



Supporting Online Material for

Positive Natural Selection in the Human Lineage

P. C. Sabeti, S. F. Schaffner,* B. Fry, J. Lohmueller, P. Varilly,
O. Shamovsky, A. Palma, T. S. Mikkelsen, D. Altshuler, E. S. Lander

*To whom correspondence should be addressed. E-mail: sfs@broad.mit.edu

Published 16 June 2006, *Science* **312**, 1641 (2006)
DOI: 10.1126/science.1124309

This PDF file includes:

Figs. S1 and S2
Tables S1 to S4

Figure S1: Decay of detection power with increasing time since the selective sweep for 3 population genetic signatures for selection

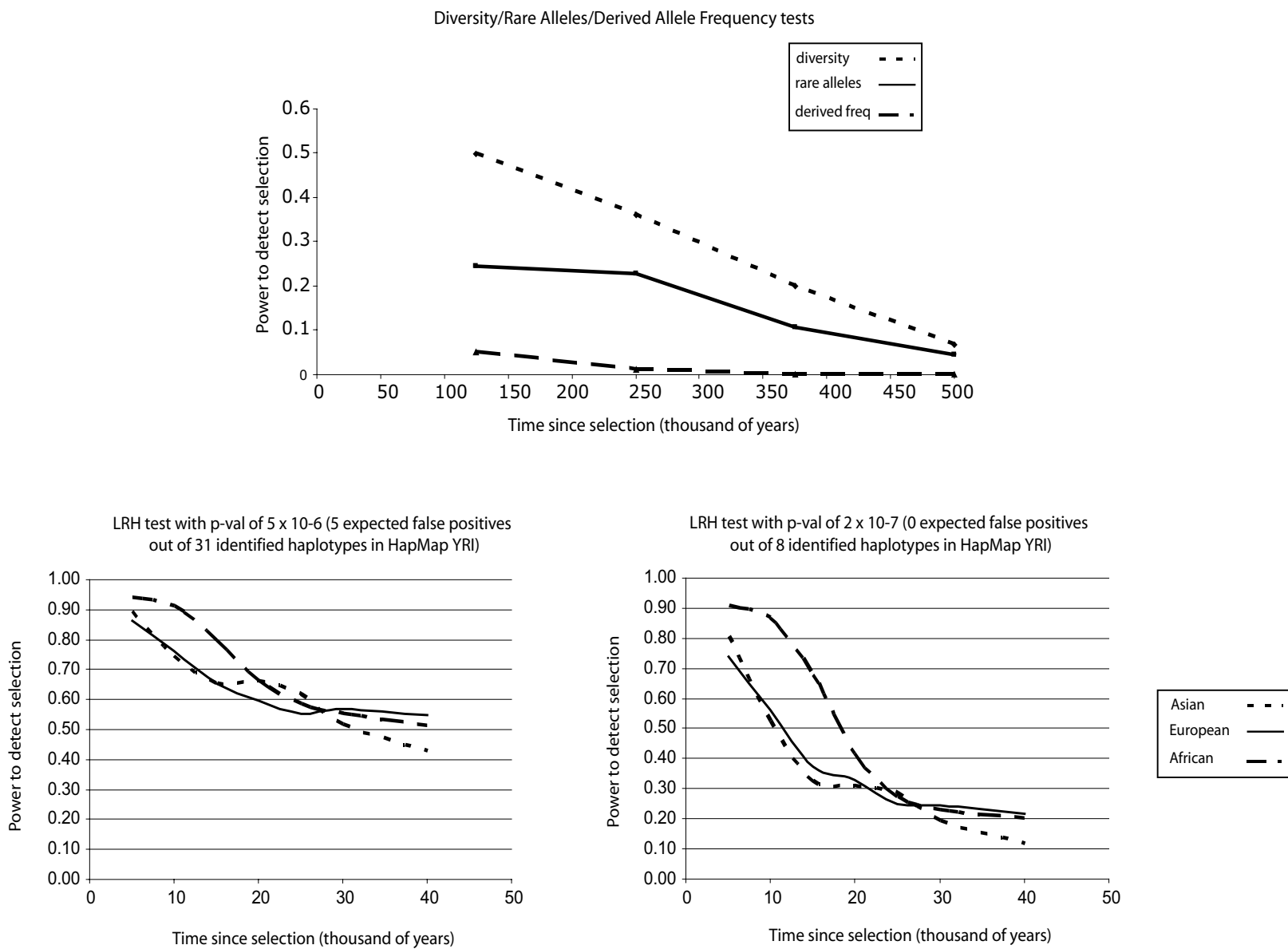


Figure S2: The previous candidates for selection identified by limited empirical data (red triangles) in comparison to the newly available genome wide empirical data sets (gray diamonds). This is an extension of results presented in Figure 7.

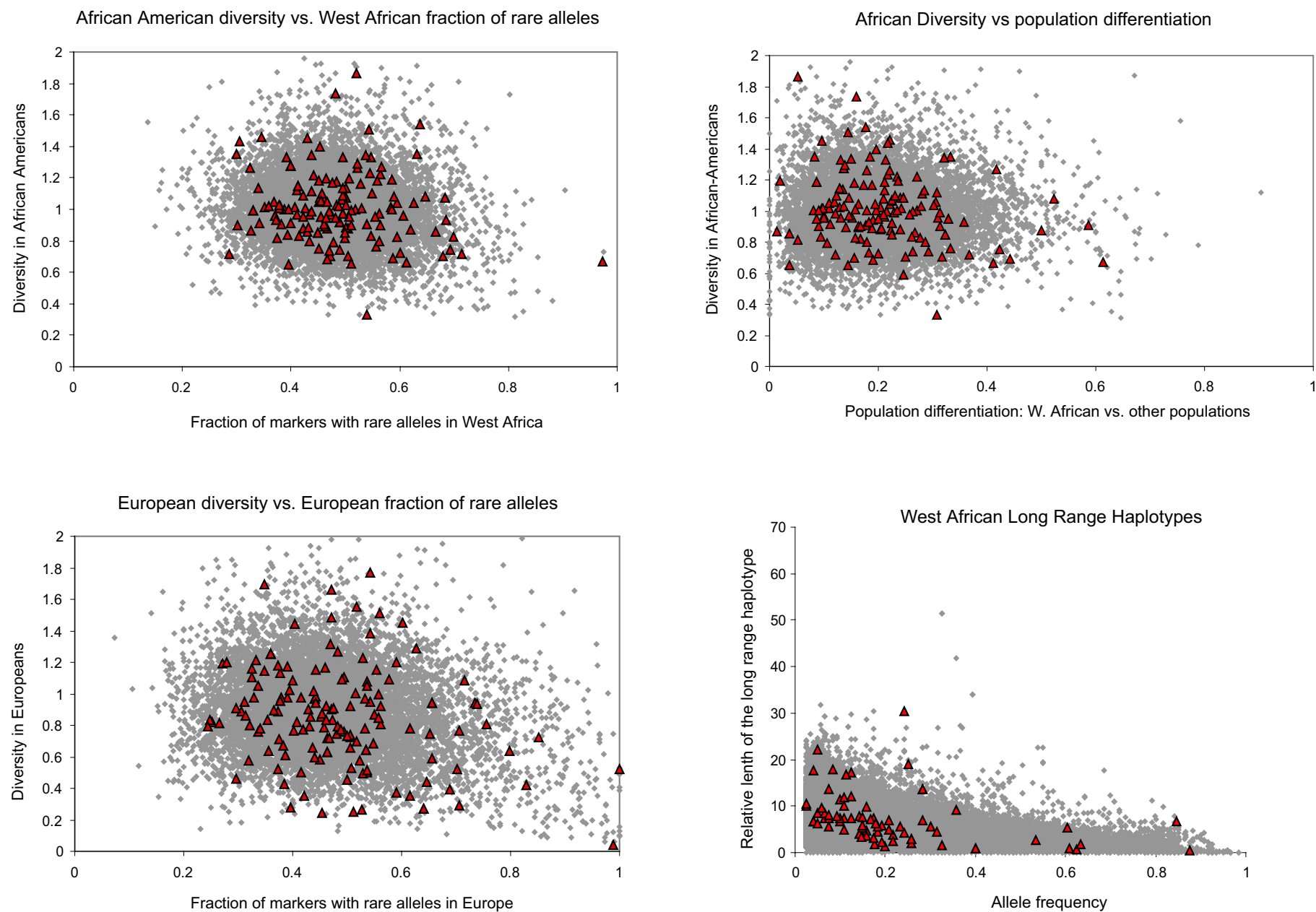


Table S1. Review of the previous reports of selection in the human lineage. We report only genes where selection was identified specifically in the human lineage since the divergence of humans and chimpanzees. The gene is given along with the chromosomal position based on the Human Genome Build 16, the signal of selection, a brief description of gene function, the type of selection proposed, the tests used with significant tests in bold, and the references.

