Genes and Immunity (2004) 5, 215–220

www.nature.com/gene

FULL PAPER

Defining the contribution of the HLA region to cis DQ2-positive coeliac disease patients

MJ van Belzen[1,2], BPC Koeleman[1], JBA Crusius[3], JWR Meijer[4], AFJ Bardoel[1], PL Pearson[1], LA Sandkuijl*, RHJ Houwen[2] and C Wijmenga[1]

[1]*Complex Genetics Group, Department of Biomedical Genetics, University Medical Centre Utrecht, Utrecht, The Netherlands;* [2]*Department of Paediatric Gastroenterology, University Medical Centre Utrecht, Utrecht, The Netherlands;* [3]*Laboratory for Immunogenetics, VU University Medical Centre, Amsterdam, The Netherlands;* [4]*Department of Pathology, Rijnstate Hospital, Arnhem, The Netherlands*

*The major genetic susceptibility to coeliac disease is contributed by the human leukocyte antigen (HLA) region. The primary association is with the HLA-DQ2 molecule, encoded by the DQA1*05 and DQB1*02 alleles, which is expressed by over 90% of patients. The aim of our study was to perform an extensive scan of the entire HLA region to determine whether there is evidence for the presence of additional HLA susceptibility genes for coeliac disease in the Dutch population, acting independently of DQ2. In all, 16 microsatellite markers and the DQA1 and DQB1 genes were genotyped in simplex cis DQ2-positive coeliac disease families and cis DQ2-positive control families. Allele frequencies of markers on phase-known DQ2-positive haplotypes transmitted to patients were compared to a combined group of DQ2-positive nontransmitted and control haplotypes, thereby controlling for the DQ2 contribution. No significant differences at any of the marker loci were detected, suggesting that DQ2 is the major HLA risk factor for coeliac disease. Individuals homozygous for DQ2 or heterozygous for DQA1*05-DQB1*02/DQA1*0201-DQB1*02 were found to be at five-fold increased risk for development of coeliac disease ($P < 10^{-8}$). This risk seems to be conferred by the presence of a second DQB1*02 allele next to one DQA1*05-DQB1*02 haplotype, independently of the second DQA1 allele.*
Genes and Immunity (2004) **5,** 215–220. doi:10.1038/sj.gene.6364061
Published online 11 March 2004

Keywords: *coeliac disease; HLA region; association analysis; linkage disequilibrium*

Introduction

Coeliac disease is a common food intolerance in humans, with a prevalence estimated to be as high as 0.5–0.3% in the Netherlands.[1,2] The disease is characterized by lesions of the small intestine with partial to total villous atrophy, crypt hyperplasia and invasion of lymphocytes into the gut epithelium and lamina propria.[3,4] The main clinical symptoms include chronic diarrhoea and growth retardation, but abdominal pain, anaemia, osteopenia and chronic fatigue may also occur.[5] However, most patients show only some of these symptoms, while others are monosymptomatic or have no symptoms at all. Coeliac disease is caused by dietary intake of gluten peptides from wheat and related proteins from barley and rye. A gluten-free diet usually results in the recovery of the small intestinal lesions and disappearance of the clinical symptoms.

Correspondence: Professor C Wijmenga, Complex Genetics Group, Department of Biomedical Genetics, Stratenum, University Medical Centre Utrecht, PO Box 80030, Utrecht 3508 TA, The Netherlands. E-mail: t.n.wijmenga@med.uu.nl.
*Dedicated to the memory of Lodewijk Sandkuijl (1953–2002).
Received 09 October 2003; revised 18 December 2003; accepted 22 December 2003

Coeliac disease is strongly associated to the human leukocyte antigen (HLA) region. It has been well established that the primary association is with HLA-DQ2, with over 90% of patients expressing this molecule.[6] There is substantial evidence for involvement of DQ2 in coeliac disease pathogenesis. Gluten-derived peptides are modified by the enzyme tissue transglutaminase, which improves binding to DQ2 on the surface of antigen-presenting cells. These complexes are recognized by gluten-specific T cells isolated from small intestinal tissue of coeliac disease patients.[7,8] Most of the DQ2-negative patients express the HLA-DQ8 molecule, which is also capable of binding gluten-derived peptides with subsequent activation of gluten-specific T cells.[7]

