Annals of the Rheumatic Diseases 1991; **50**: 343–346

Annals of the
RHEUMATIC DISEASES

Leader

Polygenic susceptibility in rheumatoid arthritis

Many common disorders, including such scourges of Western society as ischaemic heart disease[1] and many forms of cancer,[2] result in part from heritable factors. Of particular interest to the rheumatologist is the group of complex, multifactorial, polygenic disorders comprising the 'autoimmune diseases', of which rheumatoid arthritis (RA) is probably the commonest. Similarities between this form of inflammatory arthritis and others, known to be triggered by infectious agents, have inevitably fuelled speculation that RA may have a similar cause. To date, however, the identification of infectious triggers has been conspicuously lacking. In contrast, over the past two decades there has been considerable progress in identifying at least some of the genetic factors contributing to susceptibility to RA. Thus the immune response genes of the HLA class II region of the major histocompatibility complex (MHC) are quite clearly implicated in the pathogenesis of this condition, though their precise role is, as yet, unclear.

What is the evidence for heritable influences in RA?

Rheumatoid arthritis has a worldwide distribution and has a strikingly similar prevalence (about 1%) in most populations studied. The observation that rural Africans have much lower prevalence of the disease than their counterparts who migrate to the towns illustrates the importance of environmental factors in the pathogenesis of the disease.[3 4] At least 11% of probands with RA have an affected first degree relative,[5–7] but such familial clustering might reflect either genetic or environmental influences. There is probably a polygenic influence in susceptibility to RA, with subjects carrying the threshold number of requisite genes being at risk of developing the disease. Possibly, also, other subjects carrying yet more of the genes associated with RA might be at risk of developing particularly severe disease. Supporting evidence of this concept comes from the following observations.

Firstly, familial clustering is most obvious where the probands have severe disease, suggesting that a greater number of the requisite genes are segregating within these families. This makes it more likely that other members of the same family might have inherited a sufficient number of the susceptibility genes to develop at least mild RA. Lawrence noted that an excess sibling recurrence risk (K_s) was confined to families in which the probands had severe disease—that is, seropositive or erosive RA.[8] In cases where the proband had mild disease K_s was barely above the general population risk (K), but the ratio K_s:K (or λ_s) gradually increased with increasing disease severity in the proband. Thus the value for λ_s was about 3 where the proband was seropositive and nearly 7 where the proband had seropositive, erosive RA, in keeping with polygenic susceptibility.

Secondly, comparisons of concordance rates in monozygotic and dizygotic twins strongly support a heritable contribution to susceptibility. Although twin studies are notoriously difficult to interpret, not least because of ascertainment bias, there is remarkable consistency in the relative levels of concordance for monozygotic and dizygotic twins. Thus monozygotic twins are at least four times more likely to be concordant than dizygotic twins, who have a similar recurrence risk to their non-twin siblings.[8 9] Such differences almost certainly reflect the genetic differences between dizygotic, but not monozygotic, twins as both types of twin are likely to share a similar environment.

How important are genetic factors in RA?

The absolute prevalences suggested by population studies inevitably vary as a consequence of the diagnostic criteria applied and the study design. The generally accepted concordance rates in monozygotic twins (12–30%) suggest a major extrinsic influence.[8 9] A major criticism of these studies, however, is their cross-sectional design. This probably leads to a serious underestimate of the true level of concordance because there may be a lapse of several years, or even decades, between the first and second twins developing RA.[10] Therefore the genetic contribution to RA is probably considerably greater than has previously been estimated and might be 50% or even more.

What is known of the genetic loci responsible for susceptibility to RA?

