

Detecting and Quantifying Changing Selection Intensities from Time-Sampled Polymorphism Data

Hyunjin Shim,^{*,†} Stefan Laurent,^{*,†} Sebastian Matuszewski,^{*,†} Matthieu Foll,^{*,†,*}
and Jeffrey D. Jensen^{*,†}

^{*}School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland, [†]Swiss Institute of Bioinformatics (SIB), 1015 Lausanne, Switzerland, and [‡]International Agency for Research on Cancer Genetic Cancer Susceptibility, International Agency for Research on Cancer, 69372, Lyon, France

ORCID ID: 0000-0002-7052-0971 (H.S.)

ABSTRACT During his well-known debate with Fisher regarding the phenotypic dataset of *Panaxia dominula*, Wright suggested fluctuating selection as a potential explanation for the observed change in allele frequencies. This model has since been invoked in a number of analyses, with the focus of discussion centering mainly on random or oscillatory fluctuations of selection intensities. Here, we present a novel method to consider nonrandom changes in selection intensities using Wright-Fisher approximate Bayesian (ABC)-based approaches, in order to detect and evaluate a change in selection strength from time-sampled data. This novel method jointly estimates the position of a change point as well as the strength of both corresponding selection coefficients (and dominance for diploid cases) from the allele trajectory. The simulation studies of this method reveal the combinations of parameter ranges and input values that optimize performance, thus indicating optimal experimental design strategies. We apply this approach to both the historical dataset of *P. dominula* in order to shed light on this historical debate, as well as to whole-genome time-serial data from influenza virus in order to identify sites with changing selection intensities in response to drug treatment.

KEYWORDS

fluctuating selection
change point analysis
time-sampled data
approximate Bayesian computation
Wright-Fisher model

The common assumption of constant selection intensity through time utilized in many tests of selection is often criticized as unrealistic in natural and experimental populations, both owing to environmental changes (e.g., fluctuations in climate, predator density, or nutrient availability) as well as to genetic changes (e.g., epistasis, clonal interference). Despite this, such considerations are not accounted for in most population genetic models, since inferring changing selection coefficients (s) from single-time point polymorphism data are difficult. However, owing to recent technological advances, time-sampled polymorphism data are increasingly available, and time-serial analytical

methods are expanding (Malaspinas *et al.* 2012; Mathieson and McVean 2013; Foll *et al.* 2014; Lacerda and Seoighe 2014; and see review of Bank *et al.* 2014), allowing for an empirical evaluation of the importance of changing s models.

Fluctuating selection in natural populations was suggested by Wright (1948) with regards to the phenotypic time-serial data of *Panaxia dominula* (scarlet tiger moth) to account for its observed annual fluctuations in coloring polymorphism (Fisher and Ford 1947). Since, there have been several theoretical considerations of fluctuating selection (Kimura 1954; Karlin and Levikson 1974; Karlin and Lieberman 1974; Chevin 2013; Gossmann *et al.* 2014; Gompert 2015), as well as many observations of fluctuating selection in natural populations (for a review, see Bell 2010). Nonetheless, until recently, analyses of fluctuating selection centered on *random* or *seasonal* oscillations of selection strength through time, as the mathematical complexity of analytical methods only allowed the simplest cases to be considered.

Approximate Bayesian Computation (ABC) has the advantage of being flexible in integrating complex models due to computational efficiency and the lack of likelihood computation (Beaumont 2010). Recently, a hierarchical ABC-based method based on the Wright-Fisher model was developed in order to infer genome-wide effective

Copyright © 2016 Shim *et al.*

doi: 10.1534/g3.115.023200

Manuscript received September 29, 2015; accepted for publication January 28, 2016; published Early Online February 10, 2016.

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Supplemental Material is available online at www.g3journal.org/lookup/suppl/doi:10.1534/g3.115.023200/-/DC1

[†]Corresponding author: Ecole Polytechnique Fédérale de Lausanne SV IBI-SV Ujpsen, AAB 0 45 (Bâtiment AAB), Station 15, CH-1015 Lausanne, Switzerland. E-mail: hyunjin.shim@epfl.ch

population size and per-site selection coefficients from whole-genome multiple-time point datasets (Foll *et al.* 2014, 2015). While the initial approach performs well overall, the authors noted the possibility for observations inconsistent with a single-*s* Wright-Fisher model; this was indeed observed at certain sites in their analysis of the influenza virus genome, which were simply excluded from consideration. Thus, as a natural extension, we here investigate the presence of changing selection intensities in these outlier single nucleotide polymorphisms (SNPs); in doing so, we also develop an extended Wright-Fisher ABC-based method capable of detecting and quantifying changing selection intensities through time (termed CP-WFABC).

METHODS

Wright-Fisher ABC-based method

This approach first relies on a previously developed Wright-Fisher ABC method (WFABC - Foll *et al.* 2014, 2015) in order to estimate a constant effective population size (N_e). WFABC infers an effective population size from time-serial datasets through the average fluctuation of neutral trajectories over time (following Wright 1948). If whole-genome multi-time point datasets are available, the estimated value is referred to as the genome-wide effective population size. The posterior of the N_e estimated from WFABC is used as a prior for the following extended method. The trajectory X of a given allele with a known N_e consists of time-serial allele frequencies f_t ($t = 1, \dots, T$) where T is the total number of generations (with $T > 4$ to allow for a change point to be realizable with the Wright-Fisher model), from which a sample n_m is taken at sampling time points $m = 1, \dots, M$ (with $M > 4$ to allow for a change point to be detectable). Parameters to be inferred include the selection coefficient prior to the change in selection intensity (s_1), the selection coefficient subsequent to the change in selection intensity (s_2), the change point in selection (CP), and the dominance coefficient (h) for diploid models. The joint posterior distribution of these parameters can be estimated by

$$P(s_1, s_2, CP, h|X) \propto P(X|s_1, s_2, CP, h)P(s_1)P(s_2)P(CP)P(h), \quad (1)$$

$$P(s_1, s_2, CP|X) \propto P(X|s_1, s_2, CP)P(s_1)P(s_2)P(CP) \quad (2)$$

for the diploid model and the haploid model, respectively. The ABC approach allows these parameters to be inferred using Wright-Fisher model simulations without calculating the likelihood $P(X|s_1, s_2, CP, h)$ or $P(X|s_1, s_2, CP)$.

The Wright-Fisher model simulator with a change point in selection strength is used to simulate the data X , with relative fitnesses $w_{AA} = 1 + s$, $w_{Aa} = 1 + sh$ and $w_{aa} = 1$ for the diploid model, and $w_A = 1 + s$ and $w_a = 1$ for the haploid model (Ewens 2004). Initially, the random sampling of an allele from generation 1 to generation $CP-1$ is simulated using s_1 , and onwards from the change point (CP) using s_2 . In order to simulate realistic allele trajectories with changing selection coefficients, the allele needs to be segregating at the time of the change point. This condition is necessary since the change in selection coefficient cannot occur if the allele is either lost or fixed beforehand, assuming the infinite-site model with no back mutations. Thus, only alleles segregating at the change point are accepted as simulated datasets.

The associated summary statistic for these time-serial data are Fs' , an unbiased estimator of N_e based on the allele frequency change between two sampling time points that remains unbiased even in the presence of highly skewed allele frequencies and small sample size (Jorde and

Ryman 2007), an important property for inferring change points along the allele trajectory. It is given as:

$$Fs = \frac{(f_m - f_{m+1})^2}{f_i(1 - f_i)} \quad (3)$$

$$Fs' = \frac{1}{t_i} \frac{Fs \left[1 - \frac{1}{2\tilde{n}} \right] - \frac{2}{\tilde{n}}}{\left(1 + \frac{Fs}{4} \right) \left[1 - \frac{1}{n_{m+1}} \right]} \quad (4)$$

where f_m and f_{m+1} are the allele frequencies at two consecutive sampling time points, m and $m+1$ (separated by t_i generations), $f_i = (f_m + f_{m+1})/2$, and \tilde{n} is the harmonic mean of the chromosome sample sizes n_m and n_{m+1} at two consecutive time points. Unlike the WFABC approach that summarizes time-serial trajectories into only two summary statistics (increasing and decreasing Fs' ; Foll *et al.* 2014), here Fs' is summarized at every pair of consecutive time points as Fs'_i with $i = 1, \dots, M-1$, where M is the number of sampling time points. This modification allows additional information, such as the timing of increase or decrease in allele frequency, to be captured. In order to retain information about directionality that is lost by using Fs' as a summary statistic in the original WFABC approach, increasing allele frequencies are made positive and decreasing allele frequencies are made negative with regards to the absolute value.

