Promoter Variation in the DC-SIGN–Encoding Gene *CD209* Is Associated with Tuberculosis

Luis B. Barreiro[1,2], Olivier Neyrolles[2], Chantal L. Babb[3], Ludovic Tailleux[2], Hélène Quach[1], Ken McElreavey[4], Paul D. van Helden[3], Eileen G. Hoal[3], Brigitte Gicquel[2], Lluís Quintana-Murci[1*]

1 CNRS FRE2849, Unit of Molecular Prevention and Therapy of Human Diseases, Institut Pasteur, Paris, France, 2 Unité de Génétique Mycobactérienne, Institut Pasteur, Paris, France, 3 Faculty of Health Sciences, Stellenbosch University, Tygerberg, South Africa, 4 Reproduction, Fertilité et Populations, Institut Pasteur, Paris, France

ABSTRACT

Background

Tuberculosis, which is caused by *Mycobacterium tuberculosis,* remains one of the leading causes of mortality worldwide. The C-type lectin DC-SIGN is known to be the major *M. tuberculosis* receptor on human dendritic cells. We reasoned that if DC-SIGN interacts with *M. tuberculosis,* as well as with other pathogens, variation in this gene might have a broad range of influence in the pathogenesis of a number of infectious diseases, including tuberculosis.

Methods and Findings

We tested whether polymorphisms in *CD209,* the gene encoding DC-SIGN, are associated with susceptibility to tuberculosis through sequencing and genotyping analyses in a South African cohort. After exclusion of significant population stratification in our cohort, we observed an association between two *CD209* promoter variants (−871G and −336A) and decreased risk of developing tuberculosis. By looking at the geographical distribution of these variants, we observed that their allelic combination is mainly confined to Eurasian populations.

Conclusions

Our observations suggest that the two −871G and −336A variants confer protection against tuberculosis. In addition, the geographic distribution of these two alleles, together with their phylogenetic status, suggest that they may have increased in frequency in non-African populations as a result of host genetic adaptation to a longer history of exposure to tuberculosis. Further characterization of the biological consequences of DC-SIGN variation in tuberculosis will be crucial to better appreciate the role of this lectin in interactions between the host immune system and the tubercle bacillus as well as other pathogens.

Introduction

One-third of the world's population is estimated to be infected with *Mycobacterium tuberculosis*, the etiological agent of tuberculosis (TB). This disease tops the World Health Organization list of deaths due to a single infectious agent, with the death toll between 2 and 3 million people per year [1]. A perplexing, and yet unsolved, feature of TB is that less than 10% of infected individuals develop the disease. Substantial epidemiological evidence supports that host-related factors, such as sex, age, HIV infection, malnutrition, and BCG (bacille Calmette-Guérin) vaccination, influence the balance between the tubercle bacilli and host immune defences [2,3]. In addition, there is increasing evidence that host genetic factors determine differences in host susceptibility to mycobacterial infection and might contribute therefore to the pattern of clinical disease [4–7]. From a host perspective, the innate immunity system acts as the first line of host defense against microbial pathogens [8]. Initial recognition of pathogens by the innate immunity system is mediated by phagocytic cells, such as dendritic cells (DCs) or macrophages, through germ-line-encoded receptors, known as pattern recognition receptors [9]. DCs bear a range of pattern recognition receptors, such as C-type lectins and Toll-like receptors, involved both in recognition of conserved products of microbial metabolism and in the induction of adaptive immunity [8,10–12]. In particular, C-type lectins detect pathogens by their characteristic carbohydrate structures and internalise them for further antigen processing and presentation [13]. We have recently shown that a prototypic C-type lectin, DC-SIGN (dendritic cell–specific ICAM-3 grabbing nonintegrin), is the major *Mycobacterium tuberculosis* receptor on human DCs [14]. DC-SIGN is specifically, though not exclusively, expressed on DCs and functions both as a cell adhesion and as a pathogen recognition receptor [15]. As an adhesion receptor, it plays an important role in many DC functions, such as DC-T cell interaction and DC migration [16,17]. Besides its cellular recognition role, DC-SIGN serves as pathogen uptake receptor and mediates interactions with a plethora of pathogens other than *M. tuberculosis* [18]. Indeed, it has been shown that DC-SIGN allows DCs to capture other bacteria such as *Helicobacter pylori* and certain *Klebsiella pneumonia* strains, but also viruses such as HIV-1, Ebola, cytomegalovirus, hepatitis-C, dengue, and SARS-coV, and parasites like *Leishmania pifanoi* and *Schistosoma mansoni* [19–27]. In addition, recent data suggest that DC-SIGN may mediate intracellular signalling events leading to cytokine secretion and, on this basis, it has been proposed that the lectin could be used by pathogens, including *M. tuberculosis,* as a part of an immune evasion strategy to their own advantage [28,29].

