

Interplay between genetics and the environment in the development of celiac disease: perspectives for a healthy life

George K. Papadopoulos,¹ Cisca Wijmenga,² and Frits Koning³

¹The Laboratory of Biochemistry and Biophysics, Faculty of Agricultural Technology, Technological Educational Institute of Epirus, Arta, Greece

²Department of Medical Genetics, University Medical Centre Utrecht, Utrecht, The Netherlands

³Department of Immunohaematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands

Address correspondence to: Frits Koning, IHB, E3-Q, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands. Phone: 31-71-5266673; Fax: 31-71-5216751; E-mail: fkoning@lumc.nl.

J. Clin. Invest. **108**:1261–1266 (2001). DOI:10.1172/JCI200114344.

Progress in celiac disease (CD) research has been spectacular in recent years. The nature of gluten-derived peptides that are recognized by gut-derived T cells from patients, and that are therefore likely to be involved in disease development, has been elucidated (1–5). It has further been established that a ubiquitous enzyme modifies such peptides, leading to improved binding of such peptides to HLA-DQ2/8 molecules (6, 7). This now establishes a molecular basis for the well-known association between CD and the expression of these HLA-DQ molecules (8, 9). CD is therefore the best-characterized HLA-associated disease to date. However, several issues remain unclear. Why does the disease develop in only a small percentage of the individuals who carry the predisposing HLA-DQ molecules? Why does this disease severely affect some patients while others have only mild or no clinical symptoms? Can we use our current knowledge to prevent or cure this disease?

CD pathogenesis

CD is caused by intolerance to gluten, a common protein present in wheat. Recent studies indicate that CD develops in 1 of 200–300 individuals in Western Europe and the USA (10–14). The disease often starts shortly after the first introduction of wheat into the diet. Symptoms include diarrhea, malabsorption, and failure to thrive, as a result of an immunological reaction to gluten in the small intestine, which eventually leads to a flat intestinal epithelium, prohibiting efficient uptake of nutrients. Removal of gluten from the diet is an effective way to stop the disease process. It is, at present, also the only treatment for the disease. Reintroduction of gluten in the patient's diet, even decades later, invariably leads to the reappearance of the symptoms. With few exceptions, CD is limited to genetically predisposed individuals, those who express the HLA-DQ2 (A1*0501/B1*0201) and/or -DQ8 (A1*0301/B1*0302) heterodimers (8, 9). This observation correlates well with the overwhelming evidence that CD is a T cell-mediated immune disease.

Gluten-specific T cells are commonly found in small intestinal biopsies of patients but not in the biopsies from controls (1–4). Without exception, these T cells respond to gluten-derived peptides bound to the disease-associated HLA-DQ2 or HLA-DQ8 molecules (1–6). In response, the T cells produce IFN- γ , which is presumed to induce tissue damage. Several gluten-derived T cell-stimulatory peptides have now been characterized, and this work has also revealed a novel protein modification that can lead to the induction and/or enhancement of T cell-mediated immune responses (3–7). This modification, the selective conversion of specific glutamine residues in gluten peptides into glutamic acid, is mediated by the enzyme tissue transglutaminase (tTG). The presence of acidic residues in the modified peptides helps explain the strong correlation between CD and HLA-DQ2/8, since these HLA-molecules bind preferentially to peptides with negative charges at characteristic positions (Figures 1 and 2, Table 1). Indeed, in the case of the three DQ2-specific gliadin epitopes identified thus far, T cell recognition is completely dependent on tTG transformation (3–5). In contrast, the two known DQ8-specific epitopes, one gliadin- and one glutenin-derived, induce T cell proliferation as native peptides (1, 2, 6). Transformation by tTG, however, results in strongly enhanced T cell recognition of the gliadin epitope (6), while it has much less effect in the case of the glutenin epitope (2).

