Detecting and quantifying changing selection intensities from time-sampled polymorphism data

Hyunjin Shim[1,2], Stefan Laurent[1,2], Matthieu Foll[§,1,2,3] and Jeffrey D. Jensen[§,1,2]

[1]School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

[2]Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland

[3] International Agency for Research on Cancer, Lyon, France

Corresponding author: Hyunjin Shim (hyunjin.shim@epfl.ch)

Address for author:

EPFL SV IBI-SV UPJENSEN

AAB 0 45 (Bâtiment AAB)

Station 15

CH-1015 Lausanne

SWITZERLAND

Keywords: population genetics; fluctuating selection; change point analysis; time-sampled data; approximate Bayesian computation; Wright-Fisher model; experimental design

Running title: Changing selection in time-sampled data

1

1   Abstract

2   During his well-known debate with Fisher regarding the phenotypic dataset of

3   *Panaxia dominula*, Wright (1948) suggested fluctuating selection as a potential

4   explanation for the observed change in frequency. This model has since been invoked in

5   a number of analyses, with the focus of discussion centering mainly on random or

6   oscillatory fluctuations of selection intensities. Here, we present a novel method to

7   consider non-random changes in selection intensities using Wright-Fisher approximate

8   Bayesian (ABC)-based approaches, in order to detect and evaluate a change in selection

9   strength from time-sampled data. This novel method jointly estimates the position of a

10   change point as well as the strength of both corresponding selection coefficients (and

11   dominance for diploid cases) from the allele trajectory. The simulation studies of CP-

12   WFABC reveal the combinations of parameter ranges and input values that optimize

13   performance, thus indicating optimal experimental design strategies. We apply this

14   approach to both the historical dataset of *Panaxia dominula* in order to shed light on this

15   historical debate, as well as to whole-genome time-serial data from influenza virus in

16   order to identify sites with changing selection intensities in response to drug treatment.

17

18   Introduction

19   The common assumption of constant selection intensity through time utilized in

20   many tests of selection is often criticized as unrealistic in natural and experimental

21   populations – both owing to environmental changes (e.g., fluctuations in climate,

22   predation, or nutrition) as well as to genetic changes (e.g., epistasis, clonal interference).

23   Despite this, such considerations are not accounted for in most population genetic

24   models, since inferring changing selection coefficients (s) from single-time point

25   polymorphism data is difficult. However, owing to recent technological advances, time-

26   sampled polymorphism data are increasingly available, and time-serial analytical

27   methods are expanding (Malaspinas et al. 2012; Mathieson and McVean 2013; Foll et al.

28   2014a; Lacerda & Seoighe 2014; and see review of Bank et al. 2014) – allowing for an

29   empirical evaluation of the importance of changing s models.

30   Fluctuating selection in natural populations was suggested by Wright (1948)

31   with regards to the phenotypic time-serial data of *Panaxia dominula* (scarlet tiger moth)

32   to account for its observed annual fluctuations (Fisher and Ford 1947). Since, there have

33   been several theoretical considerations of fluctuating selection (Kimura 1954; Karlin &

34   Levikson 1974; Karlin & Lieberman 1974; Gossmann et al. 2014; Gompert 2015), as well

35   as many observations of fluctuating selection in natural populations (for a review, see

36   Bell 2010). Nonetheless, until recently, analyses of fluctuating selection centered on

37   *random* or *seasonal* oscillations of selection strength through time, as the mathematical

38   complexity of analytical methods only allowed the simplest cases to be considered.

39   Approximate Bayesian Computation (ABC) has the advantage of being flexible in

40   integrating complex models due to computational efficiency and the lack of likelihood

41   computation (Beaumont 2010). Recently, a hierarchical ABC-based method based on the

42   Wright-Fisher model was developed in order to infer genome-wide effective population

43   size and per-site selection coefficients from whole-genome multiple-time point datasets

44   (Foll et al. 2014a, b). While the initial approach performs well overall, the authors noted

45   the possibility for observations inconsistent with a single-s Wright-Fisher model; this

46   was indeed observed at certain sites in their analysis of the influenza virus genome. In

47   their analyses, these trajectories are simply excluded from consideration. Thus, as a

48   natural extension, we here investigate the presence of changing selection in these outlier

49   SNPs; in doing so, we also develop an extended Wright-Fisher ABC-based method

50   capable of detecting and quantifying changing selection intensities through time.

51

52 Materials and Methods

53 Wright-Fisher ABC-based method

54   This approach first relies on a previously developed Wright-Fisher ABC method

55 (WFABC; Foll et al. 2014a,b) in order to estimate effective population size (N_e). The

56 posterior of the N_e estimated from WFABC is used as a prior for the following extended

57 method. The trajectory X of a given allele with a known N_e consists of time-serial allele

58 frequencies f_t ($t = 1,...,T$) where T is the total number of generations (with $T > 4$ to allow

59 for a change point to be realizable with the Wright-Fisher model), from which a sample

60 n_i is taken at sampling time points $i = 1,...,I$ (with $I > 4$ to allow for a change point to be

61 detectable). Parameters to be inferred include the selection coefficient prior to the

62 change in selection intensity (s_1), the selection coefficient subsequent to the change in

63 selection intensity (s_2), the time of change (CP), and the dominance coefficient (h) for

64 diploid models. The joint posterior distribution of these parameters can be estimated by

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

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

65 for the diploid model and the haploid model, respectively. The ABC approach allows

66 these parameters to be inferred using Wright-Fisher model simulations without

67 calculating the likelihood $P(X|s_1, s_2, CP, h)$ or $P(X|s_1, s_2, CP)$.

68   The Wright-Fisher model simulator with a change point in selection strength is

69 used to simulate the data X, with relative fitnesses $w_{AA}=1+s$, $w_{Aa}=1+sh$ and $w_{aa}=1$ for the

70 diploid model, and $w_A=1+s$ and $w_a=1$ for the haploid model (Ewens 2004). Initially, the

71 random sampling of an allele from generation 1 to generation $CP-1$ is simulated using s_1,

72 and onwards from the change point (CP) using s_2. In order to simulate realistic allele

73 trajectories with changing selection coefficients, the allele needs to be segregating at the

74 time of the change point. This condition is necessary since the change in selection

4

75   coefficient cannot occur if the allele is either lost or fixed beforehand, assuming the

76   infinite-site model with no back mutations. Thus, only alleles segregating at the change

77   point are accepted as a data censoring procedure.

78   The associated summary statistic for these time-serial data is Fs', an unbiased

79   estimator of N_e that measures the allele frequency change between two sampling time

80   points without bias in cases of highly skewed allele frequencies and cases of small

81   sample size (Jorde and Ryman 2007). It is given as

$$Fs = \frac{(x-y)^2}{z(1-z)} \quad (3)$$

$$Fs' = \frac{1}{t_{xy}} \frac{Fs\left[1 - \frac{1}{2\tilde{n}}\right] - \frac{2}{\tilde{n}}}{\left(1 + \frac{Fs}{4}\right)\left[1 - \frac{1}{n_y}\right]} \quad (4)$$

82   where x and y are the allele frequencies at two consecutive time points separated by t_{xy}

83   generations, $z = (x+y)/2$, and $\tilde{n}$ is the harmonic mean of the chromosome sample sizes n_x

84   and n_y at two consecutive time points. Unlike the WFABC approach that summarizes

85   time-serial trajectories into only two summary statistics (increasing and decreasing Fs';

86   Foll et al. 2014a), here Fs' is summarized at every pair of consecutive time points as

87    $Fs'_1,...,Fs'_{I-1}$, where I is the number of sampling time points. This modification allows

88   additional information such as the timing of increase or decrease in allele frequency to

89   be captured - an important factor for detecting the change point. In order to retain

90   information about directionality, increasing allele frequencies are made positive and

91   decreasing allele frequencies are made negative with regards to the absolute value.

92   The joint posterior distribution of the parameters of interest is obtained using the

93   algorithm described in Beaumont et al. (2002). The approximate posterior density

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

95   with $\theta=(s_1, s_2, CP, h)$ for the diploid model and $\theta=(s_1, s_2, CP)$ for the haploid model, is

96   obtained using an ABC algorithm as follows:

97   i.   Simulate K trajectories from the Wright-Fisher model with a change in

98   selection intensity, with θ randomly sampled from its prior $P(\theta)$,

99   conditional on the allele segregating at the change point.

100   ii.   Compute $U(x_k)$ for each accepted trajectory using the Fs' summary statistic

101   between all consecutive sampling time points i: $U(x_{k,i})$ where $i = 1,...,I\text{-}1$

102   where I is the last sampling time point.