The joint posterior distribution of the parameters of interest is obtained using the algorithm described in Beaumont *et al.* (2002). The approximate posterior density

$$P(\theta|U(X)) \approx P(\theta|X) \quad (5)$$

with $U(X)$ as data-specific summary statistics, $\theta = (s_1, s_2, CP, h)$ for the diploid model and $\theta = (s_1, s_2, CP)$ for the haploid model, is obtained using an ABC algorithm as follows:

- i. Simulate K trajectories from the Wright-Fisher model with a change in selection intensity, with θ randomly sampled from its prior $P(\theta)$, conditional on the allele segregating at the change point.
- ii. Compute $U(x_k)$ for each trajectory using the Fs'_i summary statistic between all consecutive sampling time points i : $U(x_{k,i})$ where $i = 1, \dots, M-1$ where M is the last sampling time point.
- iii. Retain the 1000 best simulations with the smallest Euclidian distance between $U(x_{k,i})$ (from the simulated) and $U(X_i)$ (from the observed) to obtain an approximate posterior density of $P(\theta|X)$.

For the first step, simulations are performed with the same initial conditions as the observed data, including effective population size, initial allele frequency, and the sampling points and sizes. In addition, a minimum allele frequency in one of the sampling time points is imposed on simulated trajectories as is done in observed data. This ascertainment scheme takes into account the nonrandom criterion of considering only the trajectories reaching values above the sequencing error threshold in the observed data (Foll *et al.* 2015).

For the second step, it is important to note that the Fs' summary statistic is calculated between every pair of consecutive sampling time points (Figure 1). This construction of the summary statistic enables information on both the timing and strength of the allele frequency change to be captured, as the timing of the change is essential in detecting the change point and the strength of the change is essential

s1=0.2, s2=-0.2, CP=50

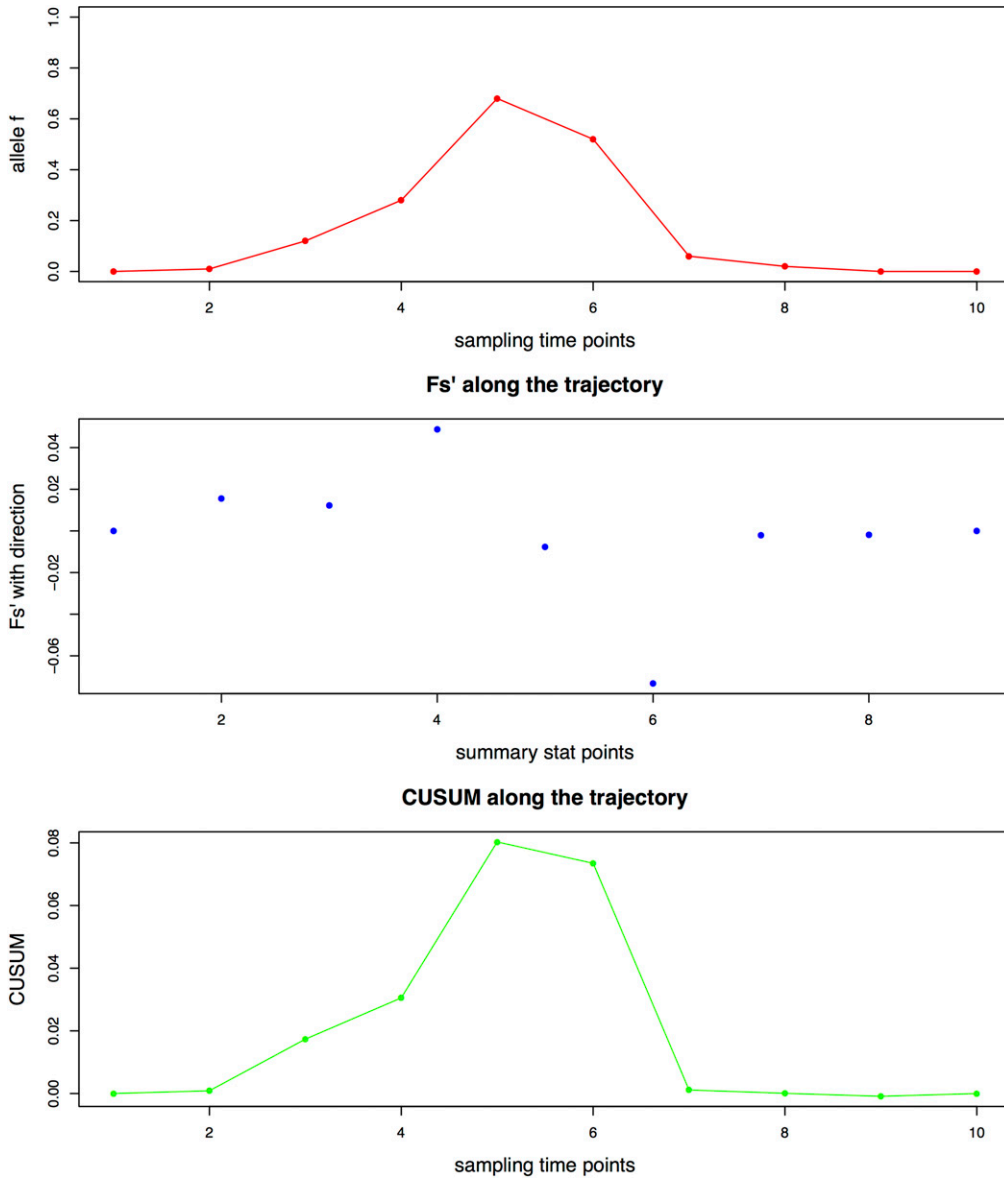


Figure 1 Illustration of F_s' calculated between every pair of consecutive sampling time points with 10 generations and the maximal CUSUM (cumulative sum control chart) value S_{CP} as summary statistics, using a haploid population of $N_e = 1000$ with a *de novo* mutation and the sample size as 100. (A) Allele frequency trajectory (B) F_s' values along the trajectory (C) CUSUM values along the trajectory

in estimating the corresponding selection coefficients. For the diploid model, an additional parameter h is inferred jointly with the other three parameters, as its value is one of the determining factors in the timing of allele frequency change (Haldane 1932).

For the third step, the simulated F_s' summary statistics $U(x_{k,i})$ between every pair of consecutive sampling time points are compared with the corresponding observed F_s' summary statistics $U(X_i)$ – allowing a small fraction of the simulated trajectories with allele frequency changes that best match the observed trajectory (in terms of both timing and strength) to be retained.

Wright-Fisher ABC-based method with change point analysis

In order to increase computational efficiency and sensitivity in change point detection, an additional summary statistic is integrated into the Wright-Fisher ABC-based method. This novel summary statistic is derived from change point analysis – statistical techniques

developed and used in many disciplines ranging from finance to quality control in order to detect and estimate change (e.g., Chen and Gupta 2001). Among the techniques available, the cumulative sum control chart (CUSUM) developed by Page (1954) is able to detect small and sustained shifts in the statistics β obtained from a sample (Ryan 2011). Instead of using the entire CUSUM procedure as a separate method for detecting change, the CUSUM value is integrated into the Wright-Fisher ABC-based method as an additional summary statistic that characterizes the time-sampled trajectory of an allele:

$$S_i = S_{i-1} + (\beta_i - \bar{\beta}), \quad i = 1, \dots, M-1 \quad (6)$$

where $\bar{\beta} = \text{mean}$ and $S_0 = 0$. The CUSUM value S is accumulated only when the statistic β is different from its average value in the dataset.

