HLA class II gene frequencies in Crohn's disease: a population based analysis in Germany

M Reinshagen, C Loeliger, P Kuehnl, U Weiss, B J Manfras, G Adler, B O Boehm

Abstract

Background—Ulcerative colitis is the only known inflammatory bowel disease associated with particular HLA alleles. Whereas the association with the HLA-DRB1*15 allele has been described in several independent studies for ulcerative colitis, no contribution of HLA alleles to susceptibility in Crohn's disease has yet been shown.

Aim—This study was designed to study the strength of association of HLA class II alleles as risk markers for Crohn's disease in a homogenous population in Germany.

Patients—A total of 4251 randomly selected control subjects, and 162 unrelated subjects with Crohn's disease were studied. Subjects were studied for their HLA-DRB1, HLA-DQA1, and HLA-DQB1 alleles.

Method—HLA DNA typing was performed after locus specific amplification with the polymerase chain reaction and reverse dot blot hybridisation.

Results—The HLA-DRB1*07 was the only HLA class II allele found to be significantly associated with Crohn's disease (relative risk (RR)=1·9, 95% CI: 1·66 to 2·14; p=0·0001). This association remained significant after correction for the number of DRB1 alleles compared. In patients with disease onset before 35 years the RR for the disease in HLA-DRB1*07 positive subjects was found to be higher (RR=3·1, 95% CI: 2·44 to 3·76). The HLA-DRB1*03 was significantly decreased in frequency in Crohn's disease (RR=0·5, 95% CI: 0·39 to 0·61; p=0·0028).

Conclusion—The HLA-DRB1*07 allele provides risk for the disease especially in patients with younger ages of onset. These data also provide indirect evidence for an immunogenetically based heterogeneity of the disease.

(*Gut* 1996; 38: 538–542)

Keywords: Crohn's disease, disease heterogeneity, HLA, polymerase chain reaction.

Department of
Medicine I, University
of Ulm, Germany
M Reinshagen
U Weiss
B J Manfras
B O Boehm

Department of
Transfusion Medicine,
Transplantation
Immunology,
University of
Hamburg, Germany
C Loeliger
P Kuehnl

Correspondence to:
Professor B O Boehm,
Department of Medicine I,
University of Ulm, Robert-
Koch-Strasse 8, 89070
Ulm/Donau, Germany.

Family and twin studies have provided evidence for a genetic susceptibility to inflammatory bowel disease (for a detailed review see ref 1). The most strongly associated genetic markers for inflammatory bowel disease thus far are provided by the polymorphic cell surface molecules of the major histocompatibility complex – that is, the HLA class I, and the HLA class II gene products.

Presence of the HLA-B27 specificity in inflammatory bowel disease patients with simultaneous presence of ankylosing spondylitis was reported.[1][2] Furthermore an association of ulcerative colitis with primary sclerosing cholangitis and the serologically defined specificities HLA-B8 and HLA-DR3 has been found.[1] In addition, other HLA class I and class II specificities have been described to be associated with the disease.[3][4] All these studies were carried out with classic serological HLA typing techniques and have provided inconclusive results.[1][4] Only a few studies have considered the role of HLA genes as markers for inflammatory bowel disease with novel DNA typing procedures. In a recent study, white subjects from the United States were HLA genotyped. In this study an association of ulcerative colitis with HLA-DR2 (DRB1*15 is a split of the DR2 specificity) has been reported.[5] Furthermore, the HLA-DQw5 specificity, which was provided by a HLA-DR1-DQw5 positive HLA-haplotype, contributed to Crohn's disease susceptibility in that study.[5]

While the association of HLA genes with ulcerative colitis seems now well established, the association of the serological defined HLA-DR locus with Crohn's disease is much more controversial. The combined analysis of three earlier studies in white populations showed no association of the serologically defined HLA-DR alleles with Crohn's disease.[6–8] No evidence for an association of HLA-DR antigens and Crohn's disease was provided in a segregation analysis of HLA haplotypes in multiplex Crohn's disease families, as no significant deviation from Mendelian segregation was seen.[9][10]

In our study we have compared HLA DNA typing results from 162 patients with Crohn's disease to the HLA typings taken from 4251 healthy local controls. All subjects studied were white of German origin from the same local area. We sought to discover if class II genes represent risk markers for the disease. As HLA class II antigens are well known for their important role as immune response genes, we also intended to define structural features of the class II genes with respect to disease susceptibility and protection for Crohn's disease.

