

Your Trusted Supplier of *in vivo* MAbs  
BioCell  $\alpha$ -PD-1 ·  $\alpha$ -PD-L1 ·  $\alpha$ -CTLA-4 ·  $\alpha$ -CD20 ·  $\alpha$ -NK1.1 ·  $\alpha$ -IFNAR-1  
DISCOVER MORE



## Balancing Selection Is the Main Force Shaping the Evolution of Innate Immunity Genes

This information is current as of August 8, 2022.

Anna Ferrer-Admetlla, Elena Bosch, Martin Sikora, Tomàs Marquès-Bonet, Anna Ramírez-Soriano, Aura Muntasell, Arcadi Navarro, Ross Lazarus, Francesc Calafell, Jaume Bertranpetit and Ferran Casals

*J Immunol* 2008; 181:1315-1322; ;  
doi: 10.4049/jimmunol.181.2.1315  
<http://www.jimmunol.org/content/181/2/1315>

**References** This article **cites 58 articles**, 10 of which you can access for free at:  
<http://www.jimmunol.org/content/181/2/1315.full#ref-list-1>

Why *The JI*? [Submit online.](#)

- **Rapid Reviews!** 30 days\* from submission to initial decision
- **No Triage!** Every submission reviewed by practicing scientists
- **Fast Publication!** 4 weeks from acceptance to publication

\*average

**Subscription** Information about subscribing to *The Journal of Immunology* is online at:  
<http://jimmunol.org/subscription>

**Permissions** Submit copyright permission requests at:  
<http://www.aai.org/About/Publications/JI/copyright.html>

**Email Alerts** Receive free email-alerts when new articles cite this article. Sign up at:  
<http://jimmunol.org/alerts>

*The Journal of Immunology* is published twice each month by  
The American Association of Immunologists, Inc.,  
1451 Rockville Pike, Suite 650, Rockville, MD 20852  
Copyright © 2008 by The American Association of  
Immunologists. All rights reserved.  
Print ISSN: 0022-1767 Online ISSN: 1550-6606.



# Balancing Selection Is the Main Force Shaping the Evolution of Innate Immunity Genes<sup>1</sup>

Anna Ferrer-Admetlla,\* Elena Bosch,\*<sup>‡</sup> Martin Sikora,\* Tomàs Marquès-Bonet,<sup>2\*</sup> Anna Ramírez-Soriano,\* Aura Muntasell,<sup>†</sup> Arcadi Navarro,\*<sup>‡§¶</sup> Ross Lazarus,<sup>||</sup> Francesc Calafell,\*<sup>‡</sup> Jaume Bertranpetit,\*<sup>‡</sup> and Ferran Casals<sup>3\*</sup>

The evolutionarily recent geographic expansion of humans, and the even more recent development of large, relatively dense human settlements, has exposed our species to new pathogenic environments. Potentially lethal pathogens are likely to have exerted important selective pressures on our genome, so immunity genes can be expected to show molecular signatures of the adaptation of human populations to these recent conditions. While genes related to the acquired immunity system have indeed been reported to show traces of local adaptation, little is known about the response of the innate immunity system. In this study, we analyze the variability patterns in different human populations of fifteen genes related to innate immunity. We have used both single nucleotide polymorphism and sequence data, and through the analysis of interpopulation differentiation, the linkage disequilibrium pattern, and intrapopulation diversity, we have discovered some signatures of positive and especially balancing selection in these genes, thus confirming the importance of the immune system genetic plasticity in the evolutionary adaptive process. Interestingly, the strongest evidence is found in three *TLR* genes and *CD14*. These innate immunity genes play a pivotal role, being involved in the primary recognition of pathogens. In general, more evidences of selection appear in the European populations, in some case possibly related to severe population specific pressures. However, we also describe evidence from African populations, which may reflect parallel or long-term selective forces acting in different geographic areas. *The Journal of Immunology*, 2008, 181: 1315–1322.

Human populations have adapted to diverse pathogenic environments. Evidence of positive natural selection is relatively common in immunity genes compared with other functional gene classes, suggesting that infectious pathogens may have constituted one of the major selective forces during human history. In a recent review, Sabeti et al. (1) compiled the existing information on 91 genes showing signatures of either positive or balancing selection, or both, of which several are indeed related to immune function and to pathogen susceptibility. Genomic approaches have confirmed that immunity is one of the functional classes most enriched with positively selected genes (2,

3). Furthermore, balancing selection has been invoked to explain the variability patterns in some gene systems related to the immune response, such as the MHC in vertebrates (4, 5) or the Resistance genes in plants (6). In both cases, the persistence of ancient polymorphisms is explained by balancing selection-preventing fixation (7) and maintaining high genetic diversity levels within the species.

In addition to the evolutionary interest, the description of genomic regions under positive or balancing selection is a useful tool to identify regions that have been functionally relevant, and have thus played a substantial role in adaptation. In particular, in the case of genes related to the immune function, evolutionary approaches should help both to identify genes involved in the resistance to diseases still affecting human populations and to bring valuable sight into the molecular mechanisms of pathogenesis (8). While a number of studies have considered adaptive immunity genes, only a small number of studies have focused on the effects of natural selection on genes involved in innate immunity. Innate immunity constitutes the first line of defense against invading pathogens, and provides non-specific protection against a wide variety of organisms by recognizing generic molecular structures called pathogen-associated molecular patterns (PAMPs)<sup>4</sup> that are essential for the survival of many pathogens (9, 10). Pattern recognizing receptors play a key role in PAMP recognition, since their stimulation by PAMPs triggers the activation of other immunity genes and the inflammatory response. Pattern recognizing receptors can be secreted proteins or be expressed on the cell surface or in the cytosol, where they serve to recognize invariant molecular

\*Unitat de Biologia Evolutiva and <sup>†</sup>Unitat d'Immunopatologia Molecular, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, <sup>‡</sup>Centro de Investigación Biomédica en Red Epidemiología y Salud Pública, <sup>§</sup>Institució Catalana de Recerca i Estudis Avançats, and <sup>¶</sup>National Institute for Bioinformatics, Barcelona, Spain; and <sup>||</sup>Channing Laboratory, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115

Received for publication December 24, 2007. Accepted for publication May 9, 2008.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>1</sup> This work was supported by Genoma España (Proyectos Piloto del Centro Nacional de Genotipado), by Grants BFU2005-00243, BFU2006-15413-C02-01 and SAF-2007-63171 awarded by Ministerio de Educación y Ciencia (Spain), by the Direcció General de Recerca of Generalitat de Catalunya (Grup de Recerca Consolidat 2005SGR/00608). M.S. is supported by a Ph.D. fellowship from the Programa de becas Formació del Professorado Universitario del Ministerio de Educación y Ciencia, Spain (AP2005-3982).

<sup>2</sup> Current address: Department of Genome Sciences, University of Washington School of Medicine and the Howard Hughes Medical Institute, 1705 Northeast Pacific Street, Seattle, WA 98195.

<sup>3</sup> Address correspondence and reprint requests to Dr. Ferran Casals, Unitat de Biologia Evolutiva, Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Doctor Aiguader 88, 08003, Barcelona, Catalonia, Spain. E-mail address: ferran.casals@upf.edu

<sup>4</sup> Abbreviations used in this paper: PAMP, pathogen-associated molecular pattern; SNP, single nucleotide polymorphism; DC, dendritic cell; MBL, mannose-binding lectin; LD, linkage disequilibrium.

