



Review

Genetic Basis for Rheumatoid Arthritis

DHARAM P. SINGAL, JIANPING LI and YAOHUI ZHU

Department of Pathology and Molecular Medicine, McMaster University, 1200 Main Street West, Hamilton, ON, Canada L8N 3Z5

Abstract. Rheumatoid arthritis (RA) is a common disabling disorder of unknown etiology. In the past 2 decades, a number of studies have examined the genetic basis for RA. One major focus of these studies has been to identify genes within the MHC class II (HLA-DR) chromosomal region, which confer susceptibility/resistance to RA. A strong association between HLA-DR4 and adult seropositive RA has been observed in majority of populations. In addition, there is evidence of a positive association between HLA-DR1 and RA. On the basis of prevalence of DR1 (B1*0101) and of subtypes of DR4 (B1*0401, B1*0404 and B1*0405), it has been suggested that a five amino acid sequence motif (QKRAA/QRRRA) from position 70 to 74 in the third hypervariable region of DR β 1 molecules is associated with susceptibility to RA. These associations between RA and HLA-DR genes are however incomplete in that about 1/4 of patients do not carry RA-susceptibility DRB1 epitope. Since MHC class III region contains genes that are involved in immune response, we have recently examined the role of a number of microsatellites (D6S273, Bat2, TNFa) and HSP70 promoter region alleles in susceptibility to RA. The results demonstrate that two regions in MHC, class II (DR β 1) and class III (D6S273, HSP70, Bat2, TNFa) more completely define the risk for development of RA.

Key words: rheumatoid arthritis; HLA-DRB1; D6S273; Bat2; TNFa; HSP70.

Rheumatoid arthritis (RA) is a chronic inflammatory joint disease caused by autoreactive tissue-destructive immune response(s). RA is a relatively common disease with worldwide distribution and an adult prevalence of approximately 1%. The disease occurs as a chronic articular inflammatory disorder characterized by symmetric joint involvement and a variable frequency of extra-articular features such as rheumatoid nodules, vasculitis and scleritis. Although the precise etiology of the disease is not yet known, it is clearly multifactorial, with contributions from both genetic and non-genetic factors. The development of RA, therefore, depends on interactions between genetic elements in a susceptible individual as well as non-genetic (e.g. infective) factors. In the past two decades, considerable progress has been made towards elucidation of the genetic basis for susceptibility to RA. One major focus

of these studies has been to identify genes in the major histocompatibility complex (MHC), particularly the immune response genes within the HLA class II region, which confer susceptibility to RA.

Major Histocompatibility Complex

The human MHC, located on the short arm of chromosome 6 (6p21.3), includes genes in class I and class II regions that encode for class I (HLA-A, B, C) and class II (HLA-DR, DQ, DP) antigens⁵. The class III region is located between class I and class II regions and contains genes that encode for proteins involved in immune response, e.g. TNF-A and B, three members of the heat shock 70 (HSP70) family and the complement genes. The MHC region is characterized both by

a high degree of polymorphism and by strong linkage disequilibrium between different class I, class II and class III region alleles. For example, a number of extended haplotypes that carry different genes belonging to these regions have been defined.

HLA class I and class II antigens are heterotrimeric membrane glycoproteins expressed on the cell surface¹². The class I and class II antigens differ with respect to their structure, distribution and function. For example, class I antigens are comprised of a polymorphic α chain of about 44 kDa and an invariant β_2 -microglobulin (12 kDa) encoded by a gene on chromosome 15. In contrast, class II molecules are composed of an α chain of about 33 kDa and a polymorphic β chain of 28 kDa. In addition, there is an antigenic peptide associated with class I and class II molecules. Class I antigens, present on all nucleated cells and platelets, function in antigen presentation to cytotoxic T lymphocytes and are primarily involved in cell-mediated cytotoxicity. On the other hand, class II molecules are expressed on antigen-presenting cells, such as B lymphocytes and macrophages. They present antigens to T helper lymphocytes and thus play a crucial role in initiating specific immune responses and for the generation of the T cell repertoire in the thymus. The DR antigens of class II region are comprised of a non-polymorphic α chain encoded by a single DRA gene, and in most cases of two functional polymorphic β chains encoded by two DRB genes. Of these, one gene (DRB1) is highly polymorphic, while the other (DRB3, DRB4, DRB5) shows little polymorphism. Nucleotide sequence analysis has shown that most allelic variability is restricted to 3 major diversity regions: the first hypervariable region is clustered around residues 9–13, the second around residues 25–38, and the third hypervariable region is clustered around residues 57–88. In the third diversity region, most of the polymorphism is centred around position 70. These polymorphic residues correlate with functionally important sites either on the β sheets or on the floor of the cleft or along the α helices, as shown by crystallographic analysis of class II molecules². For example, most polymorphic residues in the first and second hypervariable regions correlate with the β sheet conformation and lie at the base of the cleft pointing into the cleft from the β sheets, whereas the polymorphic residues in the third diversity region are located on the edge of the peptide binding cleft. These residues may face towards the cleft or upward away from the cleft, depending on their position in the α helix.