103   iii.   Retain the simulations with the smallest Euclidian distance between $U(x_{k,i})$

104   (from the simulated) and $U(X_i)$ (from the observed) to obtain an

105   approximate posterior density of $P(\theta|X)$.

106

107   For the first step, simulations are performed with the same initial conditions as

108 the observed data - including effective population size, initial allele frequency, and the

109 sampling points and sizes. In addition, a minimum allele frequency in one of the

110 sampling time points is imposed on simulated trajectories as is done in observed data.

111 This ascertainment scheme takes into account the non-random criterion of considering

112 only the trajectories reaching values above the sequencing error threshold in the

113 observed data (Foll et al. 2014b).

114   For the second step, it is important to note that the Fs' summary statistic is

115 calculated between every pair of consecutive sampling time points (Figure 1) – thus

116 there are I-1 summary statistics for each simulated and observed trajectory. This

117 construction of the summary statistic enables information on both the timing and

118 strength of the allele frequency change to be captured, as the timing of the change is

119 essential in detecting the change point and the strength of the change is essential in

120 estimating the corresponding selection coefficients. For the diploid model, an additional

121 parameter h is inferred jointly with the other three parameters, as its value is one of the

122 determining factors in the timing of allele frequency change (Haldane 1932).

123   For the third step, the simulated *Fs'* summary statistics U*(x_{k,i})* between every pair

124   of consecutive sampling time points are compared with the corresponding observed *Fs'*

125   summary statistics U*(X_i)* – allowing a small fraction of the simulated trajectories (less

126   than 0.1%) with allele frequency changes that best match the observed trajectory (in

127   terms of both timing and strength) to be retained.

128

129   **Wright-Fisher ABC-based method with Change-point analysis**

130   In order to increase computational efficiency and sensitivity in change point

131   detection, an additional summary statistic is integrated into the Wright-Fisher ABC-

132   based method. This novel summary statistic is derived from change point analysis –

133   statistical techniques developed and used in many disciplines ranging from finance to

134   quality control in order to detect and estimate change (e.g., Chen and Gupta 2001).

135   Among the techniques available, the cumulative sum control chart (CUSUM) developed

136   by Page (1954) is able to detect small and sustained shifts in the statistics β obtained

137   from a sample (Ryan 2011). Instead of using the entire CUSUM procedure as a separate

138   method for detecting change, the CUSUM value is integrated into the Wright-Fisher ABC-

139   based method as an additional summary statistic that characterizes the time-sampled

140   trajectory of an allele:

141    $$S_i = S_{i-1} + \left(\beta_i - \bar{\beta}\right), \quad i = 1, \dots, I \qquad (6)$$

142   where $\bar{\beta} = mean$ and $S_0 = 0$. The CUSUM value S is accumulated only when the statistic

143   β is different from its average value in the dataset.

144   The change point S_{CP} is the sampling time point with the maximal absolute value

145   of S_m, which is the furthest point from the initial value zero attaining the maximal

146   accumulation of difference from the average value:

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

148

7

149   Here, we calculate Fs' at each pair of consecutive sampling time points as the statistic β,

150   since it is a time-serial measure of the allele frequency change - which is indicative of the

151   selection strength change. Thus, when Fs' is used as the statistic β in the CUSUM, the

152   maximal CUSUM value S_{CP} is the potential change point of the allele trajectory, as

153   illustrated with an example in Figure 1.

154   In the Change-Point Wright-Fisher ABC (CP-WFABC), an additional summary

155   statistic S_{CP} with an infinite weight is used to characterize observed and simulated allele

156   frequency trajectories for detecting a change point. In the third step of the ABC

157   algorithm, the Euclidean distance between $U(x_{k,i})$ and $U(X_i)$ is calculated only if the

158   maximal CUSUM value $S_{CP,k}$ of the simulated data matches the maximal CUSUM value S_{CP}

159   of the observed data:

160
$$D = \begin{cases} \|U(X_i) - U(x_{k,i})\|, & \text{if } S_{CP} = S_{CP,k} \\ \infty, & \text{otherwise.} \end{cases} \quad (8)$$

161   This additional step allows the computation to be more efficient – especially

162   when there is a large number of time points sampled – as the Euclidean distance is

163   calculated for a fraction of simulated trajectories whose maximal CUSUM value is equal

164   to that of the observed (i.e., with the same time-sampled characteristic). Furthermore,

165   as the CUSUM is sensitive to small and sustained changes, integrating the CUSUM into

166   the Wright-Fisher ABC increases its sensitivity for detecting small and sustained

167   changes in selection strength. The potential bias in the calculation of the maximal

168   CUSUM value is counteracted by the fact that the bias would be present in both the

169   observed and the simulated trajectories.

170

171   **Simulated data with constant selection and with changing selection**

172   We generated simulated datasets of different effective population sizes using the

173   Wright-Fisher model for two scenarios: (1) trajectories of constant selection with only s

8

174   and h (for diploid models) as parameters, and (2) trajectories of changing selection with

175     $s1$, $s2$, CP, and h (for diploid models) as parameters. For selection coefficients, uniform

176   priors of [-1,1] were used. The uniform prior of CP was set to occur between the second

177   generation and the second-to-last generation [2,T-1], where T is the number of

178   generations of the population in the time-serial data. The dominance coefficient h for the

179   diploid model was randomly drawn from one of three values: complete recessiveness,

180   co-dominance, or complete dominance [0,0.5,1]. Although these prior ranges are

181   uninformative, the constraint on the trajectories to be segregating at the change point

182   shapes the distribution of the prior ranges according to the input parameters such as

183   ploidy, effective population size, initial allele frequency, and number of generations; the

184   updated priors for the haploid population of N_e = 100 and the diploid population of N_e

185   =50 are shown as examples (Figure S1 and S2).

186   The other input values – such as the number of generations (T=100), the

187   sampling time points (I=10), the sample size (n=100), the initial allele arising as a new

188   mutation, and the ascertainment of observing a minimum frequency at 2% – were kept

189   constant for the two scenarios. We retained the best 0.1% of 1,000,000 simulations for

190   each pseudo-observable trajectory using the rejection algorithm based on the Euclidean

191   distance as described above. The mode of the posterior distribution from the best

192   simulations (Sunnåker et al. 2013) was used to evaluate the estimated parameter value

193   against the true parameter value.

194

195   **Results**

196   **ABC model choice in the Change-Point Wright-Fisher ABC method**

197   The first step of CP-WFABC is to be able to distinguish changing selection

198   trajectories from constant selection trajectories. ABC model choice was constructed to

9

199   choose between two models: M_0 with a single selection coefficient, and M_1 with two

200   selection coefficients and a change point. The relative probability of M_1 over M_0 can be

201   computed through the model posterior ratio as the Bayes factor $B_{1,0}$ (Sunnåker et al.

202   2013):

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

203   when the model prior $p(M_0)$ is equal to $p(M_1)$. In practice the model priors are made

204   equal by producing the same number of simulations for each model and retaining the

205   best simulations from the lot. The posterior ratio is computed as the number of accepted

206   simulations from M_1 over those of M_0 – giving the Bayes factor $B_{1,0}$ which is an indicator

207   of the support for a specific model. The performance study was conducted with a

208   haploid population of N_e = (100, 1000, or 10000) and a diploid population with N_e = (50,

209   500, or 5000) using the simulated datasets of the two scenarios described in the

210   previous section as M_0 and M_1, respectively.

211   We considered two cases for the pseudo-observables to test the sensitivity and

212   specificity of the ABC model choice: the first case when the pseudo-observed trajectories

213   have a single selection coefficient, and the second case when they have changing

214   selection coefficients with a change point. One thousand pseudo-observable trajectories

215   were generated for each case with the data ascertainment minimum frequency set to 2%

216   for at least one of the sampling time points. Additionally for the second case, pseudo-

217   observable trajectories were accepted only when the allele was segregating at the time

218   of the change point – a constraint for realistic combinations of selection coefficients,

219   change points, and dominance (for diploids) – in order to reproduce changing selection

220   trajectories in real datasets. All other input values were kept constant as in the

221   simulated datasets described in the previous section.

222   The results of the ABC model choice from a haploid population with N_e = 100 and

223   a diploid population with N_e = 500 are represented as ROC curves (Robin et al. 2011) in

224   Figure 2. Specificity is given on the x-axis showing the true negative rate, while

225   sensitivity is given on the y-axis showing the true positive rate of the Bayes factor $B_{1,0}$

226   calculated from 1000 pseudo-observables of changing selection (where $B_{1,0}$ should be

227   large) and 1000 pseudo-observables of constant selection (where $B_{1,0}$ should be small).