Copyright © 2008 by The American Association of Immunologists, Inc. 0022-1767/08/\$2.00

constituents of pathogens. Genes belonging to the *TLR* family play a key role in the induction of the immune response, recognizing viral nucleic acids and bacterial products that are often components of the cell wall (10). The *TLR* genes were first described in fruit flies and can even be found in phylogenetically diverse organisms including sponges (11). In humans, the *TLR* family includes eleven genes coding for integral membrane glycoproteins, and their ligand specificities are being exhaustively analyzed (12, 13).

The aim of this work is to study the role of innate immunity genes in the adaptation of human populations to the specific pathogenic environments. We analyzed single nucleotide polymorphism (SNP) and resequencing data on 15 genes related to this system, spanning nine genomic regions. In this study, we choose representative genes involved in different stages of the innate immune response. We included genes codifying for *TLR1*, *TLR6*, *TLR9*, *TLR10*, *CD14*, and *CD209* as a first-line molecules with the ability to directly recognize and signal to trigger an inflammatory response upon interaction with pathogens. They are expressed by professional APCs, such as dendritic cells (DCs) and macrophages in tissues, with an extremely high capacity to activate the innate immune response (14). The second category of genes includes *IL6* and *IL10*, representing inflammatory mediators produced by DCs and macrophages as a consequence of the signaling post ligation of the microbial sensors. These molecules are involved in the orchestration and regulation of the immune response and functionally they can be divided into proinflammatory (*IL6*) and anti-inflammatory (*IL10*) molecules. In a third category, we included the *CD1A\_E* cluster of genes, codifying for lipid-presenting molecules that contribute to the innate immune response triggered against the pathogen by a second-line mechanism. They require the pathogen-derived lipids to be processed in the interior of the infected cell to be associated to CD1 and presented to the effector cells of the immune system (15). Finally, we also analyzed the evolution pattern of *MBL2*, a gene codifying for the human collectin, mannose-binding lectin (MBL). With multiple carbohydrate-recognition domains, MBL is able to bind to sugar groups displayed on the surfaces of a wide range of microorganisms and activates the complement system. Thereby MBL provide first-line defense by participating in the pathogen clearance with no role in directly triggering the inflammatory immune response (16). Although far from exhaustive, these genes give a broad view on essential components and interactions in the human innate immune system.

The goal is to elucidate the selective pressures acting on these genes during modern human history, studying the genetic diversity pattern in different populations. To do this, we searched for signatures of natural selection through the analysis of the following: 1) interpopulation differentiation, since environmental pressures acting on a given population may lead to local adaptation, and thus, to a change in allele frequencies and an increase in between-population differences (17); 2) the comparison of linkage disequilibrium (LD) across populations, given that the action of positive selection in a given population may increase LD; and 3) the analysis of the variability at the intrapopulation level, which has precise expected values under the neutral evolution model; significant decreases in genetic diversity point to positive selection (selection favoring a particular allele), while significant increases are interpreted as the footprint of balancing selection, in which increased diversity per se has adaptive value.

Through these analysis it is possible to show whether innate immunity genes have been targets of positive or balancing selection as a consequence of the diverse pathogenic pressures existing

in the different geographic areas around the world or if, in contrast, due to the crucial role that these genes play in the immune response, purifying selection, to maintain mostly a unique variant, has been the main force shaping their diversity pattern, thus preventing them from accumulating mutations and differentiating across populations.

## Materials and Methods

### Samples

We used the 1,049 samples included in the HGDP-CEPH Human Genome Diversity Panel (18). DNA samples were grouped according to geographic and ethnic criteria into 39 different populations with a worldwide distribution. For some analysis, the 39 populations were grouped into seven continental regions (19). Two samples were excluded from the analysis because of their uncertain population origin (HGDP00770 and HGDP00980). One control sample (CEPH0001) was duplicated in each plate to check for reproducibility and genotyping quality, and 13 DNA samples were duplicated and were used as genotyping controls. At the end, the sample had a composition close to the 971 set proposed by Rosenberg (20).

### Single nucleotide polymorphisms

SNPs were selected from the HapMap Phase I (<http://www.hapmap.org>) and dbSNP (<http://www.ncbi.nlm.nih.gov/SNP>) databases. We have used two browsers in the selection process: PupaSNP (<http://pupasnp.bioinfo.ocha.fib.es>) and SNPbrowser (Applied Biosystems). We preferentially selected SNPs from HapMap with minor allele frequency >10%, followed by those SNPs reported by dbSNP build 125 with minor allele frequency >10%. SNPs in the gene region were selected at ~1 SNP/5 Kb density, with one non-synonymous SNP in *CD14*, one in *CD1C*, and one in *TLR1*. At the 5' and 3' regions, we selected one SNP up to 5 Kb, one up to 10 Kb, one up to 20 Kb, and one up to 30 Kb, adding one SNP up to 100–200 bp at 5' end and, another up to 1 Kb at 3' end when possible. For some of the analysis, contiguous genes were grouped in a single region: *CD1A*, *CD1B*, *CD1C*, *CD1D* and *CD1E* were clustered into the *CD1A\_E* region; and *TLR1*, *TLR6*, and *TLR10* into the *TLR1\_6\_10* region.

### Genotyping

Genotyping was performed using the ABI SNPLex technology (Applied Biosystems), which is based on an oligonucleotide ligation assay. Detection of the alleles is done using two competitive primers (one for each allele) joined to their ZIPcode (labeled nucleotidic fragment of particular length) and to a universal PCR primer. The combination of different lengths and fluorescent dyes for each ZIPcode allows allele discrimination in multiplex reactions (up to 48 SNPs in a single nucleotide reaction). Samples were typed in three assays of 384-well plates. Oligonucleotide ligation assay reactions require 1 ng of DNA for each SNP. Allele detection is achieved by capillary electrophoresis on ABI3730 (Applied Biosystems). Genotypes were assigned with GeneMapper software v3.5 (Applied Biosystems). One hundred out of 118 designed SNPs produced reliable genotypes. The rest (18 SNPs) were removed from the study (missing data below a reasonable threshold). Thus, the SNPLex conversion rate for our set of SNPs is 84.75%.

### SNP data analysis

Basic analysis of the genotyping data, including allele and genotype frequency calculations and Hardy-Weinberg equilibrium calculations were performed with SNPator (<http://bioinformatica.cegen.upf.es/public/principal/index.php>). Two SNPs (hCV7481405 and rs2228049) were removed from the study because they were not in Hardy-Weinberg equilibrium in most of the populations. The allele assignment of these SNPs was manually revised and it was determined that the absence of Hardy-Weinberg equilibrium was due to genotyping errors.  $F_{ST}$  values and a locus by locus analysis of molecular variance test were performed with the Arlequin software (21), using the 39 defined populations and the seven continental groups. Permutation tests were performed by subsampling each of the three datasets 10,000 times.

