Correlation between Disease Phenotype and Genetic Heterogeneity in Rheumatoid Arthritis

Cornelia M. Weyand, Timothy G. McCarthy, and Jörg J. Goronzy
Department of Medicine, Division of Rheumatology, Mayo Clinic and Foundation, Rochester, Minnesota 55905

Abstract

RA is a heterogeneous group of disorders characterized by variations in clinical manifestations, disease course, and probably response to therapeutic interventions. We have addressed the question whether genetically and potentially etiologically more homogeneous subgroups of RA patients can be defined based upon the expression of the RA-linked sequence motif in the third hypervariable region of the HLA-DRB1 gene. Genetic comparison of patients classified upon clinical manifestation and disease course demonstrated that patients with mild disease were genetically distinct from those progressing to severe and destructive disease. Specifically, rheumatoid factor (RF) negative patients preferentially expressed RA-linked HLA-DRB1 alleles with an arginine substitution in position 71, whereas the alleles with a lysine substitution in position 71 accumulated in RF$^+$ patients. RF$^-$ patients were further subdivided based on clinical markers (time of onset of erosive disease and requirement for aggressive therapy). Clinical heterogeneity correlated with genetic heterogeneity. Patients with early erosive disease and patients requiring aggressive therapy frequently typed HLA-DRB1*04$^+$. Patients with late erosive/nonerosive disease or a benign disease course manageable with nonaggressive treatment preferentially expressed HLA-DRB1*01 or lacked an RA-linked haplotype. These data indicate that the heterogeneity of RA reflects genetic differences. Sequence variations within the disease-linked sequence motif, as well as polymorphisms surrounding the candidate genetic element, affect pattern, course, and treatment response of RA. Amino acid position 71 in the HLA-DRB1 gene has a unique role, the understanding of which may provide important clues to disease etiology. (*J. Clin. Invest.* 1995. 95:2120–2126.) Key words: HLA-DRB1 alleles • disease severity • rheumatoid factor • erosion • HLA and disease association

Introduction

RA is a chronic inflammatory disease with a strong genetic influence (1). HLA-DRB1 alleles sharing a sequence motif in

Address correspondence to Cornelia M. Weyand, M.D., Mayo Clinic, 401 Guggenheim Building, Rochester, MN 55905. Phone: 507-284-1650; FAX: 507-284-1086.

1. *Abbreviations used in this paper:* HVR, hypervariable region, RF, rheumatoid factor.

the third hypervariable region (HVR3)[1] of the B1 gene are highly enriched in patients with RA (2). The HLA-DR association of the disease is the most common denominator among patients and includes patients of distinct ethnic origin (3, 4). Association with different HLA-DRB1 alleles in patients with different ethnic backgrounds led to the identification of the shared sequence motif, which is now believed to be the critical genetic element in the disease. The shared sequence can be integrated into at least 10 different HLA-DRB1 alleles (B1*0101, *0102, *0401, *0404, *0405, *0408, *0409, *0410, *1402, and *1001) and spans the amino acid positions 70 to 74 in the HVR3 (5, 6). The concept that a shared sequence stretch found in multiple disease-linked alleles confers the risk to develop RA has been a productive working hypothesis in studying the role of HLA-DR molecules in RA. The concept has however underemphasized the fact that the disease-linked sequence motif exhibits subtle allelic sequence variations. In particular, it has been assumed that the substitution in position 71 (lys→arg) is conservative and does not influence the role of the sequence motif in the pathogenesis of the disease.

We have recently shown that the HLA-DR association of RA is unique in the sense that allelic combinations are linked to clinical manifestations (7). Patients who inherit a disease-linked HLA-DRB1 allele on both haplotypes tend to have more aggressive disease with extra-articular manifestations and increased frequency of joint surgery (8). We have also identified a particular role of HLA-DRB1*0401 homozygosity which accumulates among patients with rheumatoid vasculitis. The unique role of HLA-DRB1*0401 homozygosity has raised the question whether the functional contribution of the different alleles of the HLA-DR4 family might be distinct. We have, therefore, addressed the question whether sequence variations which distinguish the different disease-associated HLA-DRB1*04 alleles influence the clinical presentation of the disease. Here, we demonstrate that RA-linked alleles with a lysine in position 71 accumulate in seropositive RA patients, whereas an arginine substitution in position 71 is dominantly found in seronegative RA patients. While the exchange of lysine versus arginine appears to be important for the RF status of the disease, determinants in the first and second HVR of the HLA-DRB1 gene modulate the disease course of seronegative disease. In particular, disease-linked alleles of the HLA-DRB1*04 family are associated with more aggressive treatment and predispose patients to early erosive disease.

Methods

Study population. 142 patients with seropositive RA, 60 patients with seronegative RA, and 81 normal individuals were entered into the study. All patients fulfilled the 1987 American College of Rheumatology (formerly, American Rheumatism Association) criteria (9). Patients with seropositive disease had minimal disease duration of at least 3 yr. The study sample of seropositive patients partially overlapped with a disease cohort which has recently been published (8). Patients with seronegative

RA had a minimal disease duration of 5 yr, a negative family history of psoriatic arthritis, inflammatory bowel disease, spondylarthropathy, or connective tissue disease. All patients were negative for rheumatoid factor (RF⁻) tested on at least three occasions, the first being before the initiation of any DMARD therapy. To exclude patients with seronegative RA in the elderly, which may represent a distinct subset, only patients with seronegative RA who had disease onset before the age of 60 were included. All patients and control individuals were Caucasians.

