GENETICS OF SUSCEPTIBILITY TO INFECTIOUS DISEASES: TUBERCULOSIS AND LEPROSY AS EXAMPLES

SANDRINE MARQUET AND ERWIN SCHURR

McGill Centre for the Study of Host Resistance, McGill University Health Centre Research Institute, Montreal General Hospital, Montreal, Québec, Canada

This paper is available online at http://dmd.aspetjournals.org

In the majority of infectious diseases only a proportion of individuals exposed to a pathogen become infected and develop clinically evident disease. At least in part, this interindividual variability is determined by the combined effect of host proteins encoded by a series of genes that control the quantity and quality of host-parasite interaction and host immune responses. Identification of the most important host susceptibility/resistance genes will allow a better understanding of infectious disease pathogenesis and likely facilitate the development of new therapeutic strategies.

Several approaches can be used to map and identify a host infectious disease susceptibility gene. Three of the most widely used strategies, i.e., mouse models, candidate gene approach, and genome scanning, are briefly presented. To date, at least 11 genes have been implicated in susceptibility/resistance to mycobacterial infection and a short discussion of the experiments implicating individual genes in infectious disease susceptibility is given.

Mouse Models

One approach to identify human disease resistance and susceptibility genes is to identify murine resistance/susceptibility genes. In this strategy, it is assumed that the basic pathology of the infectious disease is similar in the animal model and the human host. Consequently, orthologous genes in mouse and humans are assumed to be important for variable susceptibility/resistance to infection with the same pathogen. Experimental models present several advantages, including the ease of control of the environment, the ready access to strains of defined and genetically homogenous backgrounds, the availability of genetically engineered animals, and the breeding at will of appropriately chosen progenitor strains. A well known example for a susceptibility gene that has been identified in the mouse is the “natural resistance associated macrophage protein 1” (Nramp1). Natural resistance to infection with several intracellular pathogens belonging to the genera Mycobacterium, Leishmania, and Salmonella has been shown to be under control of a single G169D amino acid substitution in the Nramp1 protein (Vidal et al., 1995, 1996). A potential problem of the cross-species homology approach is that allelic variants of orthologues can be highly divergent among two species. Hence, the most powerful use of mouse models is the identification of genes and their respective biochemical pathways that are involved in disease susceptibility rather than the identification of specific susceptibility/resistance gene variants.

Candidate Gene Approach

Candidate genes are generally selected on the basis of their known or speculated relevance to disease pathogenesis and the presence of intragenic polymorphisms of possible biological significance. Candidate genes can also be derived based on experiments in mouse models of infectious diseases thereby exploiting the identification of murine resistance/susceptibility loci. Variants within a candidate gene can be analyzed in linkage studies (family studies) and/or in association studies (case-control studies), but in most cases, association studies are used to study the possible biological relevance of polymorphisms in specific candidate genes. With a growing number of gene polymorphisms appearing in public databases each month, the candidate gene strategy has gained tremendously in popularity. Nevertheless, problems remain because it is unlikely that all genes important for susceptibility can be found a priori, and genes with major effects but unknown function can easily be missed. The interpretation of positive results on genetic associations with infectious diseases is frequently complicated by the lack of appropriate corrections for multiple comparisons. Moreover, undetected population admixture or poor choice of control populations probably explain part of the difficulties in reproducing significant marker disease associations.

An alternative approach that has been recently proposed is to scan the whole human genome with a large number of single nucleotide polymorphisms for whole genome association studies. The power to detect marker-phenotype associations following this strategy is presently a matter of controversy. However, it is clear that several million genotypes will need to be generated for such studies. Present genotyping techniques cannot be used for such a task, and until new technologies become available, family studies offer a more efficient and feasible strategy.

