Convergent evolution in European and Roma populations reveals pressure exerted by plague on Toll-like receptors

Hafid Laayouni^{a,1}, Marije Oosting^{b,c,1}, Pierre Luisi^a, Mihai Ioana^{b,d}, Santos Alonso^e, Isis Ricaño-Ponce^f, Gosia Trynka^{f,2}, Alexandra Zhernakova^f, Theo S. Plantinga^{b,c}, Shih-Chin Cheng^{b,c}, Jos W. M. van der Meer^{b,c}, Radu Popp^g, Ajit Sood^h, B. K. Thelmaⁱ, Cisca Wijmenga^f, Leo A. B. Joosten^{b,c}, Jaume Bertranpetit^{a,3}, and Mihai G. Netea^{b,c,3,4}

^aInstitut de Biologia Evolutiva (Consejo Superior de Investigaciones Científicas–Universitat Pompeu Fabra), Universitat Pompeu Fabra, 08003 Barcelona, Spain; ^bDepartment of Medicine and ^cNijmegen Institute for Infection, Inflammation and Immunity, Radboud University Nijmegen Medical Centre, 6525 GA, Nijmegen, The Netherlands; ^dUniversity of Medicine and Pharmacy Craiova, 200349 Craiova, Romania; ^eDepartment of Genetics, Physical Anthropology and Animal Physiology, University of the Basque Country, Barrio Sarriena s/n, 48940 Leioa, Spain; ^fDepartment of Genetics, University of Groningen/University Medical Center Groningen, 9700 RB, Groningen, The Netherlands; ^gDepartment of Medical Genetics, “Iuliu Hatieganu” University of Medicine and Pharmacy, 400023 Cluj-Napoca, Romania; ^hDepartment of Gastroenterology, Dayanand Medical College and Hospital, Ludhiana, Punjab 141001, India; and ⁱDepartment of Genetics, University of Delhi South Campus, New Delhi 110 021, India

Edited* by Charles A. Dinarello, University of Colorado Denver, Aurora, CO, and approved January 2, 2014 (received for review September 19, 2013)

Recent historical periods in Europe have been characterized by severe epidemic events such as plague, smallpox, or influenza that shaped the immune system of modern populations. This study aims to identify signals of convergent evolution of the immune system, based on the peculiar demographic history in which two populations with different genetic ancestry, Europeans and Roma (Gypsies), have lived in the same geographic area and have been exposed to similar environments, including infections, during the last millennium. We identified several genes under evolutionary pressure in European/Romanian and Roma/Gypsy populations, but not in a Northwest Indian population, the geographic origin of the Roma. Genes in the immune system were highly represented among those under strong evolutionary pressures in Europeans, and infections are likely to have played an important role. For example, Toll-like receptor 1 (*TLR1*)/*TLR6*/*TLR10* gene cluster showed a strong signal of adaptive selection. Their gene products are functional receptors for *Yersinia pestis*, the agent of plague, as shown by overexpression studies showing induction of proinflammatory cytokines such as TNF, IL-1 β , and IL-6 as one possible infection that may have exerted evolutionary pressures. Immunogenetic analysis showed that *TLR1*, *TLR6*, and *TLR10* single-nucleotide polymorphisms modulate *Y. pestis*-induced cytokine responses. Other infections may also have played an important role. Thus, reconstruction of evolutionary history of European populations has identified several immune pathways, among them *TLR1*/*TLR6*/*TLR10*, as being shaped by convergent evolution in two human populations with different origins under the same infectious environment.

immunity | pattern recognition receptors | pandemics | migration

By recognition and elimination of pathogenic microorganisms during infection, the immune system has allowed mankind to survive. Genetic variation in the immune system is a major factor influencing susceptibility to infections. Subsequently, genes of the immune system are under constant evolutionary pressure (1), and this pressure can change based on local conditions and migration routes of human populations (2).

In time, changes induced in the immune system by infectious pressures can shape not only the host defense and susceptibility to infections but also susceptibility to autoimmune or inflammatory diseases of modern human populations (2), with balancing selection proposed as a main force shaping the innate immunity reaction (3). It has been suggested that a predominantly proinflammatory profile in the immune system, induced by infections, predisposes modern human populations to autoimmune diseases (4, 5), whereas selection of certain genetic variants during epidemics [e.g., selection of C-C chemokine receptor type 5 (*CCR5*) variants presumably by plague] reduces susceptibility to HIV

infection in modern Europeans compared with Africans (6). All these studies have investigated candidate genes selected on the basis of biological assumptions, but comprehensive genome-wide approaches to identify the immune pathways under evolutionary pressure by infections are missing.