### Haplotype estimation and analysis

Haplotype frequencies were estimated using the Bayesian algorithm implemented in the PHASE 2 software (22) considering the criteria suggested by Graffelman et al. (23). Network 4.1.0.0 software was used to generate median-joining networks describing possible genealogical relationships

Table I. Population differentiation in the nine gene regions analyzed

Region	SNPs <sup>a</sup>	F <sub>ST</sub> <sup>b</sup>	F <sub>CT</sub> <sup>c</sup>	F <sub>SC</sub> <sup>d</sup>	F <sub>SC</sub> AME	F <sub>SC</sub> CSASIA	F <sub>SC</sub> EASIA	F <sub>SC</sub> EUR	F <sub>SC</sub> MENA	F <sub>SC</sub> OCE	F <sub>SC</sub> SSAFR
<i>CD14</i>	8	0.07	0.05	0.03	0.18	0.02	0.00	0.01	0.01	0.09	0.01
<i>CD14_E</i>	16	0.12	0.09	0.05	0.13	0.02	0.02	0.03	0.02	0.04	0.06
<i>CD209</i>	7	0.14	0.13	0.03	0.14	0.01	0.02	0.01	0.01	0.05	0.05
<i>IL6</i>	9	0.18	0.18	0.03	0.04	0.02	0.01	0.01	0.02	0.13	0.06
<i>IL10</i>	9	0.11	0.09	0.04	0.06	0.02	0.02	0.05	0.00	0.04	0.06
<i>MBL2</i>	9	0.14	0.14	0.03	0.11	0.01	0.01	0.02	0.04	0.03	0.03
<i>TLR1_6_10</i>	21	0.12	0.10	0.04	0.10	0.01	0.01	0.10	0.00	0.01	0.05
<i>TLR4</i>	11	0.13	0.12	0.03	0.06	0.02	0.04	0.00	0.01	0.15	0.03
<i>TLR9</i>	10	0.05	0.02	0.04	0.10	0.02	0.02	0.00	0.05	0.06	0.07
Average	–	0.12	0.10	0.03	0.10	0.02	0.02	0.03	0.02	0.07	0.05

<sup>a</sup> Successfully genotyped SNPs; <sup>b</sup> between the 39 populations; <sup>c</sup> between continental groups; <sup>d</sup> within continents F<sub>ST</sub>, F<sub>CT</sub>, and F<sub>SC</sub> values are the average of all the SNPs in each gene. SSAFR = Sub-Saharan Africa; MENA = Middle East-North Africa; EUR = Europe; CSASIA = Central-South Asia; EASIA = East Asia; OCE = Oceania; AME = America.

among haplotypes in terms of mutational differences (24). Linkage disequilibrium patterns and measures ( $D'$  and  $r^2$ ) were calculated using the Haploview 3.32 program (25).

### Sequence data and analysis

Resequencing data for nine of the fifteen genes analyzed in this work is available in public databases. Sequences for *CD14*, *CD209*, *IL10*, *TLR1*, *TLR4*, *TLR6*, *TLR9*, and *TLR10* were retrieved from the Innate Immunity Program in Genomics Applications (IIPGA, <http://innateimmunity.net>) (26), and sequences for *IL6* were retrieved from the SeattleSNPs database (<http://pga.mbit.washington.edu/>). The data are from the same 23 European-American and 24 African-American samples included in the Coriell CEPH/AA panel. All coding sequences were completely sequenced in both directions, but for genes greater than 20 Kb, some intronic sequences were not available. Multiple sequence alignments were performed with ClustalW (27). Diversity statistics and neutrality tests to detect signals of natural selection in resequencing data (Tajima's  $D$ , Fu and Li's  $F^*$  and  $D^*$ , Fu and Li's  $F$  and  $D$ , Fay and Wu's  $H$ , Fu's  $F_s$ ) were calculated using DnaSP 4.0 (28). The significance of these tests was calculated by means of coalescent simulations using the COSI software (29), using the model that takes into consideration the demographic history of humans. The demographic model is particularly useful in humans, since it includes information about the demographic processes that have differentially affected West African, European, and East Asian populations (29). Simple neutral models that do not include this information may be also useful as primary approaches to data, but due to the unrealistic demographic assumptions, may lead to wrong conclusions. In brief, 10,000 replicates were performed by using the local recombination rate estimate of each region obtained from HapMap.

## Results

Genotyping data for 100 SNPs from fifteen genes related to innate immunity were obtained for 1,049 individuals belonging to 39 populations representative of most of extant human population diversity (18). The populations from America and Oceania show a wider range of frequencies and a higher number of fixed alleles, leading to a lower level of heterozygosity. This is in agreement with the smaller size of these populations with a more important effect of genetic drift due to founder effects (30) and the small sample size in some populations.

### Differentiation among populations

F<sub>ST</sub> was used to measure the allele frequency differentiation among populations (31). We calculated this parameter among the 39 populations (F<sub>ST</sub>), among the seven continental groups (F<sub>CT</sub>), and among the populations included in the same continental group (F<sub>SC</sub>). Average values for each gene or gene region are shown in Table I. As expected, variability is mainly due to differences between continents, and there is relatively little difference between populations in the same continental groups (Table I). At the intracontinental level, America presents the highest differentiation among its populations, except for some

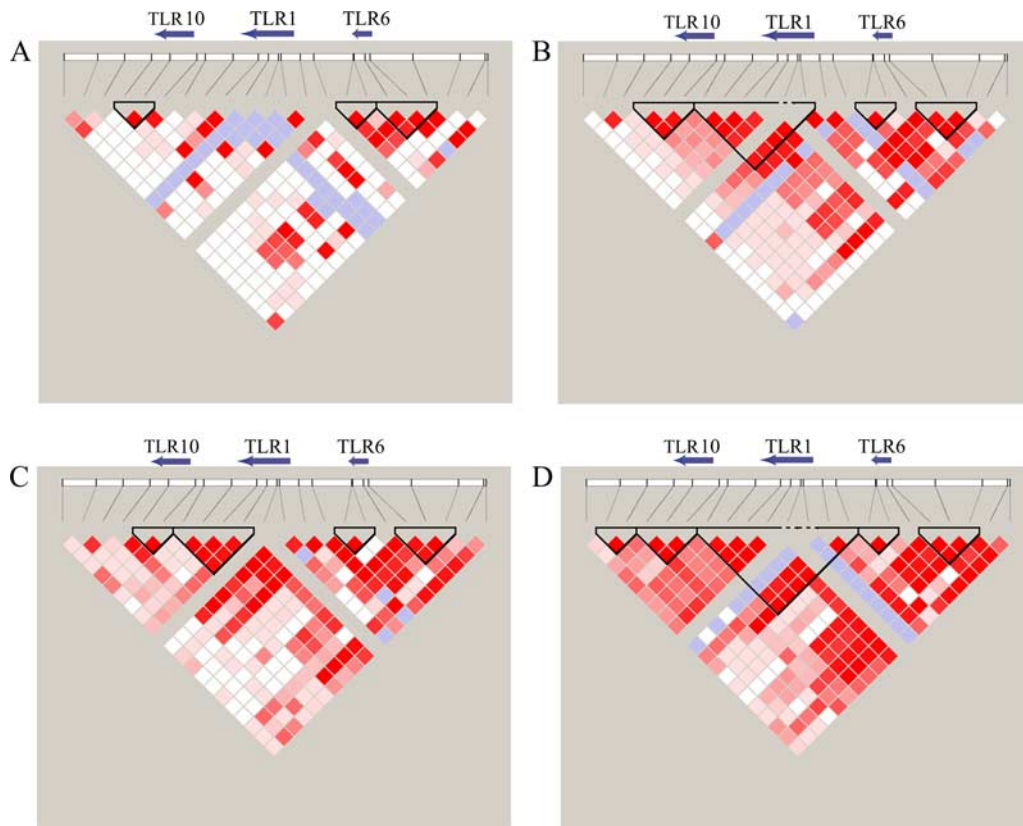
cases found in Oceanic populations. Again, this fact is consistent with the smaller size and the demographic history of these populations (30).