In light of the ability of DC-SIGN to interact with *M. tuberculosis* and other pathogens, it is plausible that variation in its gene may influence the pathogenesis of a number of infectious diseases, including TB. We have therefore explored the relationship between *CD209* polymorphisms and susceptibility to TB by determining *CD209* sequence variation in a cohort of South African Coloured origin.

Methods

Patients and Methods

The study was conducted in a cohort of 711 individuals, including 351 TB patients and 360 healthy controls, living in the Cape Town area. Certain suburbs of metropolitan Cape Town have some of the highest reported incidence rates of TB in the world, despite extensive BCG vaccination. Indeed, our study population comes from two suburbs that have been extensively studied because of their uniform ethnicity (known as South African Coloured) and socio-economic status as well as high incidence of TB and low prevalence of HIV [30]. In addition, our study group represents a present-day homogenous population [31] that previously received genetic input from Khoisan, Malaysian, Bantu, and European descent populations [32]. Thus, it represents a community originating from populations with different susceptibilities to TB and offers a unique opportunity to dissect the contributing genetic variants and their probable geographic/ethnic origins. TB patients were bacteriologically-confirmed (smear-positive and/or culture-positive) to present pulmonary tuberculosis. Their mean age (± standard deviation) was 36.7 ± 10.9 y, and 51.8% were male. Controls were unrelated healthy individuals from the same community, with the same socio-economic status, access to health facilities, and chance of diagnosis, and with neither signs nor previous history of TB (mean age 34.6 ± 12.5 y, 22% male). The annual risk of infection in this suburb was estimated at 2.5% in 1987 and at 2.8%–3.5% in 1999, and it is therefore highly likely that, in such an environment, the vast majority of controls have been exposed to *M. tuberculosis* [33,34]. All subjects were HIV-negative and older than 18 y. Informed consent was obtained from all participants, and the study was approved by the ethics committee of the Faculty of Health Sciences, Stellenbosch University (South Africa).

Laboratory Procedures and Statistical Analysis

To identify informative *CD209* single nucleotide polymorphisms (SNPs) and to avoid ascertainment bias in the choice of markers to be tested, we first sequenced the whole *CD209* genomic region (seven coding exons, flanking intronic regions, and 1,000 base pairs situated 5′ of the start codon) in 28 randomly chosen individuals (56 chromosomes). Using polymorphisms with a minimum allele frequency of 0.05, unphased genotypic data were converted into haplotypes using the accelerated EM (Expectation Maximization) algorithm implemented in Haploview v3.1 [35]. To evaluate the accuracy of the EM algorithm, haplotype reconstruction was performed in parallel using the Bayesian statistical method [36] implemented in Phase v.2.1.1. Equivalent results were obtained using both methods, with all haplotypes presenting high levels of statistical support. In order to define a minimal number of SNPs explaining most haplotypic diversity, we used the BEST v1.0 software [37]. Eight haplotype-tagging SNPs were then selected to genotype the entire panel of 711 individuals. Further, potential population stratification between cases and controls was tested by genotyping 25 unlinked SNP markers in the entire study cohort. DNA samples were genotyped by either fluorescence polarization (VICTOR-2TM technology; PerkinElmer, Wellesley, California, United States) or TaqMan (ABI Prism-7000 Sequence Detection System; Applied Biosystems, Foster City, California, United States) assays. Statistical testing for genotypic and haplotypic associations were performed using STATA 8.2 and Haploview v3.1, respectively. The haplotype frequencies were obtained by summing the fractional likelihood of each haplotype for each individual (i.e., if a particular individual has been determined to have a 40% likelihood of haplotype A

and 60% likelihood of haplotype B, 0.4 and 0.6 would be added to the counts for A and B, respectively) [35].