Antibodies against gliadin and tTG are typically present in the sera of most patients (15, 16) and can be used for diagnostic purposes. The finding that tTG not only selectively modifies gluten peptides but is also a direct target of autoimmune antibodies is intriguing. In addition to its ability to modify gluten peptides, tTG is known to cross-link proteins via the formation of covalent bonds between lysine and glutamine amino acids. Sollid and colleagues have therefore proposed that tTG forms covalent complexes with gliadin, which could be internalized by tTG-specific B cells, leading to presentation of gliadin peptides in the HLA-DQ molecules on

Table 1

Peptide binding motifs of disease associated (DQA1*0501/B1*0201, DQA1*0301/B1*0302) and non-associated (DQA1*0201/B1*0202, DQA1*0301/B1*0302, DQA1*0301/B1*0303) alleles alongside the sequences of Cd-specific T cell-sensitizing gluten peptides

Allele	P1	P2	P3	P4	P5	P6	P7	P8	P9
DQA1*0501/B1*0201 Preferred ^A	F, W, Y, L, I, V, N, Q, E			D, E		P, E, D, A	D, E, W, A		W, Y, F, L, I, M, V, A, E
Antigen 1 α -II Gliadin ^B	P	Q	P	E*	L	P	Y	P	Q
Antigen 2 α -I Gliadin ^B	P	F	P	Q	P	E*	L	P	Y
Antigen 3 γ -I Gliadin ^B	P	Q	Q	S	F	P	E*	Q	Q
DQA1*0201/B1*0202 Preferred ^A	F, W, Y, L, I, V			D, E		P, A, D	D, E, W, A		W, Y, F, L, I, M, V D, E
DQA1*0301/B1*0302 Preferred ^A	D, E, I, L, F, Y								
Antigen 4 α -I Gliadin ^C	Q(E)	G	S	F	Q	P	S	Q	Q(E)
Antigen 5 I-Glutenin ^C	Q(E)	G	Y	Y	P	T	S	P	Q(E)
DQA1*0301/B1*0303 Preferred ^A	D, E, I, L, F, Y								A, V, S

^APreferred residues for pockets indicated are shown. ^BIn these epitopes the binding of the native peptide to the HLA-DQ2 allele is very weak, while that of the transformed peptide (Q→E residue shown by asterisk) is significant. ^CIn these epitopes the native peptide binds to the HLA-DQ8 allele and is recognized by specific restricted T cell clones. However, transformation at the specified positions (Q→E) results in more sensitive T cell recognition of the gliadin peptide, but not of the glutenin peptide.

the B cell. Thus, gliadin-specific T cells would drive the production of tTG-specific antibodies (17). It should be kept in mind, however, that the gliadin and tTG-specific antibodies themselves are not responsible for the disease symptoms, since these disappear rapidly upon withdrawal of gluten from the diet, whereas the antibodies persist much longer.

Genetics

The main evidence that genetic factors contribute to the development of CD comes from twin studies and the observation of familial clustering. The concordance rate for CD in monozygotic twins is estimated to be 75%, implying a large degree of heritability for CD (18). Moreover, approximately 15% of children with CD have a first-degree relative with *symptomatic* CD. Given the recent demonstration of the frequent occurrence of *asymptomatic* CD (10–14), this percentage may be substantially higher. The relative risk for CD (that is, the risk for a relative of an individual with CD compared with the population risk) is, therefore, at least 30.

Until recently, only one genetic factor had been identified as playing an important role in the development of CD, namely HLA-DQ, encoded by the MHC genes *DQA1* and *DQB* on chromosome 6 (8, 9). More than 90% of CD patients carry a combination of *HLA-DQA1*0501* and *HLA-DQB1*02* alleles, in either the *cis* or the *trans* configuration, and the gene products together form the HLA-DQ2 heterodimer. Most of the HLA-DQ2–negative CD patients carry the *HLA-DQA1*03* and *HLA-DQB1*0302* alleles, encoding the HLA-DQ8 heterodimer (8, 9). This HLA association is now readily explained by the observation that the disease-associated HLA-DQ molecules present gluten peptides to T cells.