Number	Gene	Chromosomal position (HG16)	Signal(s)	Brief function	Type of selection	Tests	References
1	ABCC1	chr16:16009872-16201711	long haplotype	one of family of organic anion transporters, can protect cell against drugs, also linked to Strep pneumo resistance in mouse models	Positive	Fst, LRH , MAF threshold	E. C. Walsh et al., Hum Genet, 1 (Dec 14, 2005).
2	ABO	chr9:131406607-131426675	diversity (high)	A & B code for glycosyltransferases, O null allele	Balancing	Tajima's D , Fu and Li's F , Fu and Li's D , Fu and Li's F , Fay and Wu's H , Fst	F. B. Livingstone, Hum Biol 32, 17 (Feb, 1960). N. Saitou, F. Yamamoto, Mol Biol Evol 14, 399 (Apr, 1997). J. M. Akey et al., PLoS Biol 2, e286 (Oct, 2004).
3	ACE2	chrX:14940581-14980563	diversity (high)	protein metabolism and modification, associated with hypertension	Balancing	Tajima's D , Fu and Li's D , Fu and Li's F , Fay and Wu's H	J. M. Akey et al., PLoS Biol 2, e286 (Oct, 2004).
4	ACTN3	chr11:66089750-66106157	diversity (high)	skeletal muscle alpha-actinin in fast (type2) fibers	Balancing	HW equilibrium	N. Yang et al., Am J Hum Genet 73, 627 (Sep, 2003).
5	ADAM2	chr8:39695889-39713143	function-altering mutations	heterodimeric sperm protein, potential role in fertilization	Positive, Purifying	dN/dS, PAML	A. Civetta, Mol Biol Evol 20, 21 (Jan, 2003).
6	ADH	chr20:3060203-3061058	long haplotype, population differentiation	alcohol metabolism, associated with alcoholism	Positive	LD , Fst	M. V. Osier et al., Am J Hum Genet 71, 84 (Jul, 2002).
7	AGT	chr1:227873893-227881601	derived allele frequency	salt regulation pathway, associated with hypertension	Positive	Fst, LD, Tajima's D , Fu and Li's D , Fu and Li's F , Fay and Wu's H , correlation of allele frequency with latitude	T. Nakajima et al., Am J Hum Genet 74, 898 (May, 2004). E. E. Thompson et al., Am J Hum Genet 75, 1059 (Dec, 2004).
8	ALDH2	chr12:110616298-110659735	population differentiation	alcohol metabolism, associated with alcoholism	Positive	Fst , LD	H. Oota et al., Ann Hum Genet 68, 93 (Mar, 2004).
9	APCS	chr1:156774407-156775450	long haplotype	serum amyloid P component gene	Positive	Fst, LRH , MAF threshold	E. C. Walsh et al., Hum Genet, 1 (Dec 14, 2005).

Table S1.

Number	Gene	Chromosomal position (HG16)	Signal(s)	Brief function	Type of selection	Tests	References
10	APOBEC3G	chr22:37716174-37726786	function-altering mutations	broad antiviral enzyme	Positive	dN/dS	J. Zhang, D. M. Webb, Hum Mol Genet 13, 1785 (Aug 15, 2004).
11	ASPM	chr1:193785110-193847224	function-altering mutations, long haplotype, population differentiation	controls human brain size, mutations lead to primary microcephaly	Positive	Ka/Ks, dN/dS , Tajima's D, Fu and Li's F, Fay and Wu's H, MK, LD, Fst	N. Kouprina et al., PLoS Biol 2, E126 (May, 2004). N. Mekel-Bobrov et al., Science 309, 1720 (Sep 9, 2005). J. Zhang, Genetics 165, 2063 (Dec, 2003).
12	ATP6	mtDNA	function-altering mutations	many genes involved in metabolism and other functions	Positive	Ka/Ks	J. L. Elson, D. M. Turnbull, N. Howell, Am J Hum Genet 74, 229 (Feb, 2004). D. Mishmar et al., Proc Natl Acad Sci U S A 100, 171 (Jan 7, 2003).
13	BRCA1	chr17:38450366-38530994	function-altering mutations, long haplotype	DNA repair; mutations cause risk of breast cancer	Positive	LD, Ka/Ks , Hardy-Weinberg	M. A. Fleming, J. D. Potter, C. J. Ramirez, G. K. Ostrander, E. A. Ostrander, Proc Natl Acad Sci U S A 100, 1151 (Feb 4, 2003). G. A. Huttley et al., Nat Genet 25, 410 (Aug, 2000).
14	C6	chr5:41187836-41259113	diversity (high)	membranolytic activity important for defense against invading microorganisms	Balancing	LD, Fu and Li's D, Fu and Li's F, Tajima's D, HKA, Fst	M. Soejima et al., Ann Hum Genet 69, 239 (May, 2005).
15	CAPN10	chr2:241846923-241859260	population differentiation, long haplotype	associated with type 2 diabetes	Positive, Balancing	Fst, LD , Tajima's D, MK	S. M. Fullerton et al., Am J Hum Genet 70, 1096 (May, 2002). J. Vander Molen et al., Am J Hum Genet 76, 548 (Apr, 2005).
16	CCR5	chr3:46372224-46378283	function-altering mutations, long haplotype	protection from recent epidemic such as bubonic plague	Positive, Balancing	LD, dN/dS, LRH	M. J. Bamshad et al., Proc Natl Acad Sci U S A 99, 10539 (Aug 6, 2002). M. Carrington, M. Dean, M. P. Martin, S. J. O'Brien, Hum Mol Genet 8, 1939 (1999). P. C. Sabeti et al., PLoS Biol 3, e378 (Nov 1, 2005). J. C. Stephens et al., Am J Hum Genet 62, 1507 (Jun, 1998). S. Wooding et al., Am J Hum Genet 76, 291 (Feb, 2005).
17	CD40L	chrX:134435969-134448155	long haplotype	protection from P.falciparum malaria	Positive	LRH	P. C. Sabeti et al., Nature 419, 832 (Oct 24, 2002).
18	CD45/ PTPRC	chr1:195896481-195950589	function-altering mutations	transmembrane protein tyrosine phosphate	Positive	dN/dS	L. C. Filip, N. I. Mundy, Mol Biol Evol 21, 1504 (Aug, 2004).
19	CD59	chr11:33688866-33722333	function-altering mutations	terminal complement inhibitor	Positive	Ka/Ks	N. Osada et al., Genomics 79, 657 (May, 2002).
20	CDY	chrY:18935329-18937289	function-altering mutations	a chromodomain implicated in chromatin binding, and a catalytic domain involved in acylation reactions	Positive	Ka/Ks	S. Dorus, S. L. Gilbert, M. L. Forster, R. J. Barndt, B. T. Lahn, Hum Mol Genet 12, 1643 (Jul 15, 2003).
21	CGB	chr19:54217938-54219405	function-altering mutations	glycoprotein critical as a signal in establishing pregnancy in humans and some other primates	Positive	dN/dS	G. A. Maston, M. Ruvolo, Mol Biol Evol 19, 320 (Mar, 2002).

Table S1.

Number	Gene	Chromosomal position (HG16)	Signal(s)	Brief function	Type of selection	Tests	References
22	COX7C	chr5:85949539-85952337	function-altering mutations	a multisubunit enzyme complex that catalyzes final electron transfer through mitochondrial inner membrane	Positive	Ka/Ks	N. Osada et al., Genomics 79, 657 (May, 2002).
23	COX8	chr11:63498654-63500590	function-altering mutations	a multisubunit enzyme complex that catalyzes final electron transfer through mitochondrial inner membrane	Positive	Ka/Ks	N. Osada et al., Genomics 79, 657 (May, 2002).
24	CYP1A2	chr15:72757888-72764754	derived allele frequency	drug transport and metabolism	Positive	Tajima's D, Fay and Wu's H , Fu and Li's D, Fu and Li's F, HKA	S. P. Wooding et al., Am J Hum Genet 71, 528 (Sep, 2002).
25	CYP3A4	chr7:98965979-98993166	diversity (low), derived allele frequency, long haplotype, population differentiation	involved in metabolism of exogenous substances	Positive	Tajima's D, Fst, Fay and Wu's H, Hudson's haplotype test, HKA	E. E. Thompson et al., Am J Hum Genet 75, 1059 (Dec, 2004).
26	CYP3A5	chr7:98857219-98889018	diversity (low), long haplotype, population differentiation	involved in metabolism of exogenous substances, mutation associated with increased blood pressure	Positive	Tajima's D, Fst, Fay and Wu's H, correlation of allele frequency with geographic distribution, Hudson's haplotype test, HKA	E. E. Thompson et al., Am J Hum Genet 75, 1059 (Dec, 2004).
27	DAZ	chrY:25247346-25389955	function-altering mutations	functional role in fertility, specific function unknown but suggested it controls cell cycle switch from mitotic to meiotic cell division	Positive, Purifying	dN/dS, PAML	J. P. Bielawski, Z. Yang, Mol Biol Evol 18, 523 (Apr, 2001).
28	DAZL1	chr3:16603306-16621932	function-altering mutations	functional role in fertility, specific function unknown but suggested it controls cell cycle switch from mitotic to meiotic cell division	Positive, Purifying	dN/dS, PAML	J. P. Bielawski, Z. Yang, Mol Biol Evol 18, 523 (Apr, 2001).
29	DCN	chr12:90042171-90078922	diversity (low), derived allele frequency	signal transduction, associated with renal disease	Positive	Tajima's D, Fu and Li's D, Fu and Li's F, Fay and Wu's H	J. M. Akey et al., PLoS Biol 2, e286 (Oct, 2004).
30	DMD	chrX:32294370-33117215	diversity (low), long haplotype, population differentiation	mutation causes Duchenne's Muscular Dystrophy	Positive	Tajima's D, Fu and Li's D, HKA, LD, Fst	M. W. Nachman, S. L. Crowell, Genetics 155, 1855 (Aug, 2000).

