Supplementary Material

Targets of balancing selection in the human genome

Aida M Andrés[1,2], Melissa J Hubisz[3], Amit Indap[4], Dara G Torgerson[1], Jeremiah D Degenhardt[4], Adam R Boyko[4], Ryan N Gutenkunst[4], Thomas JWhite[5], Eric D Green[2], Carlos D Bustamante[4], Andrew G Clark[1] and Rasmus Nielsen[6].

Supplementary Methods

Data generation.

19 African-Americans (AA) and 20 European-Americans (EA) were sequenced by direct PCR and sequencing for all well-annotated predicted exons of over 20,000 genes in the human genome (Bustamante et al 2005). A strict bioinformatics pipeline ensured true homology of sequences and use of only well-supported SNPs (Boyko et al. 2008). Briefly, reads were mapped to the human reference sequence hg18 and only genes/reads with high sequence identity (>98.5%) with the human genomic sequence and high coverage (>90%) were maintained. Refseq genes were mapped to the human reference sequence hg18 under the same filtering conditions. SNPs from the amplicons mapping a gene model and a unique in-frame Refseq gene were kept provided they laid on a region with good syntenic correspondence with the chimpanzee reference sequence PanTro2. When more than one transcript mapped to the same gene, the longest transcript was used. We also discarded genes with one or more non-specific results from *in-silico* PCR (http://genome.ucsc.edu), run with perfect match=15 and maximum product size=800bp. This process checks for multiple genomic matches of the amplification primers, and detects cases of putative non-specific amplification.

Neutrality tests

Neutrality tests were performed by the method of Nielsen et al. (2009) designed to minimize the effects of demography in neutrality tests and described in the main text:

1. Inference of admixture proportions of individuals using a maximum likelihood method and using the complete dataset. This step represents an effort to take into account the likely admixed nature of African Americans individuals.

2. The demographic parameters that best fit the data are estimated using a (composite) maximum likelihood approach through coalescent simulations, and considering the estimated admixture proportions. This step ascribes as much of the variability observed as possible to demography.

3. For each gene, neutrality tests are then performed and their statistical significance is assessed by comparing the test statistics with neutral coalescent simulations under the inferred demographic scenario. Neutral simulations were performed with ms (Hudson, 2002), with the number of segregating sites and missing data of the gene, a recombination rate of 7.5×10^{-4} per base pair (Nielsen et al. 2005), and the demographic parameters detailed below.

Note that steps 1 and 2 are not intended to infer the exact demographic history of human populations, but to obtain the demographic model that best explains the data observed in the two samples. The model does not necessarily represent the exact demographic history of the populations (as no inference from a genetic analysis does), but its application as the null model in neutrality tests represents a conservative approach: only genes with unusual patterns when compared with the rest of the genome, and according to the demographic history of the sample, will show significant results. This approach is considerably more conservative than using an equilibrium model or using demographic scenarios inferred from different datasets –based on different samples, individuals, and genomic regions.

The best demographic scenario gives maximum likelihood estimates of the parameters as follow: $T = 0.099$ (divergence time between the two populations); $\alpha_A = 9.5$, $\alpha_E = 21.1$ (rate of expansion of African and European populations since divergence time); $m = 6.67$ (gene flow rate of migrants per generation between the two populations); bottleneck in European population 0.1 generations ago lasting 0.01 generations, with a reduction in population size (β) = 0.018; $\gamma = 1.82$ (ratio of the current African to European population size) (Nielsen et al. 2009). These parameters describe the demographic model assumed in the simulations used to estimate the p-values of neutrality tests.

Two alternative demographic scenarios were considered as a way to assess the influence of the demographic model in our results. Model 1 corresponds to the originally inferred demographic scenario, described above. Model 2 corresponds to the same model, but with all genetic admixture between the two populations explained by recent admixture rather than

migration between populations: 20% European ancestry admixture proportion into AA individuals and 0% migration rate ($m = 0$). Model 3 corresponds to the best demographic model inferred from this data using an independent method, δaδi (Gutenkunst R, *in preparation*): $T = 0.142$, $\alpha_A = 10.2$, $\alpha_B = 14.3$, $\gamma = 1.95$, $\beta = 0.021$, $m = 4.6$, f (Admixture proportion EA to AA)= 0.18.

Supplementary Discussion

Signatures of purifying and positive selection, ancestral admixture (or ancestral population structure), and long-range LD of extreme genes.