The heterodimeric DQ2 protein is encoded by the HLA-DQA1*05 and HLA-DQB1*02 alleles, in either the *cis* or the *trans* configuration. In North European populations, the DQA1*05 and DQB1*02 alleles are frequently present on the extended HLA-B8-DR3-DQ2 haplotype.[6,9] This haplotype has also been shown to be associated with other autoimmune disorders, including type I diabetes mellitus, systemic lupus erythematosus, Graves' disease, Hashimoto's disease and myasthenia gravis, suggesting that the genes on this haplotype are involved in autoimmunity in general (for a review, see

Candore *et al*[10]). In coeliac disease, it was shown that different DQ2 genotypes account for different disease risks. In particular, the homozygous DQA1*05-DQB1*02/DQA1*05-DQB1*02 (DR3/3) and the heterozygous DQA1*05-DQB1*02/DQA1*0201-DQB1*02 (DR3/7) genotypes were shown to be associated with increased risk.[11–14]

The extended HLA-DR3-DQ2 haplotype includes many other genes that play a role in the immune response and it cannot be excluded that another HLA gene also confers increased risk to coeliac disease. The HLA region is known to display extensive linkage disequilibrium (LD). Therefore, one may expect that specific alleles at various loci in this region will show an increased frequency in coeliac disease patients, not because these alleles enhance or complement the DQ2 risk, but simply because they are in LD with DQ2. However, it has been suggested that non-class II loci also predispose to coeliac disease, independently of DQ2.[15–21]

The aim of this study was to test whether there was evidence for the presence of additional HLA susceptibility loci for coeliac disease on DQ2-positive haplotypes from patients of Dutch origin. In all, 16 markers, covering the entire HLA region and flanking regions, were genotyped in simplex coeliac disease and in control families. In this way, phase-known DQ2-positive haplotypes from cases and controls could be generated and tested for association. In addition, the effect of different DQ2 genotypes to coeliac disease risk was evaluated.

Results

Comparison of microsatellite marker loci on DQ2-positive haplotypes

Allele frequencies of all 16 microsatellite markers were determined on the 150 transmitted (T) and 36 non-transmitted (NT) DQ2-positive haplotypes from the case families and on the 103 DQ2-positive haplotypes from the control families. The overall allele distribution was not significantly different between T haplotypes and the combined group of NT and control haplotypes at any of the marker loci (data not shown). The most frequent allele on the T haplotypes was determined and the frequency of this allele was compared to the combined group of NT and control haplotypes (Table 1). These alleles are markers of the B8-DR3-DQ2 haplotype, characterized by alleles TNFa*2, MICB*10, MICA*3 and MIB*11.[22,23] This haplotype is the major DQ2 haplotype in patients and controls, with a frequency >75%. No evidence for significant association of these B8-DR3-DQ2-specific alleles was found. Other alleles with a frequency of 10% or more were also tested for association with coeliac disease (Table 2). These less frequent alleles could be compared for eight markers, but no significant evidence for association with coeliac disease was found. Only one marginally significant result was obtained for allele 4 of marker D6S2707, but the significance was lost after correction for multiple testing.

Conditional extended TDT (CETDT)

CETDT analysis in the case families, conditioned on DQ2, did not detect any alleles at the marker loci that showed significant distortion of random transmission.

Effect of different DQ2 genotypes

The DQ2 genotypes carried by the 110 DQ2-positive cases and 93 DQ2-positive controls are shown in Table 3. The homozygous DQA1*05-DQB1*02/DQA1*05-DQB1*02 (DR3/3) and heterozygous DQA1*05-DQB1*02/DQA1*0201-DQB1*02 (DR3/7) genotypes both conferred a five-fold increased risk for coeliac disease. Homozygosity for DQB1*02 was strongly associated with coeliac disease ($P<10^{-8}$).