Total Genome Scanning

In this approach, a large number of microsatellites (around 300) evenly spaced across the whole genome are used for linkage analysis employing families with multiple sibs affected by the studied disease. Two types of linkage analysis can be performed: parametric and nonparametric analysis. Parametric linkage analysis by the lod score method requires a defined model specifying the relationship between the phenotype and the factors (environmental and genetic), which have an effect on phenotype expression. For example, such a model can be provided by complex segregation analysis. Genetic linkage analysis tests whether the marker segregates with the disease in pedigrees with multiple affected according to a Mendelian mode of inheritance. The test is formulated as logarithm of the ratio \( L(\theta)/L(\theta = 0.5) \) or lod score, i.e., the likelihood of observing the experimentally determined segregation pattern at a given recombination...
The possible role of NRAMP1 as a genetic factor in leprosy has been extensively studied. However, the first direct evidence of its involvement came from the study of Roy et al. (1997), who found that NRAMP1, a gene encoding an ABC transporter, was associated with lepromatous leprosy in an Indian population. NRAMP1 encodes a protein that is expressed in macrophages and is thought to play a role in the regulation of intracellular pH, which affects the growth of mycobacteria. The association between NRAMP1 and lepromatous leprosy suggests that this gene plays a role in the host immune response to the disease.

Despite the advances in understanding the genetic basis of leprosy, many questions remain unanswered. The complexity of the disease, with its multiple clinical manifestations and varying degrees of severity, presents a significant challenge for geneticists. Furthermore, the host genetic susceptibility to leprosy is likely to be influenced by a combination of genetic and environmental factors. Understanding the role of these factors will be crucial in developing effective strategies for disease prevention and control.
dence for a role of \textit{NRAMP1} in susceptibility to leprosy was obtained only recently in a large familial study in South Vietnam. Six \textit{NRAMP1} polymorphisms and four closely linked microsatellites were genotyped in 168 patients of 20 multiplex leprosy families, and significant evidence for linkage was obtained ($P<0.005–0.02$; Abel et al., 1998) (Table 1). Segregation analysis performed on 285 Vietnamese and 117 Chinese families detected evidence for a major codominant gene only in the Vietnamese population suggesting that the control of susceptibility to leprosy could be genetically heterogeneous according to the ethnic origin of the families. It is possible that genetic heterogeneity will partly explain the lack of linkage between \textit{NRAMP1} and leprosy susceptibility in previous reports (Shaw et al., 1993; Roger et al., 1997).

\textbf{VDR.} The binding of the active form of vitamin D (1,25-dihydroxy vitamin D) to the vitamin D receptor present on monocytes, macrophages, and activated lymphocytes plays an immunoregulatory role. A single base polymorphism in codon 352 of the \textit{VDR} gene can be detected as TaqI restriction fragment polymorphism with the two alleles designated “T” and “t”, respectively. The less common allele “t” has been associated with higher levels of \textit{VDR} mRNA expression in transient transfection assays (Morrison et al., 1992). In a case-control study in Calcutta, the codon 352 \textit{TaqI} \textit{VDR} polymorphism was found associated with lepromatous leprosy ($P=0.04$) and tuberculous leprosy ($P=0.004$; Roy et al., 1999) (Table 1). The frequency of the tt genotype was significantly increased in tuberculoid patients (21.5%) as compared with controls (7.8%). In contrast, the TT genotype frequency was higher in the lepromatous leprosy group (52.4%) compared with the controls (39.8%). However, in this population no significant association was found between \textit{NRAMP1} polymorphisms and leprosy susceptibility.

\textbf{Tuberculosis (HLA, NRAMP, VDR, IL-1Ra, IL-1b).} Numerous studies have analyzed a possible contribution of genetic factors to tuberculosis susceptibility. A consistent conclusion from twin, family, and adoption studies was that the genetic background of the host is an important control element for susceptibility to tuberculosis. Employing candidate gene studies a number of gene variants have been identified that contribute to tuberculosis risk.

