

# Detecting Selection in Experimental Evolution Experiments

Arya Iranmehr<sup>1</sup>, Ali Akbari<sup>1</sup>, and Vineet Bafna<sup>2</sup>

<sup>1</sup>Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093, USA.

<sup>2</sup>Computer Science & Engineering, University of California, San Diego, La Jolla, CA 92093, USA

## Abstract

Experimental evolution (EE) studies are powerful tools for observing molecular evolution "in-action" in wild and controlled environments. This paradigm of experiment was infeasible until recently when the whole-genome and whole-population was made possible by next-generation sequencing technologies. However, one of the primary constraints of the EE studies is the limited time for the experiment, which primarily depend on the organism's generation time. This constraint impedes adaptation and optimization (evolvability) studies, where the population can only evolved and re-sequenced in a small number of generations, relative to the number of generations required for fixation of adaptive allele. Although a powerful library of tests-of-selection has already been developed, they are mainly designed for static data to identify adaptation when the sample is taken close enough (before/after) to the fixation of adaptive allele. In this article, we study the problem of identifying selective sweep in short-term experimental evolution of sexual organisms and propose a Dynamic Composite Likelihood Ratio (DCLR) statistic which computes its score by averaging likelihood ratios of polymorphisms for a genomic region. The likelihood of null (neutral) and alternative (selection) hypotheses calculated using the Wright-Fisher Markov chain model for each variant. Extensive simulation study shows that DCLR achieves higher detection power methods on both soft and hard sweep simulations for various selection strengths. Finally, we apply the DCLR statistic to the controlled experimental evolution of *D. melanogaster* to detect adaptive genes/alleles under alternating cold and hot temperatures.

## 1 Introduction

Genetic adaptation is *the* central evolutionary process and is at the core of some of the greatest challenges facing humanity. For example, HIV's rapid evolution of drug resistance during treatment, makes treatments ineffective and even by prescribing a combination of drugs the speed of adaptation only slowed down [22]. Cancer would be much more straightforward to treat if not for tumor's ability to adapt to anti-cancer drugs [23, 69]. Malaria could be treated with cheap drugs such as quinine instead of being one of the world's worst killers [3, 40]. Crop pests would be manageable with small doses of safe insecticides instead of requiring applications of ever increasing amounts of a diverse array of powerful chemicals[12]. Finally, antibiotic resistance [56] to new antibiotics could return us to the pre-Antibiotic age. Despite enormous examples of rapid adaptation, Messer and Petrov [39] argued that "many, if not most, cases of adaptation are yet to be discovered".

Model for describing genetic adaptations, *selective sweep* [30, 55], take into account of associations of the beneficial mutation with its surrounding loci. The extent of genetic loci that are in association with the adaptive allele depends on the amount of accumulated recombination events between adaptive allele and the rest of genome. In the asexual populations, where no recombination event occurs, the whole genome is perfectly linked to the adaptive allele and "clonal interference"

[15, 33] describes the adaptation process when more than one beneficial allele exist in the population. On the other hand, in sexual population, the favored mutation is only in linkage-disequilibrium (LD) with its nearby polymorphisms. Hence, methods for identifying selective sweep in sexual populations often analyze polymorphism data of a population of in a genomic region, rather than a single site.

Adaptation leaves a variety of signatures in different kinds of genomic data, and methods for identifying natural selection are essentially *data-driven*. For instance, reduction in genetic diversity[20, 51, 59] in allele-frequency data, prevalence of long haplotypes [52, 63] in haplotype (phased) data, population differentiation [11, 25] in multiple-population data and rapid increase in allele frequencies [7] in the dynamic data are different signatures of selective sweep in the polymorphism data. In the experimental evolution with pooled-sequencing, only allele frequency of the population, and sometimes initial population’s haplotypes are available.

Traditionally, given static frequency data, Tajima’s  $D$  [59], Fay and Wu’s  $H$  [20], Composite Likelihood Ratio [42], SFSelect [51] test statistics can be computed using the distribution of polymorphisms, site-frequency spectrum (SFS), in a genomic region<sup>1</sup>. Despite their simplicity and clarity, it has been shown that SFS-based tests often fail to distinguish demographic processes from adaptation. They are also prone to pathological false-positive/negatives due to low linkage of the adaptive allele to its surrounding variation and ascertainment bias [2, 39, 41, 46, 47].

On the other hand, tests-of-selection for dynamic data is less studied. Feder et. al [21] proposed Frequency Increment Test(FIT) for dynamic frequency data and Likelihood Ratio Test (LRT) for frequency/haplotype dynamic single-population data. Burke et al. [11] applied Fisher exact test to the last observation of data on case and control populations. Bergland et. al [7] applied  $F_{st}$  to populations throughout time to signify their differentiation from ancestral as well as geographically different populations. Jha et al. [29] computed test statistic of generalized linear-mixed model (GLMM) directly from read counts. Song et al. [60] computed LRT statistic by fitting parameters to a Gaussian process model. Nevertheless, their detection power under short-term experimental evolutions are yet to be known.

**Experimental Evolution** Recent advances of whole genome sequencing has enabled us to sequence populations at a reasonable cost, such that one could design experiments to study different forces of evolution in real-time. *Experimental evolution* is the study of the evolutionary processes of a model organism in a controlled [6, 8, 24, 32, 33, 43, 44] or natural [4, 7, 13, 14, 35, 48, 68] environment. In this setting, constraints such as small population sizes, limited timescales and oversimplified laboratory environments limits interpreting experimental evolution results, yet, they can be used to test different hypotheses [31] regarding mutation and adaptation, genetic drift and inbreeding, environmental variability, sexual selection and conflict, kin selection and cooperation life history and sex allocation, sexual reproduction and mating systems, behavior and cognition, hostparasite interactions, speciation repeatability of evolution and make more accurate inferences than static data analysis [10, 16, 53]. For example, dynamic data has been used to estimate model parameters including population size [45, 60, 64–66] strength of selection [9, 26, 27, 34, 34, 37, 57, 60], allele age [34] recombination rate [60], mutation rate [5, 60] and test neutrality hypotheses [7, 11, 21, 60].

Among different types of evolution experiments [5, 54] in this paper we only focus on adaptive evolution of multicellular sexual organisms with continues culture, fixed population size. For this setting, *D. melanogaster* is usually the model organism and it has been used to identify adaptive genes in longevity and aging [11, 49] (600 and generations respectively), courtship song [62] (100 generations), hypoxia tolerance [70] (300 generations), adaptation to new temperatures[43, 61] (37

---

<sup>1</sup>The extent of genomic region is mainly depend on the amount background linkage in the genome.

generations), egg size [29] (40 generations), C virus resistance [36] (20 generations), and dark-fly [28] (49 generations) experiments.

## 2 Background

**Notation** Let  $\mathbf{X} \triangleq (X_{ijk}) \in [0, 1]^{T \times M \times R}$  denote the population frequency where  $T$  is the number of samples in time,  $M$  is the number of segregating sites, and  $R$  is the number of replicates. Samples in time are taken such that  $\tau_1 < \tau_2 < \dots < \tau_T$ . To simplify our notation, we define  $x_t$  as a allele frequency of a site in a replicate, and we denote it by  $\nu_t$  if such a site is directly under selection with the strength  $s$ . I.E.,  $\nu_t$  can be regarded as frequency of the carrier in the population.

Should we use  $\backslash\nu_t(s)$

**Simulations** For each experiment a diploid population is created and evolved as follows.

- I. **Creating initial founder line haplotypes** First using msms program we created neutral populations for  $F$  founding haplotypes with *default* parameters `./msms <F> 1 -t <2μLNe> -r <2rNeL> <L>` where  $F = 200$  is number of founder lines,  $N_e = 10^6$  is effective population size,  $r = 2 * 10^{-8}$  is recombination rate and  $\mu = 1 * 10^{-9}$  is mutation rate and  $L = 50K$  is the window size in base pairs which gives  $\theta = 2\mu N_e L = 200$  and  $\rho = 2N_e r L = 2000$ . For default parameter, the expected number of segregating sites in a window is

$$\mathbb{E}[M] = \theta \sum_{i=1}^{F-1} \frac{1}{i} = 1175$$

This detail is out of place here.  
You should have generalities  
here, and move the rest to  
Methods

- II. **Creating initial diploid population** To implement similar setting for experimental evolution of diploid organisms, initial haplotypes first cloned to create  $F$  diploid homozygotes. Then each diploid individual is cloned  $N/F$  times to yield diploid population of size  $N$ .

- III. **Forward Simulation** Given initial diploid population, position of the site under selection, selection strength  $s$ , number of replicates  $R = 3$ , recombination rate  $r = 2 \times 10^{-8}$  and sampling times  $\mathcal{T} = \{10, 20, 30, 40, 50\}$ , simuPop is used to perform forward simulation and compute allele frequencies for all of the  $R$  replicates. Also, to avoid spurious simulation samples, simulation results are constrained to those that the beneficial allele escapes stochastic loss of genetic drift and *establishes* in all the replicates.

**Hard and Soft Sweep** In this paper, we only study (positive) selection with single beneficial allele under selective sweep, i.e. favored allele is in linkage disequilibrium (LD) with its nearby loci. Selective sweeps are classified by the amount of variation exist in the individuals carrying adaptive allele. By definition, hard sweep is the case when all the carriers coalesce after onset of selection and thus diversity between carriers is at its minimum. In general, soft sweep is difficult to detect. We conduct experiments, so that evaluate prediction performance of the proposed method under hard and soft sweeps with different selection strengths. Since here we consider random mating with no de novo mutations, soft sweep can only happen in standing variation where the site under selection is at frequency is larger than  $1/F$ . In contrast, hard sweep experiments are those that their adaptive allele is at its minim frequency,  $1/F$ , at the onset of selection.

**Challenges** Instead, they observed incomplete fixation ('soft-sweeps') due in part to standing variation,

The main constraint in the experimental evolution experiments, is the sampling-time-span (STS)  $\tau = \tau_T - \tau_0$ , the number of generations between the first and last generation in the time-series data,

## What does this mean? reword it a bit

which primarily determined organism generation time. In other words, for wide range of values of  $s$  (not very strong selection), STS will be significantly smaller than fixation time. As a result, methods for detecting selection on static data, which work the best when the beneficial allele is close to fixation, would not work properly for dynamic data.

evolved and inbred

Moreover, in controlled experimental evolution, a population is evolve and inbreed under ~~a~~ selective pressure. This scenario ~~in~~ which population size ~~is~~ effectively reduces from large number of wild type (e.g.D. melanogaster  $N_e \approx 10^6$  ) to a small number (typically  $F$  is between 100-1000) of founder lines for EE, resembles a severe population bottleneck. This phenomena makes genetic drift quite strong and confounds polymorphism summary statistics and makes difficult to detecting selection weaker selections.

Complicated. Explain with an example about why drift becomes strong

Let us re-word. The points of this paragraph are well recognized in the community, and we do not want to irritate them by stating the obvious.

### 3 Results

Selective sweep is a complex random process which under infinite site assumption only depends on selection strength, population size and the underlying genealogy at the onset of selection. As a result, the change in different population statistics, including *site allele frequencies*(AF), *site frequency spectrum*(SFS) and *linkage disequilibrium*(LD), can be attributed to the selection event. However, other independent processes such as genetic drift, population expansion/bottleneck and recombination cause AF,SFS and LD to change, respectively. Therefore, obtaining a clear understanding of dynamics of selective sweep requires disentangling effects of natural selection from those of neutral evolution and demography.

In latex use two back quotes ``

In the following, we examine the dynamics of AF and SFS in the experimental evolution and evaluate their power in identification of selection events. The "identification" of a selection event can be done in different levels of detail. At coarsest level, identification can be done by determining whether a region (e.g. a window of 50Kbp, or a gene) on genome is under selection. In the rest, we consider this task to be the task of *detection*. Then, finding the causal mutation/allele would be a more elaborate identification of selection, henceforth, *locating* selection. Finally, estimating model parameters such as strength of selection and overdominance at the site fully describes the selective sweep.