Discussion

We investigated whether there was support for the presence of additional risk factors for coeliac disease in the HLA region, independent of DQ2 (ie DQA1*05-DQB1*02). As the HLA region exhibits strong LD, several approaches have been suggested to correct for this. Initially, a homozygous parent transmission disequili-

Table 1 Comparison of the most frequent allele of each microsatellite marker on DQ2-positive haplotypes transmitted (T) and non-transmitted (NT) to patients and on control haplotypes

Locus	*Allele (bp)*	*T (N=150)*	*NT (N=36)*	*Controls (N=103)*	*NT+controls (N=139)*	*P-value T vs NT+controls*[a]
D6S291	2 (201)	57	13	44	57	0.60
D6S2414	2 (170)	93	24	53	77	0.19
TAP1	3 (192)	106	26	66	92	0.27
STR2	6 (108)	127	28	81	109	0.19
D6S273	7 (140)	110	22	82	104	0.62
TNFa	2 (104)	121	24	86	110	0.67
MICB	10 (292)	114	24	77	101	0.68
MICA	3 (187)	117	26	79	105	0.83
MIB	11 (354)	106	26	78	104	0.64
STR1	6 (127)	110	26	76	102	0.73
D6S2700	3 (227)	109	30	72	102	0.36
D6S2704	7 (165)	75	17	56	73	0.73
D6S2707	9 (312)	83	20	56	76	0.78
D6S105	5 (119)	65	13	44	57	0.70
D6S2223	3 (170)	110	25	75	100	0.75
D6S1281	4 (194)	39	7	27	34	0.67

[a]No significant differences were observed when comparing T *vs* NT haplotypes or T *vs* control haplotypes separately.

brium (TDT) was applied, in which only transmissions from parents homozygous for DQ2 and heterozygous at the test locus were included in the analysis.[24] This approach was later extended to a case–control design, in which only DR3 homozygous cases and controls were included.[25] Although both methods control elegantly for the existing LD, they have very little power. Only individuals homozygous for DQ2 or DR3 are informative, which leads to the exclusion of the majority of the data set. Recently, a TDT approach using affected family-based controls (AFBAC) was applied to two large data sets of coeliac disease families.[15,19] This approach allows for the construction of phase-known haplotypes, in which transmitted (T) DQ2-positive haplotypes are compared to non-transmitted (NT) DQ2-positive haplotypes. A major advantage is that all DQ2-positive haplotypes are included in the analysis. However, less than 20% of DQ2-positive haplotypes were NT, so it requires a rather large data set to provide sufficient power.[15,19]

In view of all this, we chose to use a combined case–control and AFBAC approach, as well as CETDT, to maximize the power of our data set, which is of moderate size. By also genotyping the parents of cases and children of controls, we were able to obtain phase-known haplotypes in both groups. Hence, all the DQ2-positive haplotypes could be used in the case–control analysis. The AFBAC DQ2-positive haplotypes were combined with the DQ2-positive control haplotypes into one control group. This is a valid approach, as AFBAC controls were shown to be comparable to population controls.[26] Also, no significant differences in the allele frequencies between the AFBAC and population controls were observed at any of the 16 marker loci in our data set (data not shown). This case–control design greatly increased the power of our study, as an AFBAC case–control approach would have resulted in only 36 DQ2-positive control haplotypes, compared to 139 when using the combined group of AFBAC and control haplotypes.

We performed an extensive scan of the HLA region using microsatellite markers, but no evidence for independent association between any of the loci and coeliac disease was found. Likewise, no significant differences were present when comparing T and NT haplotypes, or T and control haplotypes separately (data not shown). When analysing the data by the homozygous parent TDT, which included only transmissions by parents homozygous for DQ2, only 29 transmissions could be scored. This resulted in too little transmissions of each allele to obtain sufficient power for statistical analysis, but increased transmission of marker alleles in the *TNF/MIC* gene region was observed (MICB*10, 7T *vs* 1 NT; TNFa*2, 7T *vs* 1 NT; data not shown).