We interrogated our data in two ways to ascertain whether innate immunity genes showed a distinct pattern of differentiation among populations. First, F<sub>ST</sub> distribution for the selected genes was compared with that obtained in three different, independent, and comparable sets: SNPs in the Alfred database (<http://alfred.med.yale.edu/alfred/>) (32), SNPs from a gene free region on chromosome 22 (33), and an insertion/deletion set (34). These three sets have been obtained in either the same or similar sample sets to that used in this work. Although the innate immunity group of genes shows the lowest F<sub>ST</sub> value (mean F<sub>ST</sub> = 0.1096, against 0.1267, 0.1149, and 0.1182 for the Alfred database, the neutral region on chromosome 22, and the insertion/deletion set, respectively) these differences are not statistically significant either when data was compared through parametrical (ANOVA) and non-parametrical tests (Mann-Whitney and Kolmogorov-Smirnov). Only marginal significance ( $p = 0.014$ ) was obtained in a permutation test comparing our F<sub>ST</sub> values with those from the Alfred database.

Second, we checked whether some SNPs of a particular gene show unusually high F<sub>ST</sub> values, which could indicate the existence of different selective pressures on it (17). The highest average values were found in *IL6* (0.18) and *CD209* (0.14), and some relatively high F<sub>ST</sub> values at single SNPs (above 0.20) were found in *TLR1\_6\_10*, *TLR4* and the *IL6* region. However, only three SNPs (rs4804797 in *CD209* with F<sub>ST</sub> = 0.29; rs1800796 in *IL6* with F<sub>ST</sub> = 0.33; rs4833103 in *TLR1* with F<sub>ST</sub> = 0.34) showed values higher than the 95th percentile value for SNPs analyzed from the Alfred database (32). Aside from not being extremely high, the fact that these high F<sub>ST</sub> values were observed only in a few sparse SNPs makes the hypothesis that this observation is consequence of positive selection unreliable, and suggests that allele frequency differences among populations are more probably caused by random genetic drift (35). On the other hand, *TLR9* shows an unusually low average F<sub>ST</sub> (0.05). In the *TLR9* region, eight of the ten genotyped SNPs show F<sub>ST</sub> < 0.06, and the fraction of variation explained by differences among continental groups is only 1.8%, lower than the average value of ~10% (36, 37).

### LD pattern analysis

We compared the LD patterns for each gene region in the seven continental groups, through analysis of LD structure and LD



**FIGURE 1.** LD pattern in the *TLR1\_6\_10* region in Africa (A), East Asia (B), Central and South Asia (C), and Europe (D) continental groups. The approximate position of *TLR10*, *TLR1*, and *TLR6* is indicated. Arrows indicate the orientation of the genes. Vertical lines at the top correspond with the genotyped SNPs. Each square represents the magnitude of LD for a single pair of markers, with different intensities of red (bright red, LD = 1).

amount (38). The structure of LD is similar among the seven continental groups. To analyze this, we calculated the  $r^2$  values between all SNPs from the same gene region and computed the correlation between pairs of continental groups assessing the significance with Mantel's test. In general, the LD patterns in all continental groups are highly and significantly correlated, as expected by the underlying common genomic structure (data not shown).

To obtain a measure of the amount of LD, we used  $r^2$  between adjacent SNPs. First, we tested whether these values are similar in the seven continental groups by means of the Friedman test. Second, we conducted a pairwise comparison between continental groups using the Wilcoxon's test. Three regions show significant differences in the amount of LD found in each continental group: *CD1A\_E*, *TLR1\_6\_10* and *TLR9* (data not shown). In the case of *CD1A\_E*, a higher amount of LD was observed in Oceania,

whereas in the case of *TLR1\_6\_10*, the European populations showed a significantly higher LD level (Fig. 1). The *TLR9* region shows a more complex and diverse pattern, with most of the comparisons being statistically significant.

#### Full sequence variation

We analyzed public resequencing data available for nine of the genes included in this work (Table II). These nine regions have been resequenced in the same 24 African-American and 23 European-American Coriel samples (see *Materials and Methods*). Diversity indexes were calculated for each region in the two populations (Table II). The observed diversity values fall within the expected range, and are in agreement with the "out-of-Africa" model, which predicts that higher diversity values are expected in the African populations (39).

Table II. Diversity statistics of nine genes in African- and European-American populations<sup>a</sup>

	Africa					Europe			
	m	S	$\pi$	$\theta$	Hd	S	$\pi$	$\theta$	Hd
<i>CD14</i>	7,967	34	0.0010	0.0010	0.924	17	0.0008	0.0005	0.790
<i>CD209</i>	9,723	84	0.0018	0.0020	0.984	42	0.0009	0.0010	0.884
<i>IL6</i>	7,986	38	0.0008	0.0011	0.957	24	0.0008	0.0007	0.860
<i>IL10</i>	7,027	20	0.0008	0.0006	0.785	18	0.0009	0.0006	0.834
<i>TLR1</i>	10,310	71	0.0009	0.0016	0.958	34	0.0006	0.0008	0.742
<i>TLR4</i>	9,026	33	0.0006	0.0008	0.952	16	0.0003	0.0004	0.669
<i>TLR6</i>	6,751	44	0.0014	0.0015	0.926	24	0.0012	0.0008	0.756
<i>TLR9</i>	8,360	11	0.0003	0.0003	0.754	12	0.0003	0.0003	0.762
<i>TLR10</i>	7,096	70	0.0023	0.0022	0.895	54	0.0010	0.0017	0.483

<sup>a</sup> m = Sequence length; S = segregating sites;  $\pi$  = average number of nucleotide differences per site between two sequences;  $\theta$  = Watterson estimator; Hd = haplotype diversity.

Table III. Neutrality tests for nine genes in African- and European-American populations<sup>a</sup>

	African-American						
	Tajima's D	Fu and Li's D*	Fu and Li's F*	Fu and Li's D	Fu and Li's F	Fay and Wu's H	Fu's Fs
<i>CD14</i>	0.09	0.43	0.37	0.45	0.38	0.21	-1.49
<i>CD209</i>	-0.28	0.25	0.06	0.27	0.02	-4.30	-7.15
<i>IL6</i>	-0.99	-0.71	<b>0.96*</b>	-0.85	-1.09	0.22	-6.57
<i>IL10</i>	<b>0.97*</b>	0.22	0.57	0.20	0.58	0.56	-0.05*
<i>TLR1</i>	-1.52*	0.04	-0.63	0.00	-0.73	-4.78	-6.71
<i>TLR4</i>	-1.03	-1.07	-1.26	-1.26	-1.42	-1.36	-9.48
<i>TLR6</i>	-0.22	0.21	0.07	0.64	0.37	-0.63	-0.65
<i>TLR9</i>	-0.07	-0.27	-0.24	-0.32	-0.28	-0.86	-1.24
<i>TLR10</i>	0.11	0.52*	0.44*	0.53*	0.43*	-5.62	-0.40*