A hallmark of HLA-associated diseases, however, is that the HLA association alone is not sufficient to explain the occurrence of the disease; CD is no exception to this rule. In fact, only a small percentage of HLA-DQ2/8 carriers develop CD, and the HLA-DQ2/8 genotype accounts for approximately a four- to fivefold increased risk (19–21). Therefore, additional genetic markers may predispose to disease development, and several studies have been carried out to elucidate the nature of such genes. Since *HLA-DQ2* is carried on the DR3-DQ2 haplotype, which extends far beyond the *HLA-DQ2* locus (22, 23), it has been suggested that further loci within the MHC region itself may also be involved in CD. Although various loci from the MHC region have been investigated for their putative role in CD, these results have been difficult to interpret due to extensive linkage disequilibrium within this region. Because of this linkage disequilibrium, a direct association, which might reflect the biological effects of known polymorphisms, cannot be readily distinguished from an indirect association in which the identified polymorphism (i.e., a “hitchhiker”) is close to an unidentified but functional polymorphism. A recent family-based study in type 1 diabetes, however, showed that statistical approaches can be developed that enable us to distinguish a true risk locus from a hitchhiker (24). For CD, a similar approach has been followed by van Belzen et al. (25), who have implicated the *MICB* gene as a potential candidate gene for CD (M.J. van Belzen and C. Wijmenga, unpublished results).

The human MHC region has now been sequenced completely (26) and all genes have been defined. We now know that the conserved CD haplotype contains classical class I, II, and III genes and other important immune response genes, many of which are potential

candidate genes for CD. It is expected that many more genetic studies will follow to investigate their roles in CD. However, to implicate additional risk loci and the effect of certain functional polymorphisms on CD, experimental studies will be needed in addition to genetic evidence.

Secondary determinants, outside the HLA region, could exert still stronger effects on disease susceptibility. Based on the pathophysiology of CD, genes involved in the T cell immune response could potentially be considered as functional candidate genes for the disease. Association studies have, for example, been conducted between the gene encoding the cytotoxic T lymphocyte-associated protein-4 (CTLA-4) and CD. These studies, however, have yielded conflicting results, and so far the *CTLA4* gene has not been implicated as a major risk factor for CD (27–29).

In the absence of strong functional candidate genes, many research groups have conducted genome-wide scans. Due to the lack of a clear genetic model for CD (i.e., the lack of a clear inheritance pattern, disease gene frequency, or penetrance), the majority of these studies had to be conducted in affected pairs of siblings using so-called nonparametric analysis. Highly polymorphic markers are used to identify chromosomal regions in which affected sib pairs with CD show more allele sharing than expected by chance. However, despite a number of genome-wide screens, no additional susceptibility loci have been identified so far (30–34).

Although all studies showed highly significant linkage to the HLA region on chromosome 6, the results of non-MHC loci were not extremely significant and still await replication in independent cohorts. There are several possibilities why these independent studies have not yet replicated each other's findings: First, most studies were based on a relatively small number of families or affected sib pairs. Second, since each of the studies to date examined individuals in different populations, any population-specific gene variants would not be confirmed from one another. Finally, differences in diagnostic criteria might have led to the mapping of diagnostic subtypes and obscured the evidence for common contributing genes.

Although no definitive results have been obtained with this approach to CD, evidence for additional predisposing genes has come from studies undertaken in other immune or autoimmune diseases. A recent study which compared 23 genome-wide screens showed that non-MHC loci clustered in a defined number of chromosomal loci in human autoimmune diseases (35). These data suggest that there are shared genetic factors predisposing to autoimmunity and that more disease-specific genes exist. In this respect, it is particularly relevant to consider the search for type 1 diabetes genes. As in CD, the major genetic component maps to the HLA-DQA/B alleles (36). A second type 1 diabetes-associated gene, which is far less important than the HLA locus, is the *Insulin* gene itself, in particular in its regulatory region, where a variable number of tandem repeat (VNTR) polymorphisms is linked to differences in insulin expression level. While one class of

VNTR alleles, VNTRI, renders one susceptible to the disease relative to the general population, the VNTRI-II class of alleles confers resistance. A lower expression of insulin in the thymus and a higher expression of the hormone in the pancreas characterize the susceptibility class of alleles (37, 38). This expression pattern may result in a limited tolerance induction to insulin and the susceptibility to type 1 diabetes associated with this class of alleles. The identification of CD susceptibility genes may likewise broaden our understanding of the specific and more general factors associated with autoimmune disorders.