HLA-DRB1 genotyping. The HLA-DRB1 alleles were determined by allele specific amplification by PCR with primer sets specific for the polymorphism of the HVR1 and subsequent oligonucleotide hybridization with probes specific for sequence polymorphism in the HVR2 and HVR3, as previously described. Genomic DNA was isolated from PBMC. Amplification was carried out for 30 cycles consisting of 1 min denaturation at 94°C, 2 min annealing at 55°C, and 2 min extension at 72°C. The following primers were used: HLA-DRB1*01, TTGTGG-CAGCTTAAGTTTGAA; DRB1*15/16, TTCCTGTGGCAGCCTAA-GAGG; DRB1*03,11,13,14, TTCTTGGAGTACTCTACGTCTGAG; DRB1*04, GTTTCTTGGAGCAGGTTAAAC; DRB1*07, TTCCT-GTGGCAGGGTAAGTATA; DRB1*08,12, CTTGGAGTACTCTAC-GGG; DRB1*09, TTCTGGAAGCAGGATAAGTTTG; DRB1*10, CGGTTGCTGGAAAGACGCG; and DRB1 nonpolymorphic, GCC-GCTGCACTGTGAAGCTCTC. PCR products were subsequently separated by gel electrophoresis and stained with ethidium bromide. Each PCR included DNA from lymphoblastoid cell lines from the Xth International Histocompatibility Workshop as positive controls. PCR products were analyzed for efficiency and specificity of the reaction by electrophoresis and adjusted to comparable concentrations. The amplified products were blotted and hybridized to allele specific digoxygenin-UTP labeled oligonucleotides representing the polymorphism of the HVR2 and HVR3. The following oligonucleotides were used to detect polymorphism of the HLA-DRB1*04 alleles and to define the HLA-DRB1*03, B1*08, B1*11, B1*12, and selected B1*13 and B1*14 alleles: DRB1*0401, B1*0409, GGAGCAGAAGCGGGCCGCG; DRB1*0103, B1*0402, B1*1102/3, B1*1301/2/4, CGCGGCCCGCTCGTCTTCC; DRB1*0403, B1*0407, B1*0411, CACCTCGGCCCGCCTCTGC; DRB1*0404, B1*0405, B1*0408, B1*0410, B1*1402, GTAGGTGTC-CACCGCGGCCCGCCT; DRB1*03, AAGCGGGGCCGGGTGGAC-AAC; DRB1*08, CAGGGCCCGCCT; DRB1*11, CAAGAGGAG-TACGTGCGCTTCGACAGC; DRB1*12, CGCAGGAGCTCC, and CTACGGGGTTGGTGAGAGC (codons 83–88, with glycine at position 86). Blots were washed and developed with digoxygenin specific alkaline phosphatase-labeled antibodies as described by the manufacturer (Genius; Boehringer Mannheim, Indianapolis, IN).

Clinical evaluation. Clinical records of all 60 patients with seronegative RA were reviewed with particular attention paid to the following variables: age at disease onset, presence or absence of extra-articular manifestations, and onset of erosions on sequential hand and/or foot x rays. In 45 out of the 60 patients, it could clearly be determined whether or not these patients had developed erosions within the first five years of disease. All patients were evaluated for the number and type of second line agents used during the disease course. A minimal treatment duration of three months was required for each second line agent. We did not attempt to analyze patient responses to disease remitting therapy.

Statistical analysis. Frequency of HLA-DRB1 alleles was compared by using a chi-squared test or a Fisher exact probability test if appropriate.

Results

HLA-DRB1 alleles in seropositive and seronegative RA. Results of HLA-DRB1 genotyping for 81 normal control individuals, 142 patients with seropositive RA, and 60 patients with seronegative RA are shown in Table I. In patients with seropositive RA, the polymorphism of HLA-DRB1 alleles was severely restricted. The dominant alleles were variants of the HLA-DRB1*04 family which were identified in 91% of the patients.

Most of the known HLA-DRB1 alleles showed the expected compensatory reduction in frequencies, with the exception of DRB1*01, B1*03, and B1*14. HLA-DRB1*01 and B1*14 have been described as disease associated alleles and carry the disease implicated sequence polymorphism in the HVR3. No contribution of HLA-DRB1*03 in RA has been described so far.

Compared with normal controls, the distribution of HLA-DRB1 alleles in RA patients with the seronegative variant of the disease showed a preferential usage of HLA-DRB1*01 and DRB1*04, as also seen in the patient cohort with seropositive disease, however, the association was much less striking. The most frequent allele in seronegative patients was HLA-DRB1*01, accounting for 42% of all haplotypes, compared with 27% in normal individuals, and 23% in seropositive RA patients. HLA-DRB1*04 alleles were also more frequent in seronegative RA when compared to normal individuals but were not enriched as significantly as in the cohort of seropositive patients. Taken together, 68% of all seronegative patients carried the RA implicated sequence, versus 41% of normal individuals ($P = .001$), and 97% of patients with seropositive RA.

