complex traits from summary association data. *The American Journal of Human Genetics*, 99, 139–153.
- Siepielski, A. M., DiBattista, J. D., & Carlson, S. M. (2009). It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecology Letters*, 12, 1261–1276.
- Siepielski, A. M., Gotanda, K. M., Morrissey, M. B., Diamond, S. E., DiBattista, J. D., & Carlson, S. M. (2013). The spatial patterns of directional phenotypic selection. *Ecology Letters*, 16, 1382–1392.
- Siepielski, A. M., Morrissey, M. B., Buoro, M., Carlson, S. M., Caruso, C. M., Clegg, S. M., Coulson, T., DiBattista, J., Gotanda, K. M., Francis, C. D., Hereford, J., Kingsolver, J. G., Augustine, K. E., Kruuk, L. E. B., Martin, R. A., Sheldon, B. C., Sletvold, N., Svensson, E. I., Wade, M. J., & MacColl, A. D. C. (2017). Precipitation drives global variation in natural selection. *Science*, 355, 959–962.
- Sisson, S. A., Fan, Y., & Beaumont, M. (2018). *Handbook of approximate Bayesian computation*. Chapman and Hall/CRC.
- Speidel, L., Forest, M., Shi, S., & Myers, S. R. (2019). A method for genome-wide genealogy estimation for thousands of samples. *Nature Genetics*, 51, 1321–1329.
- Sprout, J. S., & Maddison, D. R. (2017). Sequencing historical specimens: successful preparation of small specimens with low amounts of degraded dna. *Molecular Ecology Resources*, 17, 1183–1201.
- Szpiech, Z. A., & Hernandez, R. D. (2014). *selscan*: an efficient multi-threaded program to perform EHH-based scans for positive selection. *Molecular Biology and Evolution*, 31, 2824–2827.
- Takahashi, Y., Yoshimura, J., Morita, S., & Watanabe, M. (2010). Negative frequency-dependent selection in female color polymorphism of a damselfly. *Evolution*, 64, 3620–3628.
- Thompson, J. N. (2013). *Relentless evolution*. University of Chicago Press.
- Tuda, M., Kagoshima, K., Toquenaga, Y., & Arnqvist, G. (2014). Global genetic differentiation in a cosmopolitan pest of stored beans: Effects of geography, host-plant usage and anthropogenic factors. *PLoS One*, 9, e106268.
- Tufto, J. (2015). Genetic evolution, plasticity, and bet-hedging as adaptive responses to temporally autocorrelated fluctuating selection: a quantitative genetic model. *Evolution*, 69, 2034–2049.
- Turchin, M. C., Chiang, C. W. K., Palmer, C. D., Sankararaman, S., Reich, D., & Hirschhorn, J. N. (2012). Evidence of widespread selection on standing variation in europe at height-associated SNPs. *Nature Genetics*, 44, 1015.

- Turelli, M., Schemske, D. W., & Bierzychudek, P. (2001). Stable two-allele polymorphisms maintained by fluctuating fitnesses and seed banks: protecting the blues in *Linanthus parryae*. *Evolution*, 55, 1283–1298.
- Walsh, B., & Lynch, M. (2018). *Evolution and Selection of Quantitative Traits*. Oxford University Press.
- Wittmann, M. J., Bergland, A. O., Feldman, M. W., Schmidt, P. S., & Petrov, D. A. (2017). Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift. *Proceedings Of the National Academy of Sciences of the United States of America*, 114, E9932–E9941.
- Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., & Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42, 565–569.
- Zhou, X., Carbonetto, P., & Stephens, M. (2013). Polygenic modeling with Bayesian sparse linear mixed models. *PLoS Genetics*, 9, e1003264.

- Zhou, X., & Stephens, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nature Genetics*, 44, 821–824.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Gompert Z. A population-genomic approach for estimating selection on polygenic traits in heterogeneous environments. *Mol Ecol Resour.* 2021;21:1529–1546. <https://doi.org/10.1111/1755-0998.13371>