SUPPLEMENTARY FIGURES AND TABLES

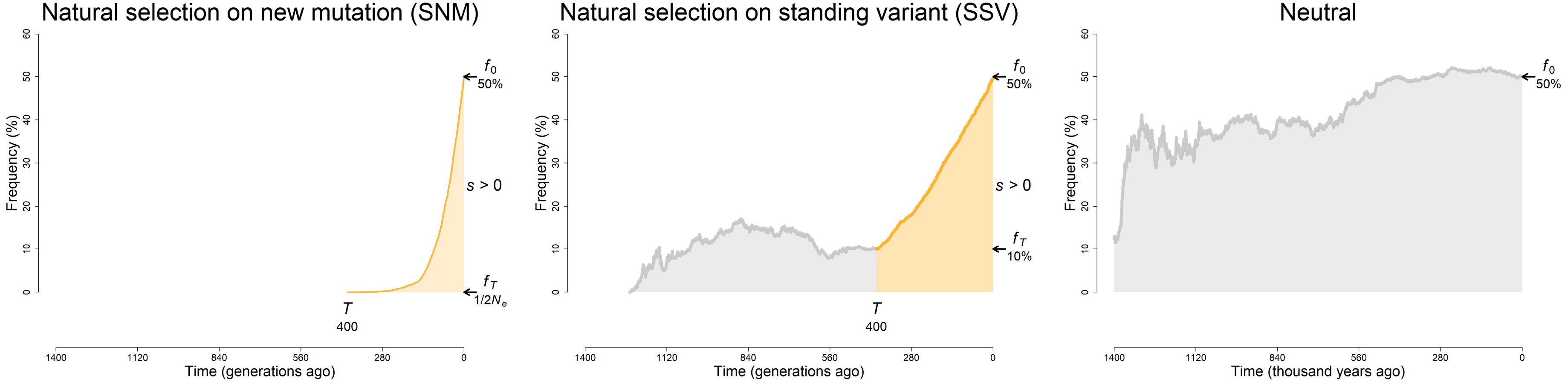

Figure S1. Examples of trajectory in three different evolutionary models. The selective and neutral processes are highlighted with different colours of orange and grey. In the SNM model, a beneficial mutation appeared at $T = 400$ with $f_T = 1/2N_\mathrm{e}$ and was increased to $f_0 = 50\%$ by $s > 0$. In contrast, a mutation that already existed as neutral allele became advantageous at $T = 400$ with $f_T = 10\%$. The neutral model is only dependent on f_0.