Although seropositive and seronegative RA were associated with the same disease-linked epitope, they appeared to be immunogenetically different. The disease associated alleles B1*01 and B1*04 constitute families of allelic variance (6). For the allelic subtypes of the DR4 family, it has been established that selected members (B1*0401, *0404, *0405, and *0408) are associated with seropositive RA, whereas other members (B1*0402 and *0403) are not enriched. In the study cohorts analyzed here, the HLA-DRB1*04 subtypes B1*0401(9), *0404(10), and *0405(8) accounted for more than 90% of the seropositive as well as the seronegative patients with an HLA-DRB1*04 haplotype. However, HLA-DRB1*0401(9) was the dominant allelic variant in DR4⁺ seropositive patients (71%), whereas DR4⁺ seronegative patients expressed each of the B1*0401(9), *0404(10), and *0405(8) variants in a frequency of 33%. HLA-DRB1*01 alleles were mainly found in seronegative patients. In these patients, the allelic subtypes B1*0101 and B1*0102 were equally represented, the B1*0103 was not detected at all (data not shown).

Sequence variations in position 71 of the disease-linked sequence cassette correlate to RF production. Genotyping revealed that although seropositive and seronegative RA are associated with the same set of DRB1 alleles, they differ in the representation of these alleles with an accumulation of HLA-DRB1*0401(9) among RF producing patients and increased frequencies of DRB1*0404(5/8/10) and DRB1*0101(2) among patients who did not produce RF. Amino acid positions 70 to 74 defining the RA-linked sequence motif are identical for DRB1*0101(2) and DRB1*0404(5/8/10) and comprise the sequence QRRAA characterized by an arginine in position 71. This sequence motif can be distinguished from the amino acid cassette QKRAA in HLA-DRB1*0401(9) by the arg → lys substitution in position 71. We compared the expression of these two different variants of the RA-linked sequence motif in seropositive and seronegative patients. The results are summarized in Table II. In normal individuals, the sequence cassette QKRAA is detected in 16% and QRRAA in 31% of donors. Most of these donors express the cassette on only one haplotype. Seropositive RA patients show a dramatic enrichment of QKRAA. 65% of these patients carry the disease-linked sequence characterized by lysine in position 71. The frequency of the QRRAA cassette is slightly less among seropositive RA

Table I. Distribution of HLA-DRB1 Alleles in Patients with Seropositive and Seronegative RA

HLA-DR alleles	Frequency (%)			P		
	Normal donors (N) n = 81	Seropositive RA (RF+) n = 142	Seronegative RA (RF−) n = 60	N vs RF+	N vs RF−	RF+ vs RF−
B1*01	27.2	23.2	41.7	NS	0.07	0.08
B1*03	21.0	18.3	23.3	NS	NS	NS
B1*04	22.2	90.8	35.0	< 0.0001	0.09	< 0.0001
B1*07	19.8	7.8	20.0	0.008	NS	0.01
B1*08	13.6	5.6	11.7	0.04	NS	NS
B1*09	4.9	0.7	1.7	0.04	NS	NS
B1*10	2.5	0.0	0.0	NS	NS	—
B1*11	22.2	6.3	15.0	0.0005	NS	0.05
B1*12	1.2	0.0	0.0	NS	NS	—
B1*13	21.0	5.6	10.0	0.0005	0.08	NS
B1*14	3.7	4.2	1.7	NS	NS	NS
B1*15/16	30.9	11.3	21.7	0.0003	NS	0.05

patients compared to QKRAA but still high with 57%. About 45% of the patients express the disease-linked sequence motif on both haplotypes. In these double dose patients, the QKRAA as well as the QRRAA cassette are equally represented.

The distinguishing feature in patients with and without RF is the expression of QKRAA. This sequence motif is rarely identified in seronegative patients. Rather, these patients frequently type positive for QRRAA. 20% of seronegative patients have inherited a double dose and all of them express the QRRAA motif on both haplotypes. Patients expressing two copies of QKRAA are uniquely found in the seropositive cohort. Individuals carrying the combination of QKRAA on one haplotype and QRRAA on the other haplotype also accumulate among RA patients producing RF.

Role of sequence polymorphisms surrounding the disease-linked sequence cassette in modulating clinical presentation. All 60 patients with seronegative disease were evaluated by hand or foot x rays for the presence of bony erosions. A subset of 35% of the patients had nonerosive disease. To analyze at which period in the disease the 39 patients with erosions had developed bony destruction, serial x rays were compared. For 24 of the 39 patients, serial hand or foot x-rays were available

such that the onset of bony erosions could be unequivocally assigned to the first 5 yr of disease or later. 17 of the 24 patients developed erosions within the first 5 yr of disease and were defined as having early erosive disease. Seven patients had negative x-rays for the first 5 yr but developed typical erosive changes later in the disease course. Most of the patients with late onset erosions developed bony destruction between 8 and 12 years after disease onset (data not shown). Some patients had more than 20 yr of disease before erosions were detected.