The environment

The environment clearly plays a crucial role in the development of CD: No gluten, no disease! The characterization of a number of gluten-derived peptides that stimulate T cells

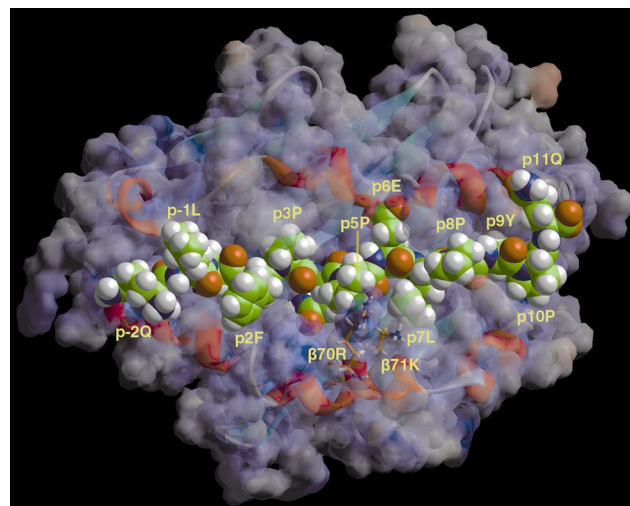


Figure 1

T cell receptor view of the $\alpha 1\beta 1$ domain of the modeled structure of the HLA-DQA1*0501/B1*0201 (DQ2) complexed with the α -I gliadin peptide, one of the dominant epitopes in adult DQ2-restricted CD. The protein molecule is in solid surface representation with colors according to the surface electrostatic potential (gray, neutral; blue, positive; red, negative). The antigenic peptide residues are in van der Waals space filling form, while the few DQ residues shown are in stick representation (color convention: red, oxygen; blue, nitrogen; green, carbon; white, hydrogen). Anchor residues at p1 and p9 point into the plane of the paper and are thus not visible. Secondary structure elements of the DQ2 are also shown for orientation purposes: α -helix in red, β -pleated sheet in turquoise, and random coil in gray. A transparency function has been included so that the secondary structural elements of DQ2 and the peptide residues buried by DQ2 residues become slightly visible. DQ2 is unique among MHC II alleles as it contains $\beta 70\text{Arg}/\beta 71\text{Lys}$, a combination responsible for the preference of acidic residues at p4/p7. Also, positions $\beta 28$ and $\beta 30$ are occupied by serine residues and $\beta 37$ by isoleucine (instead of Tyr for nearly all other DQB alleles), thus allowing ample space at p9 for the bulky and inflexible aromatic residues (here Tyr), the only DQ allele documented with this binding property to date. In this peptide, residue p6 has been transformed by tTG from Gln to Glu, making the peptide a better binder of DQ2 (3, 9, 39). Homology modeling, based on the coordinates of DQ8-insulin B9-23 complex, kindly provided by Kon Ho Lee (40), was performed via the program Discover II (MSI, San Diego, California, USA) and depiction on WebLabViewer 3.5 (MSI).

