

## Supplementary

### S1 Introduction of the genetic model with more details

For a given bi-allelic locus of a population, we denoted the frequencies of the wild-type and mutated alleles at time  $t$  as  $p_t^w$  and  $p_t^m$ , respectively. Their corresponding frequencies at time  $t-1$  are  $p_{t-1}^w$  and  $p_{t-1}^m$ . In this study, we defined fitness of wild-type and mutated alleles  $\lambda_w$  and  $\lambda_m$ , so we have  $p_t^m / p_t^w = (\lambda_m / \lambda_w)(p_{t-1}^m / p_{t-1}^w)$ . The ratio of the allele fitness could be represented as  $\lambda_m / \lambda_w = e^{-s}$ . When population dynamics  $g$  is considered, we have  $\lambda_m = e^{-(g+s)}$  and  $\lambda_w = e^{-g}$ , where  $s$  indicates differences in fitness between mutated and wild-type alleles.

For any time  $t$ , the number of chromosomes  $N_t^m$  carrying the mutated allele could be written as  $N_t^m = N_{t-1}^m \exp[-(g+s+\omega^m)]$ , where  $N_{t-1}^m$  is the number of chromosomes at time  $t$ . Effect of genetic drift is indicated as a random variable  $\omega^m$  with  $E(\omega^m) = 0$ . When the population dynamics  $g$  and difference in fitness  $\phi$  remain the same during time  $0$  to  $t$ , we have the number of chromosomes at time  $t$

$N_t^m = N_0^m \exp\{-(g+s)t + \sum_{i=1}^t \omega_i^m\}$ . Following similar algebraic principles, we have

the number of chromosomes carrying the wild-type allele at time  $t$

$$N_t^w = N_0^w \exp\{-(gt + \sum_{i=1}^t \omega_i^w)\}.$$

In the scenario with two populations, we assumed that populations  $A$  and  $B$  have the same ancestral population  $O$  at time  $0$ . Then we have

$$R = \frac{N_A^m}{N_A^w} = \frac{N_O^m}{N_O^w} \exp\left\{-\left[s_A t + \sum_{i=1}^t (\omega_{A,i}^m - \omega_{A,i}^w)\right]\right\}$$

$$S = \frac{N_B^m}{N_B^w} = \frac{N_O^m}{N_O^w} \exp\left\{-\left[s_B t + \sum_{i=1}^t (\omega_{B,i}^m - \omega_{B,i}^w)\right]\right\}$$

Consequently, the strength of natural selection could be compared between populations *A* and *B* as follows:

$$\Phi = s_B - s_A = \frac{\log(R) - \log(S) + \sum_{i=1}^t [(\omega_{A,i}^w - \omega_{A,i}^m) - (\omega_{B,i}^w - \omega_{B,i}^m)]}{t}$$

## S2 Performance with simulated data

To further evaluate our method, we conducted the method on simulated data in multiple scenarios. Since theoretical analysis showed our method are insensitive to demographic changes, we used only a simple two-population demographic model with divergence time  $t=2000$  generations. We assumed the specified loci are always neutral in one population but may under directional selection in the other population where the selection coefficient was given as 0, 0.00125, 0.0025, 0.00375, or 0.005, respectively. For each scenario, 10000 data set were simulated with initial benefit allelic frequencies  $\text{Freq}=0.05, 0.1, \text{ or } 0.2$ , respectively. The simulated data sets included 100 individuals from each of the two populations. Data sets were marked from 1 to 15 for the 15 different scenarios (Table S1).

In our simulation of infinite-population-size model, the method performed well when the loci are under small or medium selection differences. Box plot showed little bias of the 10000 estimations in each scenario (Figure S1). The estimated

difference could be smaller than the given selection difference if the selection are strong enough. The bias happened when the frequency of benefit allele is extremely high in the present population and consequently absent from the samples. The bias can be due to a continuity correction for calculation of log-odds ratio. In these scenarios with strong selection, estimation can be improved by using a larger size of samples.

We also evaluated this method on finite-population-size model using the widely recognized simulator *msms* with constant population size  $N=10000$  individuals. Other demographic parameters are the same as that given in Table S1. Because *msms* cannot work with only allelic fitness, we assigned fitnesses of genotypes (AA, Aa and aa) as  $1+2s$ ,  $1+s$ , and  $1$ , respectively. Although our method was built on an infinite-population-size model, it still works well on the simulated data sets from the finite-population-size model. The results suggest our method is robust to demographic models with infinite or finite population sizes.

