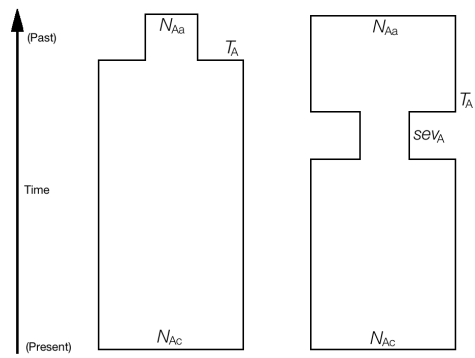


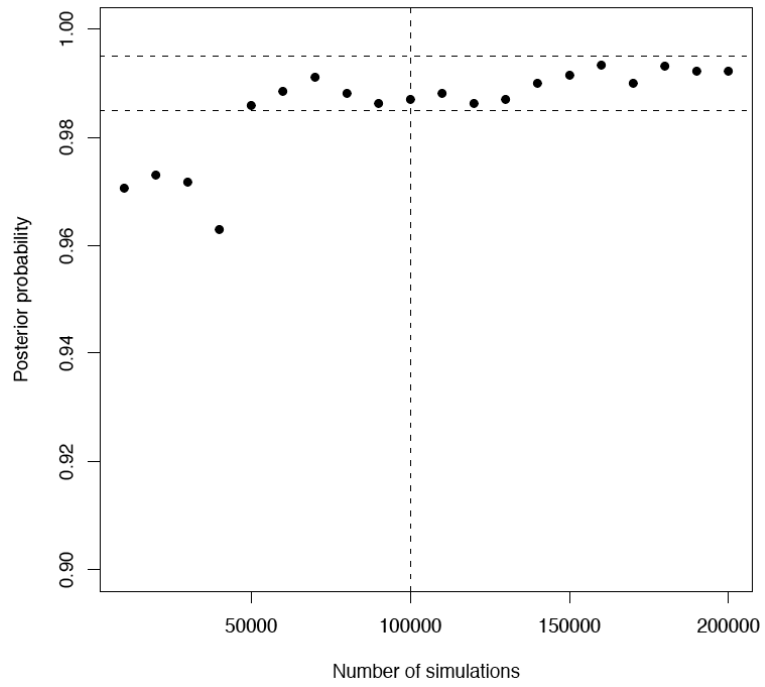
**Figure S1** Phred quality scores of individual base calls belonging to the first six classes of the site frequency spectrum in the North American population (calculated from the DPGP1 raw fastq files). The middle bands indicate the median values, the boxes the upper and lower quartiles and the whiskers the minimum and maximum values.

	0	1	2	...	$n_2-2$	$n_2-1$	$n_2$
0	X	W2					W3
1	W1	W4					W1
2							
...							
$n_1-2$							
$n_1-1$							
$n_1$	W3	W2					X

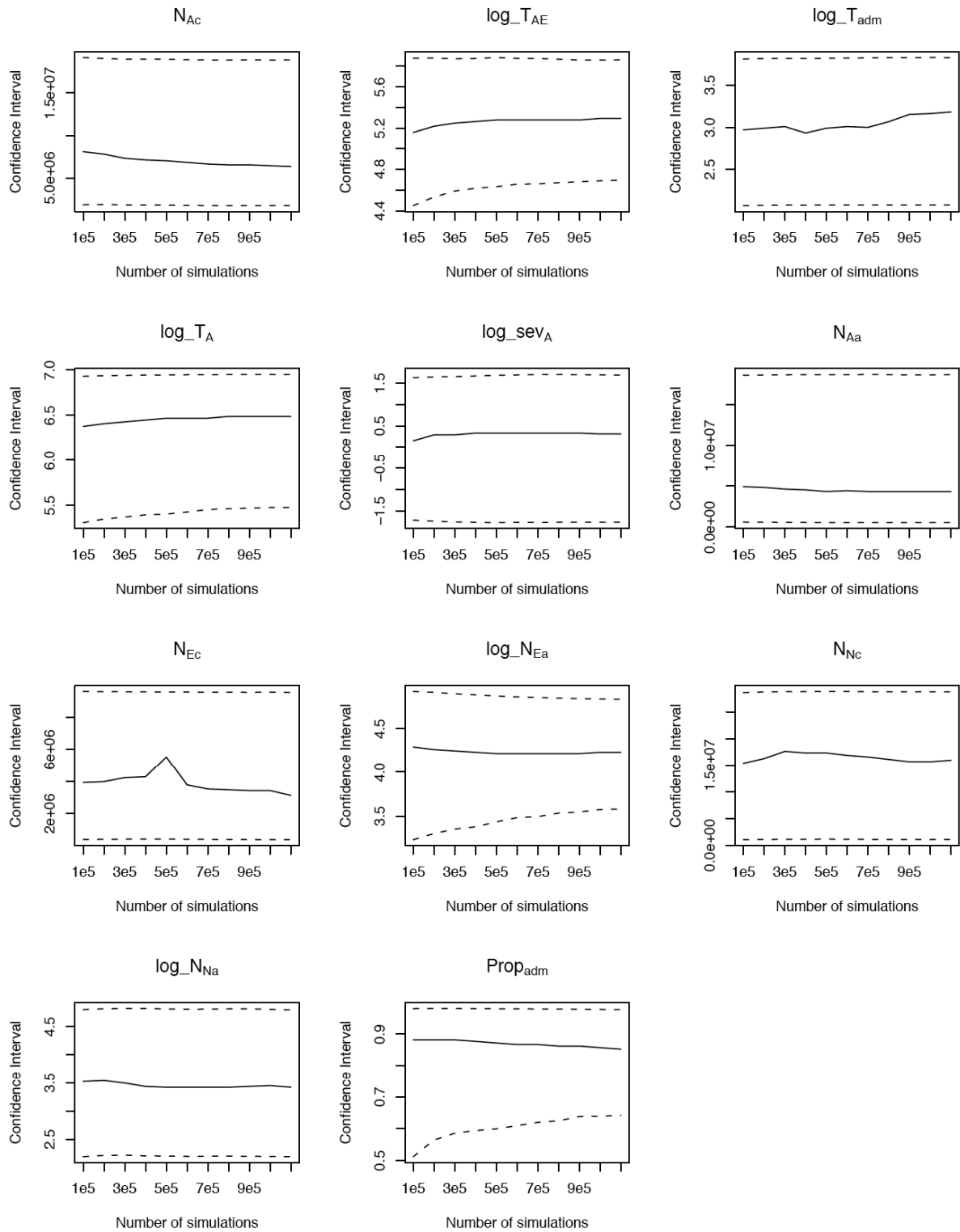
**Figure S2** Joint Site Frequency Spectrum (JSFS) classes, according to the Wakeley-Hey model. On left most column we have the sample size  $n_1$  of population 1. On the upper most row we have the sample size  $n_2$  of population 2. The summary statistics proposed by Wakeley-Hey (1997) are represented by the letters W1 to W4.