228   The overall ROC curves in black (all trajectories) show that when the specificity

229   threshold is most conservative in detecting no false positives (i.e. $B_{1,0}$ = infinite), the

230   Bayes factor $B_{1,0}$ has a sensitivity of around 30% for all populations. Considering that the

231   pseudo-observable trajectories were simulated randomly from a wide range of prior

232   values, the Bayes factor $B_{1,0}$ from CP-WFABC is in general sensitive and specific. The

233   ROC curves in black for the other haploid and diploid populations (Figure S3) also

234   indicate that the Bayes factor $B_{1,0}$ is sensitive and specific as they are above the diagonal

235   line of no-discrimination. The area under the ROC curve (AUC) is used to assess how

236   reflective the Bayes factor $B_{1,0}$ is of the true model, as summarized for all pseudo-

237   observable populations in Table 3. The AUC values show that the Bayes factor $B_{1,0}$ is

238   ~80% more probable to rank a randomly chosen changing selection case above a

239   randomly chosen constant selection case. Additionally, the distribution of Bayes factors

240     $B_{1,0}$ under the null model M_0 (i.e., case 1) was used to compute the significance level α at

241   1% (Good 1992). For both diploids and haploids, the significance threshold is higher for

242   smaller population sizes (Table 4), and the calculation of these thresholds will be

243   important in any given data application.

244   Following the detection of changing selection trajectories using ABC model choice,

245   the quality of parameter estimation by the model chosen was evaluated. The cross-

246   validation results from the haploid population of N_e = 100 are shown in Figure 3 and

247   those from the diploid population of N_e = 500 in Figure 4 (see Figures S4-S7 in

248   Supporting information for additional results). For the case where the pseudo-

249   observables were of constant selection, the estimation for a single s (and the dominance

11

250   h for diploid) using CP-WFABC is very accurate, as the mode of the best simulations

251   from the M_0 model for each pseudo-observable lies along the red diagonal line.

252   Exceptions include uninformative trajectories where the allele surpasses the minimum

253   frequency of 2% in the first sampling and is lost immediately due to genetic drift or

254   negative selection and therefore not observed in subsequent samplings. Such

255   trajectories will always keep the same set of best simulations from the M_0 model since

256   their selection strength is indistinguishable, and they result in horizontal lines along the

257   estimated negative value. This phenomenon is particularly pronounced when

258   population size is small as shown in Figures 3 and S6, since the role of genetic drift is

259   more significant.

260   For the second case when the pseudo-observables are of changing selection

261   intensity, the joint estimation of the parameters is also effective for a restricted range of

262   values. In Figures 3-4 and S4-S7, the mode estimation of each pseudo-observable is

263   color-coded according to the three categories of trajectory shape. The green dots are

264   pseudo-observable trajectories that change from positive s_1 to positive s_2. The blue dots

265   are those that change from positive s_1 to negative s_2, while the magenta colors include all

266   other cases (e.g., neutral or negative s_1 to any value of s_2). There is a clear clustering by

267   category – with the best estimation being of positive values of s_1 below 0.5, moderate

268   values of s_2 between -0.5 and 0.5, and CP values for the blue category of positive s_1 to

269   negative s_2. In trajectories other than those with positive s_1 to negative s_2, the change

270   point is difficult to detect, particularly for diploid populations where the additional

271   dominance parameter h was estimated (Figure 4 and Figure S6-S7). This trend is also

272   observed when the ROC curves are generated according to these three categories

273   (Figure 2 and Figure S3). For all populations, the Bayes factor $B_{1,0}$ is more sensitive and

274   specific for trajectories changing from positive s_1 to negative s_2 (ROC curves in blue),

275   reaching above 60% of the true positive rate when there are no false positives. Despite

276   the restricted range of good parameter estimation in s_1, s_2 and CP, the estimation of

277   dominance is robust for both cases of constant and changing selection (except for the

278   small population size of N_e = 50; Figure S6).

279   In order to evaluate the performance of the joint parameter estimation, the

280   coefficient of determination R^2 is used to assess the cross-validation between the

281   estimated values (y_m) and the true values (f_m), compared with the simple average of the

282   estimated values ($\bar{y}$). The closer the R^2 value is to 1, the better the parameter estimation

283   as shown:

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

284   Tables 2 and 3 summarize the performance of the joint parameter estimation for

285   all cases as the R^2 values for the haploid and diploid populations, respectively. The first

286   case is when the pseudo-observables are of constant selection intensity, in which case

287   the true model (M_0) performs only slightly better than the false model (M_1) for

288   estimating s. This discrepancy in parameter estimation of M_0 is mainly owing to

289   uninformative pseudo-observable trajectories with constant selection (which have been

290   lost or fixed) being associated with the true model (M_0) of constant selection; this is due

291   to the constraint for the allele to be segregating at the change point in the (false) model

292   M_1 of changing selection. In the cross-validation of the constant selection case, the

293   parameters estimated form horizontal lines at negative estimated values for those

294   trajectories that are lost, and cluster at the top right corner for those trajectories that

295   are fixed (Figure 3-4, Figure S4-S7).

296   For the second case in which the pseudo-observables have changing selection

297   coefficients, parameter estimation from the true model (M_1) performs better than that

298   from the false model (M_0) for all parameters – particularly when population sizes are

299   large. As expected, there is a trend of better parameter estimation as population size

13

300   increases. Additionally, it has been shown that the value of the Bayes factor $B_{1,0}$ is a good

301   indicator of the parameter estimation performance (results not shown).

302

303   ### Data Application

304   ### Historical dataset of *Panaxia dominula*

305   A long-running dataset based on the *medionigra* morph responsible for darker

306   wing color in wild populations of *Panaxia dominula* (Figure S8) began in 1939 with

307   collections by Fisher (Fisher and Ford 1947) and continued through 1999 (Cook and

308   Jones 1996; Jones 2000). Despite this phenotypic time-serial data having been analyzed

309   previously from various angles (O'Hara 2005; Mathieson and McVean 2013; Foll et al.

310   2014b), it is still relevant to consider a model of changing selection in time, as Wright

311   (1948) originally suggested.

312   The recent reconsiderations of the dataset tend to favor a lethal-recessive model

313   with an effective population of $2N_e = 1000$ (Mathieson and McVean 2013; Foll et al.

314   2014b) – however, the biological question of how the *medionigra* morph could have

315   reached the initial frequency of 11% in the dataset remains unanswered with this

316   conclusion of constant strong negative selection. Wright asserted that the trajectory of

317   the *medionigra* morph during this period could be explained by fluctuating selection

318   with "no net selective advantage or disadvantage". Although this alternative hypothesis

319   has been considered as a random fluctuation of selection by estimating selection

320   coefficients between every sampling time point (see O'Hara 2005), the quantitative

321   plausibility of a directional change-in-*s* model over a single-*s* model lacks thorough

322   investigation. Thus, we re-analyze this dataset using the CP-WFABC method in order to

323   investigate the possibility of changing selection in the *medionigra* morph during the 60-

324   year data collection.

14

325   Using the ABC model choice introduced here as a test for a change in selection

326 strength, and to estimate the parameters of interest for the chosen model, we assume

327 the *medionigra* allele is a single co-dominant locus responsible for the homozygous and

328 heterozygous expressions of the phenotypic forms *bimacula* and *medionigra*,

329 respectively (Cook and Jones 1996). The model M_0 assumes a single selection coefficient,

330 thus the only parameter to estimate is *s*. The M_1 model assumes a change in selection

331 strength, thus the parameters of interest are *s1*, *s2* and *CP*. Both M_0 and M_1 take the prior

332 range of [-1,1] for the selection coefficients and the prior range of [2, 59] for the change

333 point in the M_1 model. For the M_1 model, these uninformative priors are updated with

334 the constraint that the allele must be segregating at the time of change point. Here, we

335 create 10,000,000 simulated datasets for each M_0 and M_1, and apply the rejection

336 algorithm of the CP-WFABC method to retain the best 1000 simulations compared with

337 the observed trajectory. The effective population size is assumed to be $2N_e = 1000$ as in

338 previous studies (Wright 1948; Cook and Jones 1996; O'Hara 2005), with an initial allele

339 frequency of 11% and a minimum frequency ascertainment of 2%.