Three other extensive screens of the HLA region in search for additional risk loci have been performed in

Table 2 Comparison of other alleles with frequencies $\geqslant 10\%$ on DQ2-positive haplotypes in cases T and a combined group of NT and control haplotypes (NT+controls)

Locus	*Allele (bp)*	*T (N = 150)*	*NT+controls (N = 139)*	*P-value T vs NT+controls*
D6S291	1 (199)	38	34	0.86
	6 (209)	15	16	0.68
D6S2414	3 (174)	34	39	0.30
TAP1	2 (190)	28	34	0.15
MICB	1 (272)	15	13	0.89
D6S2704	8 (167)	23	17	0.42
D6S2707	4 (302)	18	6	0.023[a]
D6S105	6 (121)	15	8	0.18
	7 (123)	21	28	0.16
	8 (125)	18	10	0.17
D6S1281	5 (198)	37	27	0.23
	6 (202)	30	35	0.33

[a]This *P*-value is not corrected for multiple testing.

Table 3 DQ2 genotype frequencies in *cis* DQ2-positive coeliac disease patients and controls

DQ2 genotype[a]	*DR type*[b]	*Cases (N = 110)*	*Controls (N = 93)*	*P-value*[c]	*OR (95% CI)*[d]
DQA1*05-DQB1*02/ DQA1*05-DQB1*02	DR3/3	40 (36%)	10 (11%)	0.00002	5.54 (2.5–12.1)
DQA1*05-DQB1*02/ DQA1*0201-DQB1*02	DR3/7	24 (22%)	6 (6%)	0.002	5.31 (2.1–13.5)
DQA1*05-DQB1*02/ DQA1*0301-DQB1*0302	DR3/4	7 (6%)	13 (14%)		
DQA1*05-DQB1*02/ DQA1*05-DQB1*0301	DR3/5	3 (3%)	10 (11%)		
DQA1*05-DQB1*02/ DQA1*X-DQB1*X	DR3/X	36 (33%)	54 (58%)		Reference

[a]Haplotypes are phase known as parents or children were also genotyped. DQA1*X-DQB1*X refers to any haplotype except those listed in this table.
[b]DR genotype was not acquired but derived from alleles at the DQA1 and DQB1 locus. X refers to anything except DR3, 4, 5 or 7.
[c]*P*-values for association with coeliac disease were calculated by testing each DQ2 genotype against all other DQ2 genotypes. *P*-value for the presence of two DQB1*02 alleles $<10^{-8}$.
[d]Odds ratios were calculated relative to the DR3/X type as reference. OR for the presence of two DQB1*02 alleles = 5.68 (95% CI 2.9–11.2).

coeliac disease. Two of them did not find significant association, independently of DQ2, at any of the microsatellite loci either.[19,27] However, a trend towards association within the *TNF*/*MIC* gene region was observed.[27] The presence of an additional risk locus in the *MIC* gene region was also suggested by the third study.[15] Association of the *MICA* gene polymorphism in a DQ2-positive population has been reported twice, but the control DQ2 groups were small in both studies.[16,17] In addition, independent association of a single-nucleotide polymorphism (SNP) in the *TNF* gene region has been reported by several groups.[18–21] These results indicate the possibility of additional HLA risk loci in the *TNF*/*MIC* gene region. However, there is a possibility that a SNP conferring increased disease risk may not be detected by the association of nearby microsatellite markers because of their high mutation rate. Therefore, SNPs may be the preferred type of polymorphism for studying the presence of additional HLA susceptibility loci. These SNPs should be genotyped in large collections of phase-known DQ2-positive patients and controls to provide unambiguous evidence.