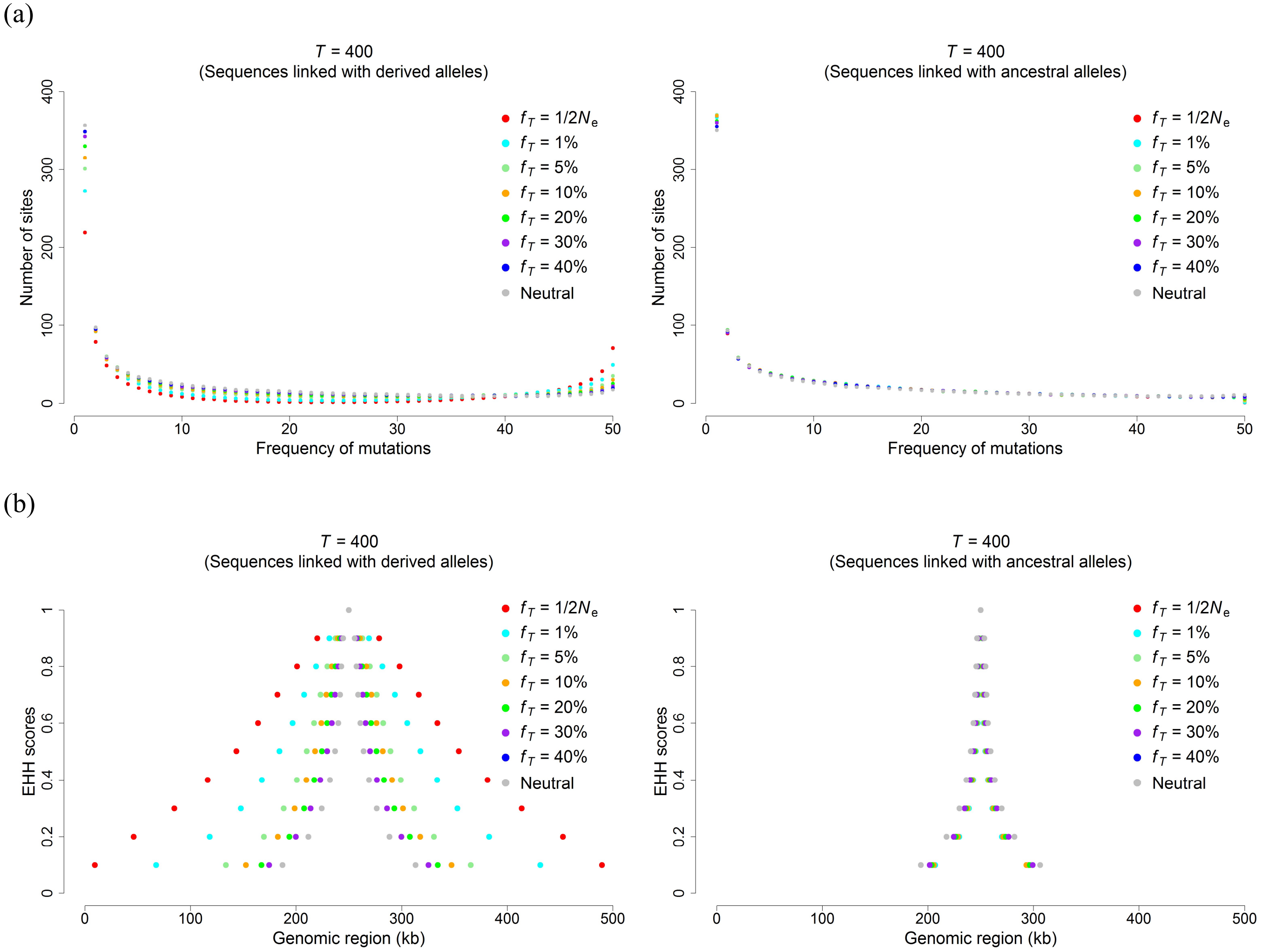

Figure S2. Examples of summary statistics under natural selection with $T = 400$ and different f_T or neutral model. Each plot represents a mean of mutation frequency in (a) full SFS or of physical position in (b) the decay of EHH from the sequences carrying derived or ancestral alleles across 1,000 simulation data. Different colours show the selection models with different f_T and the neutral model.

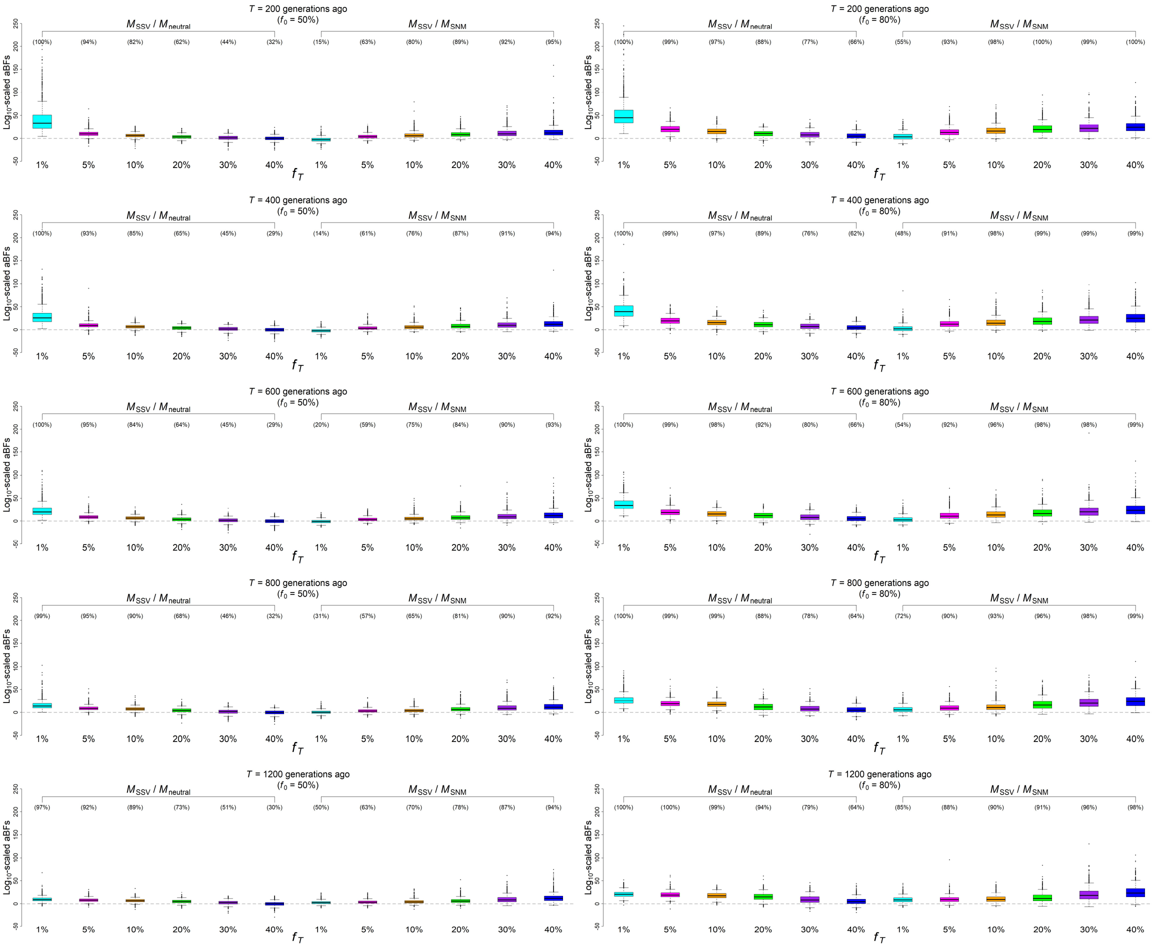

Figure S3. Accuracy of model selection in choosing the SSV models against the neutral or the SNM models under the varying population size demography. Boxplots show $\log_{10}$-scaled approximate Bayes factors (aBFs) that take ratios of approximate marginal likelihoods (aMLs) between a true model (*i.e.*, SSV) and an alternative model (*i.e.*, neutral or SNM). The percentage with the parentheses above each boxplot shows the probability that the method can correctly identify the true model with $\log_{10}(\text{aBF}) > 2$. The total 1,000 pseudo-observations under the SSV models are generated with $T = 200, 400, 600, 800, 1200$, $f_T = 1\%, 5\%, 10\%, 20\%, 30\%, 40\%$, and $f_0 = 50\%$ or 80%.