An association between RA and the immune response genes found within the MHC on the short arm of chromosome 6 has been known for many years. Initially, this association between certain class II HLA antigens and RA was defined by cellular typing methods in the mixed lymphocyte reaction (Dw4)[11] and by serological reagents (DR4).[12] During the 1980s these associations were confirmed repeatedly in many different groups,[13–20] whereas other, generally weaker, associations were described with DR1[21–25] and DRw10.[26] Furthermore, the DR4 serotype can be subdivided by cellular methods or DNA oligonucleotide typing, and only certain of these subtypes are associated with the tendency to develop RA.[27–30]

These associations must have been surprising when they were first described 15 years ago but much has subsequently been learnt of the structure and function of the HLA class II molecules to improve our understanding of the likely nature of these associations. The class II HLA molecules have a crucial role in the selection of the mature T lymphocyte receptor repertoire during thymic development and also in the self-restricted presentation of foreign and self antigens to these same T lymphocytes in later life.[31 32] By analogy with the known crystallographic structure of the HLA class I molecule HLA-A2,[33] the likely structure of the HLA class II molecules, including those of the DR series, can be inferred.[34] Antigen is processed within specialised cells (including macrophages and B lymphocytes) and 'presented' to T lymphocytes as a short peptide (about 13 amino acids in length), nestling in a cleft on the surface of the HLA molecule. The exact mechanisms by which T lymphocytes become tolerant of self antigens presented in this way remain to be clarified, but a breakdown in tolerance is thought to underlie the autoimmune disorders.[35]

The polymorphic residues around the antigen binding cleft account for many of the different specificities recognised by both serological and cellular typing reagents. Therefore we would expect different HLA molecules to select different T cell receptor repertoires and to be particularly efficient at presenting certain antigens. The observation that all DR molecules associated with RA share a linear amino acid sequence along one side of the antigen binding cleft (derived from the β chain of the DR molecule) lends credence to the concept that these molecules may be involved in the presentation of a limited series of peptides associated with RA within the joint.[27 28 30] Furthermore, the repertoire of T cell receptors found in the joint will already have been heavily influenced by the same DR epitope during thymic selection of T lymphocytes. One obvious corollary of this hypothesis is that a contribution towards the susceptibility to RA should also come from the T cell receptor α and β chains. Studies in multiple sclerosis, another chronic disorder with many immunogenetic similarities to RA, have provided corroborative evidence for this point of view. In multiple sclerosis there is a strong population association with the HLA-DR2 antigen, and associations have been shown with both T cell receptor α[36] and β[37] loci. In RA some studies have suggested the presence of specific T cell receptor rearrangements within the synovial tissues, but others have not.[38 39] Definitive studies have not yet been performed.

The role of the DQ locus in overall susceptibility to RA has been somewhat controversial. The DQ3 serotype found almost invariably on DR4 haplotypes in white subjects can be split into the DQw7 and DQw8 subtypes, but neither of these is preferentially associated with RA.[27 28] In contrast, in severe forms of RA (Felty's syndrome or those with vasculitis) there seems to be a preferential association with DQw7, in addition to an association with DR4 which approaches 100%.[40 41] Currently it is unclear whether this reflects a true association with DQw7 (as well as DR4), a preferential association with the Dw4 subtype of DR4 (which is in linkage disequilibrium with DQw7), or whether DQw7 is a marker for another locus within the MHC. The last hypothesis is particularly attractive in view of the recent demonstration of many novel genes in this region, some of which have immunological significance—for example, heat shock protein 70[42] and tumour necrosis factor.[43] Likewise, the presence of a 110 kilobase insertion of DNA on DR4 haplotypes, demonstrable by pulse field gel electrophoresis, might be relevant.[44] To date, studies have not suggested a major role for other loci, such as complement or tumour necrosis factor, within the MHC. Loci outside the MHC have not yet been studied in a systematic fashion, but the possibility of associations between RA and loci on chromosome 14 has been suggested.[45]

What proportion of the genetic contribution to RA comes from outside the MHC?

