The Genetics of Rheumatoid Arthritis: Influences on Susceptibility, Severity, and Treatment Response

S. Louis Bridges, Jr, M.D., PhD

Introduction

Rheumatoid arthritis (RA) is a chronic systemic disease that affects approximately 1% of human populations throughout the world. It is characterized by lymphocytic infiltration of the synovial membrane of diarthrodial joints and often has extra-articular manifestations. Despite intensive investigation for many years, the etiology of the disease remains elusive. There is a genetic component to susceptibility to RA, as there is with virtually every form of arthritis, including familial osteoarthritis (type II collagen) [1], ankylosing spondylitis (MHC Class I gene HLA-B27) [2], gout (hypoxanthine-guanine phosphoribosyl-transferase) [3], and systemic lupus erythematosus (HLA DR3, complement C4a, Fc receptors) [4,5].

Compared with the general population, individuals who have siblings with RA are at increased risk of developing RA. Estimates of the sibling risk $\lambda_2$ range from 2- to 17-fold times that in the general population [6••]. The best-studied genetic association in RA is that of major histocompatibility complex (MHC) class II alleles. The contribution of MHC genes in RA susceptibility is difficult to quantitate, but it appears to be in the range of 15% to 40% of the total disease risk [7,8].

This article focuses on genetic loci that may predispose to development of RA, that may have an effect on disease manifestations, such as erosions or extra-articular manifestations, or that may influence response to particular treatments, such as tumor necrosis factor (TNF) inhibitors. The major benefits of identifying genetic markers of susceptibility and severity are better understanding of the pathogenesis of RA and in developing strategies for prevention or treatment.

Difficulties in Studying the Genetics of Complex Diseases

Concomitant with advances in our knowledge of the human genome and in tools used to study it is the appreciation that genetic factors contribute to some degree to virtually all illnesses. Despite remarkable progress in molecular genetics and epidemiology, however, there are major problems in identifying genes that contribute to complex, multifactorial diseases such as RA, asthma, and diabetes mellitus. First, a combination of genetic, environmental, historical, and evolutionary factors, as well as stochastic processes that have influenced populations, are all important etiologic considerations [6••]. Second, in complex diseases, a single allele of a particular gene may increase susceptibility but does not invariably cause disease [9]. Individual genes may have only a small effect on the disease, whereas sets of genes are likely to contribute to pathophysiology [6••]. Third, even though a particular gene may show an impact in one patient population, it may not have an effect in other populations. Such is the case with the association between RA and particular MHC class II alleles in whites, but not in African-Americans [10•]. A functional IL-6 polymorphism associated with susceptibility to systemic-onset juvenile rheumatoid arthritis [11] is found in markedly different proportions of normal whites and African-Americans [12]. Indeed, many single nucleotide polymorphisms (SNPs, see below) are relatively specific for individuals of particular race [13]. Fourth, the outbred human population requires the study of a large number of affected relatives to show a difference resulting from genetics [6••]. Finally, for complex diseases
not inherited in a completely Mendelian fashion, there is often no defined genetic model, which limits the usefulness of linkage studies. As with many other complex illnesses, there is no consensus regarding the mode of inheritance of RA. There is evidence to support a recessive model [7,14,15], a dominant model with reduced penetrance [16,17], and a multilocus model [8,18].

Development of Informative Genetic Markers

Many types of DNA sequence variations in the human genome have been used to identify genetic regions involved in complex diseases. With the advent of molecular biology, DNA restriction fragment length polymorphisms were discovered, followed by minisatellites and microsatellites. Microsatellites are tandem repeats of simple sequence (10–50 copies of motifs from 1 to 6 base pairs) that occur frequently and at random at ~10,000 base pair intervals throughout virtually all eukaryotic genomes. Because they are usually embedded in stretches of DNA with unique sequence, they can usually be easily polymerase chain reaction (PCR)-amplified and analyzed. Microsatellite alleles containing various numbers of dinucleotide repeats are inherited as typical genetic traits, making them useful in genetic mapping. Microsatellite studies can identify previously undiscovered genes associated with disease, based on the premise that the closer microsatellite lies to a disease gene, the more likely they are to be inherited together.

The most frequent and stable genetic variations in the human genome are single base pair differences. When found in ≥1% of the population, these are referred to as singlenucleotide polymorphisms (SNPs). SNPs are found at ~700 to 1000 base pair intervals, with an estimated 3 x 10^6 common variants in humans. In addition to being more abundant than microsatellites, SNPs have many properties that make them attractive in the study of human sequence variation and disease mapping studies [19]. SNPs have much lower mutation rate than do repeat sequences such as microsatellites, which makes them very stable. Although the proportion of SNPs that lead to changes in amino acid sequence of coding regions is lower than initially thought [13,20••], there are still likely to be a large number of SNPs responsible for functional changes. Because of their potential usefulness in genetic studies, detailed maps of SNPs are currently being developed.

The Role of MHC Class II Alleles in RA Susceptibility, Severity, and Treatment Response

An association between MHC class II DR alleles and RA [21] has been confirmed in multiple studies [22]. It is now generally accepted that particular DR4 subtypes Dw4 (DRB*0401), Dw14 (DRB1*0404), and Dw15 (DRB1*0405), and some DR1 alleles are associated with RA. Other MHC regions, including the DQA1, DQB1, DPB1, and DM loci may also play a role in susceptibility to RA. Nucleotide sequence analysis of HLA DR genes led to the hypothesis that these alleles confer susceptibility to RA based on shared homology at amino acid residues 70-74 of the third hypervariable region of the DRB1 chain, the so-called shared epitope (SE) [Fig 1] [23].

There are several possible mechanistic explanations for the association between particular MHC alleles and RA. These include: 1) a particular arthritogenic peptide antigen is presented to T cells by MHC molecules bearing the SE; 2) the SE is itself an antigentic peptide bound by other MHC molecules; and 3) molecular mimicry [24••]. Recently, an alternative explanation has been proposed, in which the SE selects for TCR through direct interaction between residues 70 to 74 of the DRB1 chain with the second complementarity-determining region of the TCR β chain [24••].

Associations between the HLA-DR SE and RA have not been found in all racial groups. For example, the predisposition to and severity of RA in African-Americans appears to be independent of the presence and dose of the SE [10•]. Among Indian RA patients, DR1 alleles appear to be the most important MHC susceptibility genes [29]; in Spaniards [30] and Israeli Jews [31], DR10 alleles are most prominent. Thus, it is critical to take race, and possibly ethnicity, into account during genetic analyses of susceptibility, severity, and treatment response in RA.

An alternative explanation for the association between the MHC class II locus and RA in whites is the presence of a recessive disease-susceptibility gene in linkage disequilibrium with the DRB1 locus [25], such as the prolactin gene [26]. The tumor necrosis factor-lymphotoxin-α (TNF-LT) locus is located near the MHC class II locus on chromosome 6p21. Several studies implicate TNF microsatellite polymorphisms in susceptibility to RA [27,28]. There is conflicting evidence, however, as to whether these associations are independent of the SE.

In addition to having a role in susceptibility to RA, MHC class II DR4 alleles have been reported to have an effect on disease severity (eg, more erosions on radiographs) [32,33]. Rheumatoid factor (RF)-positive whites with RA who bear two susceptibility alleles have been shown to be more likely