Table S1.

Number	Gene	Chromosomal position (HG16)	Signal(s)	Brief function	Type of selection	Tests	References
31	DRD4	chr11:627536-630933	function-altering mutations, long haplotype	dopamine receptor associated with novelty seeking and ADHD	Positive, Balancing	Ka/Ks, LD	Y. C. Ding et al., Proc Natl Acad Sci U S A 99, 309 (Jan 8, 2002). E. Wang et al., Am J Hum Genet 74, 931 (May, 2004).
32	ECP	chr14:19349689-19350634	function-altering mutations	eosinophil cationic protein with an anti-pathogen function	Positive	Bn/Bs	J. Zhang, H. F. Rosenberg, M. Nei, Proc Natl Acad Sci U S A 95, 3708 (Mar 31, 1998).
33	EPHB6	chr7:142023752-142039777	diversity (low)	an ephrin receptor, part of a group of receptors and ligands that mediate numerous developmental processes, particularly in the nervous system	Positive	Tajima's D, Fu and Li's D, Fu and Li's F, Fay and Wu's H	J. M. Akey et al., PLoS Biol 2, e286 (Oct, 2004).
34	F9	chrX:137318530-137351224	diversity (low)	important in process of hemostasis and normal blood clotting	Positive, Purifying	Fu and Li's D, Tajima's D, HKA	E. E. Harris, J. Hey, Curr Biol 11, 774 (May 15, 2001).
35	FOXP2	chr7:113609778-113884658	derived allele frequency	associated with speech	Positive	Tajima's D, PAML	W. Enard et al., Nature 418, 869 (Aug 22, 2002). J. Zhang, D. M. Webb, O. Podlaha, Genetics 162, 1825 (Dec, 2002).
36	FUT2	chr19:53875071-53899057	diversity (high), population differentiation	encodes a protein responsible for determining an individual's A, B, and O blood protein secretor status	Balancing	Fst, LRH, MAF threshold	E. C. Walsh et al., Hum Genet, 1 (Dec 14, 2005). Y. H. Liu et al., J Hum Genet 44, 181 (1999). Y. Koda et al., Hum Genet 113, 534 (Nov, 2003).
37	FY	chr1:156391389-156393081	population differentiation, long haplotype, derived allele frequency	protection from P. vivax malaria	Positive	Fst, Tajima's D, HKA, LD	F. B. Livingstone, Hum Biol 56, 413 (Sep, 1984). M. T. Hamblin, A. Di Rienzo, Am J Hum Genet 66, 1669 (May, 2000). M. T. Hamblin, E. E. Thompson, A. Di Rienzo, Am J Hum Genet 70, 369 (Feb, 2002). P. A. Zimmerman et al., Proc Natl Acad Sci U S A 96, 13973 (Nov 23, 1999).
38	G6PD	chrX:152227156-152243020	long haplotype	protection from P.falciparum malaria vs. G6PD deficiency	Positive, Balancing	LD, LRH	M. A. Saunders, M. F. Hammer, M. W. Nachman, Genetics 162, 1849 (Dec, 2002). S. A. Tishkoff et al., Science 293, 455 (Jul 20, 2001). B. C. Verrelli et al., Am J Hum Genet 71, 1112 (Nov, 2002). P. C. Sabeti et al., Nature 419, 832 (Oct 24, 2002).
39	GRK4	chr4:3002347-3079479	population differentiation, long haplotype	involved in blood pressure regulation, associated with hypertension	Positive	Fst, Pexcess, LD	K. E. Lohmueller et al., Annals of Human Genetics (2005).
40	GYPA	chr4:145608083-145639471	function-altering mutations, diversity (high)	Decoy Receptor helps with Pathogen defense, principal receptor for malarial parasite	Positive (primates), Balancing (within human populations)	dN/dS, HKA, Tajima's D, Fu and Li's D, Fu and Li's F	J. Baum, R. H. Ward, D. J. Conway, Mol Biol Evol 19, 223 (Mar, 2002).

Table S1.

Number	Gene	Chromosomal position (HG16)	Signal(s)	Brief function	Type of selection	Tests	References
41	HAVCR1	chr5:156437331-156466865	function-altering mutations, diversity (high)	encodes T-cell immunoglobulin and mucin-domain containing proteins	Positive, Balancing	Pa/Ps, π_a/π_s , heterozygosity	T. Nakajima et al., Genes Immun (May 12, 2005).
42	HBB	chr11:5211004-5212610	population differentiation, long haplotype	protection from P.falciparum malaria vs. Sickle cell anemia	Positive, Balancing	LD, age simulations, HS , correlation of allele frequency with geographical distribution	A. C. Allison, Exp Parasitol 6, 418 (Jul, 1957). A. C. Allison, Br Med J 4857, 290 (Feb 6, 1954). A. C. Allison, Ann N Y Acad Sci 91, 710 (Jun 7, 1961). M. Currat et al., Am J Hum Genet 70, 207 (Jan, 2002). J. G. Mears et al., J Clin Invest 68, 606 (Sep, 1981). J. Ohashi et al., Am J Hum Genet 74, 1198 (Jun, 2004). J. Pagnier et al., Proc Natl Acad Sci U S A 81, 1771 (Mar, 1984).
43	HFE	chr6:26195426-26205035	long haplotype, population differentiation	decreased risk of anemia in pregnancy vs. hemochromatosis	Positive	Tajima's D, Fay and Wu's H, Fu and Li's D, Fu's Fs, HKA, Fst, LD	R. S. Ajioka et al., Am J Hum Genet 60, 1439 (Jun, 1997). J. Rochette et al., Am J Hum Genet 64, 1056 (Apr, 1999). C. Toomajian, R. S. Ajioka, L. B. Jorde, J. P. Kushner, M. Kreitman, Genetics 165, 287 (Sep, 2003). C. Toomajian, M. Kreitman, Genetics 161, 1609 (Aug, 2002).
44	HLA	chr6:28918812-33377873	long haplotype, diversity (high), function-altering mutations	key regulator of immune recognition	Balancing	heterozygosity, LD, LRH, dN/dS, pNC/pNR, Hardy-Weinberg, Watterson's F , Identity excess	B. O. Bengtsson, G. Thomson, Tissue Antigens 18, 356 (Nov, 1981). P. W. Hedrick, Am J Hum Genet 35, 1055 (Sep, 1983). P. W. Hedrick, G. Thomson, Genetics 104, 449 (Jul, 1983). A. V. Hill, A. Jenson, M. Plebanski, S. C. Gilbert, Philos Trans R Soc Lond B Biol Sci 352, 1317 (Sep 29, 1997). A. V. Hill et al., Philos Trans R Soc Lond B Biol Sci 346, 379 (Nov 29, 1994). A. L. Hughes, M. K. Hughes, C. Y. Howell, M. Nei, Philos Trans R Soc Lond B Biol Sci 346, 359 (Nov 29, 1994). A. L. Hughes, M. Nei, Proc Natl Acad Sci U S A 86, 958 (Feb, 1989). A. L. Hughes, M. Nei, Nature 335, 167 (Sep 8, 1988). A. L. Hughes, T. Ota, M. Nei, Mol Biol Evol 7, 515 (Nov, 1990). D. A. Lawlor, J. Zemmour, P. D. Ennis, P. Parham, Annu Rev Immunol 8, 23 (1990). M. M. Miretti et al., Am J Hum Genet 76, 634 (Apr, 2005). T. Ohta, Proc Natl Acad Sci U S A 88, 6716 (Aug 1, 1991). O. Olerup, M. Troye-Blomberg, G. M. Schreuder, E. M. Riley, Proc Natl Acad Sci U S A 88, 8480 (Oct 1, 1991). H. Salamon et al., Genetics 152, 393 (May, 1999). Y. Satta, C. O'HUigin, N. Takahata, J. Klein, Proc Natl Acad Sci U S A 91, 7184 (Jul 19, 1994). N. Takahata, M. Nei, Genetics 124, 967 (Apr, 1990). N. Takahata, Y. Satta, J. Klein, Genetics 130, 925 (Apr, 1992). E. C. Walsh et al., Am J Hum Genet 73, 580 (Sep, 2003). J. A. Young, D. Wilkinson, W. F. Bodmer, J. Trowsdale, Proc Natl Acad Sci U S A 84,

Table S1.

