

Supplemental Data

Analysis of Genomic Admixture in Uyghur and Its Implication in Mapping Strategy

Shuhua Xu, Wei Huang, Ji Qian, and Li Jin

Figure S1. Relationships of Marker Information Estimated from 20,177 SNPs

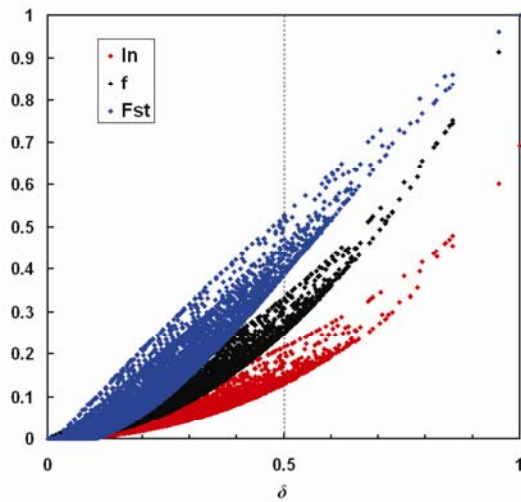


Figure S2. Distribution of Marker Information for 20,177 SNPs on Chromosome 21 Based on Allele Frequencies in CEU and CHB

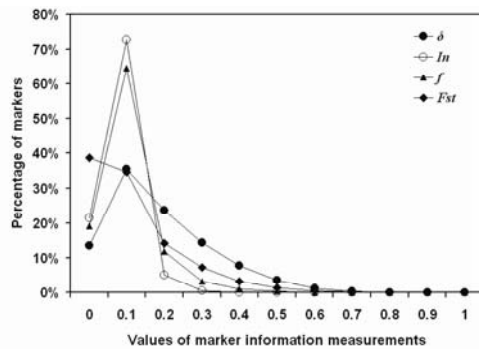
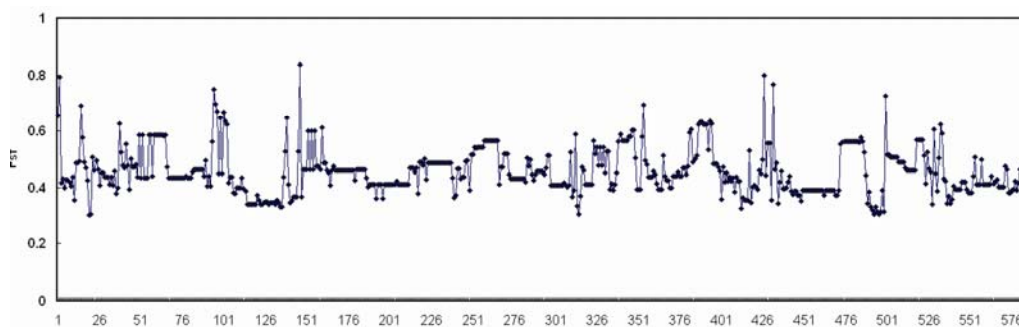
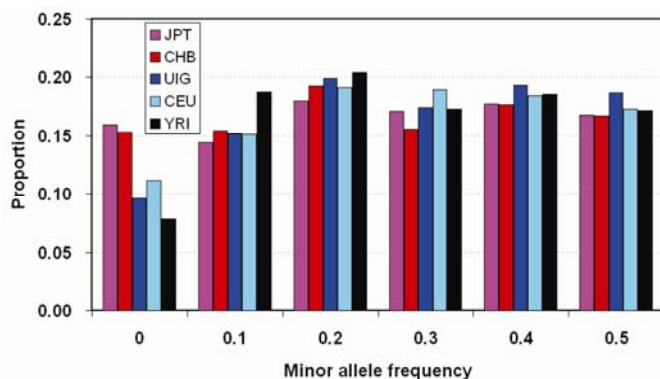


Figure S3. Distribution of F_{ST} for 602 AIMs on Chromosome 21 Based on Allele Frequencies in CEU and CHB

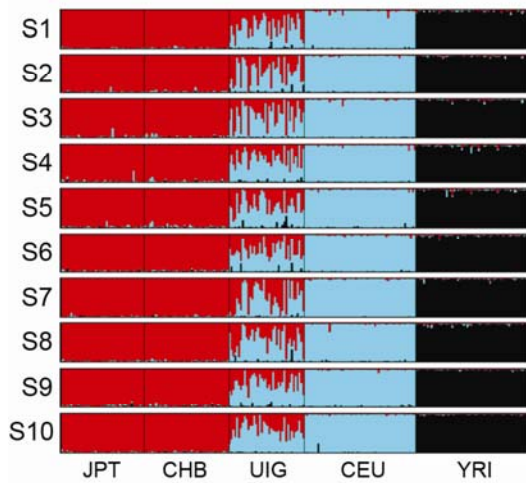


Many adjacent SNPs have the same or similar F_{ST} values and formed “blocks” of F_{ST} .