\textit{HLA.} A large number of associations of \textit{HLA} type with tuberculosis have been reported for different populations. However, a substantial proportion of these tuberculosis associations could not be reproduced in independent studies. One of the most consistent findings that has been reproduced in several populations is that susceptibility to pulmonary tuberculosis appears associated with the \textit{HLA-DR2} serotype (Table 1) (Bothamley et al., 1989; Brahmajothi et al., 1991). However, a two-stage case-control study has shown the importance of the \textit{HLA-DQB1}*0503 allele in tuberculosis progression ($P=0.005$) in a Cambodian population while no significant effect of \textit{HLA DR2} alleles was detected (Goldfeld et al., 1998). A recent association study on 126 patients with pulmonary tuberculosis and 87 endemic controls from India indicated that \textit{HLA-DRB1}*1501 ($P=0.013$) and \textit{HLA-DQB1}*0601 ($P=0.008$) were associated with pulmonary tuberculosis (Ravikumar et al., 1999). Interestingly, no connection has been found between \textit{TNFA} polymorphisms and tuberculosis risk (Shaw et al., 1997). These results suggest that the variability in the major histocompatibility complex and its relationship to tuberculosis susceptibility deserves further clarification, preferably using \textit{HLA} analysis on the level of the nucleotide.

\textit{NRAMP1.} Recently a large case-control study conducted in The Gambia (West Africa) has shown that four alleles of \textit{NRAMP1} were significantly associated with susceptibility to tuberculosis (Table 1) (Bellamy et al., 1998). This genetic analysis was performed on 410 smear-test-positive tuberculosis patients and 417 ethnically matched healthy controls. The polymorphic variants of the \textit{NRAMP1} gene analyzed correspond to a dinucleotide CA repeat in the 5’ region, a single nucleotide polymorphism in intron 4 (469 + 14G/C), a non-conservative single-base change at codon 543 (D543N), and a 4-base pair deletion in the 3’ region. Combined analysis of the polymorphisms in intron 4 and in the 3’ region detected a strong association with tuberculosis ($P<0.001$; Bellamy et al., 1998). Subjects heterozygous for these two variants were four times overrepresented among patients with tuberculosis as compared with those bearing the most common \textit{NRAMP1} genotype (odds ratio: 4.07, 95% CI: 1.86–9.12). In a more recent genetic study of a tuberculosis outbreak in a Canadian Aboriginal Community, \textit{NRAMP1} was strongly linked to tuberculosis susceptibility (lod score 4.2) (Greenwood et al., 2000). These reports confirm that \textit{NRAMP1} is important in modulating susceptibility to tuberculosis and demonstrate that mouse models of infectious disease can identify relevant candidate genes for human disease.

\textit{VDR.} Variations in the vitamin D receptor gene were analyzed in the same Gambian population that was enrolled to demonstrate the \textit{NRAMP1} association with tuberculosis. Homozygous patients for a polymorphism at codon 352 (genotype tt) were significantly under-represented among those with tuberculosis ($P=0.01$; Bellamy et al., 1999) suggesting a role of \textit{VDR} in the pathogenesis of tuberculosis (Table 1). Recently it has also been shown that serum vitamin D deficiency may contribute to the high occurrence of tuberculosis among Gujarati Asian subjects ($P=0.008$) (Wilkinson et al., 2000). Moreover, the \textit{VDR} genotypes of 91 untreated tuberculosis patients and 116 healthy people who had been sensitized to tuberculosis indicated that \textit{VDR} polymorphisms are involved in tuberculosis susceptibility (Wilkinson et al., 2000).

\textit{MBL (or MBP).} Mannose binding lectin (MBL) also called mannose binding protein (MBP) activates the classical complement pathway and phagocytosis leading to neutralization of the pathogen. To investigate the role of \textit{MBL} gene polymorphisms in tuberculosis susceptibility, a recent study was carried out with Indian patients. Two hundred and two pulmonary tuberculosis patients and 109 controls were genotyped for three \textit{MBL} (codons 52, 54, and 57) functional variants affecting the structure of MBL and associated with low serum levels. A significantly increased genotype frequency of mutant homozygotes was noted in pulmonary tuberculosis patients compared with healthy controls ($P=0.008$) (Table 1) (Selvaraj et al., 1999). On the other hand, a protective effect of the G54D MBL allele on tuberculous meningitis was noted in a South African population (Hoal-van Helden et al., 1999).