Number	Gene	Chromosomal position (HG16)	Signal(s)	Brief function	Type of selection	Tests	References
45	IL13	chr5:132070080-132073017	population differentiation, derived allele frequency	believed to play important role in the pathogenesis of atopy and allergic asthma	Positive	Fst , HKA, Tajima's D, Fu and Li's D, Fay and Wu's H	G. Zhou et al., Mol Biol Evol 21, 29 (Jan, 2004).
46	IL1A	chr2:113626783-113637193	diversity (high)	immunity and defense, associated with nasal polyposis, atopy, systemic sclerosis, and Alzheimer's disease	Balancing	Tajima's D , Fu and Li's D, Fu and Li's F , Fay and Wu's H	J. M. Akey et al., PLoS Biol 2, e286 (Oct, 2004).
47	IL4	chr5:132086030-132094495	population differentiation	cytokine, stimulates: T cells to take TH2 phenotype, B cell type switching, and IgE secretion	Positive	Fst , maximum likelihood analysis	M. V. Rockman, M. W. Hahn, N. Soranzo, D. B. Goldstein, G. A. Wray, Curr Biol 13, 2118 (Dec 2, 2003).
48	IL4RA	chr16:27232818-27283596	diversity (low), function altering mutations	encodes alpha chain for IL4 receptor	Balancing, Purifying	MK , Tajima's D	X. Wu, A. Di Rienzo, C. Ober, Genes Immun 2, 128 (May, 2001).
49	KEL	chr7:142109132-142130433	diversity (low), derived allele frequency	a type II transmembrane glycoprotein that is the highly polymorphic Kell blood group antigen	Positive	Tajima's D , Fu and Li's D , Fu and Li's F , Fay and Wu's H	J. M. Akey et al., PLoS Biol 2, e286 (Oct, 2004).
50	KIR3DL1	chr19:59989604-60034044	diversity (high)	expressed by cells of the immune system and regulates their function	Positive	Watterson's F , LD , Fst, pNC/pNR, dN/dS	A. L. Hughes, Mol Phylogenet Evol 25, 330 (Nov, 2002). P. J. Norman et al., Immunogenetics 56, 225 (Jul, 2004).
51	KIR3DL3	chr19:59927795-59939815	function-altering mutations	expressed by cells of the immune system and regulates their function	Balancing	Watterson's F, LD, Fst, pNC/pNR, dN/dS	A. L. Hughes, Mol Phylogenet Evol 25, 330 (Nov, 2002). P. J. Norman et al., Immunogenetics 56, 225 (Jul, 2004).
52	LCT	chr2:136762675-136781298	long haplotype, population differentiation, diversity (low)	nutrition from milk, increased fecundity	Positive	LRH , pexcess , diversity	T. Bersaglieri et al., Am J Hum Genet 74, 1111 (Jun, 2004). N. S. Enattah et al., Nat Genet 30, 233 (Feb, 2002). E. J. Hollox et al., Am J Hum Genet 68, 160 (Jan, 2001).
53	LDLR	chr19:11061131-11105490	diversity (high)	cell surface glycoprotein, plays a role in plasma and cholesterol maintenance	Balancing	Tajima's D , Fu and Li's D , Fu and Li's F , Fst, haplotype diversity	N. J. Fagundes, F. M. Salzano, M. A. Batzer, P. L. Deininger, S. L. Bonatto, Ann Hum Genet 69, 389 (Jul, 2005).
54	LILRA3	chr19:59491668-59496050	population differentiation	expressed by cells of the immune system and regulates their function	Positive	Watterson's F, LD, Fst	P. J. Norman et al., Immunogenetics 56, 225 (Jul, 2004).
55	LMP7	chr6:32916472-32920458	function-altering mutations	encodes a member of the proteasome B-type family	Positive	dN/dS	D. H. Bos, J Mol Evol 60, 221 (Feb, 2005).

Table S1.

Number	Gene	Chromosomal position (HG16)	Signal(s)	Brief function	Type of selection	Tests	References
56	MAOA	chrX:42636592-42652049	diversity (low), long haplotype, derived allele frequency	mutations cause broderline mental retardation and abnormal behavior	Positive	Tajima's D, Fay and Wu's H, HKA, LD, Snn, Fu and Li's D, MK, PAML, Ka/Ks	A. M. Andres et al., Hum Genet 115, 377 (Oct, 2004). Y. Gilad, S. Rosenberg, M. Przeworski, D. Lancet, K. Skorecki, Proc Natl Acad Sci U S A 99, 862 (Jan 22, 2002).
57	MBL2	chr10:54195146-54201466	diversity (high), long haplotype, population differentiation	plays an important role in innate immunity	Balancing	LD, Tajima's D, Fu and Li's D, Fu and Li's F	T. Bernig et al., Genes Immun 5, 461 (Sep, 2004).
58	MC1R	chr16:89727518-89729615	population differentiation, diversity (low)	phenotypic variation in skin and hair pigmentation	Positive (non-Africans), Purifying (Africans)	MK, HKA, Tajima's D, Ewens-Watterson's F, Fu and Li's D, Fu and Li's F, Fst	R. M. Harding et al., Am J Hum Genet 66, 1351 (Apr, 2000). (Counterpoint, suggests drift) K. D. Makova, M. Ramsay, T. Jenkins, W. H. Li, Genetics 158, 1253 (Jul, 2001).
59	MCPH1	chr8:6251529-6488548	function-altering mutations, diversity (low)	controls human brain size, mutations lead to primary microcephaly	Positive	Ka/Ks, MK , Tajima's D, Fu and Li's D, Fu and Li's F, Fay and Wu's H, Fu's Fs	P. D. Evans, J. R. Anderson, E. J. Vallender, S. S. Choi, B. T. Lahn, Hum Mol Genet 13, 1139 (Jun 1, 2004). Y. Q. Wang, B. Su, Hum Mol Genet 13, 1131 (Jun 1, 2004).
60	MDR1	chr7:86745208-86954599	long haplotype	multi-drug transregulates traffic of drugs, peptides and xenobiotics into the body and into sensitive tissues	Positive	LD, adapted LRH (SNPs)	K. Tang, L. P. Wong, E. J. Lee, S. S. Chong, C. G. Lee, Hum Mol Genet 13, 783 (Apr 15, 2004).
61	MGC2217	chr8:57174317-57181135	function-altering mutations	hypothetical protein	Positive	Ka/Ks	N. Osada et al., Genomics 79, 657 (May, 2002).
62	MMP3	chr11:102244249-102252037	diversity (low), population differentiation	associated with risk of coronary artery disease	Positive	Fst, Depaulis Veuille's K and H, Fu's Fs, Hudson's Haplotype test	M. V. Rockman et al., Curr Biol 14, 1531 (Sep 7, 2004).
63	MRG	chr6:mutiple-multiple	function-altering mutations	neuron specific gene family, pain response	Positive	Ka/Ks with sliding window	S. S. Choi, B. T. Lahn, Genome Res 13, 2252 (Oct, 2003).
64	OPN1LW	chrX:151877661-151892305	diversity (low), population differentiation, function-altering mutations	long wave "red" opsin in vision	Positive, Purifying	LD, Tajima's D, Fst, HKA, MK	B. C. Verrelli, S. A. Tishkoff, Am J Hum Genet 75, 363 (Sep, 2004).
65	OR51A6P (Horde annotation)	chr11:4875117-4876135	diversity (low), derived allele frequency	receptors involved in olfaction	Positive (humans), Purifying (chimps)	Tajima's D, Fay and Wu's H, Ka/Ks, PRF Model, PAML, MK	Y. Gilad, C. D. Bustamante, D. Lancet, S. Paabo, Am J Hum Genet 73, 489 (Sep, 2003).
66	OR51G2 (Horde annotation)	chr11:4,900,261-4,901,202	diversity (low), derived allele frequency	receptors involved in olfaction	Positive (humans), Purifying (chimps)	Tajima's D, Fay and Wu's H, Ka/Ks, PRF Model, PAML, MK	Y. Gilad, C. D. Bustamante, D. Lancet, S. Paabo, Am J Hum Genet 73, 489 (Sep, 2003).

Table S1.

