Available online http://arthritis-research.com/content/4/2/084

Commentary

Commentary on "Genetic linkage and transmission disequilibrium of marker haplotypes at chromosome 1q41 in human systemic lupus erythematosus", by RR Graham *et al.*

Anne C Barton and Jane Worthington

Arthritis and Rheumatism Campaign Epidemiology Unit, University of Manchester, Manchester, UK

Correspondence: Anne C Barton, ARC-EU, Stopford Building, University of Manchester M13 9PT, UK. Tel: +44 161 275 5037; fax: +44 161 275 5043; e-mail: ABarton@fs1.ser.man.ac.uk

Received: 14 August 2001
Revisions requested: 18 October 2001
Revisions received: 30 October 2001
Accepted: 5 November 2001
Published: 19 November 2001

Arthritis Res 2002, **4**:84-86

(Print ISSN 1465-9905; Online ISSN 1465-9913)

Abstract

Genome-wide linkage analysis studies in families with systemic lupus erythematosus (SLE) have revealed consistent evidence of linkage to several regions of the genome. In a previous issue of this journal, Graham and colleagues described their approach to following up the linkage data for one of these regions, 1q41–42. Using methods based on the transmission disequilibrium test, the region likely to harbour a SLE disease gene was refined to 2.3 Mb. This commentary discusses their approach and identifies lessons that may be applicable to the investigation of other complex diseases.

Keywords: association, linkage, systemic lupus erythematosus, transmission disequilibrium test, whole-genome scan

Introduction

Because linkage analysis approaches had been successful in the identification of disorders inherited as Mendelian traits, it was expected that the genetic basis of common diseases would be identified using a similar approach, but results to date may seem disappointing. As for most common diseases, susceptibility to autoimmunity is thought to be determined by both genetic and environmental factors. These autoimmune diseases tend not to be inherited in simple Mendelian fashion, but exhibit complex patterns of segregation. Investigation of these diseases can often be hampered by factors such as late age at disease onset, variable penetrance, variable phenotypic expression (different combinations of genes may predispose to different patterns of disease), unknown gene–gene and gene–environment interactions, genetic heterogeneity (different genes may produce the same phenotype), and misclassification of clinical phenotypes. Hence, the task of identifying susceptibility genes for complex disorders is enormous.

Investigation of genetic susceptibility loci for systemic lupus erythematosus

Twin and family studies suggest that systemic lupus erythematosus (SLE) has a substantial genetic susceptibility component [1–3]. Whole-genome scans of SLE families with affected sibling pairs have now been published, and, despite the relatively small sizes of the individual studies and the ethnic heterogeneity of the populations studied, there appears to be a surprising degree of overlap between findings [4–8]. All the studies have reported linkage to regions of the long arm of chromosome 1. In volume 3 issue 5 of this journal, Graham *et al.* described their approach to following up this linkage data for one of these regions, mapping to 1q41–42 [9].

Linkage analysis identifies genomic regions that are shared, identical-by-descent, by siblings affected by disease more often than would be expected by chance alone. However, linkage typically extends for 10 cM or more and such a region could contain 500 genes. Varia-

bp = base pairs; SLE = systemic lupus erythematosus; TDT = transmission disequilibrium test.

tion in any one of these genes could be responsible for the observed linkage. Association is the nonrandom cosegregation of alleles and assumes that populations are descended from a small founder group and that repeated recombinations over generations reduce the shared chromosomal segments to very small regions. Therefore, in order to detect an association, the marker and disease gene must be in linkage disequilibrium [10]. Because linkage disequilibrium extends for shorter distances (~60 Kbp from common coding variants in the North American population) [11], demonstration of association refines the region likely to harbour the disease gene. Linkage disequilibrium mapping can be carried out either by directly testing potential candidate genes or by using microsatellite markers mapping to a region of linkage. Going directly to candidate genes is fraught with danger. Virtually any gene could be a candidate, and sometimes functional genes appear to have an obscure role, e.g. *APOE* gene polymorphism and Alzheimer's disease [12].