340   The Bayes factor for M_1 over M_0 is calculated as 0.952, indicating that the single

341 coefficient M_0 model cannot be rejected in favor of the changing selection M_1 model

342 (Table 4). From the parameter estimation of the model M_0 (Figure S9), the mode of the

343 posterior distribution for *s* is given as -0.15 as asserted by Fisher and Ford (1947).

344 When the ABC model choice was repeated with a smaller population size of $2N_e = 100$ as

345 suggested by Wright (1948) and O'Hara (2005), the Bayes factor increases to 1.87 (i.e.,

346 the changing selection model is twice as likely as the constant selection) – however, this

347 value is not large enough to be significant for a diploid population of $N_e = 50$ (Table 4).

348

349 **Experimental evolution of Influenza virus with drug treatment**

350   The evolution of pathogens within a host is one of the most important cases in

351 which the possibility of fluctuating selection must be considered – as they may

352 experience drastically changing selective pressures due to host immune response,

353 specific drug treatments, and/or pathogenic cooperation or competition (Tanaka and

354 Valckenborgh 2011; Hall et al. 2011). Thus, how these pathogens adapt to these rapid

355 external and internal changes is of major concern to the biomedical community.

356   The time-serial experimental dataset of influenza A conducted by Renzette et al.

357 (2014) and Foll et al. (2014a) is an interesting case study on the impact of drug

358 treatment on influenza virus evolution. The dataset consists of 13 sampling points from

359 which population-level whole-genome data were collected. Drug treatment with a

360 commonly used neuraminidase inhibitor (oseltamivir) began after the collection of the

361 third sample and continued, at increasing concentrations, until the final passage. Using

362 WFABC, the genome-wide effective population size across the sampling time points was

363 estimated (N_e = 176) and the SNPs under selection were identified.

364   Here, we apply the CP-WFABC method on two cases of interest from this study to

365 consider a possible change in selection strength under drug treatment: the first case

366 includes trajectories identified as being driven by positive selection, while the second

367 includes outlier trajectories (i.e., trajectories not fitting a single *s* Wright-Fisher model).

368 For all cases, we test the model M_0 (i.e., a single selection coefficient) and M_1 (i.e., a

369 changing selection coefficient), with parameters of interest (*s*) and (*s1*, *s2*, *CP*),

370 respectively. The number of generations per passage is assumed to be 13, and the

371 minimum frequency of 2% is set as an ascertainment for observing the minor allele in

372 the data. *De novo* mutations are assumed to occur at the first sampling time point for the

373 SNPs whose allele frequency reached more than 2% before the drug administration

374 (except for trajectories whose initial frequency is above 2%, which are assumed to be

16

375   standing variation), and at the fourth sampling point for those whose frequency did not.

376   This assumption is based on the high mutation rate, large population bottlenecks

377   associated with passaging, and large census population size between passages.

378   10,000,000 datasets were simulated for each M_0 and M_1; the best 1000 trajectories from

379   the lot were retained using the rejection algorithm described in the Methods. The

380   uniform prior ranges for the selection coefficients were set as [-1,1], for the change point

381   as [2,157] or [2,105] depending on the appearance of the mutation, and with the

382   constraint of segregating alleles at the change point for M_1.

383   The results for the Bayes factors and the parameter estimates are summarized in

384   Table 5 for all trajectories of interest. The Bayes factors of most trajectories show strong

385   support for the changing selection model: the stronger the selection strength change, the

386   larger the Bayes factor. Using the Bayes factors from the simulation studies as guidance

387   (Table 4), the significance threshold to reject M_0 is computed as 3.7 for a small haploid

388   population. As expected, the trajectories identified as outliers of the single s Wright-

389   Fisher model (NP 159, PB1 33) all reject the constant selection model M_0 with a large

390   Bayes factor. We also note that the Bayes factor for the drug-resistant mutation H275Y

391   (NA 823) does not support the changing selection model strongly, confirming that the

392   experimental evolution procedure kept the selective pressure of the drug constant by

393   adjusting the drug concentration to reduce viral plaque numbers to 50% at each passage.

394   The change points are estimated to be mostly between the seventh and eighth passages,

395   a notable result since three of these trajectories (HA 48, HA 1395, NA 582) are

396   increasing rapidly after the drug-resistant mutation H275Y appears, whereas one

397   trajectory (NP 159) from a different segment decreases rapidly. This result may indicate

398   that positive selective for the three SNPs (including HA 1395; a known compensatory

399   mutation encoded also as D112N) increased along with the drug-resistant mutation

400   H275Y, potentially due to epistatic interactions, whereas another SNP decreased at that

17

401   time, potentially owing to clonal interference. The single selection estimates from

402   WFABC (Foll et al. 2014b) are similar to the M_0 estimates of the constant selection

403   coefficient only when the Bayes factor does not reject M_0 – strong evidence that an

404   alternative model of changing selection must be considered for some trajectories in

405   order to correctly estimate selection coefficients.

406   We also applied CP-WFABC to the control case of SNPs increasing in frequency

407   without drug as a comparison to the case with drug. The effective population sizes of the

408   viral populations were averaged to be 226 in the absence of drug from the previous

409   study (Foll et al. 2014a). For the control case, *de novo* mutations are assumed to occur at

410   the first sampling time point for all SNPs, but the other inputs and the ABC model choice

411   were kept the same as in the drug case. The Bayes factor results summarized in Table 5

412   demonstrate that three out of the four SNP trajectories under selection in the control

413   experiment cannot reject M_0 (i.e., constant selection). Interestingly, the only SNP

414   trajectory to support M_1 (i.e., changing selection) is HA 1395 – a known compensatory

415   mutation that also appeared under drug treatment. The parameters estimated from the

416   model chosen indicate there was a change in selective pressure from a slightly positive

417   value to a strongly positive value between the seventh and eighth passage as shown in

418   Figure 5b.

419

420   **Discussion**

421   These simulations demonstrate that the novel CP-WFABC approach presented

422   here is able to detect changing selection trajectories via ABC model choice, and also to

423   estimate a wide range of parameters of interest. Performance was analyzed separately

424   for three categories of allele trajectories according to the nature of the change in

425   selection strength: (1) a change from positive s_1 to positive s_2, (2) a change from positive

426   s_1 to negative s_2, and (3) all other changes. The datasets for each possible combination

427   were generated using the Wright-Fisher model with a change in selection strength, using

428   the most general prior ranges for all parameters s_1, s_2, CP, and h for diploids, with the

429   only constraint being segregation of the allele at the change point. For both the detection

430   and parameter estimation, CP-WFABC performs the best when the change is large,

431   particularly for the second category of change (positive s_1 to negative s_2), as shown in

432   the ROC curves (Figure 2, S3) and the cross-validation graphs (Figure 3-4, S4-S7). For

433   the first category (positive s_1 to positive s_2) and the third category (any other changes),

434   the change point is difficult to estimate, particularly for diploids where the additional

435   parameter h is also estimated. The ABC model choice of CP-WFABC has the best

436   sensitivity for full specificity, for larger population sizes (N_e>500 for diploids), and

437   haploid populations.

438   The parameter estimates of s_1 and s_2 perform best when the values are moderate.

439   For s_1, the optimal parameter range for estimation is [0,0.5], where a *de novo* mutation

440   that survives negative selection and segregates until the change point is

441   indistinguishable from other drifting mutations with similar trajectories with

442   uninformative low allele frequency. These trajectories naturally arise more frequently

443   when population size is small and in diploids where the dominance effect plays a role, as

444   shown in the third category of change (Figure 3-4, S4-S7; magenta points). When an

445   initial frequency of 10% is used instead of a *de novo* mutation, the advantage of having a

446   more informative trajectory at the beginning is counteracted by the effect of more cases

447   under negative selection or genetic drift segregating until the change point. Thus, the

448   performance of CP-WFABC for standing variation is similar to that of *de novo* mutation

449   (results not shown). For s_2, the optimal parameter range for estimation is [-0.5,0.5], as

450   trajectories with extreme values are less informative since they are lost or fixed directly

451   after the change point (explaining the clustering of the change points at earlier times).

452   For diploid populations, estimates of h are accurate to the level of determining

453   dominance from co-dominance or recessiveness, particularly for large population sizes

454   (Figure S7), given the difficulty of joint estimation with the three other parameters.

455   Indeed, estimation of this additional parameter comes at the cost of worse performance

456   for the other parameters, as can be seen in the ROC curves and cross-validation graphs:

457   the detection and parameter estimation of changing selection cases is always better for

458   haploids. Thus, in diploid cases, we recommend fixing the dominance parameter if

459   known, in order to improve the performance of CP-WFABC.