Number	Gene	Chromosomal position (HG16)	Signal(s)	Brief function	Type of selection	Tests	References
67	PRM1	chr16:11282199-11282685	function-altering mutations	influence sperm's morphology defining competitiveness, substitutions for histones in sperm	Positive	dN/dS, MK, HKA	A. P. Rooney, J. Zhang, Mol Biol Evol 16, 706 (May, 1999). G. J. Wyckoff, W. Wang, C. I. Wu, Nature 403, 304 (Jan 20, 2000). W. J. Swanson, Z. Yang, M. F. Wolfner, C. F. Aquadro, Proc Natl Acad Sci U S A 98, 2509 (Feb 27, 2001). Nielsen et al., PLoS Biol 3, e170 (Jun, 2005).
68	PRM2	chr16:11276997-11277838	function-altering mutations	influence sperm's morphology, defining competitiveness	Positive	dN/dS, MK	G. J. Wyckoff, W. Wang, C. I. Wu, Nature 403, 304 (Jan 20, 2000). W. J. Swanson, Z. Yang, M. F. Wolfner, C. F. Aquadro, Proc Natl Acad Sci U S A 98, 2509 (Feb 27, 2001).
69	PRNP	chr20:4615068-4630233	function-altering mutations, diversity (high)	heterozygosity for a common polymorphism confers relative resistance to prion diseases	Balancing	MK, Tajima's D	M. Soldevila, F. Calafell, A. Helgason, K. Stefansson, J. Bertranpetit, Trends Genet 21, 389 (Jul, 2005). M. Soldevila et al., Genome Res (Dec 20, 2005). S. Mead et al., Science 300, 640 (Apr 25, 2003).
70	Proto cadherin alpha cluster	chr5:140181817-140432988	diversity (high)	expressed in synaptic junctions in the developing brain	Balancing	Tajima's D, LD, Fu's Fs	J. P. Noonan et al., Am J Hum Genet 72, 621 (Mar, 2003).
71	PTC/ TAS2R38	chr7:141079764-141080766	diversity (high)	bitter taste receptor, population differentiation (low)	Balancing	Tajima's D, Fu and Li's D, Fu and Li's F, Fst, MK, Ka/Ks	S. Wooding et al., Am J Hum Genet 74, 637 (Apr, 2004).
72	Pyrin/MEFV	chr16:3292312-3306912	function-altering mutations	critical role in the innate immune response, associated with Familial Mediterranean fever	Positive	dN/dS	P. Schaner et al., Nat Genet 27, 318 (Mar, 2001).
73	RHCE	chr1:25359819-25419545	function-altering mutations	RH blood group	Positive	PAML	T. Kitano, N. Saitou, J Mol Evol 49, 615 (Nov, 1999).
74	RHD	chr1:25154732-25212646	function-altering mutations	RH blood group	Positive	PAML	T. Kitano, N. Saitou, J Mol Evol 49, 615 (Nov, 1999).
75	SCA2	chr12:110302207-110449433	long haplotype	CAG repeat associated with autosomal dominant cerebellar ataxia	Positive	LRH	F. Yu et al., PLoS Genet 1, e41 (Sep, 2005).
76	SEMG1	chr20:44521102-44523837	function-altering mutations	predominant structural proteins in human semen	Positive	Ka/Ks, PAML	M. I. Jensen-Seaman, W. H. Li, J Mol Evol 57, 261 (Sep, 2003).
77	SIGLEC9	chr19:56319976-56325378	function-altering mutations	sialic-acid binding receptors of innate immune cells	Positive, Balancing	dN/dS	J. L. Sonnenburg, T. K. Altheide, A. Varki, Glycobiology 14, 339 (Apr, 2004).
78	SRY	chrY:2300025-2300922	function-altering mutations	the mammalian sex-determining locus in the Y chromosome	Positive	Ka/Ks	L. S. Whitfield, R. Lovell-Badge, P. N. Goodfellow, Nature 364, 713 (Aug 19, 1993). P. Pamilo, R. J. O'Neill, Mol Biol Evol 14, 49 (Jan, 1997).
79	TAS2R16	chr7:122189315-122190191	derived allele frequency	bitter taste receptor	Positive	Fay and Wu's H	N. Soranzo et al., Curr Biol 15, 1257 (Jul 26, 2005).

Table S1.

Number	Gene	Chromosomal position (HG16)	Signal(s)	Brief function	Type of selection	Tests	References
80	TRPV5	chr7:142092820-142101750	diversity (low), derived allele frequency	step in kidney, intestine, and placenta, calcium absorption	Positive	Tajima's D, Fu and Li's D, Fu and Li's F, Fay and Wu's H	J. M. Akey et al., PLoS Biol 2, e286 (Oct, 2004).
81	TRPV6	chr7:142039887-142040438	diversity (low), derived allele frequency	step in kidney, intestine, and placenta, calcium absorption	Positive	Tajima's D, Fu and Li's D, Fu and Li's F, Fay and Wu's H	J. M. Akey et al., PLoS Biol 2, e286 (Oct, 2004).
82	V1RL1	chr19:62658604-62659666	function-altering mutations	codes for pheromone receptor	Positive	dN/dS	N. I. Mundy, S. Cook, Mol Biol Evol 20, 1805 (Nov, 2003).
83	VAV3	chr1:107500397-107585800	long haplotype	a hemopoietic cell specific guanine nucleotid exchange factor involved in actin rearrangement	Positive	Fst, LRH , MAF threshold	E. C. Walsh et al., Hum Genet, 1 (Dec 14, 2005).
84	ZP2	chr16:21116273-21130369	function-altering mutations	codes for female fertilization zona pellucida glycoprotein	Positive	dN/dS	W. J. Swanson, Z. Yang, M. F. Wolfner, C. F. Aquadro, Proc Natl Acad Sci U S A 98, 2509 (Feb 27, 2001).
85	ZP3	chr7:75698922-75716038	function-altering mutations	codes for female fertilization zona pellucida glycoprotein	Positive	dN/dS	W. J. Swanson, Z. Yang, M. F. Wolfner, C. F. Aquadro, Proc Natl Acad Sci U S A 98, 2509 (Feb 27, 2001).
86	IGHG gene family	chr14:105161971-106270487	function-altering mutations	antigen binding	Positive	dN/dS	J. Ellison, L. Hood, Proc Natl Acad Sci U S A 79, 1984 (Mar, 1982). T. Ota, M. Nei, Mol Biol Evol 11, 469 (May, 1994). T. Tanaka, M. Nei, Mol Biol Evol 6, 447 (Sep, 1989).
87	KIR gene family	chr19:multiple-multiple	function-altering mutations, diversity (high)	cell-surface receptors expressed on natural killer cells (killer inhibitory receptor)	Positive, Balancing	dN/dS, pNC/pNR, Watterson's F, LD, Fst	R. D. Salter, H. W. Chan, R. Tadikamalla, D. A. Lawlor, Immunol Rev 155, 175 (Feb, 1997). P. J. Norman et al., Immunogenetics 56, 225 (Jul, 2004). A. L. Hughes, Mol Phylogenet Evol 25, 330 (Nov, 2002).
88	Male reproductive genes	chrmultiple:mutiple multiple	function-altering mutations	influence sperm's morphology defining competitiveness, substitutions for histones in sperm	Positive	dN/dS, MK	G. J. Wyckoff, W. Wang, C. I. Wu, Nature 403, 304 (Jan 20, 2000). R. Nielsen et al., PLoS Biol 3, e170 (Jun, 2005).
89	OR gene family	chrmultiple:mutiple multiple	function-altering mutations, diversity (low), derived allele frequency	20 genes that are receptors involved in olfaction	Positive (humans), Purifying (chimps)	Tajima's D, Fay and Wu's H, Ka/Ks, PRF Model, PAML, MK	Y. Gilad, C. D. Bustamante, D. Lancet, S. Paabo, Am J Hum Genet 73, 489 (Sep, 2003).
90	SPANX gene family	chrX:multiple-multiple	function-altering mutations	encode cancer/testis-specific antigens that are potential targets for cancer immunotherapy	Positive	dN/dS	N. Kouprina et al., PLoS Biol 2, E126 (May, 2004).
91	TAS2R cluster	chr12:11029826-11230810	function-altering mutations	bitter taste perception, poison recognition	Positive	dN/dS	P. Shi, J. Zhang, H. Yang, Y. P. Zhang, Mol Biol Evol 20, 805 (May, 2003).

Table S2. The five examples below demonstrates how emperical data from different studies generally involve different statistical tests, populations sampled, sizes of genomic regions, and methods of data collection. The great differences in methodology makes it difficult to compare across studies.

Gene	Reference	Tests	Populations Sampled	Data Collection	Variant positions
ALDH2	H. Oota et al., Ann Hum Genet 68, 93 (Mar, 2004).	Fst, LD	3930 individuals in 37 worldwide populations (b/w 46 and 232 chromosomes per population)	Genotyping: 5 SNPs and 1 STRP site spanning 123kb total	5 SNPs and 1 STRP site
DMD	M. W. Nachman, S. L. Crowell, Genetics 155, 1855 (Aug, 2000).	Tajima's D, Fu and Li's, HKA, LD, Fst	41 males from 4 continental populations (African, Europe, Asia, Americas)	Sequencing: 5.4kb in 2 introns that were 500kb apart.	5 SNPs, 2 1bp indels, and 1 5bp indel
G6PD	P. C. Sabeti et al., Nature 419, 832 (Oct 24, 2002).	LD, LRH	504 males in 3 African populations (Beni, Yoruba, Shona)	Genotyping: 25 SNPs spanning 440kb	25 SNPs
LCT	T. Bersaglieri et al., Am J Hum Genet 74, 1111 (Jun, 2004).	p _{excess} , LRH	360 individuals from 3 populations (European-American, African-American, East Asian)	Genotyping: 101 SNPs spanning 3.2Mb	101 SNPs
MAOA	Y. Gilad, S. Rosenberg, M. Przeworski, D. Lancet, K. Skorecki, Proc Natl Acad Sci U S A 99, 862 (Jan 22, 2002).	Tajima's D, Fay and Wu's, HKA, LD,	56 individuals in 8 populations (Pygmy, Aboriginal Taiwanese, Chinese, Japanese, Mexican, Russian, Ashkenazi Jew, Bedouins)	Sequencing: 18.8kb in 5 regions spanning 90kb, and identified 41 SNPs	41 SNPs

Table S3. Review of the previously published tests for selection categorized by the 5 signatures of selection described in the main text.