The presence of an independent, additional risk locus for coeliac disease, located telomeric to the HLA class I region, has been suggested in a case–control study using DR3 homozygous patients and controls.[25] Allele 3 of marker D6S2223 was significantly less frequent in cases. However, we were unable to confirm these findings in our data set. Furthermore, we found no evidence for the association of allele 3 in an unstratified analysis of both case–control (76 *vs* 74%, $P = \text{NS}$) and TDT (43T *vs* 39 NT, $P = \text{NS}$) data (data not shown). Three other studies, all using large data sets, were also unable to confirm the association of D6S2223 with coeliac disease.[15,19,27] These results indicate that it is unlikely that a gene near D6S2223 predisposes to coeliac disease. The previously reported association may have occurred by chance due to the rather small sample size of 46 patients.

The majority of the DQ2-positive haplotypes in cases and controls consisted of B8-DR3-DQ2 haplotypes (see also Table 1).[22,23] We were therefore unable to determine whether there was a locus specifically associated with coeliac disease on other backgrounds. For example, the B18-DR3-DQ2 extended haplotype, characterized by alleles TNFa*1, MICB*1, MICA*1 and MIB*1, was rare in our cohorts with nine T, one NT and five control haplotypes (data not shown).[22,23] This haplotype is carried by 84% of DQ2-positive Sardinian coeliac disease patients, and this population is therefore more suitable for studying the presence of additional HLA risk loci on B18-DR3-DQ2 haplotyes.[28]

In addition to our search for the presence of additional non-class II HLA loci, our data set also enabled us to establish the risk of different DQ2 genotypes to coeliac disease in the Dutch population. Homozygous DQA1*05-DQB1*02/DQA1*05-DQB1*02 and heterozygous DQA1*05-DQB1*02/DQA1*0201-DQB1*02 individuals were at five-fold increased risk. This risk seems attributed by the presence of a second DQB1*02 allele and appears to be independent of the second DQA1 allele, since the odds ratio for homozygosity of DQB1*02 is almost equal to those for both risk genotypes. Similar risks were observed in other populations, although these studies were conducted in an ordinary case–control setting, in which haplotypes had to be estimated.[12–14] Recently, a family-based study using phase-known haplotypes also demonstrated increased risk for these two genotypes.[11] A possible explanation for the increased risk may reside in the number of DQ molecules capable of gluten presentation that arise from each genotype. The homozygous DQA1*05-DQB1*02 genotype produces 100% DQ2 molecules. The heterozygous DQA1*05-DQB1*02/DQA1*0201-DQB1*02 genotype produces only 50% DQ2 molecules and the other 50% is comprised by the α1*0201-β1*02 heterodimer. However, the α1*0201-β1*02 molecule was shown to be able to present certain gluten epitopes to T cells as well, thereby implicating this molecule in the pathogenesis of coeliac disease.[29] The DQA1*05 DQB1*02/DQA1*X-DQB1*X genotype results in just 25% DQ2 molecules, which may account for the lower disease risk of this genotype.

In conclusion, this study is the first case–control study in coeliac disease using phase-known DQ2-positive case and control haplotypes. We were not able to find support for the presence of an additional HLA susceptibility gene, acting independently of DQ2. We were able to confirm the increased risk conferred by homozygosity for the DQB1*02 allele, which is most likely due to a combination of the percentage of DQ2 molecules expressed and the gluten-presenting capacity of the heterodimer encoded by DQA1*0201 and DQB1*02.

Subjects and methods

Subjects

The case families consisted of 120 unrelated DQ2-positive coeliac disease patients with both parents available. The patients had a mean age of 17 years and 65% were female. The diagnosis of all patients was confirmed by histological re-evaluation of the initial small intestinal biopsy specimens (JWRM). All patients presented with partial to total villous atrophy in the presence of intraepithelial lymphocytosis and crypt hyperplasia. Control DQ2 haplotypes were derived from 86 control families without a history of coeliac disease. These families contained patients suffering from either attention-deficit hyperactivity disorder, schizophrenia or vesicouretheral reflux. The family history for other autoimmune disorders is unknown. The control families were selected for the presence of at least one parent carrying DQ2 in *cis* and also consisted of one child and both parents. The DQ2-positive controls had a mean age of 48 years and 50% were female. Individuals from the case and control families were all Caucasians of Dutch origin. The study was approved by the Medical Ethics Committee of the University Medical Centre Utrecht and written informed consent was obtained from all the participants.