Based on polymorphisms in 3 hypervariable regions of the β chains, a number of DR specificities (DR1

through DR18), each with a number of alleles, have been identified. Since DR molecules bind and present antigenic peptides to T cells, the polymorphisms in DRB genes have functional consequences. For example, polymorphism in the third hypervariable region influences antigenic peptide binding as well as interaction with a T cell receptor molecule. It has been shown by responses in mixed lymphocyte culture assay that the serologically defined determinant DR4 is heterogeneous and comprises a number of alleles (e.g. Dw4, Dw10, Dw13, Dw14, Dw15) that differ for residues in the third hypervariable region. These results suggest that the third diversity region differences even at a single position on the peptide binding cleft in the α helical region can influence T cell responses.

There is strong linkage disequilibrium between the alleles of the DR and DQ subregions. For example, DR3 is associated with DQ2, and DR4 is associated with DQ7 and with DQ8.

The MHC class III region contains genes that encode proteins, such as the complement proteins C2, factor B and C4, members of the HSP70 family, and the cytokines tumor necrosis factor (TNF) and lymphotoxin A and B. In addition, over 50 genes, including microsatellite markers, have been located in the class III region.

Class II Region and Susceptibility to Rheumatoid Arthritis

In the past 20 years, hundreds of original research papers, some with conflicting results, and a number of review articles have been published which address the role of class II antigens in the development of RA. To meet the space requirements of the present review, we have chosen to cite only those review articles, which contain references to all the original papers and which address pertinent issues such as association of DR genes with susceptibility to RA in different ethnic groups and the identification of the RA-susceptibility QKRAA/QRRAA epitope. The role of class II region antigens in susceptibility to RA will be reviewed in 3 areas: 1) HLA-DR genes and RA; 2) HLA-DM genes and RA; and 3) level of expression of DR genes and RA.

HLA-DR genes and RA. The association of RA with both HLA-DR4 and its subtype Dw4 was first defined by STASTNY²². Subsequently, the association between adult seropositive RA and DR4 has been observed in a majority of populations belonging to different ethnic groups, including North American and Northern European Caucasians, North American Blacks, Japanese and Hispanics, although some exceptions, e.g. Jews, Arabs,

Table 1. Amino acid differences in the third hypervariable region of HLA-DR4, DR1, DR14-Dw16 and DR10 DRB1 genes

DR specificity	DRB1 gene	Amino acid position							
		67	68	69	70	71	72	73	74
DR4 Dw4	B1*0401	L	L	E	Q	K	R	A	A
DR4 Dw10	B1*0402	I	–	–	D	E	–	–	–
DR4 Dw13	B1*0403	–	–	–	–	R	–	–	E
DR4 Dw14	B1*0404	–	–	–	–	R	–	–	–
DR4 Dw15	B1*0405	–	–	–	–	R	–	–	–
DR1 Dw1	B1*0101	–	–	–	–	R	–	–	–
DR14 Dw16	B1*1402	–	–	–	–	R	–	–	–
DR10	B1*1001	–	–	–	R	R	–	–	–

L – leucine, E – glutamic acid, Q – glutamine, K – lysine, R – arginine, A – alanine, I – isoleucine. Dashes indicate identity with DR4 Dw4 (B1*0401) allele.

Yakima Indians, have been documented^{7, 10, 14, 15, 24}. The most probable explanation for these differences is the prevalence of subtypes of DR4 (Dw4, Dw10, Dw13, Dw14, Dw15) in these ethnic groups. These subtypes show variations in the third hypervariable region of the DR β 1 chain (Table 1). It is evident that the Dw10 (B1*0402) sequence, present in high frequency in the Israeli population, is different from other DR4 subtypes in that it carries 2 negatively charged residues (D and E) at positions 70 and 71. Similarly, Dw13 (B1*0403) has a negatively charged residue (E) at position 74. On the other hand, Dw4 (B1*0401), Dw14 (B1*0404) and Dw15 (B1*0405) have similar sequences. Of these, Dw4 and Dw14 are associated with RA in Caucasians, and the Dw15 subtype found predominantly in Japanese is also associated with RA. In contrast, the Dw10 and Dw13 subtypes do not confer risk for development of RA. This region of the DR4 molecule (QKRAA/QRRAA), therefore, seems to be associated with susceptibility to RA.