The change point S_{CP} is the sampling time point with the maximal absolute value of S_i , which is the furthest point from the initial value

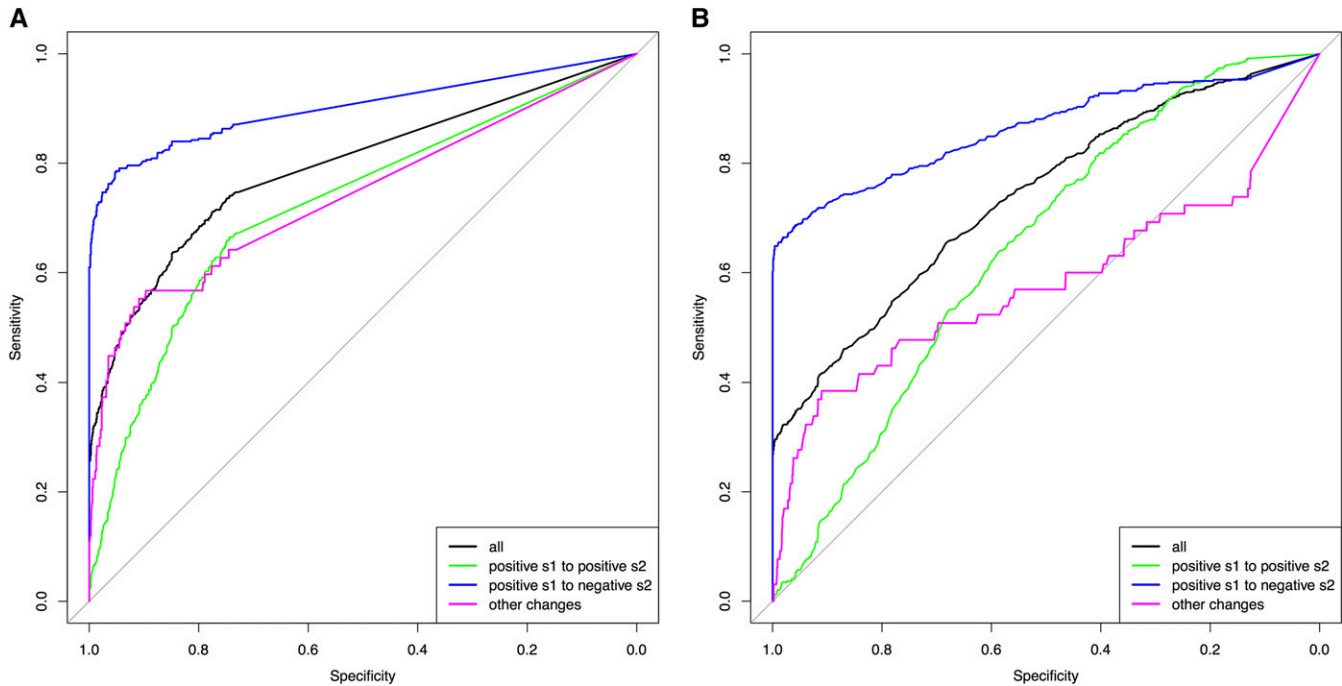


Figure 2 16. ROC (Receiver operating characteristic) curve of the Bayes factor $B_{1,0}$ from the ABC (approximate Bayesian computation) model choice (A) a haploid population with $N_e = 100$, and (B) a diploid population with $N_e = 500$.

zero attaining the maximal accumulation of difference from the average value:

$$S_{CP} = \arg \max_{i=0, \dots, I} |S_i|. \quad (7)$$

Here, we take Fs' at each pair of consecutive sampling time points as the statistic β , since it is a time-serial measure of the allele frequency change, which is indicative of the selection strength change. Thus, when Fs' is used as the statistic β in the CUSUM, the maximal CUSUM value S_{CP} is the potential change point of the allele trajectory, as illustrated with an example in Figure 1.

In the change point Wright-Fisher ABC (CP-WFABC), an additional summary statistic S_{CP} is used to characterize observed and simulated allele frequency trajectories for detecting a change point. In the third step of the ABC algorithm, the Euclidean distance between $U(x_{k,i})$ and $U(X_i)$ is calculated only if the maximal CUSUM value $S_{CP,k}$ of the simulated data matches the maximal CUSUM value S_{CP} of the observed data:

$$D = \begin{cases} \|U(X_i) - U(x_{k,i})\|, & \text{if } S_{CP} = S_{CP,k} \\ \infty, & \text{otherwise.} \end{cases} \quad (8)$$

This additional step allows the computation to be more efficient – especially when there is a large number of time points sampled – as the Euclidean distance is calculated for a fraction of simulated trajectories whose maximal CUSUM value is equal to that of the observed (*i.e.*, with the same time-sampled characteristic). Furthermore, as the CUSUM is sensitive to small and sustained changes, integrating the CUSUM into the Wright-Fisher ABC method increases its sensitivity for detecting small and sustained changes in selection strength. The potential bias in the calculation of the maximal CUSUM value is counteracted by the fact that the bias would be present in both the observed and the simulated trajectories.

Simulated data with constant selection and with changing selection

We generated simulated datasets of different effective population sizes using the Wright-Fisher model for two scenarios: (1) trajectories of constant selection with only s and h (for diploid models) as parameters, and (2) trajectories of changing selection with $s1$, $s2$, CP , and h (for diploid models) as parameters. For selection coefficients, uniform priors of $(-1, 1)$ were used. The uniform prior of CP was set to occur between the second generation and the second-to-last generation $(2, T-1)$, where T is the number of generations of the population in the time-serial data. The dominance coefficient h for the diploid model was randomly drawn from one of three values: complete recessiveness, codominance, or complete dominance $[0, 0.5, 1]$. Although these prior ranges are uninformative, the constraint on the trajectories to be segregating at the change point shapes the distribution of the prior ranges according to the input parameters such as ploidy, effective population size, initial allele frequency, and number of generations; the updated priors for the haploid population of $N_e = 100$ and the diploid population of $N_e = 50$ are shown as examples (Supplemental Material, Figure S1 and Figure S2).

The other input values – such as the number of generations ($T = 100$), the sampling time points ($M = 10$), the sample size ($n = 100$), the initial allele arising as a new mutation, and the ascertainment of observing a minimum frequency at 2% – were kept constant for the two scenarios. We retained the best 0.1% of 1,000,000 simulations for each pseudoobservable trajectory using the rejection algorithm based on the Euclidean distance as described above. The mode of the posterior distribution from the best simulations (Sunnåker *et al.* 2013) was used to evaluate the estimated parameter value against the true parameter value.

Data availability

The R package of CP-WFABC is available on jensenlab.epfl.ch.

■ **Table 1** AUC values and confidence intervals for ROC curves

ROC (Receiver operating characteristic) curves			
	100	1000	10000
AUC	0.7936	0.7988	0.7943
CI	0.7756–0.8124	0.7797–0.8181	0.7750–0.8140
ROC curves for haploid populations			
	50	500	5000
AUC	0.8233	0.7378	0.7470
CI	0.8051–0.8402	0.7164–0.7593	0.7257–0.7683

ROC, ; AUC, area under the ROC curve; CI, confidence interval.

RESULTS

Model choice in the change point Wright-Fisher ABC method

The first step of CP-WFABC is to be able to distinguish changing selection trajectories from constant selection trajectories. ABC model choice was constructed to choose between two models: M_0 with a single selection coefficient, and M_1 with two selection coefficients and a change point. The relative probability of M_1 over M_0 can be computed through the model posterior ratio as the Bayes factor $B_{1,0}$ (Sunnåker *et al.* 2013):

$$\frac{p(M_1|D)}{p(M_0|D)} = \frac{p(D|M_1)p(M_1)}{p(D|M_0)p(M_0)} = B_{1,0} \frac{p(M_1)}{p(M_0)} \quad (9)$$

when the model prior $p(M_0)$ is equal to $p(M_1)$. In practice, the model priors are made equal by producing the same number of simulations for each model and retaining the best simulations from the lot. The posterior ratio is computed as the number of accepted simulations from M_1 over those of M_0 , giving the Bayes factor $B_{1,0}$, which is an indicator of the support for a specific model. The performance study was conducted with a haploid population of $N_e = (100, 1000, \text{ or } 10,000)$ and a diploid population with $N_e = (50, 500, \text{ or } 5000)$ using the simulated datasets of the two scenarios described in the previous section as M_0 and M_1 , respectively.

We considered two cases for the pseudoobservables to test the sensitivity and specificity of the ABC model choice: the first case when the pseudoobserved trajectories have a single selection coefficient, and the second case when they have changing selection coefficients. One thousand pseudoobservable trajectories were generated for each case with the data ascertainment minimum frequency set to 2% for at least one of the sampling time points. Additionally, for the second case, pseudoobservable trajectories were accepted only when the allele was segregating at the time of the change point – a constraint for realistic combinations of selection coefficients, change points, and dominance (for diploids) – in order to reproduce changing selection trajectories in real datasets. All other input values were kept constant as in the simulated datasets described in the previous section.