Genotyping the HLA loci

A total of 18 loci were genotyped in the case and control families (Figure 1). The DQA1 and DQB1 genes were typed as described before, and primer sequences are provided in Web Table A (see Supplementary Information).[30] In all, 14 microsatellite markers, spanning the entire HLA region, and two flanking markers (D6S291 and D6S1281) were amplified in multiplex PCR reactions (Figure 1). Primers sequences were obtained from the Genome Database (www.gdb.org), except for markers

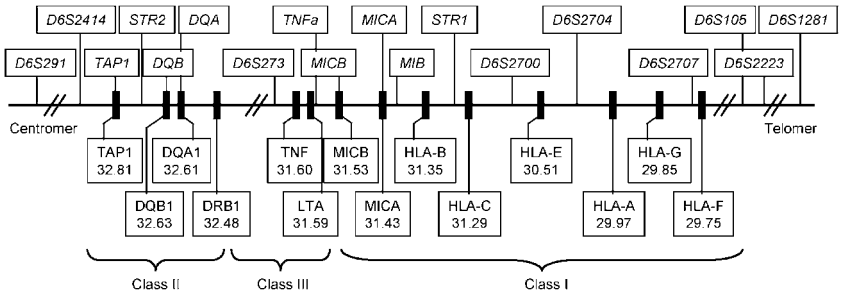

Figure 1 Overview of major genes in the HLA region and marker loci from this study. Vertical bars indicate genes, with names and locations depicted in the lower half of the figure. The location of the genes is in Mb and was based on the June 2003 release of the Ensembl human genome map. Marker loci are depicted in italics in the upper half of the figure.

STR1, STR2, MICB, MIB, TNFa and TAP1, which were designed in our lab (Web Table A, see Supplementary Information). Primer sequences for all loci were blasted against the Ensembl database (June 2003 release) to determine the correct order of the loci. The PCR reaction volume of 10 μl contained 25 ng of DNA, 0.2 mM dNTPs, 2.5 mM $MgCl_2$, 50 ng of each fluorescence-labelled primer and 0.4 U AmpiTaq Gold (PE Applied Biosystems, Foster City, CA, USA). The PCR products were pooled and separated on a 3700 DNA sequencer and analysed by Genescan 3.5 and Genotyper 2.0 software (all from PE Applied Biosystems, Foster City, CA, USA). All genotypes were checked independently by two researchers (MJvB and AFJB).

Statistical analysis

Phase-known haplotypes from the case and control families were constructed with the SHOWHAPLO program (F Dudbridge; available from ftp.hgmp.mrc.ac.uk/pub/linkage). Four patients turned out to be DQ2 positive, but in the *trans* configuration. The aim of this study was to test for the presence of additional risk loci on DQA1*05-DQB1*02 haplotypes, and the *trans* DQ2-positive patients were therefore excluded from the analysis. Six of the 116 *cis* DQ2-positive patients were not informative at either the DQA1 or DQB1 locus, resulting in 110 informative case families that were eventually included in the analysis. These 110 cases carried in total 150 DQ2-positive haplotypes (T haplotypes), while 36 DQ2-positive haplotypes were not transmitted to the cases (NT haplotypes). The control families contained 93 *cis* DQ2-positive parents who carried together 103 DQ2-positive haplotypes (control haplotypes).

Allele frequencies at the 16 microsatellite marker loci were determined on the T, NT and control haplotypes. The overall allele distribution on NT haplotypes was compared to the control haplotypes. There were no significant differences at any of the marker loci and these two groups of control haplotypes were therefore combined into one large control group to increase the power of the analysis. The marker loci were tested for association with coeliac disease by comparing the overall allele distribution on T haplotypes to the combined group of NT and control haplotypes. In addition, alleles with a frequency $\geqslant 10\%$ were tested separately for association. Statistical significance was determined by χ^2 analysis. TDT analysis, conditioned on the presence of DQ2, was performed in the case families by the CETDT program.[31]

The effect of DQ2 genotype to coeliac disease risk was determined in cases and controls. Association of the different DQ2 genotypes was tested for significance by χ^2 analysis with 1 df, by testing each DQ2 genotype against all other DQ2 genotypes. Odds ratios were calculated using Woolf's method with Haldane's correction relative to DQA1*05-DQB1*02/DQA1*X-DQB1*X (DR3/X) as reference genotype.