We find a strongly supported set of genes with signals of long-term balancing selection. Their double signature of excess of polymorphism and intermediate frequency alleles is difficult to reconcile with forces other than balancing selection, including other types of selection. For example, purifying and background selection increase the polymorphism to divergence ratio by preventing fixation of deleterious alleles; but those variants are maintained at low frequencies, biasing the SFS towards rare alleles. Recent relaxed constraint can also increase the levels of variability and slightly reduce the bias toward rare alleles, but it cannot explain the specific increase in intermediate-frequency alleles in extreme genes. As for directional selection, both ongoing and partial sweeps could potentially produce a temporal bias toward intermediate-frequency alleles. For example, Przeworski et al. (Przeworski et al. 2005) reported that some simulation runs of sweeps from standing variation result in a bias toward intermediate-frequency alleles. Nevertheless, no case of directional selection is expected to significantly increase the ratio of polymorphism to divergence. If anything, positive selection would increase divergence if the gene has undergone subsequent sweeps. Showing extreme patterns of both excess polymorphism and excess intermediate-frequency variants, genes in Table 1 are strong candidate targets of balancing selection. Since weak overdominance does not increase polymorphism (Williamson et al. 2004) selection must be strong to lead to the patterns observed.

A possible neutral explanation for the presence of very long genealogies in the genome is ancestral admixture between modern humans and ancestral populations (Garrigan et al. 2005; Plagnol and Wall 2006), equivalent to ancient population structure. Nevertheless, the genomic signature of ancestral admixture differs from that of balancing selection. The main genomic signal of ancestral admixture is extended LD as a consequence of the long genealogical time the two haplotypes (in two non-mating populations) were unable to recombine (Wall 2000; Garrigan

et al. 2005). We tested this possibility by comparing the average LD (r^2) in HapMap SNPs (CEU and YRI) for regions of 20kb centered on every gene in our dataset. Extreme genes do not systematically fall in regions of long-rage LD (T-test P(AA) = 0.1497, P(EA) = 0.1604), confirming that the signal is specific to extreme genes and not to the genomic regions they lay in. Similar results were obtained for regions of 50 kb (T-test P(AA) = 0.9942, P(EA) = 0.3751). Maybe more important, the presence of alleles or haplotypes at intermediate frequencies cannot be explained solely by ancestral admixture. It would be surprising for a newly introduced haplotype to be driven to and maintained at intermediate frequencies for such long time in the absence of selective forces. Note that the same logic applies for large-coalescence regions deriving from putative hybridization between ancestral humans and chimpanzees (Patterson et al. 2006).

Supplementary Tables

Supplementary Table 1: Extreme genes

AA&EA	
ADAM11	ADAM metallopeptidase domain 11
ALPK2	Alpha-kinase 2
BTN1A1	Butyrophilin, subfamily 1, member A1
DEPDC2	DEP domain containing 2
KRT14	Keratin 14 (epidermolysis bullosa simplex, Dowling-Meara, Koebner)
LGALS8	Lectin, galactoside-binding, soluble, 8 (galectin 8)
LILRB4	Leukocyte immunoglobulin-like receptor, subfamily B, member 4
LINS1	Lines homolog 1 (Drosophila)
RCBTB1	Regulator of chromosome condensation and BTB containing protein 1
RPS7	Ribosomal protein S7
RTP4	Receptor (chemosensory) transporter protein 4
TRIM22	Tripartite motif-containing 22
WDR40C	WD repeat domain 40C
AA	
ADAMTS7	ADAM metallopeptidase with thrombospondin type 1 motif, 7
C14orf124	Chromosome 14 open reading frame 124
CLCNKB	Chloride channel Kb
COL27A1	Collagen, type XXVII, alpha 1
COPE	Coatomer protein complex, subunit epsilon
FGF6	Fibroblast growth factor 6
FLJ40243	Hypothetical protein
KRT6B	Keratin 6B
KRT84	Keratin 84
LRRN6A / LINGO1	Leucine rich repeat and Ig domain containing 1
PPP1R15A	Protein phosphatase 1, regulatory (inhibitor) subunit 15A
SERPINH1	Serpin peptidase inhibitor, clade H, member 1,
TARBP1	Tar (HIV-1) RNA binding protein 1
TNS1	Tensin 1
TRPV6	Transient receptor potential cation channel, subfamily V, member 6
EA	
ALDH4A1	Aldehyde dehydrogenase 4 family, member A1
ARHGEF3	Rho guanine nucleotide exchange factor 3
C20orf186	Antimicrobial peptide RY2G5
CAMK2B	Calcium/calmodulin-dependent protein kinase (CaM kinase) II beta
CD200R1	CD200 receptor 1
CDSN	Corneodesmosin
FLJ90650	Laeverin
FUT2	Fucosyltransferase 2 / secretor factor (se)
GM632 / ZNF512B	Zinc finger protein 512B