Methods

Patients and controls

A total of 4251 unrelated control subjects (2268 men, 1983 women) were recruited from the same local area and were of the same ethnic origin (white subjects). The mean age was 35·2 (range 18 to 58 years). These control subjects
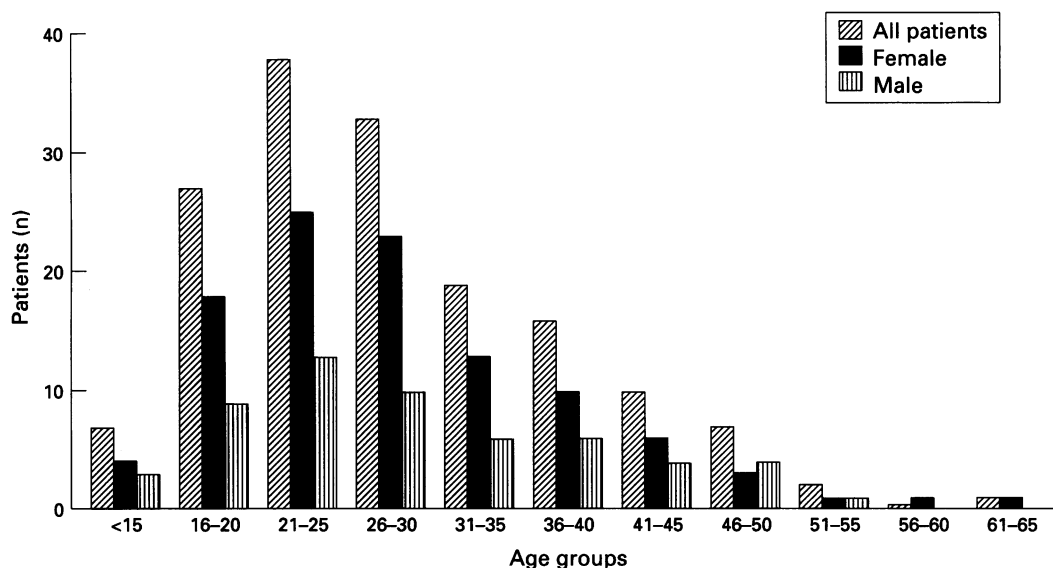
Figure 1: Ages at manifestation of the disease.

were healthy blood donors and healthy volunteers for bone marrow donation. The cohort of unrelated patients with Crohn's disease (n=162; 106 women, 56 men) included 1272 patient years. The average duration of the disease was 7·8 years (range 1–27 years). The mean age of the patients was 35·6 (range 9–65 years). The diagnosis of Crohn's disease and the disease localisation was confirmed in all patients by standard diagnostic procedures, which included endoscopy, histology, and radiological small bowel studies. Figure 1 summarises the ages as the start of the disease. The distribution of the disease localisation was: 88 of 162 (54%; 54 women, 34 men) patients had an inflammation of the ileum, whereas 74 (46%; 49 women, 25 men) had an inflammatory process also of the large bowel. From the 74 patients with a disease localisation in the colon 31 patients had Crohn's colitis, whereas in 43 patients active disease was found in the ileum and the colon.

Genomic HLA class II genotyping

The HLA-DRB1, -DQA1, and -DQB1 alleles were determined by use of a locus specific amplification procedure as described elsewhere.[11] In brief, 1 μg of genomic DNA, which had been extracted from peripheral blood leucocytes, was added to 10 μl of prepacked biotinylated primer mixes each of which corresponded to a given HLA class II locus, for example, HLA-DRB1, HLA-DQA1, and HLA-DQB1, respectively. After 35 cycles of polymerase chain reaction (PCR) amplification in an automatic thermocycler (Thermocycler 9600, Perkin Elmer Cetus) the amplificate was denatured by sodium hydroxide and hybridised to sequence specific oligonucleotides, which were fixed after oligo-dT tailing to a nylon membrane (Hybond N, Biodyne B, Pall Filtration Ltd, Portsmouth, UK). Allele specific hybridisation of the locus specific amplificate was detected by adding streptavidine coupled peroxidase and a colourimetric reaction of

tetra-methyl-bencidine. As a control of the reverse dot blot protocol the standard DNA typing procedure of the 11th International HLA Workshop was used in randomly collected samples.[12][13] One hundred and forty eight subjects were typed with both methods. One hundred per cent congruence of the typing results could be shown. In all typing procedures, DNA from homozygous typing cell lines (10th International HLA Workshop panel: number of HTCs included: n=27) were also included as internal controls. Our laboratory took part in the recent 11th International HLA Workshop and in the ongoing 12th International HLA Workshop; the workshop identification numbers are GER/BOE 332 and GER/KUE 338.

