

Supplementary Table 2: Frequency of *APOBEC3B* deletion grouped by broad geographic location

Geographic Region	Allele Counts		Deletion Frequency
	Deletion	Insertion	
Oceania	65	5	0.9286
East Asia	174	298	0.3686
Central South Asia	55	333	0.1418
Middle East	26	314	0.0765
Europe	20	288	0.0649
Africa	2	220	0.0090
America	102	112	0.4766
Total	444	1570	0.2205

Supplementary Table 3: HGDP Genotype Frequencies

Population	Geographic Location	Genotype Counts			HWE <sup>a</sup>
		D/D	D/I	I/I	
Melanesian	Bougainville	16	2	1	0.160
Papuan	New Guinea	15	1	0	1.000
Cambodian	Cambodia	2	8	1	0.257
Dai	China	3	7	0	0.220
Daur	China	2	5	3	1.000
Han	China	0	28	16	0.002
Hezhen	China	0	2	7	1.000
Japanese	Japan	1	12	15	0.646
Lahu	China	0	6	4	0.480
Miao	China	1	6	3	1.000
Mongola	China	0	8	2	0.173
Naxi	China	2	4	4	0.573
Oroqen	China	0	6	4	0.480
She	China	2	7	1	0.520
Tu	China	1	6	3	1.000
Tujia	China	2	4	4	0.573
Xibo	China	2	2	5	0.167
Yakut	Siberia	2	13	10	0.663
Yi	China	1	8	1	0.200
Balochi	Pakistan	0	8	16	1.000
Brahui	Pakistan	0	5	20	1.000
Burusho	Pakistan	1	9	15	1.000
Hazara	Pakistan	0	8	14	1.000
Kalash	Pakistan	0	3	21	1.000
Makrani	Pakistan	0	5	15	1.000
Pathan	Pakistan	2	6	13	0.540
Sindhi	Pakistan	0	4	19	1.000
Uygur	China	0	1	9	1.000
Bedouin	Israel-Negev	0	5	42	1.000
Druze	Israel-Carmel	1	5	39	0.223
Mozabite	Algeria-Mزاب	0	2	27	1.000
Palestinian	Israel-Central	1	10	38	0.535
Adygei	Russia-Caucasus	0	4	13	1.000
Basque	France	0	3	21	1.000
French	France	0	1	23	1.000
Italian	Italy-Bergamo	1	0	12	0.040
Orcadian	Orkney Islands	0	3	13	1.000
Russian	Russia	0	6	19	1.000
Sardinian	Italy	0	1	27	1.000
Tuscan	Italy	0	0	7	-
Bantu	Kenya	0	0	12	-

Kenya					
Biaka Pygmy	Central African Republic	0	0	31	-
Mandenka	Senegal	0	2	21	1.000
Mbuti Pygmy	Congo	0	0	15	-
San	Namibia	0	0	6	-
Yoruba	Nigeria	0	0	24	-
Colombian	Colombia	2	8	3	0.602
Karitiana	Brazil	8	11	5	0.698
Maya	Mexico	1	9	15	1.000
Pima	Mexico	6	7	11	0.083
Surui	Brazil	13	7	1	1.000

The number of individuals genotyped as homozygous for the deletion (D/D), hemizygous (I/D), and homozygous for the insertion (I/I) are given. Homozygous deletion status was confirmed by two independent breakpoint assays for the insertion allele.

<sup>a</sup>Probability of Hardy-Weinberg equilibrium

Supplementary Table 5: HapMap Genotype Frequencies.

	Genotype Counts			
Population	D/D	D/I	I/I	HWE
YRI	0	3	57	1
CEU	0	8	52	1
CHB	7	19	19	0.5323
JPT	7	21	17	1
JPT+CHB	14	40	36	0.6544

Columns as indicated in Supplementary table 4.

Supplemental Table 6:  
HapMap Phase II SNPs used in haplotype network analysis.

BD35 Coordinate	SNP ID
37673148	rs5750716
37673559	rs5757401
37673846	rs6001337
37674117	rs5995640
37674276	rs6001338
37674506	rs6001339
37674722	rs17370615
37675450	rs5995641
37675537	rs5995642
37675966	rs17298479
37676118	rs9611054
37677672	rs12157810
37713920	rs2019907
37716133	rs6001361
37716696	rs738308
37716750	rs11089916
37716796	rs2142833
37716980	rs6001363
37718597	rs8140586
37720818	rs10470278
37721871	rs9611073

### Supplemental Table 7: Observed Haplotypes

I/D indicates presence of insertion or deletion allele. SNP positions as given in Supplemental Table 6

Haplotype Number	Sequence
1	CTAGTCGCCGGCIGCTAGCGC
2	CTAGTCGCCGGCIGCTGGCGC
3	CTAGTCGCCGGCIGTTGGCGC
4	CTAGTCGTTGGAIACTGATGC
5	CTAGTCGTTGGAIGCTAGCGC
6	CTAGTCGTTGGAIGCTGGCGC
7	CTGACCGCCGGAIACTGATGC
8	CTGACCGCCGGAIGCTGGCGC
9	CTGACCGCCGGCIGCTGGCGC
10	CTGACCGTTGGAIACTGATGC
11	CTGACCGTTGGAIGCTAGCGC
12	CTGACCGTTGGAIGCTGGCGC
13	CTGACTGCCGGCIGCTGGCGC
14	CTGGCCGCCGGAIGCTAGCGC
15	CTGGCCGCCGGAIGCTGGCGC
16	CTGGTCACCGGAIACCGATGC
17	CTGGTCACCGGAIACCTGATGC
18	CTGGTCACCGGAIAGCTAGCGC
19	CTGGTCACCGGAIAGCTGGCGC
20	CTGGTCACCGGCACTGATGC
21	CTGGTCACCGGCIGCTGGCGC
22	CTGGTCACTGGAIGCTGGCGC
23	CTGGTCGCCAGAIACCGATGC
24	CTGGTCGCCAGAIACCTGATGC
25	CTGGTCGCCAGAIAGCTGGCGC
26	CTGGTCGCCAGAIAGCTAGCGC
27	CTGGTCGCCAGAIAGCTGGCGC
28	CTGGTCGCCGGADACTGATGC
29	CTGGTCGCCGGADATTGATGC
30	CTGGTCGCCGGAIACCGATGC
31	CTGGTCGCCGGAIACCTGATGC
32	CTGGTCGCCGGAIAGCTAGCGC
33	CTGGTCGCCGGAIAGCTGGCGC
34	CTGGTCGCCGGCDGCTGGCGC
35	CTGGTCGCCGGCACTGATGC
36	CTGGTCGCCGGCACTGATGC
37	CTGGTCGCCGGCIGCTAGCGC
38	CTGGTCGCCGGCIGCTGATGC
39	CTGGTCGCCGGCIGCTGGCGC
40	CTGGTCGCTGGAIACTGATGC
41	CTGGTCGCTGGAIGCTAGCGC
42	CTGGTCGCTGGAIGCTGGCGC
43	CTGGTCGCTGGCACTGATGC
44	CTGGTCGCTGGCIGCTAGCGC
45	CTGGTCGCTGGCIGCTGGCGC
46	CTGGTCGCTGGCIGTTGGCGC
47	CTGGTCGTTGGAIACTGATGC
48	CTGGTCGTTGGAIGCTAGCGC
49	CTGGTCGTTGGAIGCTGGCGC