Instead of 50kbp, say "a small region with no or low recombination"

Use

period at Allele Frequencies Selection and genetic drift are the two main *independent* evolutionary forces that cause change in allele frequencies, each generation. In the simplest case, models for neutral evolution of a single-locus can be used to derive a null distribution in order to detect the region under selection or locate the adaptive allele. For example, Feder et al. [21] proposed a (continues-time continues-state) Brownian motion process for modeling variation of allele frequencies under genetic drift. Basically, given current allele frequency  $\nu_t$  at a site, Brownian motion approximation assumes future-generations allele frequencies are drawn from the Gaussian distribution

$$\nu_{t+\tau} \sim \mathcal{N}\left(\nu_t, \frac{2\nu_t(1-\nu_t)}{N_e}\tau\right) \quad \text{why } \backslash \text{nu and not } x? \quad (1)$$

Using (1), Frequency Increment Test (FIT) [21] computes p-value of time-series data by a Student's t-test. However, Brownian motion is a poor approximation of the actual genetic drift process when 1)  $\nu_t$  is not close to 0.5, 2) when  $\tau$  is large. At the start of selection site frequencies are distributed according to SFS of a neutral population, and thus most of the sites are at low frequencies Figure 1. Thus if the selecting site is chosen randomly with high probability its frequency will be far 0.5. Moreover, in many experimental evolution experiments STS  $\tau$  is chosen be large, e.g. of the order of 10 to 100 generations.

Instead, probabilities of trajectories can be computed using a (discrete-time discrete-state-space) Markov Chain with transition matrix  $P$

$$P_{i,j} = \Pr\left(\nu_{t+1} = \frac{j}{2N} \mid \nu_t = \frac{i}{2N}\right) = \binom{2N}{j} \nu_t^j (1-\nu_t)^{2N-j} \quad (2)$$

which transition probabilities for  $\tau$  generations ahead can be computed by powering the transition matrix to  $\tau$ . Figure 3 illustrates expected and observed distribution of frequencies after 1, 10 and 100 generations when starting frequency is 0.005 (top) and 0.1 (bottom) for Brownian motion and Markov chain. Brownian motion exhibits poor prediction of distributions of allele frequencies when starting frequency is far from 0.5. On the other hand, Markov chain provides an accurate distribution of allele frequencies for future generations, under genetic drift.

On the other hand, alternative models for single-locus selection can be utilized to estimate strength of selection and robustly detect and locate selection by evaluating two hypotheses in order

repeated phrase

to perform likelihood ratio tests. In fact, it has been shown that the power of identifying selection can be improved by exploiting the mechanisms that selection changes allele frequencies [42] <sup>2</sup>.

In a pure selection process with no drift, i.e. infinite population size, dynamic of allele frequencies can be well approximated by the logistic function (see Appendix 6.1 for derivation)

$$\nu_t = \sigma(st + \eta(\nu_0)) \quad \text{Only true for one type of soft sweep} \quad (3)$$

where  $\sigma(x) = 1/(1 + e^{-x})$  is the logistic function, and  $\eta(x) = \log(x)/\log(1 - x)$  is inverse of the logistic, aka logit, function. Figure 2 depicts the behavior of the logistic model for site allele frequencies and the default sampling time span(STS). Without genetic drift, soft sweeps are easier to detect, because the logistic function happens to have steeper slope in the STS than those of hard sweeps. In addition, even in infinite population size, it is difficult to differentiate between weak selections and genetic drift (a horizontal line), in 50 generations.

As shown in Figures 8 and 7 (first row), the approximate logistic model is consistent with simulated data and we use it to estimate the strength of selection for each site by a simple regression (see 4.1 for details).

#### Funny break between previous and next paragraph

**Site Frequency Spectrum** SFS (Figure 1) is the distribution of allele frequencies in a genomic region. Historically, SFS of static data has been extensively used to "detect" genetic adaption and demographic changes in a population. Basically, SFS is measure of diversity in a genomic region computed from allele frequencies. In general, reduction in diversity is a signal of selection, and detecting selection based on the reduction in genomic diversity is a subtle task, because

- (i) in soft sweeps the genomic diversity is not necessarily reduced.
- (ii) even in hard sweeps with no recombination, the reduction diversity is significant only when the SFS sample is taken close to fixation (not too far prior or after fixation)
- (iii) SFS change in the same way for both selection and demography changes.

Although conditions (hard sweep with no recombination, not far from fixation, and random-mating and constant size population ) for detecting selection based on SFS are very restrictive, SFS-based tests are simple and inexpensive to use and often used in combination with other tests [].

In practice, SFS distribution is scalarized by a weighted linear combination[1]. More precisely, the weights should be positive and sum to one, which called a convex combination, aka discrete expectation. For example, test statistics for Tajima's  $D$ [59], Fay Wu's  $H$  [20] and SFSelect[51] can be obtained by a dot product of SFS vector with their weight vector. cite achasz?

Whether the dynamic of SFS in times-series can improve the detection power in the conditions (i-iii) is not well-studied yet. Evans et al. [17] developed diffusion equations for evolution of SFS in time series, they are difficult to solve. Hence, we analyze and visualize scalar test statistics  $D$ ,  $H$  and SFSelect in the dynamic data under different selection regimes.

To adapt SFS-based test for dynamic data, one can simply design statistical tests to examine whether the observations of  $D$ ,  $H$  and SFSelect statistics are significantly deviated from the null distribution, i.e. constant values in time. On the other hand, we can model the dynamics of SFS as a function of selection strength to 1)have a model for estimating model parameter 2)obtain an insight regarding how does SFS statistics behave under different stages of selection.

As mentioned before, restricting population to  $F$  founder lines ( $F \ll N_e$ ) resembles a severe bottleneck event, which confounds SFS. Figure 6 demonstrates effect of controlled experimental

<sup>2</sup>Basically what they did is quite similar to ours. First they came up with a neutrality test, then they extend it by a parametric formulation which they can estimate  $s$  and compute likelihood.

This looks out of place so let us discuss where to put it. We shouldn't mix results and methods, even though this is a result that motivates a method. Let's discuss

evolution on different SFS statistics under neutral evolution for 1000 simulations. The mean of neutral simulations can be used to empirically filter out the effect of bottleneck in dynamic data.

I. **Tajima's D.** As shown in the Appendix 6.2, dynamic of Tajima's D in hard sweep is

$$D_t = D_0 - \log(1 - \nu_t) \frac{W_0}{\log(2N)} - \nu_t^2 \Pi_0 \quad (4)$$

where  $W_0$  and  $\Pi_0$  are Watterson and Tajima estimate of  $\theta$  at the first generation. Given that  $\nu_t$  defined in term of  $\nu_0$  and  $s$  in (3), we can use (4) to estimate the strength of selection by performing a nonlinear regression to find  $s$ . Then, a likelihood ratio, similar to (12), provides<sup>3</sup> a predictor for detecting selection in each window. Figures 8 and 7 (last row, left), shows that the parametric model with the actual value of  $s$  provides is consistent with 1000 simulations, in strong and weak selections. Also, (4) represents dynamic of  $D$  in hard sweep, and the model becomes invalid as initial carrier frequency increases, Figure 8 and 7 (last row, right).

Finally, even in a infinite population size setting, differentiating the between selection and drift is difficult in early generations. W.l.o.g when  $D_0 = 0$  and  $\Pi_0 = W_0 = 1$ ,  $D_t$  is sum of the logarithmic  $\frac{-\log(1-\nu_t)}{\log(2N)}$  and the squared term  $\nu_t^2$ , blue and green curves in Figure 5. As shown in the Figure 5 right, in early generations of hard sweep where carrier frequency is low,  $D_t$  is either positive or close zero. In other words, the reduction in diversity become significant when carrier frequency is high.

II. **Fay Wu's H** As shown in the Appendix 6.1.1, dynamic  $H$  statistic is directly related to average HAF [50] and can be written

$$nH_t = \theta\nu_t \left( \frac{\nu_t + 1}{2} - \frac{1}{(1 - \nu_t)n + 1} \right) + \theta(1 - \nu_t) \left( \frac{n + 1}{2n} - \frac{1}{(1 - \nu_t)n + 1} \right) \quad (5)$$

which is a good approximation only in hard sweep with strong selection, Figures 8 and 7 (second row, left). Similar to  $D$ , it can be shown that  $H$  statistics for natural selection behave similar to that of neutral evolution in early generations of hard sweep, where carrier frequency is not high.

III. **SFSelect** statistic proposed by Ronen et al. [51] to predict selection by empirically learning the so-called *optimal* weights using Support Vector Machines. Figure 8 and 7 (forth row), depicts the dynamics of SFSelect for selection and neutral evolution in time, which essentially behave similar to  $D$  and  $H$ , i.e. poor power of detection in soft sweep, and early generations of hard sweep.

### 3.1 Detecting Selection

We need to choose a better name,  
and we need to formally define LS  
here

To evaluate detection power of the proposed method, least-squares (LS), we compare it with Gaussian process (GP) [60], FIT [21], SFselect [50],  $D$  [59], and  $H$  [20] statistics. We also take time-series models of  $D$  and  $H$  into comparison by fitting models (4) (5) to data. Finally, since SFSelect is positive and monotone, we considered the aggregation of SFSelect statistics along time-series as score for detecting selection.

For every setting, i.e. strength of selection, initial carrier frequency, sampling time span and number of replicates, we conduct 200 simulations which half of them are neutral and the rest are

<sup>3</sup>The likelihood of the data to the model is the least-squares loss between model  $D_t(\hat{s})$  and the observed  $D$ .

of which half are neutral, ..

under selection. Then prediction score of each method for all the simulations is used to compute Receiver Operating Characteristic (ROC) curve [19], which is the primary tool for evaluating binary decision making hypotheses. Area under ROC curve (AUC), is measure of predictive performance of a binary hypothesis, but it computes the score under all possible setting of false-positive rate (FPR). However, here we restrict AUC [38] to those region that has  $\text{FPR} \leq 0.1$ , and henceforth we call  $0.1\text{-AUC} \times 1000$  as power of a method<sup>4</sup>. Note that power ranges between 0 to 100 and power of 5 corresponds to random prediction, i.e. null hypothesis.

First we compare all the methods under default parameter settings, i.e.,  $T = 5$ ,  $\tau = 50$  and  $R = 3$ , Figure 13. It is evident that LS method provides better detection power in the hard and soft sweep regimes for different values of selection strength. Also, except for H in strong selections, SFs-based scores exhibit poor performance, comparable to random prediction. Also, GP and FIT methods which are single-locus methods, reveal their best performance in softest sweep scenario,  $\nu_0 = 0.1$ .

Also, we assessed the behavior of LS method for different values of  $\tau \in \{10, 50, 100\}$ , Figure 14. For the default set of parameters, and  $\tau = 100$ , LS method provides significantly better performance, while for  $\tau = 10$  detection power is comparable to those of other methods for  $\tau = 50$  in Figure 13. Effects of number of replicates on the power of LS method is examined for  $R \in \{3, 5, 10\}$ , Figure 15. Interestingly, increasing number of replicates does not necessarily improve power, yet they provide comparable power to the case with  $R = 3$ . This observation implies that estimates of power for each method tightly depend on the replicates used for evaluation, and computed power of a method potentially under/over-estimated due to finite sample size, i.e. *estimation error*. Thus, comparison using small number of replicate is meaningful only when difference in performance is significant and consistent for different configuration, which is the case for LS method in Figure 13. Finally, Figure 16, demonstrates the power for  $R = \{3, 10\}$  and  $\tau = 50, 100$ , which shows power for  $R = 3$  and  $R = 10$  exhibits the similar behavior under different regimes for  $\nu_0$ ,  $s$  and  $\tau$ .

### 3.2 Locating the Adaptive Mutation

The secondary task in identifying selection is to locate the position of the adaptive allele. Since de novo mutations are excluded, and linkage is reinforced after selection, it is difficult to incorporate phylogeny into this task. We consider the site with highest score as the locus of the beneficial allele. For each setting of  $\nu_0$  and  $s$ , we conducted 100 simulations and computed the rank of the beneficial mutation in the sorted scores. Then for each configuration distribution of ranks of 100 simulation is presented via their Cumulative Distribution Function (CDF), Figure 18. LS works the best in hard sweep regime. For example, when  $\nu_0 = 0.005$  and  $s = 0.01$  (weakest selection), the beneficial allele is ranked first, in more than 60 experiments, and take rank of less than 5 in more than 90% of experiments of hard sweep. Accuracy of the locating the adaptive allele diminished as sweep become softer, i.e. larger values of  $\nu_0$ . Yet, in the worst case the beneficial allele is tanked among top 50 SNPs.