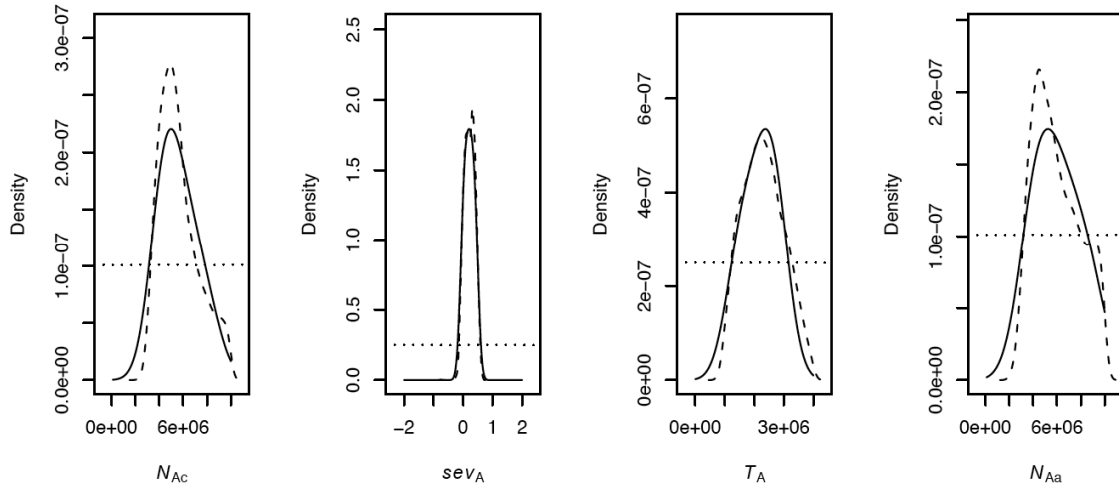
**Figure S3** Population expansion (left) versus Bottleneck (right) model in Africa. The posterior probability of the Expansion model is 0.013. The posterior probability of the Bottleneck model is 0.987. Parameters are explained in Table 3.



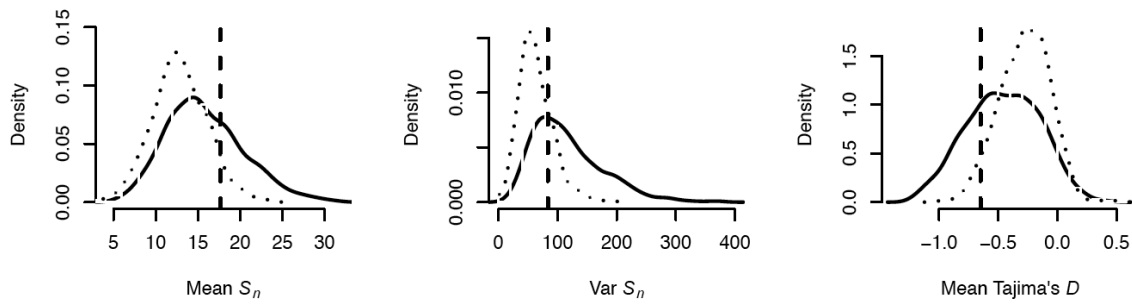
**Figure S4** Behavior of the posterior probabilities of the Bottleneck model for different numbers of simulations. In the case of the Admixture model (model C) the posterior probability is always above 0.999 for different numbers of simulations.