Name		Technique	Selective signal	Papers
<i>Function-altering mutations</i>				
1	Ka/Ks	Ratio of nonsynonymous to synonymous substitutions, per available site	elevated rate of amino acid substitutions relative to synonymous substitutions	W. H. Li, C. I. Wu, C. C. Luo, Mol Biol Evol 2, 150 (Mar, 1985). W. H. Li. J. Mol. Evol. 36, 86 (1993).
2	dN/dS	Ratio of nonsynonymous to synonymous substitutions, per available site (simplified method of estimating nonsynonymous and synonymous sites)	elevated rate of amino acid substitutions relative to synonymous substitutions	M. Nei, T. Gojobori, Mol Biol Evol 3, 418 (Sep, 1986). A. L. Hughes, M. Nei, Nature 335, 167 (Sep 8, 1988). A. L. Hughes, M. Nei, Proc Natl Acad Sci U S A 86, 958 (Feb, 1989).
3	pNC vs. pNR	Comparison of conservative and radical nonsynonymous change rates	elevated rate of radical nonsynonymous substitutions relative to conservative nonsynonymous substitutions	A. L. Hughes, T. Ota, M. Nei, Mol Biol Evol 7, 515 (Nov, 1990). A. L. Hughes, J. A. Green, J. M. Garbayo, R. M. Roberts, Proc Natl Acad Sci U S A 97, 3319 (Mar 28, 2000). J. Zhang, J Mol Evol 50, 56 (Jan, 2000).
4	McDonald/Kreitman (MK)	Comparison of nonsynonymous & synonymous rates between and within 2 species	Elevated rate of amino acid fixation, relative to polymorphism	J. H. McDonald, M. Kreitman, Nature 351, 652 (Jun 20, 1991).
5	PRF Model	advancement of MK test, frequencies of variant sites in two species are modeled as a Poisson random field distribution	Elevated rate of amino acid fixation, relative to polymorphism	S. A. Sawyer, D. L. Hartl, Genetics 132, 1161 (Dec, 1992). C. D. Bustamante et al., Nature 416, 531 (Apr 4, 2002).
6	Bn/Bs	Modification of dN/dS, to examine typology of a phylogenetic tree	elevated rate of amino acid substitutions relative to synonymous substitutions	J. Zhang, H. F. Rosenberg, M. Nei, Proc Natl Acad Sci U S A 95, 3708 (Mar 31, 1998).
7	implemented in PAML	Maximum likelihood estimate of dN/dS, with selection varying by site	dN/dS > 1 in portion of gene	Nielsen, R., and Z. Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929–936. Yang, Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13:555–556.
<i>Diversity and frequency spectrum</i>				
8	Ewens-Watterson Homozygosity (F)	Sum of squared allele frequencies (primarily used to detect balancing selection)	High value indicates excess of low frequency alleles	G. A. Watterson, Theor Popul Biol 7, 256 (Apr, 1975). G. A. Watterson, Genetics 88, 405 (1978).
9	Heterozygosity (π)	Mean number of differences per bp between two copies of locus	Low value (must be normalized for mutation rate for sequence data)	M. Nei, Molecular evolutionary genetics (Columbia University Press, New York, 1987), pp. x, 512 p.
10	HKA	Ratio of polymorphism to divergence from outgroup, compares at two or more loci	low ratio, relative dearth of polymorphism	Hudson, Kreitman, Aguade 1987 (Genetics)
11	Tajima's D	Comparison of π to number of segregating sites	Negative value indicates excess rare alleles	F. Tajima, Genetics 123, 585 (Nov, 1989).
12	Fu and Li's D, D^*	Measures the number of variants observed only once in a sample versus total number of variants sites	excess of relative singleton variants	Y. X. Fu, W. H. Li, Genetics 133, 693 (Mar, 1993).
13	Fu and Li's F, F^*	Measures the number of variants observed only once in a sample versus the mean pairwise differences between sequences	excess of relative singleton variants	Y. X. Fu, W. H. Li, Genetics 133, 693 (Mar, 1993).

Table S3.

Name		Technique	Selective signal	Papers
<i>Diversity and frequency spectrum (continued)</i>				
14	Fu's F_S	Comparison of π and number of genic alleles	Negative value	Y. X. Fu, Genetics 147, 915 (Oct, 1997).
15	Z_{nS}	Average pairwise LD	High value compared to simulation (neglecting recombination)	J. K. Kelly, Genetics 146, 1197 (Jul, 1997).
16		Small number of segregating sites	Low value compared to model of population growth	M. Slatkin, G. Bertorelle, Genetics 158, 865 (Jun, 2001).
17		Degree of two-marker LD; declines with recombination and mutation	High value compared to model of expanding population	M. Slatkin, G. Bertorelle, Genetics 158, 865 (Jun, 2001).
18	RV	Multisatellite length diversity	Reduced allele diversity	C. Schlotterer, Genetics 160, 753 (Feb, 2002).
19	Kim and Stephan	Likelihood ratio of neutral and hitchhiking models, using allele freqs and spatial information	High likelihood for hitchhiking model	Kim, Y., and W. Stephan, Genetics 160, 765 (2002)
20	MAF threshold	Fraction of alleles with low minor allele frequency	High fraction of rare alleles	D. Altshuler et al., Nature 437, 1299 (Oct 27, 2005).
21	CRTR	Contiguous Region of Tajima's D Reduction	Large region with low Tajima's D	C. S. Carlson et al., Genome Res 15, 1553 (Nov, 2005).
<i>High frequency derived alleles</i>				
22	Fay and Wu's H	Comparison of high freq derived alleles to intermediate frequency	Excess high frequency derived alleles	J. C. Fay, C. I. Wu, Genetics 155, 1405 (Jul, 2000).
23	DAF threshold	Fraction of markers with high derived allele frequency	High fraction of derived alleles	D. Altshuler et al., Nature 437, 1299 (Oct 27, 2005).
<i>Differences Between Populations</i>				
24	F_{ST}	Inbreeding coefficient, measures population differentiation	High value	S. Wright, Nature 166, 247 (Aug 12, 1950). S. Wright, Ann Eugen 15, 323 (1951). R. C. Lewontin, J. Krakauer, Genetics 74, 175 (May, 1973).
25	P_{excess}	Frequency change in 1 population relative to another	High value	B. O. Bengtsson, G. Thomson, Tissue Antigens 18, 356 (Nov, 1981). J. Hastbacka et al., Cell 78, 1073 (Sep 23, 1994).
26	CONTML	Maximum likelihood estimate of branch lengths (genetic drift) in bifurcating population model	Inconsistency in amount of drift between selected and neutral loci	J. Felsenstein, Evolution 35, 1229 (1981).
27	FDIST	F_{ST} conditional on heterozygosity	Outlier in F_{ST} vs heterozygosity distribution	M.A. Beaumont, and R. A. Nichols, Proc. R. Soc. Lond. Ser. B 263, 1619 (1996).
28	Snn	Compare nearest neighbors of a haplotype (in sequence space), see if it is found more often in same population as haplotype then expected under panmixia	Nearest neighbors are more often in same population as test haplotype suggesting population structure	R. R. Hudson, Genetics 155, 2011 (Aug, 2000).
29		Joint distribution of inbreeding coefficients conditional on number of alleles	Outlier in joint distribution of inbreeding coefficients	R. Vitalis, K. Dawson, P. Boursot, Genetics 158, 1811 (Aug, 2001).
30	R_{ST}	Population differentiation, assuming stepwise microsatellite mutation	High value	M. Slatkin 1995. Genetics 139:457–462. M. Kayser, S. Brauer, M. Stoneking, Mol Biol Evol 20, 893 (Jun, 2003).

Table S3.