Table S1. Data sets with different initial frequencies and selection differences.

	<b>s=0</b>	<b>s=0.00125</b>	<b>s=0.00250</b>	<b>s=0.00375</b>	<b>s=0.005</b>
<b>Freq=0.05</b>	1	4	7	10	13
<b>Freq=0.10</b>	2	5	8	11	14
<b>Freq=0.20</b>	3	6	9	12	15

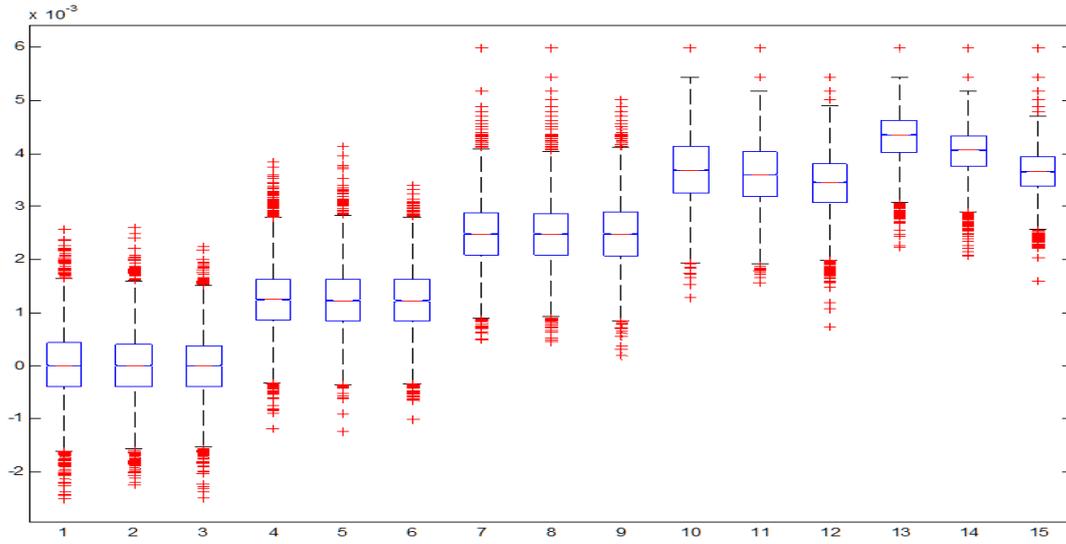


Figure S1. Boxplot showed estimations in 15 simulations with infinite-population-size model. Difference of selection coefficient is present on y-axis; code of different data set is shown on x-axis

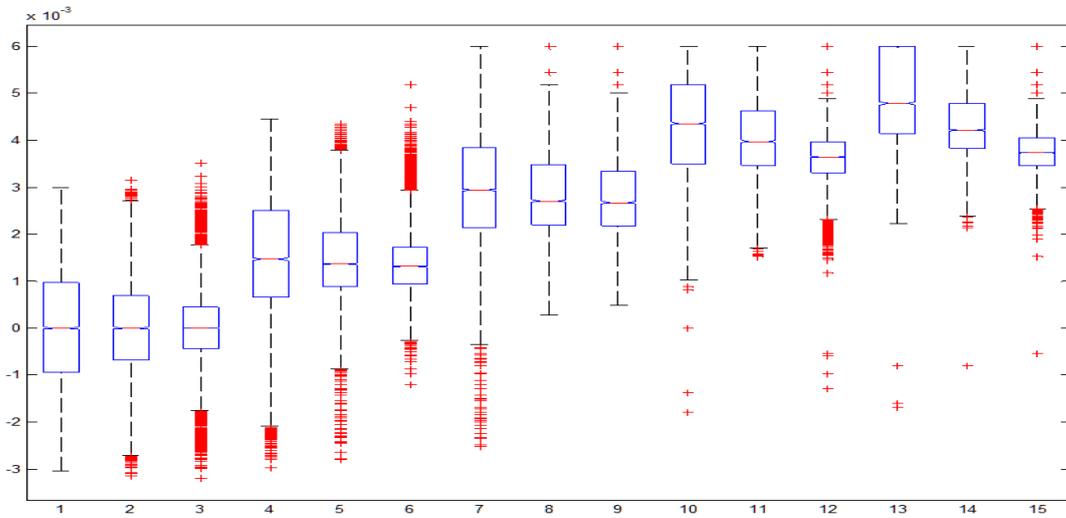


Figure S2. Estimations from our method. Difference of selection coefficient is present on y-axis; code of different data set is shown on x-axis.