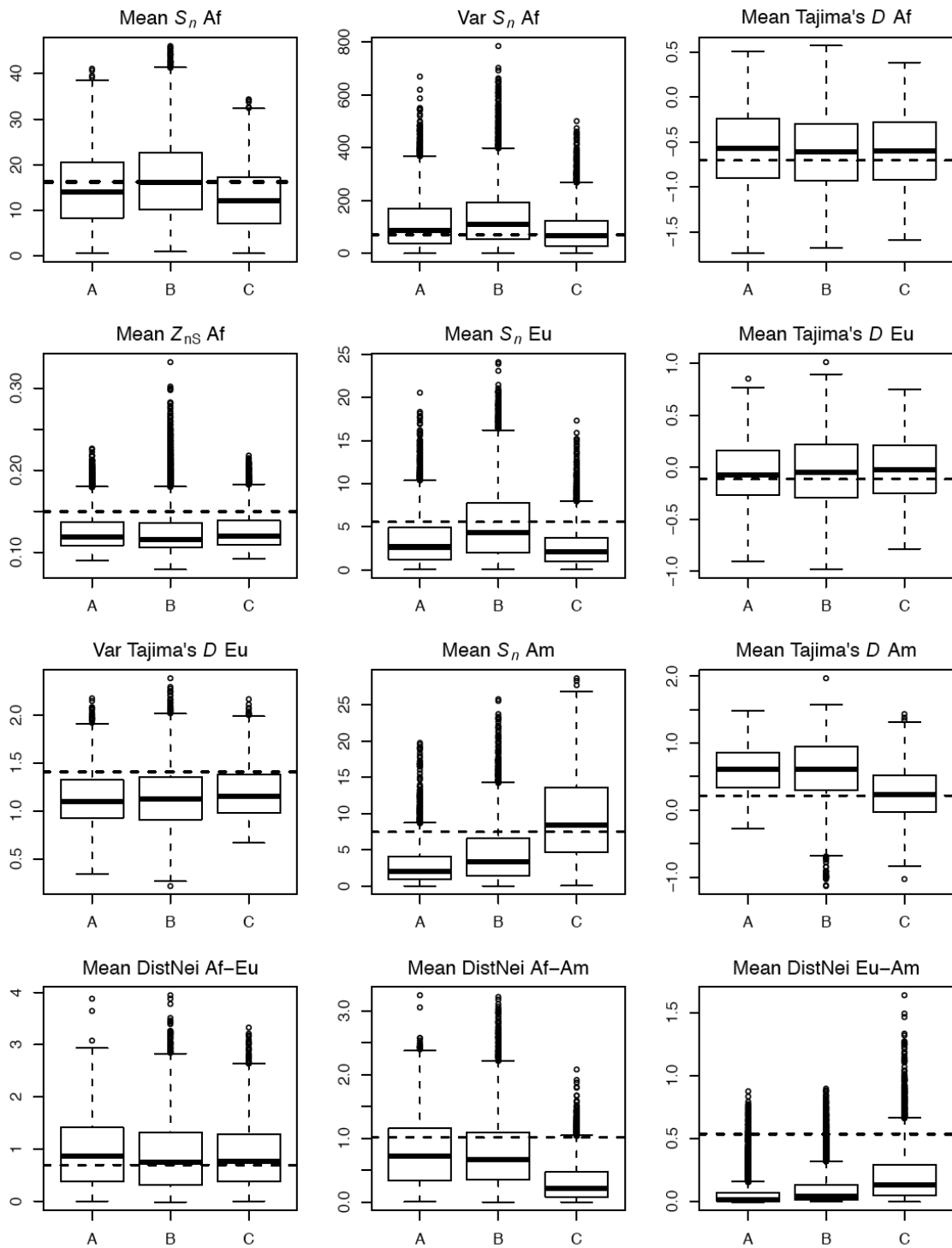
**Figure S5** Behavior of the modes and 95% confidence intervals of the estimates of the parameters of the Admixture model (model C) for different numbers of simulations. Solid line: mode, dashed lines: upper and lower confidence intervals.



**Figure S6** Posteriors of the Bottleneck model in Africa. Posteriors are represented by the rejection method (dashed line) and the regression method (solid line). Parameter abbreviations are explained in Table 3. Mode and confidence interval for each parameter are shown in Table 4.