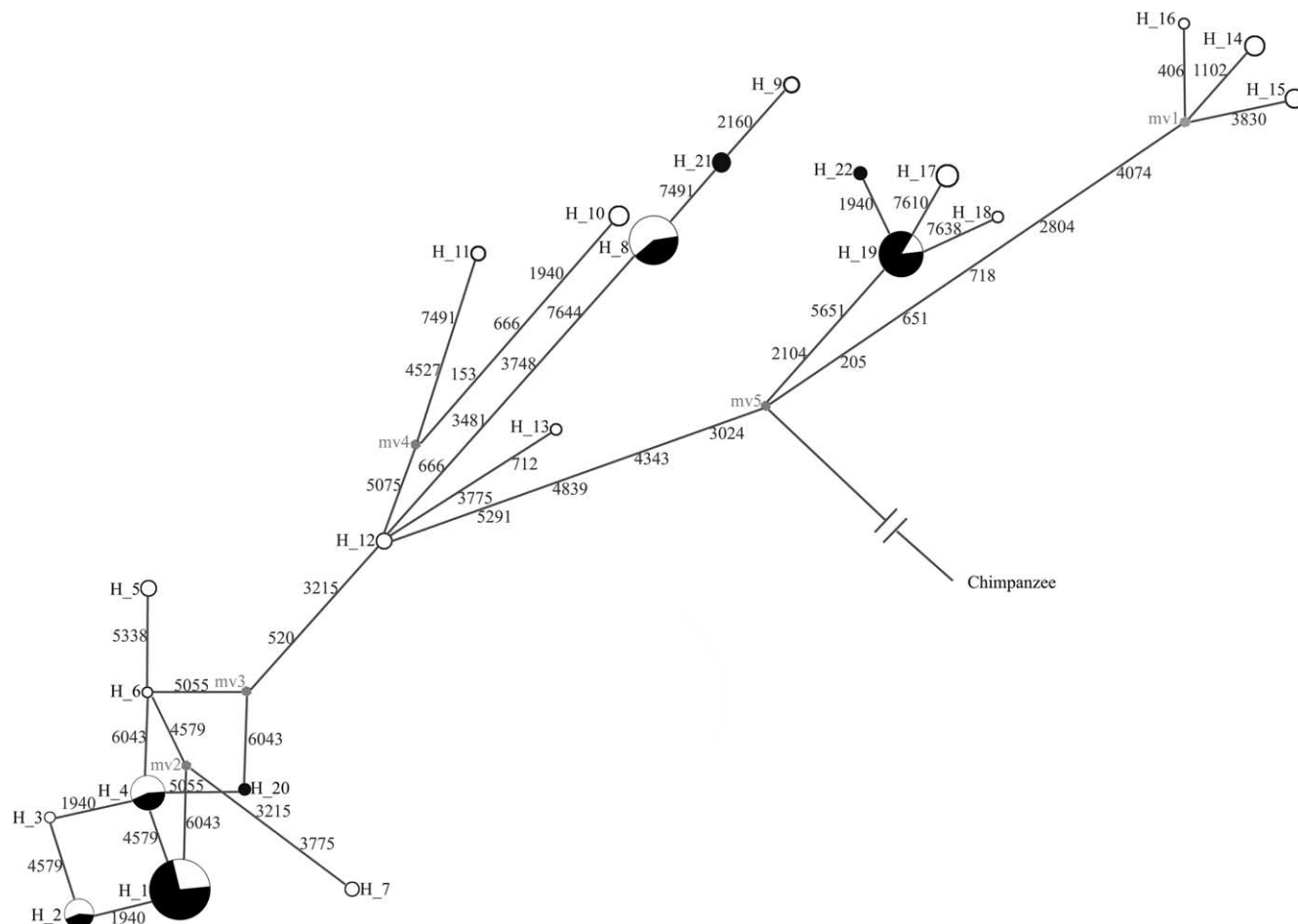
  

	European-American						
	Tajima's D	Fu and Li's D*	Fu and Li's F*	Fu and Li's D	Fu and Li's F	Fay and Wu's H	Fu's Fs
<i>CD14</i>	<b>2.20*</b>	<b>1.62**</b>	<b>2.15**</b>	1.73**	2.28**	0.05	5.08
<i>CD209</i>	-0.32	-0.05	-0.17	0.02	-0.11	-10.03*	-2.41
<i>IL6</i>	0.44	-0.13	0.08	-0.19	0.05	1.89	-0.53
<i>IL10</i>	<b>1.60*</b>	0.86	1.31	0.90	1.38	2.60**	3.56*
<i>TLR1</i>	-0.56	1.15*	0.65	1.52*	0.90	-9.32*	2.18
<i>TLR4</i>	-0.81	0.31	-0.08	0.30	-0.11	-2.64	-1.60
<i>TLR6</i>	<b>1.58*</b>	1.12	<b>1.52*</b>	1.21	1.62*	-1.87	2.66
<i>TLR9</i>	-0.64	<b>-2.23*</b>	-2.01	-2.44*	-2.18	-0.87	-2.65
<i>TLR10</i>	<b>-1.50*</b>	<b>-2.89*</b>	<b>-2.84*</b>	-1.30	-1.65	-23.41***	4.46*

<sup>a</sup> \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$  (see *Materials and Methods*). Tajima's D, Fu and Li D\* and F\*, and Fay and Wu's values greater or smaller to the 95 percentile of the genes analyzed in Seattle SNP are indicated in bold.

To search for signatures of positive or balancing selection, we performed several neutrality tests. In several regions, the results are significantly different to those expected under neutral evolution (Table III; see *Materials and Methods* for details on the calculation of statistical significance). *CD14* in the European population had the greatest higher statistical significance, indicating an excess of intermediate variants. Although the statistical significance is less conclusive, other loci including *TLR1*, *TLR6*, and *IL10* also show an excess of intermediate frequencies in this population. On the other hand, *TLR9* and especially *TLR10* show an excess of rare alleles in the European population. In the African population, *IL10* and *TLR10* show also an excess of intermediate frequencies, whereas *TLR1* show an excess of rare variants. However, in these three cases statistical significance is weaker, particularly when the number of tests performed is considered. Alternatively, significance of the results can be estimated. We compared the results obtained here to 132 genes included in the SeattleSNPs database, for which Tajima's D, Fu and Li's D\*, and F\* and Fay and Wu's H are available (40, 41). Several of the significant values obtained in our work for these four statistics are greater than the 95th percentile or smaller than the 5th percentile of the distribution for the SeattleSNPs genes (Table III), thus providing further evidence of true departures of neutrality. By far, the most striking evidence of natural selection acting on a single gene detected was in *CD14*. The obtained values for *CD14* are close to those for *ABO*, the gene showing the highest values in the neutrality tests in the SeattleSNPs database, and which has been extensively studied and proposed to be under balancing selection (42). Fig. 2 shows the median-joining network of the *CD14* region, which describes the most plausible mutational relationships between inferred haplotypes. Consistent with the results obtained in the neutrality tests (Table III), the phylogenetic structure shown by the European population consists of three well defined groups of haplotypes separated by a large number of mutations. This pattern, together with the neutrality test results, strongly suggests that this region has

