Integration of Genetic and Immunological Insights into a Model of Celiac Disease Pathogenesis

Valérie Abadie,[1] Ludvig M. Sollid,[2] Luis B. Barreiro,[3] and Bana Jabri[1,4,5]

Departments of [1]Medicine, [3]Human Genetics, [4]Pathology, and [5]Pediatrics, University of Chicago, Chicago, Illinois 60637; email: lbarreir@bsd.uchicago.edu, bjabri@bsd.uchicago.edu

[2]Centre for Immune Regulation, Institute of Immunology, University of Oslo and Oslo University Hospital–Rikshospitalet, N-0372 Oslo, Norway

Annu. Rev. Immunol. 2011. 29:493–525

First published online as a Review in Advance on January 3, 2011

The *Annual Review of Immunology* is online at immunol.annualreviews.org

This article's doi: 10.1146/annurev-immunol-040210-092915

0732-0582/11/0423-0493$20.00

Keywords

gluten intolerance, genome-wide association studies, inflammatory disorder, autoimmune disease, HLA

Abstract

Celiac disease (CD) is a gluten-sensitive enteropathy that develops in genetically susceptible individuals by exposure to cereal gluten proteins. This review integrates insights from immunological studies with results of recent genetic genome-wide association studies into a disease model. Genetic data, among others, suggest that viral infections are implicated and that natural killer effector pathways are important in the pathogenesis of CD, but most prominently these data converge with existing immunological findings that CD is primarily a T cell–mediated immune disorder in which $CD4^+$ T cells that recognize gluten peptides in the context of major histocompatibility class II molecules play a central role. Comparison of genetic pathways as well as genetic susceptibility loci between CD and other autoimmune and inflammatory disorders reveals that CD bears stronger resemblance to T cell–mediated organ-specific autoimmune than to inflammatory diseases. Finally, we present evidence suggesting that the high prevalence of CD in modern societies may be the by-product of past selection for increased immune responses to combat infections in populations in which agriculture and cereals were introduced early on in the post-Neolithic period.

CD: celiac disease

HLA: human leukocyte antigen

TG2: transglutaminase 2

T1D: type 1 diabetes

HISTORY AND KEY FEATURES OF CELIAC DISEASE

The first clear description of celiac disease (CD) was found in the writings of the Greek physician Aretaeus between the first and second centuries AD (reviewed in References 1 and 2). CD was described as an intestinal disorder associated with diarrhea and malabsorption occurring in children and adults, more frequently in women than in men (3). The idea that the disease was linked to food ingestion was brought forward in 1888 by Gee (4). This idea was confirmed in the 1950s, when Dicke and colleagues (5) established that the consumption of wheat and rye brought on CD and that removing these grains from the diet resulted in a marked improvement in the condition. In 1954, Paulley (6) became the first to report that the clinical manifestations of CD are linked to destruction of the lining of the small intestine. Many years later, Marsh (7) established a histological classification of celiac lesions, which range from hyperproliferative crypts with intraepithelial lymphocytosis to total villous atrophy.

Around the period in which important progress was made in the clinical and histological definition of CD, epidemiological studies showed that twins and first-degree relatives have a much higher incidence of CD than do members of the general population, indicating a genetic component in CD (8–11). A link to human leukocyte antigen (HLA) alleles was suggested in pioneering studies by the Strober (12) and Cooke (13) groups, who showed that 88% of adult CD patients in the United States and England have the HLA-B8 antigen, compared with 22–30% of controls. Later, it was found that the association is stronger with HLA-DR3 (14, 15) and HLA-DQ2 (16). Cloning of major histocompatibility (MHC) class II genes finally revealed that the genes encoding an HLA-DQ2 variant and HLA-DQ8 are the causative genes for CD (17, 18). Insights into the molecular basis of the association of CD with MHC class II molecules were provided more than 30 years after the first evidence for an HLA association was found (19–26).

Interestingly, during the establishment of the association of CD with HLA genes, Ferguson et al. (27) and Ferguson & MacDonald (28) reported that CD is associated with a lymphocyte-mediated immunity to gluten within the small intestinal mucosa and that T cell–mediated immunity causes villous atrophy and crypt hyperplasia in an allograft rejection model. These observations were fully appreciated when immunological studies performed on intestinal biopsies showed that inflammatory gluten-reactive T cells recognized gluten selectively in the context of HLA-DQ2 or HLA-DQ8 molecules (29, 30) and were present only in the small intestinal mucosa of individuals with CD (31). In addition to providing a molecular basis for the association with MHC class II molecules, these studies suggested that HLA-DQ2- or HLA-DQ8-restricted $CD4^+$ T cells are critical to the pathogenesis of CD. Why gluten peptides, which are very rich in glutamines but very poor in acidic residues, bind to HLA-DQ2 and HLA-DQ8 molecules that have a preference for negatively charged peptides remained enigmatic until it was found that gluten is an excellent substrate for transglutaminase 2 (TG2; also known as tissue transglutaminase) (32, 33). This enzyme converts glutamine residues into negatively charged glutamate residues in a process termed deamidation. Strikingly, Schuppan and colleagues (34) found that CD patients develop autoantibodies against the same enzyme, which suggests that CD has an autoimmune component despite its induction by a dietary antigen (see discussion of the mechanism underlying the induction of autoantibodies in the section entitled Immunological Model of Celiac Disease Pathogenesis). Accompanying antibodies to TG2 as well as to gluten is a massive plasmacytosis in the lamina propria with dominance of immunoglobulin A (IgA) plasma cells in the overt celiac lesion (35). The relationship between CD and autoimmunity is further supported by epidemiological studies that show a link between CD and autoimmune disorders, in particular type 1 diabetes (T1D) and autoimmune thyroiditis (36).

Although the role of $CD4^+$ T cells in CD pathogenesis is well established, these cells' effector role in mediating tissue damage has been questioned by studies in human and mouse suggesting that $CD4^+$ T cell–mediated adaptive antigluten immunity is necessary but not sufficient to induce intestinal damage, specifically villous atrophy. This idea had been proposed in 1993 by Ferguson et al. (37, p. 150), who wrote:

> Although mucosal immunological sensitization is an invariable feature of celiac disease, it is not the precipitating factor for the expression of the full intestinal lesion; a second factor drives the enteropathy from minimal (latent) to overt. . . Candidate factors include an episode of hyperpermeability, nutrient deficiency, increased dietary gluten, impaired intraluminal digestion of ingested gluten, adjuvant effects of intestinal infection and a non-HLA associated gene.

We know now that CD is a complex multigenic disorder that involves HLA and non-HLA genes, adaptive and innate immunity, and environmental factors.

In addition to increased numbers of T cells and plasma cells in the lamina propria, there is, early on in the disease process, a marked increase in the number of intraepithelial lymphocytes (IELs) (7). The role of IELs in CD pathogenesis had long been disregarded because no link to MHC class I genes could be established (38) and no gluten-specific IELs could be identified. However, the increase in cytotoxic T cell receptor (TCR)$\alpha\beta^+$ IELs typically correlates with the presence of villous atrophy (39), and the malignant transformation of IELs is a hallmark of CD (40, 41), which suggests that IELs are implicated and abnormally activated in CD. A breakthrough came with studies showing that cytotoxic $CD8^+$ TCR$\alpha\beta^+$ IELs expressing activating natural killer (NK) cell receptors induced the killing of intestinal epithelial cells expressing stress- and inflammation-induced nonclassical MHC class I molecules (42–46). These findings unraveled the key role played by $CD8^+$ IELs in inducing villous atrophy (47).

Despite the numerous advances in the field of immunological disorders, many questions remain unanswered. CD is the only human immune-mediated disease for which we have comprehensive immunological information on the target tissue under normal and diseased conditions, in addition to extensive epidemiological and genetic data. In this review, we discuss CD pathogenesis and its relationship to other autoimmune and inflammatory disorders in light of the knowledge obtained from these different fields of investigation.

IEL: intraepithelial lymphocyte

NK: natural killer

EPIDEMIOLOGY OF A COMPLEX DISEASE

The Celiac "Iceberg" or the Clinical and Pathological Spectrum of Celiac Disease

CD can occur at all ages following the introduction of gluten to the diet. Similar to most autoimmune disorders, CD is more frequently (twice as often) found in women than in men (48). The clinical expression of the disease is very eclectic: The most typical manifestations are related to nutrient malabsorption (diarrhea, failure to thrive in children, anemia, etc.) (49). The diagnosis of CD is made based on the presence of anti-TG2 antibodies and intestinal villous atrophy (see sidebar). CD has a wide biological, histological, and clinical spectrum. Some healthy family members of CD patients show a local increased inflammatory response to rectal gluten challenge (50). Other individuals with anti-TG2 antibodies have normal intestinal morphology but can present with gluten-sensitive skin lesions in the context of a disease known as dermatitis herpetiformis (47). Still other patients present typical CD features with severe malabsorption and total villous atrophy. Finally, patients with the most severe form of the disease become refractory to a gluten-free diet and develop enteropathy-associated T cell lymphomas. Due to the heterogeneity of CD manifestations, investigators (51–55) have proposed a representation of CD as an iceberg reflecting different forms and/or

KEY FEATURES OF CELIAC DISEASE

1. Gluten and gluten-related proteins present in wheat, rye, and barley are the causative antigens of CD.
2. Histological lesions are characterized by the presence of crypt hyperplasia, intraepithelial lymphocytosis, and destruction of the surface epithelial lining of the small intestine.
3. Clinical presentation is eclectic, but the most characteristic presentations are linked to the malabsorption of nutrients.
4. The presence of autoantibodies directed against TG2 suggests that CD has an autoimmune component.
5. Epidemiological studies show a high prevalence of autoimmune disorders in CD patients and, conversely, a high incidence of CD in autoimmune patients.
6. CD occurs almost exclusively in patients who express the MHC class II HLA-DQ2 and HLA-DQ8 molecules.
7. Posttranslational modifications of gluten by TG2 result in the introduction of acidic residues and better binding of gluten peptides to the HLA-DQ2 and HLA-DQ8 molecules.
8. $CD4^+$ T cells in the lamina propria of CD patients recognize gluten peptides in the context of HLA-DQ2 or HLA-DQ8, and the preferential presentation of gluten peptides by these molecules explains the HLA association.
9. In CD patients, there is an expansion of cytotoxic IELs that express activating NK cell receptors, which recognize stress- and inflammation-induced nonclassical MHC class I molecules. These NK receptors mediate epithelial cell destruction by lowering the TCR-activation threshold of IELs or by mediating direct TCR-independent killing.