The contribution of MHC linked genes to RA can be estimated by the method of Risch.[46] This allows an estimate of the recurrence risk of RA in siblings, sharing no HLA haplotypes identical by descent (IBD), to be made according to the formula:

$$P(\text{IBD}=0/\text{sib pair})=\Phi_s \div \lambda_s$$

where Φ_s is the expected proportion of sibling pairs sharing no HLA alleles (IBD)—that is, 0·25 and λ_s is the excess sibling recurrence risk over that seen in the general population. The value of λ_s is highly dependent on the criteria used to define RA but probably lies between 5 and 10. Taking Lawrence's figure of 7 where the proband has seropositive, erosive RA[8] we have:

$$P(\text{IBD}=0/\text{sib pair})=0{\cdot}25 \div 7=0{\cdot}035$$

This is less than one third of the observed figure of 0·11 from the combined analysis of 206 sibling pairs reported in five separate studies,[15 47–50] implying that non-MHC genes are responsible for a major contribution towards susceptibility to RA. A similar estimate of the contribution arising from the MHC (28%) has been derived using a mathematical formula which essentially contrasts the recurrence risk in genetically identical subjects (monozygotic twins) and subjects who are just HLA identical.[51] It is therefore likely to prove worthwhile to instigate a systematic search for non-MHC genes in RA.

How can one identify new genes contributing towards RA?

Genetic linkage analysis has proved to be remarkably effective in defining DNA markers for a large number of monogenic disorders, including many where virtually nothing was known of the underlying biochemical abnormality or chromosomal localisation of the causal gene. Remarkably, it has proved possible, by a process of 'reverse genetics', to move from a series of tightly linked DNA markers to identify the actual genes responsible for certain conditions, including chronic granulomatous disease[52] and cystic fibrosis.[53] Dramatic success has also been achieved in immunologically mediated diseases with the definition of DNA markers on chromosome 11 tightly linked to a major locus contributing to atopy.[54] Unfortunately, a traditional approach to linkage analysis in RA is not possible because of the paucity of extended pedigrees with the disease. Consequently, an alternative technique must be adopted when trying to show linkage of RA to a particular locus, that of non-random sharing of haplotypes between affected siblings.[47–49] This method has been successfully used to show linkage between RA and HLA and could equally well be applied to the analysis of other regions of interest in the genome. There are three major problems to be overcome before this method can be applied successfully. Firstly, the ascertainment of a sufficiently large number of suitable families with at least two affected siblings to make the analysis statistically valid. Secondly, the development of highly polymorphic DNA markers to maximise the information which can be gained from pedigrees, which will inevitably be small. Thirdly, the development of powerful computer based multipoint analysis, which allows the construction of high resolution genetic linkage maps of regions of interest to be constructed.

Fortunately, most of these problems can now be resolved, opening the way for a systematic search to begin. The Arthritis and Rheumatism Council is investing in a national repository at the ARC Epidemiology Research Unit in Manchester for the storage of material, including cell lines and DNA from families with multiple cases of RA. A nationwide search is now on for families with at least two affected siblings and both parents alive (which maximises the chance of allocating parental haplotypes). Because of the small amount of information which can be harvested from such small families individually it will be necessary to analyse about 100 families in this way to be sure of establishing whether specific loci are indeed linked to RA.

Highly polymorphic regions of DNA abound throughout the genome as variable nucleotide tandem repeats ('mini-satellites')[55] and dinucleotide repeats ('microsatellites'),[56] providing a wealth of genetic markers for use in these studies. As progress is made in the effort to map the human genome we can confidently expect even more to become available. Another potential spin off from this project may be the identification and chromosomal localisation of genes of potential importance in the pathogenesis of RA ('candidate genes'). These could then be analysed for linkage to RA in the same way as other candidates, such as T cell receptors α and β, by identifying markers in the flanking regions of DNA.

Paterson has described the methodology for resolving quantitative traits to discrete genetic loci by multipoint linkage analysis.[57] Others have adapted this approach to the analysis of the human polygenic disease insulin dependent diabetes mellitus using a high resolution genetic linkage map of a region of interest (chromosome 11q) and analysis of affected sibling pairs.[58] This approach is directly applicable to the study of potential susceptibility loci in RA and could be extended to produce a 20 chromosome linkage map of the complete genome in time.

Over the past few years dramatic progress has been made in mapping and characterising the genes responsible for many of the major monogenic disorders in man. It is perhaps in the field of polygenic, multifactorial disorders, however, that the molecular biologist will find his greatest challenges over the next decade. We now have the technology to dissect the genetics of susceptibility to RA as never before, and this is an opportunity which the rheumatology community should not miss.