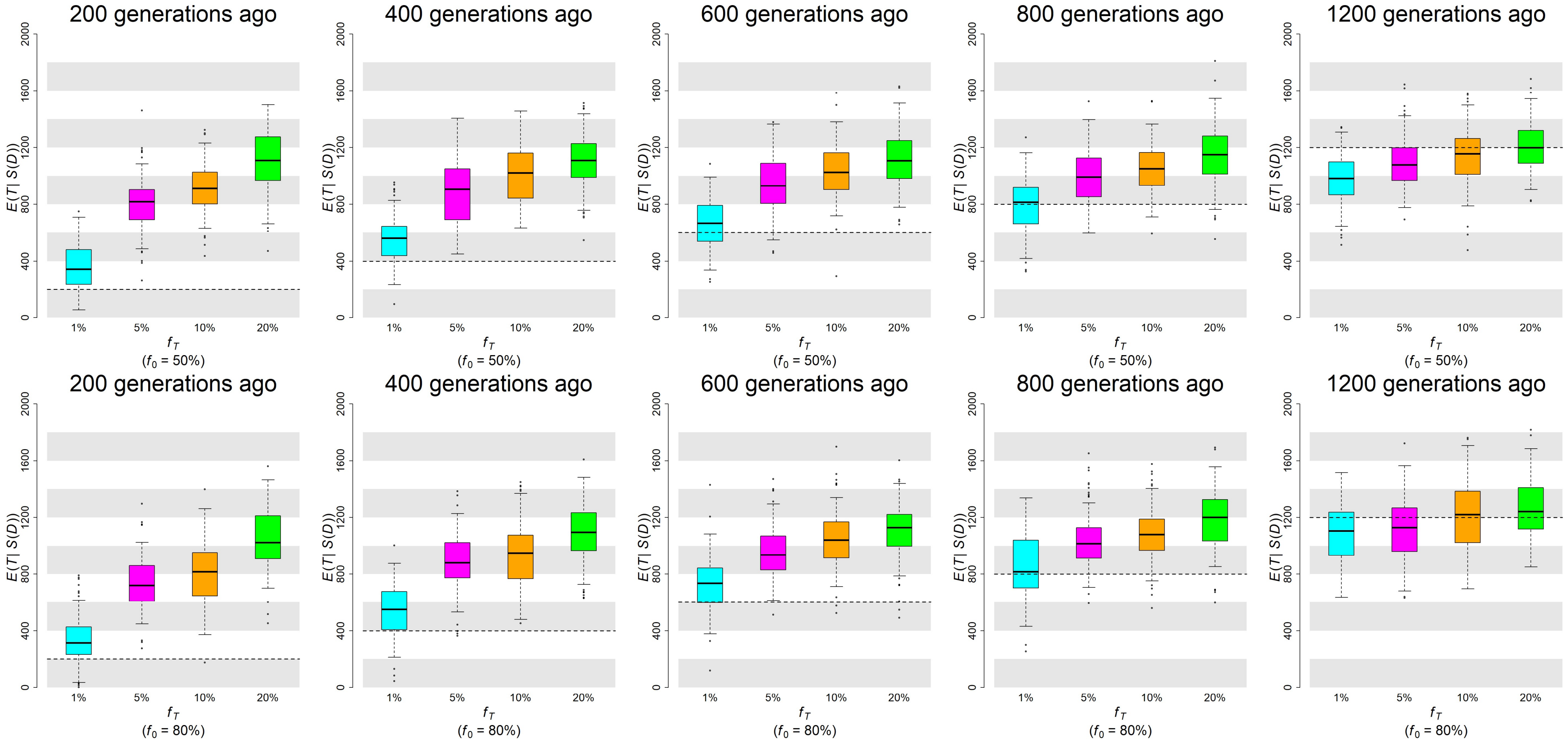

Figure S4. Accuracy of kernel ABC in estimating T under the SSV models. Boxplots show posterior means of T given $S(\mathcal{D})$. 100 pseudo-observations are randomly chosen from the 1,000 observations used in Figure S3 for each combination of the different parameters: $T = 200, 400, 600, 800, 1200$, $f_T = 1\%, 5\%, 10\%, 20\%$, and $f_0 = 50\%, 80\%$. The horizontal dashed lines indicate the true ages.