stages of antigluten immunity. Expanding on Ferguson et al.'s hypothesis, we proposed that to develop villous atrophy, patients must have an intestinal stress response that, in association with adaptive antigluten immunity, leads to the activation of IELs and villous atrophy (56). Patients who have only the adaptive antigluten immune response would have anti-TG2 antibodies and possible intraepithelial lymphocytosis but would conserve a normal intestinal architecture (57). Conversely, patients with epithelial stress and no adaptive antigluten immunity would show signs of gluten sensitivity in the absence of anti-TG2 antibodies (56). In any case, the celiac iceberg suggests that multiple "hits" are required to develop the classical features of CD, namely adaptive antigluten immunity and villous atrophy. Future studies will help delineate the different clinical and pathological representations of dysregulated immune responses to gluten, as well as their underlying genetic risk factors. One day we may conclude that, depending on the genetic background and environmental factors, some CD patients will develop villous atrophy, whereas other patients will suffer from gluten-induced irritable bowel syndrome or from a neurological disease such as gluten ataxia.

Celiac Disease Is a Multifactorial Disorder Whose Development Is Controlled by a Combination of Genetic and Environmental Risk Factors

CD is a complex disorder, the development of which is controlled by a combination of genetic and environmental risk factors. The primary environmental factor associated with the development of CD is gluten consumption. The critical role played by wheat gluten (consisting of gliadins and glutenins) and the related proteins of rye and barley (58) is illustrated by the fact that, under a gluten-free diet, clinical symptoms of disease, anti-TG2 antibodies, and villous atrophy typically recede. From a genetic perspective, susceptibility to CD is strongly associated with the MHC class II molecules HLA-DQ2 and HLA-DQ8. Indeed, almost all patients with CD express at least one of these HLA molecules (17, 18).

Given the key roles played by gluten (environment) and HLA-DQ2 and HLA-DQ8 (genetics) in the development of CD, one might predict that the regions of the globe where these risk factors are found at higher frequencies should present elevated rates of CD. We compiled the prevalence of CD, the levels of wheat consumption, and the frequencies of HLA-DQ2 and HLA-DQ8 for different regions of the globe (**Figure 1**). Overall, our analyses reveal that although these two factors are required for the development of the disease, in the absence of other factors they are not strong

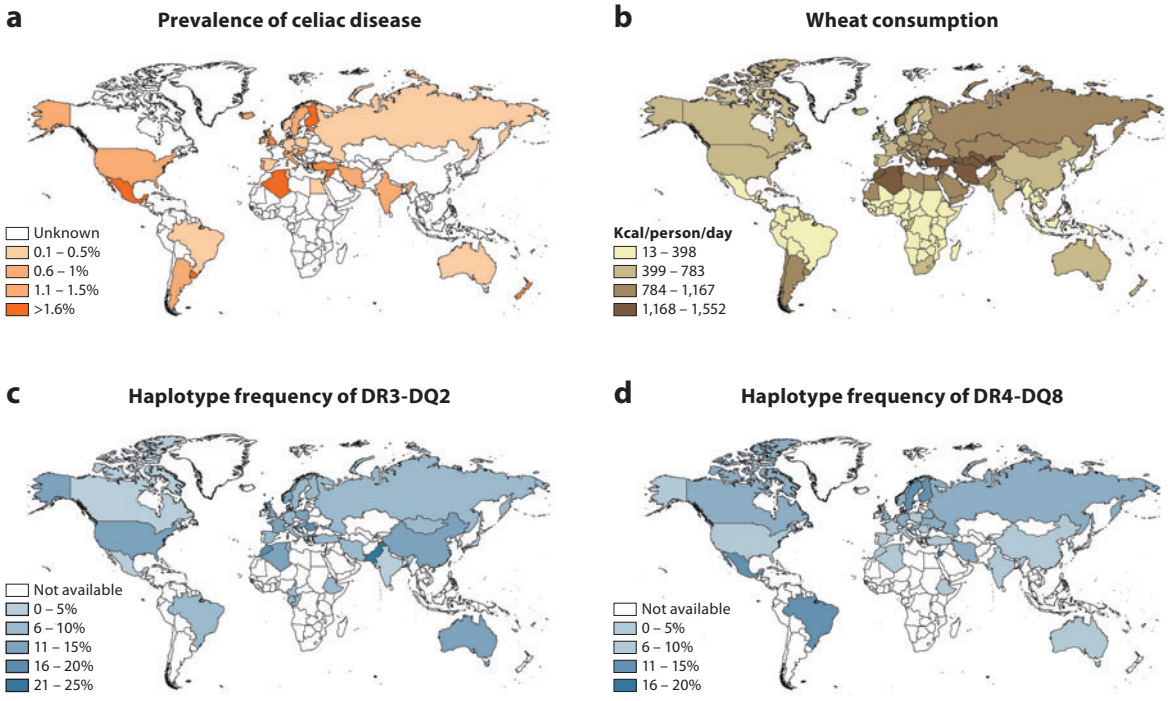


Figure 1

Prevalence of celiac disease (CD), wheat consumption, and frequencies of the DR3-DQ2 and DR4-DQ8 haplotypes worldwide. (*a*) Prevalence of CD based on the screening of adult populations or on the screening of child populations when the prevalence in adults has not been determined. (*b*) Worldwide distribution of wheat consumption levels. Data were obtained from the Food and Agriculture Organization of the United Nations (**http://www.fao.org**). (*c*) Frequency of the DR3-DQ2 haplotype (DRB1*0301-DQA1*0501-DQB1*0201). (*d*) Frequency of the DR4-DQ8 haplotype (DRB1*04-DQA1*03-DQB1*0302).

predictors of the prevalence of CD in most parts of the world. Indeed, and possibly surprisingly at first glance, we do not observe a significant correlation between the prevalence of CD and the levels of wheat consumption, the sum of the frequencies of DR3-DQ2 and DR4-DQ8, or the product of both factors [i.e., the frequency of (DQ2+DQ8) × wheat consumption] (**Figure 2*a–c***). However, the dual requirement of wheat and HLA for the development of the disease is well illustrated in Burkina Faso, where the prevalence of CD is zero, probably due to a very low frequency of HLA-DQ2 or HLA-DQ8 genes and low levels of wheat consumption (59).

Further analysis reveals that the overall lack of correlation between wheat consumption and CD-predisposing HLA expression with CD prevalence is driven primarily by a small number of clear outlier populations spread over most of the continents: Algeria, Finland, Mexico, north India, and Tunisia (**Figure 2**). Indeed, by excluding these populations, we observe a significant correlation between the combination of both risk factors and the incidence of CD worldwide (correlation coefficient $R^2 = 0.4$; P value $= 0.002$) (**Figure 2*d***). We also observe a significant correlation between the prevalence of CD and wheat consumption ($R^2 = 0.14$, P value $= 0.03$) and the prevalence of CD and the frequency of DQ2+DQ8 haplotypes ($R^2 = 0.24$, P value $= 0.03$). The existence of clear outlier populations, together with the fact that the observed correlations are far from complete (i.e., $R^2 = 1$), suggests that other environmental and genetic factors must contribute to the development or pathogenesis of CD.

