

Supplementary Material

To accompany Han et al. *Considerable Haplotype Diversity within the 23kb Encompassing the ADH7 Gene.*

This material will be posted on Kidd Lab web site if the Journal does not have a web site for Supplementary Material.

	Polymorphic Site	Method	PCR Primers or sequence flanking site	PCR Condition [cycles]
A D H 7	S1:rs1154469 (intron 1)	Taqman ^c	CTGACATTGGATATACTATTTCCAA [C/T] TTACCCCTGCTTCCACTGACATTCT	93°C (15s), 60°C (60s) [40] ^a
	S2: rs971074 <i>Hinf</i> I (exon 6)	RFLP (<i>Hinf</i> I ^b)	A7BUE5F 5'- TTCAGGTTTCACTTTTTGACCC-3' A7BUE5R 5'-CAAATTTGGGGTCTTACC- 3'	94°C (40s), 56°C (45s), 72°C (60s) [35] ^a
	S3: rs1154456 <i>Sty</i> I (intron 6)	RFLP (<i>Sty</i> I ^e)	A7IN5DW2 5'- TATTAATTTATTGCTTAATAACTGG-3' A7IN5UP1 5'-TTCCTGTGTCTCTTACAGTG- 3'	95°C (15s), 54°C (15s), 72°C (60s) [40] ^a
	S4:rs2851011 (intron 8)	TaqMan ^d	CTGTAAAATAAGGAAATTTGTGGC [T/C] TACTTGAAATAAATAGCAGCTTTAT	93°C (15s), 60°C (60s) [40] ^a
	S5:rs284784 (intron 8)	TaqMan ^d	CCAGTTATTTATTTGCTTTCACAAC [A/C] TTGGTCTTCATCTAAAAATGGTAAA	93°C (15s), 60°C (60s) [40] ^a
	S6:rs284786 (3' downstream)	TaqMan ^d	CTTTATACGCTATATCTGTTTTGAG [A/T] TATATATCTGCAGCAGCTGGTTAAA	93°C (15s), 60°C (60s) [40] ^a
	S7:rs729147 (3' downstream)	FP	A7DWSDW1 5'- TTGAAGCGCTGAAAACATA-3' A7DWSUP1 ^e 5'- AGTGTCCAATCTATATTAT <i>TCCCT</i> -3'	94°C (30s), 55°C (30s), 72°C (60s) [25] ^a
A D H I B	rs1159918 (5' upstream)	TaqMan ^d	ATCCCTTTGCTGTGACTAATTGGG [G/T] GTCTGATTCTCTGCTGTAAAGACA	93°C (15s), 60°C (60s) [40] ^a
	rs1229984 Arg47His (exon 3)	RFLP ^f	A2FXNFOR 5'- ATTCTAAATTGTTTAATTCAAGAAG-3' A2FXNREV 5'- ACTAACACAGAATTACTGGAC-3'	95°C (30s), 56°C (30s), 72°C (60s) [35] ^a
	Rs2066701 <i>Rsa</i> I (intron 3)	RFLP ^g	A2IN3DW3 5'- ATATTTATTTTACCCTAAACTTATG-3' A2IN3UP2 5'- GAGCTAAAACATACTTTGGATAG-3'	94°C (40s), 56°C (30s), 72°C (60s) [30] ^a

Supplementary Table 1: Primers and protocols for 10 SNPs including 7 SNPs in *ADH7* and 3 SNPs in *ADH1B*

^a all cycling protocols were performed on a Perkin Elmer 9600, with an initial bold at 95°C, for 5 mins, and a final hold at 72°C, for 10 mins.

^b PCR product was digested with 4 U of *Hinf* I using the buffer that was recommended by the manufacturer.

^c PCR product was digested with 6 U of *Sty* I using the buffer that was recommended by the manufacturer.

^d The PCR final volumes for all the Taqman sites were 5 µl.

^e The mismatched base is bold and italicized

^f PCR product was digested with 6 U of *Msl* I using the buffer that was recommended by the manufacturer.

^g PCR product was digested with 8 U of *Rsa* I using the buffer that was recommended by the manufacturer.

The PCR primers for the intron 6 *Sty* I site were the same as described in Osier et al., (2002), and the PCR primers for *ADH1B* Arg47His and *Rsa* I were the same as described in Osier et al., (1999). The PCR primers for the *Hinf* I site were designed based on the primers previously reported in Buervenich et al., (2000). For the 3' downstream rs729147 site, we designed PCR primers appropriate for the FP method. The program “mfold” (Santalucia 1998) predicted a secondary structure that would likely inhibit the primer extension reaction. Therefore, we introduced an artificial mismatch in the downstream primer to disrupt the secondary structure. The upstream PCR primer was used as a detection primer for the single nucleotide base extension (SBE), giving very tight homo- and hetero- zygote genotype clusters. The TaqMan assay IDs for rs1154469, rs2851011, rs284784, rs284786, rs1159918 are C_8934019, C_16129902, C_1492617, C_714911, and C_2688471 respectively (detailed information is available in the Website <http://www.appliedbiosystems.com>).