Results and Discussion

Two variants located in the *CD209* promoter region (−871 A/G and −336 A/G) exhibited a frequency distribution significantly distorted between TB patients and controls, as indicated by a Chi-square test (Table 1). For the −871 variant, genotypes GG and GA were less frequently observed in cases (16.8%) compared to the control group (27.2%) ($p = 8.2 \times 10^{-4}$). For the −336 variant, genotypes GG and GA were more frequent in cases (70.6%) than in controls (61.9%) ($p = 0.01$). These observations suggest that the alleles −871A (odds ratio [OR]: 1.85; 95% CI: 1.29–2.66) and −336G (OR: 1.48; 95% CI: 1.08–2.02) increase the risk of developing TB in our South African cohort. At the haplotype level (Table 2), a Chi-square test first revealed that the global distribution of haplotype frequencies was significantly different between cases and controls ($p = 1.2 \times 10^{-3}$). One haplotype (H3) turned out to be the main haplotype responsible for such a distorted frequency distribution (Table 2). This haplotype, which contains both −871G and −336A, was found to be strongly associated with the control group ($p = 1.6 \times 10^{-3}$; OR: 1.7; 95% CI: 1.22–2.38). The associations with this haplotype, and with −871, remained highly significant ($p = 1.3 \times 10^{-2}$ and 6.6×10^{-3} respectively), even after the conservative Bonferroni correction for multiple testing.

Although our cohort is considered a present-day homogeneous community that has received genetic contribution from different populations multiple generations ago [31,32], population stratification between cases and controls can be a confounding factor leading to a spurious positive association. Indeed, the use of admixed populations in association-mapping studies can be very useful to identify disease-causing genetic variants that differ in frequency across parental populations. However, when the admixture event is too recent, allelic frequencies can differ coincidentally among cases and controls, reflecting a nonuniform genetic contribution from the parental populations to each subpopulation (i.e., cases and controls), rather than a genuine association between a given genetic variant and the phenotype under study. In this case, the study-cohort is said to present population stratification. To formally test and quantify the levels of background genetic differences [38], if any, between cases and controls, we genotyped the entire cohort for a panel of 25 independent SNPs markers which are (1) not in linkage disequilibrium with the candidate *CD209* locus and with any other known gene, (2) randomly distributed along the genome, and (3) polymorphic among the major ethnic groups (Table 3). The mean χ^2 statistic among the 25 SNPs for the comparison of allele frequencies between cases and controls, which represents the levels of stratification (μ) between the two groups [39], was 1.25 ($p = 0.26$), implying that the two groups were not significantly stratified. As an additional correction for stratification, we divided the χ^2 values obtained for our candidate gene *CD209* by the level of stratification detected (1.25) [39]. Even after such a conservative correction, the associations observed with −336 and −871 as well as with H3 remained significant (−336 $p = 2.8 \times 10^{-2}$; −871 $p = 2.7 \times 10^{-3}$; H3 $p = 4.8 \times 10^{-3}$). These observations support therefore the idea that the −871G and −336A variants are indeed genuinely associated with a protective role against TB.

In order to gain insights into the frequency distribution of these two SNPs, we genotyped them in 254 human chromosomes from sub-Saharan Africa, Europe, and East Asia as well as in eight chimpanzee chromosomes. We observed that the −871G and −336A forms, which we propose as offering protection against TB, corresponded to the derived allele in humans; we also observed that these forms are present at higher frequencies in Eurasians as compared to Africans (Table 4). Indeed, the −871G is absent in African populations whereas it reaches high frequencies (20%–40%) in European and Asian populations. Given the absence of the haplotypic combination of −871G and −336A among sub-Saharan Africans, its presence among South African Coloureds suggests that it was introduced through the historically well-known admixture with Europeans and Asians [31]. This observation highlights the power of using admixed populations to better understand historical issues associated with the geographic/ethnic origin of disease-affecting alleles, provided that their prevalence varies in the ancestors of the admixed population (i.e., different frequency of H3 in Africans versus non-Africans; Table 4).