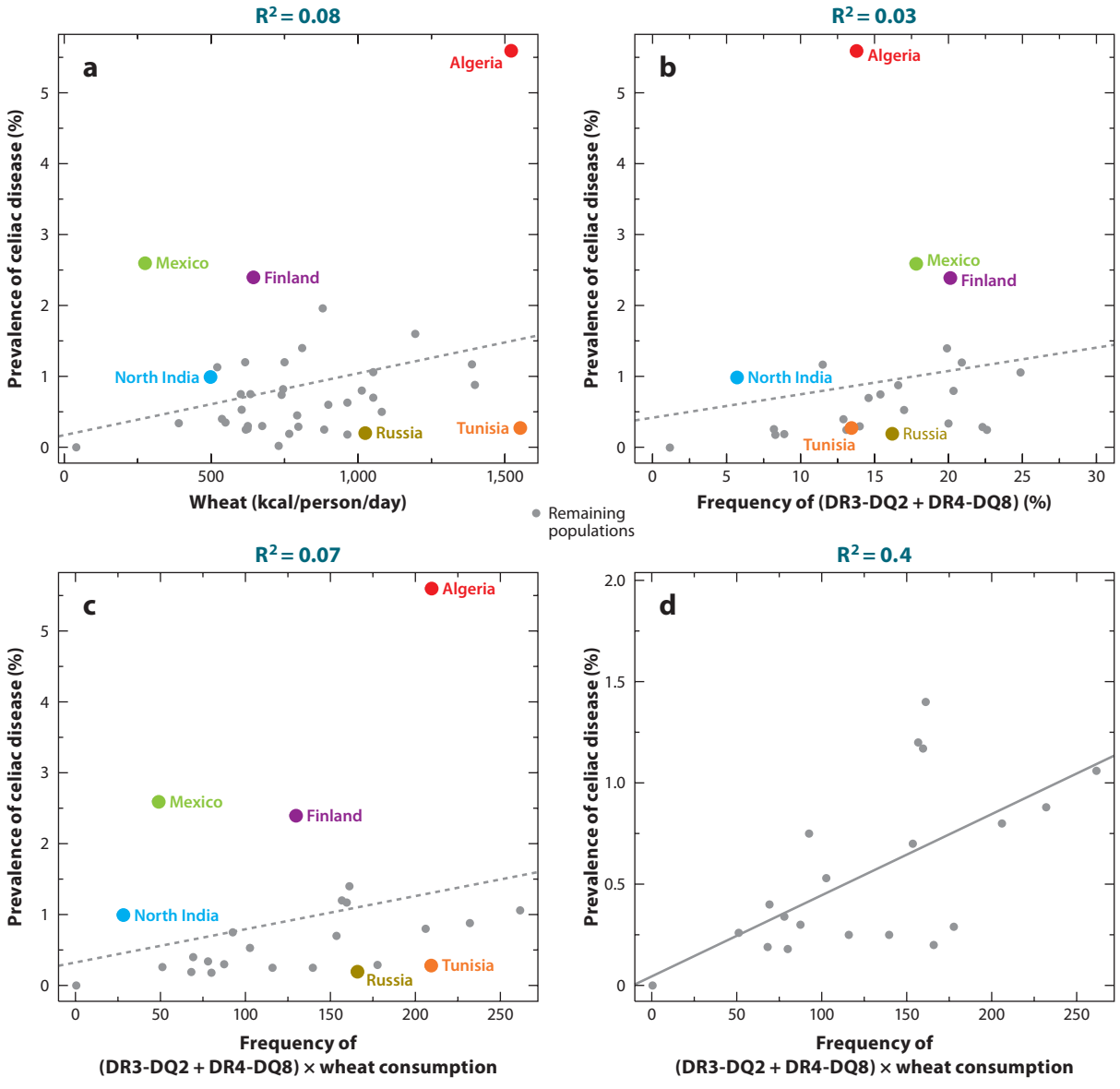


Figure 2

Correlations between the prevalence of celiac disease (CD), wheat consumption, and the frequencies of the DR3-DQ2 and DR4-DQ8 haplotypes. (*a*) Correlation between the prevalence of CD (*y* axis) and wheat consumption (*x* axis). (*b*) Correlation between the prevalence of CD (*y* axis) and the sum of the frequencies of the DR3-DQ2 and DR4-DQ8 haplotypes (*x* axis). (*c*) Correlation between the prevalence of CD (*y* axis) and the product of the frequencies of DR3-DQ2+DR4-DQ8 and the amounts of wheat consumption (*x* axis). (*d*) Correlation between the prevalence of CD (*y* axis) and the product of the frequencies of DR3-DQ2+DR4-DQ8 and the amounts of wheat consumption (*x* axis) after excluding the following outlier populations: Algeria, Finland, Mexico, north India, and Tunisia.

In the Maghreb area, where wheat and barley are the major staple foods, there is a remarkable disparity between the incidences of CD in the neighboring countries of Algeria and Tunisia. Indeed, despite similar frequencies of the DR3-DQ2 and DR4-DQ8 haplotypes (60–62), the prevalence of CD in Algeria (5.6%) is by far the highest reported worldwide (63), whereas the prevalence of CD in Tunisia (0.28%) remains one of the lowest (**Figures 1** and **2**) (64). A similar pattern is observed between two other adjoining countries, Finland and Russia (**Figures 1** and **2**). Although these two countries have similar wheat consumption levels and comparable HLA haplotype frequencies, the prevalence of CD in Finland is 1–2.4% (65–67), whereas in the adjacent Russian republic of Karelia, the prevalence of CD is considerably lower (0.2%) (68). Mexico is an interesting example where, despite a very low level of wheat consumption, a high prevalence of serological CD has been reported (69). Altogether, these observations suggest that other environmental and/or genetic factors can significantly impact disease outcome. Such factors could, for instance, influence the microbiome of individuals, which in turn could change the immunological responses to oral antigens. It would be of great interest to obtain more information on CD for distinct regional areas that could display different dietary habits and/or genetic features, as the data available do not necessary reflect the whole-country situation. For example, in northern China, where there are a high prevalence of CD-associated HLA and a high level of wheat consumption, there may be a high prevalence of CD (70) that would be missed when looking at the country as a whole. This possibility remains to be investigated, as we have only very limited information on CD prevalence in China. Altogether, these observations suggest that similar levels of wheat consumption and predisposing HLA expression can be associated with strikingly different levels of CD prevalence, which highlights the role of environmental factors and other genetic risk factors in CD pathogenesis.

ROLE OF THE HUMAN LEUKOCYTE ANTIGEN LOCUS IN CELIAC DISEASE PATHOGENESIS

Genetic Insights

HLA is the single most important susceptibility locus for CD (71). As mentioned above, the primary genetic factors associated with CD are the MHC class II genes that encode HLA-DQ2 and HLA-DQ8 (**Figure 3**). HLA-DQ2 is, however, more strongly associated with CD than HLA-DQ8 is (38). For example, 89% of CD patients from France have one or two copies of HLA-DQ2.5, compared with 21% in a matched control population (72). That HLA-DQ2 and HLA-DQ8 molecules are also commonly found in healthy individuals demonstrates that they contribute to but are not sufficient for disease development.

Several haplotypes encoding for the risk HLA-DQ2.5 heterodimer have consistently been associated with CD in several populations (**Figure 3**). Indeed, the risk heterodimer HLA-DQ2.5 can be encoded in *cis*, when both DQA1*0501 and DQB1*0201 are located on the same DR3-DQ2 haplotype, or in *trans*, when these two molecules are located on different haplotypes, namely DR5-DQ7 and DR7-DQ2 (**Figure 3**). The resulting *cis* and *trans* HLA-DQ2.5 heterodimers differ by only one residue in the leader peptide of the DQ α-chains (DQA1*0501 versus DQA1*0505) and by one residue in the membrane-proximal domain of the DQ β-chains (DQB1*0201 versus DQB1*0202) (73). It is unlikely that these differences have any functional consequence, and they are considered to confer a similar disease risk. In contrast, there is a dramatic difference in the genetic risk conferred by HLA-DQ2.5 and by HLA-DQ2.2; see below (**Figure 3**) (74).

Disease susceptibility depends on the dosage effect of the DQ2.5 heterodimer (75, 76). Homozygous individuals for the DR3-DQ2 haplotype or heterozygous DR3-DQ2/DR7-DQ2 express the highest levels of DQ2.5 heterodimers (77). These two

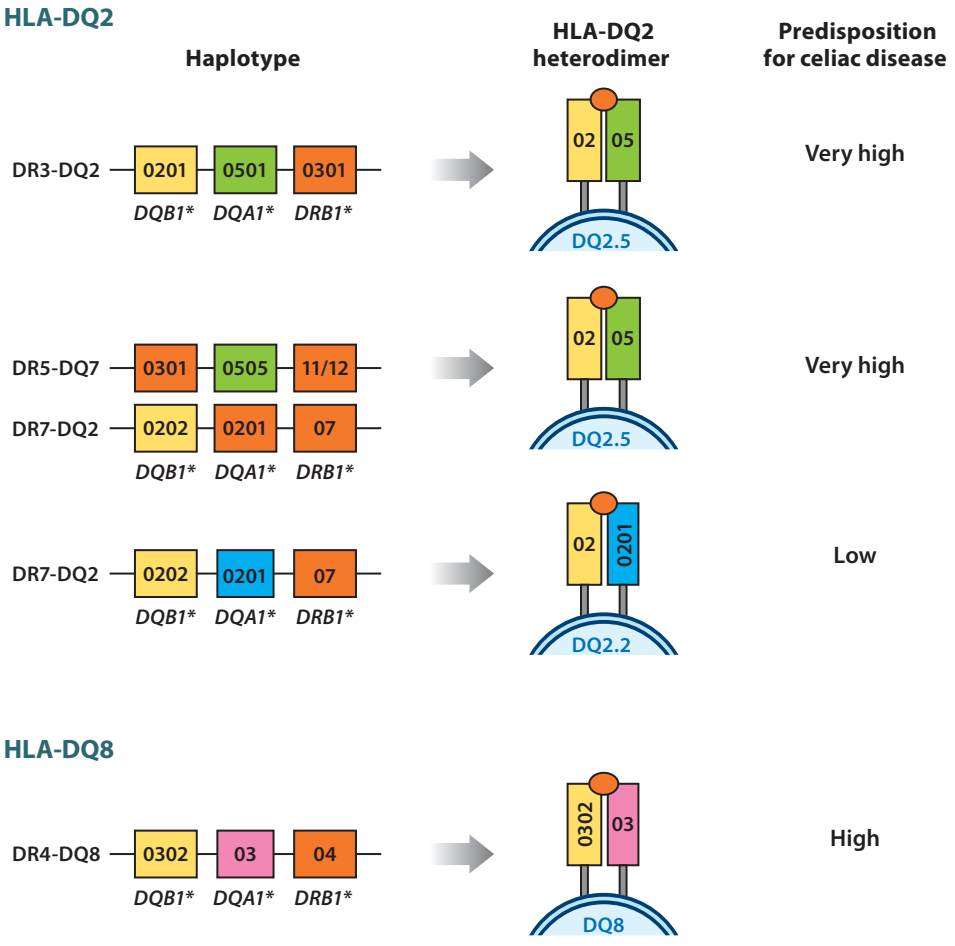


Figure 3

Human leukocyte antigen (HLA) associations in celiac disease (CD). HLA-DQ2 is the strongest genetic risk factor associated with CD. The great majority of CD patients express the HLA-DQ2.5 heterodimer encoded by the HLA-DQA1*05 (α-chain) and HLA-DQB1*02 (β-chain) alleles. These two alleles are carried either in *cis* on the DR3-DQ2.5 haplotype or in *trans* in individuals who are DR5-DQ7 and DR7-DQ2.2 heterozygous. HLA-DQ2.2, another variant of the HLA-DQ2 molecule, is encoded by the HLA-DQA1*0201 and HLA-DQB1*02 alleles and confers a very low risk for CD on its own. DQ2-negative patients express HLA-DQ8, which is encoded by the DR4-DQ8 haplotype.

genotypes are associated with the highest risk of CD. A dosage effect for DQ8 molecules has also been suggested (78). Furthermore, refractory CD patients who do not respond to a gluten-free diet and have aberrant intestinal T cells have greatly increased levels of homozygosity for the DR3-DQ2 haplotype (44–62%), compared with other CD patients (20–24%) (79).