S1: rs1154469		Residual ^a	A	A	A	G	G	A	A	A	G	A	A	A	G	A	A	G	A ^b	A	A	
S2: <i>Hinf</i> I			G	G	G	G	G	A	A	G	G	A	G	G	G	G	A	G	G	A	G	
S3: <i>Sly</i> I			G	G	C	C	C	C	C	G	C	C	G	C	G	G	C	C	C	C	G	G
S4: rs2851011			C	C	C	C	C	C	C	C	C	T	T	T	T	T	C	T	C	C	C	
S5: rs284784			C	C	C	C	C	C	C	A	A	C	C	C	C	C	A	C	C	C	C	
S6: rs284786			A	T	T	T	A	A	T	A	A	T	T	T	T	T	A	T	A	A	T	
S7: rs729147			A	A	A	A	A	A	A	A	A	G	G	G	G	A	A	G	A	A	G	
Population	2N																					
Biaka	134	0.057	0.14	0.12	0.07	0.08	0.06	0.03	0.28	0.06	0.01	0.03	0.01	0.05	0.01	0	0	0	0	0	0	
Mbuti	76	0.028	0.08	0.23	0.08	0	0.02	0.06	0.2	0	0	0.09	0	0	0	0.11	0.03	0.09	0	0	0	
Yoruba	156	0.024	0.4	0.2	0	0	0.03	0.04	0.05	0.01	0.02	0.03	0.07	0.03	0	0	0.06	0	0.04	0	0	
Ibo	96	0.02	0.47	0.15	0.03	0	0.01	0.08	0	0.07	0.02	0.04	0.03	0.02	0	0	0.03	0	0.03	0.01	0	
Hausa	78	0.136	0.24	0.17	0.03	0	0.07	0.05	0.02	0.05	0	0.05	0.08	0	0	0.02	0.04	0	0.05	0	0.01	
Chagga	90	0.039	0.24	0.21	0.01	0.06	0.04	0.05	0.04	0.1	0	0.02	0.09	0	0	0	0.07	0	0.03	0	0	
Ethiopians	60	0.027	0.02	0.05	0	0.02	0	0	0	0	0	0.41	0.24	0.1	0	0.02	0.03	0	0.01	0	0.09	
Afr American	182	0.022	0.22	0.21	0.01	0.03	0.04	0.05	0.09	0.08	0.01	0	0.06	0.02	0	0.01	0.03	0.06	0.05	0	0	
Yemenites	84	0.077	0.05	0.08	0.14	0.02	0	0	0.08	0.04	0.15	0.19	0.04	0	0	0	0.04	0.11	0	0	0	
Druze	204	0.058	0.05	0.19	0.05	0.02	0.01	0	0.03	0.1	0.18	0.15	0.12	0	0	0	0.02	0.02	0	0	0	
Samaritans	80	0	0	0.01	0	0	0	0	0.06	0.03	0.27	0.23	0.11	0	0	0	0.17	0.13	0	0	0	
Adygei	108	0	0.05	0.11	0.13	0.08	0.04	0	0.08	0.19	0.22	0.02	0.07	0	0.01	0	0	0	0	0	0	
Chuvash	82	0.019	0	0.24	0.13	0.07	0.03	0.04	0.03	0.12	0.1	0.02	0.15	0	0.02	0	0.01	0.02	0.01	0	0	
Russians	96	0.01	0.03	0.28	0.15	0.13	0.03	0.05	0.06	0.03	0.05	0	0	0.05	0.01	0	0.03	0.1	0	0	0	
Ashkenazi	162	0.006	0.04	0.24	0.11	0.11	0.02	0.01	0.01	0.05	0.1	0.13	0.14	0	0	0	0.01	0.05	0	0	0	
Finns	70	0.014	0	0.11	0.14	0.27	0	0.04	0.06	0.04	0.04	0.09	0.06	0	0.05	0	0	0.09	0	0	0	
Danes	102	0.001	0	0.14	0.13	0.14	0.02	0.01	0.09	0.13	0.06	0.05	0.1	0.01	0.02	0	0	0.04	0.06	0	0	
Irish	216	0	0.02	0.2	0.1	0.12	0.04	0.02	0.04	0.07	0.1	0.07	0.15	0.02	0	0	0	0.05	0	0	0	
Euro American	176	0.034	0.04	0.14	0.19	0.15	0.02	0	0.07	0.08	0.12	0.02	0.06	0	0	0	0	0.08	0	0	0	
Komi	94	0.075	0.06	0.21	0.1	0.12	0	0	0.08	0.07	0.1	0	0.04	0.01	0.02	0	0.04	0.04	0.06	0	0	
Khanty	100	0.053	0	0.23	0	0	0	0.17	0	0.07	0.09	0	0.04	0.01	0.19	0	0.01	0.09	0	0.06	0	
SF Chinese	108	0	0	0	0.02	0.03	0.05	0.03	0	0.4	0.05	0.01	0.15	0.07	0	0	0.05	0.14	0	0	0	
TW Chinese	94	0.012	0	0.01	0.01	0	0	0	0.03	0.42	0.03	0.04	0.07	0.09	0	0	0.03	0.26	0	0	0	
Hakka	82	0	0	0.03	0	0	0.02	0.03	0.01	0.45	0.04	0.02	0.15	0.06	0	0	0.02	0.16	0.01	0	0	
Japanese	90	0.011	0	0.01	0	0.01	0.09	0.06	0.01	0.6	0.05	0	0.04	0.01	0	0	0.02	0.09	0	0	0	
Ami	80	0.025	0	0.01	0	0	0.01	0.05	0	0.07	0.16	0	0.13	0	0	0	0.01	0.51	0	0	0.01	
Atayal	84	0.012	0	0.02	0	0.01	0	0	0	0.13	0.27	0	0.35	0	0	0	0	0.21	0	0	0	
Cambodians	50	0.08	0	0	0	0	0	0	0	0.16	0.12	0.04	0.18	0.17	0	0	0	0.16	0.06	0.04	0	
Nasioi	44	0.026	0	0.07	0	0	0	0.11	0.14	0.11	0	0.08	0.07	0	0	0	0.05	0.32	0.03	0	0	
Micronesians	68	0.03	0.12	0.03	0.03	0.02	0	0	0	0.16	0.1	0.12	0.06	0	0	0	0	0.34	0	0	0	
Yakut	102	0.012	0	0.06	0.05	0.04	0.01	0.15	0	0.16	0.04	0.04	0.29	0	0.04	0	0.01	0.11	0	0	0	
Cheyenne	110	0	0.01	0	0	0	0.04	0.2	0	0.03	0.05	0.04	0.33	0	0	0	0	0.23	0.09	0	0	
Pima Arizona	100	0	0.01	0.09	0	0	0	0.04	0	0	0.16	0	0.13	0	0	0	0.05	0.51	0.01	0	0	
Pima Mexico	104	0	0	0	0	0	0	0.24	0	0	0.14	0	0.14	0.01	0	0	0.05	0.4	0	0	0.01	
Maya	92	0	0	0.01	0	0	0.03	0.04	0	0.02	0.06	0.07	0.4	0.03	0.01	0	0.01	0.31	0	0	0	
Ticuna	130	0	0	0.02	0	0.02	0	0.06	0	0.06	0.05	0.41	0	0	0	0	0	0.37	0	0	0	
R Surui	90	0.018	0	0	0.01	0	0	0.06	0	0.01	0.01	0.03	0.6	0	0	0	0	0.28	0	0	0	
Karitiana	102	0	0	0	0	0	0	0	0	0.01	0.36	0	0.43	0.02	0.01	0	0	0.17	0	0	0	