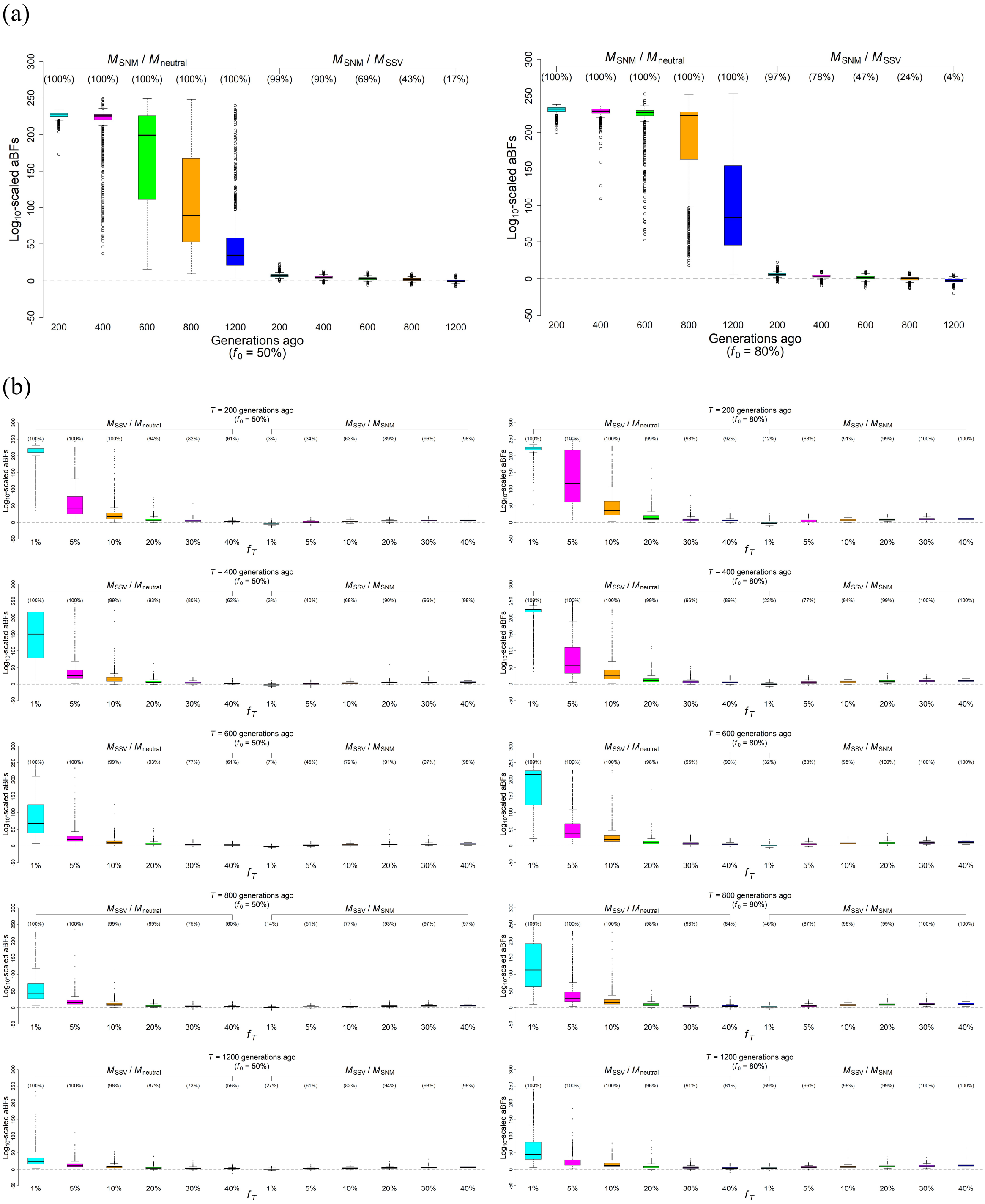

Figure S5. Accuracy of model selection in choosing (a) the SNM models against the neutral or the SSV models and (b) the SSV models against the neutral or the SNM models under the constant size model ($N_e = 12000$). Boxplots show $\log_{10}$-scaled aBFs that take ratios of aMLs between a true model (*i.e.*, SSV) and an alternative model (*i.e.*, neutral or SNM). The percentage with the parentheses above each boxplot shows the probability that the method can correctly identify the true model with $\log_{10}(\text{aBF}) > 2$. The total 1,000 pseudo-observations under the SSV models are generated with $T = 200, 400, 600, 800, 1200$, $f_T = 1\%, 5\%, 10\%, 20\%, 30\%, 40\%$, and $f_0 = 50\%$ or 80%.