Molecular Immunology Group,
Institute of Molecular Medicine,
John Radcliffe Hospital, Oxford OX3 9DU

PAUL WORDSWORTH
JOHN BELL

Addendum

The association of RA with HLA-DRw10 has recently been confirmed in English and Spanish populations. Furthermore, study of the Yakima Amerindians has shown a strong association with the rare DR6/Dw16 allele, which shares the same amino acid sequence near the antigen binding cleft as other RA associated DR molecules.

1 Berg K. Genetics of coronary heart disease and its risk factors. In: Brock G, Collins G M, eds. *Molecular approaches to human polygenic disease*. Chichester: Wiley, 1987: 14–33. (Ciba Foundation Symposium 130.)
2 Faeron E R, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990; **61**: 759–67.
3 Brighton S W, De La Hope A L, van Staden D J, Badenhorst J H, Myers O L. The prevalence of rheumatoid arthritis in a rural African population. *J Rheumatol* 1988; **15**: 405–8.
4 Myers O L, Daynes G, Beighton P. Rheumatoid arthritis in a tribal Xhosa population in the Transkei, Southern Africa. *Ann Rheum Dis* 1977; **36**: 62–5.
5 Wolfe F, Kleinheksel S M, Khan M A. Prevalence of familial occurrence in patients with rheumatoid arthritis. *Br J Rheumatol* 1988; **27** (suppl 2): 150–2.
6 Deighton C M, Wensel J, Roberts D F, Walker D J. The concordance rates for rheumatoid arthritis in same-sexed and HLA-identical sibships [Abstract]. *Br J Rheumatol* 1989; **28** (suppl 2): 80.
7 Sanders P A, Grennan D M, Dyer P A, Thomson W, Delange G G. A comparison of clinical and immunogenetic features in familial and sporadic rheumatoid arthritis. *J Rheumatol* 1987; **14**: 718–22.
8 Lawrence J S. Rheumatoid arthritis: nature of nurture? *Ann Rheum Dis* 1970; **29**: 357–79.
9 Aho K, Koskenvuo M, Tuominen J, Kaprio J. Occurrence of rheumatoid arthritis in a nationwide survey of twins. *J Rheumatol* 1986; **13**: 899–902.
10 Tarp V, Grauda H K. Seropositive rheumatoid arthritis in monozygotic twin sisters carrying HLA-DR4 associated with B27. *J Rheumatol* 1989; **16**: 1530–2.
11 McMichael A J, Sasazuki T, McDevitt H O, Payne R O. Increased frequency of HLA-Cw3 and HLA-Dw4 in rheumatoid arthritis. *Arthritis Rheum* 1977; **20**: 1037–42.
12 Stastny P. Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. *N Engl J Med* 1978; **298**: 869–71.
13 Jaraquemada D, Ollier W, Awad J, *et al.* HLA and rheumatoid arthritis: a combined analysis of 440 British patients. *Ann Rheum Dis* 1986; **45**: 627–36.
14 Maeda H, Juji T, Mitsui H, Sonozaki H, Okitsu K. HLA-DR4 and rheumatoid arthritis in Japanese people. *Ann Rheum Dis* 1981; **40**: 299–302.
15 Mehra N K, Kaul R, Chaudhuri T K, Jhinghan B, Malaviya A N. HLA-linked susceptibility to rheumatoid arthritis in North India. *Tissue Antigens* 1987; **31**: 220–3.