Acknowledgements

We are grateful to Winny van der Hoeven for her help in collecting the case families, to Jessica Bader for DQA1 and DQB1 genotyping of the case and control families and to Frank Dudbridge for his advice with regard to the analysis of the data. We thank Jackie Senior for improving the manuscript. The study was financially supported by a grant from the Dutch Digestive Diseases Foundation (WS 97-44). This paper is dedicated to the memory of Lodewijk Sandkuijl who unexpectedly died on 4 December 2002. He participated in this project from the beginning and his contribution to the design of the study and analysis of the data was of great importance to us.

References

1 Csizmadia CG, Mearin ML, von Blomberg BM, Brand R, Verloove-Vanhorick SP. An iceberg of childhood coeliac disease in the Netherlands. *Lancet* 1999; **353**: 813–814.
2 Rostami K, Mulder CJ, Werre JM *et al.* High prevalence of celiac disease in apparently healthy blood donors suggests a high prevalence of undiagnosed celiac disease in the Dutch population. *Scand J Gastroenterol* 1999; **34**: 276–279.
3 Rostami K, Kerckhaert J, Tiemessen R, von Blomberg BM, Meijer JW, Mulder CJ. Sensitivity of antiendomysium and antigliadin antibodies in untreated celiac disease: disappointing in clinical practice. *Am J Gastroenterol* 1999; **94**: 888–894.
4 Marsh MN. Gluten, major histocompatibility complex, and the small intestine. A molecular and immunobiologic approach to the spectrum of gluten sensitivity ('celiac sprue'). *Gastroenterology* 1992; **102**: 330–354.
5 Farrell RJ, Kelly CP. Celiac sprue. *N Engl J Med* 2002; **346**: 180–188.