The alternative approach taken by Graham *et al.* was to try to refine the ~16 cM region of linkage likely to harbour the disease gene by first investigating association with a number of microsatellite markers mapping to the region in 210 families with affected sibling pairs and 122 families with three affected members. Using extensions of the family-based association method, the transmission disequilibrium test (TDT) [13], they found strong evidence for association with one marker, D1S490, by all the TDT methods used. Haplotype analysis not only can increase the power to detect association but also can be used to localise the genetic region harbouring the disease gene. Association with three haplotypes spanning ~9 cM was demonstrated using two-marker approaches. When three-marker haplotypes were investigated, however, association with two different combinations of markers, spanning just 3 cM, was demonstrated. The equivalent physical distance is ~2.3 Mb. Reassuringly, linkage to the 1q41–42 region was largely accounted for by families carrying either of two risk haplotypes spanning the five markers. Even though the results presented in the study provide consistent and compelling evidence to support association to the region using a number of tests, it must be remembered that no correction has been applied for multiple testing, and confirmation of these findings in other data sets is required.

Lessons that can be drawn from this study

The study teaches us several important lessons. Firstly, it demonstrates the usefulness of animal models of disease in implicating candidate susceptibility regions in humans. The 1q41–42 region is homologous to a locus linked to a mouse model of lupus, and linkage in humans was first demonstrated after this area was targeted as a candidate susceptibility region using information from the mouse model [14]. Secondly, it is salutary to note that the linkage results for this region from analysis of whole-genome scans might have been discounted if stringent criteria had been applied [15]. In both whole-genome scans reporting linkage to the region, the LOD scores (logarithms of odds ratios) barely achieved statistically significant evidence for linkage [4–7]. However, replication of findings by independent groups is strong evidence that the findings are not due to a type-1 error. Identification of association with specific haplotypes of markers and demonstration that families with these haplotypes are largely responsible for the evidence of linkage support the hypothesis that true susceptibility genes may map to the region. Thirdly, this study demonstrates the superior ability of haplotype analysis to detect association over single-point methods. The gain in power from haplotyping arises in two main ways: analysis of single markers for tests of association using TDT-based methods can only use information from families in which transmissions are informative, i.e. when either the known or the inferred parental genotype is heterozygous at the locus under investigation. Haplotype methods can be more powerful, because transmission of a combination of markers is assessed, so that even if the inferred parental genotype is homozygous at one locus, it may not be at a second, third, or subsequent locus. The increase in power provided by haplotype methods also arises because there may be preferential allele transmission at two loci which, when analysed separately, do not achieve statistical significance, whereas a haplotype of alleles from the combination of markers may be strongly associated with disease.

Conclusion

Thus, from a linkage result that implicated an ~16 cM region, Graham *et al.* have refined the region likely to harbour an SLE disease gene to a manageable 2.3 Mb. A region this size is still likely to contain many candidate genes, so the task of identifying which is the disease gene is still huge. Demonstration of association with polymorphisms mapping to potentially functional domains of a gene may implicate it as the disease gene, but association does not necessarily imply causation (the association could arise due to linkage disequilibrium with a disease mutation in a nearby gene) and confirmation will require functional studies. Alternatively, the animal model in which the homologous region was first implicated may help in the identification of the disease gene.

Acknowledgements

Dr A Barton is in receipt of an MRC Clinical Research Fellowship. Dr J Worthington is funded by the Arthritis and Rheumatism Campaign.

References

1. Block SR, Winfield JB, Lockshin MD, D'Angelo WA, Christian CL: **Studies of twins with systemic lupus erythematosus: A review of the literature and presentation of 12 additional sets.** *Am J Med* 1975, **59**:533-552.
2. Deapen D, Escalante A, Weinrib L, Horwitz D, Bachman B, Roy-Burman P, Walke A, Mack TM: **A revised estimate of twin concordance in systemic lupus erythematosus.** *Arthritis Rheum* 1992, **35**:311-318.