16 Martell R W, Du Toit E D, Kalla A A, Myers O L. Association of rheumatoid arthritis and HLA in three South African populations—whites, blacks and a population of mixed ancestry. *S Afr Med J* 1989; **76**: 189–90.
17 Ueno Y, Iwaki Y, Terasaki P, *et al.* HLA-DR4 in negro and Mexican rheumatoid arthritis patients. *J Rheumatol* 1981; **8**: 804–7.
18 Harvey J, Arnett F S, Bias W B, Hsu S H, Stevens M B. Heterogeneity of HLA-DR4 in the rheumatoid arthritis of a Chippewa band. *J Rheumatol* 1981; **8**: 797–803.
19 Sathar M A, Al-Saffar M, Guinsi R T, Sugathan T N, White A G, Behbehani K. Histocompatibility antigens (A, B, C, and DR) in Arabs with rheumatoid arthritis. *Dis Markers* 1990; **8**: 11–15.
20 Massardo L, Jacobelli S, Rodriguez L, Rivero S, González A, Marchetti R. Weak association between HLA-DR4 and rheumatoid arthritis in Chilean patients. *Ann Rheum Dis* 1990; **49**: 290–2.
21 Jajic Z, Jajic J. Antigens of HLA-DR locus in rheumatoid arthritis. *Br J Rheumatol* 1988; **27** (suppl 2): 139–40.
22 Nunez-Roldan A, Arquer Zanazua E, Arroya E. Estudio de las antigens HLA-DR en la arteritis reumatoided. *Revista Espanola Reumatologica* 1982; **9**: 23–7.
23 Vischer T. HLA-DR antigens in rheumatoid arthritis: a Swiss collaborative study. *Rheumatol Int* 1986; **6**: 89–92.
24 Woodrow J A, Nichol F E, Zaphiropoulos G. DR antigens and rheumatoid arthritis: a study of two populations. *BMJ* 1981; **283**: 1287–9.
25 Schiff B, Mizrachi Y, Orgad S, Gazit E. Association of HLA-Aw31 and HLA-DR1 with adult rheumatoid arthritis. *Ann Rheum Dis* 1982; **41**: 403–4.
26 Stevens C, Ollier W, Awad J, *et al.* Class II HLA antigens associated with rheumatoid arthritis in Asians [Abstract]. *Br J Rheumatol* 1989; **28** (suppl 2): 80.
27 Wordsworth B P, Lanchbury J S S, Sakkas L I, Welsh K I, Panayi G S, Bell J I. HLA-DR4 subtype frequencies in rheumatoid arthritis indicate that DRB1 is the major susceptibility locus within the HLA class II region. *Proc Natl Acad Sci USA* 1989; **86**: 10049–53.
28 Nepom G T, Byers P, Seyfried C, *et al.* HLA genes associated with rheumatoid arthritis. *Arthritis Rheum* 1989; **32**: 15–21.
29 Ohta N, Nishimura Y K, Tanimoto K, *et al.* Association between HLA and Japanese patients with rheumatoid arthritis. *Hum Immunol* 1982; **5**: 123–32.
30 Watanabe Y, Tokunago K, Matsuki K, *et al.* Putative amino acid sequence of HLA-DRB chain contributing to rheumatoid arthritis susceptibility. *J Exp Med* 1989; **169**: 2263–8.
31 Blackman M, Kappler J, Marrack P. The role of the T-cell receptor in positive and negative selection of developing T-cells. *Science* 1990; **248**: 1335–41.
32 Von Boehmer H, Kisielow P. Self-nonself discrimination by T cells. *Science* 1990; **248**: 1369–73.