The following HLA-DRB1 alleles could be determined with the HLA genotyping procedures: HLA-DRB1*01, *15, *16, *03, *04, *11, *12, *13, *14, *07, *08, *09, and *10 (13 alleles). Very rare HLA-DRB1 alleles or splits have not been considered. The following HLA-DQA1 alleles were determined: HLA-DQA1*0101–*0103, *0201, *0301, *0401, and *0501 (n=7 alleles). The following HLA-DQB1 alleles were determined: HLA-DQB1*0501 to *0503, *0601 to *0605, 0201, *0301–*0303, and *04 (13 alleles).

Amino acid comparison of class II antigens

Published sequences of HLA class II antigens were analysed. The amino acid sequences were taken from a computer based source kindly provided by S Marsh, Tissue Antigen Laboratory, Imperial Cancer Research Fund, London ('HLA class II region nucleotide sequences', Version October 1994).

Statistical analysis

The χ^2 test was used as described.[14] Relative risk (RR) was calculated using Woolf's formula, the aetiological fraction, and the preventive fraction, which provide an estimate of the relative contribution of a marker to the

TABLE I   HLA-DRB1 gene frequencies in patients with Crohn's disease and local controls

	Crohn's disease a=320		Controls a=8502				
	a	gf(%)	a	gf(%)	RR	χ^2	p Value
DRB1-allele							
*01	27	8·4	920	10·8			
*15	43	13·4	1040	12·2			
*16	4	1·3	136	1·6			
*03	21	6·6	1020	11·9	0·5 (0·39 to 0·61)	8·9	0·0028
*04	42	15·0	1463	17·1			
*11	36	11·3	1048	12·3			
*12	5	1·6	126	1·5			
*13	50	15·6	1202	14·1			
*14	8	2·5	289	3·4			
*07	59	18·4	886	10·4	1·9 (1·66 to 2·14)	20·7	0·0001
*08	13	4·1	289	3·3			
*09	1	0·3	50	0·6			
*10	5	1·6	33	0·4			
Total	320		8502				

a=number of alleles; gf=gene frequency; RR=relative risk; 95% confidence intervals are given in brackets. The threshold for significance of a deviation was given by a corrected p value of ≤0·0038.

disease, were calculated as described.[15–17] The Bonferoni inequality method was used to correct for each antigen comparison.[18] The threshold for significance of a deviation for the HLA-DRB1 and HLA-DQB1 alleles was given by a corrected p value of ≤0·0038, and for the HLA-DQA1 alleles by a corrected p value of ≤0·0071.

Results

HLA-DRB1 alleles

Table 1 shows the gene frequencies in patients with Crohn's disease. A total of 320 alleles from Crohn's patients and 8502 alleles from control subjects have been compared. HLA-DRB1*03 was found to be significantly decreased in frequency in Crohn's disease patients when compared with the controls (gene frequency: 6·6% versus 11·9%; p=0·0028). A RR of 1·9 (95% CI: 1·66 to 2·14; p=0·0001) was found for the presence of HLA-DRB1*07 (gene frequency: 18·4% in Crohn's disease versus 10·4% in controls). This association remained significant after correction for the number of DRB1-alleles compared. The aetiological fraction of the DRB1*07 allele could be calculated as 9%. The reduction of both, the HLA-DRB1*01 and the DRB1*04 allele, did not reach a level of significance. There were six HLA-DRB1*07 homozygous patients. On the basis of a gene frequency of 0·184 the expected number of homozygous patients was 5·4.

TABLE II   HLA-DQA1 gene frequencies in patients with Crohn's disease and local controls

	Crohn's disease a=320		Controls a=4882				
	a	gf(%)	a	gf(%)	RR	χ^2	p Value
DQA1-allele							
*0101	39	12·3	721	14·8			
*0102	68	20·1	1025	20·9			
*0103	20	6·3	449	9·2			
*0201	60	18·9	540	11·1	1·9 (1·8 to 2·0)	14·8	0·0001
*0301	49	15·5	738	15·1			
*0401	13	4·1	145	2·9			
*0501	67	21·2	1264	25·9	0·8 (0·59 to 1·0)	3·8	0·0321
Total	320		4882				

The 95% confidence intervals are given in brackets. The threshold for significance of a deviation was given by a corrected p value of ≤0·0071. Abbreviations as Table I.