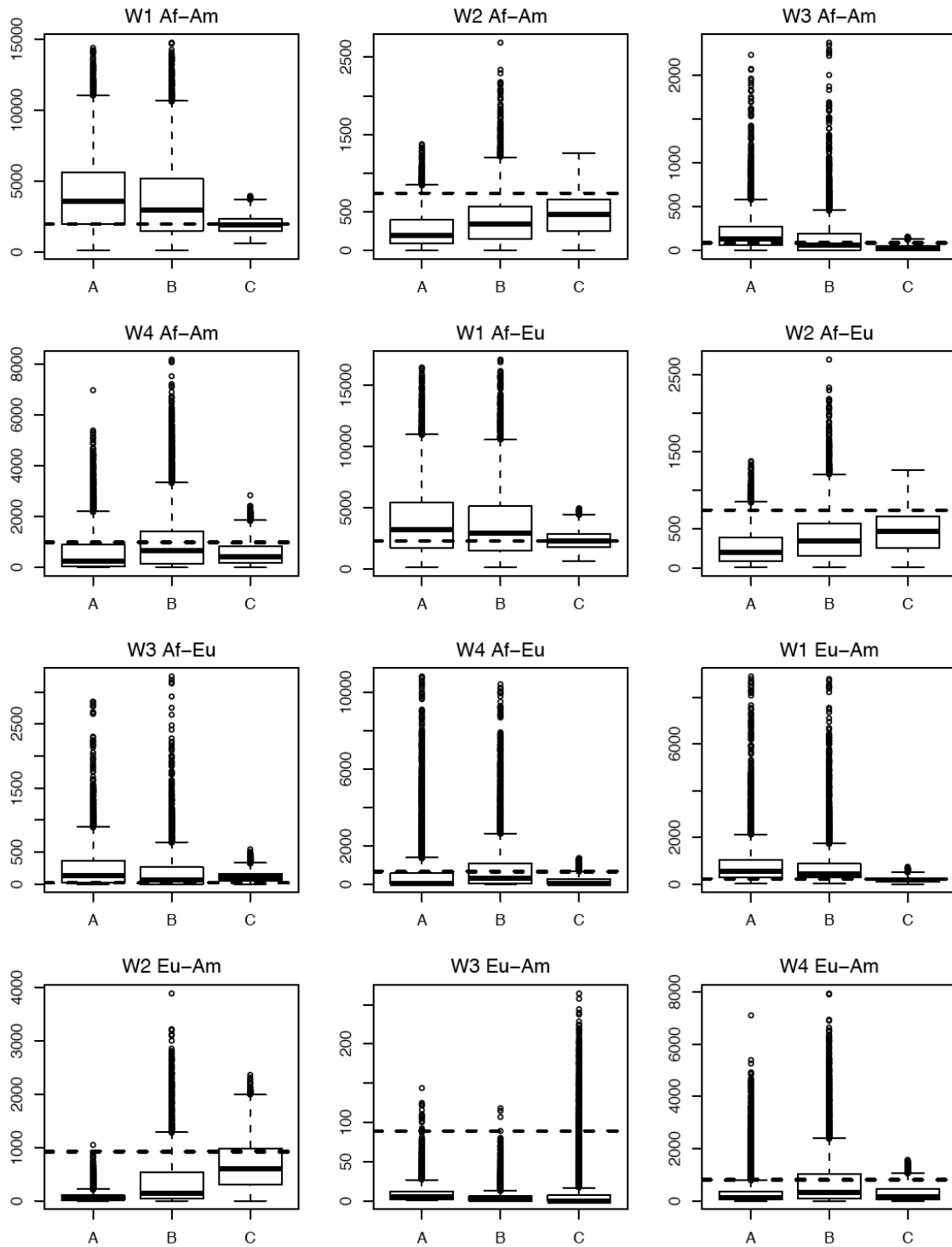


**Figure S7** Predictions of the Bottleneck versus Population Expansion in Africa. Solid line: Bottleneck, dotted line: Population expansion, vertical dashed line: observed value. Parameters for predictive simulations are drawn from the posterior distributions generated by the regression method (see Materials and Methods).

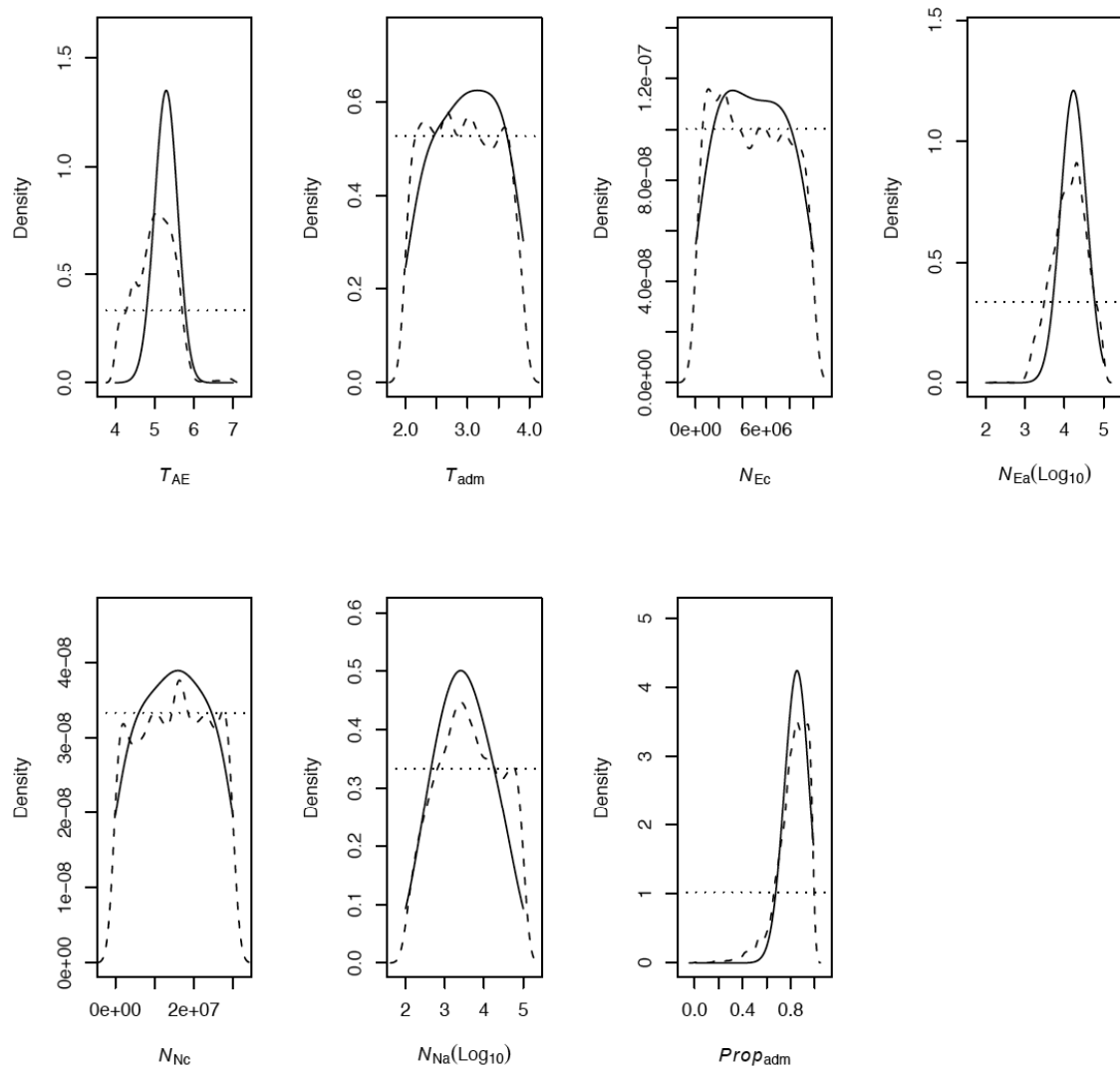


**Figure S8** Predictions of summary statistics for models A, B and C based on the rejection method. The horizontal dashed line represents the observed value.