In this study, we make use of the opportunity that a special historical demographic situation is present in Europe—that is, ancient European populations living together with Roma in the same geographic locations. Roma (traditionally called Gypsies) are a population from Northwest India that has migrated in Europe one millennium ago (7). We hypothesized that despite their different ethnic and genetic backgrounds, the strong infectious pressure exerted by the major epidemics of the last millennium (of which epidemics of plague are probably the most significant) has led to convergent evolution: specific immune genes, selected during these European epidemics, become signatures that differ from

Significance

This article gives a unique perspective on the impact of evolution on the immune system under pressure by infections, using the special demographic history of Europe in which two populations with different genetic ancestry, Europeans and Roma (Gypsies), have lived in the same geographic area and have been exposed to similar environmental hazards, including infections. We identified convergent evolution signals in genes from different human populations. Reconstruction of evolutionary history of European populations has identified Toll-like receptor 1 (*TLR1*)/*TLR6*/*TLR10* as a pattern recognition pathway shaped by convergent evolution by infections, among which plague is a likely cause, influencing the survival of these populations during the infection.

Author contributions: H.L., J.W.M.v.d.M., A.S., B.K.T., C.W., L.A.B.J., J.B., and M.G.N. designed research; H.L., M.O., P.L., M.I., S.A., I.R.-P., G.T., A.Z., T.S.P., S.-C.C., R.P., A.S., and L.A.B.J. performed research; M.O., P.L., M.I., S.A., I.R.-P., G.T., A.Z., T.S.P., S.-C.C., and R.P. contributed new reagents/analytic tools; H.L., M.O., P.L., M.I., S.A., I.R.-P., G.T., A.Z., T.S.P., S.-C.C., R.P., and M.G.N. analyzed data; and H.L., M.O., M.I., S.A., J.W.M.v.d.M., A.S., B.K.T., C.W., L.A.B.J., J.B., and M.G.N. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹H.L. and M.O. contributed equally to this work.

²Present address: Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115; and Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, MA 02142.

³J.B. and M.G.N. share senior authorship.

⁴To whom correspondence should be addressed. E-mail: Mihai.Netea@radboudumc.nl.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1317723111/-DCSupplemental.

Table 1. Genes with extreme values of XP-CLR statistic and TreeSelect test, indicative of putative signals of positive selection

Genes	Chromosome	Test	Populations
SLC45A2, ADAMTS12, AMACR, RXFP3	chr5	XP-CLR TreeSelect	Roma and Romanians vs. Indians All
TLR1, TLR6, TLR10, FAM114A1	chr4	XP-CLR TreeSelect	Roma and Romanians vs. Indians Roma and Romanians
FBXL19, SETD1A, STX1B, STX	chr16	TreeSelect	Indians
BTNL2, HLA-DRA	chr6	TreeSelect	Roma and Romanians
ANK3	chr1	TreeSelect	Roma and Romanians
BAZ1A, SRP54	chr14	XP-CLR	Romanians vs. Indians
KCNK10	chr14	XP-CLR	Roma vs. Indians
NEK7	chr1	XP-CLR	Roma vs. Romanians
Ataxin2	chr12	XP-CLR	Romanians vs. Indians

Genes that appear in the same row belong to the same chromosomal regions and are in a linkage disequilibrium block. Using XP-CLR statistic, the interest is in genes with signals in Romanian compared with Indians, and in Rroma compared with Indians, but not present in Romanian compared with Rroma; for TreeSelect, the interests are signals in Rroma and Romanians but not Indians. We report other cases even if they do not fulfill the above criteria and are not of direct interest to this study.

values of undoubted value for the present framework. Notably, this SNP (intergenic between TLR1 and TLR6) was reported to be associated with an expression quantitative trait loci of the expression of three genes, TLR1, TLR6, and TLR10, in lymphoblastoid cell lines (LCLs) (13).

We also performed an additional analysis using genotype data from the Illumina Omni 2.5M Chip for the 1000 Genome Project for individuals (14) in an Indian (Gujarati) and European (Northern Europeans from Utah, CEU) population. XP-CLR statistic was used to detect selection in this Indian population. Results show that there is clear signal of selection in the European population (CEU) compared with the Indian (Gujarati) population, but no signal of selection was detected in this Indian population compared with the European population (Fig. S2 A and B).