GPR111	G protein-coupled receptor 111
GRIN3A	Glutamate receptor, ionotropic, N-methyl-D-aspartate 3A
HLA-B	Major histocompatibility complex, class I, B
KIAA0753	KIAA0753
KIAA1303 / RAPTOR	Raptor
KRT6E	Keratin 6E/C
LHB	Luteinizing hormone beta polypeptide
LOC197322 / ACSF3	Acyl-CoA synthetase family member 3
LRAP	Leukocyte-derived arginine aminopeptidase
MYO1G	Myosin IG
NALP13	NLR family, pyrin domain containing 13
PCDHB16	Protocadherin beta 16
RABEP1	Rabaptin, RAB GTPase binding effector protein 1
RIOK2	RIO kinase 2 (yeast)
SAMM50	Sorting and assembly machinery component 50 homolog (S. cerevisiae)
SERPINB5	Serpin peptidase inhibitor, clade B (ovalbumin), member 5
SLC2A9	Solute carrier family 2 member 9
SMARCAD1	SWI/SNF-related, matrix associated actin-dependent regulator of chromatin
TMEM171	Transmembrane protein 171
TSPAN10	Tetraspanin 10
UNC5C	Unc-5 homolog C (C. elegans)
VARSL	Valyl-tRNA synthetase 2, mitochondrial (putative)
ZNF415	Zinc finger protein 415

Supplementary Table 2.1. Well-established targets of balancing selection detected.

Gene	Selection	Observations
HLA-B	Long-term balancing selection (Hedrick et al. 1991) (Sánchez-Mazas 2007)	Only HLA gene with previously reported signatures of balancing selection in humans present in our dataset.
FUT2 (Secretor Factor)	Long-term balancing selection (Koda et al. 2000) (Soejima et al. 2007)	ABO-secretor gene considered an 'honorary blood group'.

Suplementary Table 2.2. Other previously reported targets of balancing selection.

Gene	Selection	Observations
HLA genes	Long-term balancing selection (Hughes and Yeager, 1998)	Most are not present in our dataset due to technical issues in dealing with this complicated genomic region (note HLA-B as an exception)
β-globin	Recent balancing selection (Flint et al. 1998)	Our method does not detect recent selection. Contains less than 10 informative sites (recent selection does not contribute to enrichment in polymorphism): not present in our filtered dataset.
G6PD	Recent balancing selection (Verelli et al. 2002)	Our method does not detect recent selection. Contains less than 10 informative sites (recent selection does not contribute to enrichment in polymorphism): not present in our filtered dataset.
CFTR	Recent balancing selection (Quinton 1994) (Gabriel et al. 1994)	Our method does not detect recent selection. Δ508, the deletion putatively maintained by selection, is not present in our dataset . Other variants are at low frequencies (Lao et al. 2003; this dataset). The gene shows excess of polymorphism (HKAlow P(AA) = 0.0364, P(EA) = 0.1631) but not excess of intermediate-fequency alleles (MWUhigh P(AA) = 0.5472, P(EA) = 0.672)
ABO	Balancing selection (Saitou and Yamamoto, 1997)	Like CFTR, most variants are at low frequencies. The gene shows excess of polymorphism (HKAlow P(AA) = 0.0001, P(EA) = 0.0003) but not excess of intermediate-frequency alleles (MWUhigh P(AA) = 0.1663, P(EA) = 0.4334)

Supplementary Table 3: Gene categories showing the strongest excess of low *P*-values in HKAlow and MWUhigh tests, and in both populations.

category	pMWUhigh		pHKAlow	
	AA	EA	AA	EA
Extracellular matrix	0.0057	0.0387	0	0.0001
Extracellular matrix structural protein	0.0122	0.0042	0.0069	0.0111
Structural protein	0.0167	0.0135	0.0005	0.0006
Intermediate filament	0.0101	0.0115	0.0004	0.0031
Serine protease inhibitor	0.0022	0.0362	0.0043	0.0116

Note: Only categories with HKAlow $P < 0.05$ & MWUlow $P < 0.05$ in both populations are shown. All categories correspond to molecular functions since no biological process category showed consistent excess of low *P*-values in both populations. Notably, the category with the strongest signal in AA is immunity and defense, but this category shows no signals in EA (HKAlow $P = 0.3638$, MWUhigh $P = 0.4177$).