In the context of TB, it has been suggested that present-day

Table 1. DC-SIGN Genotype Distributions in Patients with Tuberculosis and in Healthy Controls

| | | | Genotype Frequencies (%) | | | | | | | |
| | | | Cases (*n* = 351) | | | Controls (*n* = 360) | | | Cases vs. Controls[b] | |
htSNP[a]	1		2	11	12	22	11	12	22	*p*	OR
−939	G	>	A	52.1	39	8.8	45.3	46.9	7.8	0.07	0.76 [0.57–1.02]
−871	A	>	G	83.2	15.7	1.1	72.8	26.1	1.1	8.2×10^{-4}	1.85 [1.29–2.66]
−336	A	>	G	29.3	50.4	20.2	38.1	43.3	18.6	0.01	1.48 [1.08–2.02]
−139	A	>	G	8.3	37.6	54.1	7.2	43.1	49.7	0.24	0.84 [0.62–1.13]
2392	G	>	A	96.6	3.4	0	95	4.7	0.3	0.29	0.67 [0.32–1.42]
3220	T	>	C	74.6	23.4	2	75.8	22.8	1.4	0.71	1.07 [0.76–1.50]
3838	A	>	C	80.1	17.9	2	84.7	14.4	0.8	0.1	1.38 [0.94–2.04]
4235	G	>	C	51.6	41	7.4	57.2	37.8	5	0.13	1.26 [0.94–1.69]

[a]All htSNPs (haplotype tagging SNPs) were in Hardy-Weinberg *equilibrium* in both the global sample and in cases and controls, separately.
[b]The homozygotes for the most frequent allele were compared with the sum of the homozygotes and heterozygotes for the rare allele.
DOI: 10.1371/journal.pmed.0030020.t001

Table 2. Haplotype Distributions in Patients with Tuberculosis and in Healthy Controls

Haplotype[a]	Haplotype Frequencies (%)		Cases vs. Controls[b]	
	Cases	Controls	p	OR
H1-AAAGGTCG	10	6.2	0.01	1.70 [1.15–2.52]
H2-AAAGGTAG	7.6	9.3	0.25	0.80 [0.55–1.16]
H3-AGAGGTAG	8.9	14.2	1.6×10^{-3}	1.70 [1.22–2.38]
H4-GAAAGTAG	26	27.9	0.41	0.91 [0.72–1.15]
H5-GAGGGTAG	17.3	15	0.25	1.17 [0.88–1.55]
H6-GAGGGTAC	12.2	8.5	0.02	1.49 [1.05–2.11]
H7-GAGGATAC	1.7	2.2	0.44	0.77 [0.36–1.63]
H8-GAGGGCAC	12.8	11.2	0.35	1.16 [0.84–1.60]

[a]Haplotypes with frequency greater than 1%. The alleles are ordered from SNP-939 until SNP4235.
[b]The frequency of each haplotype in cases and controls was compared with the sum of all the others.
DOI: 10.1371/journal.pmed.0030020.t002

susceptibility to TB is determined by previous history of exposure [40]. There is fairly convincing evidence that TB has been endemic in Europe for several hundred years, whereas in Africa it has probably been rare before contact was initiated with Europeans [41–43]. It is expected therefore that *M. tuberculosis* has exerted stronger selective pressures on European than African populations [42]. Our results lend support to this hypothesis and suggest that the protective alleles −871G and −336A increased in frequency in non-African populations as a result of genetic adaptation to a longer period of TB exposure. The potential impact of tuberculosis

on the frequency of resistant alleles in European populations has been recently addressed using epidemiological data and statistical modeling [44]. The authors have sought to evaluate the expected changes in resistant allele frequencies, during the 300-y period corresponding to the peak epidemics of TB in Europe. They concluded that if a given resistant allele was at a low frequency in the beginning of an epidemic, selection by *M. tuberculosis* alone would increase the frequency of this allele, but not enough to bring it to epidemiologically significant levels. In this context, since DC-SIGN is known to interact with a vast range of pathogens, it is indeed likely that

Table 3. Frequency Distribution in the Study Cohort of the 25 SNPs Used to Test for Population Stratification