### 3.3 Strength of Selection

Finally, estimating the model parameters such as strength of selection provides the finest detail of identification of a selective sweep process, because given the model parameters we can predict the state of population, e.g. fixation time, in future generations. We computed bias,  $|s - \hat{s}|$  for each experiment of LS, and GP. The distribution of the bias is presented in Figure 17 for different configurations. In general, both GP and LS are biased for weak selections, which genetic drift

---

<sup>4</sup>Power can be interpreted as average true positive rate when  $\text{FPR} \leq 0.1$ .

dominates. However, for stronger selections, e.g.  $s = 0.1$ , LS provides estimates with smaller bias and variance.

### 3.4 Analysis of Real Data

We finally apply LS method to the controlled experimental evolution experiment of [43], which evolves 5 replicates of a population of Drosophila melanogaster for 37 generations under alternating 12-hour cycles of hot (28°C) and cold (18°C) temperatures. Three replicates are sampled at the first generation, 2 replicates at generation 15, one replicate at generation 23, one replicate at generation 27 and three replicates at generation 37.

Figure 6.1 demonstrates Manhattan plot for the scan of data. Genes will be extracted ASAP.

## 4 Materials and Methods

### 4.1 Least-Squares Estimate of $s$

why introduce  $T$  as a new parameter? Is  $T$  the number of observations, why is it not in the main text

In a simplest case, we have 2 observations,  $\nu_0$  and  $\nu_t$  and we wish to find  $\hat{s}$  so that  $\nu_t = \sigma(\hat{s}t/2 + \eta(\nu_0))$  which can be solved analytically:

$$\hat{s} = \frac{2}{t} \log \left( \frac{\nu_t(1 - \nu_0)}{\nu_0(1 - \nu_t)} \right) = \frac{2}{t} (\eta(\nu_t) - \eta(\nu_0)) \quad (6)$$

which selection strength can be interpreted as the normalized frequency increment in log-odds space.

However, with more than two observations, the system becomes over determined and we need to solve  $T$  set of equations. By re-arranging (6), we can write a set of linear equations

$$ts = 2(\eta(\nu_t) - \eta(\nu_0)) \quad \forall t \in \mathcal{T} \quad (7)$$

which can be written in matrix form  $\mathbf{t}s = \mathbf{b}$  with

$$\mathbf{t} = \begin{bmatrix} \tau_1 \\ \vdots \\ \tau_T \end{bmatrix} \mathbf{b} = \begin{bmatrix} 2(\eta(\nu_{\tau_1}) - \eta(\nu_0)) \\ \vdots \\ 2(\eta(\nu_{\tau_T}) - \eta(\nu_0)) \end{bmatrix}, \quad (8)$$

and the solution for each replicate can be computed using pseudo inverse

$$\hat{s}_r = \mathbf{t}^+ \mathbf{b} = (\mathbf{t}^T \mathbf{t})^{-1} \mathbf{t}^T \mathbf{b} = \frac{\mathbf{t}^T \mathbf{b}}{\|\mathbf{t}\|^2} \quad (9)$$

and for multiple independent replicates with the same sampling times we have

$$\hat{s} = \frac{1}{R} \sum_r^R \hat{s}_r = \mathbf{t}^+ \bar{\mathbf{b}} \quad (10)$$

where  $\bar{\mathbf{b}}$  is the average of  $\mathbf{b}$  over  $R$  replicates. Finally, for each replicate, instead of  $T$  equations where each observation is compared with  $x_0$ , we can take all possible pairs and solve  $\binom{T}{2}$  set of equations.

Better to just start with  $\text{choose } 2$

## 4.2 Composite Likelihood Ratio Test

Given an estimate for  $s$ , the Markov chain with transition probabilities ([18], eq 1.58-59)

$$Q_{i,j}(s) = \Pr\left(\nu_{t+1} = \frac{j}{2N} \mid \nu_t = \frac{i}{2N}\right) = \binom{2N}{j} \hat{\nu}_{t+1}^j (1 - \hat{\nu}_{t+1})^{2N-j}, \quad (11)$$

$$\hat{\nu}_{t+1} = \sigma(\hat{s} + \eta(\nu_t))$$

computes probabilities of time-series under the alternative model. Figure 4 shows that the predicted distributions of frequencies using Markov chain 11 is consistent with simulation data.

Likelihood of the trajectories under null and alternative hypothesis can be computed to perform likelihood ratio tests which provide predictors that are robust to pathological false-positives [21]. Specifically, for each site, we define the likelihood-ratio predictor

$$\Lambda(\hat{s}) = \hat{s} \log \left( \frac{\mathcal{L}(\mathbf{x}|s = \hat{s})}{\mathcal{L}(\mathbf{x}|s = 0)} \right) \quad (12)$$

where  $\mathbf{x}$  is the vector of trajectories and  $\mathcal{L}$  is the likelihood function defined by the Markov chain with transition (11).

Given likelihood ratio (12) for each SNP we can further exclude false positives by combining likelihood ratio scores of the polymorphisms in a genomic region by computing Composite Likelihood Ratio(CLR)[42, 63, 67]. For each window, CLR is computed by taking average of top one percentile of the likelihood ratios.

## 5 Discussion

## 6 Appendix

### 6.1 Allele frequencies under selective sweep

Let in a dialelic diploid Wright-Fisher model with single locus selection, with genotypes 0—0, 0—1 and 1—1,  $w_{00}$ ,  $w_{01}$  and  $w_{11}$  be the relative fitness of each genotype at a locus. Also, consider  $x$  to be frequency of the allele 1 and allele 1 be the favored allele, then

$$w_{00} = 1, \quad w_{01} = 1 + us, \quad w_{11} = 1 + s \quad (13)$$

$$x_{t+1} = \frac{w_{11}x_t^2 + w_{01}x_t(1 - x_t)}{w_{11}x_t^2 + 2w_{01}x_t(1 - x_t) + w_{00}(1 - x_t)^2} = x_t + \frac{s(h + (1 - 2u)x_t)x_t(1 - x_t)}{1 + sx_t(2u + (1 - 2u)x_t)} \quad (14)$$

where  $s \in \mathbb{R}$  is the selection coefficient and  $o \in [0, 1]$  is the overdominance parameter which for  $u = 0.5$  we have

$$x_{t+1} = x_t + \frac{sx_t(1 - x_t)}{2 + 2sx_t}. \quad (15)$$

we also have

$$\frac{dx_t}{dt} = \frac{sx_t(1 - x_t)}{2 + 2sx_t} \quad (16)$$

which is a differential equation that is difficult to solve. However if take the approximation  $2 + 2sx_t \approx 2$ , it becomes an ordinary differential equation that can be readily solved

$$\nu_t = \frac{1}{1 + \frac{1-x_0}{x_0} e^{-st/2}} = \sigma(st/2 + \eta(x_0)) \quad (17)$$

where  $\sigma(\cdot)$  is the logistic function and  $\eta(\cdot)$  is logit function (inverse of the logistic function).

### 6.1.1 Fay Wu's H

**Lemma 1.** In any finite population size of  $n$  with  $m$  segregating sites, allele frequencies take discrete values, i.e.,  $x_j \in \{\frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}\}$ ,  $\forall j \in 1, \dots, m$  and we can write

$$\|\mathbf{x}\|^2 = \sum_{j=1}^m x_j^2 = \sum_{i=1}^{n-1} \left(\frac{i}{n}\right)^2 \xi_i = \frac{(n-1)}{2n} H \quad (18)$$

where  $\xi_i$  is the number of sites with frequency  $i/n$  and  $H$  is the Fay & Wu's estimate of  $\theta$ .

Recently, Ronen et al. [50] devised the 1-HAF statistic for identifying selection on static data, which has the expected value related to the  $\|\mathbf{x}\|^2$ :

$$\mathbb{E}[1\text{-HAF}(t)] = n\|\mathbf{x}_t\|^2 \approx ng(\nu_t) \quad (19)$$

where

$$g(\nu_t) = \theta\nu_t \left( \frac{\nu_t + 1}{2} - \frac{1}{(1-\nu_t)n+1} \right) + \theta(1-\nu_t) \left( \frac{n+1}{2n} - \frac{1}{(1-\nu_t)n+1} \right) \quad (20)$$

which easily follows that

$$\theta_H(t) = \frac{n-1}{2} g(\nu_t) \quad (21)$$

## 6.2 Tajima's D

Let  $D_0, \Pi_0, W_0$ , be Tajima's D, Tajima's estimate of  $\theta$ , and Watterson's estimate of  $\theta$  at time zero and  $D_0 = \Pi_0 - W_0$ . In order to compute,  $D_t = \Pi_t - W_t$  we compute  $\Pi_t$  and  $W_t$  separately as follows.

Let  $P$  be the  $n \times n$  matrix of pairwise heterozygosity if individuals, then  $\Pi = \frac{1}{n^2} \sum P_{ij}$ . So, if the population consist of  $\nu n$  identical carrier haplotype (due to lack of recombination), their pairwise hamming distance is zero and should be subtracted from the total  $\Pi_t$ :

$$\Pi_t = (1 - \nu_t^2)\Pi_0 \quad (22)$$

To compute  $W_t$ , first remember that  $W_t = \frac{m_t}{S_n}$  where  $m_t$  is the number of segregating sites at time  $t$  and  $S_n = \sum_i^n 1/i \approx \log(n)$ . Also we have

$$\frac{W_t}{W_0} = \frac{\frac{m_t}{S}}{\frac{m_0}{S}} \Rightarrow W_t = \frac{m_t}{m_0} W_0 \quad (23)$$

where  $m_t$  to be interpreted as the expected number of segregating sites at time  $t$ , under neutral evolution. At time  $t$ , the number of individuals that undergone neutral evolution is  $(1 - \nu_t)n + 1$ , which leads to

$$\frac{m_t}{m_0} = \frac{\log((1 - \nu_t)n + 1)\theta}{\log(n)\theta} \approx \frac{\log((1 - \nu_t)n)}{\log(n)} = \frac{\log(1 - \nu_t) + \log(n)}{\log(n)} = 1 + \frac{\log(1 - \nu_t)}{\log(n)} \quad (24)$$

putting all together

$$D_t = (1 - \nu_t^2)\Pi_0 - (1 + \frac{\log(1 - \nu_t)}{\log(n)})W_0 = D_0 - \log(1 - \nu_t) \frac{W_0}{\log(n)} - \nu_t^2 \Pi_0 \quad (25)$$

### 6.3 Linkage Disequilibrium

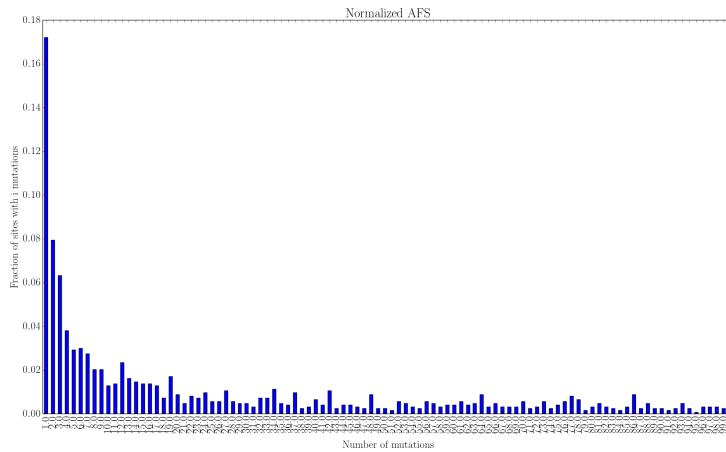
Nonrandom associations between polymorphisms are established in the substitution process according to the phylogeny, broken by recombination events and reinforced by selection. Although in EE the experiments with pooled sequencing, LD can not be measured throughout evolution, it is still worthwhile to examine the behavior of LD as a result of the interaction between recombination and natural selection, to take into account of some of EE implicit constraints.