Name		Technique	Selective signal	Papers
<i>Differences Between Populations(continued)</i>				
31		F_{ST} , using bootstrap resampling to estimate null distribution	High value relative to reconstructed neutral distribution	Porter A.H., Molecular Ecology 12, 903–915 (2003).
32		correlation of allele frequency to geographical distribution (eg. latitude, malaria)	High correlation	A. C. Allison, Ann N Y Acad Sci 91, 710 (Jun 7, 1961). E. E. Thompson et al., Am J Hum Genet 75, 1059 (Dec, 2004).
33		Regression-based likelihood estimate of F_{ST} (Bayesian)	Locus- or population-specific increase in inbreeding coeff	M. A. Beaumont, D. J. Balding, Mol Ecol 13, 969 (Apr, 2004).
<i>Haplotype Length and Diversity</i>				
	Identity Excess	A measure of linkage disequilibrium	Excess of LD	T. Ohta, Genet Res 36, 181 (Oct, 1980).
34	Hudson' Haplotype Test	Detection of subset of very similar haplotypes	Large cluster of similar haplotypes	R. R. Hudson, K. Bailey, D. Skarecky, J. Kwiatowski, F. J. Ayala, Genetics 136, 1329 (Apr, 1994).
35	Depaulis & Veuille's K and H	Low haplotype number (K) or diversity (H)	Low value (reflects both haplotype length and low diversity)	F. Depaulis, M. Veuille, Mol Biol Evol 15, 1788 (Dec, 1998).
36		Probability of detecting observed haplotype diversity in an arbitrary window	Low diversity	P. Andolfatto, J. D. Wall, M. Kreitman, Genetics 153, 1297 (Nov, 1999).
37		Length of unrecombined segment	Length greater than predicted by model of population growth	M. Slatkin, G. Bertorelle, Genetics 158, 865 (Jun, 2001).
38	LRH	Relative length of long-range haplotypes (using EHH measure)	High frequency alleles with long-range LD	P. C. Sabeti et al., Nature 419, 832 (Oct 24, 2002).
39		Number of haplotypes normalized by number of segregating sites	Low value (when only intermediat frequency markers used)	M. Przeworski, Genetics 160, 1179 (Mar, 2002).
40	shared ancestral	Extent of haplotype sharing at linked sites	High frequency alleles with long-range LD	C. Toomajian, R. S. Ajioka, L. B. Jorde, J. P. Kushner, M. Kreitman, Genetics 165, 287 (Sep, 2003).
41	Kim and Nielsen	Kim and Stephan method, with addition of LD to likelihood	High likelihood for hitchhiking model	Y. Kim, R. Nielsen, Genetics 167, 1513 (Jul, 2004).
42	Haplotype Similarity (HS)	Similarity of long-range haplotypes, with sliding-window and resampling to produce single LRH score	High frequency alleles with long-range LD	N. A. Hanchard et al., Am J Hum Genet 78, 153 (Jun, 2006).
43	iHS	Relative length of long-range haplotypes (using integral of EHH measure)	High frequency alleles with long-range LD	B. F. Voight, S. Kudravalli, X. Wen, J. K. Pritchard, PLoS Biol 4, e72 (Mar 7, 2006).
44		Relative length of long-range haplotypes using only homozygous individuals in analysis	High frequency alleles with long-range LD	E. T. Wang, G. Kodama, P. Baldi, R. K. Moyzis, Proc Natl Acad Sci U S A 103, 135 (Jan 3, 2006).

Table S4. Percentiles for the previously published candidate loci in comparison to the genome-wide empirical distribution, for 6 tests of natural selection. The gene names are given along with chromosomal position. The column 'signal' gives the signature of selection the gene was first identified by. The signatures are described in the main text and are indicated by 1) function-altering mutations 2a) low diversity, excess of rare alleles 2b) high diversity 3) excess of high frequency derived alleles 4) population differentiation and 5) length of haplotypes. More details on each gene can be found in Table S1.

Gene	Chromosome position (build HG16)	Signal	Function-Altering Mutations (R. Nielsen et al., PLoS Biol 3, e170 (Jun, 2005), C. D. Bustamante et al., Nature 437, 1153 (Oct 20, 2005))			Heterozygosity (D. Altshuler et al., Nature 437, 1299 (Oct 27, 2005))		Excess of rare alleles (D. Altshuler et al., Nature 437, 1299 (Oct 27, 2005))		pexcess (D. Altshuler et al., Nature 437, 1299 (Oct 27, 2005))		LRH test (D. Altshuler et al., Nature 437, 1299 (Oct 27, 2005))			
			Nielsen et al.	Bustamante et al.	Chimpanzee Consortium	AA	CEU	YRI	CEU	YRI	CEU	YRI	CEU	YRI	CEU
1	ABCC1	chr16:16009872-16201711	5	3	na	na	100.0	92.6	75.0	76.6	81.3	86.6	93.5	94.6	54.7
2	ABO	chr9:131406607-131426675	2b, 3	38.7	10	na	67.9	67.6	34.4	14.0	16.3	37.1	97.9	90.9	66.8
3	ACE2	chrX:14940581-14980563	2b	69.2	98	88.8	82.1	63.8	45.0	8.6	25.1	8.4	78.8	70.2	91.3
4	ACTN3	chr11:66089750-66106157	2b	14.6	49	na	90.7	81.3	76.4	98.5	21.7	83.9	71.2	77.3	89.0
5	ADAM2	chr8:39695889-39713143	1	96.9	92	na	38.0	14.6	13.2	34.0	28.7	81.4	44.1	24.9	8.0
6	ADH	chr20:3060203-3061058	4, 5	36.7	24	59.3	59.7	56.9	63.0	79.7	29.6	48.5	65.3	83.8	20.2
7	AGT	chr1:227873893-227881601	5	4.4	11	84	44.8	60.8	80.3	88.8	77.4	15.4	76.2	83.8	41.1
8	ALDH2	chr12:110616298-110659735	4, 5	39.3	0	8.1	84.5	99.3	91.0	83.5	99.6	89.1	71.5	95.6	96.5
9	APCS	chr1:156774407-156775450	5	93	93.3	89.6	88.8	68.7	67.5	93.8	66.3	99.0	59.0	76.5	86.5
10	APOBEC3G	chr22:37716174-37726786	1	99.9	97	97.5	96.1	87.3	45.6	65.8	17.6	63.6	48.7	16.8	na
11	ASPM	chr1:193785110-193847224	1, 5	60.6	45	76.1	81.1	67.7	96.1	90.0	86.3	99.2	87.3	97.5	99.0
12	ATP6	mtDNA	1	na	na	na	na	na	na	na	na	na	na	na	na
13	BRCA1	chr17:38450366-38530994	1, 5	na	na	90.6	89.9	70.5	91.9	48.6	63.0	99.6	91.3	58.7	60.6
14	C6	chr5:41187836-41259113	5	74.3	na	79.3	12.1	31.7	79.8	78.9	65.2	59.9	52.5	50.9	99.8
15	CAPN10	chr2:241846923-241859260	4, 5	40.3	99.2	80.7	36.1	68.7	21.1	59.3	57.1	86.2	63.4	66.5	75.4
16	CCR5	chr3:46372224-46378283	1, 5	43.6	41	60.5	47.9	99.3	51.2	90.3	37.7	85.7	90.7	38.8	94.7
17	CD40L	chrX:134435969-134448155	5	65.4	83	23	58.5	74.0	68.5	64.5	72.2	87.5	99.2	98.9	79.6
18	CD59	chr11:33688866-33722333	1	na	na	96.9	68.1	60.1	4.7	84.3	13.6	96.9	15.0	40.0	96.7
19	CDY	chrY:18935329-18937289	1	na	na	na	na	na	na	na	na	na	na	na	na
20	CGB	chr19:54217938-54219405	1	na	na	75.5	53.5	54.9	79.6	73.4	86.1	56.9	58.4	64.4	62.9
21	COX7C	chr5:85949539-85952337	1	na	0	27	48.1	12.6	60.6	44.8	72.4	63.0	33.8	53.3	74.5
22	COX8	chr11:63498654-63500590	1	na	0	86.8	92.3	97.4	94.3	99.0	91.7	94.1	94.4	42.5	54.9
23	CYP1A2	chr15:72757888-72764754	3	42	na	na	96.4	96.0	97.9	18.2	99.7	9.4	36.4	86.9	23.3
24	CYP3A4	chr7:98965979-98993166	2a, 3, 4, 5	89.2	96	91.7	98.7	94.0	100.0	98.6	99.6	98.7	25.7	49.0	12.5
25	CYP3A5	chr7:98857219-98889018	2a, 4, 5	na	0	66.8	98.7	94.0	100.0	98.6	99.6	98.7	12.3	2.4	48.4
26	DAZ	chrY:25247346-25389955	1	na	na	na	na	na	na	na	na	na	na	na	na
27	DAZL1	chr3:16603306-16621932	1	56.8	na	46	31.0	27.7	63.8	62.2	83.0	29.5	1.6	3.7	84.5
28	DCN	chr12:90042171-90078922	2a, 3	na	17	8.3	36.4	66.7	89.6	71.0	82.1	25.8	90.3	51.7	9.2

Table S4.