To examine whether the onset of erosive disease was influenced by the absence or presence of disease-linked HLA-DRB1 alleles, the patient cohort with early erosive disease (first 5 yr) and with nonerosive or late erosive disease (after 5 yr) were compared. The results are given in Fig. 1. All patients in the early erosive disease cohort expressed disease-linked HLA-DRB1 alleles. Two thirds of the patients typed positive for HLA-DRB1*04 and 41% carried the HLA-DRB1*01. A different picture emerged for the patients with nonerosive/late erosive disease. All seronegative patients lacking an RA-linked HLA-DRB1 allele accumulated in that subset and accounted for 61% of the nonerosive/late erosive patients. While the frequency of HLA-DRB1*01 was similar in early erosive and nonerosive/

Table II. Correlation of Seropositive and Seronegative RA with the Sequence Polymorphism (Arg vs Lys) in Position 71 of the HLA-DRB1 Allele

Sequence motif	Individuals percent			P		
	Normal controls (N)	RF+ RA	RF− RA	N vs RF+	N vs RF−	RF+ vs RF−
QKRAA*	16.0	64.8	6.0	< 0.0001	NS	< 0.0001
QK/QK	0.0	9.9	0.0	0.004	NS	0.01
QK/QR	6.2	24.6	0.0	0.0005	0.05	< 0.0001
QK/X	9.9	30.3	6.0	0.0005	NS	0.002
QRRAA	30.9	57.0	58.3	0.0002	0.001	NS
QR/QR	2.5	10.6	20.0	0.03	0.0006	NS
QR/QK	6.2	24.6	0.0	0.0005	0.05	< 0.0001
QR/X	22.2	21.7	38.3	NS	0.03	0.016

Amino acid positions 70–74 in HLA-DRB1*0401 and *0410 (QKRAA) and HLA-DRB1*0101, *0102, *0404, *0405, *0408, *0410, and *1402 (QRRAA).
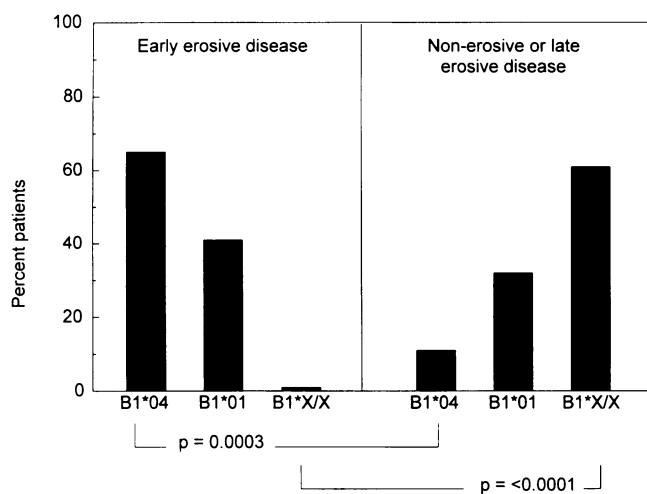
Figure 1. Correlation between HLA-DRB1 alleles and joint destruction in patients with seronegative RA. Patients with seronegative arthritis were subsetted into patients who did ($n = 17$) or did not ($n = 28$) develop bony erosions within the first five years of disease and compared for the distribution of HLA-DRB1 alleles. Patients with early erosive disease predominantly expressed an HLA-DRB1*04 allele whereas patients with late or nonerosive disease frequently lacked a disease-associated haplotype (HLA-DRB1*X/X).

Figure 2. Correlation between requirement for disease-remitting therapy and HLA-DRB1 alleles in seronegative RA. Clinical records from 60 patients with seronegative RA were retrospectively analyzed to determine whether they ever had been treated with a disease remitting agent (parental gold, D-penicillamine, azathioprine, or methotrexate) or only with a nonsteroidal anti-inflammatory drug and/or hydroxychloroquine (non-aggressive therapy). Patients treated with aggressive ($n = 32$) and non-aggressive ($n = 28$) therapy were compared for the distribution of HLA-DRB1 alleles.

late erosive disease, this was not the case for HLA-DRB1*04. HLA-DRB1*04[+] patients represented the majority within the early erosive cohort but accounted for only 11% of the subset with less aggressive disease. The equal representation of HLA-DRB1*01 and the enrichment of HLA-DRB1*04[+] patients in the early erosive subset indicates that not only the absence or presence of the disease associated sequence motif but also the polymorphisms distinguishing HLA-DRB1*01 and B1*04 alleles modulates disease presentation.

Differences in the therapeutic management of patients with seronegative RA indicates a role of the entire HLA-DRβ1 chain. To assess whether expression of HLA-DRB1 alleles was correlated to the aggressiveness of treatment, we have subsetted the patients with seronegative RA into two subsets, patients managed with nonaggressive treatment (NSAID or hydroxychloroquine) only and patients treated with potentially toxic medications (gold, D-penicillamine, azathioprine, or methotrexate), and have compared the distribution of HLA-DRB1 alleles in these two subsets.