Name	rs Number	Location	ObsHET[a]	PredHET[b]	HWP[c]	MAF Cases (%)	MAF Controls (%)	χ^2	p-Value
SNP1	rs2048022	Chr4	0.509	0.497	0.588	43.4	48.3	3.469	0.0625
SNP2	rs1380229	Chr8	0.429	0.468	0.034	36.1	38.4	0.815	0.3667
SNP3	rs650389	Chr10	0.276	0.287	0.369	15.2	19.5	4.518	0.0335
SNP4	rs870384	Chr12	0.511	0.499	0.591	48.3	46.8	0.311	0.5771
SNP5	rs695982	Chr12	0.416	0.419	0.888	32.1	27.7	3.376	0.0662
SNP6	rs708682	Chr15	0.216	0.212	0.826	11.8	12.3	0.086	0.7693
SNP7	rs715774	Chr15	0.244	0.257	0.209	14.3	16.1	0.883	0.3473
SNP8	rs1433456	Chr15	0.313	0.317	0.784	19.9	19.7	0.011	0.9169
SNP9	rs807131	Chr17	0.478	0.453	0.183	35.5	33.9	0.398	0.5279
SNP10	rs11672183	Chr19	0.214	0.209	0.649	12	11.7	0.018	0.8947
SNP11	rs2024628	Chr20	0.513	0.494	0.333	42.2	46.5	2.627	0.1051
SNP12	rs1028184	Chr2	0.461	0.464	0.9	34.2	39	3.53	0.0603
SNP13	rs2056773	Chr3	0.492	0.473	0.32	39.5	37.1	0.808	0.3687
SNP14	rs1479067	Chr5	0.398	0.395	0.972	25.9	28.4	1.099	0.2945
SNP15	rs327747	Chr7	0.374	0.399	0.111	25.8	29.2	2.084	0.1488
SNP16	rs12665321	Chr6	0.239	0.234	0.693	14.2	12.9	0.462	0.4969
SNP17	rs1566838	Chr9	0.475	0.497	0.265	46.5	45.8	0.081	0.7762
SNP18	rs12785524	Chr11	0.452	0.483	0.106	39	42.4	1.708	0.1912
SNP19	rs975423	Chr13	0.429	0.464	0.056	35.1	37.8	1.134	0.2868
SNP20	rs914904	Chr13	0.397	0.409	0.462	29.2	28.2	0.179	0.6723
SNP21	rs876287	Chr14	0.459	0.484	0.182	41.3	40.9	0.024	0.8765
SNP22	rs1582598	Chr16	0.378	0.394	0.321	27.5	26.5	0.164	0.6852
SNP23	rs1364198	Chr16	0.361	0.364	0.908	25.2	22.7	1.242	0.2652
SNP24	rs739259	Chr22	0.469	0.469	1	36.1	39	1.217	0.27
SNP25	rs169479	Chr21	0.217	0.217	1	13.3	11.5	1.064	0.3024

[a]Observed heterozygosity.
[b]Predicted heterozygosity.
[c]Hardy–Weinberg equilibrium probability.
MAF, minimum allele frequency.
DOI: 10.1371/journal.pmed.0030020.t003

Table 4. Frequency Distribution of the Eight *CD209* SNPs Genotyped in the Multi-Ethnic Panel of 127 Individuals (*n* = 254 Chromosomes) as Well as in the 711 Individuals of the South African Cohort (*n* = 1,422 Chromosomes)

SNP	Variant[a]	Population Frequencies (%)			
		African (*n* = 82)	Asian (*n* = 86)	European (*n* = 86)	SAC (*n* = 1,422)
−939	A	45.1	29.1	54.7	29.8
	G	54.9	70.9	45.3	70.2
−871	A	100	79.1	61.6	88.4
	G	0	20.9	38.4	11.6
−336	G	37.8	5.8	20.9	42.8
	A	62.2	94.2	79.1	57.2
−139	G	87.8	33.7	75.6	72.1
	A	12.2	66.3	24.4	27.9
2,392	G	100	98.8	91.9	97.8
	A	0	1.2	8.1	2.2
3,220	C	14.6	0	0	13.2
	T	85.4	100	100	86.8
3,838	A	70.7	100	90.7	90.5
	C	29.3	0	9.3	9.5
4,235	G	76.8	96.5	86	74.1
	C	23.2	3.5	14	25.9

[a]The allele in bold corresponds to the derived allele when compared with the sequence of the chimpanzee.
SAC, South African coloured population.
DOI: 10.1371/journal.pmed.0030020.t004

the increased frequencies observed today for both −871G and −336A in non-African populations (specially for −871G which is absent in sub-Saharan Africans) may have been driven, not only by the selective pressures imposed by *M. tuberculosis,* but also by other infectious agents. Indeed, two independent studies have recently reported a genetic association between the −336A variant and protection against parenteral HIV infection [45] and severity of dengue pathogenesis [46]. Although HIV infection, for example, is too recent to have left any signature of selection on *CD209,* these observations emphasize the possible action of other pathogens in shaping the patterns of variability of this gene.