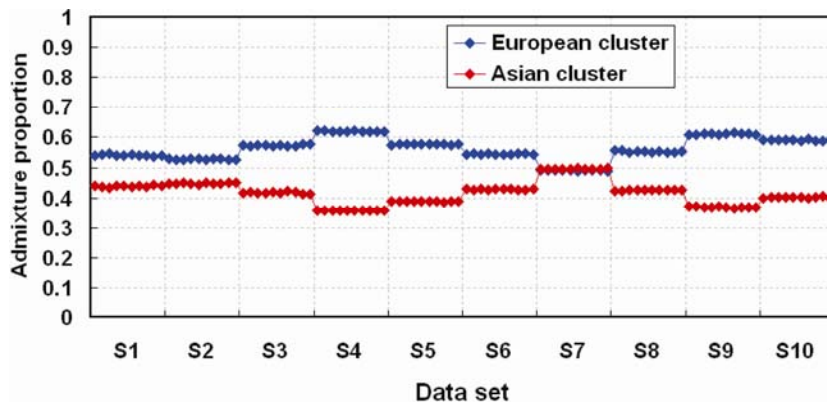
Figure S4. Distribution of Minor-Allele Frequency (MAF) of 20,177 SNPs in Five Populations



UIG has an even higher proportion of common alleles (high-MAF SNPs) than does YRI, and average heterozygosity of UIG is also larger than that of YRI, which is one characteristic of admixture populations.

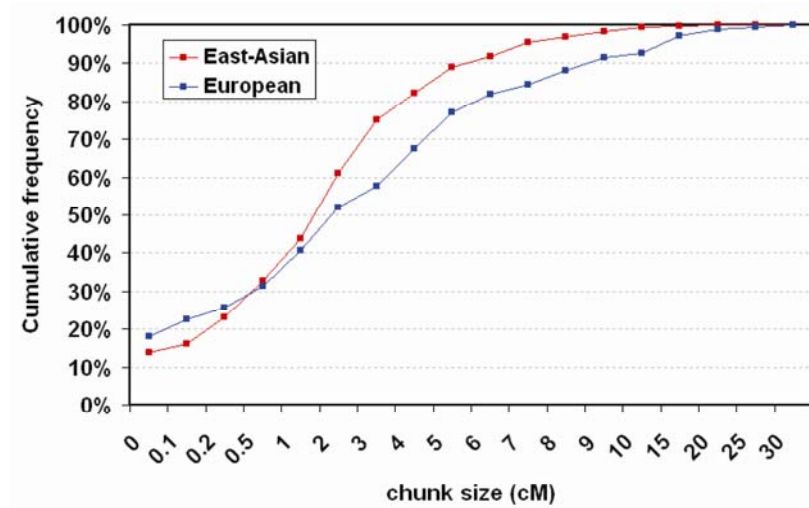
Figure S5. Summary Plot of Individual Admixture Proportions

The STRUCTURE analysis results ($K = 3$) of individual admixture proportions estimated from ten data sets of randomly selected markers (S1~S10). Each individual is represented by a single vertical line broken into three colored segments, with lengths proportional to each of the two inferred clusters. Red indicates East-Asian ancestry proportion, blue indicates European ancestry proportion, and black indicates African ancestry proportion. The predefined population IDs (JPT, CHB, UIG, CEU, and YRI) are presented in the abscissa. The ordinate indicates the proportion unit.

Figure S6. Admixture Proportions of UIG Samples at Population Level Estimated from Ten Data Sets of Randomly Selected Markers

Within each data set, estimations of ten repeat runs are consistent, but there are fluctuations of estimation between different data sets. On average, the UIG population has 56.2% of admixture from European ancestry and 41.7% of admixture from East-Asian ancestry.

Figure S7. Cumulative Distribution of the Ancestral Chromosome Block Sizes in UIG



The sizes of ancestral chromosome segments were calculated separately for East-Asian-derived segments (red square) and European-derived (blue square) segments.