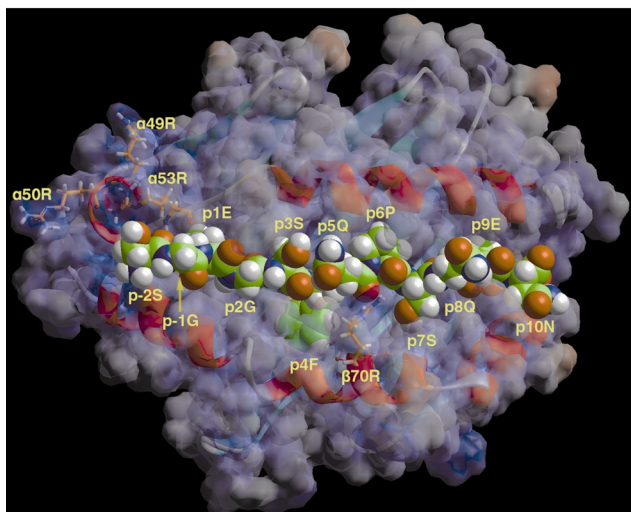


Figure 2

T cell receptor view of the $\alpha 1\beta 1$ domain of DQA1*0301/B1*0302 (DQ8) complexed with the α -I gliadin peptide, identified as a DQ8-specific epitope (Table 1). Color conventions and representations are as in Figure 1. DQ8 is also unique in its preference for acidic residues at p1 and p9 (40–43). These anchors also point into the plane of the paper and are not visible to the viewer. Four arginines around p1 ($\alpha 49$, $\alpha 50$, $\alpha 52$ [not shown], and $\alpha 53$) are responsible for the preference of an acidic residue at p1, and $\alpha 76$ Arg is responsible for a similar preference at p9. The peptide shown here is the transformed gliadin peptide (p1/9Glu), which is recognized by the same T cell clone, but at a higher sensitivity than the native one (p1/9Gln) (1, 6, 9, 43).

isolated from CD patients provides a mechanistic basis for this simple observation. In addition, although a particular gliadin sequence has been reported to be immunodominant in adult CD patients, several T cell-stimulatory gluten peptides have been found and there is evidence of many additional T cell-stimulatory peptides. We have also identified an immunogenic glutenin peptide, variants of which are found in several copies in glutenin. Similarly, variants of other T cell-stimulatory gluten sequences are often present in gluten, and it is becoming increasingly clear that gluten contains a relatively large number of T cell-stimulatory sequences. Because gluten is present in relatively large amounts in a variety of common food products, the daily gluten intake in a Western diet is high. In combination, we see that every HLA-DQ2- and/or -DQ8-positive individual is exposed to a large repertoire of immunogenic and abundant gluten peptides, and this may be an important factor determining disease development. There is, at present, no evidence linking additional environmental factors to CD.

Variability of CD presentation

Not only does CD develop in only a small percentage of HLA-DQ2/8-positive individuals, but the severity of the disease is quite variable, ranging from very severe in some patients to hardly noticeable or completely unnoticeable in others. In fact, several studies have indicated that many cases of undiagnosed CD occur due to the lack of complaints that can be linked to CD.

It is important to consider several aspects that are unique to CD and that distinguish it from the most

common autoimmune diseases. First, the nature and the time when individuals first become exposed to the disease-inducing antigen is well known. In addition, due to the high exposure to immunogenic gluten peptides, gluten-specific T cell responses may be quite common, both in CD patients and in healthy subjects with the right HLA background. In fact, gliadin-specific IgG antibodies are found in healthy individuals as well as patients, an indication of gliadin-specific Th cell responses.

Why, then, do not all HLA-DQ2/8-positive individuals develop CD? Is there another environmental factor – besides gluten – involved, or is this due to unknown genetic factors that fail to regulate the gluten-specific immune response? Given the short time period between exposure to gluten and disease development, an environmental factor would have to be a very common pathogen, a possibility that is difficult to reconcile with the limited proportion of affected individuals carrying HLA-DQ2/8. We suggest that the existence of a number of additional predisposing genetic alleles, in addition to HLA-DQ2/8, provides a more likely explanation. When these additional susceptibility loci are present in an HLA-DQ2/8-positive individual, disease will almost certainly develop, since there is a failure to regulate the developing gluten-specific T cell response. Conversely, when these susceptibility genes are not present in HLA-DQ2/8-positive individuals, the T cell response will be properly regulated and no disease will develop. Consequently, the number of susceptibility loci and the number of the predisposing alleles at each of these loci might influence severity of the disease and provide an explanation for the observed variation in disease severity.