Supplementary Table 2: Haplotype frequency for 7-SNP in 38 populations

^a Haplotypes with a frequency < 5% in all populations.

^b The AGCCCA haplotype is ancestral (i.e., each of the seven sites is represented by the ancestral allele)

The “residual” column in the tables gives the total frequency in the specific population of all haplotypes that never exceeded the criterion frequency (5%) in any population studied.

S1:rs1154469		Residual ^a	A	A	G	G	A	G	A	A ^b	A	A	A	G	A	A	A	G	A	G
S2:HinfI			G	G	G	G	A	G	G	G	A	A	G	G	G	A	G	G	G	G
S:StyI			G	C	C	C	C	C	G	C	C	C	G	G	C	C	G	C	C	C
S4:rs1159918			C	C	C	C	C	A	A	A	A	C	C	A	C	C	C	C	C	A
S5:Arg47His			G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A
S6:RsaI			C	C	T	C	C	C	C	C	C	T	T	C	T	C	T	T	C	C
POPULATION	2N																			
Biaka	132	0.041	0.114	0.02	0.036	0.078	0.168	0.041	0.214	0.132	0.156	0	0	0	0	0	0	0	0	0
Mbuti	78	0.016	0.031	0	0	0	0.236	0.113	0.376	0.08	0.124	0.024	0	0	0	0	0	0	0	0
Yoruba	154	0	0.122	0.005	0	0.031	0.049	0.034	0.558	0.068	0.12	0.006	0.007	0	0	0	0	0	0	0
Ibo	96	0.011	0.079	0.002	0	0.018	0.068	0.024	0.628	0.082	0.068	0.009	0.012	0	0	0	0	0	0	0
Hausa	78	0.053	0.113	0	0	0.038	0.051	0.066	0.433	0.054	0.137	0.016	0.025	0.014	0	0	0	0	0	0
Chagga	90	0.019	0.155	0.023	0	0.011	0	0.07	0.438	0.029	0.189	0	0.04	0	0.026	0	0	0	0	0
Ethiopians	62	0.042	0.081	0.017	0	0	0.144	0	0.301	0.024	0	0	0.019	0	0.025	0.305	0.024	0.016	0	0
Afr American	180	0.007	0.13	0.036	0	0.022	0.034	0.139	0.366	0.04	0.127	0.017	0.075	0	0.003	0	0	0.006	0	0
Yemenites	80	0.072	0.111	0	0	0.03	0.085	0.12	0	0.092	0.045	0.032	0.042	0	0	0.117	0	0.013	0.037	0.046
Druze	196	0.023	0.071	0.035	0.042	0.047	0.032	0.101	0.147	0.015	0.031	0.011	0.177	0	0.008	0.134	0	0	0	0.034
Samaritans	80	0	0	0	0.039	0	0.081	0	0	0	0	0.056	0.15	0	0	0.151	0	0.018	0.125	0.275
Adygei	106	0	0.135	0.05	0.063	0.122	0.092	0.1	0.14	0	0.012	0	0.116	0.01	0.029	0	0.033	0.033	0.045	0.021
Chuvash	78	0.025	0.12	0	0.026	0.063	0.047	0.101	0.202	0.015	0.057	0.005	0.063	0.024	0.083	0	0.103	0	0.028	0.037
Russians	96	0.088	0.148	0.033	0.04	0.098	0.034	0.172	0.168	0.021	0.03	0.051	0	0.012	0.104	0	0	0	0	0
Ashkenazi	162	0.07	0.156	0.083	0.092	0.054	0.014	0.057	0.164	0	0	0	0.066	0	0.013	0.12	0.019	0	0	0.046
Finns	70	0.029	0.113	0.032	0.027	0.127	0.139	0.208	0.093	0	0.061	0	0	0.051	0.12	0	0	0	0	0
Danes	98	0.052	0.234	0.043	0.062	0.057	0.054	0.151	0.074	0.022	0.05	0.034	0.054	0	0.112	0	0	0	0	0
Irish	182	0.013	0.164	0.081	0.069	0.111	0.016	0.133	0.193	0.031	0.064	0.021	0.105	0	0	0	0	0	0	0
Euro American	178	0.047	0.111	0.077	0.093	0.149	0	0.098	0.117	0.054	0.039	0.039	0.101	0	0.041	0.017	0	0	0	0.017
Komi	94	0.024	0.16	0.102	0.073	0.111	0.062	0.062	0.148	0.029	0.031	0.019	0.067	0.034	0.08	0	0	0	0	0
Khanty	98	0.114	0.135	0	0.008	0.051	0	0.127	0.11	0	0.157	0.019	0.09	0.189	0	0	0	0	0	0
SF Chinese	104	0.131	0.01	0	0.092	0.01	0	0.071	0.047	0.008	0.037	0	0.039	0	0	0	0.469	0.087	0	0
TW Chinese	96	0.111	0	0	0.031	0	0	0	0.093	0.01	0.02	0.04	0.029	0	0.016	0	0.378	0.26	0	0.012
Hakka	82	0.053	0	0	0	0	0	0.057	0.048	0.021	0.069	0	0.117	0	0.005	0	0.467	0.162	0	0
Japanese	94	0.076	0	0	0.013	0.021	0	0.087	0.029	0.012	0.01	0	0.063	0	0	0	0.576	0.113	0	0
Ami	80	0.018	0.012	0	0	0	0	0.111	0.076	0	0	0.037	0	0	0	0.007	0.136	0.458	0	0
Atayal	84	0	0	0	0.024	0.024	0	0	0.119	0	0	0	0	0.012	0	0	0.357	0.44	0	0
Cambodians	50	0.103	0	0.043	0.021	0.085	0	0.086	0.087	0.067	0.04	0	0	0	0.151	0	0.229	0.088	0	0
Nasioi	46	0.077	0	0	0.055	0	0	0.107	0.185	0.023	0.327	0.086	0	0	0	0	0.021	0.119	0	0
Micronesians	72	0.014	0	0	0.059	0.017	0	0.151	0.338	0	0.08	0.031	0.037	0	0.028	0	0	0.246	0	0
Yakut	98	0.034	0	0.049	0.037	0.022	0.036	0.118	0.316	0	0.131	0.027	0.047	0.023	0	0	0.134	0	0.001	0.023
Cheyenne	110	0.002	0	0.009	0	0.009	0.019	0.3	0.349	0.074	0.222	0	0.012	0	0.005	0	0	0	0	0
Pima Arizona	92	0	0	0	0	0	0	0.652	0.227	0.012	0.109	0	0	0	0	0	0	0	0	0
Pima Mexico	106	0.038	0	0	0	0	0	0.499	0.132	0.009	0.292	0	0.019	0	0	0	0	0	0	0.01
Maya	96	0	0	0.031	0	0	0	0.395	0.426	0.011	0.052	0	0.01	0.011	0	0.063	0	0	0	0
Ticuna	130	0	0	0	0	0.015	0	0.446	0.431	0	0.108	0	0	0	0	0	0	0	0	0
R Surui	84	0	0	0	0	0	0	0.298	0.607	0.012	0.083	0	0	0	0	0	0	0	0	0
Karitiana	106	0.005	0.005	0	0	0	0	0.537	0.424	0.02	0	0	0	0.01	0	0	0	0	0	0