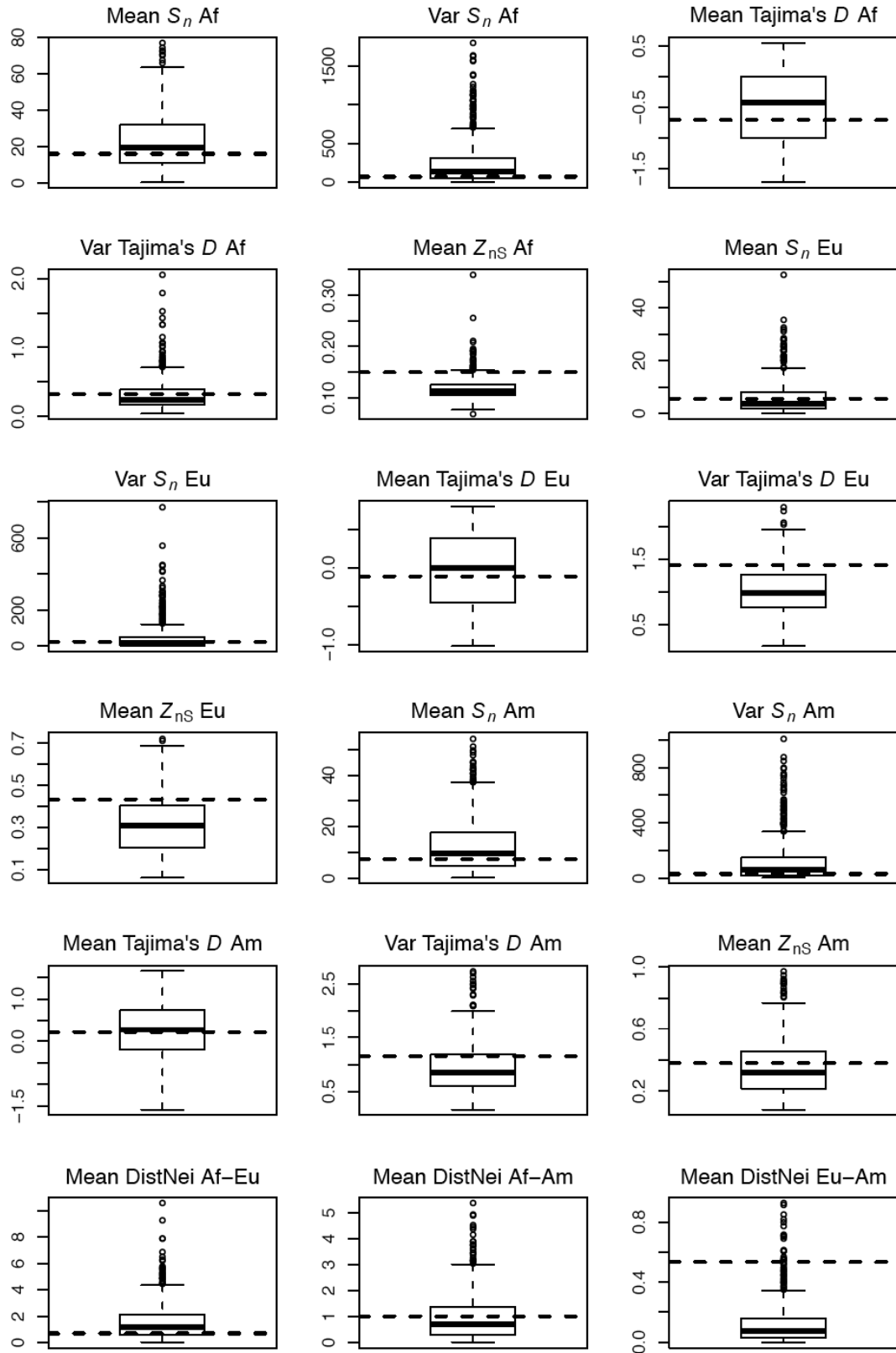




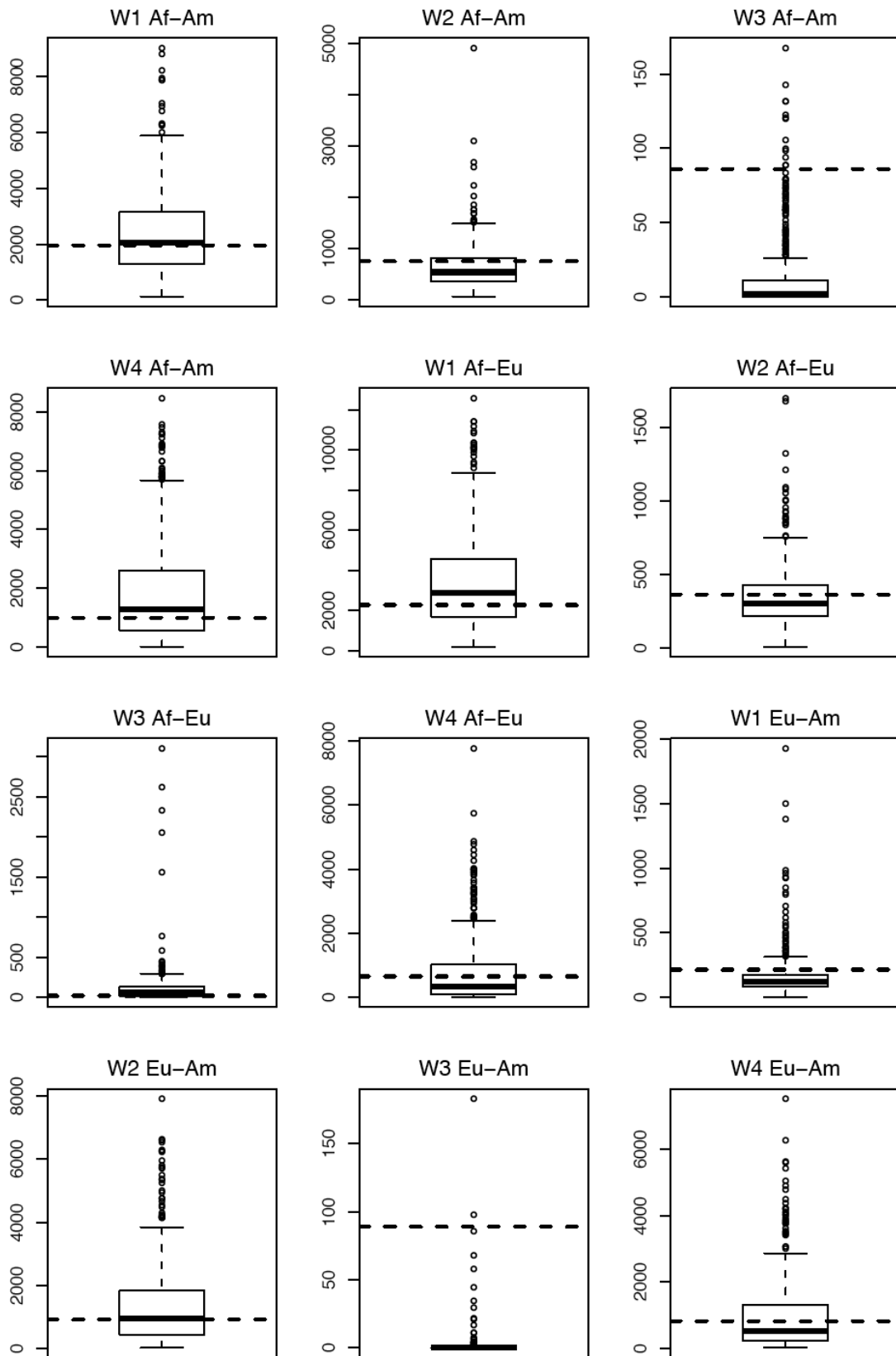
**Figure S9** Predictions of the JSFS for models A, B and C based on the rejection method. The horizontal dashed line represents the observed value.