been under balancing selection. Alternatively, the excess of intermediate frequencies detected at the *CD14* region could also have been originated by a hitchhiking effect with balancing selection acting in a close gene. To test this hypothesis, we analyzed the LD pattern of this region in the European population using the data from the International HapMap project ([www.hapmap.org](http://www.hapmap.org)). Although *CD14* is in a region of high LD, especially in the European population, *CD14* and its 5' region is located in a small LD block of ~8 kb, located between two larger LD blocks. Moreover, a recombination hotspot has been described from SNP genotyping data (Perlegen) to be at ~3.5 kb of the *CD14* 5' end, and the nearest gene to the *CD14* 3' end is located at 63 kb. All these data suggest that the observed diversity pattern at the *CD14* locus is relatively independent from its genomic context, giving further support to the hypothesis of balancing selection at *CD14* itself as the main force to explain this pattern. Among the 17 variable nucleotides described in *CD14* in the European population, only one (7491) is located in the coding region of *CD14*. However, this is a synonymous substitution only present in two chromosomes, and differentiates two close haplotypes. Three more substitutions (1940, 4579, and 5055) separate haplotypes within one of the main groups, and 13 differentiate the three main groups. Three potential regulatory regions have been predicted in the 5' region by detecting conserved regions from alignments of five mammal species, and three of the substitutions (3748, 5391, and 6043) fall within these regions. The position 6043 corresponds to the previously described polymorphism CD14/-159. This mutation is located at the promoter of the gene, and has been related with higher CD14 concentrations as a consequence of increased transcription (43, 44). Although more recent analyses have questioned the functional importance of this SNP (45, 46), our results of balancing selection suggest the presence of more than one functional form of CD14, even if it is still unclear which polymorphisms are indeed functional.



**FIGURE 2.** Median-joining network of *CD14* haplotypes in African (white) and European (black) populations. The circle areas are proportional to the frequency of the haplotypes. Haplotype number is shown next to circles and mutations are indicated on the branches linking the haplotypes. Recurrent mutations have been described at positions 666, 1940, 3215, 3775, 4579, 5055, 6043, and 7491; of those, positions 666, 3215, 5055, and 6043 are in hypermutable CpG motifs. The chimpanzee *CD14* sequence is included as an ancestral haplotype.

## Discussion

Our results strongly suggest that the innate immune system has been evolving under forces other than demography and purifying selection. Several of the genes included in our study clearly show departures from neutrality through the different analyses used, with different levels of statistical significance.

Two genes, *TLR9* and *TLR10*, show evidence of positive selection in Europeans. The *TLR1\_6\_10* region shows a higher amount of LD in Europeans (see *Results*), and this difference is mainly due to the region containing *TLR10* (Fig. 1). Moreover, neutrality tests also show signatures of positive selection in *TLR10* in European-Americans (Table III). *TLR9* also showed some significant results, although their statistical significance was less convincing. In addition, *TLR9* is located in a high LD and gene-rich region, and in absence of other evidence, it cannot be determined whether the observed diversity pattern is a consequence of evolutionary forces acting on it, or the effect of a hitchhiking with nearby genes.

The most interesting point is the large number of genes showing an excess of intermediate variants, which is compatible with the action of balancing selection. We have found signatures of balancing selection in six out of the nine analyzed genes. The evidence is stronger in the European population, where *CD14*, *TLR1*, *TLR6*, and *IL10* show significant results in more than one neutrality test, with the case of *CD14* being extremely evident. In the African-American population, significant results have been ob-

tained for *IL6*, *IL10*, and *TLR10*, although evidence is weaker. In previous work, signatures of balancing selection have been identified in the promoter region of *IL10* in Europeans, although based on only marginal statistical significance (47). Similar evidence of balancing selection at this locus has been found in our study in both European and African populations. For the rest of the genes, this is the first published evidence of balancing selection. This large number of genes showing evidence of balancing selection may be due to the great diversity of ligand-binding affinities and functions (10, 12, 13). In the case of *IL10*, its diversity pattern has been attributed to conflicting forces acting on the promoter region of this locus, because although high levels of this protein protect against pathological effects from the inflammatory response, low levels favor the elimination of pathogens (47).

One intriguing point is the differences observed between the African and the European populations, with more cases of balancing selection and more statistical significance described in Europeans. Balancing selection needs a long time, probably a number of generations greater to those separating these two populations, to produce a detectable signal, except in cases of strong selective pressures (48, 49). In fact, some population-specific cases of strong balancing selection are well known, as those of *HBB* and thalasemias related with malaria (50). Given that genes analyzed in this study are involved in the immune response, this explanation may account for some of the cases reported here. However, since

significant cases have also been described in Africa, the most parsimonious explanation for the data is widespread balancing selection and mostly detected in Europeans because of the strength of the selective pressures. Therefore, we hypothesize that polymorphism patterns in our set of genes have been mostly shaped by recurrent episodes of balancing selection and/or partial selective sweeps in response, likely, to epidemics. Infectious agents might have been different in different places and times and might have acted with varying strengths. That would account for the general positization of the values of most statistics in most populations (explained by more ancestral selective episodes), and the very significant and positive values in certain populations (explained by balancing selection acting upon a frequency spectrum that was already biased by older episodes). Finally, other explanations implying a lack of accuracy of the demographic model in which simulations are based cannot be completely excluded for non-extreme cases.

While the effect of local positive selection increasing interpopulation differentiation is widely accepted, the effect of balancing selection is far less clear (48, 49), and it has been thought to generate low levels of population differentiation, since it maintains the genetic variation across all the populations (17). In previous works, this effect has been proposed to occur at the 5' regions of *CCR5* and *IL10* (47, 51). In those reports, an excess of intermediate variants is accompanied by extremely low  $F_{ST}$  values. However, this effect should only be produced when the same selective forces are acting on similar genetic variability background in the different populations. Therefore, such a reduction of interpopulation differentiation is not expected in all regions showing evidence of balancing selection, as would be the case for different pathogens inducing different selective forces. Alternatively, the relatively low level of interpopulation differences among the genes analyzed in this work can be explained by their relationship with the innate immune system rather than with the adaptive immune system. In contrast with the adaptive system, the innate immune system provides inborn protection against a wide variety of Ags by recognizing PAMPs, which are molecular structures present only in the pathogens and not in the host organism, and are shared by a wide spectrum of microorganisms (52). Therefore, the innate immunity activity is not specific for particular infectious agents, but is generic and effective for large groups of pathogens (53). Given that, selective forces may be acting similarly on these genes in different populations, and therefore they are expected to show high levels of conservation. However, it has to be pointed out that the innate and adaptive parts of the immune function are not two completely independent systems and in some cases, like *IL10*, genes are directly involved in both responses pathways (54).

Interestingly, all genes with evidences of balancing selection code for proteins with a strong impact in the magnitude of the inflammatory immune response. TLR1 and TLR6 function as first-line sensors for pathogens that are capable of signaling through NF- $\kappa$ B to trigger an inflammatory response (55). IL10 is a cytokine with a strong ability to modulate the development of the inflammatory response. CD14 is an adaptor molecule that binds LPS and promotes innate immunity to Gram-negative bacteria. It is important to point out that the protein encoded by CD14, the gene under the strongest prediction of balancing selection, is required to trigger the TLR4 signaling cascade. CD14 exists as a membrane-bound and soluble protein that functions by binding LPS and presenting it to the signaling complex TLR4/MD2 (56). Depending on the functional and clinical setting in which CD14 function has been analyzed, certain alleles have been associated with susceptibility to infections (57), or with the development of autoimmunity and allergies (58). These situations would reflect an inappropiate

signaling through TLR4 leading to a detrimental immune response for the host.