Within the cohort of 60 patients, 28 patients had been managed with NSAID or hydroxychloroquine only. 32 patients had received one or multiple of the potentially more toxic medications. These patients had been treated with an average of 2.4 disease modifying agents. The results of the HLA-DRB1 genotyping are summarized in Fig. 2. Patients with seronegative RA lacking an RA-linked HLA-DRB1 allele accumulated in the nonaggressive management group and accounted for 43% of that cohort. Only four of the 28 patients (14%) in the nonaggressive treatment group carried an HLA-DRB1*04 allele. Vice versa, HLA-DRB1*04[+] patients represented the majority of patients requiring potentially toxic therapy. Only a small subset of the aggressively managed patients (19%) lacked HLA-DRB1*04 as well as DRB1*01. Patients expressing DRB1*01 were found in similar frequencies among both treatment cohorts (43 and 41%). The more aggressive management of
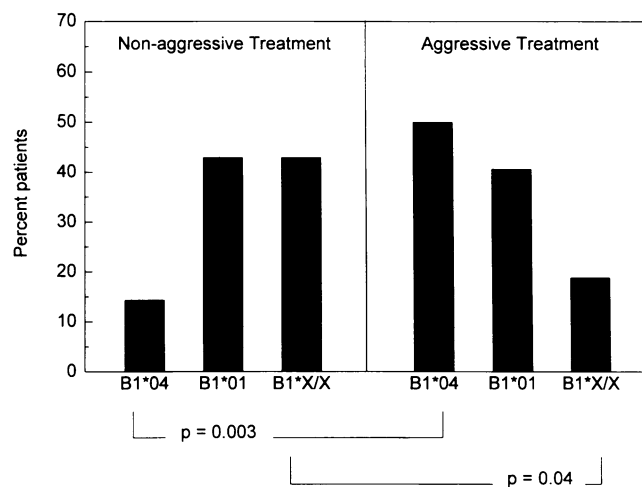
DRB1*04[+] patients and the less aggressive management of the patients expressing neither a DRB1*04 nor a DRB1*01 allele suggests that differences in clinical presentation, as indicated by the therapeutic decisions of the treating physician, correlated to the representation of HLA-DRB1 alleles. In contrast to HLA-DRB1*04 alleles, the HLA-DRB1*01 allele seems to not be indicative of the onset of erosive disease or the requirement of aggressive drug treatment.

Thus, HLA-DRB1 alleles modulate the disease process at three distinct levels, absence or presence of the RA-linked sequence motif, sequence variations at position 71 of the sequence cassette, and polymorphisms outside of the shared sequence motif but linked to HLA-DRB1*04 alleles.

Discussion

RA is a heterogeneous disease ranging from mild to life-threatening. We have previously reported that a gene dosage effect for HLA-DRB1 alleles is functional in RA patients. Patients who have inherited disease associated alleles on both haplotypes tend to develop more aggressive disease than patients with a single copy (8). This correlation of genotype and disease phenotype suggests an additive effect of both haplotypes in the mechanisms leading to disease manifestations characteristic of RA. We have now extended these studies to patients with milder forms of the disease. Here, we are providing data that genetic elements contribute to the clinical manifestations and that differences in disease severity correlate with polymorphisms of disease-linked HLA genes in Caucasian patients with seropositive and seronegative RA. By far the most powerful disease gene appears to be HLA-DRB1*0401, characterized by a lysine in position 71 of the β1 chain. The potency of the gene is reduced when lysine in position 71 is substituted by arginine. Shuffling of the disease-linked sequence stretch QRRAA into the frame-
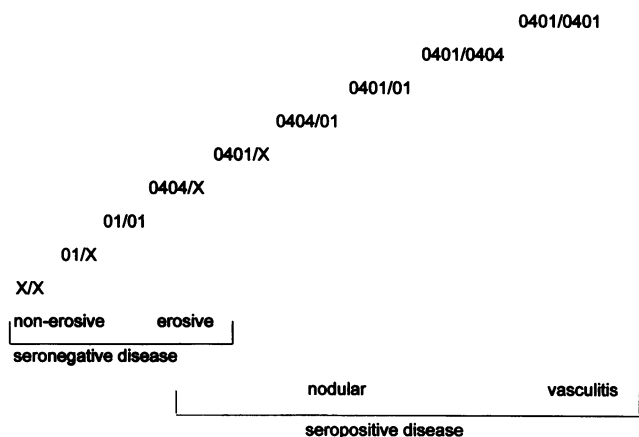
```
                                    0401/0401
                              0401/0404
                        0401/01
                  0404/01
            0401/X
      0404/X
  01/01
 01/X
X/X
|non-erosive        erosive  |
|_____|
seronegative disease

            |                nodular              vasculitis  |
            |_____|
                          seropositive disease
```

Figure 3. Proposed hierarchy of HLA-DRB1 alleles in determining
disease severity in seronegative and seropositive RA. This model implies
that the distinct RA associated HLA-DRB1 alleles differ in their impact
on the disease process. In addition, allelic combinations are functional
in modulating the disease phenotype (B1*0404 represents the group of
alleles including B1*0404, *0405, and *0408).

work of a different allele, HLA-DRB1*01, maintains disease
relevance. However, HLA-DRB1*01 is less potent than HLA-
DRB1*0404(5/8/10), indicating that the entire β1 chain con-
tributes to the disease expression. We are proposing the model
that a hierarchy exists for the disease impact of the distinct RA-
linked HLA-DRB1 alleles (Fig. 3).