Supplementary Table 3: Haplotype frequency for upstream 3 SNPs in *ADH7* and 3 SNPs in *ADH1B* in 38 populations

^a Haplotypes with a frequency < 8% in all populations.

^b The AGCAGC haplotype is ancestral (i.e., each of the six sites is represented by the ancestral allele)

S1:rs2851011		Residual ^a	C	C	C	C	C ^b	C	C	T	T	T	T	C	C	T	C	C	C	C	T
S2:rs284784			C	C	C	A	C	A	C	C	C	C	C	C	A	C	A	C	A	C	C
S3:rs284786			A	A	T	A	A	A	T	T	T	T	T	T	A	T	A	T	A	A	T
S4:rs729147			A	A	A	A	A	A	A	G	G	A	G	A	A	G	A	A	A	A	G
S5:rs1159918			C	C	C	C	A	A	C	A	A	C	C	C	C	C	C	A	A	C	C
S6:Arg47His			G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	A
S7:RsaI			C	T	C	C	C	C	C	C	C	C	T	T	T	C	T	C	C	C	T
POPULATION	2N																				
Biaka	136	0.032	0.014	0.038	0.342	0.005	0.157	0.062	0.209	0.003	0.077	0.051	0.01	0	0	0	0	0	0	0	0
Mbuti	74	0	0.088	0	0.168	0.001	0.061	0.026	0.332	0	0.176	0.122	0.027	0	0	0	0	0	0	0	0
Yoruba	156	0.017	0.051	0	0.099	0.033	0.464	0.051	0.135	0.017	0.111	0.014	0	0.007	0	0	0	0	0	0	0
Ibo	94	0.022	0.013	0	0.073	0.063	0.572	0.054	0.097	0	0.074	0.011	0	0.021	0	0	0	0	0	0	0
Hausa	76	0.073	0.061	0	0.038	0.087	0.347	0	0.157	0	0.131	0.044	0.013	0.03	0.018	0	0	0	0	0	0
Chagga	90	0.025	0.017	0.015	0.085	0.079	0.324	0.114	0.237	0	0.056	0	0.041	0	0.008	0	0	0	0	0	0
Ethiopians	58	0.176	0	0	0	0.017	0.017	0	0.048	0.104	0.213	0	0.059	0.016	0	0.35	0	0	0	0	0
AfrAmerican	180	0.023	0.033	0.014	0.134	0.034	0.311	0.055	0.165	0.014	0.119	0.018	0.01	0.046	0.019	0	0.006	0	0	0	0
Yemenites	82	0.079	0.059	0	0.07	0.001	0	0.073	0.066	0.165	0.059	0	0	0	0.059	0.178	0.014	0.144	0.032	0	0
Druze	206	0.079	0	0	0.085	0.036	0.03	0.123	0.098	0.061	0.032	0.01	0.051	0.062	0.104	0.142	0	0.017	0.034	0.034	0
Samaritans	78	0	0	0	0	0	0	0	0	0.081	0	0	0.138	0.077	0.025	0.264	0	0	0.269	0.128	0.016
Adygei	106	0.03	0.015	0.06	0.247	0.142	0.011	0.079	0.07	0	0.094	0	0	0.058	0.091	0.009	0.064	0	0.029	0	0
Chuvash	78	0.05	0	0.028	0.208	0.038	0.055	0.026	0.152	0	0.158	0	0	0.087	0.04	0	0.09	0.033	0.019	0	0.016
Russians	92	0.09	0	0.049	0.272	0.015	0.047	0.041	0.223	0.04	0.069	0	0	0.122	0.033	0	0	0	0	0	0
Ashkenazi	162	0.074	0	0.015	0.199	0.024	0.025	0.047	0.091	0.076	0.063	0	0.033	0.104	0.027	0.132	0	0.017	0.051	0.022	0
Finns	70	0.068	0	0	0.284	0.086	0.012	0	0.138	0.033	0.226	0	0	0.154	0	0	0	0	0	0	0
Danes	98	0.05	0.01	0.04	0.26	0.048	0.013	0.079	0.065	0.048	0.153	0	0	0.173	0.06	0	0	0	0	0	0
Irish	180	0.013	0.053	0.005	0.18	0.021	0.026	0.05	0.191	0.105	0.154	0	0.017	0.096	0.09	0	0	0	0	0	0
Euro American	178	0.046	0.012	0.016	0.256	0.074	0.033	0.043	0.139	0.023	0.072	0	0.067	0.111	0.075	0.012	0.004	0	0.017	0	0
Komi	94	0.034	0.002	0.088	0.22	0.171	0.031	0	0.243	0.066	0	0	0.03	0.043	0.071	0	0	0	0	0	0
Khanty	96	0.029	0	0	0.132	0.062	0.246	0.015	0.066	0.011	0.322	0	0	0.013	0.104	0	0	0	0	0	0
SF Chinese	102	0.024	0	0.018	0	0	0.019	0.025	0.022	0.02	0.101	0	0.046	0.025	0.038	0	0.457	0	0	0	0.206
TW Chinese	96	0.045	0	0	0	0	0.014	0.013	0	0.098	0	0.078	0.026	0	0	0.464	0	0	0	0	0.261
Hakka	82	0.037	0	0	0	0	0.036	0	0.024	0	0.134	0	0.079	0	0.043	0	0.469	0	0	0	0.177
Japanese	90	0.092	0	0.023	0	0	0.064	0.015	0	0.022	0.055	0	0.02	0.01	0.024	0	0.628	0	0	0	0.047
Ami	80	0.088	0	0.037	0	0.013	0	0.025	0	0	0.15	0	0	0	0.097	0.192	0	0	0.008	0.391	
Atayal	84	0.046	0	0	0	0	0	0	0.022	0.119	0.012	0.023	0	0	0.003	0.381	0	0	0	0.392	
Cambodians	50	0.06	0	0.067	0	0.09	0.033	0.032	0.02	0	0.176	0	0.16	0	0.023	0	0.136	0	0	0	0.204
Nasioi	44	0	0	0	0	0	0.136	0.121	0.205	0	0.182	0	0.151	0	0	0	0	0	0.038	0	0.167
Micronesians	68	0.015	0	0.007	0	0.015	0.11	0.189	0.015	0	0.303	0	0.038	0.058	0.076	0	0	0	0	0	0.174
Yakut	102	0.004	0.01	0.011	0.097	0	0.14	0.016	0.046	0.011	0.389	0	0.081	0	0.019	0	0.154	0	0.022	0	0
Cheyenne	110	0	0	0.018	0	0.017	0.317	0.055	0	0.019	0.572	0	0	0	0	0	0	0	0	0	0
Pima Arizona	98	0	0	0	0	0	0.061	0.204	0.092	0	0.643	0	0	0	0	0	0	0	0	0	0
Pima Mexico	106	0.047	0	0	0	0	0.245	0.164	0	0	0.514	0	0.014	0	0.005	0	0	0	0.009	0	0
Maya	94	0.004	0	0	0	0	0.085	0.081	0	0.021	0.738	0	0	0.011	0	0.06	0	0	0	0	0
Ticuna	130	0	0	0	0.015	0	0.015	0.062	0.085	0	0.823	0	0	0	0	0	0	0	0	0	0
R Surui	84	0	0	0	0	0	0.06	0.024	0.006	0	0.905	0.006	0	0	0	0	0	0	0	0	0
Karitiana	104	0	0.005	0	0	0.005	0	0.375	0	0	0.615	0	0	0	0	0	0	0	0	0	0

Supplementary Table 4: Haplotype frequencies for downstream 4 SNPs in *ADH7* and 3 SNPs in *ADH1B* in 38 populations

^a Haplotypes with a frequency < 8% in all populations.

