EDITORIAL

# A Common Genetic Fingerprint in Leprosy and Crohn's Disease?

Erwin Schurr, Ph.D., and Philippe Gros, Ph.D.

The cause-and-effect relationship between severe infections and death suggests that microbial pathogens are evolutionary sculptors of the genome. However, the genetic component of susceptibility to infections in the general population is complex and heterogeneous and is modulated by environmental factors such as determinants of microbial virulence. Thus, it is a challenge to identify specific genetic effects in human populations. Availability of the human genome sequence, combined with knowledge of genetic variation, has facilitated the genomewide association study, a powerful approach to detecting genetic associations. In this issue of the *Journal*, Zhang and colleagues[1] describe a genomewide association study of leprosy, a bacterial disease.

Leprosy manifests with a broad pathologic spectrum. At one end is the localized paucibacillary form, characterized by a small number of hypopigmented, anesthetic skin lesions; at the other end is the disseminated multibacillary form, involving numerous skin lesions with a high bacillary load. Paucibacillary infection is associated with immune responses mediated by type 1 helper T (Th1) cells (involving the production of interferon-$\gamma$ and interleukin-2) that promote granuloma formation and limit bacterial replication and dissemination. The multibacillary form, in contrast, is associated with the Th1 polarization of the immune response (and the production of interleukin-4 and interleukin-10), which promotes uncontrolled bacterial replication and more severe pathology.

Population studies and studies of twins have established that there is a genetic component to susceptibility to leprosy.[2,3] Linkage and association studies have implicated variants of the *HLA-DR* region, *PARK2* (encoding parkin), *LTA* (encoding lymphotoxin alpha), and chromosome 10p13 in conferring susceptibility to leprosy in independent populations.[3]

Zhang and colleagues describe the results of genomewide scanning in persons with paucibacillary or multibacillary forms of leprosy from eastern or southern China. The authors compared the prevalence of each genetic marker — in this case, each single-nucleotide polymorphism (SNP), which is currently the marker typically used in genomewide association studies — in 706 case patients and 1225 unaffected persons. A total of 93 SNPs were shown to have a significant association with leprosy. These SNPs were then tested in three replication sets of more than 3000 patients and nearly 6000 controls from eastern or southern China. The data implicate *CCDC12* (the gene encoding coiled-coil domain containing 122), *C13orf31* (encoding chromosome 13 open reading frame 31), *NOD2* (encoding nucleotide-binding oligomerization domain containing 2), *TNFSF15* (encoding tumor necrosis factor [ligand] superfamily member 15), *RIP2K* (encoding receptor-interacting serine–threonine kinase 2), and the *HLA-DR–DQ* locus. Several of the proteins encoded by these genes are involved in microbial sensing and in the early innate immune and inflammatory responses. NOD2 recognizes a component of the mycobacterial wall, and stimulation of NOD2 results in the recruitment of RIPK2 and indirectly prompts the activation of the transcriptional regulator nuclear factor $\kappa$B (NF-$\kappa$B)[4] — which in turn activates the transcription of genes encoding proinflammatory cytokines including TNFSF15. On the surface of phagocytes, HLA-DR molecules present bacterial antigens to CD4+ T cells to initiate Th1-cell polarization.

Human genetic studies have implicated both

*IL12B* (the gene encoding interleukin-12$\beta$) and *NOD2* in increased susceptibility to mycobacterial disease, and mouse mutants lacking either *Nod2*, *Ripk2*, or *Infg* (encoding interferon-$\gamma$) are highly susceptible to tuberculosis.[5-7] The fact that such genes are now implicated through a genomewide association study of leprosy not only validates this approach to studying the disease, but also raises interest in the newly implicated genes (e.g., *CCDC12* and *C13orf31*), the functions of which are not known.

Another interesting aspect of the study is that variation in some of the implicated genes is known to be associated with bowel inflammatory conditions. A frame-shift mutation in *NOD2* has been identified as a strong susceptibility factor for Crohn's disease; additional *NOD2* mutations have been discovered not only in persons with Crohn's disease but also in those with Blau's syndrome and in those with early-onset sarcoidosis.[8] Likewise, variants of *TNFSF15* and *IL12B* have been associated with Crohn's disease.[9] These findings are consistent with studies of mouse models that have also established a role for *Nod2*, *Ripk2*, and *Nfkb* in intestinal homeostasis and colitis.[10,11] Together, these studies establish a strong genetic and functional link between susceptibility to leprosy and predisposition to Crohn's disease.

Although the results described by Zhang and colleagues are exciting, additional experiments are required to validate and refine their conclusions. Genomewide association studies are fairly crude. For example, a regulatory SNP in the *LTA* gene, previously identified as a leprosy risk factor,[12] was not tested in the platform used by Zhang and colleagues. Moreover, linkage disequilibrium (a state in which genetic markers — typically those in close physical proximity — are more likely than not to be inherited together) can extend over large intervals, with the main genetic effect located within a region that is a considerable distance from the SNPs showing the disease association. As known SNPs increase in number, the dissection of local effects of linkage disequilibrium will become more accurate, and testing for the presence of associations in groups of persons of differing ancestries will shed light on the extent to which these findings of Zhang and colleagues apply to other populations.