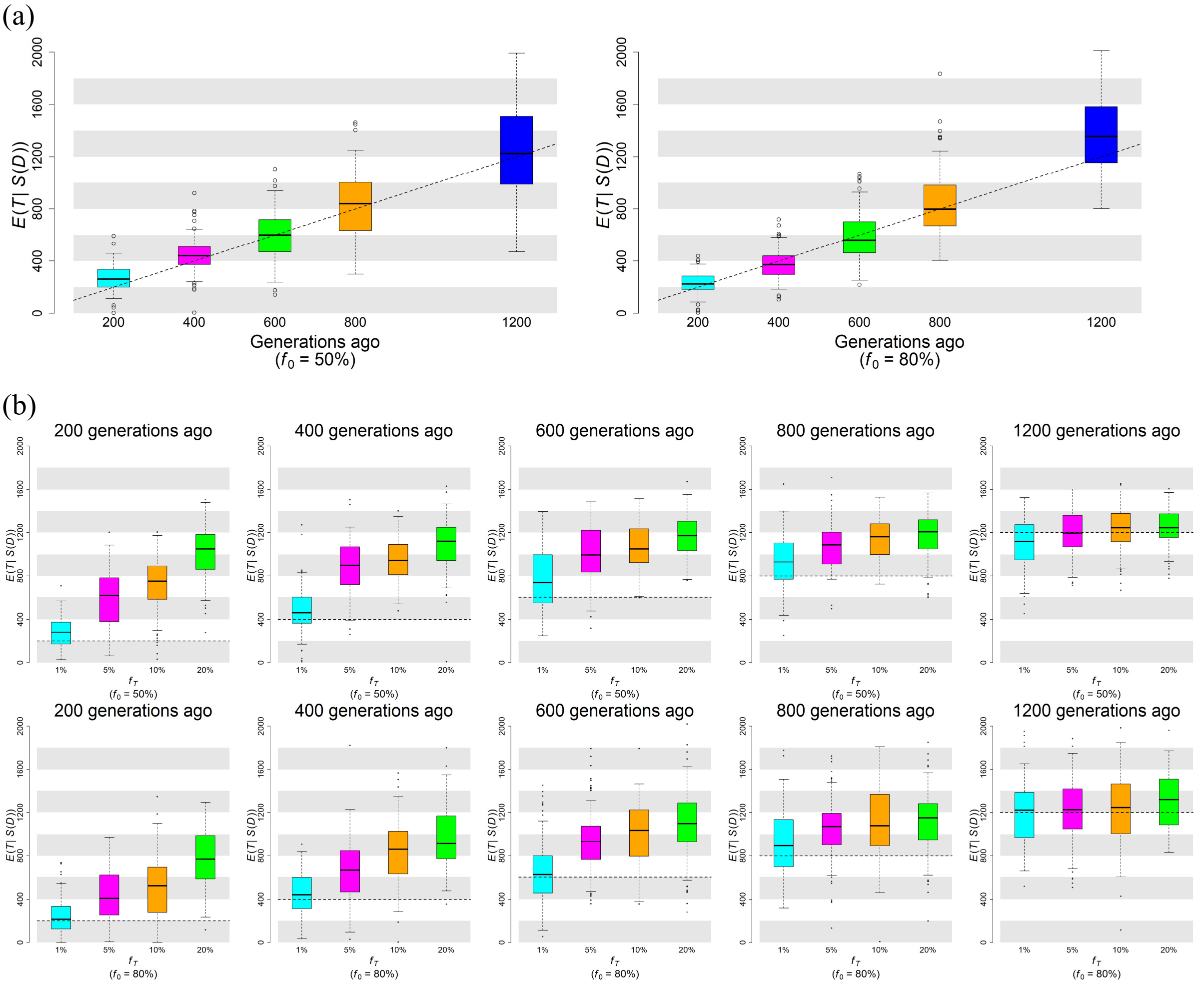

Figure S6. Accuracy of kernel ABC in estimating T under the constant size model ($N_\mathrm{e} = 12000$) with (a) SNM and (b) SSV. Boxplots show posterior means of T given $S(\mathcal{D})$. 100 pseudo-observations are randomly chosen from the 1,000 observations used in Figure S5 for each combination of the different parameters: $T = 200, 400, 600, 800, 1200$ and $f_0 = 50\%, 80\%$ with $f_T = 1/2N_\mathrm{e}$ for (a) and $f_T = 1\%, 5\%, 10\%, 20\%$ for (b). The dashed line in (a) shows a diagonal plot with slope = 1, while the horizontal dashed lines in (b) indicate the true ages.