460   Although CP-WFABC is intended to detect and evaluate changing selection

461   intensities, the simulation studies show that the method also performs well in

462   estimating parameters for cases of constant selection – as has been demonstrated by

463   Foll et al. (2014a). For haploids with large population sizes, in particular, the estimated

464   values of the single parameter s correlate almost perfectly with the true values (Figure

465   S5). However, when the population size is small for both diploids and haploids, some

466   trajectories that are lost by negative selection or genetic drift are difficult to estimate, as

467   shown in Figure 3 and Figure S6 as horizontal lines along some negative estimated

468   values. This limitation of constant selection coefficients, however, is due to the

469   simulation conditions of the de $novo$ mutation at the first generation and the minimal

470   ascertainment scheme (minimum frequency of 2% at one of the sampling time points).

471   For real datasets, the conditions are likely to be less stringent, since such uninformative

472   trajectories will not be considered for parameter estimation.

473   Finally, we utilized this approach to make inference in two very different time-

474   sampled datasets: $Panaxia$ $dominula$ (diploid) and Influenza A (haploid). The time-serial

475     $medionigra$ trajectory of $P.$ $dominula$ was re-analyzed to test for a change in selection

476   strength and/or direction during 60 years of data collection. By assuming h as co-

477   dominant, the results for N_e = 500 indicate that the model M_0 of constant selection

478   cannot be rejected according to the Bayes factor from the ABC model choice algorithm.

479   The selection coefficient from this model is estimated as -0.15, corresponding with that

480   calculated by Fisher and Ford (1947). However, when the population size is assumed to

481   be smaller (N_e = 50), the Bayes factor result supports M_1 (changing selection) twice as

482   strongly as M_0 (constant selection), but not strong enough to reject M_0 according to the

483   significance level test computed with the distribution of Bayes factors under the null

484   model M_0. This dataset of the *medionigra* morph thus demonstrates the difficulty of

485   detecting and evaluating a change in selection when the population size and the number

486   of generations are small.

487   Next, CP-WFABC was applied to SNP trajectories of interest from an experimental

488   dataset of influenza A virus. The ABC model choice test was conducted on the

489   trajectories identified as 1) being positively selected and 2) as outliers from the single-*s*

490   WFABC method. For the SNPs in the presence of drug, the Bayes factor for six out of nine

491   trajectories favored the changing selection model M_1. The change points for four out of

492   these six trajectories occurred between passages 7 and 8 – the interval during which

493   three trajectories from the segments HA and NA increased rapidly while one trajectory

494   from the segment NP decreased rapidly along with the known drug-resistant mutation

495   NA 823 (H275Y). These results appear to support the presence of epistasis and clonal

496   interference, where the selection strength of the other SNPs is influenced by the

497   appearance of a drug-resistant mutation under drug pressure. In fact, a known

498   compensatory mutation (HA 1395) was among the three trajectories increasing rapidly,

499   reinforcing the use of the method to evaluate biological hypotheses. Moreover, the

500   estimated values of selection coefficients differed greatly between the constant-selection

501   method (WFABC) and CP-WFABC. In particular, the estimates of *s* for outlier trajectories

502   from the Wright-Fisher model were inferred by WFABC as being near zero (neutral),

503   though the more robust CP-WFABC estimates here indicate fluctuation of *s* from

504   negative to positive values. Thus, these fluctuations cannot be explained by genetic drift

21

505   alone, as previously speculated. We have therefore identified some cases where an

506   alternative model of changing selection is essential for correctly estimating selection

507   parameters and identifying change points. For the SNPs in the absence of drug treatment,

508   the Bayes factor for three out of four trajectories could not reject the constant selection

509   model M_0. This result indicates that in the absence of drug, the selective pressures on the

510   population are largely constant as expected. The SNP trajectory identified as supporting

511   M_1 in the control case is a known compensatory mutation (D112N) for infectivity

512   (Thoennes et al. 2008) that also appeared in the presence of drug, further confirming

513   the increased infectivity might contribute to the tissue culture adaptation.

514   The simulation studies of changing selection reveal some important points to

515   consider from the standpoint of experimental design. Firstly, at least two sampling time

516   points are needed to estimate selection strength in time-serial methods. For CP-WFABC,

517   the parameter estimation of a single selection coefficient between two sampling time

518   points performs reasonably well for haploid population sizes above N_e=1000. However,

519   it is advisable to have three sampling time points to maximize the performance of

520   parameter estimation, particularly for diploids and in smaller population sizes

521   (N_e<1000). Thus, in order for a change to be detectable, it is required to have at least

522   four sampling time points where the change must occur between the second and the

523   third sampling time points – an important factor to consider for the design of change-

524   point experiments, such as drug administration or environmental change. The

525   simulation studies of CP-WFABC confirm that the estimation of CP performs best at the

526   intermediate range of time-sampled data, as any change happening before the second

527   and after the second-to-last sampling time point is impossible to detect (Figure 3-4, S4-

528   S7). Finally, it remains a future challenge to expand this method to consider more than

529   one change point in selection strength, as some of the trajectories in the influenza A

530   application (such as PA 2194 and PB1 33) suggest the presence of several change points

531   along the trajectory.

532

Acknowledgements

534   We thank Kristen Irwin and Sebastian Matuszewski for helpful comments. The

535   computations were performed at Vital-IT (http://www.vital-it.ch) in the Swiss Institute

536   of Bioinformatics (SIB). This work was funded by grants from the Swiss National Science

537   Foundation and a European Research Council (ERC) starting grant to JDJ.

538

Data Accessibility

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

References

541 Bank, C., Ewing, G.B., Ferrer-Admettla, A., Foll, M. & Jensen, J.D. (2014) Thinking Too

543 Positive? Revisiting Current Methods of Population Genetic Selection Inference.

544 *Trends in Genetics,* 30, 540–546.

545 Beaumont, M.A., Zhang, W. & Balding, J.D. (2002) Approximate Bayesian Computation in

546 Population Genetics. *Genetics,* 162, 2025–35.

547 Beaumont, M.A., (2010) Approximate Bayesian Computation in Evolution and Ecology.

548 *Annual Review of Ecology, Evolution, and Systematics,* 41, 379–406.

549 Bell, G. (2010) Fluctuating Selection: The Perpetual Renewal of Adaptation in Variable

550 Environments. *Philosophical Transactions of the Royal Society of London. Series B,*

551 *Biological Sciences,* 365, 87–97.

552 Chen, J., & Gupta., A.K. (2001) *Parametric Statistical Change Point Analysis.* 2nd edn.

553 Birkhauser, Berlin.

554 Cook, L.M., & Jones, D.A. (1996) The Medionigra Gene in the Moth Panaxia Dominula:

555 The Case for Selection. *Philosophical Transactions of the Royal Society B: Biological*

556 *Sciences,* 351, 1623–1634.

557 Ewens, W.J. (2004) *Mathematical Population Genetics 1 - Theoretical Introduction.* 2nd

558 edn. Springer-Velag, New York.

559 Fisher, R.A., & Ford, E.B. (1947) The Spread of a Gene in Natural Conditions in a Colony

560 of the Moth PANAXIA DOMINULA L. *Heredity,* 143–174.

561 Foll, M., Poh, Y.-P., Renzette, N., Ferrer-Admetlla, A., Bank, C., Shim, H., Malaspinas, A.-S.,

562 Ewing, G., Liu, P., Wegmann, D., Caffrey, D.R., Zeldovich, K.B., Bolon, D.N., Wang, J.P.,

563 Kowalik, T.F., Schiffer, C.A., Finberg, R.W. & Jensen, J.D. (2014a) Influenza Virus

564 Drug Resistance: A Time-Sampled Population Genetics Perspective. *PLoS Genetics.*

565 Foll, M., Shim, H., & Jensen, J.D. (2014b) WFABC: A Wright-Fisher ABC-Based Approach

566   for Inferring Effective Population Sizes and Selection Coefficients from Time-

567   Sampled Data. *Molecular Ecology Resources.*

568 Gompert, Z. (2015) Bayesian Inference of Selection in a Heterogeneous Environment

569   from Genetic Time-Series Data. *Molecular Ecology.*

570 Good, I.J. (1992) The Bayes/Non-Bayes Compromise: A Brief Review. *Journal of the*

571   *American Statistical Association*, 87, 597–606.

572 Gossmann, T.I., Waxman, D. & Eyre-Walker, A. (2014) Fluctuating Selection Models and

573   Mcdonald-Kreitman Type Analyses. *PLoS ONE*, 9, 1–5.

574 Haldane, J.B.S. (1932) A Mathematical Theory of Natural and Artificial Selection. Part IX.

575   Rapid Selection. *Mathematical Proceedings of the Cambridge Philosophical Society,*

576   28, 224-248.