Interestingly, the other gene cluster detected (first row in Table 1), with four genes in chromosome 5, contains the well-known gene *SLC45A2*, described as being under positive selection in relation to skin pigmentation in Europeans (14). Other strong signals are for the *BTNL2* gene locus in chromosome 6 coming from the TreeSelect test in Rroma and Romanian populations. This gene is highly polymorphic, with homology to the butyrophilin gene family, and is located at the border of the major histocompatibility complex (MHC) class I and class II regions in humans. This signal of positive selection may be due to the role of MHC in adaptation to pathogens in human history. Many other strong signals are shown in Fig. 2 A and B, however these signals are specific to one single population or show differentiation between Rroma and Romanians and cannot be caused by a convergent adaptation of the same evolutionary process in these two populations.

Most of the signals found in this study cluster in regions of the genome with a high linkage disequilibrium (Fig. S3 A–C for TLR group, cluster containing *SLC45A2* gene and cluster containing the *BTNL2* gene). This finding makes it difficult to pinpoint the exact target of selection in each case, a general problem of selection studies (15). Clearly, genes in the TLR1/6/10 cluster are of special interest for the present study.

Fig. 2. Manhattan plot of results of selection tests in Rroma, Romanians, and Indians using TreeSelect statistic (A) and XP-CLR statistic (B). Chromosomes ordered from chromosome 1 to chromosome 22.

TLR2 Cluster Genes Are Involved in the Recognition of *Yersinia pestis*.

TLR2 recognition of V-antigen and LcrV of *Y. pestis* is the main recognition mechanism during plague. TLR2 forms heterodimers with receptors of the same gene cluster (TLR1/TLR6) for recognition of bacterial lipopeptides (16), but it is not known whether TLR2 also collaborates with TLR10 for the recognition of *Y. pestis*. We transfected HEK cells (that normally express TLR1 and TLR6) with TLR2, TLR10, or TLR2 and TLR10. The HEK cells transfected with TLR2 alone release significantly more cytokines than untransfected cells: twofold more for *Y. pestis* and fivefold more for *Yersinia pseudotuberculosis*, the microorganisms from which *Y. pestis* evolved (Fig. 3A). Although TLR10 by itself is not able to induce cytokine production, cotransfection of TLR10 with TLR2 completely abrogates the stimulatory effect of TLR2 (Fig. 3A). These data were supported by blocking TLR2 in monocytes using monoclonal antibodies (Fig. 3B–D). Interestingly, blocking TLR10 resulted in an increase in cytokine production (Fig. 3B–D), supporting the observation that TLR10 has a modulatory effect, thus corroborating the overexpression experiments.

The modulatory effects of TLR10 seem to be exerted specifically on TLR2 signaling, as anti-TLR10 antibodies modulated cytokine production induced by palmitoyl-3-cysteine-serine-lysine-4, but not by the TLR4 agonist LPS (Fig. S4). Moreover, when cells of individuals carrying the SNP in TLR10 were exposed to either LPS, Poly I:C, CpG, or flagellin, no differences between the

Table 2. Statistical overrepresentation test of PANTHER analysis

Groups	No. genes in the database	No. genes in the dataset	Expected no. genes in the dataset	P value
By biological process				
Cytokine-mediated signaling pathway	184	3	0.31	0.0028
Visual perception	209	3	0.35	0.0040
Neurological system process	830	5	1.38	0.0062
System process	920	5	1.53	0.0097
Sensory perception	326	3	0.54	0.0139
Immune system process	1,036	5	1.72	0.0162
Signal transduction	1,642	6	2.73	0.0266
Cell communication	1,730	6	2.87	0.0344
Cell surface receptor linked signal transduction	846	4	1.41	0.0387
By molecular function				
Racemase and epimerase activity	14	1	0.02	0.0230
Receptor activity	779	4	1.29	0.0294
Transporter activity	24	1	0.04	0.0392

Biological process and molecular function enrichment for genes showing signals of selection in Roma and Romanians.

groups could be detected (Fig. S5). Interestingly, however, cross-linking of TLR10 receptors inhibited the IL-6 induction by IL-1 (Fig. S6), suggesting that TLR10 may exert inhibitory effects on the IL-1 family of cytokines (17).