The results of the ABC model choice from a haploid population with $N_e = 100$, and a diploid population with $N_e = 500$, are represented as ROC curves (Xavier *et al.* 2011) in Figure 2. Specificity is given on the x -axis showing the true negative rate, while sensitivity is given on the y -axis showing the true positive rate of the Bayes factor $B_{1,0}$ calculated from 1000 pseudoobservables of changing selection (where $B_{1,0}$ should be large) and 1000 pseudoobservables of constant selection (where $B_{1,0}$ should be small). The overall ROC curves in black (all trajectories) show that when the specificity threshold is most conservative in detecting no false positives (*i.e.*, $B_{1,0} = \text{infinite}$), the Bayes factor $B_{1,0}$ has a sensitivity of $\sim 30\%$ for all populations. Considering that

■ **Table 2** The Bayes factor (BF) thresholds for the significance level α of 1% computed using the distribution of Bayes factors under the null model M_0

Diploid populations			
	50	500	5000
BF	4.7	1.6	1.3
Haploid populations			
	100	1000	10000
BF	3.7	3.2	1.7

the pseudoobservable trajectories were simulated randomly from a wide range of prior values, the Bayes factor $B_{1,0}$ from CP-WFABC is, in general, sensitive and specific. The ROC curves in black for the other haploid and diploid populations (Figure S3) also indicate that the Bayes factor $B_{1,0}$ is sensitive and specific as they are above the diagonal line of no-discrimination. The area under the ROC curve (AUC) is used to assess how reflective the Bayes factor $B_{1,0}$ is of the true model, as summarized for all pseudoobservable populations in Table 1. The AUC values show that the Bayes factor $B_{1,0}$ is $\sim 80\%$ more probable to rank a randomly chosen changing selection case above a randomly chosen constant selection case. Additionally, the distribution of Bayes factors $B_{1,0}$ under the null model M_0 (*i.e.*, case 1) was used to compute the significance level α at 1% (Good 1992). For both diploids and haploids, the significance threshold is higher for smaller population sizes (Table 2), and the calculation of these thresholds will be important in any given data application.

Following the detection of changing selection trajectories using ABC model choice, the quality of parameter estimation by the model chosen was evaluated. The cross-validation results from the haploid population of $N_e = 100$ are shown in Figure 3 and those from the diploid population of $N_e = 500$ in Figure 4 (see Figure S4, Figure S5, Figure S6, and Figure S7 for additional results). For the case where the pseudoobservables were of constant selection, the estimation for a single s (and the dominance h for diploid) using CP-WFABC is very accurate, as the mode of the best simulations from the M_0 model for each pseudoobservable lies along the red diagonal line. Exceptions include uninformative trajectories where the allele surpasses the minimum frequency of 2% in the first sampling and is lost immediately due to genetic drift or negative selection and therefore not observed in subsequent samplings. Such uninformative trajectories will always keep the same set of best simulations from the M_0 model since their selection strength is indistinguishable, and they result in horizontal lines along the estimated negative value. This phenomenon is particularly pronounced when population size is small as shown in Figure 3 and Figure S6, since the role of genetic drift is more significant.

For the second case, when the pseudoobservables are of changing selection intensity, the joint estimation of the parameters is also effective for a restricted range of values. In Figure 3, Figure 4, Figure S4, Figure S5, Figure S6, and Figure S7, the mode estimation of each pseudoobservable is color-coded according to the three categories of trajectory shape. The green dots are pseudoobservable trajectories that change from positive s_1 to positive s_2 . The blue dots are those that change from positive s_1 to negative s_2 , while the magenta colors include all other cases (*e.g.*, neutral or negative s_1 to any value of s_2). There is a clear clustering by category – with the best estimation being of positive values of s_1 below 0.5, moderate values of s_2 between -0.5 and 0.5 , and CP values for the blue category of positive s_1 to negative s_2 . In trajectories other than those with positive s_1 to negative s_2 , the change point is

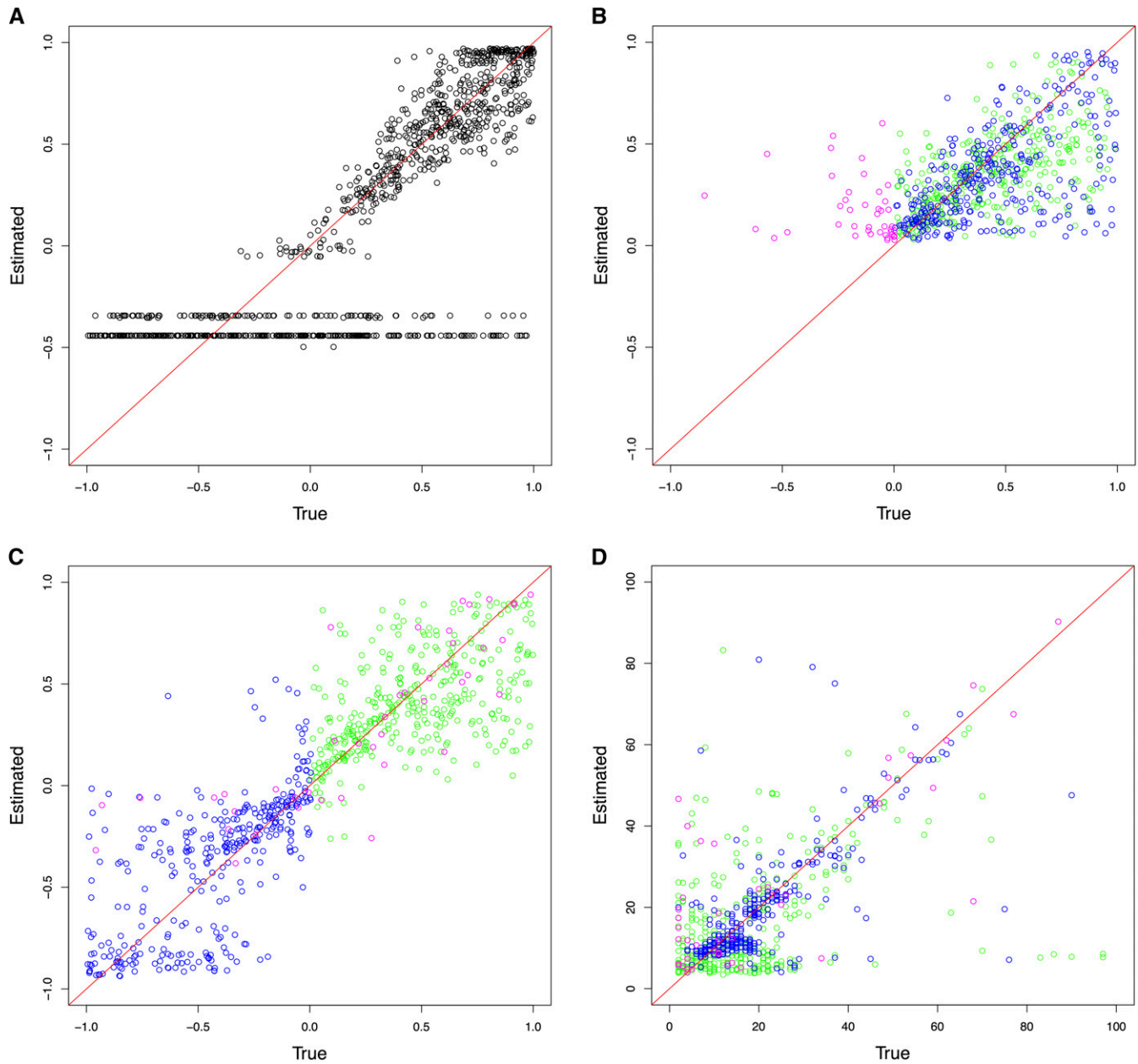


Figure 3 ABC (approximate Bayesian computation) model choice parameter estimations for 1000 pseudoobservables with a haploid population of $N_e = 100$. Each circle is the mode of the posterior distribution from the 0.1% best simulations. Case 1: Pseudoobservables with constant selection (A) M_0 estimates of s . Case 2: Pseudoobservables with changing selection (B) M_1 estimates of s_1 , (C) M_1 estimates of s_2 , and (D) M_1 estimates of CP (green: positive s_1 to positive s_2 , blue: positive s_1 to negative s_2 , magenta: other cases)

difficult to detect, particularly for diploid populations where the additional dominance parameter h was estimated (Figure 4, Figure S6, and Figure S7). This trend is also observed when the ROC curves are generated according to these three categories (Figure 2 and Figure S3). For all populations, the Bayes factor $B_{1,0}$ is more sensitive and specific for trajectories changing from positive s_1 to negative s_2 (ROC curves in blue), reaching above 60% of the true positive rate when there are no false positives. Despite the restricted range of good parameter estimation in s_1 , s_2 , and CP , the estimation of dominance is robust for both cases of constant and changing selection (except for the small population size of $N_e = 50$; Figure S6).