Along with the genes encoding the DQ molecules, the HLA locus contains additional immune-related genes that may impact susceptibility to CD. In accordance with this hypothesis, several studies have suggested that genetic variation in other HLA-associated genes, such as *MICA*, *MICB*, and *TNF*, can also predispose to CD (reviewed in Reference 38). However, the observed associations should be interpreted

cautiously because most of these studies failed to formally correct for the levels of linkage disequilibrium (LD) (i.e., genetic association) between these genes and the genes encoding the DQ risk molecules. Future resequencing or fine mapping studies of the HLA regions in large patient cohorts should help determine whether or not the HLA region contains susceptibility factors in addition to those already recognized for the DQ region.

The Immunological Role of the HLA-DQ2 and HLA-DQ8 Molecules

The genetic and epidemiological findings that position HLA-DQ2 and HLA-DQ8 molecules at the center of CD pathogenesis are supported by functional studies showing that gluten-specific $CD4^+$ T cells can be isolated from the mucosa of CD patients but not from that of healthy controls (31). Further, such $CD4^+$ T cells selectively recognize gluten in the context of HLA-DQ2 or HLA-DQ8 molecules and have a strong preference for deamidated gluten peptides over native gluten peptides, which lack negatively charged residues (**Figure 4**) (21, 29, 32, 33). Altogether, these findings indicate that the pathological response in the intestinal environment associated with the development of villous atrophy is HLA-DQ2 or HLA-DQ8 restricted and is directed mainly against deamidated gluten peptides. Deamidation is mediated by the enzyme TG2, which targets specific glutamine residues, particularly in glutamine-X-proline sequences (where X denotes any amino acid) (80, 81). Proline residues, like glutamine residues, are highly prevalent in gluten. Importantly, these residues prevent the complete digestion of gluten by intestinal enzymes. This explains how long gluten peptides that are good substrates for TG2 and can bind MHC molecules can be generated in the intestinal environment, in contrast to most other dietary proteins, which are readily fully digested (82).

The molecular basis for the association of CD with HLA-DQ2 and HLA-DQ8 is linked to the physicochemical properties of these MHC molecules. Both HLA-DQ2 and HLA-DQ8 molecules have positively charged pockets that have a preference for negatively charged peptides. HLA-DQ2 has a lysine at position β71, which confers its preference for binding peptides with negatively charged residues at positions P4, P6, and P7 (22). Both HLA-DQ2 and HLA-DQ8 are characterized by the lack of an aspartic acid at position β57 (83). This β57 polymorphism renders the P9 pocket of HLA-DQ8 basic, which explains why HLA-DQ8 has a preference for negatively charged residues at P9. Notably, the role of β57 polymorphism in HLA-DQ2 remains unclear. In addition, HLA-DQ8 has a preference for negatively charged residues at position P1, and therefore many of the HLA-DQ8-restricted gluten epitopes harbor two negatively charged glutamate residues, specifically in P1 and P9 (**Figure 4**). Overall, these observations exemplify how an enzyme present in a tissue environment can give an antigen improved binding to particular MHC class II molecules and promote pathogenic T cell responses.

Rheumatoid arthritis (RA) is another example of how posttranslational modifications by enzymes can promote T cell–mediated immune disorders by increasing the affinity of the causative antigen to the predisposing HLA molecules, in this case mainly HLA-DR4.1 (**Figure 4**) (reviewed in Reference 84). This HLA molecule has a basic P4 pocket that favors negatively charged or polar residues. This preference is well illustrated by the crystal structure of HLA-DR4.1 in complex with the 1168–1180 type II collagen peptide, which reveals aspartic acid in the P4 pocket of HLA-DR4.1 (85). No definite autoantigen has been identified in RA, but there are several candidates, including type II collagen, vimentin, and fibrinogen. The enzyme peptidylarginine deiminase can convert the positively charged guanidine group of arginine residues into the uncharged ureido group of citrulline residues through an enzymatic process known as citrullination, and citrullinated model peptides of candidate autoantigens bind with improved affinity to RA-associated HLA-DR molecules (**Figure 4**) (86).

LD: linkage disequilibrium

RA: rheumatoid arthritis

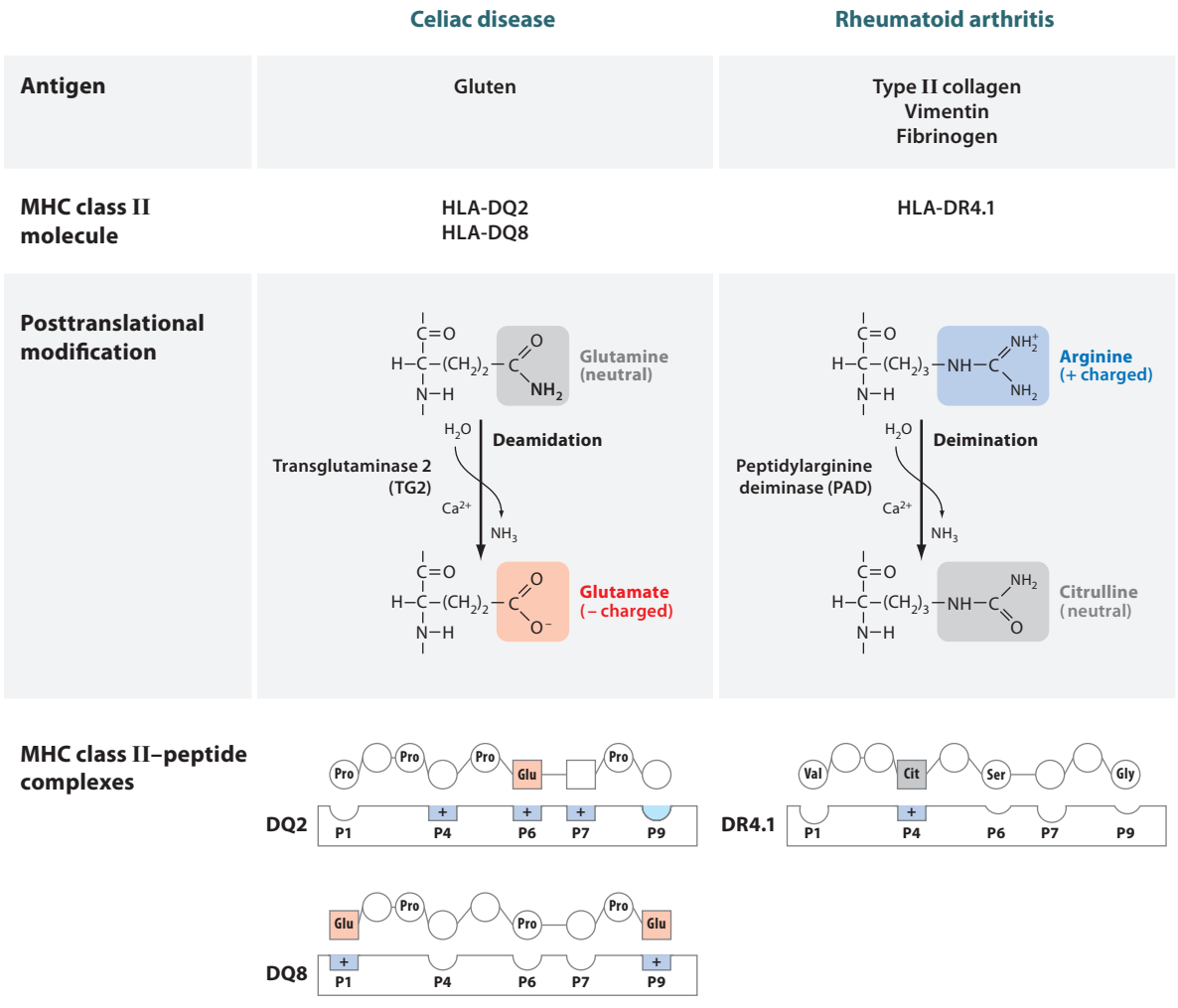


Figure 4

Posttranslational modification of antigens improves the binding of peptides to human leukocyte antigen (HLA) molecules in the context of celiac disease and rheumatoid arthritis. Gluten is a very good substrate for transglutaminase 2 (TG2), which converts glutamine residues to glutamate. This process, known as deamidation, generates peptides with negatively charged amino acid residues that bind with higher affinity to the disease-associated HLA-DQ2 or HLA-DQ8 molecules. P4, P6, and P7 pockets in HLA-DQ2 and P1 and P9 pockets in HLA-DQ8 have a preference for negatively charged anchor residues. (*Left*) Binding of a gluten peptide with glutamate in P6, and binding of a gluten peptide with glutamate residues in P1 and P9, to HLA-DQ2 and HLA-DQ8, respectively. (*Right*) In rheumatoid arthritis, the deimination of arginine to citrulline, also known as citrullination, is a posttranslational modification driven by peptidylarginine deiminase (PAD). This enzymatic conversion changes the positively charged arginine side chain to a neutral form that can be better accommodated in the P4 pocket of the HLA-DR4.1 molecule.

These observations pose the question of how these enzymes are induced and/or activated. In the case of CD, TG2 is highly expressed in the intestine but is not constitutively active (87). Studies in humanized HLA-DQ8 mice suggest that deamidation is not required for the initiation of the antigluten $CD4^+$ T cell response but that it plays a role in the

amplification of this response. Amplification of the antigluten immune response is achieved through the recruitment by HLA-DQ8 of cross-reactive TCRs that recognize native and deamidated peptides (24). The molecular basis for this process is that the polymorphism at position β57 enables HLA-DQ8 to switch from interaction with a negatively charged residue in the TCR to interaction with a negatively charged residue in the peptide. Therefore, not only can the antigluten immune response be initiated in the absence of TG2 activation, it can also trigger the activation of TG2. However, environmental factors, such as viral infections, may induce expression and activation of tissue enzymes by inducing inflammation and tissue damage. In particular, this may be the case for HLA-DQ2 individuals because, unlike HLA-DQ8-restricted T cells, HLA-DQ2-restricted T cells have an exquisite preference for deamidated peptides (88). In the case of RA, the presence of inflammatory cells increases expression of peptidylarginine deiminase (89). Finally, several recent studies suggest that posttranslational modifications may play a role in the pathogenesis of T1D, which is associated with HLA-DQ8 and HLA-DQ2 molecules (90). Interestingly, the β57 polymorphism in I-Ag7, the mouse homolog of HLA-DQ8, is required for the development of T1D (91, 92). The β57 polymorphism in I-Ag7, which is characteristic of NOD mice that spontaneously develop T1D, acts on the selection of autoreactive TCR repertoire in the same way that HLA-DQ8 acts on the selection of gluten-specific TCR (25). Future studies will determine whether cross-reactive, autoreactive TCR with a negative charge in the CDR3 plays a role in the pathogenesis of T1D.