Gene	HG16 pos	Signal	Function-Altering Mutations			Heterozygosity		Excess of rare alleles		pexcess		LRH test			
			Nielsen et al.	Bustamante et al.	Chimpanzee Consortium	AA	CEU	YRI	CEU	YRI	CEU	YRI	CEU	HCB	
29	DMD	chrX:32294370-33117215	2a, 5, 4	13.3	10	54.6	86.3	85.2	64.3	83.8	70.5	68.7	99.8	98.8	86.8
30	DRD4	chr11:627536-630933	1, 5	na	na	na	66.0	92.4	23.4	5.7	45.2	16.2	84.2	40.2	40.3
31	ECP	chr14:19349689-19350634	1	92.9	67	na	43.4	50.1	50.7	60.1	52.3	57.5	87.8	74.3	70.9
32	EPHB6	chr7:142023752-142039777	2a, 3	56.6	na	45.8	71.6	63.2	42.8	39.5	62.8	55.4	87.4	68.5	88.8
33	F9	chrX:137318530-137351224	2a	na	49	41	47.8	54.3	66.6	38.1	71.2	27.6	87.8	59.2	53.1
34	FOXP2	chr7:113609778-113884658	3	na	na	39.3	95.7	85.9	70.7	92.7	93.3	97.9	92.7	51.1	88.1
35	FUT2	chr19:53875071-53899057	2b	na	29	na	58.8	65.6	88.6	60.1	89.9	56.9	96.7	93.5	73.7
36	FY	chr1:156391389-156393081	4, 5, 3	85.5	33	75.9	44.6	38.1	46.6	23.5	46.0	86.54*	80.7	67.6	62.5
37	G6PD	chrX:152227156-152243020	5	22.6	18	na	na	na	77.3	92.5	87.2	95.5	97.7	5.1	14.3
38	GRK4	chr4:3002347-3079479	4, 5	88.1	44	53.5	93.4	99.1	57.2	50.3	14.4	72.4	91.7	86.2	63.6
39	GYPA	chr4:145608083-145639471	1, 2b	na	0	na	79.7	58.4	14.1	41.1	23.8	50.7	74.7	93.0	60.2
40	HAVCR1	chr5:156437331-156466865	1, 2b	76.5	na	na	76.6	92.3	94.8	31.2	34.2	11.3	90.3	19.1	46.0
41	HBB	chr11:5211004-5212610	4, 5	63.3	26	10.5	9.0	8.4	19.2	19.8	58.0	60.3	100.0	42.8	17.3
42	HFE	chr6:26195426-26205035	5, 4	na	0	15.5	62.8	88.5	78.9	98.2	48.7	93.4	81.8	27.2	36.4
43	HLA	chr6:28918812-33377873	2b, 5	na	na	na	33.5	98.2	95.5	98.1	86.3	90.5	99.9	99.8	98.4
44	IL13	chr5:132070080-132073017	4, 5	na	31	93.8	58.8	83.0	92.3	41.6	96.4	89.9	99.6	66.3	60.2
45	IL1A	chr2:113626783-113637193	2b	na	na	19.7	93.3	13.2	28.7	97.9	39.2	88.1	59.0	82.2	89.0
46	IL4	chr5:132086030-132094495	4, 2a	85.9	89	89.5	58.8	83.0	92.3	41.6	96.4	89.9	99.6	8.4	60.2
47	IL4RA	chr16:27232818-27283596	1, 2a	na	53	79.8	93.6	88.4	76.2	63.7	57.6	34.3	77.5	27.6	60.3
48	KEL	chr7:142109132-142130433	2a, 3	na	na	83.9	71.6	63.2	42.8	39.5	62.8	55.4	87.4	68.5	88.8
49	KIR3DL1	chr19:59989604-60034044	2b	na	na	na	72.4	87.1	61.6	46.9	75.1	49.9	1.4	31.2	5.7
49	KIR3DL3	chr19:59927795-59939815	2b	76.8	na	na	72.4	87.1	61.6	46.9	75.1	49.9	1.4	31.2	5.7
50	LCT	chr2:136762675-136781298	4, 5	12.6	85	87.5	69.2	100.0	99.9	94.5	99.4	36.8	70.5	100.0	17.1
51	LDLR	chr19:11061131-11105490	2b	40.9	na	67.7	91.9	87.6	92.3	90.9	95.6	76.7	85.4	79.1	88.7
52	LILRA3	chr19:59491668-59496050	4	na	na	na	31.3	66.5	75.3	89.4	75.7	70.6	73.0	81.8	87.6
53	LMP7	chr6:32916472-32920458	1	34.1	0	49.4	2.7	0.9	47.8	86.0	54.2	50.6	44.9	36.1	21.1
54	MAOA	chrX:42636592-42652049	2a, 3, 5	41.4	na	na	80.1	74.9	63.6	62.0	28.5	56.2	93.4	95.5	94.8
55	MBL2	chr10:54195146-54201466	2b, 4, 5	78.9	34	11.1	49.5	46.4	56.9	66.3	92.1	80.8	64.3	37.7	31.9
56	MC1R	chr16:89727518-89729615	2b	36.4	39	89.8	79.1	85.2	74.5	97.4	92.3	77.2	92.7	76.7	36.6
57	MCPH1	chr8:6251529-6488548	1, 3	79.1	2	75.1	47.5	13.6	57.8	8.1	75.5	66.5	97.7	86.8	90.8
58	MDR1	chr7:86745208-86954599	5	72.7	na	na	57.3	61.1	96.9	92.0	97.7	68.0	86.3	98.6	99.7
59	MGC2217	chr8:57174317-57181135	1	na	0	12.8	67.2	95.4	65.9	76.8	56.8	42.7	56.0	74.6	71.3
60	MMP3	chr11:102244249-102252037	2a, 4	na	na	82.6	81.9	56.5	53.5	81.6	15.7	37.9	34.0	97.2	38.5
61	MRG	chr6:mutiple-multiple	1	73.8	na	94	na	na	na	na	na	na	na	na	na
62	OPN1LW	chrX:151877661-151892305	1, 2a, 4	na	na	na	na	na	70.2	92.5	78.8	95.5	7.1	71.1	66.3
63	OR51A6P	chr11:4875117-48761135	1, 2a, 3	na	na	na	57.0	35.8	46.1	19.8	57.3	25.5	na	na	na

* FY was noted as one of the top regions by single-snp Fst analysis in the HapMap Consortium paper

Table S4.

Gene	HG16 pos	Signal	Function-Altering Mutations			Heterozygosity		Excess of rare alleles		pexcess		LRH test			
			Nielsen et al.	Bustamante et al.	Chimpanzee Consortium	AA	CEU	YRI	CEU	YRI	CEU	YRI	CEU	HCB	
64	OR51G2	chr11:4900261-4901202	1, 2a, 3	na	na	na	64.5	84.0	54.2	58.3	61.5	63.3	na	na	na
65	PRM1	chr16:11282199-11282685	1	99.9	99.7	na	64.5	84.0	54.2	58.3	61.5	63.3	71.1	64.1	93.0
66	PRM2	chr16:11276997-11277838	1	na	na	99.6	62.2	39.8	69.7	54.3	78.6	52.6	71.1	64.1	93.0
67	PRNP	chr20:4615068-4630233	1, 2b	58.8	63	62.9	92.4	95.1	75.3	17.5	9.9	23.1	74.5	23.3	55.1
68	Proto cadherin alpha cluster	chr5:140181817-140432988	2b	na	na	14.8	71.1	71.7	67.0	62.6	37.6	47.4	72.6	94.2	91.5
69	PTPRC	chr1:195896481-195950589	1	na	na	90.4	72.2	97.7	85.1	69.3	96.5	59.3	43.7	52.2	69.3
70	Pyrin/MEFV	chr16:3292312-3306912	1	31.9	62	84	na	na	na	na	na	na	25.0	74.4	92.5
71	RHCE	chr1:25359819-25419545	1	na	na	99.9	38.4	48.3	23.4	39.9	22.5	61.1	13.6	77.6	16.4
72	RHD	chr1:25154732-25212646	1	na	na	na	77.4	99.2	91.0	65.2	99.6	89.1	13.6	72.5	16.4
73	SCA2	chr12:110302207-110449433	5	na	na	na	94.5	74.3	65.5	47.2	63.5	48.6	94.7	99.9	58.3
74	SEMG1	chr20:44521102-44523837	1	92.7	na	96.9	51.0	56.5	44.9	34.7	69.8	87.8	26.4	17.6	53.5
75	SIGLEC9	chr19:56319976-56325378	1	87.6	na	97	na	na	na	na	na	na	99.3	98.7	33.9
76	SRY	chrY:2300025-2300922	1	na	na	97.2	85.6	53.7	50.5	52.7	7.4	56.6	na	na	na
77	TAS2R16	chr7:122189315-122190191	3	90.2	11	80.9	68.4	96.6	98.4	66.6	84.5	87.9	82.1	50.5	89.6
78	TAS2R38	chr7:141079764-141080766	1	na	na	89.3	47.0	63.1	25.2	29.0	23.7	66.8	21.9	24.3	19.6
79	TRPV5	chr7:142092820-142101750	2a, 3	12.9	35	57.3	71.6	63.2	42.8	39.5	62.8	55.4	87.4	68.5	88.8
80	TRPV6	chr7:142039887-142040438	2a, 3	53.7	58	72.4	71.6	63.2	42.8	39.5	62.8	55.4	87.4	68.5	88.8
81	V1RL1	chr19:62658604-62659666	1	84.1	na	96.7	73.3	42.9	80.3	63.6	38.6	69.2	38.5	78.7	93.3
82	VAV3	chr1:107500397-107585800	5	49.5	27	37.6	82.4	91.8	75.1	61.1	96.6	43.1	93.3	60.8	31.0
83	ZP2	chr16:21116273-21130369	1	75	59	60.3	84.5	73.8	54.5	32.7	47.4	43.1	69.6	65.2	92.6
84	ZP3	chr7:75698922-75716038	1	83	87	83.1	na	na	na	na	na	na	13.8	36.7	91.3