3. Vyse TJ, Todd JA: **Genetic analysis of autoimmune disease.** *Cell* 1996, **85**:311-318.
4. Gaffney PM, Ortmann WA, Selby SA, Shark KB, Ockenden TC, Rohlf KE, Walgrave NL, Boym WP, Malmgren ML, Miller ME, Kearns GM, Messner RP, King RA, Rich SS, Behrens TW: **Genome screening in human systemic lupus erythematosus: results from a second Minnesota cohort and combined analyses of 187 sib-pair families.** *Am J Hum Genet* 2000, **66**:547-556.
5. Gaffney PM, Kearns GM, Shark KB, Ortmann WA, Selby SA, Malmgren ML, Rohlf KE, Ockenden TC, Messner RP, King RA, Rich SS, Behrens TW: **A genome-wide search for susceptibility genes in human systemic lupus erythematosus sib-pair families.** *Proc Natl Acad Sci U S A* 1998, **95**:14875-14879.
6. Gray-McGuire C, Moser KL, Gaffney PM, Kelly J, Yu H, Olson JM, Jedrey CM, Jacobs KB, Kimberly RP, Neas BR, Rich SS, Behrens TW, Harley JB: **Genome scan of human systemic lupus erythematosus by regression modelling. Evidence of linkage and epistasis at 4p16-15.2.** *Am J Hum Genet* 2000, **67**:1460-1469.
7. Moser KL, Neas BR, Salmon JE, Yu H, Gray-McGuire C, Asundi N, Bruner GR, Fox J, Kelly J, Henshall S, Bacino D, Dietz M, Hogue R, Koelsch G, Nightingale L, Shaver T, Abdou NI, Albert DA, Carson C, Petri M, Treadwell EL, James JA, Harley JB: **Genome scan of human systemic lupus erythematosus: evidence for linkage to chromosome 1q in African-American pedigrees.** *Proc Natl Acad Sci U S A* 1998, **95**:14869-14874.
8. Shai R, Quismorio FP Jr, Li L, Kwon OJ, Morrison J, Wallace DJ, Neuwelt CM, Brautbar C, Gauderman WJ, Jacob CO: **Genome-wide scan for systemic lupus erythematosus susceptibility genes in multiples families.** *Hum Mol Genet* 1999, **8**:639-644.
9. Graham RR, Langefeld CD, Gaffney PM, Ortmann WA, Selby SA, Baechler EC, Shark KB, Ockenden TC, Rohlf KE, Moser KL, Brown WM, Gabriel SE, Messner RP, King RA, Horak P, Elder JT, Stuart PE, Rich SS, Behrens TW: **Genetic linkage and transmission disequilibrium of marker haplotypes at chromosome 1q41 in human systemic lupus erythematosus.** *Arthritis Res* 2001, **3**:299-305.
10. Jorde LB: **Linkage disequilibrium as a gene mapping tool.** *Am J Hum Genet* 1995, **56**:11-14.
11. Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumjian R, Farhadian SF, Ward R, Lander ES: **Linkage disequilibrium in the human genome.** *Nat Genet* 2001, **411**:199-204.
12. Corder EH, Saunders AM, Risch NJ, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, Roses AD, Haines JL, Pericak-Vamce MA: **Gene dosage of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families.** *Science* 1993, **261**:921-923.
13. Speilman RS, McGinnis RE, Ewens WJ: **Transmission disequilibrium test for linkage disequilibrium: the insulin gene region and insulin dependent diabetes mellitus (IDDM)**. *Am J Hum Genet* 1993, **52**:506-516.
14. Tsao BP, Cantor RM, Kalunian KC, Chen CJ, Badsha H, Singh R, Wallace DJ, Kitridou RC, Chen SL, Shen N, Song YW, Isenberg DA, Yu CL, Hahn BH, Rotter JI: **Evidence for linkage of a candidate chromosome 1 region to human systemic lupus erythematosus.** *J Clin Invest* 1997, **99**:725-31.
15. Lander ES, Kruglyak L: **Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results**. *Nat Genet* 1995, **11**:241-247.