6 Sollid LM, Thorsby E. HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. *Gastroenterology* 1993; **105**: 910–922.
7 van de Wal Y, Kooy Y, van Veelen P *et al.* Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *J Immunol* 1998; **161**: 1585–1588.
8 Molberg O, McAdam SN, Korner R *et al.* Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat Med* 1998; **4**: 713–717.
9 Alper CA, Fleischnick E, Awdeh Z, Katz AJ, Yunis EJ. Extended major histocompatibility complex haplotypes in patients with gluten-sensitive enteropathy. *J Clin Invest* 1987; **79**: 251–256.
10 Candore G, Lio D, Colonna Romano G, Caruso C. Pathogenesis of autoimmune diseases associated with 8.1 ancestral haplotype: effect of multiple gene interactions. *Autoimmun Rev* 2002; **1**: 29–35.
11 Louka AS, Nilsson S, Olsson M *et al.* HLA in coeliac disease families: a novel test of risk modification by the 'other' haplotype when at least one DQA1*05-DQB1*02 haplotype is carried. *Tissue Antigens* 2002; **60**: 147–154.
12 Ploski R, Ek J, Thorsby E, Sollid LM. On the HLA-DQ(alpha 1*0501, beta 1*0201)-associated susceptibility in celiac disease: a possible gene dosage effect of DQB1*0201. *Tissue Antigens* 1993; **41**: 173–177.
13 Mearin ML, Biemond I, Pena AS *et al.* HLA-DR phenotypes in Spanish coeliac children: their contribution to the understanding of the genetics of the disease. *Gut* 1983; **24**: 532–537.
14 Mearin ML, Bouquet J, Mourad N *et al.* HLA-DR antigens and phenotypes in Dutch coeliac children and their families. *Clin Genet* 1985; **27**: 45–50.
15 Bolognesi E, Karell K, Percopo S *et al.* Additional factor in some HLA DR3/DQ2 haplotypes confers a fourfold increased genetic risk of celiac disease. *Tissue Antigens* 2003; **61**: 308–316.
16 Lopez-Vazquez A, Rodrigo L, Fuentes D *et al.* MHC class I chain related gene A (MICA) modulates the development of coeliac disease in patients with the high risk heterodimer DQA1*0501/DQB1*0201. *Gut* 2002; **50**: 336–340.
17 Rueda B, Pascual M, Lopez-Nevot MA *et al.* Association of MICA-A5.1 allele with susceptibility to celiac disease in a family study. *Am J Gastroenterol* 2003; **98**: 359–362.
18 McManus R, Wilson AG, Mansfield J, Weir DG, Duff GW, Kelleher D. TNF2, a polymorphism of the tumour necrosis-alpha gene promoter, is a component of the celiac disease major histocompatibility complex haplotype. *Eur J Immunol* 1996; **26**: 2113–2118.
19 Louka AS, Lie BA, Talseth B *et al.* Coeliac disease patients carry conserved HLA-DR3-DQ2 haplotypes revealed by association of TNF alleles. *Immunogenetics* 2003; **55**: 339–343.
20 Garrote JA, Arranz E, Telleria JJ, Castro J, Calvo C, Blanco-Quiros A. TNF alpha and LT alpha gene polymorphisms as additional markers of celiac disease susceptibility in a DQ2-positive population. *Immunogenetics* 2002; **54**: 551–555.
21 de la Concha EG, Fernandez-Arquero M, Vigil P *et al.* Celiac disease and TNF promoter polymorphisms. *Hum Immunol* 2000; **61**: 513–517.
22 Bolognesi E, Dalfonso S, Rolando V, Fasano ME, Pratico L, Momigliano-Richiardi P. MICA and MICB microsatellite alleles in HLA extended haplotypes. *Eur J Immunogenet* 2001; **28**: 523–530.
23 Grimaldi MC, Clayton J, Pontarotti P, Cambon-Thomsen A, Crouau-Roy B. New highly polymorphic microsatellite marker in linkage disequilibrium with HLA-B. *Hum Immunol* 1996; **51**: 89–94.
24 Polvi A, Maki M, Partanen J. Celiac patients predominantly inherit HLA-DPB1*0101 positive haplotype from HLA-DQ2 homozygous parent. *Hum Immunol* 1997; **53**: 156–158.
25 Lie BA, Sollid LM, Ascher H *et al.* A gene telomeric of the HLA class I region is involved in predisposition to both type 1 diabetes and coeliac disease. *Tissue Antigens* 1999; **54**: 162–168.
26 Thomson G. Mapping disease genes: family-based association studies. *Am J Hum Genet* 1995; **57**: 487–498.
27 Louka AS, Moodie SJ, Karell K *et al.* A collaborative European search for non-DQA1*05-DQB1*02 celiac disease loci on HLA-DR3 haplotypes: analysis of transmission from homozygous parents. *Hum Immunol* 2003; **64**: 350–358.
28 Congia M, Frau F, Lampis R *et al.* A high frequency of the A30, B18, DR3, DRw52, DQw2 extended haplotype in Sardinian celiac disease patients: further evidence that disease susceptibility is conferred by DQ A1*0501, B1*0201. *Tissue Antigens* 1992; **39**: 78–83.
29 Vader W, Stepniak D, Kooy Y *et al.* The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. *Proc Natl Acad Sci USA* 2003; **100**: 12390–12395.
30 Carrington M, Miller T, White M *et al.* Typing of HLA-DQA1 and DQB1 using DNA single-strand conformation polymorphism. *Hum Immunol* 1992; **33**: 208–212.
31 Koeleman BP, Dudbridge F, Cordell HJ, Todd JA. Adaptation of the extended transmission/disequilibrium test to distinguish disease associations of multiple loci: the conditional extended transmission/disequilibrium test. *Ann Hum Genet* 2000; **64**: 207–213.

Supplementary Information accompanies the paper on Genes and Immunity's website (http://www.nature.com/gene).