Assignment of the seronegative and seropositive disease
variant to the same disease entity stems from similarities in
disease pattern and joint destruction. For all purposes, seronega-
tive and seropositive RA cannot be clinically distinguished with
the exception that RF producing patients usually have more
destructive disease and carry the risk of developing extra-articu-
lar manifestations (10, 11). Our immunogenetic studies showed
that both disease variants are associated with the same key
genetic element in the HVR3 of the HLA-DRB1 allele. The
impression that seronegative RA is generally less aggressive
was confirmed in our retrospective analysis of erosiveness of
the disease. Bony destruction as detected by x rays was present
in only 65% of patients and only 70% of the patients with
erosive disease developed erosions during the initial 5 yr of
disease. All of the seropositive patients had erosive disease.
While we do not know when the erosions occurred in seroposi-
tive patients, other studies have shown that up to 90% of bony
erosions occur within the first 2 yr of disease (12). The delayed
onset of osteolytic destruction is compatible with the model that
seronegative RA patients have the same disease but that it is
quantitatively less than in seropositive patients.

When comparing seropositive and seronegative disease, we
found a strong association of the polymorphism of position 71
with disease severity and RF production. A relationship between
HLA-DR polymorphism and RF production has been suggested
by some investigators (13). It is possible that RF secretion is
simply a consequence of more severe disease. Alternatively, it
could be hypothesized that position 71 plays a role in the induc-
tion of autoantibody production which subsequently causes
more rapid and more aggressive disease. Evidence has been
provided for the genetic association of HLA molecules with
the presence of autoantibody specificities in a number of other
chronic inflammatory diseases (14–16). This linkage is sugges-

tive of an antigen specific recognition event controlled through
the peptide selection function of HLA molecules in the induc-
tion of the autoantibody. It remains a possibility that HLA-
DRα-β dimers utilizing the DRB1*0401 allele selectively bind
an antigen and present it to T cells which provide help for the
induction of RF and potentially other disease mechanisms. The
finding that the variants DRB1*0404(10) and DRB1*0405(8)
are associated with seronegative as well as seropositive disease,
however, supports the interpretation that the microheterogeneity
within the disease-linked sequence determines severity of the
disease (Fig. 3), and the production of RF is a secondary event.

The association of RF and extra-articular disease has led to
the model that major organ involvement is an immune complex
mediated disease. For this reason, patients with seronegative
disease may not be at risk to develop rheumatoid vasculitis.
Immunogenetic differences between seronegative and seroposi-
tive patients, as shown in Tables I and II, provide a basis for
an alternative model. We have recently described that progres-
sion of RA to extra-articular organ diseases is linked with allelic
combinations at the HLA-DRB1 locus. Our cohort of seronega-
tive RA patients was missing allelic combinations which we
have previously linked with nodular disease and major organ
involvement. None of the seronegative patients were homo-
zygous for HLA-DRB1*0401 or heterozygous for HLA-
DRB1*0401(9) and *0404(5/8/10). Thus, it is possible that
patients with seronegative disease do not develop organ involve-
ment because patients expressing two relevant HLA-DRB1 al-
leles tend to progress to seropositive disease and subsequently
extra-articular manifestations. The linkage of extra-articular or-
gan disease with certain combinations at the HLA-DRB1 locus
could possibly indicate that extra-articular disease represents a
new dimension of the disease which can be dissected from
the joint disease (17). However, the haplotype combinations
associated with rheumatoid organ disease are also accumulated
among patients with severe joint disease. We can therefore not
exclude the possibility that extra-articular RA is the most severe
extension of RA.

Our data also provide evidence that polymorphisms outside
of the shared epitope influence the presentation of the disease.
As shown in Figs. 1 and 2, patients with seronegative disease
were heterogeneous in their erosiveness and in their requirement
for aggressive treatment with disease remitting agents. Subset-
ting the patients into mild and more severe disease by either
criteria allowed to determine the influence of HLA-DRB1 poly-
morphism on disease severity. Patients with mild disease pre-
dominantly did not express a disease-linked HLA-DRB1 allele,
whereas patients with more severe disease frequently used an
HLA-DRB1*04 variant. These findings are consistent with pop-
ulation based studies of RA which did not find an association
of the disease with HLA-DRB1*04, suggesting that milder
forms of RA may not have an HLA-DR association (18).

One of the striking features of the genetic predisposition of
RA is that the association with the shared sequence stretch is
found in patients of different ethnic origin. However, the lysine
substitution in position 71 is mostly found in Caucasians, in
particular, of Northern European decent. In the studies presented
here, we have confined our study population to Caucasian pa-
tients who were longitudinally followed at our institution. This
study population represents an immunogenetically relatively ho-
mogeneous cohort of patients. If our finding that lysine, more
than the arginine, in position 71 predisposes for more severe
disease can be generalized to other ethnic groups, then epidemi-

ological studies in other ethnic groups should show a bias toward seronegativity and milder joint disease as well as a lower frequency of extra-articular manifestations. It is also possible that our current view of the disease associated sequence stretch as a risk factor, independent of other background genes, is an oversimplification. RA is clearly a multigenic disease. It is therefore likely that an interaction between background genes and the disease associated epitope, and not the epitope alone, is important in the pathogenetic events leading to RA. Epidemiological studies assessing the disease pattern in different ethnic groups need to be done, however, this issue can finally be resolved only if the mechanisms are understood through which HLA-DRB1 alleles influence severity.