Supplementary Table 4. Extreme genes involved in immunology and response to pathogens.

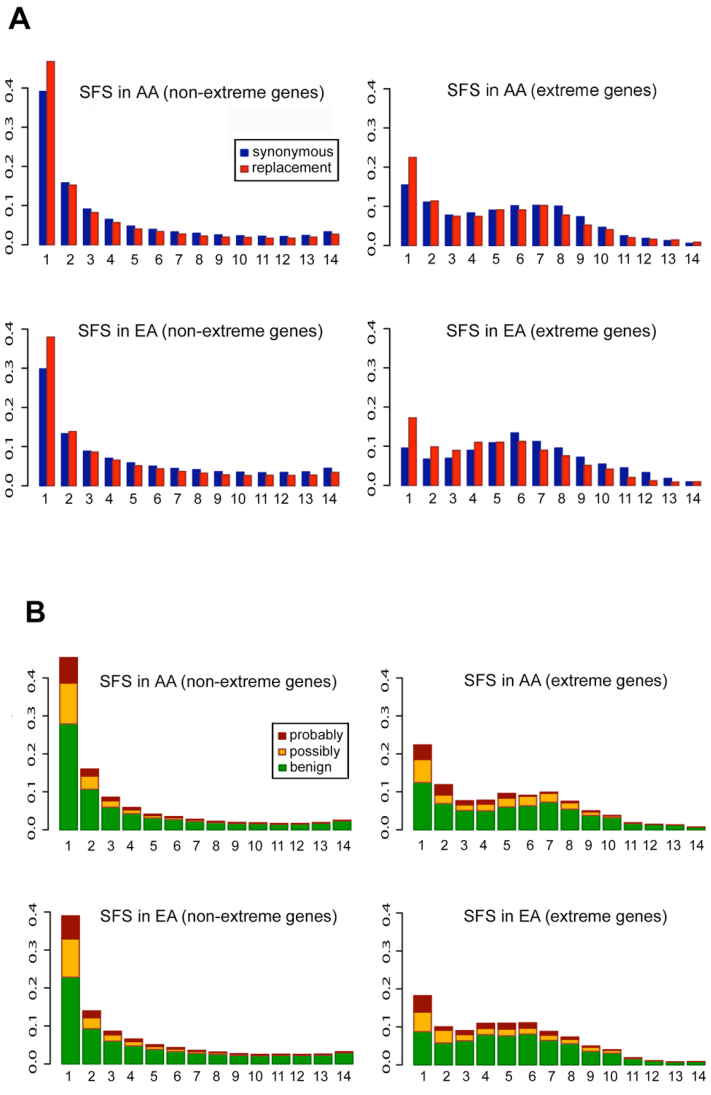
Gene	Function
HLA-B	One of the genes of the MHC complex. The MHC complex presents peptides to immunological cells and represent a first step in the immunological response against non-self peptides. The extremely high variability of these molecules ensures the presentation of diverse peptides to the immune system.
LRAP (ERAP2)	Citoplasmatic proteins (endogenous and non-endogenous) are degraded by a complex of peptidases that result in small peptides that will be further degraded into individual amino acids or transported into the endoplasmatic reticulum for MHC presentation. In the endoplasmatic reticulum ERAP1 and ERAP2 (LRAP) are the proteins that (in a concerted way) trip peptides to the size and characteristics necessary for MHC presentation (Hattori and Tsujimoto, 2004; Saveanu et al. 2005). LRAP/ERAP2 is crucial for the trimming of peptides to be presented by MHC class I and is essential for antigen presentation.
LILRB4	It encodes for a leukocyte immunoglobulin-like receptor, an immunoregulator. LILRB4 protein inhibits, by different mechanisms, immune response from mast cells and NK cells, and plays an important anti-inflammatory role against anaphylactic shock (Katz 2007). Overall, knowledge about LILRB4 function suggests that this protein provides a critical innate protection against an excessive pathologic response to bacteria.
TARBP1 (TRP-185)	HIV-1 activation by the transactivator Tat is dependent on the binding of TAR RNA. *TARBP1*, is one of the two proteins that specifically bind TAR (Wu-Baer et al. 1995; Wu-Baer et al. 1996), being one of the major cellular factors involved in TAR RNA function and, consequently, HIV-1 activation.
TRIM22	It is a tripartite motif protein that mediates interferon inhibition of HIV-1 replication (Bouazzaoui et al. 2006; Barr et al. 2008), making this an important component of the viral particle release process.
FUT2 (Secretor factor)	It encodes for a(1,2)fucosyltransferase (*Se* enzyme). The enzyme regulates the expression of the H antigen in body fluids, and its polymorphism is responsible for the difference between a secretor individual (with at least one active FUT2 allele and expression of ABO antigen in saliva) and a nonsecretor (with no active copies of the gene and no expression of ABO in saliva). Inactive and attenuated forms of FUT2 provide resistance to Norwalk virus infection to nonsecretor individuals (Lindesmith et al. 2003), and recent studies suggest that inactive forms of FUT2 due to nonsense mutation 428G->A confer resistance to HIV-1 infection (Kindberget al. 2006).
CD200R1	It encodes for a member of the immunoglobulin superfamily. CD200R1 protein binds CD200, a glycoprotein with immoregulatory role in diverse tissues (Gorczynski et al. 2005). CD200R1 acts as an inhibitory immune receptor expressed in myeloid cells, T, B, and NK cells (Rijikers et al. 2008) involved in the regulation of Th cell function (Taylor et al. 2005).
C20orf186	It encodes for antimocrobial peptide RY2G5
BTN1A1	It encodes for a member of the immunoglobulin superfamily
LRRN6A (LINGO1)	It encodes for a member of the immunoglobulin superfamily