This observation is further supported by analysis of the non-DR4 alleles: DR1 Dw1 (B1*0101), DR14 Dw16 (B1*1402) and DR10 (B1*1001), which share the sequence QKRAA/QRRAA from positions 70 to 74 with Dw4, Dw14 and Dw15 (Table 1). A number of investigators observed an increased prevalence of DR1 in RA patients of Caucasian, Jewish, Asian, Greek and Mexican origin^{7, 14}. In contrast, almost an equal number of reports did not observe an increase in frequency of DR1 in RA patients^{10, 15}. Investigations from our laboratory on relatively homogenous groups of RA patients, based on the severity of the disease and the patient's response to disease-modifying drugs, showed that DR1 was increased only in patients mildly afflicted and not in patients with severe RA^{3, 15}. Similarly, DR10 has also been found to be associated with RA. In addition, DR14 Dw16 has been associated with susceptibility to RA in Yakima Indians.

In summary, 4 groups of DRB genes: DR1, DR4 (Dw4, Dw14, Dw15), DR10 and DR14 (Dw16), have been found to be associated with RA in various populations. All these DRB1 genes share a common amino acid sequence motif, QKRAA/QRRAA, at positions 70 to 74 in the third hypervariable region. This structural motif is therefore implicated in an aberrant immune response involved in the development of RA.

Although DR4 is associated with adult seropositive RA, discrepancies exist with respect to seronegative RA. GRAN and HUSBY⁸ considered these discrepancies due to differences in the diagnostic criteria used for patient selection. These authors then set out diagnostic criteria for seronegative RA: follow-up for at least 3 years, during which at least 3 tests for IgM RF should be negative, joint erosions and exclusion of patients with clinical or radiological evidence or a family history of ankylosing spondylitis, reactive arthritis, psoriatic arthritis or inflammatory bowel disease. Studies that met these diagnostic criteria showed a remarkable concordance of results, i.e. increased prevalence of DR4 in seronegative RA⁴. Both seropositive and seronegative RA may therefore share the same immunogenetic basis.

We have described a significantly higher prevalence of DR4-associated DQ7 in RA patients who had moderately severe affliction. Similarly, a higher prevalence of DR4-associated DQ7 has been described in RA patients with severe affliction with nodules or erosions and in Felty's syndrome. Some investigators did not observe an increase in DR4-associated DQ7 in RA; the clinical status of the patients was, however, not defined in these studies. A critical analysis of the available results suggests that these differences may be due to the severity of the disease in that DR4-associated DQ7 is increased only in patients with severe affliction^{3, 15}.

HLA-DM genes and RA. Since HLA-DM (DMA and DMB) genes, located in the class II region, are

involved in the class II antigen presentation pathway, we and others examined the contribution of these gene polymorphisms in susceptibility to RA^{19, 23, 25}. No differences in the prevalence of DMA and DMB alleles in RA patients and normal controls were observed in Chinese, Japanese and Caucasian populations. These genes, therefore, do not appear to play a role in susceptibility to RA.

Level of expression of HLA-DR genes and RA. Since RA is an autoimmune disease involving tissue-destructive immune responses, it is likely that an aberrant immune response(s) due to the level of cell surface expression of DR molecules plays a role in RA susceptibility. In fact, an aberrant expression of class II molecules on synovial and on T $\gamma\delta$ cells in synovial fluid, and a higher proportion of DR- and Tac-bearing T cells have been reported in patients with RA. We therefore examined the role of *cis*-acting (promoter region) elements and *trans*-acting (nuclear proteins) factors on the expression of DR molecules. Although we observed functional allelic polymorphisms in the conserved consensus motifs (X1, X2, Y, CCAAT and TATA) and in spacers between these boxes in the DRB gene promoters, no RA-specific variation was observed in the promoter regions of DRB1 genes^{18, 19}. On the other hand, we observed differences in DNA-protein interactions between Y box oligonucleotides and *trans*-acting nuclear proteins from B cell lines from RA patients as compared to normal controls²¹. The results showed that 50% of the RA patients lacked the NF-Y protein, which binds to the Y box with an inverted CCAAT sequence. In contrast, NF-Y was detected in all normal controls. The absence of the NF-Y protein and/or the presence of the RA-susceptibility QKRAA/QRRAA epitope gave the highest RR value reported so far for susceptibility to RA. More studies are, however, necessary to elucidate the role of *trans*-acting nuclear proteins in development of RA.