577 Hall, A.R., Scanlan, P.D., Morgan, A.D. & Buckling A. (2011) Host-Parasite Coevolutionary

578   Arms Races Give Way to Fluctuating Selection. *Ecology Letters*, 14, 635–642.

579 Jones, D.A. (2000) Temperatures in the Cothill Habitat of Panaxia (Callimorpha)

580   Dominula L. (the Scarlet Tiger Moth). *Heredity,* 84, 578–86.

581 Jorde, P.E., & Ryman, N. (2007) Unbiased Estimator for Genetic Drift and Effective

582   Population Size. *Genetics,* 177, 927–935.

583 Karlin, S. & Levikson, B. (1974) Temporal Fluctuations in Selection Intensities: Case of

584   Small Population Size. *Theoretical Population Biology,* 6, 383–412.

585 Karlin, S., & Lieberman, U. (1974) Random Temporal Variation in Selection Intensities:

586   Case of Large Population Size. *Theoretical Population Biology*, 6, 355–382.

587 Kimura, M. (1954) Process Leading to Quasi-Fixation of Genes in Natural Populations

588   Due to Random Fluctuation of Selection Intensities. *Genetics,* 39, 280–295.

589 Lacerda, M. & Seoighe, C. (2014) Population Genetics Inference for Longitudinally-

590   Sampled Mutants under Strong Selection. *Genetics,* 198, 1237–50.

25

591   Malaspinas, A.-S., Malaspinas, O., Evans, S.N. & Slatkin, M. (2012) Estimating Allele Age

592   and Selection Coefficient from Time-Serial Data. *Genetics,* 192, 599–607.

593   Mathieson, I., & McVean, G. (2013) Estimating Selection Coefficients in Spatially

594   Structured Populations from Time Series Data of Allele Frequencies. *Genetics,* 193,

595   973–84.

596   O'Hara, R.B. (2005) Comparing the Effects of Genetic Drift and Fluctuating Selection on

597   Genotype Frequency Changes in the Scarlet Tiger Moth. *Proceedings. Biological*

598   *Sciences / The Royal Society*, 272, 211–217.

599   Page, B.Y.E.S. (1954) Continuous Inspection Schemes. *Biometrika,* 41, 100–115.

600   Renzette, N., Caffrey, D.R., Zeldovich, K.B., Liu, P., Gallagher, G.R., Aiello, D., Porter, A.J.,

601   Kurt-Jones, E.A., Bolon, D.N., Poh, Y.-P., Jensen, J.D., Schiffer, C.A., Kowalik, T.F.,

602   Finberg, R.W. & Wang, J.P. (2014) Evolution of the Influenza A Virus Genome during

603   Development of Oseltamivir Resistance in Vitro. *Journal of Virology,* 88, 272–281.

604   Ryan, T.P. (2011) Statistical Methods for Quality Improvement. 3rd edn. *Wiley.*

605   Sunnåker, M., Busetto, A.G., Numminen, E., Corander, J., Foll, M. & Dessimoz, C. (2013)

606   Approximate Bayesian Computation. *PLoS Computational Biology,* 9.

607   Tanaka, M.M. & Valckenborgh, F. (2011) Escaping an Evolutionary Lobster Trap: Drug

608   Resistance and Compensatory Mutation in a Fluctuating Environment. *Evolution*,

609   65, 1376–1387.

610   Thoennes, S., Li, Z.-N., Lee, B.-J., Langley, W.A., Skehel, J.J., Russell, R.J. & Steinhauer, D.A.

611   (2008) Analysis of Residues near the Fusion Peptide in the Influenza Hemagglutinin

612   Structure for Roles in Triggering Membrane Fusion. *Virology*, 370, 403–14.

613   Wright, S. (1948) On the Roles of Directed and Random Changes in Gene Frequency in

614   the Genetics of Populations. *Evolution,* 2, 279–294.

615   Xavier, R., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C. & Müller, M. (2011).

616   pROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves.

617   *BMC Bioinformatics*, 12, 77.

618

Tables

Table 1. AUC values and confidence intervals for ROC curves.

	ROC curves for haploid populations		
	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 diploid populations		
	50	**500**	**5000**
AUC	0.8233	0.7378	0.7470
CI	0.8051-0.8402	0.7164-0.7593	0.7257-0.7683

Table 2. R^2 values of the parameter estimation with the ABC model choice for haploid populations.

Case 1: 1000 constant selection pseudo-observables			
M_0 (True)	100	1000	10000
s	0.558	0.843	0.914
M_1 (False)	100	1000	10000
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 pseudo-observables			
M_0 (False)	100	1000	10000
s	-0.278	-0.00232	-0.00205
M_1 (True)	100	1000	10000
s_1	-0.323	0.363	0.415
s_2	0.704	0.768	0.777
CP	0.0946	0.413	0.303

Table 3. R^2 values of the parameter estimation with the ABC model choice for diploid populations.

Case 1: 1000 constant selection pseudo-observables			
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 pseudo-observables			
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

30

Table 4. The Bayes factor thresholds for the significance level α of 1% computed using the distribution of Bayes factors under the null model M_0.

Diploid populations			
α=1%	**50**	**500**	**5000**
BF	**4.7**	**1.6**	**1.3**

Haploid populations			
α=1%	**100**	**1000**	**10000**
BF	**3.7**	**3.2**	**1.7**

Table 5. Bayes factors and parameters estimated for the influenza trajectories in the presence and absence of drug. The estimates whose Bayes factors show strong support for M_1 are in bold.

Trajectories under selection with drug							
Segment	Position	Bayes Factor (M_1/M_0)	M_0 s estimate	M_1 s_1 estimate	M_1 s_2 estimate	M_1 CP estimate	WFABC s estimate
PA[3]	2194	999	0.029	**0.160**	**-0.168**	**21 (p1-2)**	0.09
HA[2]	48	12.0	0.179	**0.099**	**0.711**	**105 (p7-8)**	0.14
HA[2]	1395	5.06	0.173	**0.109**	**0.506**	**104 (p7-8)**	0.22
NA[2]	582	∞	-	**0.050**	**0.816**	**104 (p7-8)**	0.29
NA[2]	823	1.21	0.156	0.135	0.159	109 (p8-9)	0.15
M[2]	147	1.28	0.082	0.080	0.090	128 (p9-10)	0.08
NS[2]	820	1.03	0.051	0.045	0.060	61(p4-5)	0.12
Outlier trajectories with drug							
NP[1]	159	8.09	0.011	**0.030**	**-0.054**	**103 (p7-8)**	0
PB1[1]	33	16.2	0.047	**0.021**	**0.265**	**113 (p8-9)**	0.14

Trajectories under selection without drug							
Segment	Position	Bayes Factor (M_1/M_0)	M_0 s estimate	M_1 s_1 estimate	M_1 s_2 estimate	M_1 CP estimate	WFABC s estimate
PB1[1]	1119	0.98	0.038	0.039	0.035	10 (p0-1)	0.06
HA[1]	1395	8.52	0.067	**0.038**	**0.175**	**103 (p7-8)**	0.12
NP[1]	1104	1.91	0.042	0.039	0.057	110 (p8-9)	0.05
NP[1]	1396	1.21	0.034	0.035	0.042	9 (p0-1)	0.09

[1]A *de novo* mutation at the 1st generation (passage 0)
[2]A *de novo* mutation at the 53rd generation (passage 4)
[3]Standing variation at the 1rd generation (passage 0)