S**upplementary Table 5:** Extreme genes with known influence in human disease.

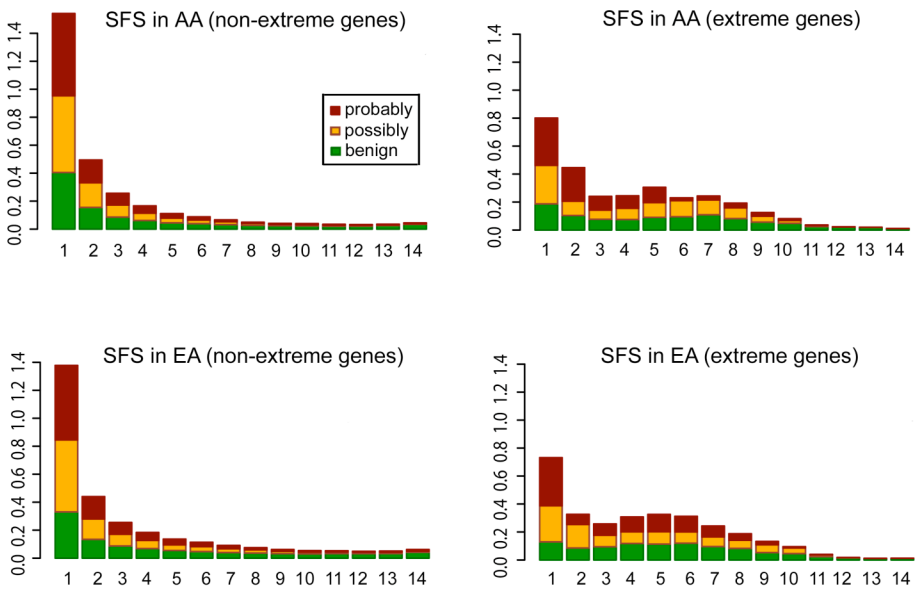
Gene	Disease
ADAM11	Candidate tumor supressor for human breast cancer
ALDH4A1	Type II hyperprolinemia, an autosomal recessive disorder
CDSN	Hypotrichosis simplex of the scalp, possible association with psoriasis
CLCNKB	Autosomal recessive Type III Bartter Syndrome
HLA-B	Progression of HIV infection, ankylosing spondylitis, Stevens-Johnson syndrome, other
KIAA1303	Association with psoriasis
KRT14	Epidermolysis bullosa simplex
KRT6B	Pachyonychia congenita
LHB	Hypogonadism, association with luteinizing hormone and ovulatory disorders
LILRB4	Inhibit autoimmunity, allergies, transplant rejection, and immune deficiencies
RABEP1	Fused with PDGFBR in a case of chronic myelomonocytic leukemia
RCBTB1	Locus for B-cell chronic lymphocytic leukemia
SERPINH1	Autoantibodies to this protein have been found in patients with rheumatoid arthritis
SLC2A9	Uric acid concentrations and gout
TNS1	Knock-out mouse preent cystic kidneys and renal malfunction
TRIM22	Down-regulates transcription from HIV-1 LTR promoter region
UNC5C	Downregulated in colorectal tumor