Several models have been proposed to explain the HLA-DR association of RA. It is possible that the correlation between disease manifestations and HLA-DRB1 alleles reflect the influence of HLA-DQ alleles in linkage disequilibrium with DR4. Stephens et al have suggested that DQw7 is a disease severity marker (19). The alternative hypothesis is that polymorphisms of the HLA-DRB1 allele, but outside of the shared epitope sequence, and the microheterogeneity in position 71 modulate disease expression. In the current paradigm, HLA molecules in RA function by binding and presenting arthritogenic antigens (20). Alternatively, we have hypothesized that disease-linked HLA-DR molecules could shape the repertoire of mature T cells thus having effects on the immunoresponsiveness (21). Available data on the peptide binding function of HLA-DR4 allelic variants are limited. Hill et al have compared the requirements for peptide binding to HLA-DRB1*0101 and DRB1*0401 and identified a single key hydrophobic side chain interacting in a conserved pocket in both DRB1*0101 and DRB1*0401 (22). No distinguishing features for peptides binding specifically to one of the alleles were described. In the crystallographic model of HLA-DR1, position 71 has been assigned to the antigen binding groove and represents a member of the residue cluster B13, 70, 71, 74, and 78 (23). This polymorphic cluster is believed to affect the binding of the peptide in the groove. It remains possible that the exchange in position 71 modulates the affinity of HLA antigen interaction in such a way that a lysine substitution increases binding affinity for a disease relevant peptide. The hierarchy model of allelic combinations (Fig. 3) suggested by us would be compatible with a gradient of affinities created by substitutions in position 71. Along the same line, it is possible that the polymorphism in position 71 affects HLA-TCR interaction. Position 71 can be localized in the middle of the α-helical wall bordering the antigen binding site. The α-helix, particularly the middle part, might be involved in contacting the TCR (24). We have studied if and how allelic variations in HLA-DR molecules affect the TCR repertoire (25). In these repertoire studies, the frequencies of Vβ-Jβ combinations in CD45RO⁻ CD4⁺ T cells from HLA-DR1/4⁺ and HLA-DR3/7⁺ donors were compared. While there was unequivocal modulation of the repertoire by HLA-DR molecules, we have not detected differences in HLA-DRB1*0401⁺ individuals compared to HLA-DRB1*0404⁺ individuals. Nevertheless, the polymorphism at position 71 could be critical in peptide binding, TCR contact, or both. Evidence suggests that the formation of the TCR-antigen-HLA complex results from a sum of contacts, and it is possible that minor variations in position 71 could lead to productive changes. Finally, it cannot be excluded that the disease associated sequence stretch represents an immunogen by itself (26, 27). The lysine/arginine

exchange might well alter the immunogenicity of the motif. Elucidation of the biological role of position 71 might provide insight into the contribution of HLA-DR molecules in the pathogenesis of RA.

Acknowledgments

References

1. Harris, E. D., Jr. 1990. Rheumatoid arthritis. Pathophysiology and implications for therapy. *N. Engl. J. Med.* 4:994–996.

2. Gregersen, P. K., J. Silver, and R. J. Winchester. 1987. The shared epitope hypothesis. An approach to understanding the molecular genetics of susceptibility to rheumatoid arthritis. *Arthritis Rheum.* 30:1205–1213.

3. Tiwari, J. L., and P. I. Terasaki. 1985. HLA and disease association. Springer-Verlag, New York.

4. Willkens, R. F., G. T. Nepom, C. R. Marks, J. W. Nettles, and B. S. Nepom. 1991. Association of HLA-Dw16 with rheumatoid arthritis in Yakima Indians. Further evidence for the "shared epitope" hypothesis. *Arthritis Rheum.* 34:43–47.

5. Winchester, R., E. Dwyer, and S. Rose. 1992. The genetic basis of rheumatoid arthritis: the shared epitope hypothesis. *Rheum. Dis. Clin. N. Amer.* 18:761–783.

6. Marsh, S. G., and J. G. Bodmer. 1991. HLA class II nucleotide sequences, 1991. *Immunogenetics.* 33:321–334.

7. Weyand, C. M., C. Xie, and J. J. Goronzy. 1992. Homozygosity for the HLA-DRB1 allele selects for extra-articular manifestations in rheumatoid arthritis. *J. Clin. Invest.* 89:2033–2039.

8. Weyand, C. M., K. C. Hicok, D. Conn, and J. J. Goronzy. 1992. The influence of HLA-DRB1 genes on disease severity in rheumatoid arthritis. *Ann. Intern. Med.* 117:801–806.

9. Arnett, F. C., S. M. Edworthy, D. A. Bloch, D. J. McShane, J. F. Fries, N. S. Cooper, L. A. Healey, S. R. Kaplan, M. H. Liang, H. S. Luthra, T. A. Medsger, Jr., D. M. Mitchell, D. H. Neustadt, R. S. Pinals, J. G. Schaller, J. T. Sharp, R. L. Wilder, and G. G. Hunder. 1988. The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum.* 31:315–324.

10. Harris, E. D., Jr. 1993. Clinical features of rheumatoid arthritis. *In* Textbook of Rheumatology. vol 1. W. N. Kelley, E. D. Harris, S. Ruddy, and C. B. Sledge, editors. W. B. Saunders Co., Philadelphia, PA. 874–911.

