Available online at www.sciencedirect.com
SCIENCE DIRECT
Microbes and Infection 8 (2006) 1647–1653
Microbes and Infection
www.elsevier.com/locate/micinf

Forum in immunology

The impact of genomics on the analysis of host resistance to infectious disease

Ashleigh Tuite [a], Philippe Gros [a,b,*]

[a] Department of Biochemistry, McGill University, 3655 Promenade Sir William Osler, Room 907, Montreal, QC H3G 1Y6, Canada
[b] Center for the Study of Host Resistance, McGill University, Montreal, QC, Canada

Available online 26 January 2006

Abstract

The advent of new technologies and resources, including the complete sequence of mammalian genomes, has had a dramatic impact on the genetic analysis of susceptibility to infections in humans and in animal models of infection. Genes responsible for simple or complex control of susceptibility to infection with different pathogens have been recently identified and characterized, and are reviewed herein.

Keywords: Genomics; Host resistance; Infectious disease

1. Introduction

Infectious diseases, including AIDS, tuberculosis, and malaria, have a major global health impact, accounting for over 25% of mortality worldwide, in addition to causing severe morbidity in a large segment of the world's population [1]. Despite the availability of drugs and vaccines against many infectious agents, the emergence of antibiotic-resistant strains and the failure of certain patients to respond to standard treatments have emphasized the need for the development of novel therapeutic agents. This involves the study of both pathogens, in terms of life cycles, invasion processes, and virulence mechanisms, and the host, whose own defense mechanisms determine the response to the infectious challenge, whether that be clearance of the microbe, asymptomatic colonization and persistence, or active infection with the development of ensuing pathology. Although pathogens have traditionally been the subject of intense scrutiny due to the relative ease of study in an experimental environment, the role of the host immune response during an infection should not be underestimated. A better understanding of the host response to various infectious diseases will help to clarify the mechanisms and pathways through which pathogens evade the host defense system and may provide additional targets for drug development.

The apparent heritability of infectious disease susceptibility has long been recognized in humans [2]. Diseases such as leprosy have long been thought to ''run'' in certain families. Over 50 years ago, Haldane proposed that the prevalence of thalessemias in malaria-endemic areas was due to the heterozygotic advantage they confer against malaria, despite their otherwise deleterious effects. The various inherited hemoglobin disorders that provide protection against malaria, such as sickle cell anemia and beta-thalassemia, represent relatively rare instances of single gene disorders that alter infection susceptibility. More frequently the contribution of host genetics is complex, with multiple genetic factors contributing to infection susceptibility, often acting at different stages during the infection process. The dissection of complex traits into single gene effects is compounded by additional factors, including environmental and pathogen interactions, incomplete penetrance, population heterogeneity, and phenotypic variance among affected individuals.

For studies in humans, the most common approach for the identification of genes important in infectious disease

* Corresponding author. Department of Biochemistry, McGill University, 3655 Promenade Sir William Osler, Room 907, Montreal, QC H3G 1Y6, Canada. Tel.: +1 514 398 7291; fax: +1 514 398 2603.
E-mail address: philippe.gros@mcgill.ca (P. Gros).

1286-4579/$ - see front matter © 2006 Elsevier SAS. All rights reserved.
doi:10.1016/j.micinf.2005.11.016

susceptibility is candidate genes studies, where genes are selected based on their known or suspected relevance to disease pathogenesis [3]. The obvious disadvantage of this approach is that it precludes the identification of presently uncharacterized genes that may nonetheless contribute in a significant way to the control of infection. Additionally, genes deemed to be significant through association or linkage studies may only modestly or minutely affect disease susceptibility once all gene effects are taken into account.

Mouse models of infection constitute a useful tool for the analysis of the genetic basis of susceptibility to infectious diseases, since single gene effects have often naturally segregated or been experimentally isolated in inbred strains of mice. A whole-genome scanning approach can be adopted, whereby genes are mapped using informative crosses followed by identification through positional cloning and validation by germline modification in vivo. Orthologous genes can subsequently be analyzed, to determine if they have a corresponding role in humans.