Supplementary Figures

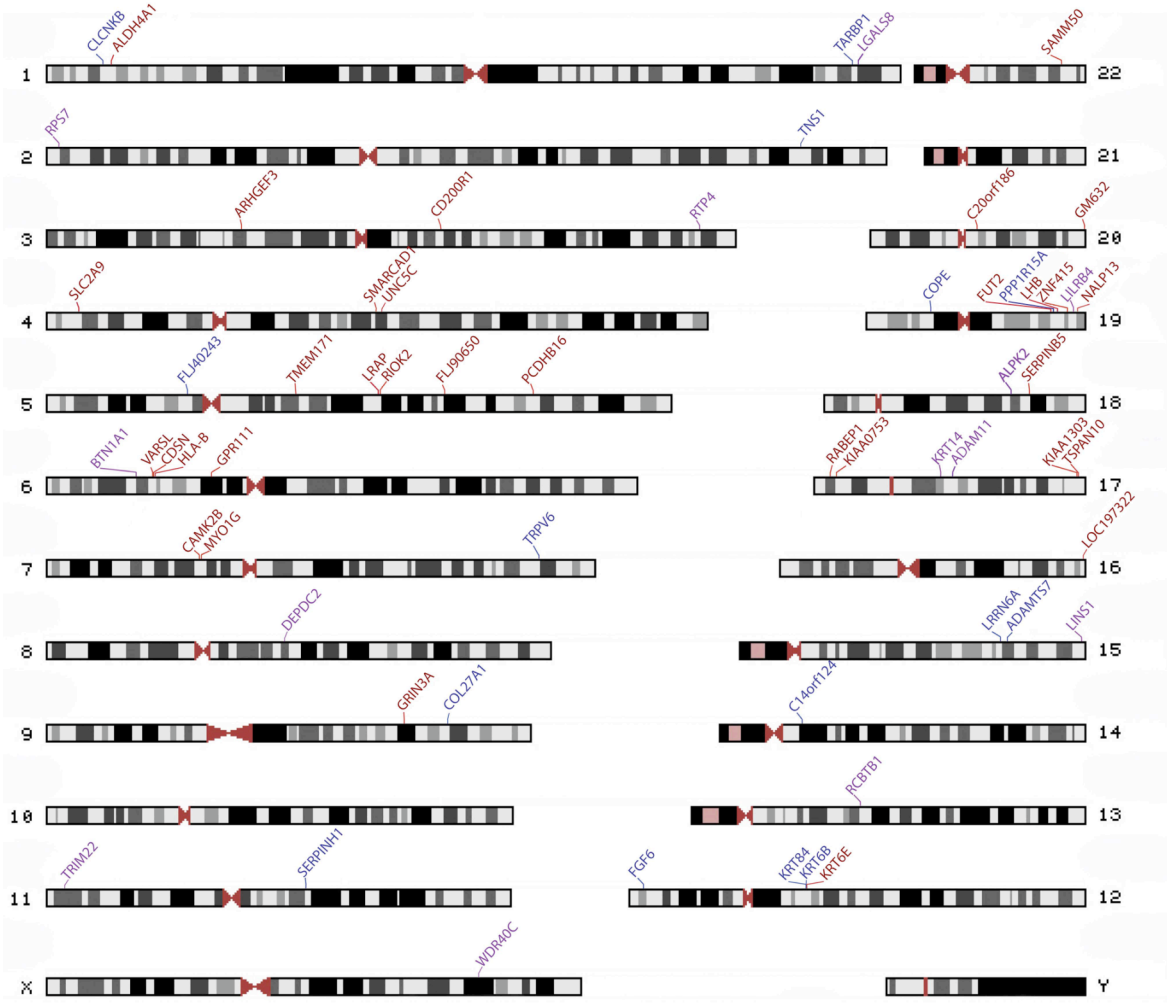
Supplementary Fig. 1: Allele site frequency spectrum (SFS) of all segregating sites in extreme and non-extreme genes. The X-axis represents the absolute allele frequency of SNPs in the sample; to account for missing data, all SFS were projected into a sample of 15 chromosomes (Nielsen et al. 2005). The Y-axis represents the frequency in the data of each respective allele frequency bins. (**A**) SFS by mutation type: synonymous sites (blue) and replacement sites (red) shown for non-extreme genes (left) and extreme genes (right), as shown for each population. (**B**) SFS by PolyPhen category: benign sites (green), sites with possible phenotypic effect (yellow), and sites with probable phenotypic effect (maroon).