HLA-DQA1 alleles

Table II summarises the HLA-DQA1 typing results. A total of 320 alleles from Crohn's patients and 4882 alleles from the controls could be compared. No difference could be seen for the HLA-DQA1 alleles *0101, *0102, and *0103 when compared with the controls. A RR of 1·9 was calculated, for the presence of the HLA-DQA1*0201 allele (p=0·0001). The correct p values also showed a significant deviation of the corresponding genotype. The HLA-DQA1*0501 allele was found to be reduced in its frequency (gene frequency: 21·1% versus 25·9%; p=0·0321), but this difference did not reach significance.

HLA-DQB1 alleles

Table III summarises the HLA-DQB1 typing in Crohn's patients and controls. A total of 320 alleles in Crohn's patients and 1456 alleles in controls were compared. A slight increase of the DQB1*0501 allele was seen in the control subjects compared with Crohn's disease patients (11% versus 9·7%). Also a slight increase of the *0602 allele was found in control (11·6% versus 13·1%), whereas the HLA-DQB1*0302 allele was found to be decreased (9·6% versus 11·9%). These differences were not statistically significant.

Heterogeneity of Crohn's disease population: HLA and age of onset of the disease

The cohort of patients with Crohn's disease was large enough to be subdivided according to the age of onset of the disease (see also Fig 1). The genotype frequency of the disease associated DRB1*07 allele differed considerably in the cohort of patients with onset of the disease before 35 years versus an older age of onset. The frequencies were 26·5% versus 14·1% (χ^2=4·2). The RR for the disease in patients with the HLA-DRB1*07 allele was also different when calculated for the two subgroups of patients (a) <35 years of age RR=3·1 (95% CI: 2·44 to 3·76; χ^2=52·5; p=0·0001), (b) ≥35 years of age RR=1·4 (not significant). Furthermore, all six DRB1*07 homozygous patients had an age of onset before age 30.

HLA association with clinical subsets: localisation of the inflammatory process

The cohort of patients with Crohn's disease could be subdivided into two subgroups – that is, with (n=74) and without (n=88) inflammation of the large bowel. No differences in both groups could be found for the prevalences of the disease associated HLA-DRB1*07 allele or the so called preventive HLA-DRB1*03 allele. We also considered whether an inflammation in the colon in Crohn's disease patients could be attributed to the presence of the HLA-DRB1*15 allele. This allele also provides a strong risk for ulcerative colitis in our study population (data not shown). The genotype frequency of DRB1*15 allele was found not to be different between the subgroups with and without inflammation of the large bowel.

TABLE III   *HLA-DQB1 gene frequencies in patients with Crohn's disease and local controls*

| | Crohn's disease a=320 | | Controls a=1·456 | | | |
	a	gf(%)	a	gf(%)	RR	x^2
DQB1-allele						
*0501	31	9·7	151	11·0		
*0502	5	1·6	44	3·0		
*0503	13	4·0	39	2·7		
*0601	1	0·3	8	0·5		
*0602	37	11·6	191	13·1		
*0603	16	5·0	111	7·6		
*0604	20	6·2	66	4·5		
*0605	0	0	5	0·3		
*0201	58	18·1	282	19·4		
*0301	67	20·9	298	20·5		
*0302	38	11·9	140	9·6		
*0303	16	5·0	82	5·6		
*04	18	5·6	36	2·5		
Total	320		1456			

The threshold for significance of a deviation was given by a corrected p value of ≤0·0038. Abbreviations as Table I.

Sequence comparison between disease associated and protective HLA-DRB1 alleles

The comparison between HLA-DRB1*01 (used as consensus sequence), DRB1*15 (associated with ulcerative colitis), and DRB1*07 (associated with Crohn's disease) and the 'protective' DRB1*03 allele showed several 'critical' positions (see Fig 2). At amino acid position 9, the hydrophobic non-polar amino acid tryptophane was found in all inflammatory bowel disease associated β chain molecules, whereas the DRB1*0301 molecule showed the acidic amino acid glutamic acid (Fig 2). At position 11, again a hydrophobic non-polar amino acid was found (leucine), whereas DRB1*03 molecule showed an uncharged polar amino acid, glycine. Also at position 26 a hydrophobic non-polar amino acid could be found in inflammatory bowel disease associated DRB1 alleles (leucine and phenylalanine), whereas HLA-DRB1*03 showed the uncharged polar amino acid tyrosine. Finally, an important difference was seen at amino acid 32, where the aromatic amino acid tyrosine, which is present in ulcerative colitis and Crohn's disease associated HLA-alleles, is substituted with the basic amino acid histidine in the 'protective' HLA-DRB1*03 molecule.