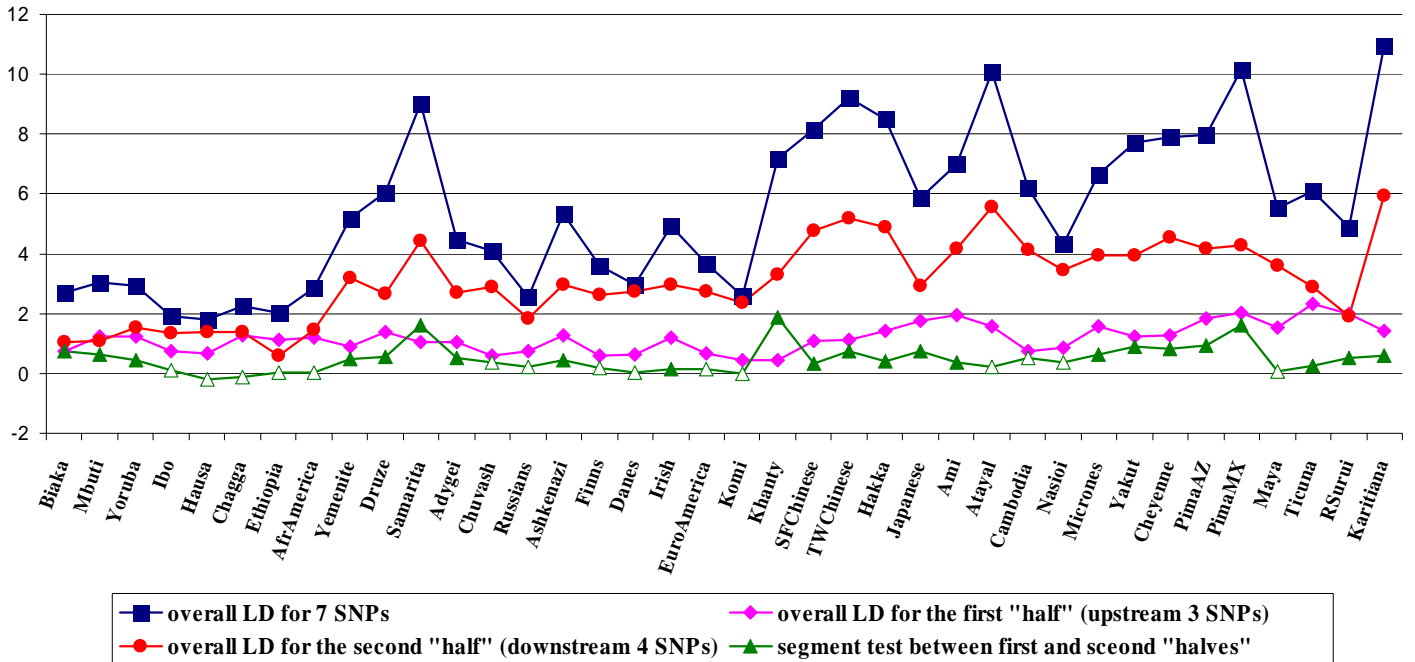
^b The CCAAAGC haplotype is ancestral (i.e., each of the seven sites is represented by the ancestral allele)

Population	Segment Test for 3 Upstream SNPs in <i>ADH7</i> and 3 SNPs in <i>ADH1B</i>		Segment Test for 4 downstream SNPs in <i>ADH7</i> and 3 SNPs in <i>ADH1B</i>	
	Xi	P value	Xi	P value
Biaka	0.15	0.054	0.51	0
Mbuti	0.45	0.001	0.16	0.095
Yoruba	-0.01	0.53	0.16	0.017
Ibo	0.01	0.366	0.28	0.009
Hausa	0.11	0.221	0.36	0.031
Chagga	0.1	0.13	0.2	0.058
Ethiopians	1.69	0	-0.04	0.539
Afri American	-0.03	0.681	0.09	0.084
Yemenite	-0.12	0.685	0.31	0.083
Druze	0.43	0	0.46	0
Samaritans	1.41	0	1.83	0
Adygei	0.3	0.021	0.45	0.003
Chuvash	0.12	0.254	0.51	0.009
Russians	0.27	0.038	0.16	0.138
Ashkenazi	0.67	0	0.4	0
Finns	0.4	0.025	0.32	0.03
Danes	0.07	0.227	0.23	0.035
Irish	-0.02	0.561	0.15	0.023
Euro American	-0.02	0.562	-0.01	0.539
Komi	0.07	0.231	0.36	0.003
Khanty	0.28	0.025	0.97	0
SF Chinese	0.27	0.015	0.17	0.03
TW Chinese	0.01	0.399	0.28	0.004
Hakka	0.13	0.094	0.3	0.003
Japanese	0.14	0.096	0.31	0.014
Ami	0.5	0.001	0.33	0.016
Atayal	0.19	0.023	0.13	0.094
Cambodians	0.31	0.096	0.44	0.06
Nasioi	0.17	0.227	0.26	0.106
Micronesians	0.54	0.005	0.28	0.069
Yakut	0.2	0.075	1.46	0
Cheyenne	-0.11	0.986	-0.05	0.972
Pima Arizona	-9.99	0	nc	nc
Pima Mexico	0.12	0.082	-0.04	0.641
Maya	0.71	0	0.12	0.079
Ticuna	0.03	0.191	0.08	0.065
R Surui	nc	nc	nc	nc
Karitiana	nc	nc	nc	nc