The immunological basis for the HLA gene dosage effect is that there are threshold effects for disease development in which HLA-DQ expression and the available number of T cell–stimulatory gluten peptides are critical limiting factors (77). Homozygous individuals express more predisposing HLA-DQ2.5 and HLA-DQ8 molecules on the surface of their antigen-presenting cells and consequently can recruit a T cell response of larger magnitude. That the T cell response must reach a certain threshold to be pathogenic may also explain why another DQ2 variant (DQ2.2), encoded by the DQA1*0201 and DQB1*02 alleles of the DR7-DQ2 haplotype, is barely associated with CD on its own. The α-chain of DQ2.5 carries a tyrosine at position 22, which in contrast to the phenylalanine of DQ2.2 forms hydrogen bonds with the peptide main chain. Consequently, DQ2.5 forms more stable complexes with gluten peptides than does DQ2.2. This stability allows DQ2.5 to better retain gluten peptides for sustained presentation to T cells, thereby increasing the likelihood that the T cell response will reach the pathogenic threshold (88).

Taken together, epidemiological, genetic, and immunological studies suggest that associations with particular MHC molecules in CD and probably other tissue-specific autoimmune disorders are driven by the fact that the T cell response must achieve a certain threshold to be pathogenic, that is, to induce tissue damage. This threshold may be achieved by selecting for HLA molecules that allow for the most stable MHC-peptide complexes (DQ2.5 versus DQ2.2), increasing the number of HLA molecules (gene dosage effect), inducing posttranslational modifications that increase the affinity of the causative antigen to the HLA molecule (TG2 and deamidation), and recruiting distinct cross-reactive TCR repertoires that can recognize native and enzymatically modified antigens (e.g., the β57 polymorphism in HLA-DQ8 that allows it to act as a switch).

ROLE OF NON–HUMAN LEUKOCYTE ANTIGEN LOCI IN CELIAC DISEASE PATHOGENESIS

Susceptibility to CD has a strong genetic basis outside the HLA locus. This hypothesis is supported by the observation that siblings of CD patients (who share 50% of the their genome) have a 30-fold-higher risk of developing the disease than do individuals in the general population (10). More importantly, HLA-identical

GWAS: genome-wide association studies

SNPs: single-nucleotide polymorphisms

IBD: inflammatory bowel disease

siblings and dizygotic twins have concordance rates in disease outcome of 30% and 10%, respectively, whereas the concordance rate—approximately 75%—is extremely elevated in monozygotic twins (93). The recent introduction of low-cost, high-throughput genotyping platforms prompted researchers to interrogate the whole genome for genetic associations with CD. These so-called genome-wide association studies (GWAS) identified a large number of genes implicated in CD and other autoimmune diseases (94). Below, we discuss (*a*) how GWAS have helped decipher the relative contributions of HLA-linked and non-HLA-linked loci to CD susceptibility and (*b*) the immunological insights gained from these studies.

Recently, several GWAS have attempted to find non-HLA genomic regions associated with CD. To date, 40 such genomic regions harboring 64 candidate genes have been identified (**Table 1**) (95). These regions correspond to LD blocks that, in most cases, contain multiple genes. Thus, the single-nucleotide polymorphisms (SNPs) that have so far been associated with CD are termed tag SNPs of the risk haplotypes, but they themselves are not the causal variants associated with the disease. Interestingly, although the causative mutations have yet to be identified, 53% of the CD-associated SNPs are genetic variants for which different genotypes correlate with differences in expression levels of at least one physically close gene; these differences are referred to as *cis* expression quantitative trait loci (*cis* eQTL) SNPs. The number of *cis* eQTL SNPs observed among CD-associated SNPs is much larger than would be expected by chance (95), which suggests that some of the identified risk variants (or other SNPs linked to them) might influence CD susceptibility through a mechanism of altered gene expression rather than through changes at the protein-coding level.

The individual impact of each of these regions on disease susceptibility is small, and together these regions explain only ~5% of the genetic heritability (95). In contrast, the HLA locus alone accounts for 35% of the genetic heritability (96). Thus, although much progress has been made, approximately 50% of the genetic heritability remains to be explained. This missing heritability can be partially accounted for by the fact that the associations found for non-HLA loci are, at least in most cases, not with the actual causal variants associated with CD, which might lead to an underestimation of the impact of these loci in the pathogenesis of the disease. However, this finding alone is probably not sufficient to explain the missing heritability, and the most likely explanation is that many other common variants of small effects and/or highly penetrant rare mutations have yet to be identified. Alternatively, epistatic interactions between risk genes may occur. Epistasis has not yet been convincingly demonstrated in CD (95), but the reason might be that gene-gene interactions occur not between a single pair of genes but rather between groups of genes, which is very difficult to test.

The loci identified so far, however, provide important clues to the pathogenesis of and immunological pathways associated with CD. To gain insight into the biological nature of the candidate genes associated with CD, we considered functional annotation based on the Gene Ontology and Kyoto Encyclopedia of Genes and Genomes databases (**Figure 5**). The set of genes associated with CD appears to be remarkably enriched for immune genes, particularly in genes coding for chemokine receptor activity, cytokine binding, T cell activation, and lymphocyte differentiation; this finding supports the idea that CD is a T cell–mediated immune disorder. There is also enrichment for genes involved in stress pathways, innate immunity, and tumor necrosis factor receptor superfamily signaling. All these enrichments are also found for other autoimmune disorders and inflammatory bowel disease (IBD) (see section entitled Overlap of Genetic Pathways and Loci with Autoimmune and Inflammatory Diseases). However, interestingly, the NK cell–activation and interferon (IFN)-γ-production gene pathways appear to be selectively enriched in CD, which suggests that these pathways may be more important in CD than in other immune-mediated disorders (**Figure 5**).

Table 1 Celiac disease (CD) susceptibility loci

Loci associated with CD			Association with other autoimmune and/or inflammatory disorders[a,b]						
Locus	**Candidate gene(s) in the region**	**Odds ratio**	**RA**	**T1D**	**SLE**	**MS**	**PSO**	**UC**	**CrD**
1p31.3	*NFIA*	1.11							
1p36.11	*RUNX3*	1.12							
1p36.23	*PARK7, TNFRSF9*	1.14							
1p36.32	*TNFRSF14, MMEL1*	1.12	144						
1q24.2	*CD247*	1.1							
1q24.3	*FASLG, TNFSF18, TNFSF4*	1.1			145				
1q31.2	*RGS1*	1.25–1.39		146					
2p14	*PLEK*	1.14							
2p16.1	*REL, AHSA2*	1.15	147					148	
2q12.1	*IL18RAP, IL18R1, IL1RL1, IL1RL2*	1.19–1.28		146				149	149
2q31.3	*ITGA4, UBE2E3*	1.13							
2q33.2	*CTLA4, ICOS, CD28*	1.14	150	151	152, 153				
3p14.1	*FRMD4B*	1.19							
3p21.31	*CCR1, CCR2, CCRL2, CCR3, CCR5, CCR9*	1.21–1.3		146					
3p22.3	*CCR4*	1.13							
3q13.33	*CD80, KTELC1*	1.13							
3q25.33	*IL12A, SCHIP1*	1.35–1.36				154			
3q28	*LPP*	1.23–1.29	155						
4q27	*KIAA1109, ADAD1, IL2, IL21*	1.44–1.59	156	151					
6p21.32	*HLA-DQA1, HLA-DQB1*	6.23–7.04	157	158–160	145, 161	154, 162, 163		164	
6p25.3	*IRF4*	1.21							
6q15	*BACH2, MAP3K7*	1.13		151, 165					
6q22.33	*PTPRK, THEMIS*	1.17							
6q23.3	*TNFAIP3*	1.23	144, 166		145, 167		168		
6q25.3	*TAGAP*	1.16–1.21		146					
7p14.1	*ELMO1*	1.14							
10q22.3	*ZMIZ1*	1.12				154			
11q24.3	*ETS1*	1.21			145, 169				
12q24.12	*SH2B3, ATXN2*	1.2	155	151, 158					
14q24.1	*ZFP36L1*	1.12							
16p13.13	*CIITA, SOCS1, CLEC16A*	1.16		151, 159		170, 171		148	
18p11.21	*PTPN2*	1.17		151, 159					158, 172, 173
21q22.3	*ICOSLG*	1.14							172
22q11.21	*UBE2L3, YDJC*	1.13			145				
Xp22.2	*TLR7, TLR8*	1.14		174					

[a]Colored boxes indicate that the locus has also been associated with another inflammatory disorder or autoimmune disease. Blue boxes refer to associations found through genome-wide association studies, and green boxes refer to associations found through gene candidate approaches. For the latter, we considered only associations that have been replicated in at least two independent cohorts. References are provided inside the boxes.

[b]Genes located in the linkage disequilibrium block associated with CD. Abbreviations: RA, rheumatoid arthritis; T1D, type 1 diabetes; SLE, systemic lupus erythematosus; MS, multiple sclerosis; PSO, psoriasis; UC, ulcerative colitis; CrD, Crohn's disease.