In order to evaluate the performance of the joint parameter estimation, the coefficient of determination R^2 is used to assess the cross-

validation between the estimated values (y_m) and the true values (f_m), compared with the simple average of the estimated values (\bar{y}). The closer the R^2 value is to 1, the better the parameter estimation as shown:

$$R^2 \equiv 1 - \frac{\sum_m (y_m - f_m)^2}{\sum_m (y_m - \bar{y})^2} \quad (10)$$

Table 3 and Table 4 summarize the performance of the joint parameter estimation for all cases as the R^2 values for the haploid and diploid populations, respectively. The first case is when the pseudoobservables are of constant selection intensity, in which case the true model (M_0) performs only slightly better than the false model (M_1) for

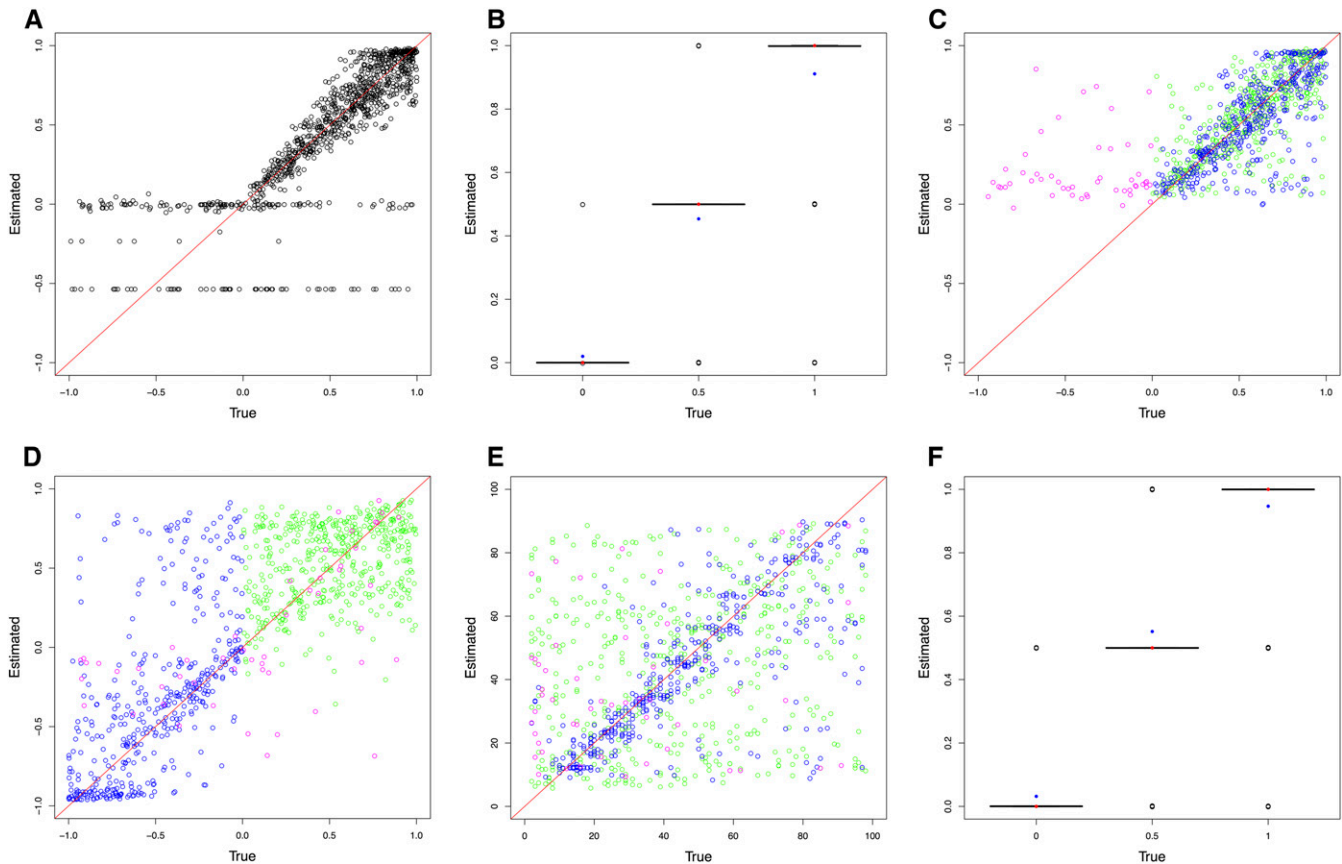


Figure 4 ABC (approximate Bayesian computation) model choice parameter estimations for 1000 pseudoobservables with a diploid population of $N_e = 500$. For cross-validation graphs, each circle is the mode of the posterior distribution from the 0.1% best simulations. For boxplots, red dots are true values and blue dots are average estimated values. Case 1: pseudoobservables with constant selection (A) M_0 estimates of s and (B) M_0 estimates of h . Case 2: pseudoobservables with changing selection (C) M_1 estimates of s_1 , (D) M_1 estimates of s_2 , (E) M_1 estimates of CP , and (F) M_1 estimates of h (green: positive s_1 to positive s_2 , blue: positive s_1 to negative s_2 , magenta: other cases)

estimating s . This discrepancy in parameter estimation of M_0 is mainly owing to uninformative pseudoobservable trajectories with constant selection (which have been lost or fixed) being associated with the true model (M_0) of constant selection; this is due to the constraint for the allele to be segregating at the change point in the (false) model M_1 of changing selection. In the cross-validation of the constant selection case, the parameters estimated form horizontal lines at negative estimated values for those trajectories that are lost, and cluster at the top right corner for those trajectories that are fixed (Figure 3, Figure 4, Figure S4, Figure S5, Figure S6, and Figure S7).

For the second case in which the pseudoobservables have changing selection coefficients, parameter estimation from the true model (M_1) performs better than that from the false model (M_0) for all parameters – particularly when population sizes are large. As expected, there is a trend of better parameter estimation as population size increases (though see Figure S4 and Figure S5). Additionally, it has been shown that the value of the Bayes factor $B_{1,0}$ is a good indicator of the parameter estimation performance (results not shown).

DATA APPLICATION

Historical dataset of *P. dominula*

A long-running dataset based on the *medionigra* morph responsible for darker wing color in wild populations of *P. dominula* (Figure S8) began in 1939 with collections by Fisher and Ford (1947) and continued

through to 1999 (Cook and Jones 1996; Jones 2000). Despite this phenotypic time-series data having been analyzed previously from various angles (O’Hara 2005; Mathieson and McVean 2013; Foll *et al.* 2015), it is still relevant to consider a model of changing selection in time, as Wright (1948) originally suggested.

The recent reconsiderations of the dataset tend to favor a lethal-recessive model with an effective population of $N_e = 500$ (Mathieson and McVean 2013; Foll *et al.* 2015) – however, the biological question of how the *medionigra* morph could have reached the initial frequency of 11% in the dataset remains unanswered with this conclusion of constant strong negative selection. Wright asserted that the trajectory of the *medionigra* morph during this period could be explained by fluctuating selection with “no net selective advantage or disadvantage.” Although this alternative hypothesis has been considered as a random fluctuation of selection by estimating selection coefficients between every sampling time point (see O’Hara 2005), the quantitative plausibility of a directional change-in- s model over a single- s model lacks thorough investigation. Thus, we reanalyzed this dataset using the CP-WFABC method in order to investigate the possibility of changing selection in the *medionigra* morph during the 60-yr data collection.

Using the ABC model choice introduced here as a test for a change in selection strength, and to estimate the parameters of interest for the chosen model, we assume the *medionigra* allele is a single codominant locus responsible for the homozygous and heterozygous expressions of the phenotypic forms *bimacula* and *medionigra*, respectively (Cook and

■ **Table 3** R^2 values of the parameter estimation with the ABC (approximate Bayesian computation) model choice for haploid populations

Case 1: 1000 constant selection pseudoobservables			
M_0 (True)	100	1000	10,000
s	0.558	0.843	0.914
M_1 (False)	100	1000	10,000
s_1	0.366	0.769	0.834
s_2	0.608	0.752	0.685
CP	-16	-5.97	-2.27
Case 2: 1000 changing selection pseudoobservables			
M_0 (False)	100	1000	10000
s	-0.278	-0.00232	-0.00205
M_1 (True)	100	1000	10,000
s_1	-0.323	0.363	0.415
s_2	0.704	0.768	0.777
CP	0.0946	0.413	0.303

Jones 1996). The model M_0 assumes a single selection coefficient, thus the only parameter to estimate is s . The M_1 model assumes a change in selection strength, thus the parameters of interest are s_1 , s_2 , and CP . Both M_0 and M_1 take the prior range of $(-1, 1)$ for the selection coefficients and the prior range of $(2, 59)$ for the change point in the M_1 model. For the M_1 model, these uninformative priors are updated with the constraint that the allele must be segregating at the time of change point. Here, we create 10^7 simulated datasets for each M_0 and M_1 , and apply the rejection algorithm of the CP-WFABC method to retain the best 1000 simulations compared with the observed trajectory. The effective population size is assumed to be $N_e = 500$ as in previous studies (Wright 1948; Cook and Jones 1996; O'Hara 2005), with an initial allele frequency of 11% and a minimum frequency ascertainment of 2%.