Based on the genome scans performed so far, there may be additional genetic loci involved in CD which each contribute only a little to the total genetic risk. If this scenario turns out to be true we will have to perform our genetic studies on much larger sample sizes. Given the high frequency of CD in most countries, it should be feasible to collect large numbers of pedigrees or affected sibling pairs. Uniform diagnostic criteria – like the European Society for Pediatric Gastroenterology and Nutrition criteria – should also allow researchers to combine data sets in order to obtain larger cohorts. Alternatively, we may need to turn to other genetic strategies, such as genome-wide association studies. Once an interesting gene locus has been identified and replicated in an independent study, the appropriate gene and functional variants have to be identified. With the completion of the Human Genome Project, the process of positional cloning of the underlying gene can now be performed in silico and the most interesting genes – and their variants by virtue of single nucleotide polymorphisms (SNPs) – can be directly selected for further studies. Genetic epidemiological studies will be required to assess the relative risk associated with each gene variant. The presence of certain polymorphisms in causal genes may be useful in predicting vulnerability to the disease (for a relative in a celiac family), the clinical course of the dis-

ease, the risk of complications, and maybe the response to therapy. Much of this research will depend on the study of the functional polymorphisms and gene variants detected by association studies using SNPs.

Much new insight is also expected to be gained from new genomic and functional genomic approaches, such as cDNA microarray technology and proteomics, which will provide us with an extensive view of the complete transcriptome or proteome. Such studies will offer unique opportunities for obtaining valuable clues about the disease mechanisms by simultaneously monitoring the expression of a large number of human genes and their products during different stages of disease pathology.

Additionally, a comprehensive understanding of the magnitude and diversity of the gluten-specific T cell response is needed. The present studies on the specificity of the gluten response have all been carried out with adult patients and may not be representative for the response early in disease development. Studies to determine the specificity of the response in young CD patients are therefore under way, and these may shed more light on the nature of disease-inducing T cell responses. Moreover, the use of HLA-tetramers will give detailed insight into the magnitude of the T cell response to the various gluten peptides, information which may be crucial for the rational design of alternative intervention protocols and the selection of less toxic wheat strains.

Prospects for a healthy life

Clearly, the recent advances in CD research have improved our understanding of the factors causing disease, but can this be translated into better prevention or the development of a cure for the disease? The identification of genes playing an important role in CD will ultimately open new ways for disease diagnosis, prognosis, prevention, and therapy. With the genes in hand, we should be able to study the complex interplay between the different genes, as well as between genes and environmental factors, and provide a platform for experimental studies to decipher the pathways involved in the disease process. This knowledge will be essential for developing new immunological or pharmacological treatment strategies.

The development of an alternative cure, however, is still a long way off. In general, it will be very difficult to induce oral tolerance to gluten when oral tolerance could not be established in the first place. This is further complicated by the fact that disease is usually diagnosed only upon the development of severe symptoms and an advanced T cell response to gluten. There is also no easy way to block an existing T cell response to gluten. While immunosuppressive drugs have such serious side effects that they cannot be used, more specific agents might be considered. Unfortunately, the altered peptide ligand approach, which is intended to block peptide-specific T cell responses, has generally been unsuccessful. Other potential interventions may become possible, however, when additional genes have been identified that play a role in CD development.

The prospects for better prevention are more hopeful. Obviously, with the identification of T cell-stimulatory gluten peptides, it may be possible to develop wheat strains free of those sequences, although this change might alter the properties of the gluten and make the wheat less suited for its normal applications. The identification of additional risk markers will help to discriminate between individuals at risk and those not at risk, especially in families with a CD history, information that would have a strong impact on prevention.

Acknowledgments

This work was supported by grants from the European Union (BIO4 CT95 0263 to F. Koning and G.K. Papadopoulos, BHM4-CT98 to F. Koning), a grant from the Hellenic General Secretariat of Research and Technology (to G.K. Papadopoulos), grants from the Dutch Digestive Disease Foundation (to F. Koning and C. Wijmenga), and the Research Committee of the Technological Educational Institute of Epirus (to G.K. Papadopoulos). We thank Antonis K. Moustakas for modeling the DQ2 and DQ8 peptide complexes and designing the Figures, Sari Iltanen for critical reading of the manuscript, and Jackie Senior for editing the manuscript.