Mounting an inflammatory response is important for effective immunity. In contrast, resolution of inflammation is essential for maintaining the balance between health and disease. The idea that the inflammatory response must be finely tuned is supported by a multiplicity of experimental data. Too strong a response causes the various adverse effects associated with infectious and autoimmune diseases (14, 55), while a weak inflammatory response attenuates the subsequent immune response (59). Thus, the tight and fine tuning required to develop and maintain a proper inflammatory response could be pointed out as the force determining the balanced selection of alleles observed in these genes. Interactions with pathogens on one side, and the development of pathogenic processes such as autoimmunity on the other side, could be the balancing pressures favoring the diversity of alleles present in the populations for these genes.

In conclusion, the fact that the diversity pattern described in most of the genes included in this study is compatible with the action of positive, and specially balancing selection, confirms the importance of pathogenic environments shaping variation in the human genome. The strongest evidence we found supports the action of balancing selection in TLR1, TLR6, and mainly CD14, and the action of positive selection in TLR10. These receptors play a pivotal role in the innate immunity system, which shows that despite its reliance on generic pathogenic patterns, it is sufficiently plastic to permit adaptation to new environmental conditions. Interestingly, apart from TLR9, where the detected signals of selection cannot be unequivocally attributed to this gene, TLR4 is the only member of the TLR receptors included in this work that do not present evidence of positive or balancing selection. A possible explanation to this observation is that, in this case, selective pressures have acted on its presenting LPS protein, CD14. Finally, these selective forces have operated in similar ways in the different human populations (and thus of environments and pathogens), so clearly the adaptation is not population (and thus, pathogen) specific, but very general to early stage exposures.

## Acknowledgments

SNP genotyping services were provided by the Spanish "Centro Nacional de Genotipado" (CEGEN; [www.cegen.org](http://www.cegen.org)) and an informatic SNP analysis platform was supplied by the Spanish "Instituto Nacional de Bioinformática" (INB; [www.inab.org](http://www.inab.org)). We thank A. Blancher for helpful advice during gene selection process.

## Disclosures

The authors have no financial conflict of interest.

## References

- Sabeti, P. C., S. F. Schaffner, B. Fry, J. Lohmueller, P. Varilly, O. Shamovsky, A. Palma, T. S. Mikkelsen, D. Altshuler, and E. S. Lander. 2006. Positive natural selection in the human lineage. *Science* 312: 1614–1620.
- Bustamante, C. D., A. Fedel-Alon, S. Williamson, R. Nielsen, M. T. Hubisz, S. Glanowski, D. M. Tanenbaum, T. J. White, J. J. Sninsky, R. D. Hernandez, et al. 2005. Natural selection on protein-coding genes in the human genome. *Nature* 437: 1153–1157.
- Nielsen, R., C. Bustamante, A. G. Clark, S. Glanowski, T. B. Sackton, M. J. Hubisz, A. Fedel-Alon, D. M. Tanenbaum, D. Civello, T. J. White, et al. 2005. A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol.* 3: e170.
- Bernatchez, L., and C. Landry. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *J. Evol. Biol.* 16: 363–377.
- Garrigan, D., and P. W. Hedrick. 2003. Perspective: detecting adaptive molecular polymorphism: lessons from the MHC. *Evolution Int. J. Org. Evolution* 57: 1707–1722.
- Bergelson, J., M. Kreitman, E. A. Stahl, and D. Tian. 2001. Evolutionary dynamics of plant R-genes. *Science* 292: 2281–2285.
- Kreitman, M., and A. Di Rienzo. 2004. Balancing claims for balancing selection. *Trends Genet.* 20: 300–304.