Let  $\rho_0$  be the LD at time zero between the site under selection and a segregating site  $l$  base-pairs away, then under natural selection we have

$$\rho_t = \alpha_t \beta_t \rho_0 = e^{-rtl} \left( \frac{H_t}{H_0} \right) \rho_0 \quad (26)$$

where  $H_T = 2\nu_0(1-\nu_0)$  is the heterozygosity at the selected site,  $r$  is the recombination rate/bp/gen. The decay factor  $\alpha_t = e^{-rtl}$  is the product of recombination and growth factor  $\beta_t$  (eq. 30-31 in [58]) is the outcome of selection. For  $s = 0.01$ ,  $l = 100\text{Kbp}$ , the log of decay, growth and product of both is depicted in Figure 9. It is evident that, for these parameters LD does not start to decay until generation 1000, which would be problematic when  $\rho_0$ . For example, in the case of hard sweep, the selection is imposed on the site with minimum AF, which is at perfect linkage ( $|D'| = 1$ ) with all the other loci.<sup>5</sup> We this phenomenon is shown in the Figures 10, 11 where at generation zero the site at position 500K is at perfect linkage with all the other sites, and linkage of the middle site with all the genome is depicted for both genetic drift and natural selection, in different generations. Also, a window of 50Kbp around the selected site is shaded in Figure 10 to demonstrate the value of LD in the window under drift and hard sweep. This implies that the precision of locating the selection on the genome is tightly dependent on a set of parameters including, recombination rate, selection strength, initial carrier frequency, and the initial linkage.

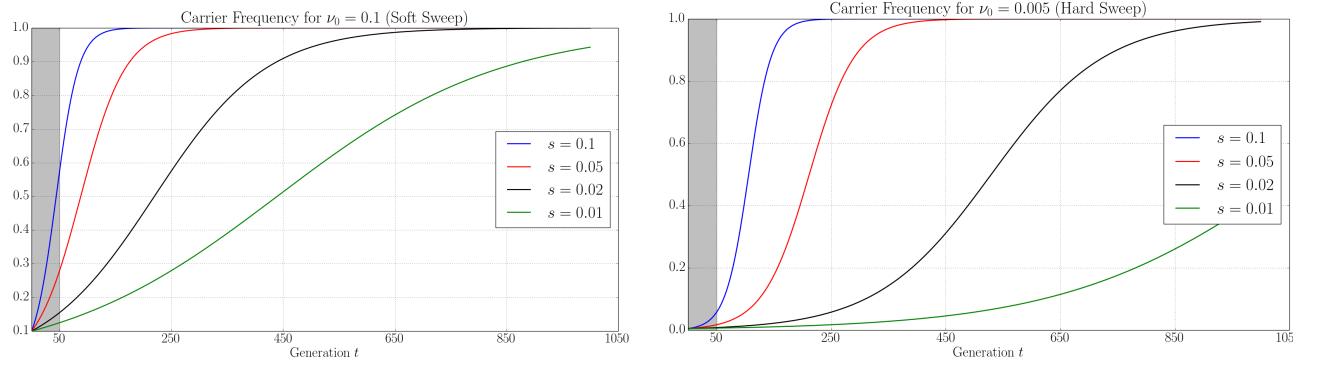
## 7 Figures



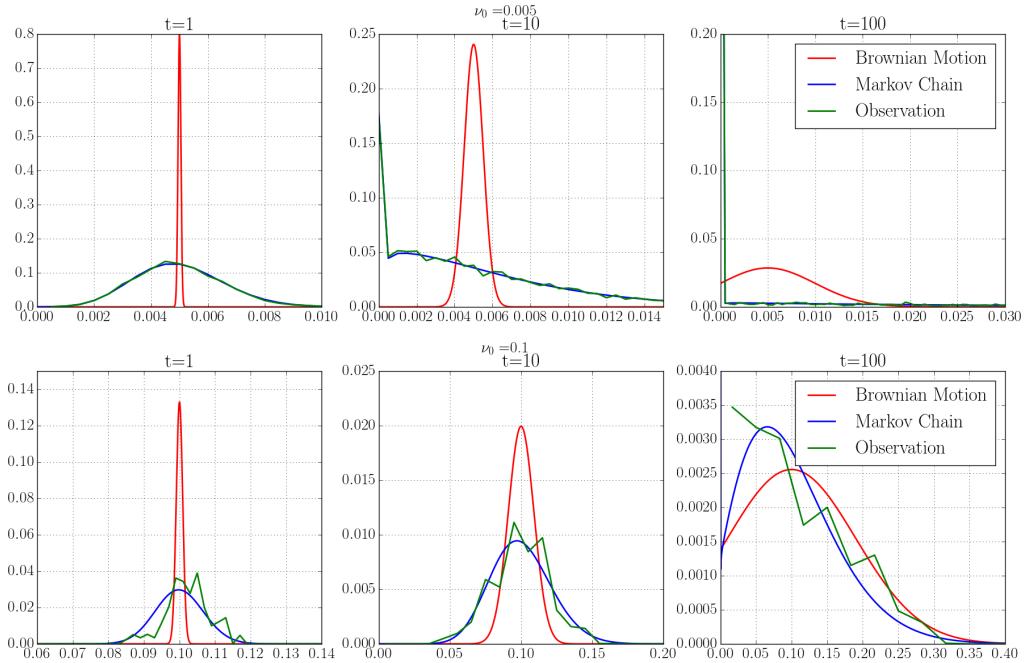
**Figure 1:** SFS

---

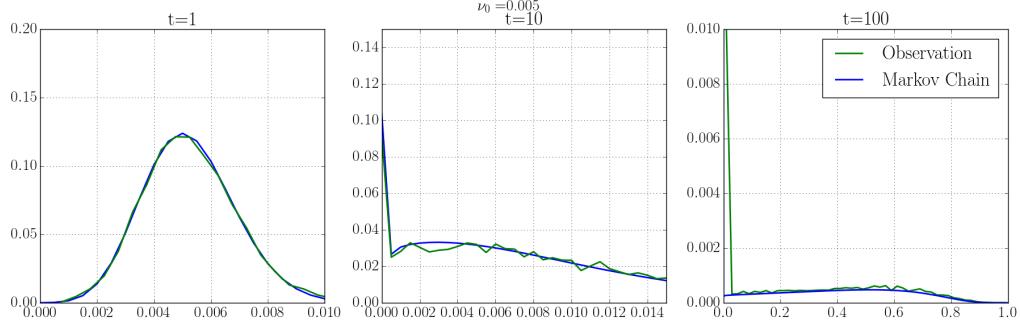
<sup>5</sup>This is because, between the selected site and all the other sites frequency of one gamete is zero.



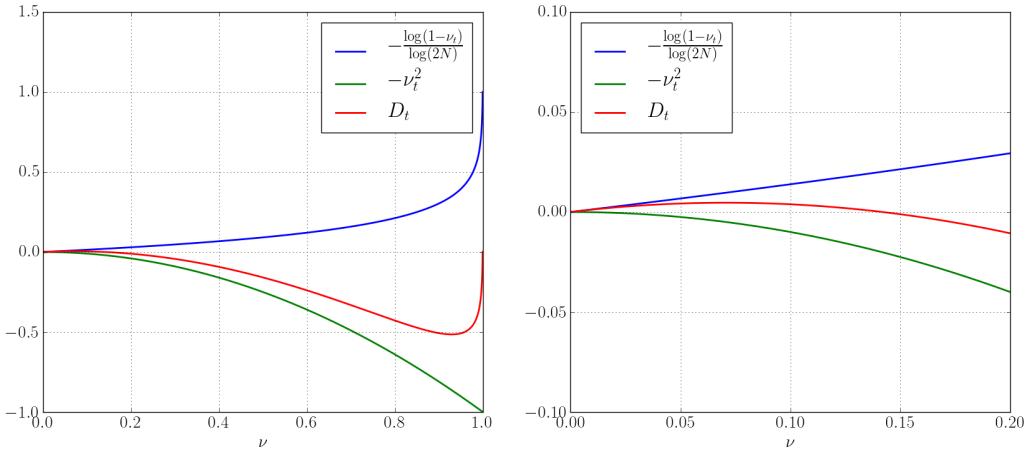
**Figure 2:** Logistic model for different selection strengths for soft (left) and hard (right) sweep as a function of time in generations. The first 50 generations, which observations are sampled is shaded.



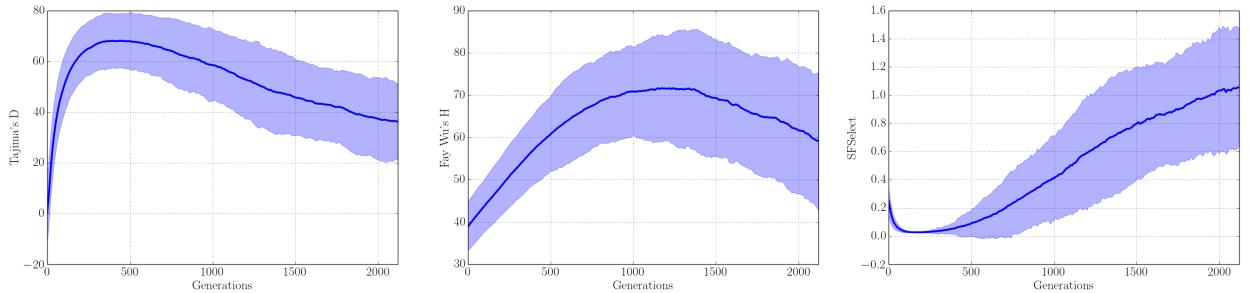
**Figure 3:** Predicted distributions of allele frequencies of sites with initial frequency  $\nu_0 = 0.005$  (top) and  $\nu_0 = 0.1$  (bottom), after  $t = \{1, 10, 100\}$  generations, under neutral evolution, by Brownian motion and Markov chain. Observed distribution of 143900 sites with  $\nu_0 = 0.005$  and 47500 with variants  $\nu_0 = 0.1$  computed from neutrally evolving simulations is depicted in green lines.



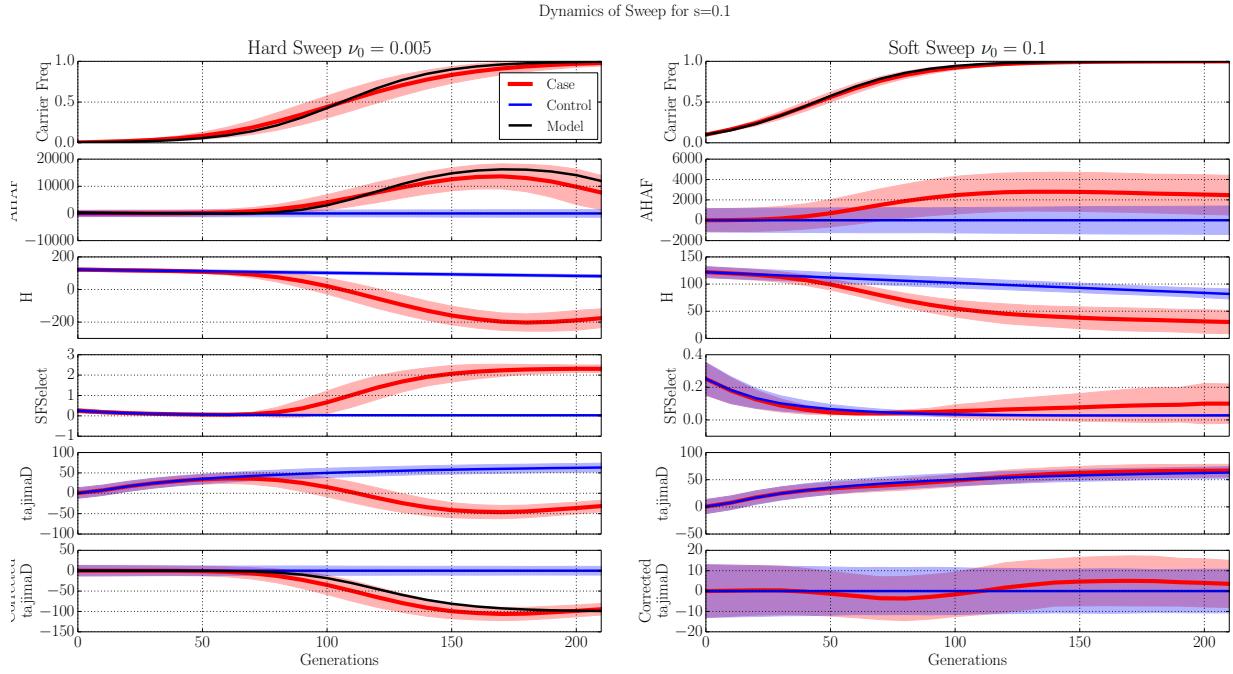
**Figure 4:** Predicted distributions of allele frequencies of sites with initial frequency  $\nu_0 = 0.005$  and selection strength  $s = 0.1$ , after  $t = \{1, 10, 100\}$  generations, by Markov chain. Observed distribution of the adaptive allele (with  $\nu_0 = 0.005$ ) in 10000 simulation is shown in green.