Common TLR1, TLR6, and TLR10 Polymorphisms in European Populations Modulate Cytokine Responses to *Y. pestis*. To demonstrate that *TLR1*, *TLR6*, and *TLR10* genetic variation in the population modulates the response to *Y. pestis*, we isolated peripheral blood mononuclear cells (PBMCs) from a group of 101 individuals of European descent and exposed them to the pathogen. SNPs in *TLR1*, *TLR6*, and *TLR10* significantly influenced cytokine production induced by *Y. pestis* and *Y. pseudotuberculosis* (Fig. 4 and Fig. S7). In contrast, known polymorphisms in *TLR4* (Asp299Gly and Thr399Ile) did not influence the response of PBMCs to *Y. pestis* or *Y. pseudotuberculosis* (Fig. S8).

Discussion

In this study, we identified a set of genes evolving under positive selection in populations of different ethnic ancestry living in Europe, but not in Northwest India. Among these genes, the region encompassing *TLR1*, *TLR6*, and *TLR10* is under selection in Europeans/Romanians and Roma/Gypsies, but not in a population from Northwest India. The common selection pressures in the Romanians and Roma may be interpreted as the same evolutionary process induced by local infectious conditions in two European populations of different genetic backgrounds. To look for more evidence on positive selection in European populations, we analyzed sequence data from the 1000 Genome Project (18). These data show a clear selective sweep in Europeans using two methods based on genetic differentiation and extended linkage disequilibrium haplotype [cross-population extended haplotype homozygosity (XP-EHH) and XP-CLR]. This signal was specific

Fig. 3. The role of TLR10 for the recognition of *Y. pestis* and *Y. pseudotuberculosis*. (A) HEK293 transiently transfected with TLR2, TLR10, or TLR2/10, and stimulated with 1×10^5 heat-inactivated *Y. pestis* or *Y. pseudotuberculosis*, respectively. Bars represent the means $\pm$ SEM of at least three separate experiments. (B) PBMCs stimulated with *Y. pestis* or *Y. pseudotuberculosis* per mL. $n = 6$; means $\pm$ SEM; * $P = 0.05$, ** $P = 0.01$. (C) TNF- α production after PBMCs stimulated with *Y. pestis* or *Y. pseudotuberculosis* in the presence or absence of 10 μ g/mL antibody. (D) IL-1 β production after 24 h of stimulation. Means $\pm$ SEM; * $P = 0.05$, ** $P = 0.01$. The data shown are from three independent experiments each performed in duplicate.

Fig. 4. Functional consequences of human *TLR1/TLR6/TLR10* SNPs for *Y. pestis*-stimulated cytokine production. PBMCs from healthy volunteers stimulated with different stimuli, including *Y. pestis* (1×10^5 /mL). Volunteers were separated into three groups: one group did not display the SNP in either *TLR1* (A/B), *TLR6* (C/D), or *TLR10* (E/F; wt, wild-type); one group was heterozygous for the polymorphism (He); and one group was homozygous (Ho). Data are means $\pm$ SEM. * $P = 0.05$, ** $P = 0.01$, *** $P = 0.001$.

in Europeans and absent in an African population (Yoruba) and in a Chinese population (Fig. S9).

Besides the *TLR2* gene cluster, other genes of interest include (i) a gene cluster with four genes in chromosome 5 that contains the well-known gene *SLC45A2* being under positive selection in relation to skin pigmentation; (ii) *FBXL19*, a gene known to be involved in the modulation of inflammation (19) in a cluster comprising three genes; and (iii) *ADAMTS12* gene, which is associated with susceptibility to autoimmune diseases (20). In the same cluster as the *SLC45A2* gene, other genes (Table 1) may be of special interest to be analyzed functionally in the future.

Linguistic and genetic studies suggested that the Roma population left India in the 5–10th centuries and started to settle in Europe during the 11th century (21). Genetic studies, focused on uniparental and Mendelian disease markers, confirmed Roma as an isolated population of Indian origin among the European majority (7). We pose that after the Roma migration, the infectious pressures to which the Roma were exposed were the same as for the Europeans, whereas for the ancestral North Indian population, they remain linked to their geographical location in India. This peculiar demographic situation in Europe, in which populations with different genetic backgrounds have been exposed for a long period to similar infection pressures, gave us the opportunity to attempt the reconstruction of recent

evolutionary events acting on the immune system of populations living in Europe.