From a functional point of view, the −336A allele has been shown to affect an Sp1-like binding site and to modulate transcriptional activity in vitro by increasing the levels of expression [46]. In the context of TB, increased DC-SIGN expression levels by DCs may result in better capture and processing of mycobacterial antigens, leading to a stronger and wider T-cell response. In addition, we have recently shown that DC-SIGN expression is markedly induced in alveolar macrophages in active TB patients and that *M. tuberculosis* is preferentially phagocytosed by DC-SIGN–expressing macrophages in these individuals [47]. Thus, the higher prevalence observed among healthy individuals of the −336A variant, which is associated with increased DC-SIGN expression, may underlie an increased efficiency of host phagocytes, such as DCs and macrophages, to control the infection. In addition to the −336A variant, our genetic data showed a strong association of the −871G allele with healthy controls, suggesting also a functional consequence of this variant that, either alone or in combination with −336A, remains to be defined.

In conclusion, the significant association found for the *CD209* promoter variants together with their phylogenetic status and frequency distribution strongly suggests that the −871G and −336A alleles may reduce the risk of developing TB. More generally, our results, together with those reporting

association of *CD209* promoter variants with both HIV susceptibility and dengue pathogenesis [45,46] suggest that variation in this lectin may be of crucial importance in the outcome of a number of infections due to DC-SIGN–interacting pathogens. Detailed in vitro and in vivo studies assessing the functional consequences of *CD209* variants on the quality of the host immune response against pathogens, including *M. tuberculosis,* are now required to eventually develop knowledge-based and effective pathway-targeted treatments.

References

1. Frieden TR, Sterling TR, Munsiff SS, Watt CJ, Dye C (2003) Tuberculosis. Lancet 362: 887–899.
2. Enarson DA, Rouillon A (1994) The epidemiological basis of tuberculosis control. In: Davies PD, editor. Clinical tuberculosis. London: Chapman and Hall. pp 19–32.
3. Lienhardt C (2001) From exposure to disease: The role of environmental factors in susceptibility to and development of tuberculosis. Epidemiol Rev 23: 288–301.
4. Bellamy R (2003) Susceptibility to mycobacterial infections: The importance of host genetics. Genes Immun 4: 4–11.
5. Blackwell JM (2001) Genetics and genomics in infectious disease susceptibility. Trends Mol Med 7: 521–526.
6. Cooke GS, Hill AV (2001) Genetics of susceptibility to human infectious disease. Nat Rev Genet 2: 967–977.
7. Casanova JL, Abel L (2002) Genetic dissection of immunity to mycobacteria: The human model. Annu Rev Immunol 20: 581–620.
8. Janeway CA Jr., Medzhitov R (2002) Innate immune recognition. Annu Rev Immunol 20: 197–216.

9. Kimbrell DA, Beutler B (2001) The evolution and genetics of innate immunity. Nat Rev Genet 2: 256–267.

10. Taylor PR, Martinez-Pomares L, Stacey M, Lin HH, Brown GD, et al. (2005) Macrophage receptors and immune recognition. Annu Rev Immunol 23: 901–944.

11. Medzhitov R (2001) Toll-like receptors and innate immunity. Nat Rev Immunol 1: 135–145.

12. Fujita T (2002) Evolution of the lectin-complement pathway and its role in innate immunity. Nat Rev Immunol 2: 346–353.

13. McGreal EP, Miller JL, Gordon S (2005) Ligand recognition by antigen-presenting cell C-type lectin receptors. Curr Opin Immunol 17: 18–24.

14. Tailleux L, Schwartz O, Herrmann J, Pivert E, Jackson M, et al. (2003) DC-SIGN is the major *Mycobacterium tuberculosis* receptor on human dendritic cells. J Exp Med 197: 121–127.

15. Soilleux EJ, Morris LS, Leslie G, Chehimi J, Luo Q, et al. (2002) Constitutive and induced expression of DC-SIGN on dendritic cell and macrophage subpopulations in situ and in vitro. J Leukoc Biol 71: 445–457.

16. Geijtenbeek TB, Torensma R, van Vliet SJ, van Duijnhoven GC, Adema GJ, et al. (2000) Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. Cell 100: 575–585.

17. Geijtenbeek TB, Krooshoop DJ, Bleijs DA, van Vliet SJ, van Duijnhoven GC, et al. (2000) DC-SIGN-ICAM-2 interaction mediates dendritic cell trafficking. Nat Immunol 1: 353–357.