Table S1. Basic Statistics of Ten Randomly Sampled Data Sets

Data Sets	# of SNPs	BMD (kb)	Het (exp)	Het (obs)
S1	148	226	.332	.298
S2	148	226	.333	.303
S3	150	224	.326	.299
S4	149	225	.331	.299
S5	149	225	.359	.328
S6	148	226	.321	.290
S7	147	227	.300	.271
S8	150	224	.335	.305
S9	148	226	.323	.289
S10	148	226	.332	.298
mean	149	226	.329	.298

Table S2. Comparison of LDs of UIG and CHB for Alleles with MAFs ≥ 0.05

	$r^2 \geq 0.8$	$r^2 \geq 0.5$	$r^2 \geq 1/3$	$r^2 \geq 0.2$	$r^2 \geq 0.1$	$r^2 < 0.1$	mean
10	0.85	1.02	0.95	1.07	1.15	1.10	1.01
20	0.81	0.96	0.88	1.05	1.05	1.08	0.95
30	0.73	0.95	0.94	1.01	1.03	1.06	0.93
40	0.62	0.94	0.91	0.96	1.05	1.05	0.90
50	0.64	0.95	0.86	0.96	1.01	1.04	0.89
60	0.58	1.00	0.81	1.01	0.98	1.03	0.88
70	0.57	1.06	0.75	0.93	1.05	1.02	0.87
80	0.64	1.00	0.73	0.91	0.99	1.02	0.85
90	0.62	1.05	0.71	0.86	0.91	1.03	0.83
100	0.66	0.99	0.76	0.83	0.99	1.02	0.85
200	0.61	0.70	0.73	1.04	0.96	1.01	0.81
300	0.38	0.38	0.60	0.76	1.39	1.00	0.70
mean	0.64	0.92	0.80	0.95	1.05	1.04	0.87

Table S3. Comparison of LDs of UIG and CEU for Alleles with MAFs ≥ 0.05

	$r^2 \geq 0.8$	$r^2 \geq 0.5$	$r^2 \geq 1/3$	$r^2 \geq 0.2$	$r^2 \geq 0.1$	$r^2 < 0.1$	mean
10	0.94	1.00	1.04	1.00	1.03	1.03	1.00
20	0.88	1.01	0.98	0.99	1.00	1.04	0.97
30	0.83	1.03	1.02	0.93	0.99	1.03	0.96
40	0.76	0.97	1.07	0.91	1.04	1.02	0.95
50	0.80	0.90	1.14	0.90	1.02	1.02	0.95
60	0.72	0.90	1.18	0.96	0.99	1.01	0.95
70	0.60	0.99	1.11	0.95	1.01	1.01	0.93
80	0.58	0.91	0.98	1.03	1.04	1.00	0.91
90	0.54	0.93	0.99	1.01	1.00	1.01	0.89
100	0.60	0.84	0.94	0.97	1.06	1.00	0.88
200	0.66	0.71	0.83	1.08	1.09	1.00	0.88
300	0.24	0.58	0.56	1.09	1.77	0.99	0.85
mean	0.68	0.90	0.99	0.98	1.09	1.01	0.93

Table S4. Comparison of LDs of UIG and CHB for Common Alleles; MAFs ≥ 0.15

	$r^2 \geq 0.8$	$r^2 \geq 0.5$	$r^2 \geq 1/3$	$r^2 \geq 0.2$	$r^2 \geq 0.1$	$r^2 < 0.1$	mean
10	0.86	1.00	1.04	1.11	1.22	1.09	1.05
20	0.80	0.95	0.94	1.05	1.07	1.12	0.96
30	0.75	0.95	0.96	1.01	1.00	1.09	0.93
40	0.66	0.97	0.88	0.98	1.07	1.06	0.91
50	0.65	0.96	0.87	1.00	1.02	1.04	0.90
60	0.61	1.01	0.85	1.03	0.98	1.03	0.90
70	0.58	1.11	0.86	0.94	1.06	1.01	0.91
80	0.69	0.95	0.83	0.93	1.00	1.02	0.88
90	0.62	1.07	0.80	0.88	0.91	1.03	0.86
100	0.67	1.02	0.77	0.87	1.00	1.02	0.87
200	0.65	0.72	0.79	1.13	0.94	1.01	0.85
300	0.58	0.38	0.59	0.76	1.38	1.00	0.74
mean	0.68	0.92	0.85	0.97	1.06	1.04	0.90