1. Wal, Y., et al. 1988. Small intestinal cells of celiac disease patients recognize a natural pepsin fragment of gliadin. *Proc. Natl. Acad. Sci. USA.* **95**:10050–10055.
2. Wal, Y., et al. 1999. Glutenin is involved in the gluten-driven mucosal T cell response. *Eur. J. Immunol.* **29**:3133–3139.
3. Sjöström, H., et al. 1998. Identification of a gliadin epitope in coeliac disease: general importance of gliadin deamidation for intestinal T cell recognition. *Scand. J. Immunol.* **48**:111–115.
4. Arentz-Hansen, H., et al. 2000. The intestinal T cell response to α -gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. *J. Exp. Med.* **191**:603–612.
5. Anderson, R.P., Degano, P., Godkin, A.J., Jewell, D.P., and Hill, A. 2000. In vivo antigen challenge in celiac disease identifies a single transglutaminase-modified peptide as the dominant A-gliadin T-cell epitope. *Nat. Med.* **6**:337–342.
6. Wal, Y., et al. 1998. Cutting edge: selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *J. Immunol.* **161**:1585–1588.
7. Molberg, Ø., et al. 1998. Tissue transglutaminase selectively modifies gliadin peptides that are recognised by gut-derived T cells in celiac disease. *Nat. Med.* **4**:713–717.
8. Wal, Y., et al. 2000. Coeliac disease: it takes three to tango! *Gut.* **46**:734–737.
9. Sollid, L.M. 2000. Molecular basis of celiac disease. *Annu. Rev. Immunol.* **18**:53–81.
10. Csizmadia, G.D.S., Mearin, M.L., von Blomberg, B.M.E., Brand, R., and Verloove-Vanhorick, S.P. 1999. An iceberg of childhood coeliac disease in the Netherlands. *Lancet.* **353**:813–814.
11. Hovdenak, N., et al. 1999. High prevalence of asymptomatic coeliac disease in Norway: a study of blood donors. *Eur. J. Gastroenterol. Hepatol.* **11**:185–187.
12. Ivarsson, A., et al. 1999. High prevalence of undiagnosed coeliac disease in adults: a Swedish population-based study. *J. Intern. Med.* **245**:63–68.
13. Korponay-Szabo, I.R., et al. 1999. High prevalence of silent celiac disease in preschool children screened with IgA/IgG antiendomysium. *J. Pediatr. Gastroenterol. Nutr.* **28**:26–30.
14. Not, T., et al. 1998. Celiac disease risk in the USA: high prevalence of antibodies in healthy blood donors. *Scand. J. Gastroenterol.* **33**:494–498.
15. Dam, M., et al. 1988. Anti-alpha-gliadin antibodies (AGA) in the serum of coeliac children and controls recognize an identical collection of linear epitopes of alpha-gliadin. *Clin. Exp. Immunol.* **114**:189–195.
16. Dietrich, W., et al. 1998. Autoantibodies to tissue transglutaminase as predictors of celiac disease. *Gastroenterology.* **115**:1317–1321.
17. Sollid, L.M., Molberg, Ø., McAdam, S., and Lundin, K.E. 1997. Autoantibodies in coeliac disease: tissue transglutaminase – guilt by association? *Gut.* **41**:851–852.