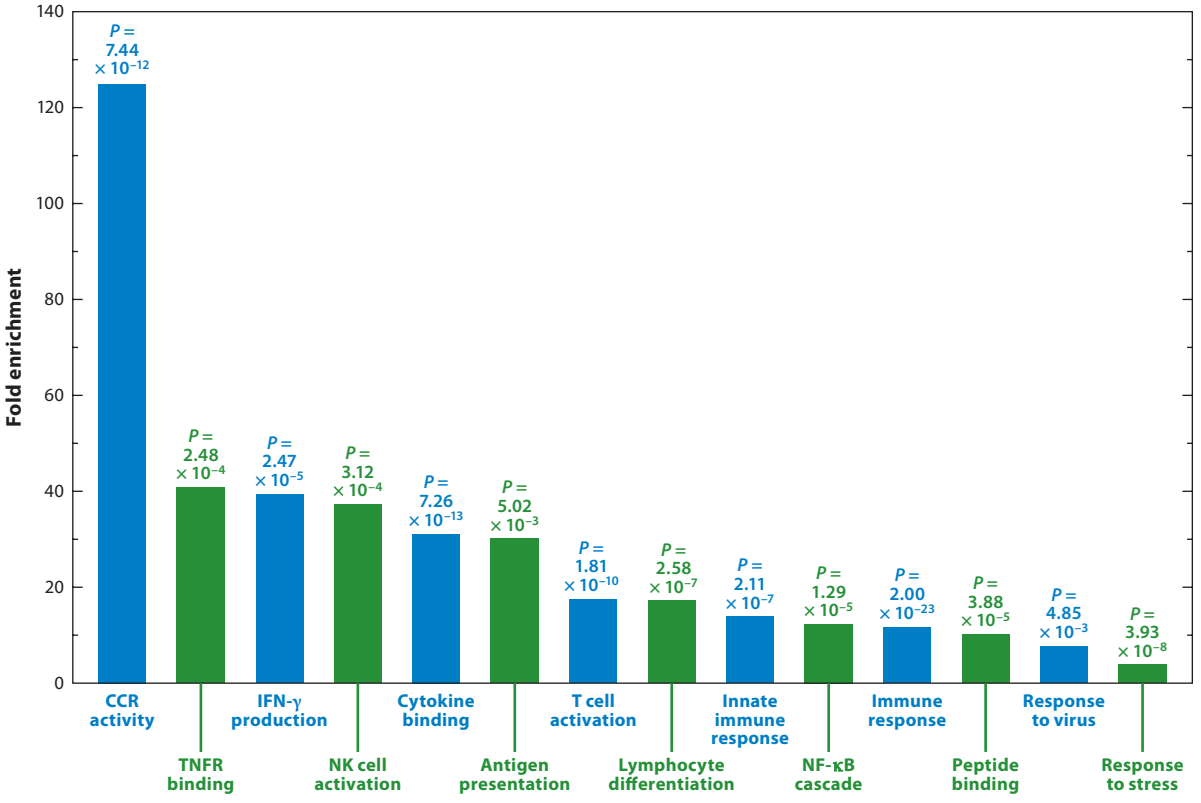


Figure 5

Gene ontology enrichment analysis for genes associated with celiac disease (CD). We used GeneTrail to test for an enrichment of functional annotations among genes associated with CD. Shown are the fold enrichments (*y* axis) observed for some of the most significantly enriched biological functions (*x* axis). Background expectations were based on all human genes. *P* values were calculated using a hypergeometric distribution, and we used the approach of Benjamini & Hochberg (143) to control the false discovery rate. Abbreviations: CCR, chemokine receptor; IFN, interferon; NF-κB, nuclear factor κB; NK, natural killer; TNFR, tumor necrosis factor receptor.

The 64 non-HLA genes (**Table 1** and **Figure 6**) identified to date can be classified according to where they exert their function in the immunological cascade, although some of them can act at several levels (e.g., *IL21*). Some genes, such as *REL*, which is part of the nuclear factor κB (NF-κB) signaling pathway, are implicated in numerous cell types and functions. Others play a role in the thymic differentiation of $CD4^+$ T cells (e.g., *THEMIS*) and CD8 T cells (e.g., *RUNX3*). Some are involved in immunological processes that take place in inductive sites such as the mesenteric lymph nodes, where they regulate T cell (e.g., *CD28* and *IL2*) and B cell (e.g., *ICOS* and *IL21*) activation and promote the differentiation of proinflammatory T cells (e.g., *IL12A*, *TLR7/TLR8*, *IRF4*, *IL1RL1*, and *IL18R1*). Finally, others are implicated in cell migration (e.g., different genes coding for chemokine receptors and *ITGA4*) and regulation of effector cell functions (e.g., *MAP3K7* and *IL21*, which are part of the c-Jun N-terminal kinase activation pathway that is critical for the function of activating NK receptors expressed by cytotoxic IELs).

Altogether, GWAS have identified a series of genes implicated in adaptive and innate

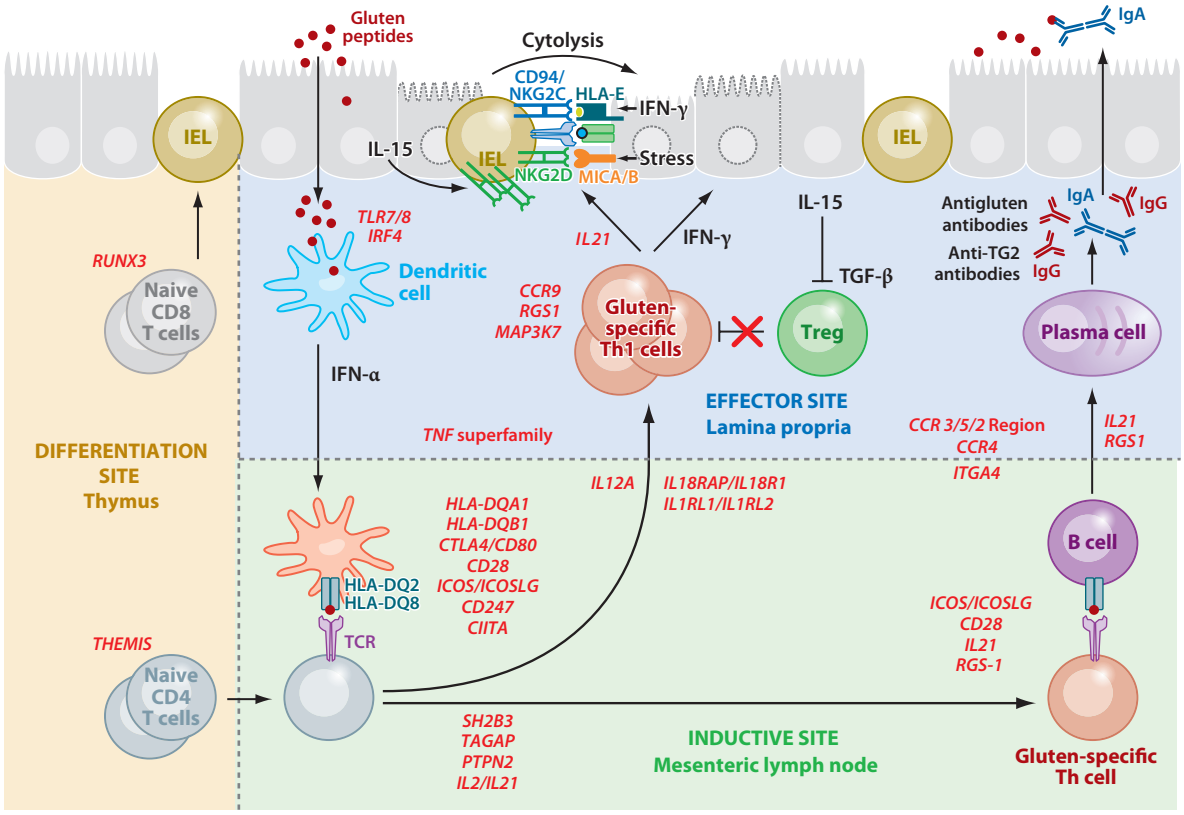


Figure 6

Integration of immunological pathways and celiac disease (CD)-associated genes into a model of CD pathogenesis. The figure is subdivided into three distinct anatomical regions in which T cell differentiation (thymus), T cell polarization (inductive site), and effector immune response (effector site) take place. Genes associated with CD by genome-wide association studies are listed in red according to their potential implication in distinct immunological pathways. *THEMIS* and *RUNX3* are involved in the thymic differentiation of CD4 and CD8 T cells, respectively. Dendritic cells located in the lamina propria acquire a proinflammatory phenotype upon viral recognition (*TLR7/8* and *IRF4*) and migrate to the mesenteric lymph nodes (inductive site). There, they present gluten peptides (*HLA-DQA1*, *HLA-DQB1*, and *CIITA*) to naive CD4 T cells and promote T cell activation (e.g., *CD28*, *CD80*, *CTLA4*, *CD247*, *PTPN2*, *SH2B3*, *TAGAP*, *IL2*, and *FASLG*) and differentiation into inflammatory effector T cells (*IL12A*, *IL18R1*, *IL18RAP*, *IL1RL1*, and *IL1RL2*). In addition, transglutaminase 2 (TG2) and gluten-specific B cells (that have internalized gluten-TG2 complexes) receive help from gluten-specific T cells, become activated, and differentiate into immunoglobulin (Ig)A- and IgG-producing plasma cells (*ICOS*, *ICOSLG*, *IL21*, and *RGS1*). Other genes regulate activation and migration of cytotoxic intraepithelial lymphocytes (IELs) (*MAP3K7*, *IL-21*, *CCR9*, and *RGS1*). Finally, some genes are involved in cell migration [e.g., genes coding for chemokine receptors (*CCR*s) and *ITGA4*], and others regulate tumor necrosis factor (TNF)-dependent pathways (*TNFAIP3*, *TNFSF4*, *TNFSF18*, *TNFRSF9*, and *TNFRSF14*). Even though their genes have not been identified by genetic studies, interleukin (IL)-15 and interferon (IFN)-α play a critical role in orchestrating the immune responses that lead to CD pathogenesis. IL-15 upregulates activating natural killer cell (NK) receptors and licenses IELs to kill epithelial cells, whereas IFN-α promotes the differentiation of proinflammatory dendritic cells. Abbreviations: HLA, human leukocyte antigen; TGF, transforming growth factor; Th, T helper cell.

immunity. As we discuss further below, unlike IBD, for which genetic studies have yielded some unexpected insights into the pathogenesis of the disease, almost all the genes identified in CD can be easily integrated into a model based on immunological studies using mainly cells from human intestinal biopsy samples.