The Bayes factor for M_1 over M_0 is calculated as 0.952, indicating that the single coefficient M_0 model cannot be rejected in favor of the changing selection M_1 model (Table 4). From the parameter estimation of the model M_0 (Figure S9), the mode of the posterior distribution for s is given as -0.15 as asserted by Fisher and Ford (1947). When the ABC model choice was repeated with a smaller population size of $N_e = 50$ as suggested by Wright (1948) and O'Hara (2005), the Bayes factor increases to 1.87 (*i.e.*, the changing selection model is twice as likely as the constant selection) – however, this value is not large enough to be significant for a diploid population of $N_e = 50$ (Table 4).

Experimental evolution of influenza virus with drug treatment

The evolution of pathogens within a host is one of the most important cases in which the possibility of fluctuating selection must be considered – as they may experience drastically changing selective pressures due to host immune response, specific drug treatments, and/or pathogenic cooperation or competition (Tanaka and Valckenborgh 2011; Hall *et al.* 2011). Thus, how these pathogens adapt to these rapid external and internal changes is of major concern to the biomedical community.

The time-serial experimental dataset of influenza A conducted by Renzette *et al.* (2014) and Foll *et al.* (2014) is an interesting case study on the impact of drug treatment on influenza virus evolution. The dataset consists of 13 sampling points from which population-level whole-genome data were collected. Drug treatment with a commonly used neuraminidase inhibitor (oseltamivir) began after the collection of the third sample and continued, at increasing concentrations, until the final passage. Using WFABC, the genome-wide effective population

■ **Table 4** R^2 values of the parameter estimation with the ABC (approximate Bayesian computation) model choice for diploid populations

Case 1: 1000 constant selection pseudoobservables			
M_0 (True)	50	500	5000
s	-0.13	0.55	0.796
h	-0.789	0.653	0.88
M_1 (False)	50	500	5000
s_1	-0.194	0.363	0.746
s_2	0.491	0.57	0.609
h	0.195	0.764	0.857
CP	-10.8	-1.65	-1.37
Case 2: 1000 changing selection pseudoobservables			
M_0 (False)	50	500	5000
s	-0.441	0.0955	0.118
h	-0.119	0.525	0.573
M_1 (True)	50	500	5000
s_1	-0.713	0.181	0.277
s_2	0.467	0.536	0.463
h	0.219	0.611	0.678
CP	-0.145	-0.258	-0.125

size across the sampling time points was estimated ($N_e = 176$) and the SNPs under selection were identified.

Here, we apply the CP-WFABC method on two cases of interest from this study to consider a possible change in selection strength under drug treatment: the first case includes trajectories identified as being driven by positive selection, while the second includes outlier trajectories (*i.e.*, trajectories not fitting a single s Wright-Fisher model). For all cases, we test the model M_0 (*i.e.*, a single selection coefficient) and M_1 (*i.e.*, a changing selection coefficient), with parameters of interest (s) and (s_1 , s_2 , CP), respectively. The number of generations per passage is assumed to be 13 (following Foll *et al.* 2014), and the minimum frequency of 2% is set as an ascertainment for observing the minor allele in the data. *De novo* mutations are assumed to occur at the sampling time point where the SNP was first sampled (except for trajectories whose initial sample frequency was first sampled above $1/N_e$, which are assumed to be standing variation). This assumption is based on the high mutation rate, large population bottlenecks associated with passaging, and large census population size between passages. A total of 10^7 datasets were simulated for each M_0 and M_1 ; the best 1000 trajectories from the lot were retained using the rejection algorithm described in the *Methods* section. The uniform prior ranges for the selection coefficients were set as $(-1, 1)$, for the change point as $(t_0, 157)$ where t_0 is the generation when the SNP was first sampled, and with the constraint of segregating alleles at the change point for M_1 .

The results for the Bayes factors and the parameter estimates are summarized in Table 5 for all trajectories of interest. The Bayes factors of most trajectories show strong support for the changing selection model: the stronger the selection strength change, the larger the Bayes factor. Using the Bayes factors from the simulation studies as guidance (Table 4), the significance threshold to reject M_0 is computed as 3.7 for a small haploid population. As expected, the trajectories identified as outliers of the single s Wright-Fisher model (NP 159, PB1 33) all reject the constant selection model M_0 with a large Bayes factor. We also note that the Bayes factor for the drug-resistant mutation H275Y (NA 823) does not support the changing selection model strongly, confirming that the experimental evolution procedure kept the selective pressure of the drug constant by adjusting the drug concentration to reduce viral plaque numbers to 50% at each passage. As shown in Figure 5, the

■ **Table 5 Bayes factors and parameters estimated for the influenza trajectories in the presence and absence of drug**

Segment	Position	Bayes Factor (M1/M0)	M0 s estimate	M1 s1 estimate	M1 s2 estimate	M1 CP estimate	WFABC s estimate
Trajectories under selection with drug							
PA ^a	2194	∞	—	0.251	-0.328	24 (p1-2)	0.09
HA ^b	48	42.5	0.11	0.031	0.309	102(p7-8)	0.14
HA ^b	1395	5.5	0.142	0.088	0.188	104 (p7-8)	0.22
NA ^b	582	999	0.127	0.019	0.41	102 (p7-8)	0.29
NA ^b	823	1.8	0.114	0.103	0.139	105 (p8-9)	0.15
M ^a	147	2.7	0.056	0.042	0.089	116 (p8-9)	0.08
NS ^b	820	1.7	0.037	0.029	0.05	146 (p11-12)	0.12
Outlier trajectories with drug							
NP ^b	159	8.09	0.011	0.027	-0.054	103 (p7-8)	0
PB1 ^b	33	16.2	0.047	0.023	0.256	113 (p8-9)	0.14
Trajectories under selection without drug							
PB1 ^a	1119	0.57	0.044	0.038	0.047	60 (p4-5)	0.06
HA ^b	1395	2.38	0.097	0.083	0.131	106 (p8-9)	0.12
NP ^a	1104	1.72	0.047	0.042	0.062	110 (p8-9)	0.05
NP ^b	1396	0.89	0.049	0.049	0.053	49 (p3-4)	0.09

The estimates whose Bayes factors show strong support for M₁ are in bold.

^a Standing variation.

^b A *de novo* mutation.

change points are estimated to be mostly between the seventh and eighth passages, a notable result since three of these trajectories (HA 48, HA 1395, NA 582) are increasing rapidly after the drug-resistant mutation H275Y appears, whereas one trajectory (NP 159) from a different segment decreases rapidly. This result may indicate that positive selective for the three SNPs (including HA 1395; a known compensatory mutation encoded also as D112N) increased along with the drug-resistant mutation H275Y, potentially due to epistatic interactions, whereas another SNP (NP 159) decreased at that time, potentially owing to clonal interference. The single selection estimates from WFABC (Foll *et al.* 2015) are similar to the estimates of CP-WFABC only when the Bayes factor does not reject M₀ – strong evidence that an alternative model of changing selection must be considered for some trajectories in order to correctly estimate selection coefficients.

We also applied CP-WFABC to the control case of SNPs increasing in frequency without drug as a comparison to the case with drug. The effective population sizes of the viral populations were averaged to be 226 in the absence of drug from the previous study (Foll *et al.* 2014). All of the other inputs and the ABC model choice conditions were kept the same as in the drug case. The Bayes factor results summarized in Table 5 demonstrate that none of the SNP trajectories under selection in the control experiment can reject M₀ (*i.e.*, constant selection).

DISCUSSION

These simulations demonstrate that the novel CP-WFABC approach presented here is able to detect changing selection trajectories via ABC model choice, and also to estimate a wide range of parameters of interest. Performance was analyzed separately for three categories of allele trajectories according to the nature of the change in selection strength: (1) a change from positive s_1 to positive s_2 , (2) a change from positive s_1 to negative s_2 , and (3) all other changes. The datasets for each possible combination were generated using the Wright-Fisher model with a change in selection strength, using the most general prior ranges for all parameters s_1 , s_2 , CP, and h for diploids, with the only constraint being segregation of the allele at the change point. For both the detection and parameter estimation, CP-WFABC performs best when the change is large, particularly for the second category of change (positive s_1 to negative s_2), as shown in the ROC curves (Figure 2 and Figure S3) and the cross-validation graphs (Figure 3, Figure 4, Figure S4, Figure S5, Figure S6, and Figure S7). For the first

category (positive s_1 to positive s_2) and the third category (any other changes), the change point is difficult to estimate, particularly for diploids where the additional parameter h is also estimated. The ABC model choice of CP-WFABC has the best sensitivity for full specificity, for larger population sizes ($N_e > 500$ for diploids), and for haploid populations.