An important question is which evolutionary pressures were common to the Romanian and Roma populations. Infections are likely to have been one of the most important evolutionary forces shaping the immune system in both Europe and India, and several candidates may be considered. An infection often associated with evolutionary effects in Europeans is plague, responsible for several large epidemics with death rates of up to 30–50% of the European population and lingering thereafter in Europe for several centuries (22), thus allowing for the exertion of selective sweeps. Based on this extreme burden of mortality, it is rational to hypothesize that plague had major evolutionary effects on the immune system of European populations. The *TLR/IL-1* functional cluster is crucial for host defense against *Y. pestis*: *TLR2* and its coreceptors *TLR1*, *TLR6*, and *TLR10* are the main pattern recognition receptors for *Y. pestis*—all localized in a single gene cluster in chromosome 4 (23), whereas *Y. pestis* *Caf1* protein is an inhibitor of *IL-1β* (24). Decreased *IL-1* responses, either through defective *TLR* signaling or release of *Caf1*, are likely to have deleterious effects on host survival. The data presented here show that the *TLR1/TLR6/TLR10* receptor cluster has been under positive selection in both Romanians and Roma, and suggest that plague is a potential infection that has exerted this selection. Our data are also supported by an earlier study that identified the *TLR1/TLR6/TLR10* gene cluster as a target of recent positive selection in non-Africans (25). We confirmed the functional impact of *TLR1*, *TLR6*, and *TLR10* polymorphisms currently present in Europeans for the immune responses to *Y. pestis*.

Although evolutionary pressure exerted by plague is a plausible cause of adaptive selection, it should be emphasized that other infections in which the receptors of the *TLR2* cluster play a central role, such as tuberculosis, leprosy, or common Gram-positive pathogens, could have also contributed to the genetic pattern observed here. Nevertheless, these infections have a generally less restricted geographical pattern as common in India as in Europe. Importantly, the impact of historical plagues in India has been a matter of debate. Out of the three main outbreaks of plague (6–7th centuries, 14th century, and turn of 19–20th century), by far the most devastating is the second, called the Black Death. This outbreak is known not to have affected India (26) and took place after the settlement of Roma in Europe. Indeed, the Indian subcontinent may have been the only part of Eurasia to have experienced steady population growth during the last half of the 14th century, and the first reports of plague are from the 17th century, with much less impact than the Black Death. During the epidemics in the Indian subcontinent, the disease behaved differentially than plague in the 14th century in Europe, with less than 5% human mortality. It is likely that the absence of the flea *Xenopsylla cheopis* due to tropical environment and the distance and geographical barriers could have prevented the entrance of the devastating outbreak of the Middle Ages into India (26).

The identification of the immune pathways and genetic variants that were specifically selected in Europe not only helps us to understand the evolutionary history of European populations, but also contributes to our understanding of the differences in susceptibility between European and other populations to modern human diseases. Evolutionary pressure exerted by plague or smallpox has been previously proposed to partly explain the increased resistance to HIV in Europeans (6). In addition, the evolution toward a proinflammatory profile induced by infections during history might explain the burden of autoimmune diseases in modern human populations (27). Genetic variation in *TLR7* and *TLR8* has been shown to protect against viral infections (25), while predisposing some to autoimmune diseases (4). Similarly, *TLR1* or *TLR10* polymorphisms can protect against infections, while being associated with autoimmune diseases such as sarcoidosis (28) and Crohn's disease (29). Although the differences in cytokine production induced by *Y. pestis* in individuals with various *TLR1*, *TLR6*, or *TLR10* polymorphisms are moderate from an immunological point of

view, they are large from an evolutionary perspective, and can lead in the long term to significant shifts in the population. It should be realized that we may not have detected other genes relevant for host defense that may be under selective pressure, as they have not been included in the Illumina immunochip array, and only future studies using genome-wide sequencing have the capacity to provide an exhaustive analysis of the entire genome.

In conclusion, by comparing genes under selection in European/Romanian and Roma/Gypsy populations, we identified several immunological pathways specifically shaped by evolutionary processes in populations living together in Europe during the last millennium. It is likely that the selection pressure at least on some of these genes has been exerted by plague epidemics, and we identify the TLR1/TLR6/TLR10 pattern recognition system as a likely candidate.