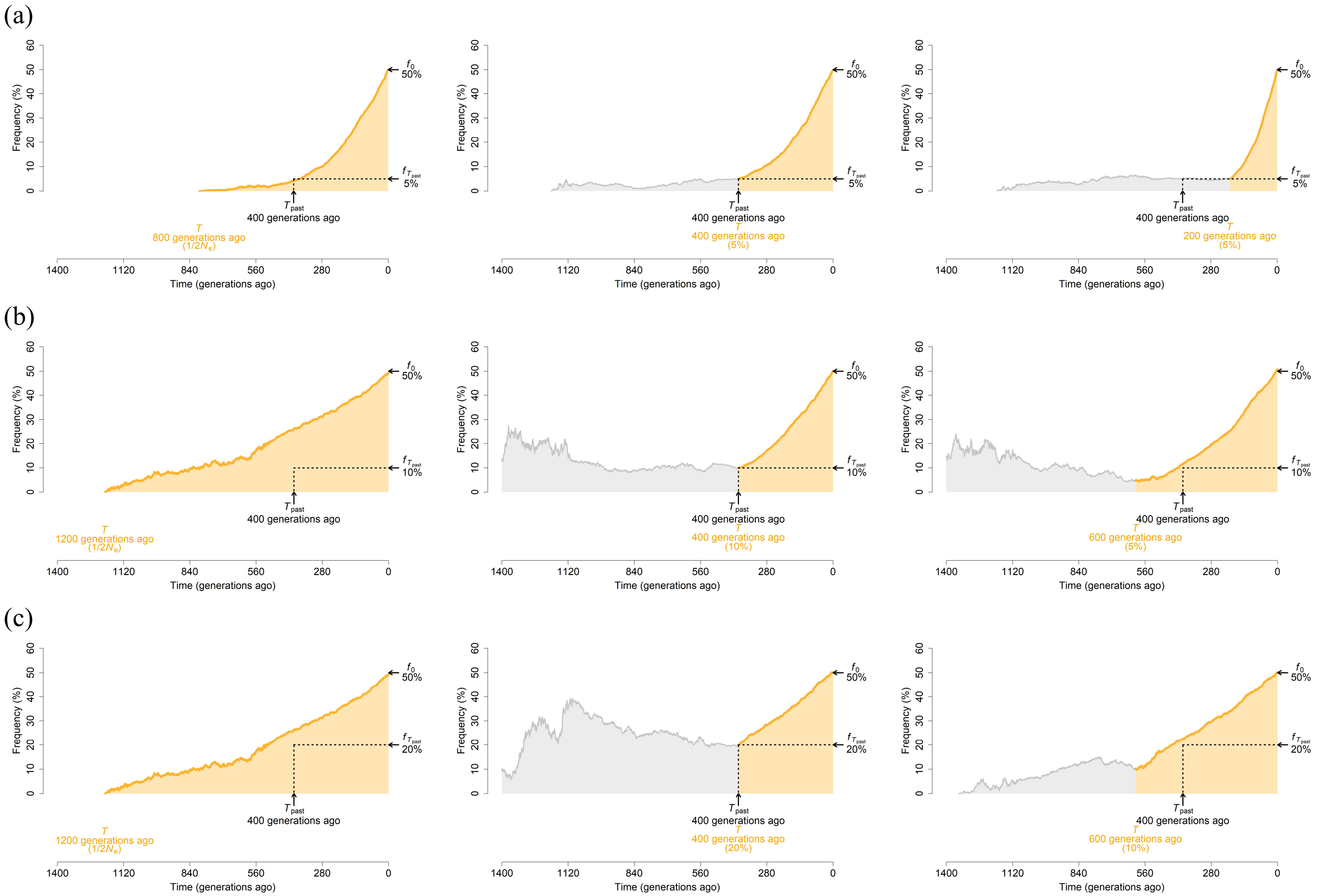

Figure S7. Examples of trajectory in different conditions of SNM and SSV models that generate pseudo-observations used in Figure 2. The selective and neutral processes are highlighted with different colours of orange and grey. To estimate T and f_T for each scenario, ancient DNA data (black dashed lines) are incorporated into kernel ABC with the conditions of $T_{\text{past}} = 400$ and (a) $f_{T_{\text{past}}} = 5\%$, (b) 10%, and (c) 20%. The values of true parameters are shown with the orange colour on the x-axis below the plots.