Figure 1. Illustration of *Fs'* calculated between every pair of consecutive sampling time points and the maximal CUSUM 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.
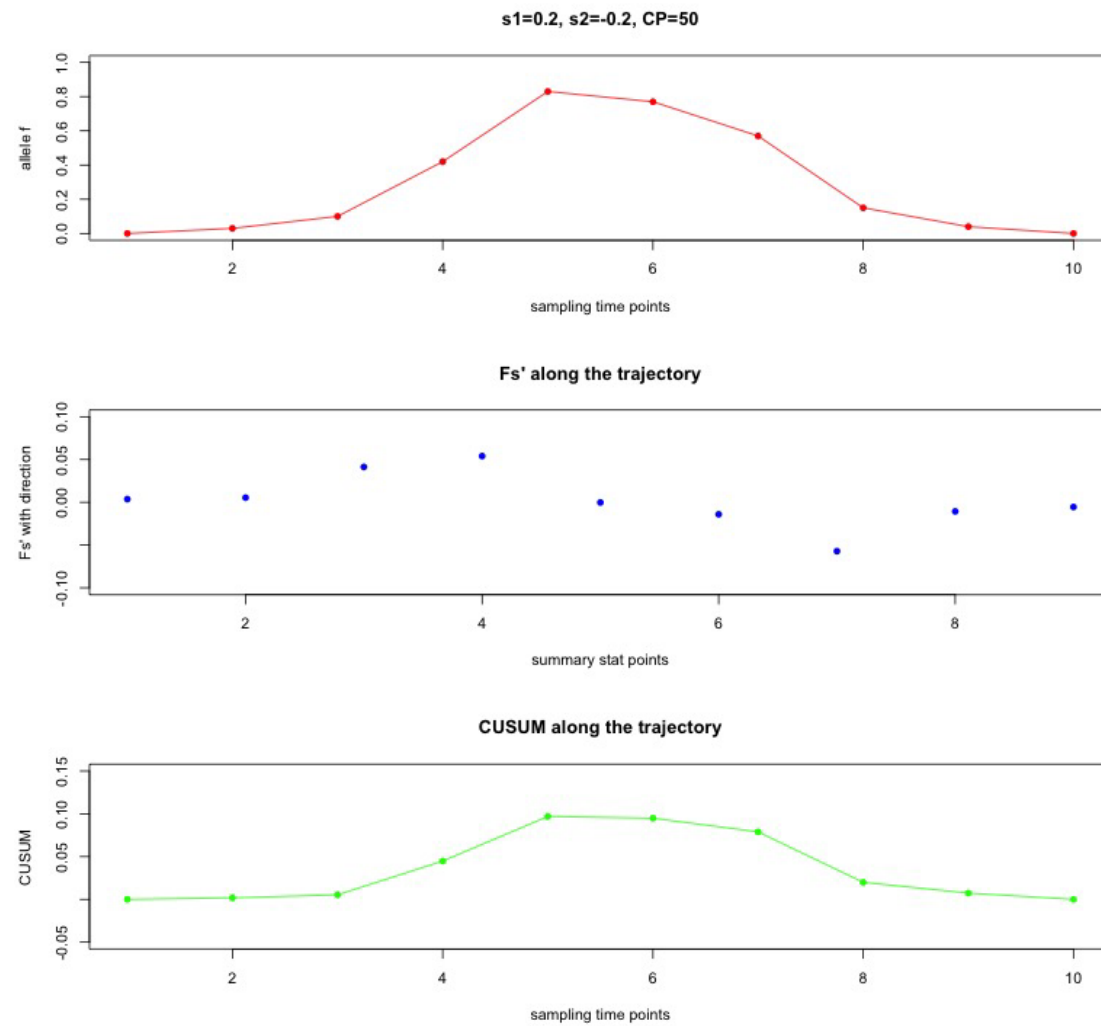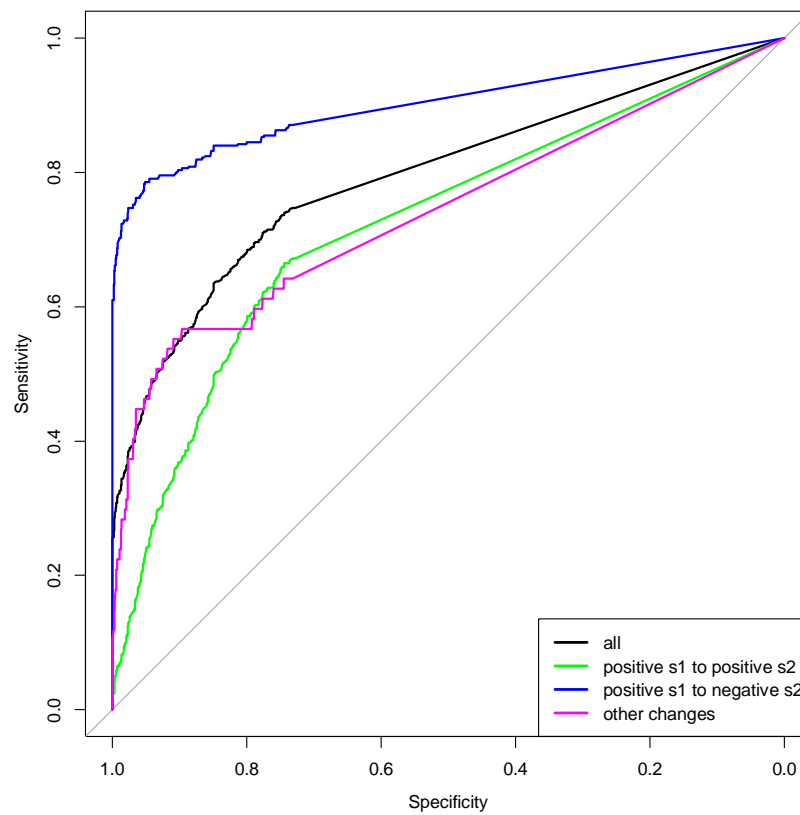

33

Figure 2. ROC curve of the Bayes factor $B_{1,0}$ from the ABC model choice of a haploid population with N_e=100 (A) and a diploid population with N_e=500 (B).
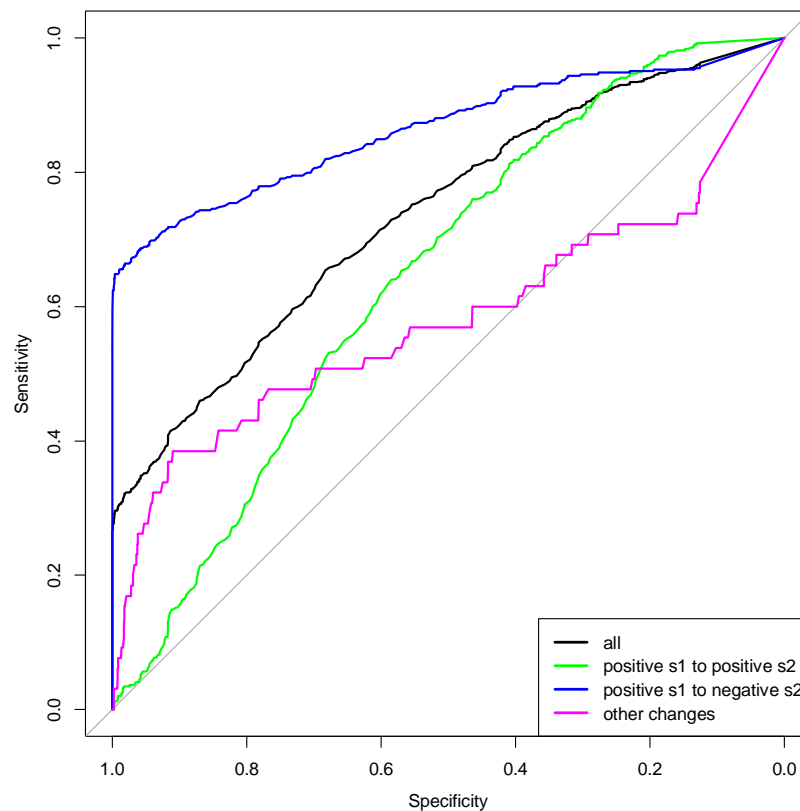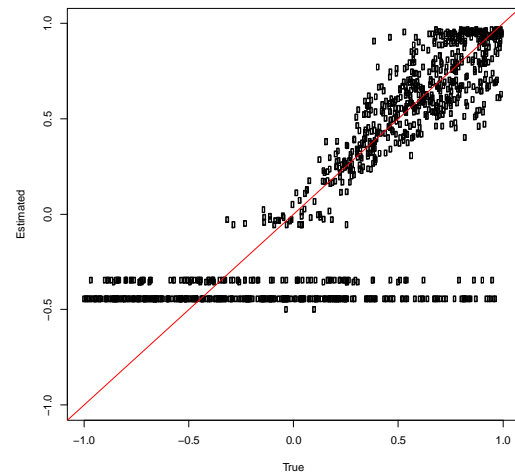
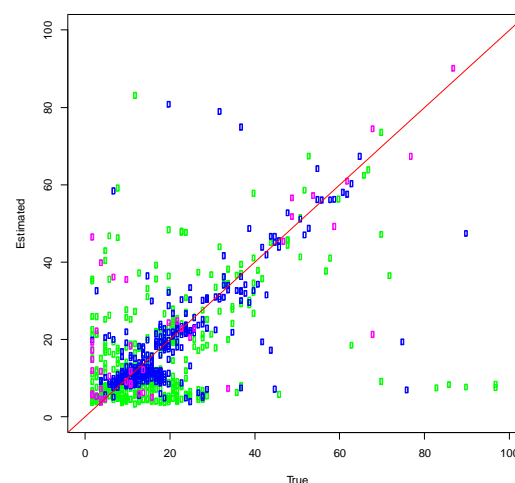**Figure 3.** ABC model choice parameter estimations for 1000 pseudo-observables with a haploid population of $N_e=100$. Each circle is the mode of the posterior distribution from the 0.1% best simulations.