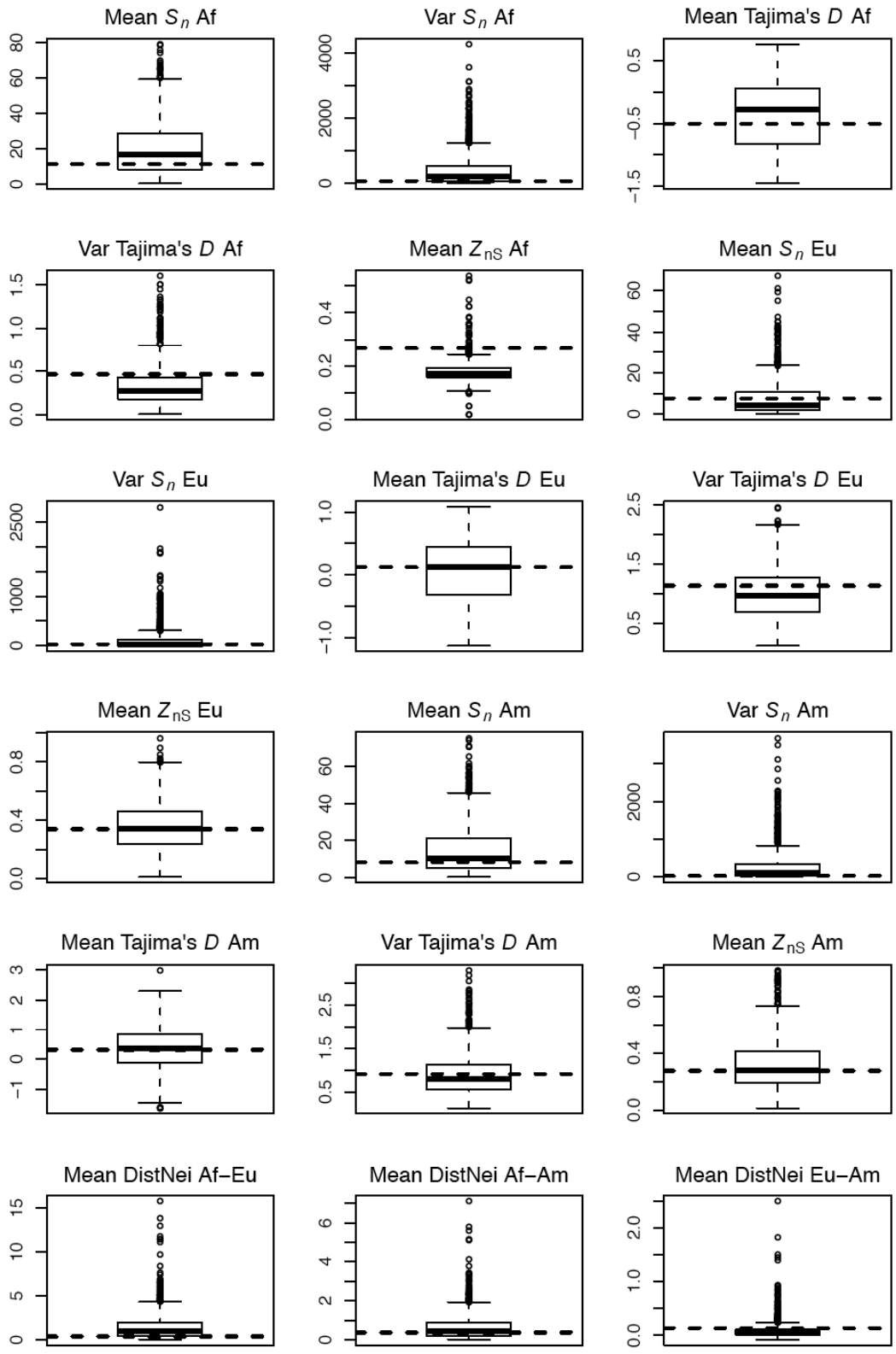
**Figure S10** Posteriors of the Admixture model C. Posteriors are represented by the rejection method (dashed line) and the regression method (solid line). Parameter abbreviations are explained in Table 3. Mode and confidence interval for each parameter are shown in Table 5.