HOW GENOME-WIDE ASSOCIATION STUDIES AND IMMUNOLOGICAL STUDIES CAN BE INTEGRATED INTO A MODEL OF CELIAC DISEASE PATHOGENESIS

Immunological Model of Celiac Disease Pathogenesis

Phenotypic and functional immunological studies in human suggest that both gluten-specific $CD4^+$ T cells and cytotoxic intraepithelial T lymphocytes play a key role in the development of CD, as defined by the presence of anti-TG2 antibodies and villous atrophy. The default immune response to an oral antigen in the intestinal environment, where transforming growth factor (TGF)-β and retinoic acid are abundant, is the induction of regulatory $Foxp3^+$ $CD4^+$ T cells that produce anti-inflammatory cytokines such as TGF-β and interleukin (IL)-10 (97–99). The induction of an inflammatory $CD4^+$ T cell response to gluten implies that dendritic cells in the intestinal mucosa of CD patients have a proinflammatory rather than tolerogenic phenotype. IFN-α, which is highly expressed in CD mucosa (100), may play a critical role in promoting the differentiation of proinflammatory dendritic cells. The role of IFN-α in CD pathogenesis is illustrated by the development of CD in hepatitis C patients treated with IFN-α (101), as well as by the increased prevalence of CD among Down syndrome patients (102). Indeed, chromosome 21 harbors the IFN-α receptor, which explains why cells of Down syndrome patients show increased levels of IFN-α receptor expression and a greater response to type 1 IFNs (103).

Gluten-specific $CD4^+$ T cells are central to all aspects of CD pathogenesis. They probably assist in the induction of anti-TG2 antibodies by providing help to anti-TG2 B cells (104). This hypothesis is based on the observation that anti-TG2 antibodies develop only in HLA-DQ2 or HLA-DQ8 individuals (105) and recede when gluten is excluded from the diet (106, 107). Gluten may form complexes with TG2, which are internalized by TG2-specific B cells. Such B cells can therefore present gluten peptides at their surface in the context of HLA-DQ2 or HLA-DQ8 molecules and can receive help from antigluten $CD4^+$ T cells to differentiate into IgA and IgG anti-TG2 plasma cells (104). However, the role of anti-TG2 and antigluten antibodies in the development of the celiac lesion remains to be defined. They may amplify the inflammatory immune response to gluten by increasing gluten uptake (108) and by inducing the activation of Fc receptors expressed on granulocytes. Gluten-specific $CD4^+$ T cells also play a role in tissue remodeling via the production of IFN-γ and metalloproteinases (109). However, this role is not sufficient to induce villous atrophy. It is thought that epithelial damage is mediated by cytotoxic IELs that express activating NK cell receptors, which recognize stress- and inflammation-induced ligands on intestinal epithelial cells (47). IL-15 upregulates the activating NKG2D receptor and confers NK-like properties—namely the ability to kill in a TCR-independent manner (43, 44, 110)—to IELs. Whether IFN-α, which promotes NK cell activity, also plays a role in the activation of IELs remains to be assessed. It is very likely that gluten-specific $CD4^+$ T cells, which produce IL-21 (100) and IFN-γ (111), also play a role in the activation of IELs. They may do so by upregulating inflammatory ligands on epithelial cells [e.g., IFN-γ promotes upregulation of the nonclassical MHC class I molecule HLA-E on epithelial cells (112)] and by promoting cytolytic activity in IELs [e.g., IL-21 promotes NK-like activity in cytotoxic T lymphocytes (113, 114)]. Refractory sprue is an extreme case in which the presence of gluten-specific $CD4^+$ T cells is no longer required for villous atrophy. This severe complication of CD cannot be treated by gluten withdrawal (115, 116). It is characterized by the presence of IELs that have acquired an inherent and aberrant highly activated NK-like phenotype (115) that is promoted and maintained by high levels of IL-15 expression in the epithelium (117, 118). Refractory sprue is mimicked in an IL-15 transgenic mouse model, in which IL-15 has

been modified such that it is secreted in an uncontrolled manner because it does not require its private IL-15Rα receptor to be expressed on the cell surface (119). In addition, it plays a role in licensing cytotoxic IELs to become effective killer cells and prevents TGF-β and regulatory $Foxp3^+$ T cells from blocking inflammatory effector responses (120, 121).

Overall, the value of the model of CD pathogenesis presented in **Figure 6** resides in its foundation on human studies. However, as in all models based on functional studies in humans, it is based more on correlations than on the demonstration of cause-effect relationships. Therefore, it is important to examine this model in view of the susceptibility genes identified by GWAS.

Interactions Between Key Immunological Markers of Celiac Disease and Susceptibility Genes

On the basis of human studies suggesting that IL-15 (42, 117, 122) and IFN-α (100) are significantly increased in the celiac mucosa and are central to CD pathogenesis (**Figure 6**) (44, 110, 117, 121), we might expect GWAS to identify mutations in the coding or regulatory regions of the genes encoding IFN-α and IL-15. Intriguingly, however, no genetic associations with CD have been found for the genes encoding IL-15 or IFN-α. The lack of association with these genes suggests that the increased levels of these cytokines in CD patients might be the by-product of the deregulation of genes that can modulate the levels of these cytokines—that is, *trans* effects. We therefore used the STRING database (123) to look for known functional interactions among CD susceptibility genes as well as between CD susceptibility genes and IL-15 or IFN-α (**Figure 7**). Our results show that 40 out of the 64 candidate genes associated with CD have a functional connection with one or more other CD genes. The 40 genes that are part of this CD susceptibility functional network (out of the 64 reported in **Table 1**) probably represent the best candidates to harbor the causative associations with CD. In addition, we noticed that several of the genes in this network have a direct association with IL-15, IFN-α, or both (**Figure 7**). This observation supports the hypothesis that the increased levels of IL-15 and/or IFN-α observed in CD patients probably result from functional variation in this network. For example, functional variation that increases the responsiveness of the transcription factor *REL*, a member of the NF-κB complex, could ultimately lead to increased levels of IL-15, as this gene is regulated by NF-κB (124, 125). However, increased signaling via Toll-like receptor (TLR)7 or TLR8, which are innate receptors involved in the detection of viral infection, would lead to increased IFN-α production.

The results from the network analysis may also explain the heterogeneity of the cellular phenotypes observed in CD patients. Indeed, this analysis demonstrates that, depending on the combination of genetic susceptibility markers present in each patient, a patient could have increased levels of IL-15, IFN-α, or both. In agreement with this hypothesis, preliminary analysis of 21 active CD patients shows that CD patients can be divided into IL-15 high expressers, IFN-α high expressers, and IL-15/IFN-α high expressers (B. Jabri, unpublished data). If confirmed, this observation would suggest that CD patients do not constitute a homogeneous group and that different immune pathways may lead to dysregulated inflammatory antigluten immunity and activation of IELs. To better delineate which genetic markers account for increased levels of each of these cytokines in CD patients, association studies could be performed to group CD patients on the basis of their inflammatory phenotypes—that is, according to whether they have high IL-15 and/or high-IFN-α expression.

OVERLAP OF GENETIC PATHWAYS AND LOCI WITH AUTOIMMUNE AND INFLAMMATORY DISEASES

Epidemiological data suggest that CD is more associated with autoimmune disorders, in