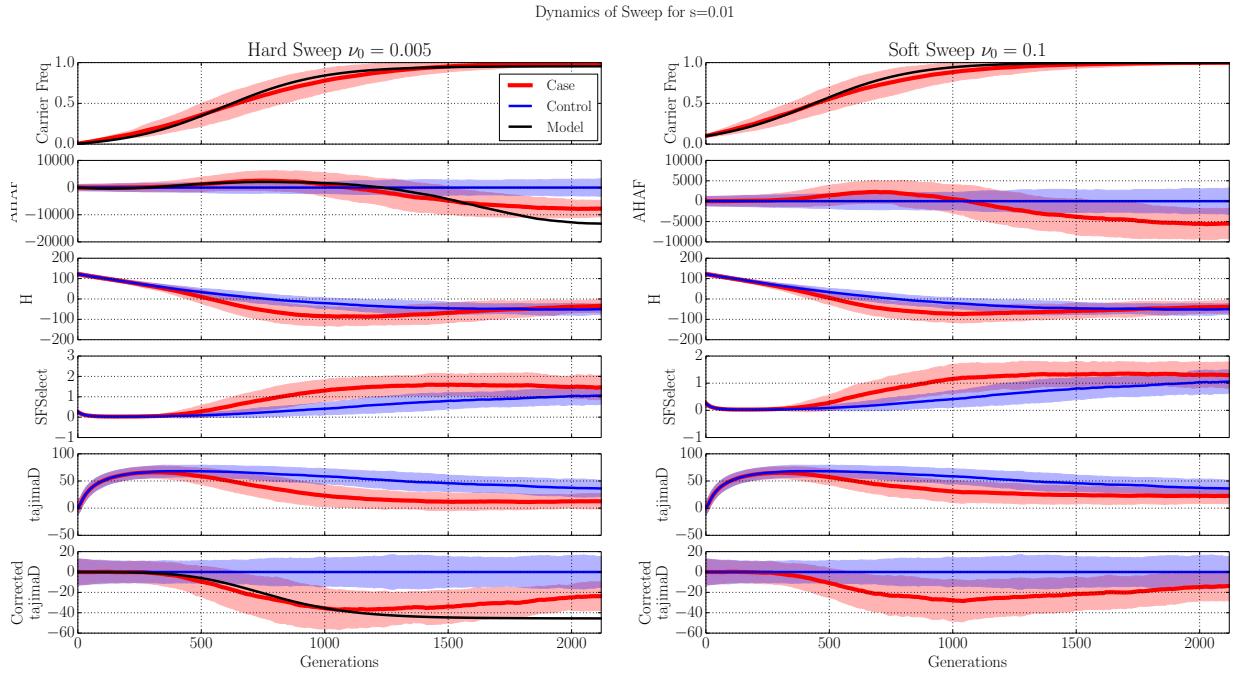
**Figure 5:** Interactions of two terms in  $D$ .



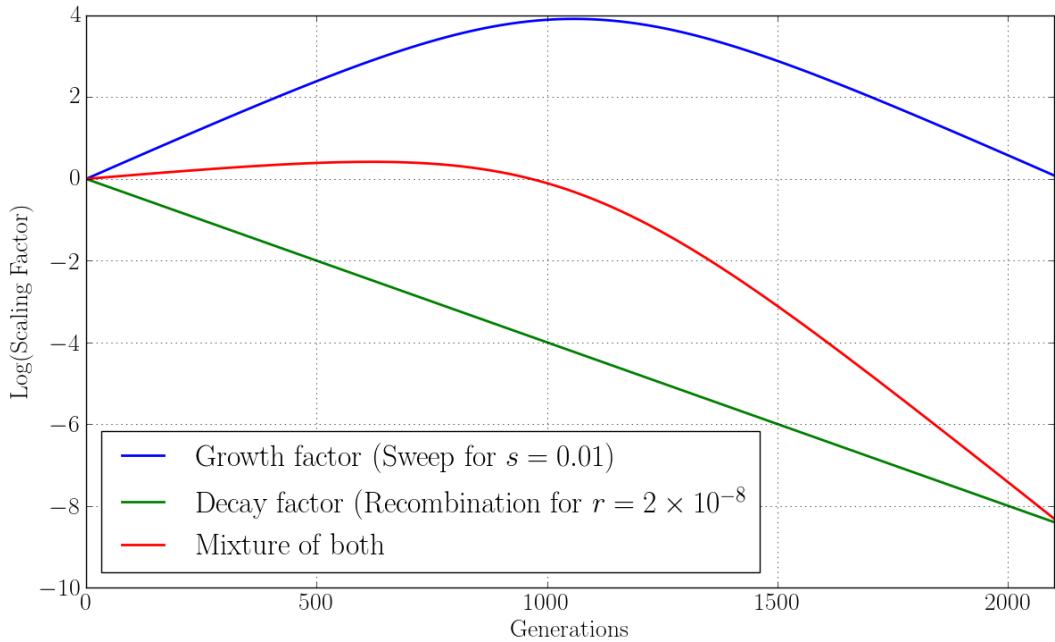
**Figure 6:** Effect of bottle neck in a typical experimental evolution experiment where a restricted number of founder lines (here  $F = 200$ ) is selected out of a larger population size ( $N_e = 10^{-6}$ ). Tajima's D (left), Fay Wu's H (middle) and SFSelect is computed for 1000 neutral simulations and mean and 95% confidence interval is plotted.



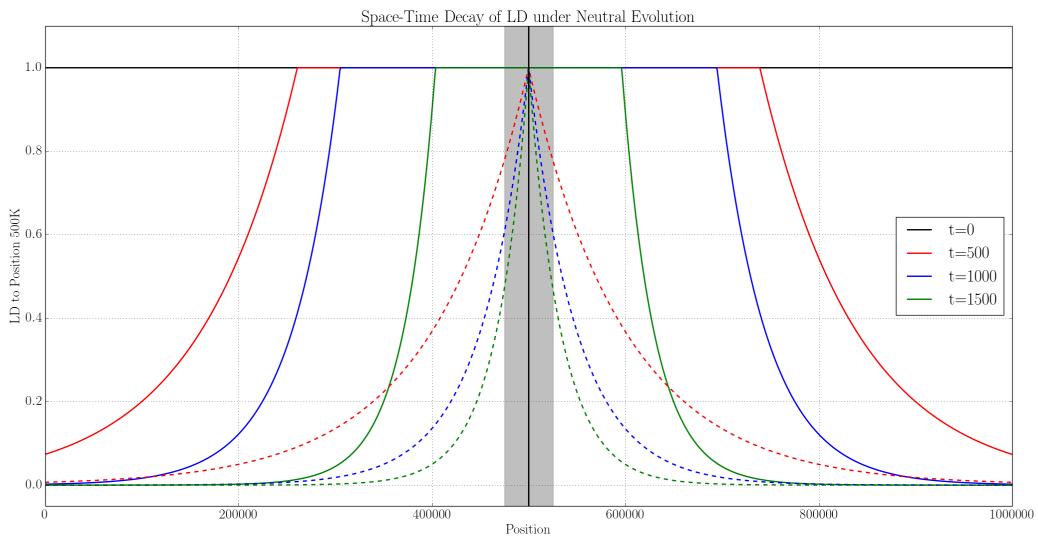
**Figure 7:** Mean and 95% CI of 1000 simulations for strong selection.



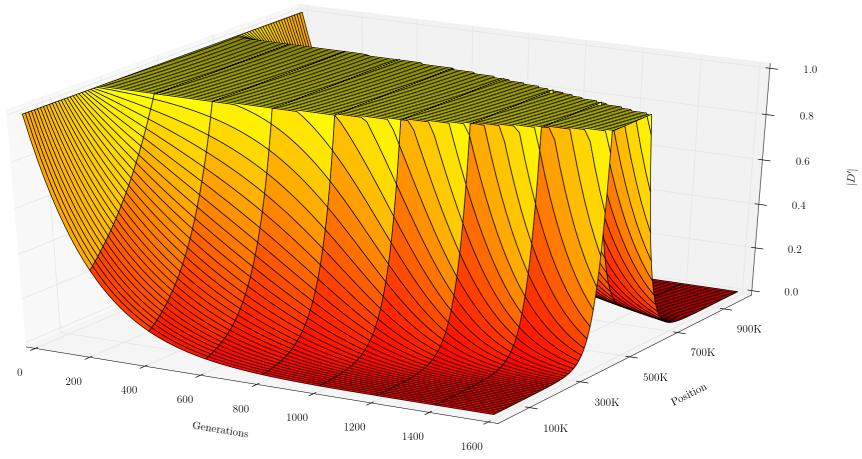
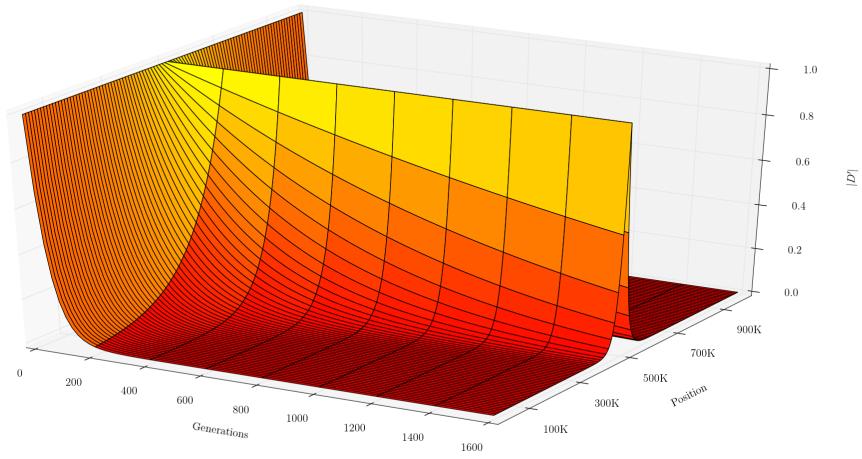
**Figure 8:** Mean and 95% CI of 1000 simulations for weak selection.



**Figure 9:** Interaction between productive factors of LD under natural selection for weak selection ( $s=0.01$ ) and a distance of 100Kb between sites. In this setting, after about 1000 generations LD start to decay (red curve).