33 Bjorkman P J, Saper M A, Samraoui B, Bennett W S, Strominger J L, Wiley D. The foreign antigen binding site and T-cell recognition regions of class I histocompatibility molecules. *Nature* 1987; **329:** 512–8.
34 Brown J H, Jardetsky T, Saper M, Samraoui B, Bjorkman P, Wiley D. A hypothetical model of the foreign antigen binding site of class I histocompatibility molecules. *Nature* 1988; **332:** 845–50.
35 Sinha A A, Lopez T, McDevitt H O. Autoimmune diseases: the failure of self tolerance. *Science* 1990; **248:** 1380–8.
36 Oksenberg J R, Sherritt M, Begovich A B, *et al.* T-cell receptor Vα and Cα alleles associated with multiple sclerosis and myasthenia gravis. *Proc Natl Acad Sci USA* 1989; **86:** 988–92.
37 Seboun E, Robinson M A, Doolittle T H, Ciulla T A, Kindt T J, Hauser S L. A susceptibility locus for multiple sclerosis is linked to the T-cell receptor β chain complex. *Cell* 1989; **57:** 1095–100.
38 Stamenkovic I, Stegagno M, Wright K A, *et al.* Clonal dominance among T-lymphocyte infiltrates in arthritis. *Proc Natl Acad Sci USA* 1988; **85:** 1179–83.
39 Duby A D, Sinclair A K, Osborne-Lawrence S L, Zelders W, Kan L, Fox D A. Clonal heterogeneity of synovial fluid lymphocytes from patients with rheumatoid arthritis. *Proc Natl Acad Sci USA* 1989; **86:** 6206–10.
40 So A K, Warner C A, Sansom D, Walport M J. DQ beta polymorphism and genetic susceptibility to Felty's syndrome. *Arthritis Rheum* 1988; **31:** 990–4.
41 Sansom D M, Amin S N, Bidwell J L, *et al.* HLA-DQ related restriction fragment length polymorphisms in rheumatoid arthritis: evidence for a link with disease severity. *Br J Rheumatol* 1989; **28:** 374–9.
42 Sargent C A, Dunham I, Trowsdale J, Campbell R D. Human major histocompatibility complex contains gene for the major heat shock protein HSP70. *Proc Natl Acad Sci USA* 1989; **86:** 1968–72.
43 Spies T, Morton C C, Nedospasor S A, Fiers N, Pious D, Strominger J L. Genes for the tumour necrosis factors α and β are linked to the human major histocompatibility complex. *Proc Natl Acad Sci USA* 1986; **83:** 8699–703.
44 Dunham I, Sargent C A, Dawkins R L, Campbell R D. An analysis of variation in the long-range genomic organization of the human major histocompatibility complex class II region by pulsed-field gel electrophoresis. *Genomics* 1989; **5:** 787–96.
45 Ollier W, Thomson W, Welch S, De Lange G G, Silman A. Chromosome 14 markers in rheumatoid arthritis. *Ann Rheum Dis* 1988; **47:** 843–8.
46 Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. *Am J Hum Genet* 1987; **40:** 1–15.
47 Payami H, Thomson G, Khan M A, *et al.* Genetics of rheumatoid arthritis. *Tissue Antigens* 1986; **27:** 57–63.
48 Ollier W, Silman A, Gosnell N, *et al.* HLA and rheumatoid arthritis: an analysis of multicase families. *Dis Markers* 1986; **4:** 85–98.
49 McDermott M, Molloy M, Cashin P, *et al.* A multicase family study of rheumatoid arthritis in South West Ireland. *Dis Markers* 1986; **4:** 103–11.
50 Walker D J, Burn J, Griffiths I D, Roberts D F, Stephenson A M. Linkage studies of HLA and rheumatoid arthritis in multicase families. *Arthritis Rheum* 1987; **30:** 31–5.
51 Deighton C M, Walker D J, Griffiths I D, Roberts D F. The contribution of HLA to rheumatoid arthritis. *Clin Genet* 1989; **36:** 178–82.
52 Orkin S H. Reverse genetics and human disease. *Cell* 1986; **47:** 845–50.
53 Karem B-S, Rommens J M, Buchanan J A, *et al.* Identification of the cystic fibrosis gene: genetic analysis. *Science* 1989; **245:** 1073–80.
54 Cookson W O, Sharp P A, Faux J A, Hopkin J M. Linkage between immunoglobulin E responses underlying asthma and rhinitis and chromosome 11q. *Lancet* 1989; **i:** 1292–5.
55 Wells R A, Green P, Reeders S T. Simultaneous genetic mapping of multiple human minisatellite sequences using DNA fingerprinting. *Genomics* 1989; **5:** 761–72.
56 Weber J L, May P E. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *Am J Hum Genet* 1989; **44:** 388–96.
57 Paterson A H, Lander E S, Hewitt J D, Peterson S, Lincoln S E, Tanksley S D. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction length polymorphisms. *Nature* 1988; **335:** 721–6.
58 Hyer R N, Julier C, Buckley J D, *et al.* High resolution linkage analysis in human polygenic disease: insulin-dependent diabetes mellitus and chromosome 11q. *Am J Hum Genet* (in press).