The focus of this review will be the effect of genomics on the mapping of host genes that contribute to resistance to infectious diseases. Specifically, we will examine several infectious agents for which the host response has been studied using a forward genetic approach, going from a specific phenotype to a genetic locus or gene, in both the mouse and human, through the use of different genomic approaches. Although the number of causative genes identified to date is relatively low, the refinement of existing and the development of new genomic tools hold much promise for continued success in the search for host resistance genes.

2. Simple traits

Mendelian gene effects, the relatively rare instances where a single gene explains most or all of the phenotypic variance of a trait, are more amenable to genetic analysis than complex traits, accounting for the higher success rates in mapping these traits in humans and mice [4]. However, identifying the causative gene often remains an arduous process, which has been recently facilitated by the advent of novel methodologies and approaches in the field of genomics.

2.1. Mycobacterium

The cloning and subsequent functional characterization of *Nramp1* represents one of the earliest demonstrations of the power of genetics in the study of host resistance to infectious diseases. Extensive reviews of this gene have been conducted elsewhere [5,6], and the cloning of this gene will only be discussed briefly, to highlight its pioneering role as the first infection susceptibility gene to be identified by a positional cloning approach. Susceptibility to infection by *Mycobacterium bovis* (BCG) and other *Mycobacterium* species, including *M. lepraemurium*, *M. avium*, and *M. intracellulare*, as measured by early in vivo replication in the liver and spleen, is under simple genetic control in inbred strains, with resistance dominant over susceptibility. The locus controlling this trait, termed *Bcg*, was mapped to proximal chromosome 1 using segregating F2 and backcross progeny derived from resistant and susceptible strains. Susceptibility loci for *Salmonella typhimurium* and *Leishmania donovani*, designated *Ity* and *Lsh*, respectively, had similarly been mapped to this region, suggesting that the three loci were either tightly linked or controlled by the same gene. It was observed that these gene effects were acting at the level of control of intracellular microbial replication in macrophages. A physical map of the area around *Bcg* was constructed using pulse field mapping data and transcription units were isolated from the minimal physical interval by exon amplification and localization of CpG islands. Of the six candidate genes identified in the region, mRNA expression analysis indicated that one gene was selectively expressed in the spleen and liver, the sites of pathogen replication, with mRNA expression enriched in macrophages isolated from these organs. This gene was designated *Nramp1* (for natural resistance associated macrophage protein 1, renamed *Slc11a1*). Sequence analysis revealed the presence of a glycine-to-aspartate substitution at position 169 in susceptible inbred strains. Demonstration of the causal relationship between *Nramp1* and *Bcg*/*Ity*/*Lsh* came in the form of a loss-of-function allele at *Nramp1* generated through gene targeting, which abrogated resistance to infection with all three infectious agents. Correspondingly, the transgenic transfer of the wild-type allele onto the mutant $Nramp1^{Asp169}$ background conferred resistance to a previously susceptible strain (reviewed in [5,6]).

Nramp1, an integral membrane protein with 12 predicted transmembrane domains, undergoes extensive phosphorylation and glycosylation. The G169D mutation, which maps to transmembrane domain 4 of the protein, seems to affect protein maturation and/or trafficking, since no mature protein is detected in macrophages with this allele. Nramp1 is expressed at the phagosomal membrane and studies of this protein and other Nramp homologues in the mouse and other species have indicated that Nramp1 most likely functions as a pH-dependent divalent cation efflux pump at the phagosomal membrane with a preference for Fe^{2+} and Mn^{2+} ions.

The long arm of human chromosome 2 harbors the *NRAMP1* orthologue and there is a high degree of conservation between the human and mouse genes. Numerous studies have been carried out to determine if *NRAMP1* has a corresponding role in infectious disease susceptibility in humans. The effect of *NRAMP1* is apparently population-specific, although an association between *NRAMP1* alleles and susceptibility to tuberculosis, leprosy, HIV, and other infectious and autoimmune diseases has been noted [6]. In populations where *NRAMP1* alleles do affect infection susceptibility with a specific pathogen, its contribution is generally modest, indicating a high level of complexity and the likely contribution of additional gene effects.

2.2. Mouse cytomegalovirus

The identification of genes that control host susceptibility to viral infections has been highly successful and has been reviewed elsewhere [7]. The identification of the gene