How do we move from P values to understanding pathogenesis and response to infection and, finally, to clinical outcomes? Although genomewide association studies rarely identify causative genetic lesions, they do point to specific genes and biologic pathways that can be targeted for pharmacologic intervention, irrespective of the mechanism underlying genetic susceptibility. A particularly attractive aspect of the study by Zhang and colleagues is the apparently narrow focus of the genetic control, which highlights early antigen sensing and signaling in the pathogenesis of both leprosy and Crohn's disease. It is tempting to speculate that these common genetic signatures support, albeit indirectly, the proposal that a proportion of Crohn's disease cases may have a mycobacterial cause.[13,14] Irrespective of its strength, such a link may broaden the therapeutic treatment options for both diseases.

In comparison with the study by Zhang and colleagues, genomewide association studies of susceptibility to malaria[15] and to infection with the human immunodeficiency virus[16] suggest that the contribution of common genetic variants is more limited. Why would this be so? The genomic variability of *Mycobacterium leprae* isolates is very small, and *M. leprae* has undergone substantial reductive evolution, possibly through adaptation to its human host. This may suggest that larger host genetic effects in infectious disease reflect decreased pathogen variability. In this view, pathogens with greater genetic variability, such as *M. tuberculosis*, will give rise to a more complex genetic architecture of host susceptibility.

Financial and other disclosures provided by the authors are available with the full text of this article at NEJM.org.

From the Research Institute, McGill University Health Centre (E.S.), and the Department of Biochemistry, McGill University (P.G.), both in Montreal.

This article (10.1056/NEJMe0910690) was published on December 16, 2009, at NEJM.org.

**1.** Zhang F-R, Huang W, Chen S-M, et al. Genomewide association study of leprosy. N Engl J Med 2009;361. DOI: 10.1056/NEJMoa0903753.
**2.** Global leprosy situation, 2007. Wkly Epidemiol Rec 2007;82:225-32.
**3.** Alter A, Alcaïs A, Abel L, Schurr E. Leprosy as a genetic model for susceptibility to common infectious diseases. Hum Genet 2008;123:227-35.
**4.** Kanneganti T-D, Lamkanfi M, Núñez G. Intracellular NOD-like receptors in host defense and disease. Immunity 2007;27:549-59.
**5.** Fortin A, Abel L, Casanova JL, Gros P. Host genetics of mycobacterial diseases in mice and men: forward genetic studies of

BCG-osis and tuberculosis. Annu Rev Genomics Hum Genet 2007;8:163-92.
**6.** Austin CM, Ma X, Graviss EA. Common nonsynonymous polymorphisms in the NOD2 gene are associated with resistance or susceptibility to tuberculosis disease in African Americans. J Infect Dis 2008;197:1713-6.
**7.** Pandey AK, Yang Y, Jiang Z, et al. NOD2, RIP2 and IRF5 play a critical role in the type I interferon response to Mycobacterium tuberculosis. PLoS Pathog 2009;5(7):e1000500.
**8.** Le Bourhis L, Benko S, Girardin SE. Nod1 and Nod2 in innate immunity and human inflammatory disorders. Biochem Soc Trans 2007;35:1479-84.
**9.** Barrett JC, Hansoul S, Nicolae DL, et al. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 2008;40:955-62.
**10.** Watanabe T, Asano N, Murray PJ, et al. Muramyl dipeptide activation of nucleotide-binding oligomerization domain 2 protects mice from experimental colitis. J Clin Invest 2008;118:545-59.
**11.** Bertrand MJ, Doiron K, Labbé K, Korneluk RG, Barker PA, Saleh M. Cellular inhibitors of apoptosis cIAP1 and cIAP2 are required for innate immunity signaling by the pattern recognition receptors NOD1 and NOD2. Immunity 2009;30:789-801.
**12.** Alcaïs A, Alter A, Antoni G, et al. Stepwise replication identifies a low-producing lymphotoxin-alpha allele as a major risk factor for early-onset leprosy. Nat Genet 2007;39:517-22.
**13.** Coulombe F, Behr MA. Crohn's disease as an immune deficiency? Lancet 2009;374:769-70.
**14.** Casanova JL, Abel L. Revisiting Crohn's disease as a primary immunodeficiency of macrophages. J Exp Med 2009;206:1839-43.
**15.** Jallow M, Teo YY, Small KS, et al. Genome-wide and fine-resolution association analysis of malaria in West Africa. Nat Genet 2009 May 24 (Epub ahead of print).
**16.** Fellay J, Shianna KV, Ge D, et al. A whole-genome association study of major determinants for host control of HIV-1. Science 2007;317:944-7.