nc: not calculable

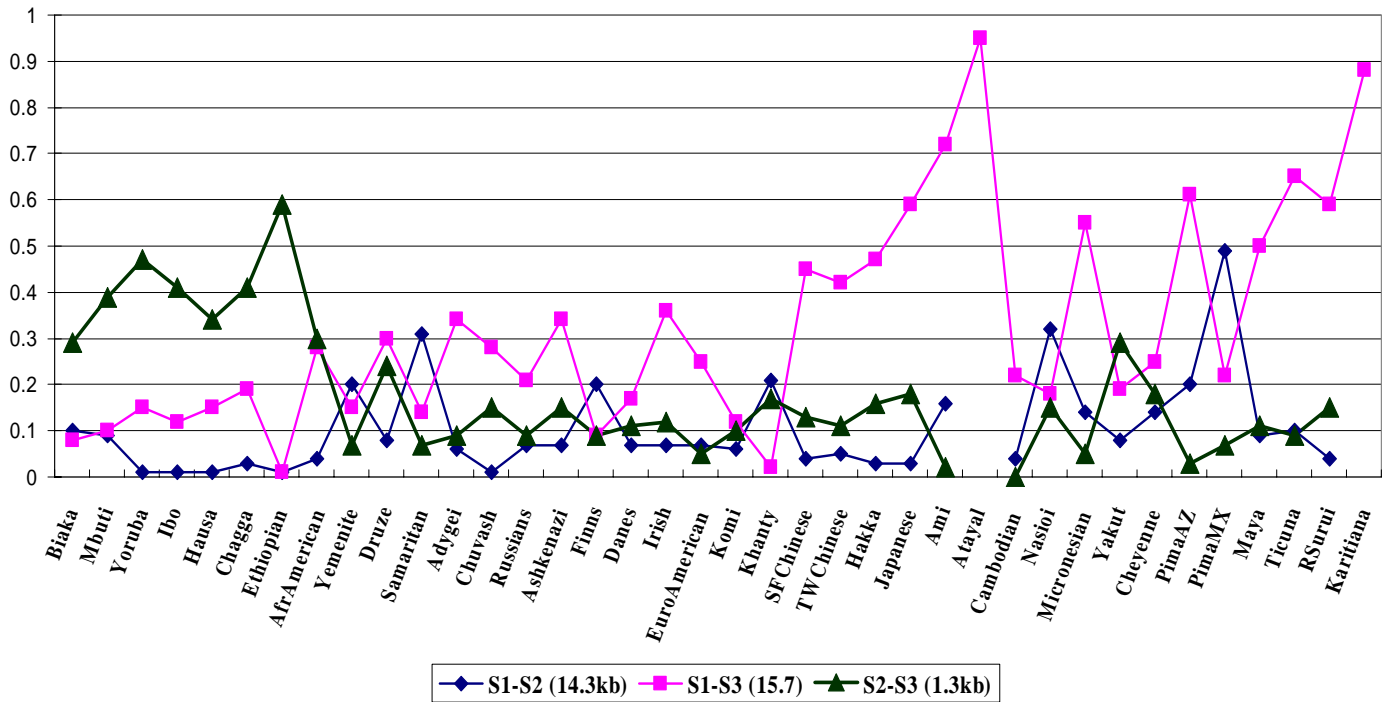
Supplementary Table 5: Segment Tests for 3 Upstream and 4 Downstream SNPs in *ADH7* with 3 SNPs in *ADH1B*

Overall LDs and Segment Test



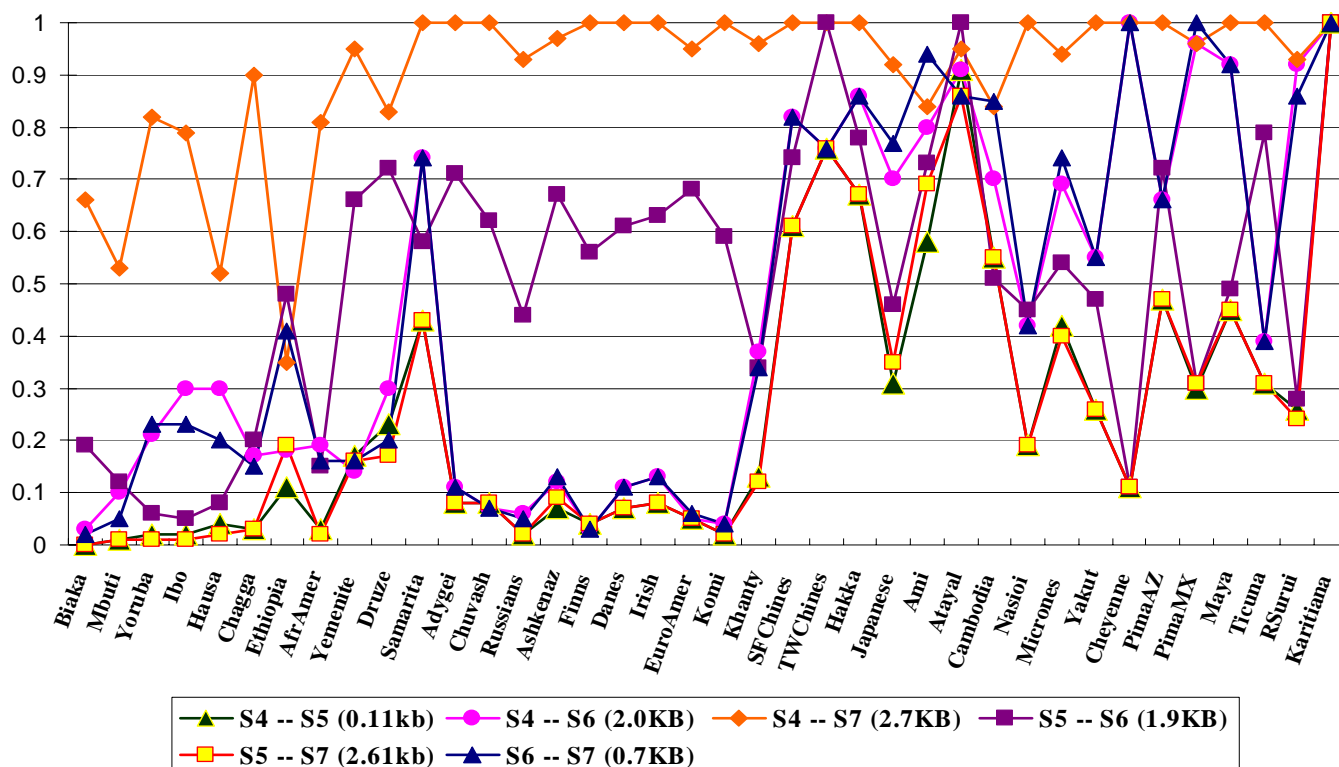
Supplementary Figure 1: The values of overall linkage disequilibrium for the seven SNPs in *ADH7*, for the first segment (site 1– site 3), for the second segment (site 4 – site 7), and the segment test between the first group and the second group. The LD values are the ξ coefficient (Zhao et al., 1999). The population names are listed at the bottom of the graph by geographic regions. The triangles represent the ξ coefficient of segment test of LD between the first group and the second group; the open triangles represent the LD values for those populations are not statistically significant ($p \geq 0.01$); all other values are significant at $p < 0.01$. LDs for the three overall tests are all statistically significant at $p \leq 0.001$

3 Pairwise LDs of S1, S2 and S3



Supplementary Figure 2: The pairwise LD values (Δ^2) for the three markers in the first “half” of *ADH7*, S1 through S3 as in Figure 1. The order of the populations is the same as in other figures. D' values were also calculated but most of them had values of 1.0, providing no information on relative levels of LD.

