Editorial

Genetic Association Studies in Systemic Sclerosis: More Evidence of a Complex Disease

Diseases that are likely to have complex genetics typically exhibit broad variation in severity of clinical manifestation and age of onset (heterogeneous phenotypes), with mixed etiological mechanisms that are likely to involve multiple biological pathways. Almost by definition they are likely to be caused by several genes, each with a small overall contribution and relative risk[1]. In particular, it has been proposed that combinations of common genetic variants, including single nucleotide polymorphisms (SNP), not only influence the susceptibility to these diseases[2] but they also may be associated with particular aspects of the disease phenotype. Thus, for a clinically heterogeneous disease the pattern or extent of organ-based complication (disease severity) may be influenced by genetic variation (severity genes).

Systemic sclerosis (SSc) clearly falls into this category of complex genetic disease, with well recognized variability in clinical and serological presentation and intricate underlying mechanisms, which involve vascular and immune activation within a fibrotic process[3]. In addition, there is convincing evidence suggesting a role for genetic factors in association with environmental exposure in the etiopathogenesis of the disease. However, precisely defining the contribution of the genetics in determining the susceptibility to or severity of SSc is problematic, in particular because the disease has a low prevalence and there is a paucity of multiplex families or identical twins with the disease[4].

Genetic association studies are generally regarded as a useful tool to overcome these problems. Such analyses compare the frequency of presentation of each of the possible variants (alleles) located in one particular site of selected candidate genes in different groups of individuals, usually patients and unaffected controls. A positive association between alleles and increased risk to develop the disease can be inferred when significant differences in the distribution of the variants are detected between the studied groups. There are many reports analyzing the relationship between a wide variety of genes and SSc susceptibility or severity. These include HLA[5-7] and non-HLA genes, such as tumor necrosis factor-α[8], endothelin[9], and fibrillin[10], which provide potential insights into the pathogenesis of the disease. Interestingly, many of these studies have demonstrated strong association between polymorphisms and presence of particular SSc hallmark autoantibodies such as anti-topoisomerase-1 or anticentromere reactivities, which are themselves generally mutually exclusive[8,11].

In this issue of *The Journal*, Barbi, *et al* report that SNP in the genes encoding interleukin 1ß (IL-1ß) and IL-2 are associated with SSc susceptibility and severity[12]. They analyzed 9 SNP in 7 cytokine (inflammation related) genes (*IL10, IL1B, ILIA, ILIRN, IL2, LTA,* and *IL6*) in a cohort of 78 patients with SSc and 692 controls. *IL1B-31C* and *IL1B-511T* were found more frequently in the SSc group, and *IL2-384G* was associated with the limited form of the disease and the presence of anticentromere antibodies. In addition to providing a clue about genetic determinants of SSc such studies also highlight potentially important cytokine mediators of pathogenesis. In this context both IL-1 and IL-2 expression and function have been reported to be altered in earlier non-genetic studies of SSc.

Although information provided by genetic association studies is valuable and might prove to be important in diagnosis, risk stratification, or therapy, one must interpret single positive reports of association with disease with great caution. Methodological limitations of genetic studies must always be considered since many have problems with the study design and/or statistical analysis. This has led to many underpowered studies yielding interesting results that could not be replicated in independent groups. A good example of this problem is results of studies on polymorphism of SPARC[13,14] or PTPN22[15-17] gene in SSc patients that fail to replicate the initial observation, or that are even contradictory. Therefore, it is important to take into consid-

See Association of polymorphisms in the *IL1B* and *IL2* genes with susceptibility and severity of SSc, *page 997*

eration key points in the design and interpretation of genetic association studies.

What is the meaning of an association in a genetic study? When an association is found, 3 possibilities might be considered: first, that the association is real, i.e., that a particular SNP influences pathogenesis of the disease (usually by modifying the protein structure or gene regulation). Second, that the SNP in question is in linkage disequilibrium with other polymorphisms that are in association: this type of association might provide useful genetic markers. The third possibility is that the association detected is not real (spurious) and this may be due to inadequate statistical analysis (i.e., lack of correction for multiple comparison), to small numbers of cases or controls with resulting low statistical power, or to confounding factors (population stratification).

Several key aspects can contribute to making results of genetic association more reliable. One is the precise definition of the disease to be studied, ensuring that included study cases actually have the disease. When the disease has a heterogeneous presentation, such as SSc, it may be advantageous to analyze more narrowly defined subtypes, such as autoantibody subtype. Adequate size of case and control cohorts is critical for detecting subtle genetic differences between populations, in order to have sufficient statistical power. The overall genetic background of the group under study should be as homogenous as possible, thereby limiting the possibility of result bias by population stratification; therefore it would be desirable to include patients and controls with similar ethnicity[18].

The selection of genetic markers (SNP) to be evaluated represents a particular challenge as there are up to 30,000 genes in the human genome, and only some are likely to have any influence on the disease phenotype. It is therefore reasonable to select genes with previous evidence of biological relevance to the disease pathogenesis[1,19]. Before any association can be validated, it is important to determine the quality of the genotype results by comparing the frequency observed with the expected normal distribution (Hardy-Weinberg equilibrium). When the results of the control group differ statistically from the expected ones, any association reported in the study should be interpreted with caution, as the values used as the base for comparison are biased and a clear explanation for this deviation should be provided[20]. The statistical analysis of the results requires careful consideration in order to limit the possible type I statistical errors arising from multiple comparisons. The degree of the adequate correction is sometimes difficult to determine, especially when many variables are analyzed. An alternative solution for this problem is to repeat polymorphism analysis in an independent second set of cases and controls[21].