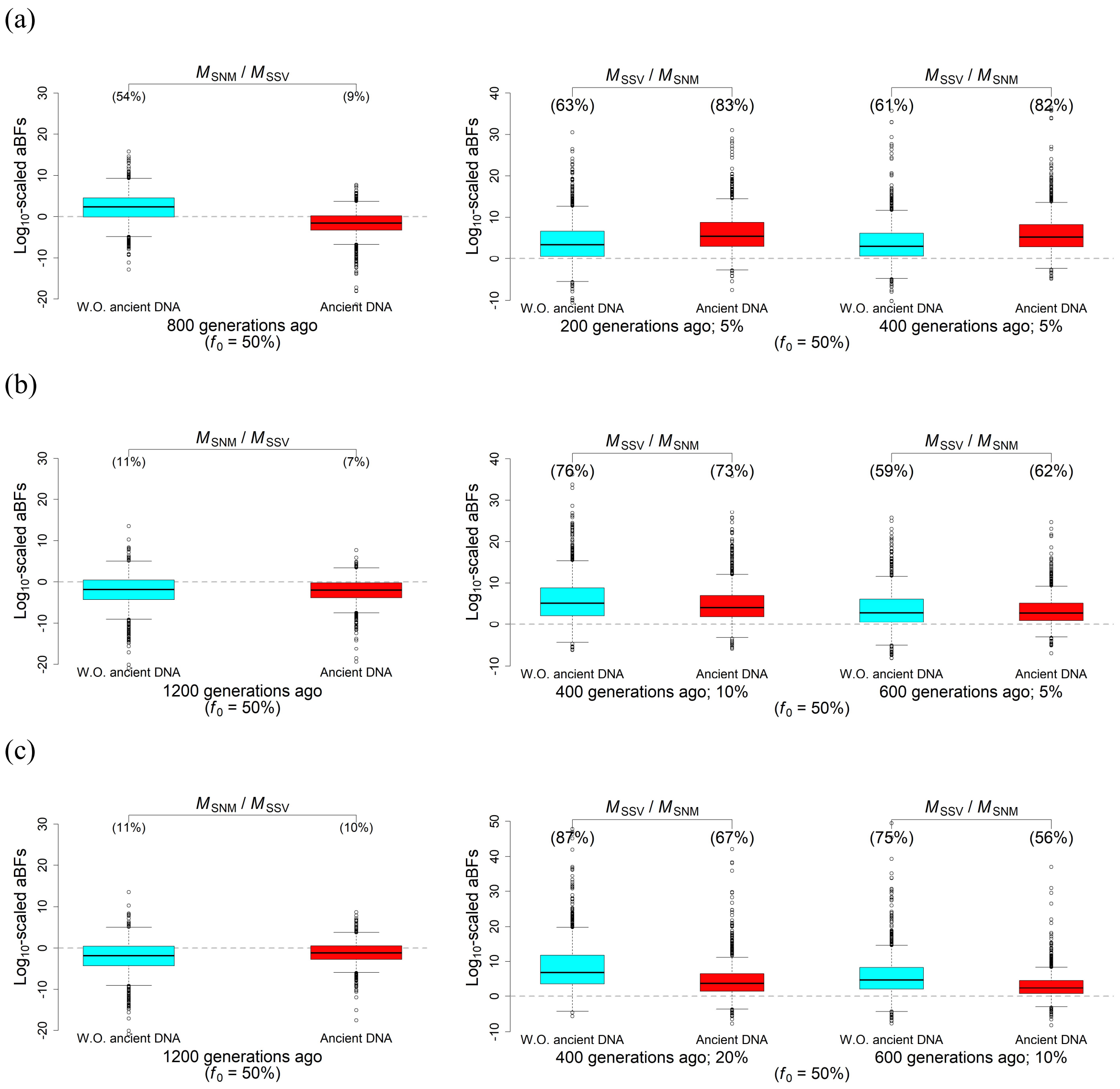

Figure S8. Differences in the performance of model selection with and without (W.O.) ancient DNA data to distinguish between the SNM and SSV models. Different frequencies at $T_{\text{past}} = 400$ are assumed as (a) $f_{T_{\text{past}}} = 5\%$, (b) 10%, and (c) 20%. Boxplots show $\log_{10}$-scaled aBFs that take ratios of aMLs between a true and an alternative model. Different pseudo-observations (total 1,000 each) under SNM or SSV models with $f_0 = 50\%$ are given for different conditions of $f_{T_{\text{past}}}$, and the information on T under SNM or T and f_T under SSV are described at the x-axis. The percentage with the parentheses above each boxplot shows the probability that the method can correctly identify the true model with $\log_{10}(\text{aBF}) > 2$.

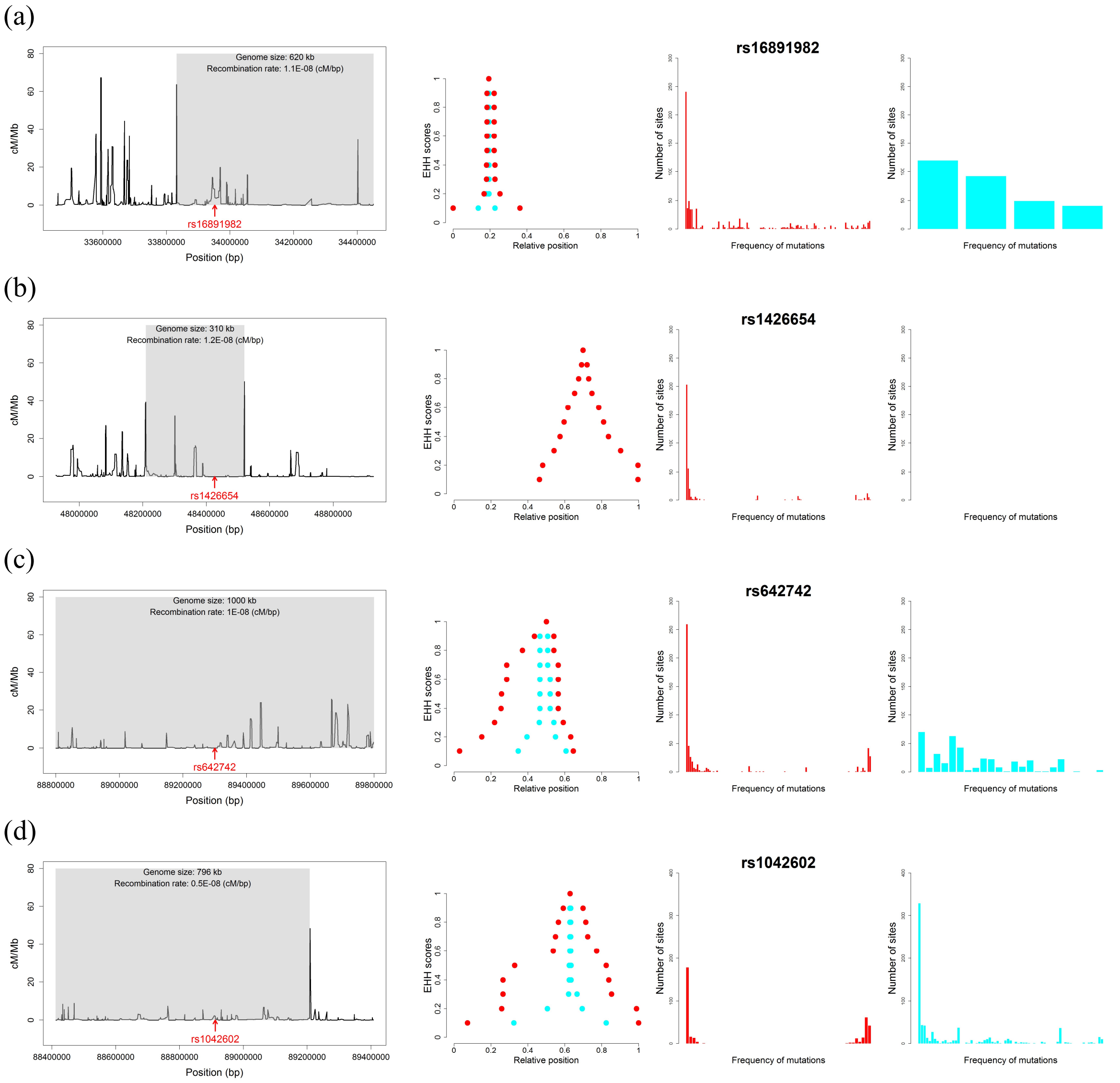

Figure S9. Genetic maps, decays of EHH, and full SFS at (a) rs16891982, (b) rs1426654, (c) rs642742, and (d) rs1042602. The focal SNPs are located at the centre of the genomic region. To remove a region containing recombination hotspots, the genomic region shaded with grey is chosen for the analysis. The genomic size and recombination rate used in the simulation are shown at the top of the shaded area (see also Table S2). Different colours of light blue and red represent the summary statistics calculated from sequences carrying ancestral or derived alleles at the focal SNP site. For rs1426654, since the derived allele is fixed in the population, there is no plot for decays of EHH, and full SFS.