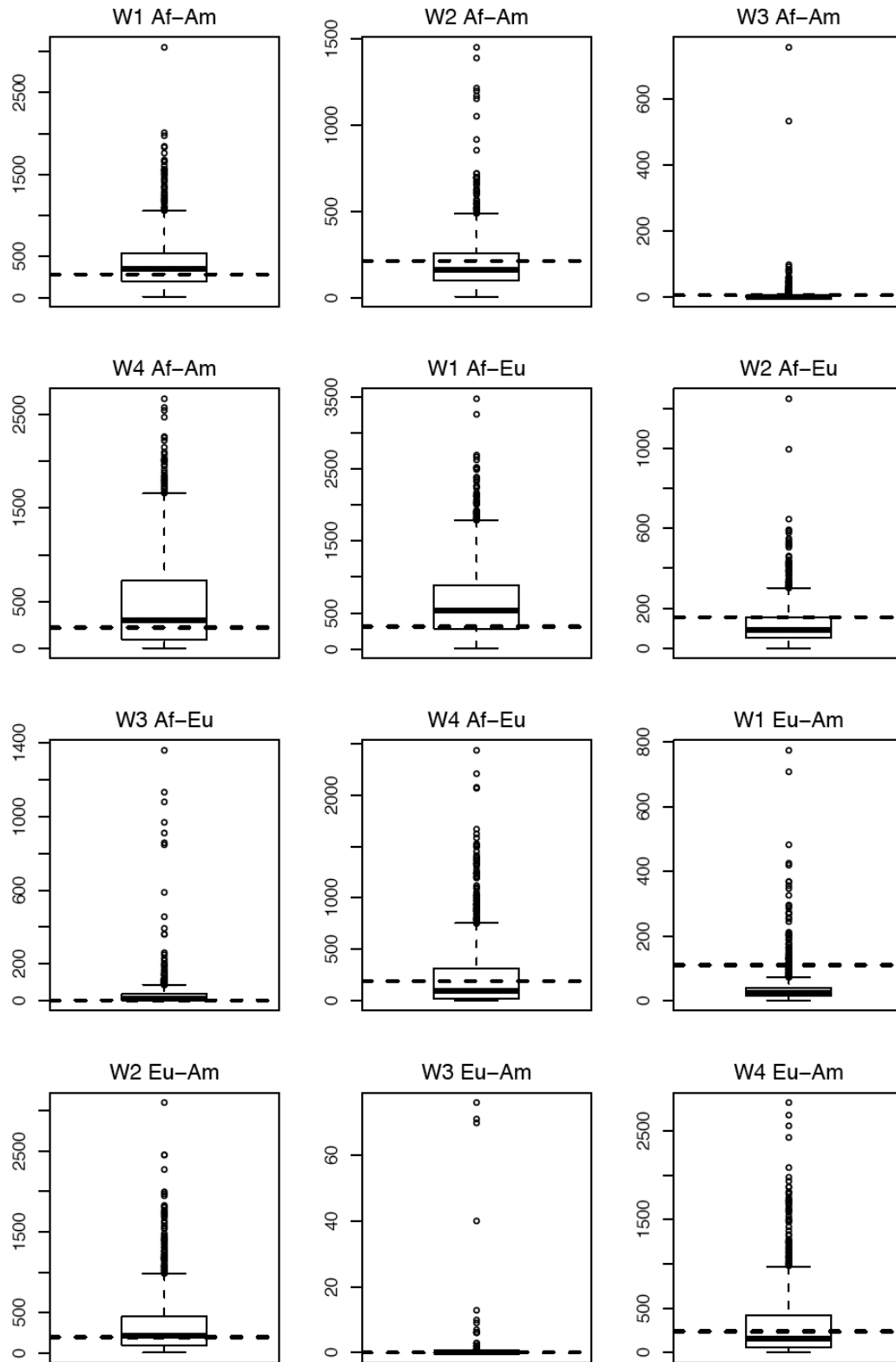
**Figure S11** Predicted statistics of model C. Predictions of the mean and variance of  $S_n$ , mean and variance of Tajima's  $D$  and mean  $Z_{nS}$  are shown for each population. Predicted mean Distance of Nei for all pairs of populations are shown as well. Statistics are predicted by sampling parameters from the posterior distributions based on the regression method (see main text for details).



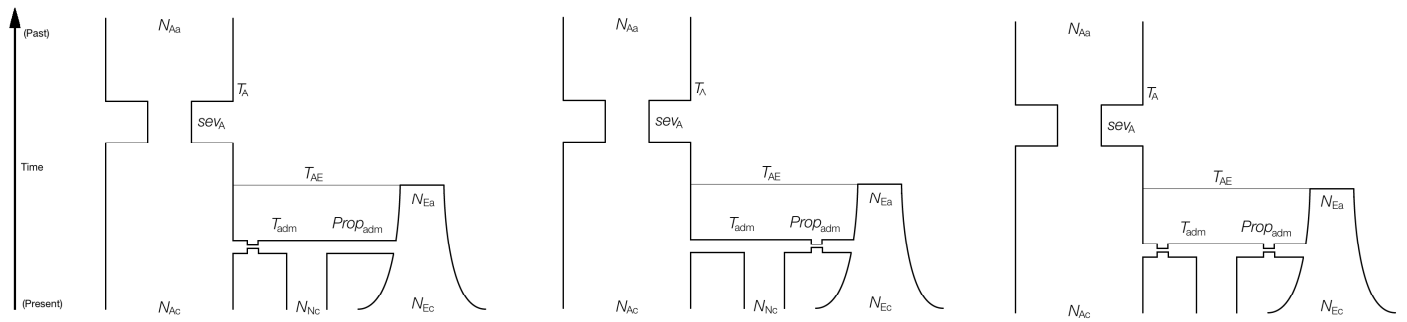
**Figure S12** Predicted JSFS of model C. Predictions of each Wakeley-Hey (1997) class are shown. Statistics are predicted by sampling parameters from the posterior distributions based on the regression method (see main text for details).