Discussion

Our population based study was able to define for the first time an association of a HLA class II allele with Crohn's disease. Our study did not have problems associated with ethnic diversity. We used a DNA based HLA typing procedure, which is highly reproducible for HLA class II typing. In our analysis a single HLA-DRB1 allele was found to be associated with Crohn's disease. This result confirmed an earlier report of our group in multiplex

families.[19] A cosegregation of a HLA-DRB7 positive haplotype with the disease had been found in familial Crohn's disease. A major contribution of HLA-DQ molecules for disease susceptibility could not be confirmed in Crohn's disease. The increase of the HLA-DQA1*0201 allele was directly related to the higher prevalence of DRB1*07-positive HLA haplotypes and can therefore be attributed to linkage disequilibrium.

Most DRB1*07 positive haplotypes showed the HLA-DQB1*0201 allele, which was also found on the 'protective' HLA-DRB1*03 positive HLA haplotype and may further substantiate our view that HLA-DQ molecules do not play a major part, at least as risk markers, for Crohn's disease.

HLA molecules function as peptide binding cell surface glycoproteins. The complex of peptide and HLA molecule is then presented to the T cell receptor.[20] Therefore, histocompatibility proteins play a major part in the functioning of the immune system. The determination of the x ray structure of a human class II molecule, and the characterisation of peptides bound to HLA class II molecules,[21-23] provided a framework for the structural requirements of peptide binding to class II molecules. Provided that the HLA-DRB1*07 molecule is directly involved in peptide presentation to T cells in Crohn's disease, differences in the peptide binding requirements of disease associated versus protective HLA alleles can probably be detected. The question may arise as to whether the different amino acids of the various β chain molecules exert any influence on the source of peptides bound to the groove of the class II molecules.[22 23] An analysis of the relative frequency and source of peptides bound to six HLA-DR alleles, HLA-DR1, -DR2, -DR3, -DR4, -DR7, and -DR8 had shown important differences between the amount of bound peptides derived from exogenous serum proteins, endogenous membrane associated/ secretory proteins, and endogenous cytosolic proteins.[22] The amount of exogenous peptides bound to the HLA-DR3 molecules (around 30%) was several times larger compared with the HLA-DR1, HLA-DR2, and HLA-DR7 molecules (5 to 10%) in that study. HLA-DR1 and HLA-DR2 molecules showed a fairly equal distribution pattern of peptides derived from all three sources studied. The amount of peptides coming from the cytosol was highest for the HLA-DR7 molecules and was the unique feature of this class II molecules compared with all other HLA class II molecules analysed thus far. These data may point to differences in the peptide binding and presentation of the various disease associated HLA-DRB1 alleles compared with the putatively protective DRB1
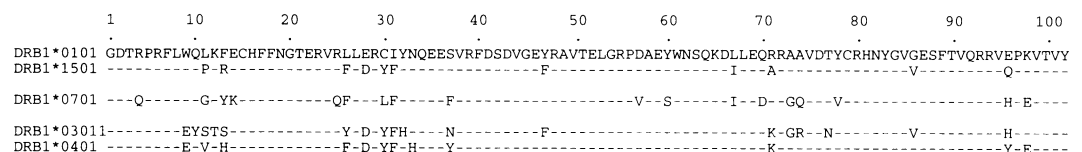
```
            1        10        20        30        40        50        60        70        80        90       100
            .        .         .         .         .         .         .         .         .         .         .
DRB1*0101   GDTRPRFLWQLKFECHFFNGTERVRLLERCIYNQEESVRFDSDVGEYRAVTELGRPDAEYWNSQKDLLEQRRAAVDTYCRHNYGVGESFTVQRRVEPKVTVY
DRB1*1501   ----------P-R------------F-D-YF--------------F--------------------I---A--------------V---------Q------

DRB1*0701   ---Q-----G-YK----------QF---LF-----F----------------V--S------I--D--GQ---V-----------------H-E----

DRB1*03011  --------EYSTS------------Y-D-YFH----N--------F----------------------K-GR--N---------V---------H------
DRB1*0401   --------E-V-H------------F-D-YF-H---Y----------------------------------K-----------------------Y-E----
```

Figure 2: The amino acid positions of the β chain molecule are given at the top of the Figure. Amino acids are given in one letter symbols.