It is vital in genetic association studies that the methodology be refined and improved so that results provide a reliable and comprehensive picture of the genetic contribution in systemic sclerosis. Only then will it be possible to robustly compare patient subgroups and studies from distinct geographic or ethnic populations to fully understand the significance of genetic factors in determining the disease phenotype.

CARMEN FONSECA, MD, PhD,
Senior Clinical Research Fellow;
CHRISTOPHER P. DENTON, PhD, FRCP,
Professor and Consultant Rheumatologist,
Centre for Rheumatology,
Royal Free and University College Medical School,
Rowland Hill Street,
London NW3 2PF, United Kingdom

Address reprint requests to Prof. Denton;
e-mail: c.denton@medsch.ucl.ac.uk

REFERENCES

1. Tabor HK, Risch NJ, Myers RM. Opinion: Candidate-gene approaches for studying complex genetic traits: practical considerations. Nat Rev Genet 2002;3:391-7.
2. Yang Q, Khoury MJ, Friedman J, Little J, Flanders WD. How many genes underlie the occurrence of common complex diseases in the population? Int J Epidemiol 2005;34:1129-37.
3. Denton CP, Black CM, Abraham DJ. Mechanisms and consequences of fibrosis in systemic sclerosis. Nat Clin Pract Rheumatol 2006;2:134-44.
4. Assassi S, Tan FK. Genetics of scleroderma: update on single nucleotide polymorphism analysis and microarrays. Curr Opin Rheumatol 2005;17:761-7.
5. Gilchrist FC, Bunn C, Foley PJ, et al. Class II HLA associations with autoantibodies in scleroderma: a highly significant role for HLA-DP. Genes Immun 2001;2:76-81.
6. Lambert NC, Distler O, Muller-Ladner U, Tylee TS, Furst DE, Nelson JL. HLA-DQA1*0501 is associated with diffuse systemic sclerosis in Caucasian men. Arthritis Rheum 2000;43:2005-10.
7. Gladman DD, Kung TN, Siannis F, Pellett F, Farewell VT, Lee P. HLA markers for susceptibility and expression in scleroderma. J Rheumatol 2005;32:1481-7.
8. Sato H, Lagan AL, Alexopoulou C, et al. The TNF-863A allele strongly associates with anticentromere antibody positivity in scleroderma. Arthritis Rheum 2004;50:558-64.
9. Fonseca C, Renzoni E, Sestini P, et al. Endothelin axis polymorphisms in patients with scleroderma. Arthritis Rheum 2006;54:3034-42.
10. Tan FK, Wang N, Kuwana M, et al. Association of fibrillin 1 single-nucleotide polymorphism haplotypes with systemic sclerosis in Choctaw and Japanese populations. Arthritis Rheum 2001;44:893-901.
11. Fanning GC, Welsh KI, Bunn C, Du BR, Black CM. HLA associations in three mutually exclusive autoantibody subgroups in UK systemic sclerosis patients. Br J Rheumatol 1998;37:201-7.
12. Mattuzzi S, Barbi S, Carletto A, et al. Association of polymorphisms in the IL1B and IL2 genes with susceptibility and severity of SSc. J Rheumatol 2007;34:997-1004.
13. Lagan AL, Pantelidis P, Renzoni EA, et al. Single-nucleotide polymorphisms in the SPARC gene are not associated with susceptibility to scleroderma. Rheumatology Oxford 2005; 44:197-201.
14. Zhou X, Tan FK, Reveille JD, et al. Association of novel polymorphisms with the expression of SPARC in normal fibroblasts

and with susceptibility to scleroderma. Arthritis Rheum 2002;46:2990-9.
15. Gourh P, Tan FK, Assassi S, et al. Association of the PTPN22 R620W polymorphism with anti-topoisomerase I- and anticentromere antibody-positive systemic sclerosis. Arthritis Rheum 2006;54:3945-53.
16. Wipff J, Allanore Y, Kahan A, et al. Lack of association between the protein tyrosine phosphatase non-receptor 22 (PTPN22)*620W allele and systemic sclerosis in the French Caucasian population. Ann Rheum Dis 2006;65:1230-2.
17. Balada E, Simeon-Aznar CP, Serrano-Acedo S, et al. Lack of association of the PTPN22 gene polymorphism R620W with systemic sclerosis. Clin Exp Rheumatol 2006;24:321-4.
18. Hattersley AT, McCarthy MI. What makes a good genetic association study? Lancet 2005;366:1315-23.
19. McCarthy MI, Smedley D, Hide W. New methods for finding disease-susceptibility genes: impact and potential. Genome Biol 2003;4:119.
20. Salanti G, Amountza G, Ntzani EE, Ioannidis JP. Hardy-Weinberg equilibrium in genetic association studies: an empirical evaluation of reporting, deviations, and power. Eur J Hum Genet 2005; 13:840-8.
21. Healy DG. Case-control studies in the genomic era: a clinician's guide. Lancet Neurol 2006;5:701-7.