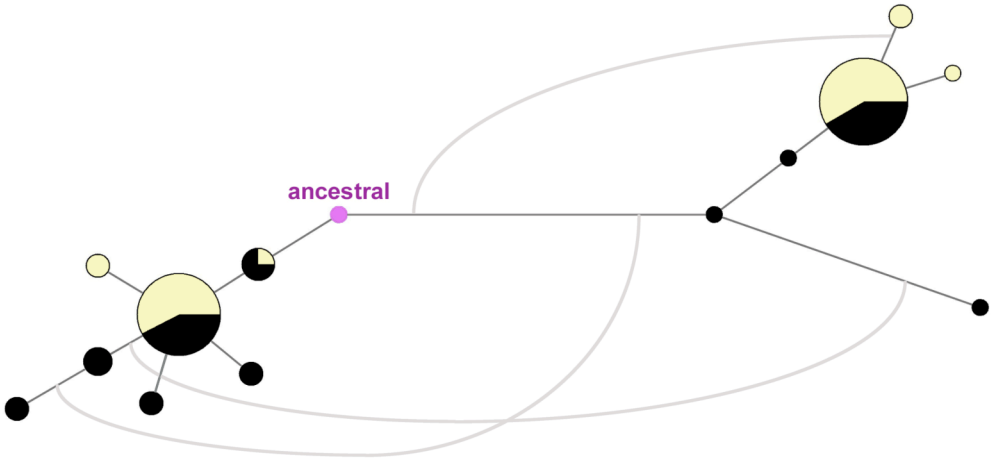
Supplementary Fig. 2: Allele site frequency spectrum (SFS) by PolyPhen category: benign sites (green), sites with possible phenotypic effect (yellow) and sites with probable phenotypic effect (maroon). The high for each color represents the proportion of sites of that PolyPhen category that fall on every allele frequency bin. Therefore the high of each bar does not equal the total number of replacement sites on that bin (see Figure 1). To account for missing data, allele frequencies were projected into a sample of 15 chromosomes (Nielsen et al. 2005).



Supplementary Fig. 3: Genomic location of extreme genes in African-Americans (blue), European-Americans (red) or both populations (purple).



Supplementary Fig. 4: Network of the inferred haplotypes for LINS1 gene. Circles represent haplotypes (size proportional to frequency), black for haplotypes in African-Americans and yellow in European-Americans. The ancestral haplotype corresponds to the ancestral allele for all SNPs, as inferred by comparison with chimpanzee. Length of the dark grey branches between haplotypes is proportional to the number of SNPs present in that branch. Curved light grey lines represent one additional possible singleton recombination/recurrent mutation event detected by manual inspection (two transitions, one transversion). Note that haplotypes were statistically inferred rather than experimentally determined.



Supplementary References

Barr SD, Smiley JR, Bushman FD (2008) The interferon response inhibits HIV particle production by induction of TRIM22. *PLoS Pathog* 4:e1000007.

Bouazzaoui A *et al*. (2006) Stimulated trans-acting factor of 50 kDa (Staf50) inhibits HIV-1 replication in human monocyte-derived macrophages. *Virology* 356:79-94.

Boyko AR, *et al*. (2008) Assessing the evolutionary impact of amino acid mutations in the human genome. *PLoS Genet*. 4(5):e1000083.

Bustamante CD *et al*. (2005) Natural selection on protein-coding genes in the human genome. *Nature* 437:1153-1157.

Flint J, Harding RM, Boyce AJ, Clegg JB (1998) The population genetics of the haemoglobinopathies. *Baillieres Clin Haematol* **11**:1–51.

Gabriel SE, Brigman KN, Koller BH, Boucher RC, Stutts MJ (1994) Cystic fibrosis heterozygote resistance to cholera toxin in the cystic fibrosis mouse model. *Science* 266:107-109.

Garrigan D, Mobasher Z, Kingan SB, Wilder JA, Hammer MF (2005) Deep haplotype divergence and long-range linkage disequilibrium at xp21.1 provide evidence that humans descend from a structured ancestral population. *Genetics* 170:1849-1856.

Gorczynski RM. (2005) CD200 and its receptors as targets for immunoregulation. *Curr Opin Investig Drugs* 6:483-8.

Hattori A, Tsujimoto M. (2004) Processing of antigenic peptides by aminopeptidases. *Biol Pharm Bull* 27:777-80.