Class III Region and Susceptibility to Rheumatoid Arthritis

Since the MHC class III region contains genes that are involved in immune response, it has been suggested that these genes are good candidates for the development of RA. For example, HSP70 proteins have been suggested as playing a role in the self-surveillance and susceptibility to autoimmune diseases, including RA¹. In fact, it has been demonstrated that HSP70 proteins specifically bind to the QKRAA motif in DRB1*0401 (DR4 Dw4) and RRRAA in DRB1*1001 (DR10). However, polymorphisms in the HSP70 gene, including

that in the HSP70 promoter, showed no independent association with RA⁶. Similarly, although an association between the TNF-lymphotoxin subregion and susceptibility to RA has been observed in multiplex families, no independent association between TNFa microsatellite alleles and RA has been found^{9, 11, 13}.

We investigated the role of microsatellite D6S273, located in the HSP70-Bat2 region, in the development of RA¹⁷. We observed that two D6S273 microsatellite alleles (132 and 138) showed significant differences in their prevalence in RA patients as compared to normal controls; allele 132 was significantly higher in total and in DRB1 QKRAA/QRRAA epitope-positive patients, and allele 138 was significantly higher in QKRAA/QRRAA-negative patients. Analysis of the data suggested that the association of D6S273 132 allele with RA was secondary to that of DRB1 genes. On the other hand, D6S273 138 allele had primary association with RA susceptibility in QKRAA/QRRAA epitope-negative patients. We extended these studies to microsatellite Bat2 alleles, also located in the HSP70-Bat2 region¹⁶. Bat2 138 allele, which is in linkage disequilibrium with D6S273 138 allele, also showed primary association with disease susceptibility in DRB1 QKRAA/QRRAA epitope-negative patients. These studies showed that the class III region contributes to susceptibility to RA, and D6S273 138 and Bat2 138 alleles provide additional markers for the development of RA.

Recently, we extended these data to additional gene (microsatellite TNFa and HSP70 promoter region) polymorphisms to obtain a fine mapping of the RA susceptibility gene(s) in the class III region and to identify the class III region haplotype that causes susceptibility to RA. Four class III region alleles, D6S273 138, HSP70c, Bat2 138 and TNFa2, showed association among themselves and with DR3 with significant linkage disequilibrium. We then identified the class I, class II and class III region extended haplotype A1-B8-DR3-D6S273 138-HSP70c-Bat2 138-TNFa2. Since the association between these class I and class II (A1, B8, DR3) and class III (D6S273 138, HSP70c, Bat2 138 and TNFa2) alleles was not absolute, we examined the relative roles in susceptibility to RA of 3 haplotypes: 1) A1-B8-DR3, 2) A1-B8-DR3-D6S273 138-HSP70c-Bat2 138-TNFa2, and 3) D6S273 138-HSP70c-Bat2 138-TNFa2. The results showed that the class III region haplotype D6S273 138-HSP70c-Bat2 138-TNFa2 has primary association with RA in DRB1 QKRAA/QRRAA epitope-negative patients. Furthermore, the RR value and the level of significance of association of this haplotype with RA are higher than those observed

with the alleles D6S273 138 and with Bat2 138. Since this haplotype has primary association with RA susceptibility and since the QKRAA/QRRAA DRB1 epitope does not provide any risk in this group of patients, our results suggest that this haplotype provides additional risk for susceptibility to RA. In addition, it is evident that the DRB1 QKRAA/QRRAA epitope and the class III region haplotype D6S273 138-HSP70c-Bat2 138-TNF α 2 more completely define the risk for development of RA. These data, therefore, demonstrate that two regions in MHC, class II (DRB1) and class III (D6S273, HSP70, Bat2, TNF α), contribute to susceptibility to RA.

Acknowledgment. The research was supported by a grant from the Arthritis Society, Canada.