18. Geijtenbeek TB, van Vliet SJ, Engering A, 't Hart BA, van Kooyk Y (2004) Self- and nonself-recognition by C-type lectins on dendritic cells. Annu Rev Immunol 22: 33–54.

19. Geijtenbeek TB, Kwon DS, Torensma R, van Vliet SJ, van Duijnhoven GC, et al. (2000) DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. Cell 100: 587–597.

20. Alvarez CP, Lasala F, Carrillo J, Muniz O, Corbi AL, et al. (2002) C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans. J Virol 76: 6841–6844.

21. Halary F, Amara A, Lortat-Jacob H, Messerle M, Delaunay T, et al. (2002) Human cytomegalovirus binding to DC-SIGN is required for dendritic cell infection and target cell trans-infection. Immunity 17: 653–664.

22. Lozach PY, Lortat-Jacob H, de Lacroix de Lavalette A, Staropoli I, Foung S, et al. (2003) DC-SIGN and L-SIGN are high-affinity binding receptors for hepatitis C virus glycoprotein E2. J Biol Chem 278: 20358–20366.

23. Tassaneetrithep B, Burgess TH, Granelli-Piperno A, Trumpfheller C, Finke J, et al. (2003) DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells. J Exp Med 197: 823–829.

24. Marzi A, Gramberg T, Simmons G, Moller P, Rennekamp AJ, et al. (2004) DC-SIGN and DC-SIGNR interact with the glycoprotein of Marburg virus and the S protein of severe acute respiratory syndrome coronavirus. J Virol 78: 12090–12095.

25. Bergman MP, Engering A, Smits HH, van Vliet SJ, van Bodegraven AA, et al. (2004) *Helicobacter pylori* modulates the T helper cell 1/T helper cell 2 balance through phase-variable interaction between lipopolysaccharide and DC-SIGN. J Exp Med 200: 979–990.

26. Appelmelk BJ, van Die I, van Vliet SJ, Vandenbroucke-Grauls CM, Geijtenbeek TB, et al. (2003) Carbohydrate profiling identifies new pathogens that interact with dendritic cell-specific ICAM-3-grabbing nonintegrin on dendritic cells. J Immunol 170: 1635–1639.

27. Colmenares M, Puig-Kroger A, Pello OM, Corbi AL, Rivas L (2002) Dendritic cell (DC)-specific intercellular adhesion molecule 3 (ICAM-3)-grabbing nonintegrin (DC-SIGN, CD209), a C-type surface lectin in human DCs, is a receptor for *Leishmania amastigotes*. J Biol Chem 277: 36766–36769.

28. Geijtenbeek T, van Vliet S, Koppel E, Sanchez-Hernandez M, Vandenbroucke-Grauls C, et al. (2003) Mycobacteria target DC-SIGN to suppress dendritic cell function. J Exp Med 197: 7–17.

29. Tailleux L, Maeda N, Nigou J, Gicquel B, Neyrolles O (2003) How is the phagocyte lectin keyboard played? Master class lesson by *Mycobacterium tuberculosis*. Trends Microbiol 11: 259–263.

30. Beyers N, Gie RP, Zietsman HL, Kunneke M, Hauman J, et al. (1996) The use of a geographical information system (GIS) to evaluate the distribution of tuberculosis in a high-incidence community. S Afr Med J 86: 40–41, 44.

31. van der Ross RE (1993) 100 questions about Coloured South Africans. Cape Town (South Africa): UWC Printing Department. 36 p.

32. Nurse GT, Weiner JS, Jenkins T (1985) The peoples of Southern Africa and their affinities. Oxford (United Kingdom): Clarendon Press. 410 p.

33. Munch Z, Van Lill SW, Booysen CN, Zietsman HL, Enarson DA, et al. (2003) Tuberculosis transmission patterns in a high-incidence area: A spatial analysis. Int J Tuberc Lung Dis 7: 271–277.

34. Beyers N, Michaelis I, Gie R, Schaaf HS, Richardson M, et al. (2001) Transmission of tuberculosis (TB) to children in a high incidence area. Int J Tuberc Lung Dis 5: S185.

35. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: Analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265.

36. Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 73: 1162–1169.

37. Sebastiani P, Lazarus R, Weiss ST, Kunkel LM, Kohane IS, et al. (2003) Minimal haplotype tagging. Proc Natl Acad Sci U S A 100: 9900–9905.