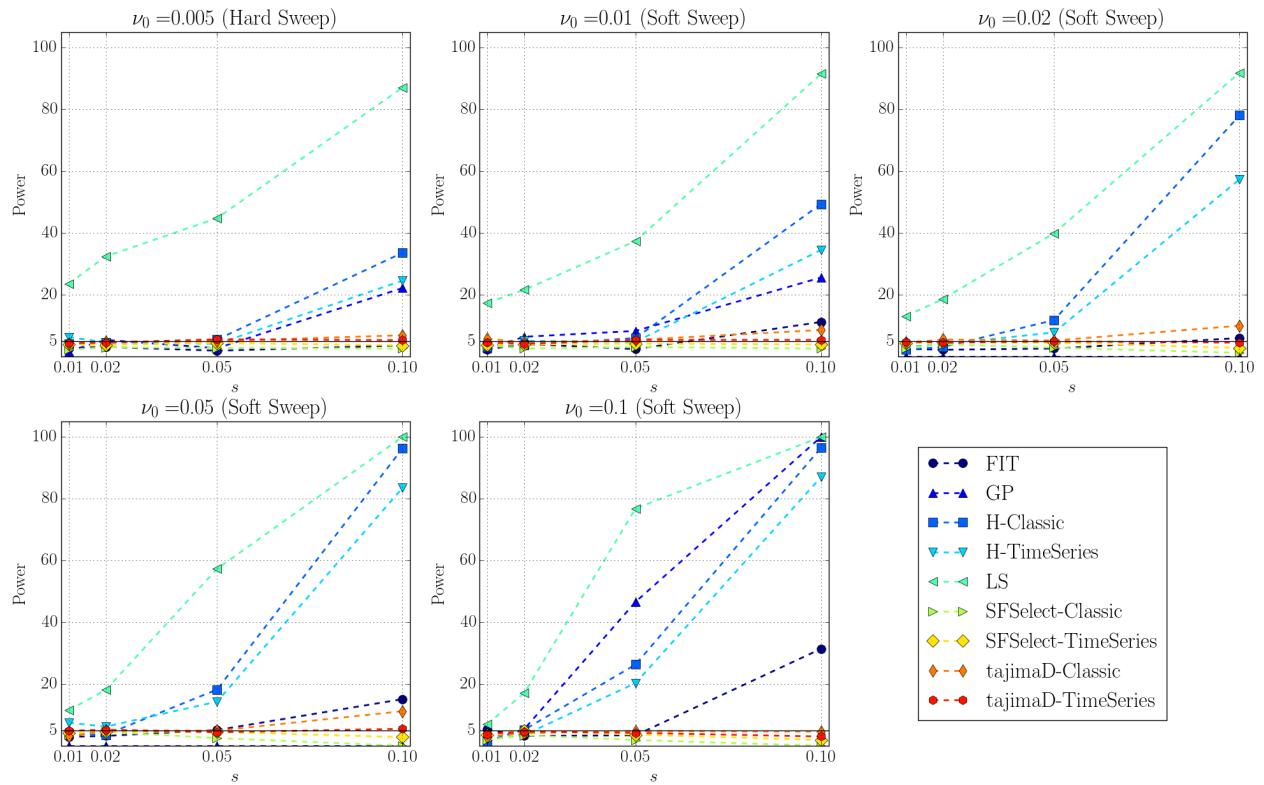


**Figure 10:** Decay of LD ( $|D'|$  measure) of the minimum AF site at position 500K with the rest of genome when  $s = 0.01$  and  $r = 2 \times 10^{-8}$ . A window of 50Kb is shaded at the center of genome to illustrate high values of linkage in both selection and drift.

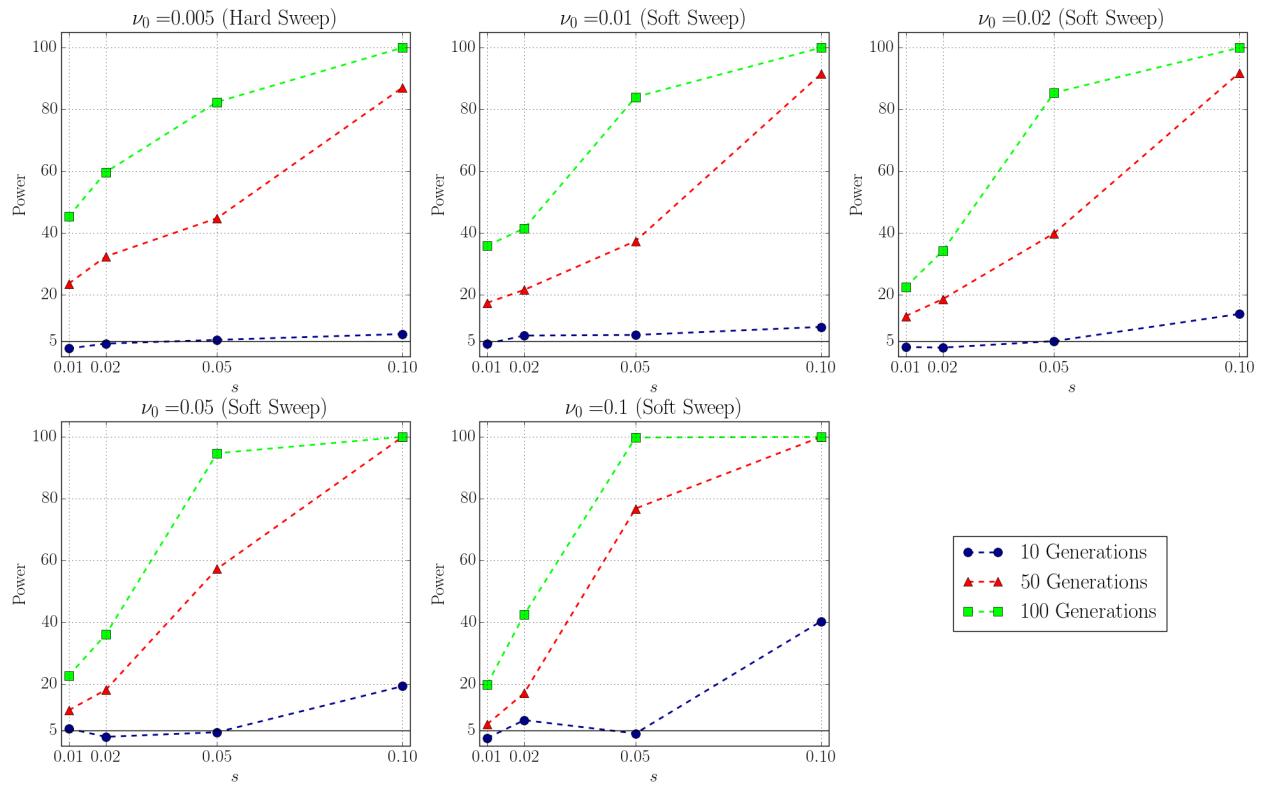


**Figure 11:** ld

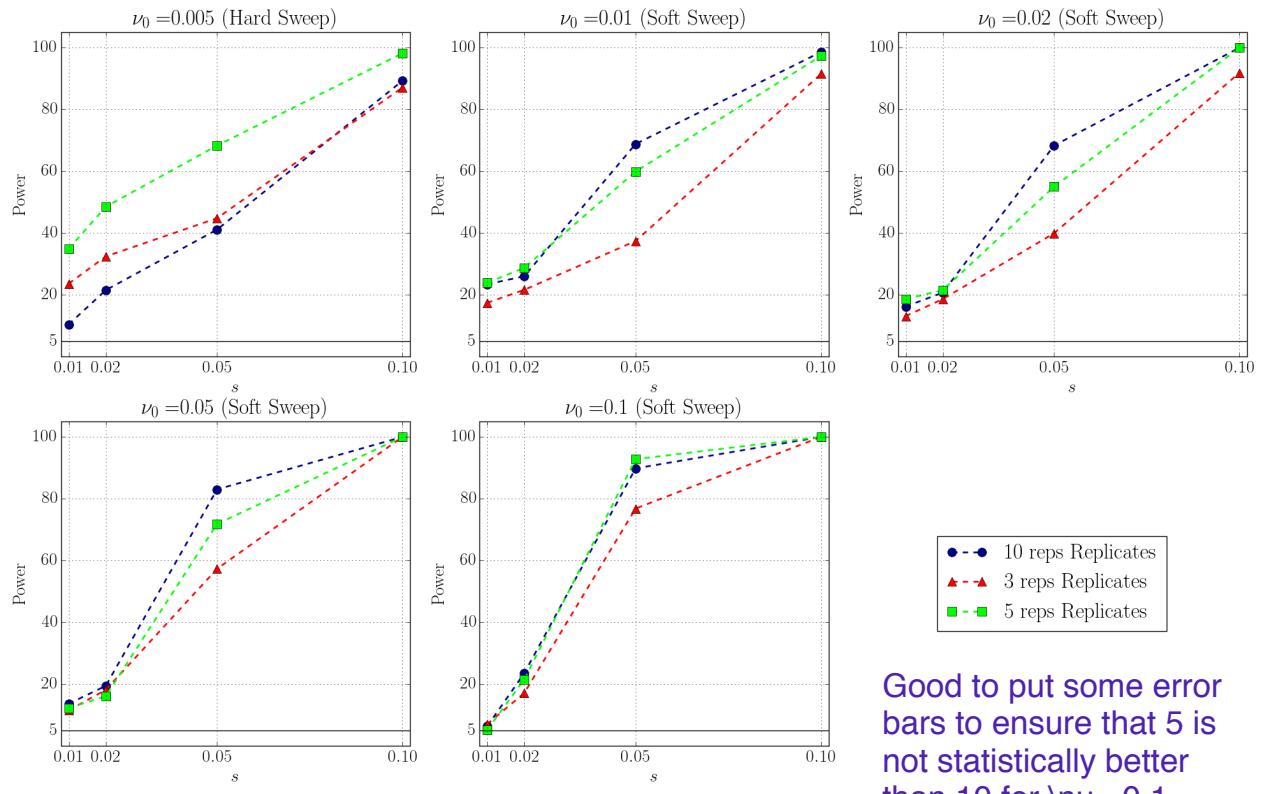
**Figure 12:** Decay of LD ( $|D'|$  measure) of the minimum AF site at position 500K with the rest of genome in genetic drift with  $r = 2 \times 10^{-8}$  (top) and hard sweep with  $s = 0.01$  (bottom).



**Figure 13:** Predictive performance of different method is evaluated on 200 simulations for different values of selection strength  $s$  and initial carrier frequency  $\nu_0$ .

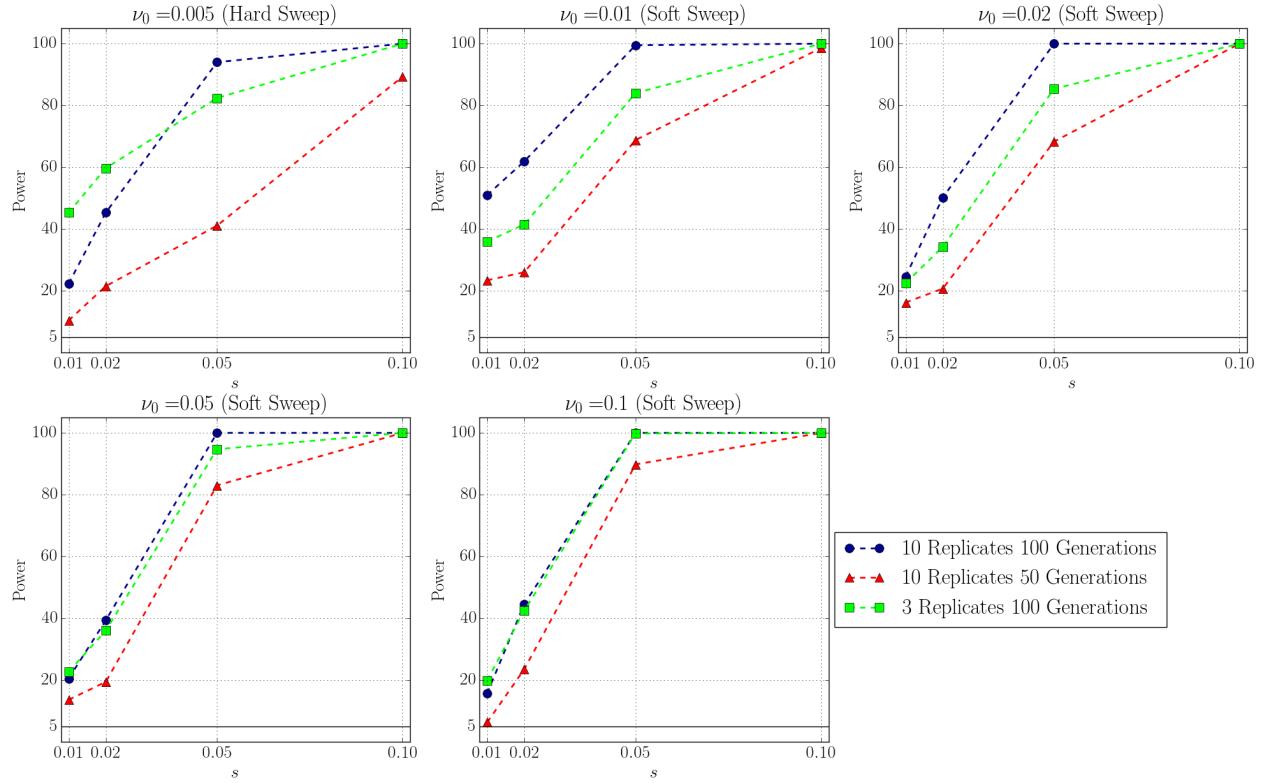


**Figure 14:** Predictive performance of the proposed method is evaluated on 200 simulations for different values of sampling-time-span  $\tau$  and initial carrier frequency  $\nu_0$ .