References

1. AUGER I. and ROUDIER J. (1997): HLA-DR and the development of rheumatoid arthritis. *Autoimmunity*, **26**, 123–128.
2. BROWN J. H., JARDETZKY T. S., GORGA J. C., STERN L. J., URBAN R. G., STROMINGER J. L. and WILEY D. C. (1993): Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, **364**, 33–39.
3. BUCHANAN W. W. and SINGAL D. P. (1990): Is there a need to reclassify adult rheumatoid arthritis? *Br. J. Rheumatol.*, **29**, 377–381.
4. BUCHANAN W. W. and SINGAL D. P. (1994): Seronegative rheumatoid arthritis: tell it as it is. *J. Rheumatol.*, **21**, 391–393.
5. CAMPBELL R. D. and TROWSDALE J. (1993): Map of the human MHC. *Immunol. Today*, **14**, 349–352.
6. CASCINO I., GALLEAZZI M., SALVETTI M. et al. (1994): HSP70-1 promoter region polymorphism tested in three autoimmune diseases. *Immunogenetics*, **34**, 291–293.
7. CHRISTIANSEN F. T., KELLY H. and DAWKINS R. L. (1984): Rheumatoid arthritis. In ALBERT E. D., BAUR M. P. and MAYER W. (eds.): *Histocompatibility testing 1984*. Springer-Verlag, New York, 378–383.
8. GRAN J. T. and HUSBY G. (1987): Seronegative rheumatoid arthritis and HLA-DR4: proposal for criteria. *J. Rheumatol.*, **14**, 1079–1082.
9. HAJEER A. H., WORTHINGTON J., SILMAN A. J. and OLLIER W. F. (1996): Association of tumor necrosis factor microsatellite polymorphisms with HLA-DRB1*04-bearing haplotypes in rheumatoid arthritis. *Arthritis Rheum.*, **39**, 1109–1114.
10. HANSEN J. A. and NELSON J. L. (1990): Autoimmune disease and HLA. *Crit. Rev. Immunol.*, **10**, 307–328.
11. HARDWICK L. J., WALSH S., BUTCHER S. et al. (1997): Genetic mapping of susceptibility loci in the genes involved in rheumatoid arthritis. *J. Rheumatol.*, **24**, 197–198.
12. KORMAN A. J., BOSS J. M., SPIES T., SORRENTINO R., OKADA K. and STROMINGER J. L. (1985): Genetic complexity and expression of human class II histocompatibility antigens. *Immunol. Rev.*, **85**, 45–86.
13. MULCAHY B., WALDRON-LYNCH F., MCDERMOTT M. F. et al. (1996): Genetic variability in the tumor necrosis factor – lymphotoxin region influences susceptibility to rheumatoid arthritis. *Am. J. Hum. Genet.*, **54**, 676–683.
14. NELSON J. L., HANSEN J. A., SINGAL D. P. et al. (1992): Rheumatoid arthritis. In TSUJI K., AIZAWA M. and SASAZUKI T. (eds.): *HLA 1991*. Oxford University Press, New York, 772–774.
15. SINGAL D. P. and BUCHANAN W. W. (1993): Human leukocyte antigens (HLA) and rheumatic diseases: HLA class II antigen-associated diseases. *Inflammopharmacology*, **2**, 47–62.
16. SINGAL D. P., LI J. and LEI K. (1999): Genetics of rheumatoid arthritis (RA): Two separate regions in the major histocompatibility complex contribute to susceptibility to RA. *Immunol. Lett.* (in press).
17. SINGAL D. P., LI J., YE M. and LEI K. (1998): D6S273 microsatellite polymorphism and susceptibility to rheumatoid arthritis. *Tissue Antigens*, **52**, 353–358.
18. SINGAL D. P. and QIU X. (1996): Polymorphism in both the X and Y box motifs controls level of expression of DRB1 genes. *Immunogenetics*, **43**, 50–56.
19. SINGAL D. P., QIU X., D'SOUZA M. and SOOD S. K. (1993): Polymorphism in the upstream regulatory regions of HLA-DRB genes. *Immunogenetics*, **37**, 143–147.
20. SINGAL D. P. and YE M. (1998): HLA-DM polymorphisms in patients with rheumatoid arthritis. *J. Rheumatol.*, **25**, 1295–1298.
21. SINGAL D. P., YE M., FLEISIG H., BUCHANAN W. W. and QIU X. (1996): Y box-binding *trans*-regulatory nuclear proteins and susceptibility to rheumatoid arthritis. *Clin. Exp. Rheumatol.*, **14**, 669–672.
22. STASTNY P. (1978): Association of the B-cell alloantigen DRw4 with rheumatoid arthritis. *N. Engl. J. Med.*, **298**, 869–871.
23. TAKEUCHI F., NABETA H., KUWATA S., TANIMOTO K. and ITO K. (1997): Association of DMA and DMB with RA in Japanese. *Clin. Exp. Rheumatol.*, **15**, 189–192.
24. WINCHESTER R. (1994): The molecular basis of susceptibility to rheumatoid arthritis. *Adv. Immunol.*, **56**, 389–466.
25. YEN J. H., CHEN C. J., TSAI W. C., TSAI J. J., CHANG J. G. and LIU H. W. (1997): HLA-DMA and DMB genotyping in patients with rheumatoid arthritis. *J. Rheumatol.*, **24**, 442–444.

Received in February 1999

Accepted in May 1999