The parameter estimates of s_1 and s_2 perform best when the values are moderate – given as the intervals (0, 0.5) and (-0.5, 0.5) respectively. For s_1 , the optimal parameter range for estimation is (0, 0.5), where a *de novo* mutation that survives negative selection and segregates until the change point is indistinguishable from other drifting mutations with similar trajectories with uninformative low allele frequency. These trajectories naturally arise more frequently when population size is small and in diploids where the dominance effect plays a role, as shown in the third category of change (Figure 3, Figure 4, Figure S4, Figure S5, Figure S6, and Figure S7; magenta points). When an initial frequency of 10% is used instead of a *de novo* mutation, the advantage of having a more informative trajectory at the beginning is counteracted by the effect of more cases dominated by negative selection or genetic drift segregating until the change point. Thus, the performance of CP-WFABC for standing variation is similar to that of *de novo* mutation. For s_2 , the optimal parameter range for estimation is (-0.5, 0.5), as trajectories with extreme values are less informative since they are lost or fixed directly after the change point (explaining the clustering of the change points at earlier times). For diploid populations, estimates of h are accurate to the level of determining dominance from codominance or recessiveness, particularly for large population sizes (Figure S7), given the difficulty of joint estimation with the three other parameters. Indeed, estimation of this additional parameter comes at the cost of worse performance for the other parameters, as can be seen in the ROC curves and cross-validation graphs: the detection and parameter estimation of changing selection cases is always better for haploids. Thus, in diploid cases, we recommend fixing the dominance parameter if known, in order to improve the performance of CP-WFABC.

Although CP-WFABC is intended to detect and evaluate changing selection intensities, the simulation studies demonstrate that the method also performs well in estimating parameters for cases of constant selection (as has been demonstrated by Foll *et al.* 2014). For haploids with large population sizes, in particular, the estimated values of the

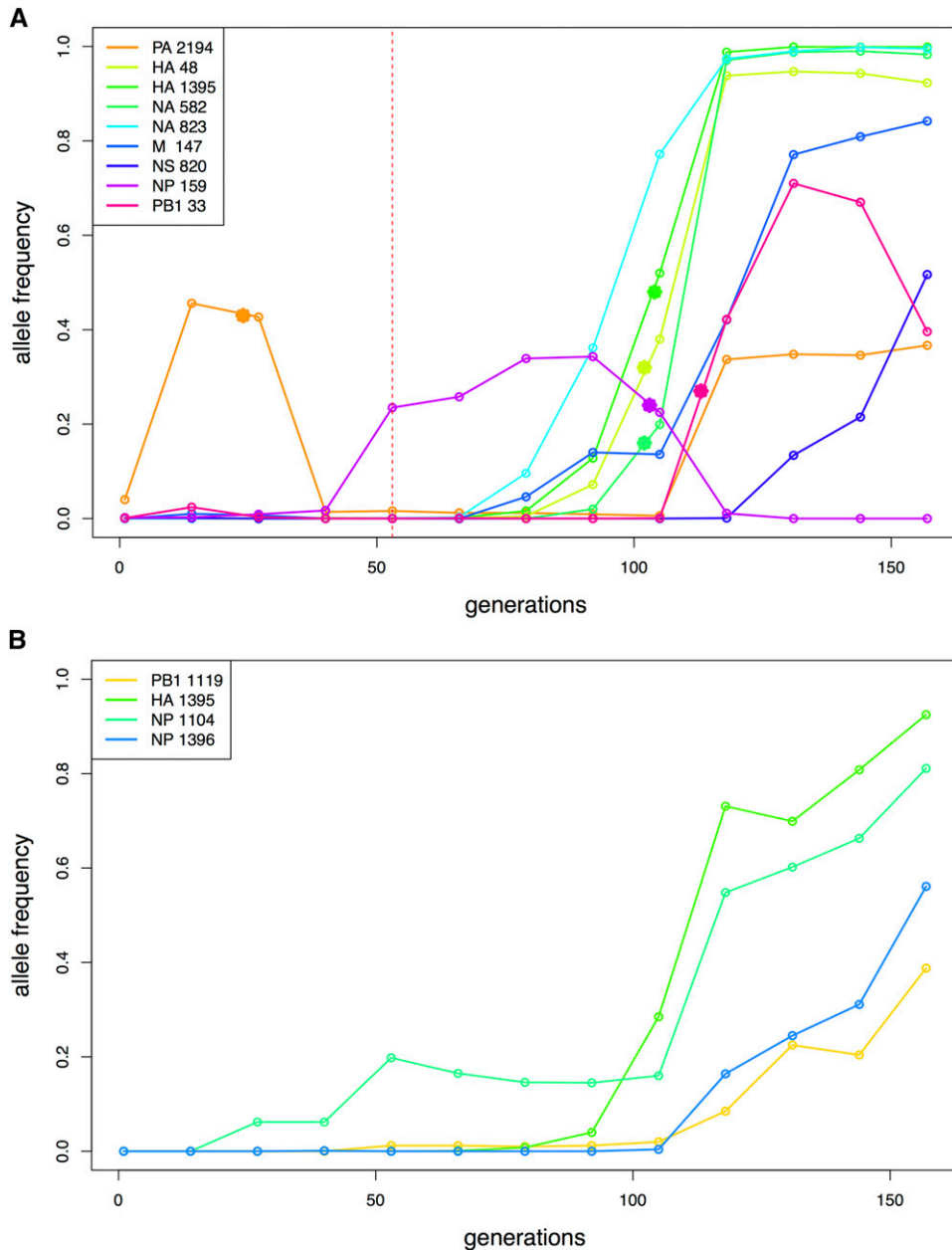


Figure 5 Change points indicated with solid stars for the trajectories of interest. (A) Increasing SNP trajectories in the presence of drug. The red vertical line indicates the sampling time of drug administration. (B) Increasing SNP trajectories in the absence of drug.

single parameter s correlate almost perfectly with the true values (Figure S5). However, when the population size is small for both diploids and haploids, some trajectories that are lost by genetic drift are difficult to estimate, as shown in Figure 3 and Figure S6 as horizontal lines along negative estimated values. This limitation of constant selection coefficients, however, is due to the simulation conditions of the *de novo* mutation at the first generation and the minimal ascertainment scheme (minimum frequency of 2% at one of the sampling time points). For real datasets, the conditions are likely to be less stringent, since such uninformative trajectories will not be considered for parameter estimation.

Finally, we utilized this approach to make inference in two very different time-sampled datasets: *P. dominula* (diploid) and Influenza A virus (haploid). The time-serial *medionigra* trajectory of *P. dominula* was reanalyzed to test for a change in selection strength and/or direction during 60 yr of data collection. By assuming h as codominant,

the results for $N_e = 500$ indicate that the model M_0 of constant selection cannot be rejected according to the Bayes factor from the ABC model choice algorithm. The selection coefficient from this model is estimated as -0.15 , corresponding with that calculated by Fisher and Ford (1947). However, when the population size is assumed to be smaller ($N_e = 50$), the Bayes factor result supports M_1 (changing selection) twice as strongly as M_0 (constant selection), but not strong enough to reject M_0 according to the significance level test computed with the distribution of Bayes factors under the null model M_0 . This dataset of the *medionigra* morph thus demonstrates the difficulty of detecting and evaluating a change in selection when the assumption of a known and constant effective population size is not fulfilled. Depending on the assumed population sizes, the conclusions may indeed change, as this study demonstrates. Furthermore, if the population size fluctuates, this method has the caveat of detecting a change in N_e^*s rather than only in selection strength.