8. Bamshad, M., and S. P. Wooding. 2003. Signatures of natural selection in the human genome. *Nat. Rev. Genet.* 4: 99–111.
9. Janeway, C. A., Jr., and R. Medzhitov. 2002. Innate immune recognition. *Annu. Rev. Immunol.* 20: 197–216.
10. Medzhitov, R., and C. Janeway, Jr. 2000. The Toll receptor family and microbial recognition. *Trends Microbiol.* 8: 452–456.
11. Wiens, M., M. Korzhev, S. Perovic-Ottstadt, B. Luthringer, D. Brandt, S. Klein, and W. E. Muller. 2007. Toll-like receptors are part of the innate immune defense system of sponges (demospongiae: Porifera). *Mol. Biol. Evol.* 24: 792–804.
12. Akira, S., and K. Takeda. 2004. Toll-like receptor signalling. *Nat. Rev. Immunol.* 4: 499–511.
13. Akira, S., S. Uematsu, and O. Takeuchi. 2006. Pathogen recognition and innate immunity. *Cell* 124: 783–801.
14. Anders, H. J., D. Zecher, R. D. Pawar, and P. S. Patole. 2005. Molecular mechanisms of autoimmunity triggered by microbial infection. *Arthritis Res. Ther.* 7: 215–224.
15. Barral, D. C., and M. B. Brenner. 2007. CD1 antigen presentation: how it works. *Nat. Rev. Immunol.* 7: 929–941.
16. Dommert, R. M., N. Klein, and M. W. Turner. 2006. Mannose-binding lectin in innate immunity: past, present and future. *Tissue Antigens* 68: 193–209.
17. Akey, J. M., G. Zhang, K. Zhang, L. Jin, and M. D. Shriver. 2002. Interrogating a high-density SNP map for signatures of natural selection. *Genome Res.* 12: 1805–1814.
18. Cann, H. M., C. de Toma, L. Cazes, M. F. Legrand, V. Morel, L. Piouffre, J. Bodmer, W. F. Bodmer, B. Bonne-Tamir, A. Cambon-Thomsen, et al. 2002. A human genome diversity cell line panel. *Science* 296: 261–262.
19. Gardner, M., A. Gonzalez-Neira, O. Lao, F. Calafell, J. Bertranpetit, and D. Comas. 2006. Extreme population differences across Neuregulin 1 gene, with implications for association studies. *Mol. Psychiatry* 11: 66–75.
20. Rosenberg, N. A. 2006. Standardized subsets of the HGDP-CEPH human genome diversity cell line panel, accounting for atypical and duplicated samples and pairs of close relatives. *Ann. Hum. Genet.* 70: 841–847.
21. Excoffier, L., G. Laval, and S. Schneider. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47–50.
22. Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68: 978–989.
23. Graffelman, J., D. J. Balding, A. Gonzalez-Neira, and J. Bertranpetit. 2007. Variation in estimated recombination rates across human populations. *Hum. Genet.* 122: 301–310.
24. Bandelt, H. J., P. Forster, and A. Rohlf. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16: 37–48.
25. Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263–265.
26. Lazarus, R., D. Vercelli, L. J. Palmer, W. J. Klimecki, E. K. Silverman, B. Richter, A. Riva, M. Ramoni, F. D. Martinez, S. T. Weiss, and D. J. Kwiatkowski. 2002. Single nucleotide polymorphisms in innate immunity genes: abundant variation and potential role in complex human disease. *Immunol. Rev.* 190: 9–25.
27. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22: 4673–4680.
28. Rozas, J., J. C. Sanchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496–2497.
29. Schaffner, S. F., C. Foo, S. Gabriel, D. Reich, M. J. Daly, and D. Altshuler. 2005. Calibrating a coalescent simulation of human genome sequence variation. *Genome Res.* 15: 1576–1583.
30. Rosenberg, N. A., J. K. Pritchard, J. L. Weber, H. M. Cann, K. K. Kidd, L. A. Zhivotovsky, and M. W. Feldman. 2002. Genetic structure of human populations. *Science* 298: 2381–2385.
31. Wright, S. 1969. The theory of gene frequencies. In *Evolution and the Genetics of Populations*. University of Chicago Press, Chicago.
32. Rajeevan, H., M. V. Osier, K. H. Cheung, H. Deng, L. Druskin, R. Heinzen, J. R. Kidd, S. Stein, A. J. Pakstis, N. P. Tosches, et al. 2003. ALFRED: the ALlele FREquency database: update. *Nucleic Acids Res.* 31: 270–271.
33. Gonzalez-Neira, A., F. Calafell, A. Navarro, O. Lao, H. Cann, D. Comas, and J. Bertranpetit. 2004. Geographic stratification of linkage disequilibrium: a worldwide population study in a region of chromosome 22. *Hum. Genomics* 1: 399–409.
34. Weber, J. L., D. David, J. Heil, Y. Fan, C. Zhao, and G. Marth. 2002. Human diallelic insertion/deletion polymorphisms. *Am. J. Hum. Genet.* 71: 854–862.
35. Gardner, M., S. Williamson, F. Casals, E. Bosch, A. Navarro, F. Calafell, J. Bertranpetit, and D. Comas. 2007. Extreme individual marker F(ST) values do not imply population-specific selection in humans: the *NRG1* example. *Hum. Genet.* 121: 759–762.
36. Barbujani, G., A. Magagni, E. Minch, and L. L. Cavalli-Sforza. 1997. An apportionment of human DNA diversity. *Proc. Natl. Acad. Sci. USA* 94: 4516–4519.
37. Romualdi, C., D. Balding, I. S. Nasidze, G. Risch, M. Robichaux, S. T. Sherry, M. Stoneking, M. A. Batzer, and G. Barbujani. 2002. Patterns of human diversity, within and among continents, inferred from biallelic DNA polymorphisms. *Genome Res.* 12: 602–612.
38. Ramirez-Soriano, A., O. Lao, M. Soldevila, F. Calafell, J. Bertranpetit, and D. Comas. 2005. Haplotype tagging efficiency in worldwide populations in *CTLA4* gene. *Genes Immun.* 6: 646–657.
39. Cavalli-Sforza, L. L., and M. W. Feldman. 2003. The application of molecular genetic approaches to the study of human evolution. *Nat. Genet.* 33 (Suppl.): 266–275.
40. Akey, J. M., M. A. Eberle, M. J. Rieder, C. S. Carlson, M. D. Shriver, D. A. Nickerson, and L. Kruglyak. 2004. Population history and natural selection shape patterns of genetic variation in 132 genes. *PLoS Biol.* 2: e286.
41. Stajich, J. E., and M. W. Hahn. 2005. Disentangling the effects of demography and selection in human history. *Mol. Biol. Evol.* 22: 63–73.
42. Fry, A. E., M. J. Griffiths, S. Auburn, M. Diakite, J. T. Forton, A. Green, A. Richardson, J. Wilson, M. Jallow, F. Sisay-Joof, et al. 2008. Common variation in the *ABO* glycosyltransferase is associated with susceptibility to severe *Plasmodium falciparum* malaria. *Hum. Mol. Genet.* 17: 567–576.
43. Baldini, M., I. C. Lohman, M. Halonen, R. P. Erickson, P. G. Holt, and F. D. Martinez. 1999. A polymorphism\* in the 5' flanking region of the *CD14* gene is associated with circulating soluble CD14 levels and with total serum immunoglobulin E. *Am. J. Respir. Cell Mol. Biol.* 20: 976–983.
44. LeVan, T. D., J. W. Bloom, T. J. Bailey, C. L. Karp, M. Halonen, F. D. Martinez, and D. Vercelli. 2001. A common single nucleotide polymorphism in the *CD14* promoter decreases the affinity of Sp protein binding and enhances transcriptional activity. *J. Immunol.* 167: 5838–5844.
45. Liang, X. H., W. Cheung, C. K. Heng, J. J. Liu, C. W. Li, B. Lim, and Y. Wang de. 2006. *CD14* promoter polymorphisms have no functional significance and are not associated with atopic phenotypes. *Pharmacogenet. Genomics* 16: 229–236.
46. von Aulock, S., J. Rupp, K. Gueinzus, M. Maass, and C. Hermann. 2005. Critical investigation of the *CD14* promoter polymorphism: lack of a role for in vitro cytokine response and membrane CD14 expression. *Clin. Diagn. Lab. Immunol.* 12: 1254–1256.
47. Wilson, J. N., K. Rockett, B. Keating, M. Jallow, M. Pinder, F. Sisay-Joof, M. Newport, and D. Kwiatkowski. 2006. A hallmark of balancing selection is present at the promoter region of interleukin 10. *Genes Immun.* 7: 680–683.
48. Barton, N. H., and A. Navarro. 2002. Extending the coalescent to multilocus systems: the case of balancing selection. *Genet. Res.* 79: 129–139.
49. Navarro, A., and N. H. Barton. 2002. The effects of multilocus balancing selection on neutral variability. *Genetics* 161: 849–863.
50. Kwiatkowski, D. P. 2005. How malaria has affected the human genome and what human genetics can teach us about malaria. *Am. J. Hum. Genet.* 77: 171–192.
51. Bamshad, M. J., S. Mummidi, E. Gonzalez, S. S. Ahuja, D. M. Dunn, W. S. Watkins, S. Wooding, A. C. Stone, L. B. Jorde, R. B. Weiss, and S. K. Ahuja. 2002. A strong signature of balancing selection in the 5' cis-regulatory region of *CCR5*. *Proc. Natl. Acad. Sci. USA* 99: 10539–10544.
52. Medzhitov, R., and C. Janeway, Jr. 2000. Innate immune recognition: mechanisms and pathways. *Immunol. Rev.* 173: 89–97.
53. Kimbrell, D. A., and B. Beutler. 2001. The evolution and genetics of innate immunity. *Nat. Rev. Genet.* 2: 256–267.
54. Mege, J. L., S. Meghari, A. Honstetter, C. Capo, and D. Raoult. 2006. The two faces of interleukin 10 in human infectious diseases. *Lancet. Infect Dis.* 6: 557–569.
55. Papadimitrakaki, E. D. 2007. Toll like receptors and autoimmunity: a critical appraisal. *J. Autoimmun.* 29: 310–318.
56. Mivake, K. 2006. Links Roles for accessory molecules in microbial recognition by Toll-like receptors. *J. Endotoxin Res.* 12: 195–204.
57. Yuan, F. F., K. Marks, M. Wong, S. Watson, E. de Leon, P. B. McIntyre, and J. S. Sullivan. 2008. Clinical relevance of TLR2, TLR4, CD14 and FcγRIIA gene polymorphisms in *Streptococcus pneumoniae* infection. *Immunol. Cell Biol.* 86: 268–270.
58. Vercelli, D. 2002. The functional genomics of *CD14* and its role in IgE responses: an integrated view. *J. Allergy Clin. Immunol.* 109: 14–21.
59. Misch, E. A., and T. R. Hawn. 2008. Toll-like receptor polymorphisms and susceptibility to human disease. *Clin. Exp. Immunol.* 114: 347–360.