Table S5. Comparison of LDs of UIG and CEU for Common Alleles; MAFs ≥ 0.15

	$r^2 \geq 0.8$	$r^2 \geq 0.5$	$r^2 \geq 1/3$	$r^2 \geq 0.2$	$r^2 \geq 0.1$	$r^2 < 0.1$	mean
10	0.94	1.00	1.08	1.00	1.04	1.05	1.01
20	0.86	1.00	1.05	0.99	0.97	1.08	0.97
30	0.81	1.05	1.08	0.92	0.95	1.06	0.96
40	0.78	0.97	1.12	0.93	0.99	1.03	0.96
50	0.81	0.89	1.17	0.99	1.03	1.00	0.98
60	0.73	0.85	1.25	1.05	0.99	1.00	0.97
70	0.61	0.94	1.23	1.02	1.02	1.00	0.96
80	0.58	0.86	1.06	1.12	1.04	1.00	0.93
90	0.53	0.89	1.04	1.05	1.00	1.01	0.90
100	0.58	0.80	1.08	1.02	1.13	1.00	0.92
200	0.75	0.68	0.89	1.20	1.17	0.99	0.94
300	1.00	0.88	0.50	1.16	1.78	0.99	1.06
mean	0.75	0.90	1.04	1.04	1.09	1.02	0.96

Table S6. Comparison of LDs of UIG and CHB for 602 AIMs

	$r^2 \geq 0.8$	$r^2 \geq 0.5$	$r^2 \geq 1/3$	$r^2 \geq 0.2$	$r^2 \geq 0.1$	$r^2 < 0.1$	mean
10	1.05	2.56	1.03	1.25	0.26	0.11	<i>1.23</i>
20	0.94	3.09	1.39	1.02	0.47	0.20	<i>1.38</i>
30	0.91	2.21	1.75	0.88	0.35	0.33	<i>1.22</i>
40	1.18	2.03	1.92	0.91	0.52	0.24	<i>1.31</i>
50	1.27	2.32	1.43	2.03	0.40	0.26	<i>1.49</i>
60	1.30	1.67	1.76	1.10	0.60	0.51	<i>1.28</i>
70	1.43	1.67	3.09	1.16	0.87	0.56	<i>1.64</i>
80	1.01	2.00	2.33	2.71	0.67	0.69	<i>1.74</i>
90	1.00	1.45	11.33	1.16	1.12	0.76	<i>3.21</i>
100	NA	2.88	1.17	1.17	1.40	0.83	<i>1.65</i>
200	NA	2.07	4.05	1.53	1.34	0.82	<i>2.25</i>
300	NA	NA	0.33	1.00	9.76	0.73	<i>3.70</i>
400	NA	NA	NA	NA	9.29	0.92	<i>9.29</i>
500	NA	NA	NA	NA	10.00	0.95	<i>10.00</i>
mean	1.12	2.18	2.63	1.33	2.65	0.56	<i>1.98</i>

Numbers in bold and italics were calculated without including “ $r^2 < 0.1$ ”

NA denotes that marker-pair proportion in CHB or both UIG and CHB is zero.

Table S7. Comparison of LDs of UIG and CEU for 602 AIMs

	$r^2 \geq 0.8$	$r^2 \geq 0.5$	$r^2 \geq 1/3$	$r^2 \geq 0.2$	$r^2 \geq 0.1$	$r^2 < 0.1$	mean
10	1.01	0.98	0.85	2.83	0.49	0.40	<i>1.23</i>
20	1.04	0.95	1.95	0.89	1.07	0.45	<i>1.18</i>
30	1.11	1.00	2.36	0.39	1.00	0.56	<i>1.17</i>
40	0.86	1.92	2.05	0.40	0.83	0.42	<i>1.21</i>
50	0.65	1.65	3.89	0.46	0.70	0.65	<i>1.47</i>
60	0.81	0.87	3.87	1.55	0.54	0.61	<i>1.53</i>
70	0.53	0.88	1.89	3.13	0.44	0.96	<i>1.37</i>
80	0.06	1.29	1.12	4.75	1.56	0.73	<i>1.75</i>
90	0.17	0.55	3.09	2.40	1.26	0.81	<i>1.49</i>
100	NA	0.77	2.63	8.20	1.40	0.76	<i>3.25</i>
200	NA	0.64	1.26	6.43	4.14	0.77	<i>3.12</i>
300	NA	NA	0.10	7.00	32.20	0.72	<i>13.10</i>
400	NA	NA	NA	4.00	26.00	0.91	<i>15.00</i>
500	NA	NA	NA	NA	3.75	0.96	<i>3.75</i>
mean	0.69	1.04	2.09	3.26	5.38	0.69	<i>2.49</i>

Numbers in bold and italics were calculated without including “ $r^2 < 0.1$ ”

NA denotes that marker-pair proportion in CHB or both UIG and CHB is zero.