Next, CP-WFABC was applied to SNP trajectories of interest from an experimental dataset of influenza A virus. The ABC model choice test was conducted on the trajectories identified as 1) being positively selected and 2) as outliers from the single-*s* WFABC method. For the SNPs in the presence of drug, the Bayes factor for six out of nine trajectories favored the changing selection model M_1 . As shown in Figure 5, the change points for four out of these six trajectories occurred between passages seven and eight – the interval during which three trajectories from the Hemagglutinin and Neuraminidase segments (HA and NA, respectively) increased rapidly while one trajectory from the segment NP decreased rapidly along with the known drug-resistant mutation NA 823 (H275Y). These results appear to support the presence of epistasis and/or clonal interference, where the selection strength of the other SNPs is influenced by the appearance of a drug-resistant mutation under drug pressure. In fact, a known compensatory mutation HA 1395 (D112N) for infectivity (Thoennes *et al.* 2008) was among the three trajectories increasing rapidly, further confirming that increased infectivity might contribute to tissue culture adaptation, thus reinforcing the use of the method to evaluate biological hypotheses. Moreover, the estimated values of selection coefficients differed greatly between the constant-selection method (WFABC) and CP-WFABC. In particular, the estimates of *s* for PA 2194 and NP 159 from the Wright-Fisher model were inferred by WFABC as being near zero (*i.e.*, neutral), though the more robust CP-WFABC estimates here indicate fluctuation of *s* from negative to positive values. Thus, these fluctuations cannot be explained by genetic drift alone, as previously speculated. We have therefore identified some cases where an alternative model of changing selection is essential for correctly estimating selection parameters and identifying change points. In the absence of drug treatment, the Bayes factor test could not reject the constant selection model M_0 for any of the identified SNPs suggesting that in, the absence of the drug, the selective pressures on the population are largely constant as expected.

The simulation studies of changing selection reveal some important points to consider from the standpoint of experimental design. For CP-WFABC, the parameter estimation of a single selection coefficient between two sampling time points performs reasonably well for haploid population sizes above $N_e = 1000$. However, it is advisable to have three sampling time points to enhance the parameter estimation by maximizing the information contained in the allele trajectory for both selection periods, particularly for diploids and in smaller population sizes ($N_e < 1000$). Thus, in order for a change to be detectable, it is required to have at least four sampling time points where the change must occur between the second and third sampling time points. The simulation studies of CP-WFABC confirm that the estimation of *CP* performs best at the intermediate range of time-sampled data, as any change happening before the second and after the second-to-last sampling time point is impossible to detect (Figure 3, Figure 4, Figure S4, Figure S5, Figure S6, and Figure S7). Finally, it remains a future challenge to expand this method to consider more than one change point in selection strength, as some of the trajectories in the influenza A application (such as PA 2194 and PB1 33) suggest the presence of several change points along the trajectory, and to resolve the problem arising from the assumption of constant population size.

ACKNOWLEDGMENTS

We thank Kristen Irwin for helpful comments. The computations were performed at Vital-IT (<http://www.vital-it.ch>) in the Swiss Institute of Bioinformatics. This work was funded by grants from the Swiss National Science Foundation and a European Research Council starting grant to J.D.J.

LITERATURE CITED

- Bank, C., G. Ewing, A. Ferrer-Admetlla, M. Foll, and J. D. Jensen, 2014 Thinking Too Positive? Revisiting Current Methods of Population Genetic Selection Inference. *Trends Genet.* 30: 540–546.
- Beaumont, M. A., 2010 Approximate Bayesian Computation in Evolution and Ecology. *Annu. Rev. Ecol. Evol. Syst.* 41: 379–406.
- Beaumont, M. A., W. Zhang, and J. D. Balding, 2002 Approximate Bayesian Computation in Population Genetics. *Genetics* 162: 2025–2035.
- Bell, G., 2010 Fluctuating Selection: The Perpetual Renewal of Adaptation in Variable Environments. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365: 87–97.
- Chen, J., and A. K. Gupta, 2001 *Parametric Statistical Change Point Analysis*, Ed. 2nd. Birkhauser, Berlin.
- Chevin, L. M., 2013 Genetic constraints on adaptation to a changing environment. *Evolution* 67: 708–721.
- Cook, L. M., and D. A. Jones, 1996 The Medionigra Gene in the Moth *Panaxia Dominula*: The Case for Selection. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 351: 1623–1634.
- Ewens, W. J., 2004 *Mathematical Population Genetics 1 - Theoretical Introduction*, Ed. 2nd. Springer-Verlag, New York.
- Fisher, R. A., and E. B. Ford, 1947 The Spread of a Gene in Natural Conditions in a Colony of the Moth *PANAXIA DOMINULA* L. *Heredity* 1: 143–174.
- Foll, M., Y.-P. Poh, N. Renzette, A. Ferrer-Admetlla, C. Bank *et al.*, 2014 Influenza Virus Drug Resistance: A Time-Sampled Population Genetics Perspective. *PLoS Genet.* 10: e1004185.
- Foll, M., H. Shim, and J. D. Jensen, 2015 WFABC: A Wright-Fisher ABC-Based Approach for Inferring Effective Population Sizes and Selection Coefficients from Time-Sampled Data. *Mol. Ecol. Resour.* 15: 87–98.
- Gompert, Z., 2015 Bayesian Inference of Selection in a Heterogeneous Environment from Genetic Time-Series Data. *Mol. Ecol.* (in press).
- Good, I. J., 1992 The Bayes/Non-Bayes Compromise: A Brief Review. *J. Am. Stat. Assoc.* 87: 597–606.
- Gossmann, T. I., D. Waxman, and A. Eyre-Walker, 2014 Fluctuating Selection Models and McDonald-Kreitman Type Analyses. *PLoS One* 9: 1–5.
- Haldane, J. B. S., 1932 A Mathematical Theory of Natural and Artificial Selection. Part IX. Rapid Selection. *Math. Proc. Camb. Philos. Soc.* 28: 224–248.
- Hall, A. R., P. D. Scanlan, A. D. Morgan, and A. Buckling, 2011 Host-Parasite Coevolutionary Arms Races Give Way to Fluctuating Selection. *Ecol. Lett.* 14: 635–642.
- Jones, D. A., 2000 Temperatures in the Cothill Habitat of *Panaxia* (*Callimorpha*) *Dominula* L. (the Scarlet Tiger Moth). *Heredity* 84: 578–586.
- Jorde, P. E., and N. Ryman, 2007 Unbiased Estimator for Genetic Drift and Effective Population Size. *Genetics* 177: 927–935.
- Karlin, S., and B. Levikson, 1974 Temporal Fluctuations in Selection Intensities: Case of Small Population Size. *Theor. Popul. Biol.* 6: 383–412.
- Karlin, S., and U. Lieberman, 1974 Random Temporal Variation in Selection Intensities: Case of Large Population Size. *Theor. Popul. Biol.* 6: 355–382.
- Kimura, M., 1954 Process Leading to Quasi-Fixation of Genes in Natural Populations Due to Random Fluctuation of Selection Intensities. *Genetics* 39: 280–295.
- Lacerda, M., and C. Seoighe, 2014 Population Genetics Inference for Longitudinally-Sampled Mutants under Strong Selection. *Genetics* 198: 1237–1250.
- Malaspinas, A.-S., O. Malaspinas, S. N. Evans, and M. Slatkin, 2012 Estimating Allele Age and Selection Coefficient from Time-Serial Data. *Genetics* 192: 599–607.
- Mathieson, I., and G. McVean, 2013 Estimating Selection Coefficients in Spatially Structured Populations from Time Series Data of Allele Frequencies. *Genetics* 193: 973–984.
- O'Hara, R. B., 2005 Comparing the Effects of Genetic Drift and Fluctuating Selection on Genotype Frequency Changes in the Scarlet Tiger Moth. *Proc. Biol. Sci.* 272: 211–217.
- Page, B. Y. E. S., 1954 Continuous Inspection Schemes. *Biometrika* 41: 100–115.

- Renzette, N., D. R. Caffrey, K. B. Zeldovich, P. Liu, G. R. Gallagher *et al.*, 2014 Evolution of the Influenza A Virus Genome during Development of Oseltamivir Resistance in Vitro. *J. Virol.* 88: 272–281.
- Ryan, T. P., 2011 *Statistical Methods for Quality Improvement*, Ed. 3rd. Wiley, New York.
- Sunnåker, M., A. G. Busetto, E. Numminen, J. Corander, M. Foll *et al.*, 2013 Approximate Bayesian Computation. *PLOS Comput. Biol.* 9: e1002803.
- Tanaka, M. M., and F. Valckenborgh, 2011 Escaping an Evolutionary Lobster Trap: Drug Resistance and Compensatory Mutation in a Fluctuating Environment. *Evolution* 65: 1376–1387.
- Thoennes, S., Z.-N. Li, B.-J. Lee, W. A. Langley, W. A. Skehel *et al.*, 2008 Analysis of Residues near the Fusion Peptide in the Influenza Hemagglutinin Structure for Roles in Triggering Membrane Fusion. *Virology* 370: 403–414.
- Wright, S., 1948 On the Roles of Directed and Random Changes in Gene Frequency in the Genetics of Populations. *Evolution* 2: 279–294.
- Xavier, R., N. Turck, A. Hainard, N. Tiberti, F. Lisacek *et al.*, 2011 pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinformatics* 12: 77.

Communicating editor: Y. Kim