18. Trier, J.S. 1998. Diagnosis of celiac sprue. *Gastroenterology*. **115**:211–216.
19. Risch, N. 1987. Assessing the role of HLA-linked and unlinked determinants of disease. *Am. J. Hum. Genet.* **40**:1–14.
20. Petronzelli, F., et al. 1997. Genetic contribution of the HLA region to the familial clustering of coeliac disease. *Ann. Hum. Genet.* **61**:307–317.
21. Bevan, S., et al. 1999. Contribution of the MHC region to the familial risk of coeliac disease. *J. Med. Genet.* **36**:687–690.
22. Alper, C.A., Fleischnick, E., Awdeh, Z., Katz, A.J., and Yunis, E.J. 1987. Extended major compatibility complex haplotypes in patients with gluten-sensitive enteropathy. *J. Clin. Invest.* **79**:251–256.
23. Sollid, L.M., and Thorsby, E. 1993. HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. *Gastroenterology*. **105**:910–922.
24. Herr, M., et al. 2000. Evaluation of fine mapping strategies for a multifactorial disease locus: systematic linkage and association analysis of IDDM1 in the HLA region on chromosome 6p21. *Hum. Mol. Genet.* **9**:1291–1301.
25. van Belzen, M.J., et al. 2000. A new locus within the MHC region strongly contributes to celiac disease. *J. Pediatr. Gastroenterol. Nutr.* **31**(Suppl.):S15. (Abstr.)
26. 1999. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. *Nature*. **401**:921–923.
27. Holopainen, P., et al. 1999. CD28/CTLA4 gene region on chromosome 2q33 confers genetic susceptibility to celiac disease. A linkage and family-based association study. *Tissue Antigens*. **53**:470–475.
28. Clot, F., et al. 1999. Linkage and association study of the CTLA-4 region in coeliac disease for Italian and Tunisian populations. *Tissue Antigens*. **54**:527–530.
29. Naluai, A.T., et al. 2000. The CTLA4/CD28 gene region on chromosome 2q33 confers susceptibility to celiac disease in a way possibly distinct from that of type 1 diabetes and other chronic inflammatory disorders. *Tissue Antigens*. **56**:350–355.
30. Zhong, F., et al. 1996. An autosomal screen for genes that predispose to celiac disease in the western counties of Ireland. *Nat. Genet.* **14**:329–333.
31. Houlston, R.S., et al. 1997. Linkage analysis of candidate regions for coeliac disease genes. *Hum. Mol. Genet.* **6**:1335–1339.
32. Brett, P.M., et al. 1998. A pedigree-based linkage study of coeliac disease: failure to replicate previous positive findings. *Ann. Hum. Genet.* **62**:25–32.
33. Greco, L., et al. 1998. Genome search in celiac disease. *Am. J. Hum. Genet.* **62**:669–675.
34. Lui, J., et al. 2000. A genome-wide mapping on celiac disease in 60 Finnish families. *J. Pediatr. Gastroenterol. Nutr.* **31**(Suppl.):S5.
35. Becker, K.S., et al. 1998. Clustering of non-major histocompatibility complex susceptibility candidate loci in human autoimmune diseases. *Proc. Natl. Acad. Sci. USA*. **95**:9979–9984.
36. Thorsby, E., and Rønningen, K.S. 1993. Particular HLA-DQ molecules play a dominant role in determining susceptibility or resistance to type 1 (insulin-dependent) diabetes mellitus. *Diabetologia*. **36**:371–377.
37. Vafiadis, P., et al. 1997. Insulin expression in human thymus is modulated by *INS* VNTR alleles at the *IDDM2* locus. *Nat. Genet.* **15**:289–292.
38. Pugliese, A., et al. 1997. Transcription levels of human insulin in the thymus correlate with allelic variation at the *IDDM2* susceptibility locus for type 1 diabetes. *Nat. Genet.* **15**:293–297.
39. Quarsten, H., Molberg, Ø., Fugger, L., McAdam, S.N., and Sollid, L.M. 1999. HLA binding and T cell recognition of a tissue transglutaminase-modified gliadin epitope. *Eur. J. Immunol.* **29**:2506–2514.
40. Lee, K.H., Wucherpfennig, K.W., and Wiley, D.C. 2001. Structure of a human insulin peptide–HLA-DQ8 complex and susceptibility to type 1 diabetes. *Nat. Immunol.* **2**:501–507.
41. Kwok, W.W., Domeir, M.E., Raymond, F.C., Byers, P., and Nepom, G.T. 1996. Allele-specific motifs characterize HLA-DQ interactions with a diabetes-associated peptide derived from glutamic acid decarboxylase. *J. Immunol.* **156**:2171–2177.
42. Godkin, A., et al. 1997. Use of eluted peptide sequence data to identify the binding characteristics of peptides to the insulin-dependent diabetes susceptibility allele HLA-DQ8 (DQ3.2). *Int. Immunol.* **9**:905–911.
43. Moustakas, A.K., et al. 2000. Structure of celiac disease-associated HLA-DQ8 and non-associated HLA-DQ9 alleles in complex with two disease-specific epitopes. *Int. Immunol.* **12**:1157–1166.