Hedrick PW, Whittam TS, Parham P. (1991) Heterozygosity at individual amino acid sites: extremely high levels for HLA-A and -B genes. *Proc Natl Acad Sci USA* 88:5897-901.

Hudson RR. (2002) Generating samples under a Wright-Fisher neutral model of genetic variation. *Bioinformatics* 18:337-8.

Hughes AL, Yeager M (1998) Natural selection at major histocompatibility complex loci of vertebrates. *Annu Rev Genet* 32:415-435.

Katz HR. (2007) Inhibition of pathologic inflammation by leukocyte Ig-like receptor B4 and related inhibitory receptors. *Immunol Rev* 217:222-30.

Kindberg E *et al.* (2006) A nonsense mutation (428G-->A) in the fucosyltransferase FUT2 gene affects the progression of HIV-1 infection. *AIDS* 20:685-9.

Koda, Y., H. Tachida, M. Soejima, O. Takenaka, and H. Kimura. 2000. Ancient origin of the null allele se428. of the human ABO-secretor locus FUT2.. J Mol Evol **50**:243-248.

Lao O, Andrés AM, Mateu E, Bertranpetit J, Calafell F. (2003) Spatial patterns of cystic fibrosis mutation spectra in European populations. *Eur J Hum Genet* 11:385-94.

Lindesmith L *et al*. (2003) Human susceptibility and resistance to Norwalk virus infection. *Nat Med* 9:548-53.

Nielsen R *et al*. (2005) Genomic scans for selective sweeps using SNP data. *Genome Res* 15:1566-75.

Nielsen R, Hubisz MJ, Torgerson D, Andrés AM, Albrechtsen A, Gutenkunst R, Adams M, Cargill M, Boyko A, Indap A, Bustamante C, Hellmann IG, Clark AG. (2009) Darwinian and demographic forces affecting human protein coding genes. Genome Res. 2009 Mar 11. [Epub ahead of print]

Patterson N, Richter DJ, Gnerre S, Lander ES, Reich D (2006) Genetic evidence for complex speciation of humans and chimpanzees. *Nature* 441:1103-1108.

Plagnol V, Wall JD (2006) Possible ancestral structure in human populations. *PLoS Genet* 2:e105.

Przeworski M, Coop G, Wall JD (2005) The signature of positive selection on standing genetic variation. *Evolution Int J Org Evolution* 59:2312-2323.

Quinton PM. (1994) Human genetics. What is good about cystic fibrosis? *Curr Biol* 4:742-3.

Rijkers ES *et al*. (2008) The inhibitory CD200R is differentially expressed on human and mouse

T and B lymphocytes. *Mol Immunol* 45:1126-35.
Saitou N, Yamamoto F. (1997) Evolution of primate ABO blood group genes and their homologous genes. *Mol Biol Evol* 14:399-411.
Sánchez-Mazas A. (2007) An apportionment of human HLA diversity. *Tissue Antigens* 69 Suppl 1:198-202.
Saveanu L, *et al*. (2005) Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. *Nat Immunol* 6:689-97.
Soejima M, Pang H, Koda Y (2007) Genetic variation of FUT2 in a Ghanaian population: identification of four novel mutations and inference of balancing selection. *Ann Hematol* 86:199-204.
Taylor N *et al*. (2005) Enhanced tolerance to autoimmune uveitis in CD200-deficient mice correlates with a pronounced Th2 switch in response to antigen challenge. *J Immunol* 74:143-54.
Verrelli BC *et al*. (2002) Evidence for balancing selection from nucleotide sequence analyses of human G6PD. *Am J Hum Genet* 71:1112-28.
Wall JD (2000) Detecting ancient admixture in humans using sequence polymorphism data. *Genetics* 154:1271-1279.
Williamson S, Fledel-Alon A, Bustamante CD (2004) Population genetics of polymorphism and divergence for diploid selection models with arbitrary dominance. *Genetics* 168: 463-475.
Wright GJ *et al*. (2000) Lymphoid/neuronal cell surface OX2 glycoprotein recognizes a novel receptor on macrophages implicated in the control of their function. *Immunity* 13:233-42.
Wu-Baer F, Lane WS, Gaynor RB. (1995) The cellular factor TRP-185 regulates RNA polymerase II binding to HIV-1 TAR RNA. *EMBO J* 14:5995-6009.
Wu-Baer F, Lane WS, Gaynor RB. (1996) Identification of a group of cellular cofactors that stimulate the binding of RNA polymerase II and TRP-185 to human immunodeficiency virus 1 TAR RNA. *J Biol Chem* 271:4201-8.