11. Van der Heijde, D, P. van Riel, M. van Rijswijk, and L. van de Putte. 1988. Influence of prognostic features on the final outcome in rheumatoid arthritis: a review of the literature. *Semin. Arthritis Rheum.* 17:284–292.

12. Sharp, J. T., F. Wolfe, D. M. Mitchell, and D. A. Bloch. 1991. The progression of erosion and joint space narrowing scores in rheumatoid arthritis during the first twenty-five years of disease. *Arthritis Rheum.* 34:660–668.

13. Nelson, J. L., C. E. Dugowson, T. D. Koepsell, L. F. Voigt, A. M. Branchaud, R. A. Barrington, M. H. Wener, and J. A. Hansen. 1994. Rheumatoid factor, HLA-DR4, and allelic variants of DRB1 in women with recent-onset rheumatoid arthritis. *Arthritis Rheum.* 37:673–680.

14. Harley, J., M. Reichlin, F. Arnett, E. Alexander, W. Bias, and T. Provost. 1986. Gene interaction at HLA-DQ enhances autoantibody production in primary Sjögren's syndrome. *Science (Wash. DC).* 232:1145–1147.

15. Harley, J. B., A. L. Sestak, L. G. Willis, S. M. Fu, J. A. Hansen, and M. Reichlin. 1989. A model for disease heterogeneity in systemic lupus erythematosus: relationships between histocompatibility antigens, autoantibodies, and lymphopenia or renal disease. *Arthritis Rheum.* 32:826–836.

16. Reveille, J. D., E. Durban, M. J. MacLeod-St. Clair, R. Goldstein, R. Moreda, R. D. Altman, and F. C. Arnett. 1992. Association of amino acid sequences in the HLA-DQB1 first domain with the antitopoisomerase I autoantibody response in scleroderma (progressive systemic sclerosis). *J. Clin. Invest.* 90:973–980.

17. Goronzy, J. J., and C. M. Weyand. 1993. Interplay of T lymphocytes and HLA-DR molecules in rheumatoid arthritis. *Curr. Opin. Rheumatol.* 5:169–177.

18. de Jongh, B. M., L. K. van Romunde, H. A. Valkenburg, G. G. de Lange, and J. J. van Rood. 1984. Epidemiological study of HLA and GM in rheumatoid

arthritis and related symptoms in an open Dutch population. *Ann. Rheum. Dis.* 43:613–619.

19. Stephens, H., L. Sakkas, R. Vaughan, I. Teitsson, K. Welsh, and G. Panayi G. 1989. HLA-DQw7 is a disease severity marker in patients with rheumatoid arthritis. *Immunogenetics.* 31:119–122.

20. Todd, J. A., H. Acha-Orbea, J. I. Bell, N. Chao, Z. Fronek, C. O. Jacob, M. McDermott, A. A. Sinha, L. Timmerman, L. Steinman, and H. O. McDevitt. 1988. A molecular basis for HLA class II-associated autoimmunity. *Science (Wash. DC).* 240:1003–1009.

21. Weyand, C. M., and J. J. Goronzy. 1990. Disease-associated human histocompatibility leukocyte antigen determinants in patients with seropositive rheumatoid arthritis. Functional role in antigen-specific and allogeneic T cell recognition. *J. Clin. Invest.* 85:1051–1057.

22. Hill, C. M., A. Liu, K. W. Marshall, J. Mayer, B. Jorgensen, B. Yuan, R. M. Cubbon, E. A. Nichols, L. S. Wicker, and J. B. Rothbard. 1994. Exploration of requirements for peptide binding to HLA-DRB1*0101 and DRB1*0401. *J. Immunol.* 152:2890–2898.

23. Brown, J. H., T. S. Jardetzky, J. C. Gorga, L. J. Stern, R. G. Urban, J. L. Strominger, and D. C. Wiley. 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature (Lond.).* 364:16–17.

24. Goronzy, J. J., C. Xie, W. Hu, S. K. Lundy, and C. M. Weyand. 1993. Restrictions in the repertoire of allospecific T cells. Contribution of the α-helical sequence polymorphism of HLA-DR molecules. *J. Immunol.* 151:825–836.

25. Walser-Kuntz, D. R., C. M. Weyand, and J. J. Goronzy. 1993. The repertoire of T cell receptor Vβ-Jβ combinations is influenced by HLA-DRB1 alleles associated with rheumatoid arthritis. *Arthritis Rheum.* 36:(297)S88 (Abstr.)

26. Roudier, J., G. Rhodes, J. Petersen, J. H. Vaughan, and D. A. Carson. 1988. The Epstein-Barr virus glycoprotein gp110, a molecular link between HLA-DR4, HLA-DR1, and rheumatoid arthritis. *Scand. J. Immunol.* 27:367–371.

27. Albani, S., J. E. Tuckwell, L. Esparza, D. A. Carson, and J. Roudier. 1992. The susceptibility sequence to rheumatoid arthritis is a cross-reactive B cell epitope shared by the *Escherichia coli* heat shock protein dnaJ and the histocompatibility leukocyte antigen DRB1*0401 molecule. *J. Clin. Invest.* 89:327–331.