Table S1. List of SNPs associated with skin pigmentation and signatures of natural selection.

SNPs	Genes	Alleles (Ancestral/Derived)	Effect of derived allele on phenotypes	Literatures	
				Skin pigmentation	Natural selection
rs16891982	*SLC45A2*	C/G (Leu/Phe)	↑Lighter skin	Fukamachi *et al*. (2001) *Nat. Genet.* 28:381-385.	Soejima *et al*. (2006) *Mol. Biol. Evol.* 23(1):179-88.
				Graf *et al*. (2005) *Hum. Mut.* 25(3):278-284.	Sabeti *et al*. (2007) *Nature* 449:913-918.
				Stokowski *et al*. (2007) *Am. J. Hum. Genet.* 81(6):1119-1132.	Myles *et al*. (2007) *Hum. Genet.* 120(5):613-621.
				Cook *et al*. (2009) *J. Invest. Dermatol.* 129(2):392-405.	Norton *et al*. (2007) *Mol. Biol. Evol.* 24(3):710-722.
					Hancock *et al*. (2008) *PLOS Genet.* 4(2):e32.
rs1426654	*SLC24A5*	G/A (Ala/Thr)	↑Lighter skin	Lamason *et al*. (2005) *Science* 310(5755):1782-1786.	Voight *et al*. (2006) *PLOS Biol.* 4(3):e72.
				Cook *et al*. (2009) *J. Invest. Dermatol.* 129(2):392-405.	Norton *et al*. (2007) *Mol. Biol. Evol.* 24(3):710-722.
				Beleza *et al*. (2013) *PLOS Genet.* 9(3): e1003372	Sabeti *et al*. (2007) *Nature* 449:913-918.
					Hancock *et al*. (2008) *PLOS Genet.* 4(2):e32.
					Coop *et al*. (2009) *PLOS Genet.* 5(6): e1000500
					Quillen *et al*. (2012) *Hum. Genet.* 131:1073-1080.
rs642742	*KITLG*	T/C	↑Lighter skin	Miller *et al*. (2007) *Cell* 131(6):1179-1189.	Williamson *et al*. (2007) *PLOS Genet.* 3(6):e90.
					Lao *et al*. (2007) *Ann. Huma. Genet.* 71:354-369.
					Pickrell *et al*. (2009) *Genome Res.* 19(5):826-837.
rs1042602	*TYR*	C/A (Ser/Tyr)	↑Lighter skin ↓Freckles	Stokowski *et al*. (2007) *Am. J. Hum. Genet.* 81(6):1119-1132.	Norton *et al*. (2007) *Mol. Biol. Evol.* 24(3):710-722.

				Sulem *et al*. (2007) *Nat. Genet.* 39(12):1443-1452.	Myles *et al*. (2007) *Hum. Genet.* 120(5):613-621
				Chaki et al. (2011) *J. Invest. Dermatol.* 131(1):260-262.	Hancock *et al*. (2008) *PLOS Genet.* 4(2):e32.

Table S2. Summary information of the parameters used on the simulations.

Parameters	Light pigmentation SNPs			
	rs16891982	rs1426654	rs642742	rs1042602
Mean of mutation rate (per bp per generation)	1.690E-08	1.941E-08	1.951E-08	1.980E-08
Recombination rate (cM/bp) (see also Figure S8)	1.1E-08	1.2E-08	1.0E-08	4.8E-9
Number of derived alleles	124	128	104	57
Number of chromosomes	128	128	128	128
Locus size of the subregion (kb) (see also Figure S8)	620	310	1,000	796
Relative position of focal SNP site	19.29%	69.84%	50.00%	62.80%

Table S3. Means of selection coefficients (s) used to generate $1{,}000$ pseudo-observations under different combinations of T, f_T, and f_0.

Selection models	f_T	T									
		$f_0 = 50\%$					$f_0 = 80\%$				
		200	400	600	800	1,200	200	400	600	800	1,200
SNM	$1/2N_\mathrm{e}$	0.104	0.049	0.024	0.017	0.007	0.121	0.057	0.029	0.021	0.010
SSV	1%	0.047	0.023	0.016	0.012	0.006	0.062	0.031	0.020	0.015	0.009
	5%	0.030	0.015	0.010	0.007	0.005	0.044	0.022	0.015	0.011	0.007
	10%	0.022	0.011	0.007	0.005	0.004	0.037	0.018	0.012	0.009	0.006
	20%	0.014	0.007	0.004	0.003	0.002	0.028	0.014	0.009	0.007	0.005
	30%	0.008	0.004	0.003	0.002	0.002	0.023	0.011	0.007	0.006	0.004
	40%	0.004	0.002	0.002	0.002	0.002	0.018	0.009	0.006	0.004	0.003