Case 1: Pseudo-observables with constant selection (A) M_0 estimates of s

A

Case 2: Pseudo-observables with changing selection (B) M_1 estimates of $s1$ (C) M_1 estimates of $s2$ (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)

B **C**

D

35
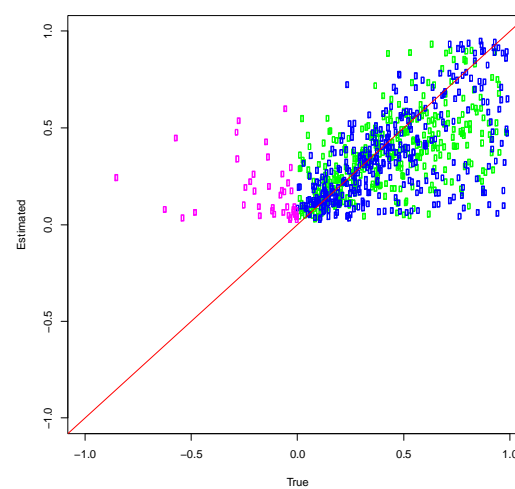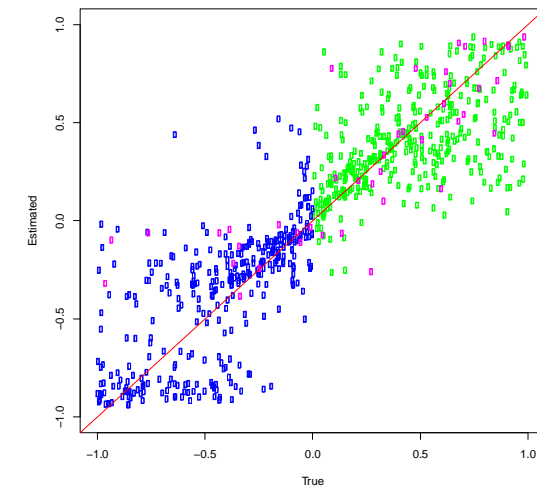
Figure 4. ABC model choice parameter estimations for 1000 pseudo-observables with a diploid population of N_e=500. For cross-validation graphs, each circle is the mode of the posterior distribution from the 0.1% best simulations. For boxplots, red dots are true values and blue dots are average estimated values.

Case 1: Pseudo-observables with constant selection (A) M_0 estimates of s (B) M_0 estimates of h
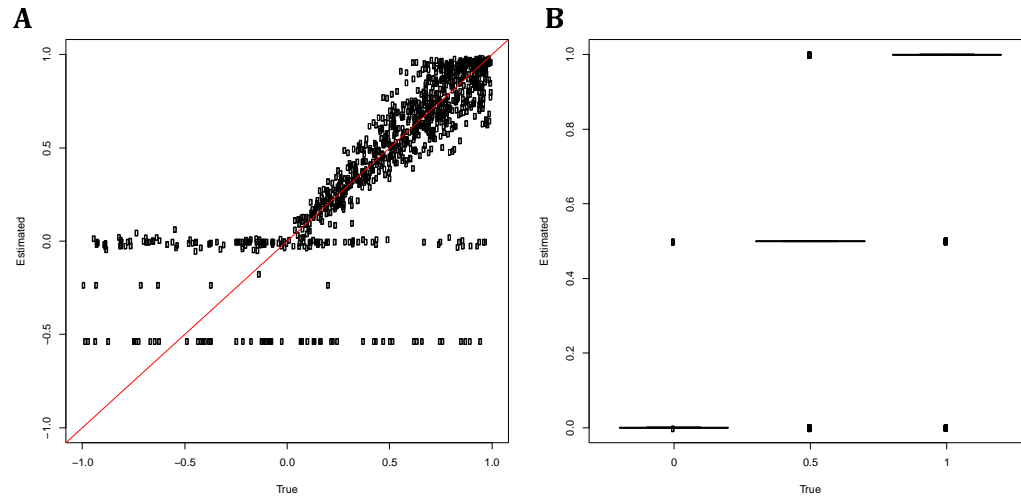
Case 2: Pseudo-observables with changing selection (C) M_1 estimates of $s1$ (D) M_1 estimates of $s2$ (E) M_1 estimates of CP (F) M_1 estimates of h (Green: positive s_1 to positive s_2, Blue: positive s_1 to negative s_2, Magenta: other cases)
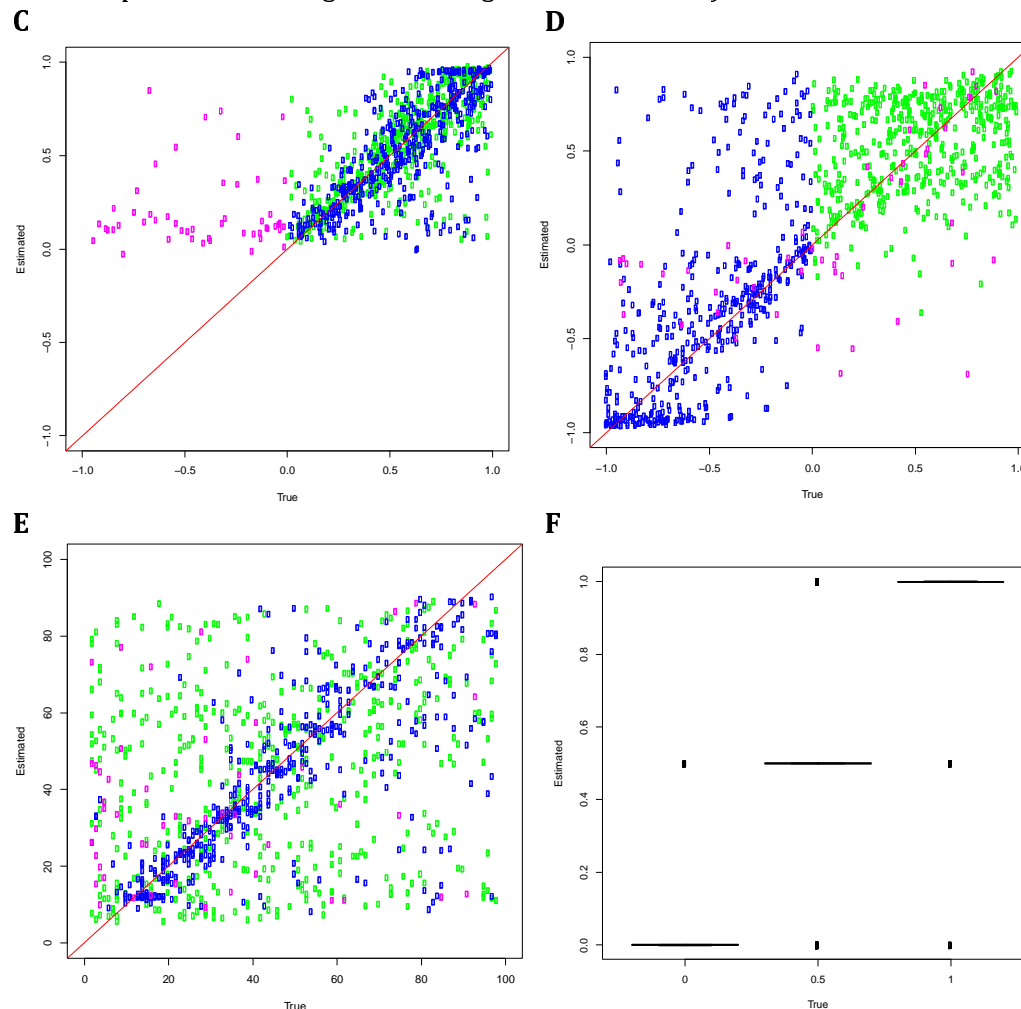
Figure 5. Change points indicated with solid stars for the trajectories of interest: (A) Increasing SNP trajectories in the presence of drug. The red vertical line indicates the sampling time of drug administration. (B) Increasing SNP trajectories in the absence of drugs.

37