**Figure S13** Predicted statistics of model C for autosomal loci (chromosome 3). Predictions of the mean and variance of  $S_n$ , mean and variance of Tajima's  $D$  and mean  $Z_{nS}$  are shown for each population. Predicted mean Distance of Nei for all pairs of populations are shown as well. Statistics are predicted by sampling parameters from the posterior distributions based on the regression method (see main text for details).



**Figure S14** Predicted JSFS of model C for autosomal data (chromosome 3). Predictions of each Wakeley-Hey (1997) class are shown. Statistics are predicted by sampling parameters from the posterior distributions based on the regression method (see main text for details).



**Figure S15** Models C1 (left), C2 (middle) and C3 (right).

**Table S1 Parameters and priors used in the one-population models and in models A, B, C, D and E.**

Parameter	Prior	Model
Current size Africa: $N_{Ac}$	$\text{unif}(1 \times 10^5, 1 \times 10^7)$	Bottleneck and Expansion
Time of bottleneck Africa: $T_A$	$\text{unif}(1 \times 10^2, 4 \times 10^5)$	Bottleneck and Expansion
Ancient size Africa: $N_{Aa}$	$\text{unif}(1 \times 10^5, 1 \times 10^7)$	Bottleneck and Expansion
Severity of bottleneck Africa: $sev_A$ (decimal log)	$\text{unif}(-2, 2)$	Bottleneck
Time of split Africa-Europe (decimal log): $T_{AE}$	$\text{unif}(4, 7)$	Model A,B,C,D,E
Time of split Europe-North America (decimal log): $T_{EN}$	$\text{unif}(4, 7)$	Model A,B
Time of split Africa-North America (decimal log): $T_{AN}$	$\text{unif}(4, 7)$	Model D,E
Time of admixture (decimal log): $T_{adm}$	$\text{unif}(2, 4)$	Model C
Current size Europe: $N_{Ec}$	$\text{unif}(1 \times 10^4, 1 \times 10^7)$	Model A,B,C,D,E
Ancient size Europe (decimal log): $N_{Ea}$	$\text{unif}(2, 5)$	Model A,B,C,D,E
Current size North America: $N_{Nc}$	$\text{unif}(1 \times 10^4, 3 \times 10^7)$	Model A,B,C,D,E
Ancient size North America (decimal log): $N_{Na}$	$\text{unif}(2, 5)$	Model A,B,C,D,E
Proportion of European admixture: $Prop_{adm}$	$\text{unif}(0.01, 0.99)$	Model C
Migration rate (decimal log): $M$	$\text{unif}(-10, -2)$	Model B,E



**Table S2 Three-population models covered in this study.**

Model	Description	Posterior Probability
A	“No migration” model. Comprises Africa as the ancestral population, colonization of Europe followed by exponential growth, and the colonization from Europe to North America with subsequent exponential growth.	< 0.001
B	“Migration” model, matches Model A but adds an equal migration rate between all populations starting at the colonization time of North America.	< 0.001
C	“Admixture” model, equals the previous models until the North American population is founded through an admixture between Africa and Europe followed by exponential growth in North America.	> 0.999
D	“No migration II” model, North America and Europe split independently from Africa, no migration.	< 0.001
E	“Migration II” model, same as model D plus one single rate of migration starting when the North American population is founded.	< 0.001

**Table S3 Mean squared error (MSE) of the ( $\log_{10}$ ) parameter estimates of model C for varying numbers of simulations.**

	100000	200000	300000	400000	500000	600000	700000	800000	900000	1000000
$N_{Ac}$	0.019	0.0101	0.00718	0.00574	0.00443	0.00296	0.00201	0.00149	0.00154	0.00125
$T_{adm}$	0.242	0.293	0.275	0.306	0.335	0.291	0.309	0.305	0.326	0.322
$T_{AE}$	0.0693	0.0388	0.0271	0.0214	0.018	0.0147	0.0128	0.011	0.00996	0.00927
$T_A$	0.0447	0.0498	0.043	0.0407	0.0352	0.0318	0.0317	0.025	0.0233	0.0203
$sev_A$	0.0178	0.03	0.03	0.0307	0.03	0.0311	0.0298	0.0291	0.03	0.0326
$N_{Aa}$	0.00114	0.00243	0.00422	0.00464	0.00661	0.00688	0.00767	0.00835	0.00869	0.00871
$N_{Ec}$	0.0221	0.0839	0.0831	0.111	0.0804	0.0818	0.069	0.0658	0.0434	0.0366
$N_{Nc}$	0.000554	0.000369	0.00059	0.000937	0.00103	0.000636	0.00054	0.000316	0.000336	0.000402
$N_{Ea}$	0.00605	0.00624	0.0075	0.00801	0.0086	0.00942	0.0104	0.011	0.0118	0.0123
$N_{Na}$	0.471	0.534	0.514	0.505	0.457	0.444	0.443	0.467	0.517	0.509
$Prop_{adm}$	0.00148	0.00149	0.00169	0.00196	0.00213	0.00222	0.00222	0.00221	0.00219	0.00214