38. Pritchard JK, Rosenberg NA (1999) Use of unlinked genetic markers to detect population stratification in association studies. Am J Hum Genet 65: 220–228.

39. Reich DE, Goldstein DB (2001) Detecting association in a case-control study while correcting for population stratification. Genet Epidemiol 20: 4–16.

40. O'Brien SJ (1991) Ghetto legacy. Curr Biol 1: 209–211.

41. Stead WW, Senner JW, Reddick WT, Lofgren JP (1990) Racial differences in susceptibility to infection by *Mycobacterium tuberculosis*. N Engl J Med 322: 422–427.

42. Stead WW (1992) Genetics and resistance to tuberculosis. Could resistance be enhanced by genetic engineering? Ann Intern Med 116: 937–941.

43. Daniel TM, Bates JH, Downes KA (1994) History of tuberculosis. In: Bloom BR, editor. Tuberculosis: Pathogenesis, protection, control. Washington, DC: ASM Press. pp 13–24.

44. Lipsitch M, Sousa AO (2002) Historical intensity of natural selection for resistance to tuberculosis. Genetics 161: 1599–1607.

45. Martin MP, Lederman MM, Hutcheson HB, Goedert JJ, Nelson GW, et al. (2004) Association of DC-SIGN promoter polymorphism with increased risk for parenteral, but not mucosal, acquisition of human immunodeficiency virus type 1 infection. J Virol 78: 14053–14056.

46. Sakuntabhai A, Turbpaiboon C, Casademont I, Chuansumrit A, Lowhnoo T, et al. (2005) A variant in the CD209 promoter is associated with severity of dengue disease. Nat Genet 37: 507–513.

47. Tailleux L, Pham-Thi N, Bergeron-Lafaurie N, Herrmann JL, Charles P, et al. (2005) DC-SIGN induction in alveolar macrophages defines a privileged target host cell population for mycobacteria in patients with tuberculosis. PLoS Med (2)12: e381.

Patient Summary

Background. Tuberculosis (TB) is an infectious disease caused by the bacterium *Mycobacterium tuberculosis*. The disease kills between 2 million and 3 million people each year, and almost one-third of the world's population, or 2 billion people, are thought to be infected with the bacterium. However, only about 10% of infected individuals become sick. In the other cases, the body's immune system is capable of controlling the growth of the bacteria. Whether a person's immune system is strong enough to keep the bacterium in check depends on a variety of factors, such as age, how well nourished the person is, other infections, and genetic make-up.

Why Was This Study Done? Understanding the genetic factors that influence whether an infected person is more or less susceptible to getting sick with TB should help doctors to fight the disease. A number of research groups are studying these genetic factors, and some genes have previously been identified that seem to increase a person's risk of TB. Some of the authors of this study work on a molecule that is part of the immune system. The molecule is called DC-SIGN, and it plays a role in infection of cells in the lung by the Mycobacterium. Together with colleagues who are genetics experts, they wanted to see whether variation in the gene (called *CD209*) that codes for DC-SIGN influences the risk of someone infected with *M. tuberculosis* getting sick.

What Did the Researchers Do and Find? They worked with a group of South African participants from an area close to Cape Town where TB is very common. Half of the participants had TB; the other half did not (even though, because TB is so common in that area, it was likely that most if not all of them had been infected with *M. tuberculosis* at some point). The researchers compared make-up of the gene for DC-SIGN in all participants and found that some variants of the gene were more common in the group of healthy individuals. In other words, having this particular type of DC-SIGN-encoding gene seemed to protect individuals from getting TB.

What Does This Mean? DC-SIGN is a central part of the immune system, and others scientists have reported links between variants in the gene for DC-SIGN and the risk of picking up other infectious diseases, including dengue fever and HIV/AIDS. The present study lends more support to the notion that DC-SIGN is a key player in the control of infectious diseases. Understanding more about DC-SIGN could help to develop better treatments for these infections.

Where Can I Find More Information Online? The following Web sites provide information about tuberculosis.

World Health Organization pages on TB:
http://www.who.int/mediacentre/factsheets/fs104/en/index.html

OMNI pages on TB:
http://omni.ac.uk/browse/mesh/D014376.html

Medline PLUS pages on TB:
http://medlineplus.nlm.nih.gov/medlineplus/tuberculosis.html

European TB Vaccine Cluster:
http://www.tb-vac.org