6 Pairwise LDs for Sites 4,5,6,7



Supplementary Figure 3: The pairwise LD values (Δ^2) for the four markers in the second “half” of *ADH7*, S4 – S7 as in Figure 1. The order of the populations is the same as in other figures. Difficult to see in the figure are the essentially identical patterns for S4 – S5 and S5 – S7, and for S4 – S6 and S6 – S7, not surprising given the nearly identical allele frequencies at site 4 and 7 (see Figure 2 and Table 2 in paper). D' values were also calculated but most of them had values of 1.0, providing no information on relative levels of LD.

Discussion of Haplotype Frequencies.

For the upstream “half” of *ADH7*, AGG is the globally most frequent in all 38 populations, ranging from 15.0% (Samaritans) to 71.8% (Ibo). The second most frequent haplotype is GGC, which is very common in non-African populations ranging from 16.5% (African American) to 71.3% (Ami), but varies within Africa. The haplotype AAC is also common in all the populations except East Asians and some Native Americans (Pima Arizona: 9.8%, R. Surui: 8.9%, and Karitiana: 0.0%). The haplotype composed of the ancestral alleles at the upstream three sites (AGC) is rare in the Pacific, Siberia, and the Americas, infrequent but present in Africans, African Americans and East Asians (except Cambodians) and more frequent in SW Asians, Europeans and NW Asians.

For the downstream “half” of *ADH7*, the ancestral haplotype for the downstream four sites—that is, CCAA—is most frequent in Africa (except Ethiopians: 3.3%) and African Americans, frequent in Northwest Asia, the Pacific, and Siberia, and infrequent to absent in Europe (except Russians: 10.9%), East Asia (except Japanese: 14.2%, and Cambodians: 10.0%), Southwest Asia (except Samaritans: 13.0%), North America (except Cheyenne: 33.6%, and Mexican Pima: 24.5%), and South America. The haplotype CCTA is common in nearly all the populations except East Asian and Native American; it is especially very frequent in Europe, ranging from 40.4% to 61.5%. The haplotype CAAA is infrequent in Africa and the Americas but common in other populations. It is particularly very common in East Asia with frequencies greater than 40% in five out of the seven East Asian populations that we studied. Another major haplotype, TCTG, is common in all the populations except the Biaka (9.2%) and Ibo

(8.3%), and this haplotype is especially very frequent in Native Americans, ranging from 55.7% to 91.1%.

In general, the 7-site haplotypes are individually uncommon in all populations. Some of the exceptions are the haplotype AGGCCAA at more than 22% frequency in African, including African American, populations; AGGCCTA at more than 21% frequency in some African including African American, European, and Northwest Asian populations; and AGGCAAA at very high frequency (more than 40%) in four of seven East Asian populations. Although the above three haplotypes have different alleles at the downstream four sites, they are all “AGG” at the upstream three sites. The haplotypes AGGTCTG and GGCTCTG are remarkably frequent in most Native Americans, and have more than 10% frequency in most East Asians. Both of the above two haplotypes have the same alleles “TCTG” at the downstream four sites.

Discussion of Haplotype frequency variation around the world

The frequency range of allele “A” of S5 rs284784 is from 0.03 to 0.20 in African populations, and the single haplotype CAAA which is the only haplotype with S5 rs284784-A among the six common haplotypes is not frequent in Africans, but has remarkably high frequency in East Asians. Both the allele frequency and haplotype frequency pattern indicate that the rs284784 (S5) polymorphism is relatively young. The haplotype frequency pattern of TCTG fits the “out-of-Africa” theory: rare in Africans, more frequent in Europeans, and higher still in East Asians and Native Americans, which suggests that TCTG is a new haplotype that is recently generated.

Reference Citations

Buervenich S, Sydow O, Carmine A, Zhang Z, Anvret M, Olson L (2000) Alcohol dehydrogenase alleles in Parkinson's disease. *Mov Disord* 15:813-8

SantaLucia J, Jr. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci U S A* 95:1460-5

Zhao H, Pakstis AJ, Kidd JR, Kidd KK (1999) Assessing linkage disequilibrium in a complex genetic system. I. Overall deviation from random association. *Ann Hum Genet* 63 (Pt 2):167-79

Osier M, Pakstis AJ, Kidd JR, Lee JF, Yin SJ, Ko HC, Edenberg HJ, Lu RB, Kidd KK (1999) Linkage disequilibrium at the ADH2 and ADH3 loci and risk of alcoholism. *Am J Hum Genet* 64:1147-57

Osier MV, Pakstis AJ, Soodyall H, Comas D, Goldman D, Odunsi A, Okonofua F, Parnas J, Schulz LO, Bertranpetit J, Bonne-Tamir B, Lu RB, Kidd JR, Kidd KK (2002) A global perspective on genetic variation at the ADH genes reveals unusual patterns of linkage disequilibrium and diversity. *Am J Hum Genet* 71:84-99