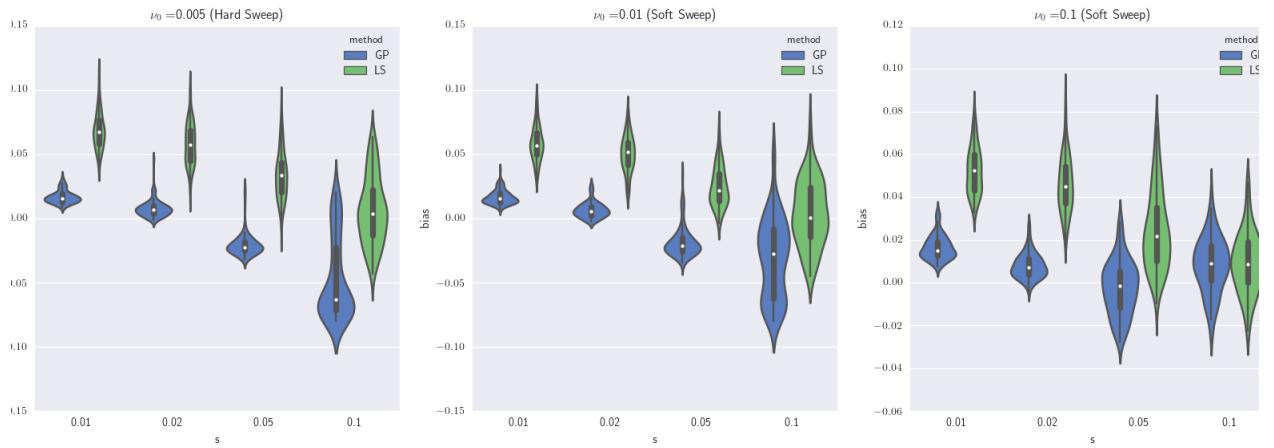


Good to put some error bars to ensure that 5 is not statistically better than 10 for  $\nu_0 = 0.1$

**Figure 15:** Predictive performance of the proposed method is evaluated on 200 simulations for different number of replicates  $R$  and initial carrier frequency  $\nu_0$ .

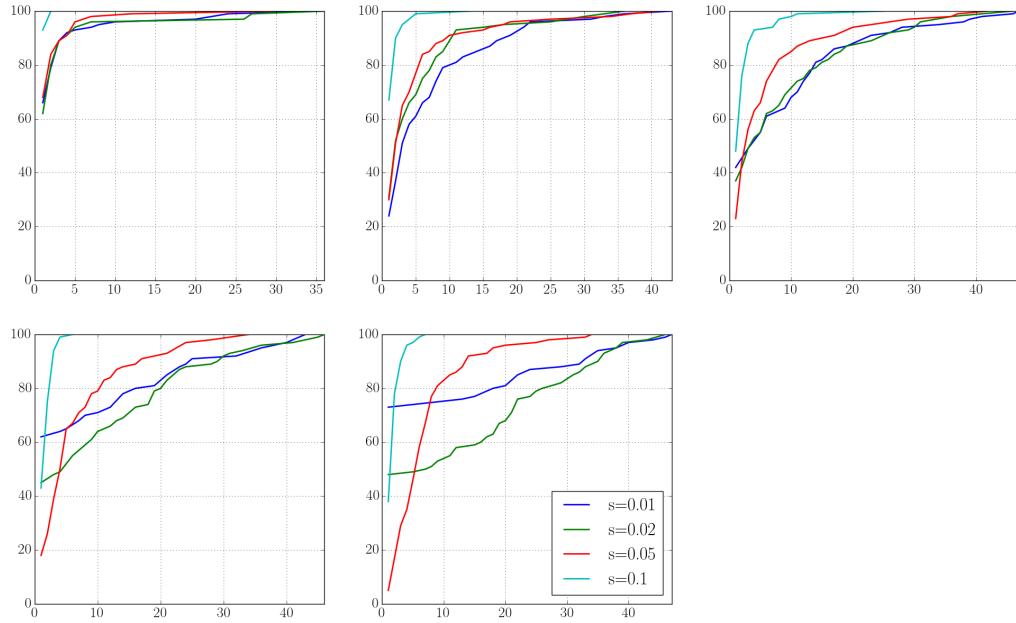


**Figure 16:** Predictive performance of the proposed method is evaluated on 200 simulations for different number of replicates  $R$ , sampling-time-span  $\tau$  and initial carrier frequency  $\nu_0$ .

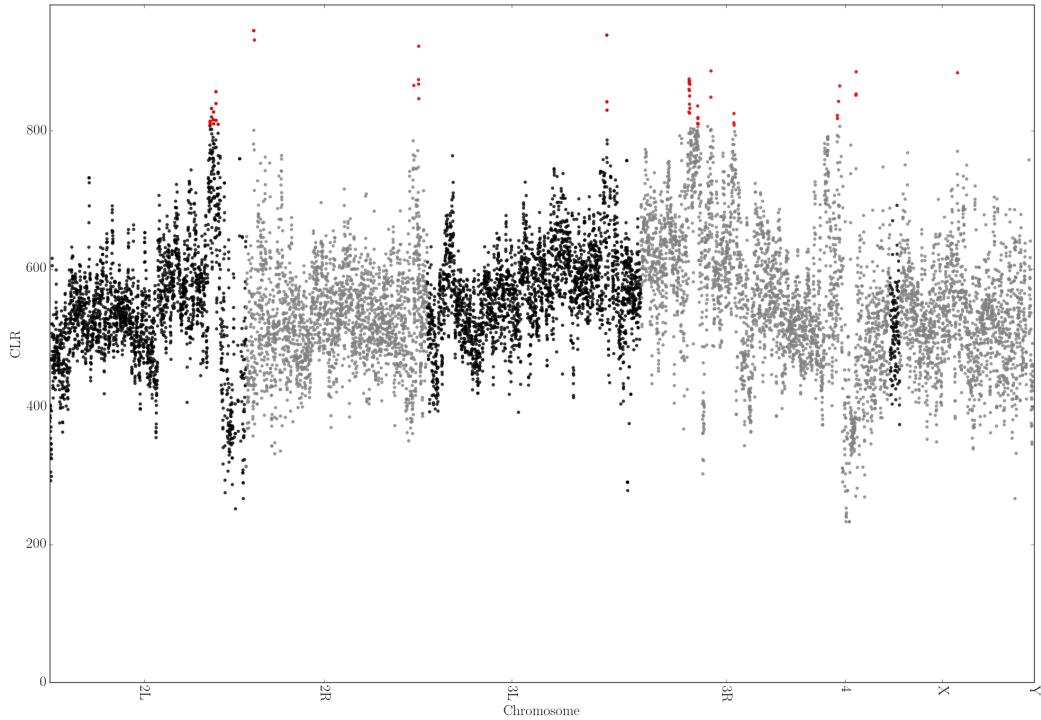


**Figure 17:** Comparison between the distribution of bias of the proposed method with Gaussian Process (GP), over 100 simulations for different values of selection strength  $s$  and soft-sweep with  $\nu_0 = 0.1$ .

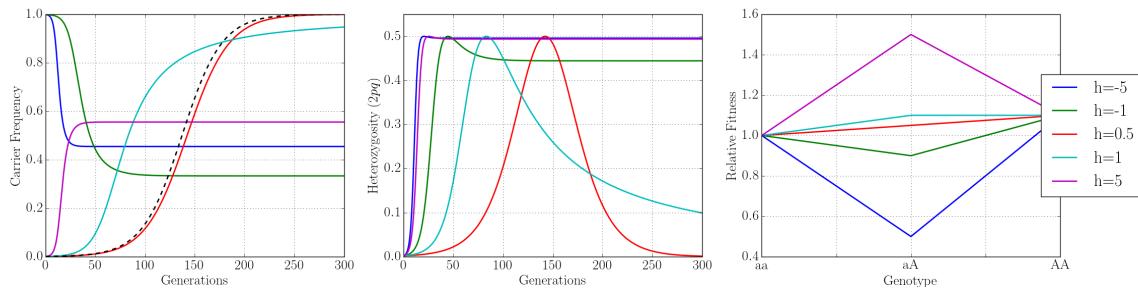
$\nu_0 = 0.1$



**Figure 18:** CDF of the rank of the adaptive allele in 100 simulations.



**Figure 19:** Manhattan plot of scan for selection using CLR on windows of size 50Kb and sliding window of 10Kb.



**Figure 20:** Dominance for  $s = 0.1$ .

## References

- [1] ACHAZ, G. Frequency spectrum neutrality tests: one for all and all for one. *Genetics* 183, 1 (2009), 249–258.
- [2] AKEY, J. M. Constructing genomic maps of positive selection in humans: Where do we go from here? *Genome research* 19, 5 (2009), 711–722.
- [3] ARIEY, F., WITKOWSKI, B., AMARATUNGA, C., BEGHAIN, J., LANGLOIS, A.-C., KHIM, N., KIM, S., DURU, V., BOUCHIER, C., MA, L., AND OTHERS. A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature* 505, 7481 (2014), 50–55.
- [4] BARRETT, R. D. H., ROGERS, S. M., AND SCHLUTER, D. Natural selection on a major armor gene in threespine stickleback. *Science* 322, 5899 (2008), 255–257.
- [5] BARRICK, J. E., AND LENSKI, R. E. Genome dynamics during experimental evolution. *Nat Rev Genet* 14, 12 (dec 2013), 827–839.
- [6] BARRICK, J. E., YU, D. S., YOON, S. H., JEONG, H., OH, T. K., SCHNEIDER, D., LENSKI, R. E., AND KIM, J. F. Genome evolution and adaptation in a long-term experiment with Escherichia coli. *Nature* 461, 7268 (2009), 1243–1247.
- [7] BERGLAND, A. O., BEHRMAN, E. L., O'BRIEN, K. R., SCHMIDT, P. S., AND PETROV, D. A. Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in *Drosophila*. *PLoS Genet* 10, 11 (2014), e1004775.
- [8] BOLBACK, J. P., AND HUELSENBECK, J. P. Clonal interference is alleviated by high mutation rates in large populations. *Molecular biology and evolution* 24, 6 (2007), 1397–1406.
- [9] BOLBACK, J. P., YORK, T. L., AND NIELSEN, R. Estimation of 2Nes from temporal allele frequency data. *Genetics* 179, 1 (2008), 497–502.
- [10] BOYKO, A. R., WILLIAMSON, S. H., INDAP, A. R., DEGENHARDT, J. D., HERNANDEZ, R. D., LOHMUELLER, K. E., ADAMS, M. D., SCHMIDT, S., SNINSKY, J. J., SUNYAEV, S. R., AND OTHERS. Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet* 4, 5 (2008), e1000083.
- [11] BURKE, M. K., DUNHAM, J. P., SHAHRESTANI, P., THORNTON, K. R., ROSE, M. R., AND LONG, A. D. Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* 467, 7315 (2010), 587–590.
- [12] DABORN, P., BOUNDY, S., YEN, J., PITTEENDRIGH, B., AND OTHERS. DDT resistance in *Drosophila* correlates with Cyp6g1 over-expression and confers cross-resistance to the neonicotinoid imidacloprid. *Molecular Genetics and Genomics* 266, 4 (2001), 556–563.
- [13] DANIELS, R., CHANG, H.-H., SÉNE, P. D., PARK, D. C., NEAFSEY, D. E., SCHAFFNER, S. F., HAMILTON, E. J., LUKENS, A. K., VAN TYNE, D., MBOUP, S., AND OTHERS. Genetic surveillance detects both clonal and epidemic transmission of malaria following enhanced intervention in Senegal. *PLoS One* 8, 4 (2013), e60780.
- [14] DENEF, V. J., AND BANFIELD, J. F. In situ evolutionary rate measurements show ecological success of recently emerged bacterial hybrids. *Science* 336, 6080 (2012), 462–466.