Methods

Populations. After informed consent was obtained, blood was collected from 100 individuals of European/Romanian descent and 100 individuals of a Roma/Gypsy ethnic background. A population of 500 individuals of North Indian descent, representing the geographic origin of the Roma/Gypsy group, was also recruited. Healthy Dutch individuals were recruited for cytokine stimulations (21–73 y old, 73% males and 27% females).

Immunochip Arrays and Analysis of Genetic Distances Between Populations.

Samples were genotyped on immunochip custom array at the Department of Genetics, University Medical Center Groningen, The Netherlands (8). To explore genetic relationships among the populations, we used PCA as implemented in the Eigensoft package (9). For a detailed description of the methods, see *SI Methods*.

Evolutionary Models. A selective sweep induces a fast spread of the beneficial allele through the population until it reaches fixation. Through hitchhiking, the selected allele carries with it neutral alleles in the linked genomic region. Thus, in comparison with the neutral expectation, one expects to observe within a region that has evolved recently under positive selection a dramatic pattern of genetic differentiation among populations within an extended genomic region. Taking advantage of these theoretical expectations, we applied two methodologies, XP-CLR (10) and TreeSelect (11) tests, to identify the genomic region under putative selection in European/Romanian and the Roma/Gypsy populations, but not in the population from North India. We focused our study on population differentiation because it has been described

as the best molecular pattern to study very recent events of positive selection after haplotype structure (30). However, the design of the immunochip with very variable SNP density across the genome does not allow us to properly study the haplotype structure (for phasing issues and haplotype informativeness differences among regions with different SNP density). We used tests that are amenable to SNP data (and thus with ascertainment bias). For an extensive description of the XP-CLR and TreeSelect tests, please consult *SI Methods*.

TLR Cloning and Transfection. TLR cloning and transfection of human embryonic kidney 293 cells that were stably transfected with hTLR2 (293-hTLR2; kindly provided by Dr. D. T. Golenbock, University of Massachusetts Medical Center, Worcester, MA) are described in detail in *SI Methods*.

Cytokine Stimulation. PBMCs were isolated after obtaining informed consent (31). PBMCs (5×10^5) in 100 μ L volume were added to round-bottom 96-well plates (Greiner) and incubated with stimuli for 24 h at 37 °C and 5% CO₂. Cytokines were measured using specific sandwich ELISA kits for IL-1 β and TNF- α (R&D Systems). IL-6, IL-8, and IL-10 were measured using PeliKine Compact ELISA kits (Sanquin).

Immunogenetic Studies. DNA was isolated from whole blood using the Gentra Pure Gene Blood kit (Qiagen), and genotype assessments of the TLR10-N241H, TLR1-N248S, and TLR6-S249P SNPs were performed using a pre-designed TaqMan SNP genotyping assay (Applied Biosystems). The software automatically plotted genotypes based on a two-parameter plot with an overall success rate of >95%. Cycling conditions were 2 min at 50 °C and 10 min at 95 °C, followed by 40 cycles of 95 °C for 15 s and 1 min at 60 °C. Fluorescence intensities were corrected using a postread/pre-read method for 1 min at 60 °C before and after the amplification.

ACKNOWLEDGMENTS. We thank Dr. Vandana Midha for recruitment of the Indian study cohort. We also thank the National Institute of Bioinformatics (www.inab.org) for computational support. M.G.N. and C.W. were supported by Vici grants of the Netherlands Organization of Scientific Research. This work was funded by Grant BFU2010-19443 (to J.B.) from the Ministerio de Ciencia y Tecnología (Spain) and the Direcció General de Recerca, Generalitat de Catalunya (Grup de Recerca Consolidat 2009 SGR 1101). P.L. was supported by a PhD fellowship from "Acció Estratègica de Salut, en el Marco del Plan Nacional de Investigación Científica, Desarrollo e Innovación Tecnológica 2008–2011" from Instituto de Salud Carlos III. B.K.T. was supported by Grant BT/01/COE/07/UDSC from the Department of Biotechnology, Government of India, New Delhi.

