

## Appendix

**Table S1.** Accuracy of PCAdmix and HAPMIX on Simulated Two-Population Data

<i>Calling Threshold</i>	<i>PCAdmix</i>	<i>HAPMIX</i>
0.5	98.1	99.2
0.8	98.6	99.3
0.9	98.8	99.3
0.95	99.0	99.3
0.99	99.2	99.5
0.999	99.5	99.8

**Table S2.** Accuracy of Our Method under Different Window Sizes

<i>SNPs per window</i>	<i>Median window length (kb)</i>	<i>Accuracy (calling threshold = 0.5)</i>
1	1 base pair	92.2
2	1.7	93.5
5	9.3	96.2
10	22.3	97.3
15	35.3	98.4
20	49.0	98.1
40	112.3	98.6
80	223.8	98.7
160	481.9	97.6

To accentuate the differences among the window sizes, a calling threshold of 0.5 was used.

**Table S3.**

<i>Estimated G</i>	<i>Accuracy (0.5)</i>	<i>Accuracy (0.9)</i>
1	98.6	99.1
2	98.5	98.9
4	98.2	98.9
<b>8</b>	<b>98.1</b>	<b>98.8</b>
16	97.7	98.7
32	97.4	98.6
64	96.7	98.3
128	95.7	98.1

Accuracy under different values of G the true value of **G**, the number of Generations since admixture, for the simulations was 8. Calling threshold is in parentheses.

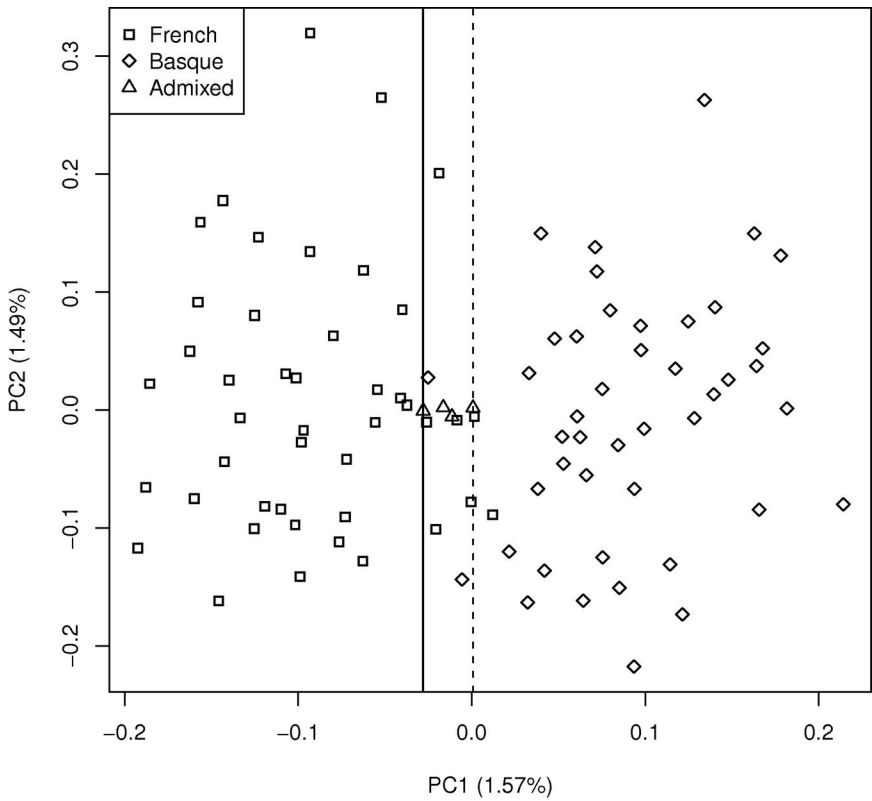
**Table S4.** Haploid Accuracy on Simulated Chromosomes with Ancestry from Three Populations, Including Yoruba and French

<i>3rd Pop</i>	<i>Overall Accuracy</i>	<i>Accuracy for SNPs Assigned to a Given Population</i>		
		<i>Assigned to Yoruba</i>	<i>Assigned to French</i>	<i>Assigned to 3rd Pop</i>
Adygei	76.08	97.80	65.81	49.15
Balochi	83.80	97.39	78.94	72.35
Bantu NE	71.75	51.42	94.89	76.05
Basque	71.50	97.54	43.08	65.10
Bedouin	81.97	95.03	70.84	75.32
Biaka Pygmies	90.40	79.47	97.15	94.23
Brahui	89.73	98.94	85.26	76.68
Burusho	87.01	98.56	71.97	90.50
Druze	83.84	96.35	67.57	82.18
Han	97.78	98.20	97.95	97.13
Hazara	85.70	97.75	70.84	93.94
Japanese	95.58	98.63	92.60	95.19
Kalash	91.01	98.15	88.53	81.95
Karitiana	97.41	98.59	94.86	98.35
Makrani	88.43	97.97	87.46	67.56
Mandenka	83.83	82.29	97.69	47.24
Maya	95.82	97.14	96.03	92.81
Mbuti Pygmies	91.67	83.75	97.43	98.76
Mozabite	82.71	95.28	72.81	82.09
North Italian	73.12	97.51	55.35	69.18
Orcadian	80.68	97.68	71.76	45.58
Palestinian	84.84	97.75	85.38	58.46
Papuan	96.98	98.23	96.57	95.42
Pathan	86.89	98.12	80.69	79.26
Pima	96.81	98.62	95.39	97.11
Russian	70.73	97.13	53.95	50.44
Sardinian	84.36	99.06	78.07	59.32
Sindhi	86.88	98.35	75.52	87.53
Yakut	97.75	98.93	96.91	97.49

All analyses were performed with a window size of 20 SNPs, LD threshold of  $r^2 \leq 0.8$ ,  $\hat{G} = 8$ , and a calling threshold of 0.8.

**Table S5.** Regions with Extreme Ancestry Proportions Reaching Bonferroni-Corrected Statistical Significance

<i>Chromosome</i>	<i>Position (Mb)</i>	<i>Ancestry</i>	<i>Population</i>	<i>Genome-Wide Ancestry Mean (SD)</i>	<i>Ancestry Proportion In Region</i>
2	35.16–35.32	NAmer	DOM	0.077 (0.044)	0.315
6	31.38–31.44	YRI	ECU	0.065 (0.048)	0.325
8	37.24–37.57	YRI	ECU	0.065 (0.048)	0.325



**Figure S1.** Example of Method C approach to filtering ancestral representatives of closely related populations. We excluded French haplotypes on the right of the solid line and Basque haplotypes on the left of the dashed line.