- [15] DESAI, M. M., AND FISHER, D. S. Beneficial mutation-selection balance and the effect of linkage on positive selection. *Genetics* 176, 3 (2007), 1759–1798.
- [16] DESAI, M. M., AND PLOTKIN, J. B. The polymorphism frequency spectrum of finitely many sites under selection. *Genetics* 180, 4 (2008), 2175–2191.
- [17] EVANS, S. N., SHVETS, Y., AND SLATKIN, M. Non-equilibrium theory of the allele frequency spectrum. *Theoretical population biology* 71, 1 (2007), 109–119.
- [18] EWENS, W. J. *Mathematical Population Genetics 1: Theoretical Introduction*, vol. 27. Springer Science & Business Media, 2012.
- [19] FAWCETT, T. An introduction to ROC analysis. *Pattern recognition letters* 27, 8 (2006), 861–874.
- [20] FAY, J. C., AND WU, C.-I. Hitchhiking under positive Darwinian selection. *Genetics* 155, 3 (2000), 1405–1413.
- [21] FEDER, A. F., KRYAZHIMSKIY, S., AND PLOTKIN, J. B. Identifying signatures of selection in genetic time series. *Genetics* 196, 2 (2014), 509–522.
- [22] FEDER, A. F., RHEE, S.-Y., HOLMES, S. P., SHAFER, R. W., PETROV, D. A., AND PENNINGS, P. S. More effective drugs lead to harder selective sweeps in the evolution of drug resistance in HIV-1. *eLife* 5 (jan 2016).
- [23] GOTTESMAN, M. M. Mechanisms of cancer drug resistance. *Annual review of medicine* 53, 1 (2002), 615–627.
- [24] HEGRENESS, M., SHORESH, N., HARTL, D., AND KISHONY, R. An equivalence principle for the incorporation of favorable mutations in asexual populations. *Science* 311, 5767 (2006), 1615–1617.
- [25] HOLSINGER, K. E., AND WEIR, B. S. Genetics in geographically structured populations: defining, estimating and interpreting FST. *Nature Reviews Genetics* 10, 9 (2009), 639–650.
- [26] ILLINGWORTH, C. J. R., AND MUSTONEN, V. Distinguishing driver and passenger mutations in an evolutionary history categorized by interference. *Genetics* 189, 3 (2011), 989–1000.
- [27] ILLINGWORTH, C. J. R., PARTS, L., SCHIFFELS, S., LITI, G., AND MUSTONEN, V. Quantifying selection acting on a complex trait using allele frequency time series data. *Molecular biology and evolution* 29, 4 (2012), 1187–1197.
- [28] IZUTSU, M., TOYODA, A., FUJIYAMA, A., AGATA, K., AND FUSE, N. Dynamics of Dark-Fly Genome Under Environmental Selections. *G3: Genes—Genomes—Genetics* (2015), g3—115.
- [29] JHA, A. R., MILES, C. M., LIPPERT, N. R., BROWN, C. D., WHITE, K. P., AND KREITMAN, M. Whole-genome resequencing of experimental populations reveals polygenic basis of egg-size variation in *Drosophila melanogaster*. *Molecular biology and evolution* 32, 10 (2015), 2616–2632.
- [30] KAPLAN, N. L., HUDSON, R. R., AND LANGLEY, C. H. The “hitchhiking effect” revisited. *Genetics* 123, 4 (1989), 887–899.

- [31] Kawecki, T. J., Lenski, R. E., Ebert, D., Hollis, B., Olivieri, I., and Whitlock, M. C. Experimental evolution. *Trends in ecology & evolution* 27, 10 (2012), 547–560.
- [32] Lang, G. I., Botstein, D., and Desai, M. M. Genetic variation and the fate of beneficial mutations in asexual populations. *Genetics* 188, 3 (2011), 647–661.
- [33] Lang, G. I., Rice, D. P., Hickman, M. J., Södergren, E., Weinstock, G. M., Botstein, D., and Desai, M. M. Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* 500, 7464 (2013), 571–574.
- [34] Malaspinas, A.-S., Malaspinas, O., Evans, S. N., and Slatkin, M. Estimating allele age and selection coefficient from time-serial data. *Genetics* 192, 2 (2012), 599–607.
- [35] Maldarelli, F., Kearney, M., Palmer, S., Stephens, R., Mican, J., Polis, M. A., Davey, R. T., Kovacs, J., Shao, W., Rock-Kress, D., and Others. HIV populations are large and accumulate high genetic diversity in a nonlinear fashion. *Journal of virology* 87, 18 (2013), 10313–10323.
- [36] Martins, N. E., Faria, V. G., Nolte, V., Schlotterer, C., Teixeira, L., Sucena, É., and Magalhães, S. Host adaptation to viruses relies on few genes with different cross-resistance properties. *Proceedings of the National Academy of Sciences* 111, 16 (2014), 5938–5943.
- [37] Mathieson, I., and McVean, G. Estimating selection coefficients in spatially structured populations from time series data of allele frequencies. *Genetics* 193, 3 (2013), 973–984.
- [38] McClish, D. K. Analyzing a portion of the ROC curve. *Medical Decision Making* 9, 3 (1989), 190–195.
- [39] Messer, P. W., and Petrov, D. A. Population genomics of rapid adaptation by soft selective sweeps. *Trends in ecology & evolution* 28, 11 (2013), 659–669.
- [40] Nair, S., Nash, D., Sudimack, D., Jaidee, A., Barends, M., Uhlemann, A.-C., Krishna, S., Nosten, F., and Anderson, T. J. C. Recurrent gene amplification and soft selective sweeps during evolution of multidrug resistance in malaria parasites. *Molecular Biology and Evolution* 24, 2 (2007), 562–573.
- [41] Nielsen, R., and Signorovitch, J. Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium. *Theoretical population biology* 63, 3 (2003), 245–255.
- [42] Nielsen, R., Williamson, S., Kim, Y., Hubisz, M. J., Clark, A. G., and Bustamante, C. Genomic scans for selective sweeps using SNP data. *Genome research* 15, 11 (2005), 1566–1575.
- [43] Orozco-Terwengel, P., Kapun, M., Nolte, V., Kofler, R., Flatt, T., and Schlotterer, C. Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. *Molecular ecology* 21, 20 (2012), 4931–4941.
- [44] Oz, T., Guvenek, A., Yildiz, S., Karaboga, E., Tamer, Y. T., Mumcuyan, N., Ozan, V. B., Senturk, G. H., Cokol, M., Yeh, P., and Others. Strength of selection pressure

- is an important parameter contributing to the complexity of antibiotic resistance evolution. *Molecular biology and evolution* (2014), msu191.
- [45] POLLAK, E. A new method for estimating the effective population size from allele frequency changes. *Genetics* 104, 3 (1983), 531–548.
  - [46] PTAK, S. E., AND PRZEWORSKI, M. Evidence for population growth in humans is confounded by fine-scale population structure. *Trends in Genetics* 18, 11 (2002), 559–563.
  - [47] RAMOS-ONSINS, S. E., AND ROZAS, J. Statistical properties of new neutrality tests against population growth. *Molecular biology and evolution* 19, 12 (2002), 2092–2100.
  - [48] REID, B. J., KOSTADINOV, R., AND MALEY, C. C. New strategies in Barrett’s esophagus: integrating clonal evolutionary theory with clinical management. *Clinical Cancer Research* 17, 11 (2011), 3512–3519.
  - [49] REMOLINA, S. C., CHANG, P. L., LEIPS, J., NUZHDIN, S. V., AND HUGHES, K. A. Genomic basis of aging and life-history evolution in *Drosophila melanogaster*. *Evolution* 66, 11 (2012), 3390–3403.
  - [50] RONEN, R., TESLER, G., AKBARI, A., ZAKOV, S., ROSENBERG, N. A., AND BAFNA, V. Predicting Carriers of Ongoing Selective Sweeps Without Knowledge of the Favored Allele. *PLoS Genet* 11, 9 (2015), e1005527.
  - [51] RONEN, R., UDPA, N., HALPERIN, E., AND BAFNA, V. Learning natural selection from the site frequency spectrum. *Genetics* 195, 1 (2013), 181–193.
  - [52] SABETI, P. C., SCHAFFNER, S. F., FRY, B., LOHMEULLER, J., VARILLY, P., SHAMOVSKY, O., PALMA, A., MIKKELSEN, T. S., ALTSCHULER, D., AND LANDER, E. S. Positive natural selection in the human lineage. *science* 312, 5780 (2006), 1614–1620.
  - [53] SAWYER, S. A., AND HARTL, D. L. Population genetics of polymorphism and divergence. *Genetics* 132, 4 (1992), 1161–1176.
  - [54] SCHLÖTTERER, C., KOFLER, R., VERSACE, E., TOBLER, R., AND FRANSSEN, S. U. Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation. *Heredity* 114, 5 (2015), 431–440.
  - [55] SMITH, J. M., AND HAIGH, J. The hitch-hiking effect of a favourable gene. *Genetical research* 23, 01 (1974), 23–35.
  - [56] SPELLBERG, B., GUIDOS, R., GILBERT, D., BRADLEY, J., BOUCHER, H. W., SCHELD, W. M., BARTLETT, J. G., EDWARDS, J., OF AMERICA, I. D. S., AND OTHERS. The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. *Clinical Infectious Diseases* 46, 2 (2008), 155–164.
  - [57] STEINRÜCKEN, M., BHASKAR, A., AND SONG, Y. S. A novel spectral method for inferring general diploid selection from time series genetic data. *The annals of applied statistics* 8, 4 (2014), 2203.
  - [58] STEPHAN, W., SONG, Y. S., AND LANGLEY, C. H. The Hitchhiking Effect on Linkage Disequilibrium Between Linked Neutral Loci. *Genetics* 172, 4 (apr 2006), 2647–2663.

- [59] TAJIMA, F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 3 (1989), 585–595.
- [60] TERHORST, J., SCHLÖTTERER, C., AND SONG, Y. S. Multi-locus Analysis of Genomic Time Series Data from Experimental Evolution. *PLoS Genet* 11, 4 (2015), e1005069.
- [61] TOBLER, R., FRANSSEN, S. U., KOFLER, R., OROZCO-TERWENGEL, P., NOLTE, V., HERMISSON, J., AND SCHLÖTTERER, C. Massive habitat-specific genomic response in *D. melanogaster* populations during experimental evolution in hot and cold environments. *Molecular biology and evolution* 31, 2 (2014), 364–375.
- [62] TURNER, T. L., STEWART, A. D., FIELDS, A. T., RICE, W. R., AND TARONE, A. M. Population-based resequencing of experimentally evolved populations reveals the genetic basis of body size variation in *Drosophila melanogaster*. *PLoS Genet* 7, 3 (2011), e1001336.
- [63] VITTI, J. J., GROSSMAN, S. R., AND SABETI, P. C. Detecting natural selection in genomic data. *Annual review of genetics* 47 (2013), 97–120.
- [64] WANG, J. A pseudo-likelihood method for estimating effective population size from temporally spaced samples. *Genetical research* 78, 03 (2001), 243–257.
- [65] WAPLES, R. S. A generalized approach for estimating effective population size from temporal changes in allele frequency. *Genetics* 121, 2 (1989), 379–391.
- [66] WILLIAMSON, E. G., AND SLATKIN, M. Using maximum likelihood to estimate population size from temporal changes in allele frequencies. *Genetics* 152, 2 (1999), 755–761.
- [67] WILLIAMSON, S. H., HUBISZ, M. J., CLARK, A. G., PAYSEUR, B. A., BUSTAMANTE, C. D., AND NIELSEN, R. Localizing recent adaptive evolution in the human genome. *PLoS Genet* 3, 6 (2007), e90.
- [68] WINTERS, M. A., LLOYD JR, R. M., SHAFER, R. W., KOZAL, M. J., MILLER, M. D., AND HOLODNIY, M. Development of elvitegravir resistance and linkage of integrase inhibitor mutations with protease and reverse transcriptase resistance mutations. *PloS one* 7, 7 (2012), e40514.
- [69] ZAHREDDINE, H., AND BORDEN, K. L. Mechanisms and insights into drug resistance in cancer. *Front Pharmacol* 4, 28.10 (2013), 3389.
- [70] ZHOU, D., UDPA, N., GERSTEN, M., VISK, D. W., BASHIR, A., XUE, J., FRAZER, K. A., POSAKONY, J. W., SUBRAMANIAM, S., BAFNA, V., AND OTHERS. Experimental selection of hypoxia-tolerant *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences* 108, 6 (2011), 2349–2354.