- Barreiro LB, Quintana-Murci L (2010) From evolutionary genetics to human immunology: How selection shapes host defense genes. *Nat Rev Genet* 11(1):17–30.
- Netea MG, Wijmenga C, O'Neill LA (2012) Genetic variation in Toll-like receptors and disease susceptibility. *Nat Immunol* 13(6):535–542.
- Ferrer-Admetlla A, et al. (2008) Balancing selection is the main force shaping the evolution of innate immunity genes. *J Immunol* 181(2):1315–1322.
- Stene LC, et al. (2006) Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: A longitudinal study. *Am J Gastroenterol* 101(10):2333–2340.
- Zhernakova A, et al.; Finnish Celiac Disease Study Group (2010) Evolutionary and functional analysis of celiac risk loci reveals SH2B3 as a protective factor against bacterial infection. *Am J Hum Genet* 86(6):970–977.
- Stephens JC, et al. (1998) Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes. *Am J Hum Genet* 62(6):1507–1515.
- Mendizabal I, et al. (2012) Reconstructing the population history of European Romani from genome-wide data. *Curr Biol* 22(24):2342–2349.
- Trynka G, et al.; Spanish Consortium on the Genetics of Celiac Disease (CEGEC); PreventCD Study Group; Wellcome Trust Case Control Consortium (WTCCC) (2011) Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet* 43(12):1193–1201.
- Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genet* 2(12):e190.
- Chen H, Patterson N, Reich D (2010) Population differentiation as a test for selective sweeps. *Genome Res* 20(3):393–402.
- Bhatia G, et al. (2011) Genome-wide comparison of African-ancestry populations from CARE and other cohorts reveals signals of natural selection. *Am J Hum Genet* 89(3):368–381.
- Mi H, Muruganujan A, Thomas PD (2013) PANTHER in 2013: Modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res* 41(Database issue):D377–D386.
- Grundberg E, et al.; Multiple Tissue Human Expression Resource (MuTHER) Consortium (2012) Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat Genet* 44(10):1084–1089.
- Lao O, de Grijter JM, van Duijn K, Navarro A, Kayser M (2007) Signatures of positive selection in genes associated with human skin pigmentation as revealed from analyses of single nucleotide polymorphisms. *Ann Hum Genet* 71(Pt 3):354–369.
- Akey JM (2009) Constructing genomic maps of positive selection in humans: Where do we go from here? *Genome Res* 19(5):711–722.
- Akira S, Uematsu S, Takeuchi O (2006) Pathogen recognition and innate immunity. *Cell* 124(4):783–801.
- Mantovani A, Locati M, Polentarutti N, Vecchi A, Garlanda C (2004) Extracellular and intracellular decoys in the tuning of inflammatory cytokines and Toll-like receptors: The new entry TIR8/SIGIRR. *J Leukoc Biol* 75(5):738–742.
- Abecasis GR, et al.; 1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491(7422):56–65.
- Zhao J, et al. (2012) F-box protein FBXL19-mediated ubiquitination and degradation of the receptor for IL-33 limits pulmonary inflammation. *Nat Immunol* 13(7):651–658.
- Nah SS, et al. (2012) Association of ADAMTS12 polymorphisms with rheumatoid arthritis. *Mol Med Rep* 6(1):227–231.
- Fraser A (1992) *The Gypsies* (Blackwell, Oxford).
- McEvedy C (1988) The bubonic plague. *Sci Am* 258(2):118–123.
- Takeuchi O, et al. (2002) Cutting edge: Role of Toll-like receptor 1 in mediating immune response to microbial lipoproteins. *J Immunol* 169(1):10–14.
- Abramov VM, et al. (2001) Structural and functional similarity between Yersinia pestis capsular protein Caf1 and human interleukin-1 beta. *Biochemistry* 40(20):6076–6084.
- Barreiro LB, et al. (2009) Evolutionary dynamics of human Toll-like receptors and their different contributions to host defense. *PLoS Genet* 5(7):e1000562.
- Sussman GD (2011) Was the black death in India and China? *Bull Hist Med* 85(3):319–355.
- Di Rienzo A (2006) Population genetics models of common diseases. *Curr Opin Genet Dev* 16(6):630–636.
- Veltkamp M, van Moorsel CH, Rijkers GT, Ruven HJ, Grutters JC (2012) Genetic variation in the Toll-like receptor gene cluster (TLR1-TLR1-TLR6) influences disease course in sarcoidosis. *Tissue Antigens* 79(1):25–32.
- Abad C, et al. (2011) Association of Toll-like receptor 10 and susceptibility to Crohn's disease independent of NOD2. *Genes Immun* 12(8):635–642.
- Sabeti PC, et al.; International HapMap Consortium (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature* 449(7164):913–918.
- Oosting M, et al. (2011) TLR1/TLR2 heterodimers play an important role in the recognition of Borrelia spirochetes. *PLoS ONE* 6(10):e25998.