\textit{IL-1Ra and IL-1b.} The proinflammatory cytokine interleukin-1b (IL-1b) and the interleukin-1 receptor antagonist (IL-1Ra), which is a specific inhibitor of IL-1 activity, have been shown to be induced in vitro by \textit{Mycobacterium tuberculosis}. Moreover patients with tuberculosis present an elevated serum concentration of IL-1Ra and IL-1Ra was identified as a marker of disease activity in tuberculosis (Juffermans et al., 1998). Within the \textit{IL-1b} gene, two biallelic polymorphisms have been identified at positions $-511$ and $+3953$, respectively. An 86-base pair variable number of tandem repeats polymorphism with five different alleles is found in the \textit{IL-1Ra} gene. Genetic analysis detected that the \textit{IL-1Ra} VNTR allele A2 was associated with a higher production of IL-1Ra in response to \textit{M. tuberculosis} infection. Indeed, individuals with the IL-1Ra A2+ allele produced 1.9-fold more IL-1Ra than patients with IL-1Ra A2– alleles (Wilkinson et al., 1999). The two polymorphisms in \textit{IL-1b} were not clearly associated with the level of IL-1b production in vitro induced by \textit{M. tuberculosis}, although expression of mRNA for IL-1b was slightly higher in patients with the \textit{IL-1b}+3953) A1+ allele ($P=0.05$).
gene was responsible for the children's immune deficiency and un-

IFNgR1 receptor at the cell surface. IFNgR1 deficiency has also been identi-

al., 1996). A second study indicated that one child with fatal dissem-

tion after BCG vaccination and the second with clinical tuberculosis

expression but the consequences of this deficiency are apparently

interestingly, patients with IL-12Rb1 deficiency develop mature BCG granulomas.

IL-12p40—complete deficiency. A large homozygous deletion (373 nucleotides) within the IL-12p40 subunit gene was found in a child with curable BCG and Salmonella enteritis infection (Table 1) (Altare et al., 1998d). Similar to IL-12Rb1 deficiency, IL-12p40 deficiency has no effect on mature BCG granuloma formation. Further studies suggested that susceptibility to mycobacterial infection of patients with genetically impaired IL-12-mediated immunity is due to insufficient IFNγ-mediated immunity. Moreover, the milder clinical phenotypes of IL-12p40- and IL-12Rb1-deficient patients compared with patients with complete IFNgR1 or IFNgR2 deficiency are explained by a residual (IL-12-independent) IFNγ secretion. Recently, Jouanguy et al. (1999) described a hotspot for small deletions in human IFNgR1 that confer dominant susceptibility to infection by poorly virulent mycobacteria. Molecular genetic analysis was performed on 18 pa-

patients from several generations of 12 unrelated families and 12 inde-

pendent mutations were found at a single site.

Taken together, these studies suggest that IFNγ and IL-12 are two

important cytokines in human defense against mycobacteria and sal-

monella. Particularly, the type-1 cytokine pathway appears essential

since it seems that type-1 deficiencies in patients are not compensated by

any other protective immune mechanism.

Conclusion

To identify susceptibility/resistance genes in infectious diseases different strategies can be used. To date, candidate gene analysis has been the method of choice in most studies. Contrary to genome-scan analysis, this approach allows the possibility of identifying genes that exert a small or moderate effect on susceptibility to infection. How-

ever, this approach is limited to known candidate genes, and the results should be confirmed in independent studies. Another challenge of candidate gene analysis is the identification of the causal mutation of the disease. This can be relatively straightforward in the case of major mutations but can be a difficult undertaking in the case of subtle polymorphic variations. The characterization of the susceptibility genes and their underlying causative mutations has important impli-

ations not only for a better understanding of disease pathogenesis but also for the control and development of new therapeutic strategies for infectious diseases.

References


GENETICS OF SUSCEPTIBILITY TO INFECTIOUS DISEASES