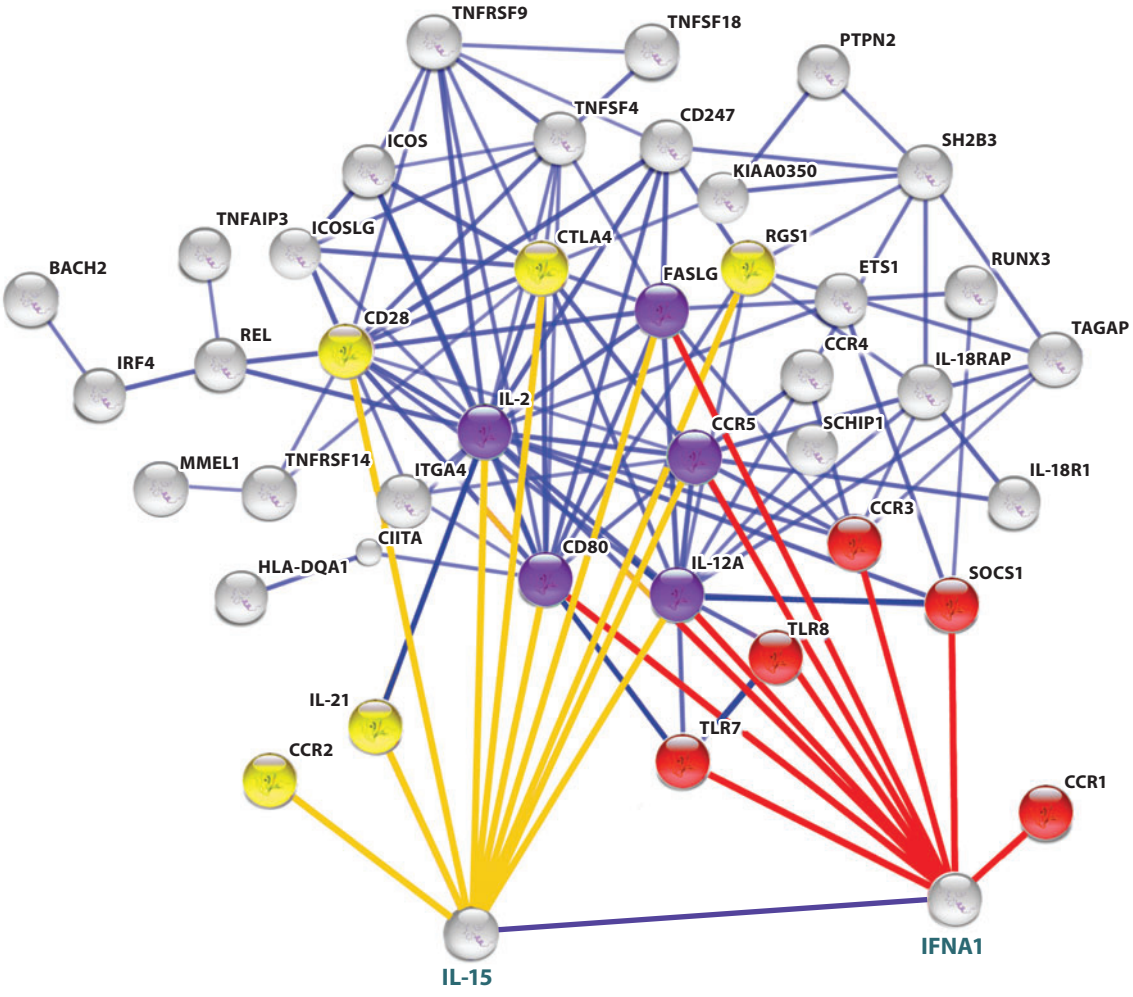


Figure 7

Network of known functional interactions between celiac disease (CD)–associated genes and key immunological markers of CD. We used the STRING database to look for known functional interactions among CD susceptibility genes, as well as functional interactions between CD susceptibility genes and interleukin (IL)-15 or interferon (IFN)-α. The STRING database assembles information about both known and predicted protein-protein interactions on the basis of numerous sources, including experimental repositories, computational prediction methods, and public text collections. Several CD susceptibility genes functionally interact with IL-15 (*yellow*), IFN-α (*red*), or both (*purple*).

particular T1D and autoimmune thyroiditis (126, 127), than with IBD, which comprises Crohn's disease and ulcerative colitis (128, 129). In accordance, when we looked for which pathways were enriched among CD susceptibility genes, the strongest enrichments were observed among disease pathways associated with T cell–mediated organ-specific autoimmune diseases such as T1D [false discovery rate (FDR) = 1.13×10^{-07}] and autoimmune thyroiditis (FDR = 3.24×10^{-07}), but not IBD. This observation is plausible, given what we know about the immunological mechanisms underlying these diseases. In particular, tissue

destruction in Crohn's disease is not cell specific and is thought to be mediated by general inflammatory effector mechanisms involving macrophages and neutrophils, whereas tissue destruction in autoimmune disorders and CD is mediated primarily by HLA-restricted T cells and cytotoxic T cells that target specific tissue cells, specifically intestinal epithelial cells in CD (130).

That CD susceptibility genes are enriched among genes known to be involved in other autoimmune pathways predicts that the genetic risk factors associated with CD also represent risk factors for other autoimmune disorders. To test this hypothesis, we compiled a list of all the regions identified by GWAS as associated with CD (**Table 1**), autoimmune diseases, and inflammatory disorders (**Figure 8**). Next,

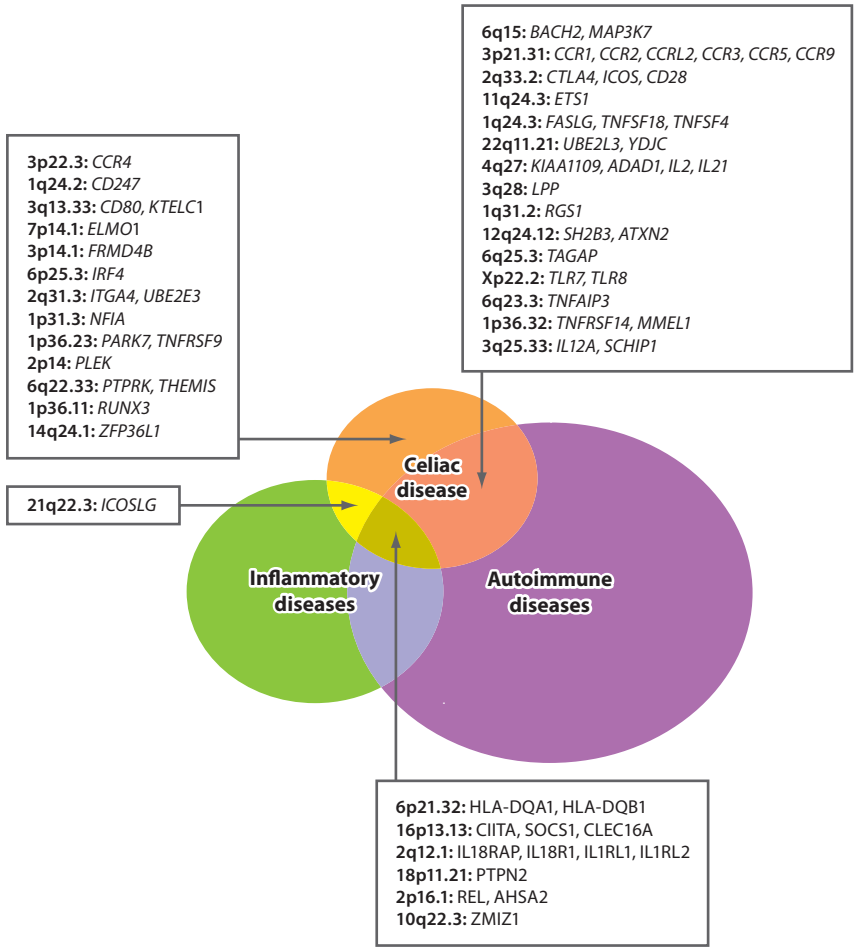


Figure 8

Overlap between celiac disease (CD) genetic risk factors and genetic risk factors identified for other autoimmune and inflammatory diseases. Shown are the overlaps between the regions identified by genome-wide association studies (GWAS) as associated with CD and the regions identified by GWAS as associated with autoimmune diseases or inflammatory disorders. The set of autoimmune diseases includes rheumatoid arthritis, systemic lupus erythematosus, type 1 diabetes, multiple sclerosis, and psoriasis. The set of inflammatory disorders includes Crohn's disease and ulcerative colitis.

we looked for overlaps between CD-associated genomic regions and those associated with at least one autoimmune or inflammatory disease (**Figure 8**). As anticipated, we observed a significantly higher overlap between CD and autoimmune disorders (12%) than between CD and inflammatory diseases (2%) (Fisher exact test; P value $= 0.025$) (**Figure 8**). T1D showed the strongest overlap with CD. This finding is well illustrated by the fact that 35% of the CD-associated genomic regions also impact susceptibility to T1D (**Table 1**). Notably, T1D and CD are the only diseases that share genes involved in immune responses against viral detection, in line with the hypothesis that viral infections may influence development of these diseases. Overall, the genes found to be common to CD and autoimmune disorders were implicated in cytokine and chemokine signaling and, importantly, T cell activation (FDR $\leq 5 \times 10^{-08}$) (**Figure 8**). In contrast, the genes found to be common to CD, autoimmune disorders, and IBD were more generally involved in immune activation; these genes include those that code for signaling molecules (e.g., *REL* and *PTPN2*) (**Figure 8**). Curiously, the genomic region that encodes for genes involved in the inflammasome pathway (*IL18RAP*, *IL18R1*, *IL1RL1*, and *IL1RL2*) affects susceptibility to CD, autoimmune disorders, and IBD (**Figure 8**). This finding is interesting in light of the current idea that activation of the inflammasome is important not only for the control of microbial infections but also to signal the presence of endogenous tissue stress and thereby enhance inflammatory immune responses (131). This analysis also allowed us to define genomic regions specifically associated with CD (**Figure 8**). These regions are particularly interesting in that they might help elucidate which immunological pathways are unique hallmarks of CD. The genomic regions selectively associated with CD are enriched for genes related to central and peripheral T cell differentiation (FDR $= 2 \times 10^{-3}$), which suggests that the pathogenic T cell response observed in CD has unique features and again stresses the central role of T cells in CD pathogenesis.

THE EVOLUTIONARY HISTORY OF CELIAC DISEASE–ASSOCIATED SUSCEPTIBILITY GENES

Even though there is incomplete knowledge of the worldwide prevalence of CD and great variance in the consumption of cereals across populations (**Figure 1*a*,*b***), there nonetheless appear to be differences among ethnic groups in terms of susceptibility for CD. The disease appears to be particularly common among Caucasians. Two alternative hypotheses could account for this fact. First, the elevated prevalence of CD in certain populations could result from the increase in frequency of CD susceptibility alleles by genetic drift (i.e., random chance). Second, CD susceptibility alleles could have increased in frequency as a result of positive selection if they confer a selective advantage to the carriers. To test these two hypotheses, we used evolutionary genetic tools to search for molecular signatures of positive selection on the genes associated with CD (**Figure 9**).

One of the most striking signatures of positive selection is an increase in the strength of LD associated with the selected allele (132, 133). Indeed, when an allele is targeted by positive selection, the beneficial allele increases in frequency in the population at a much faster rate than that of a neutrally evolving allele, and as a consequence, the haplotypes carrying the advantageous allele are longer relative to haplotypes that rise to similar frequencies solely by random genetic drift (132, 133). We used the integrated haplotype similarity test (134) to search for this molecular signature of positive selection among all the CD-associated genomic regions (**Table 1**). Four out of the 40 CD-associated regions—namely the *IL18RAP*, *IL12A*, *IL2/IL21*, and *SH2B3* loci (**Figure 9**)—show a strong signature of positive selection. The proportion of CD-associated loci showing signatures of positive selection is higher than would be expected by chance ($P = 0.04$, if one randomly samples 40 regions of the genome and tests for evidence of selection). Curiously, for all loci, with the exception of the *IL2/IL21* locus, the allele/haplotype that shows evidence

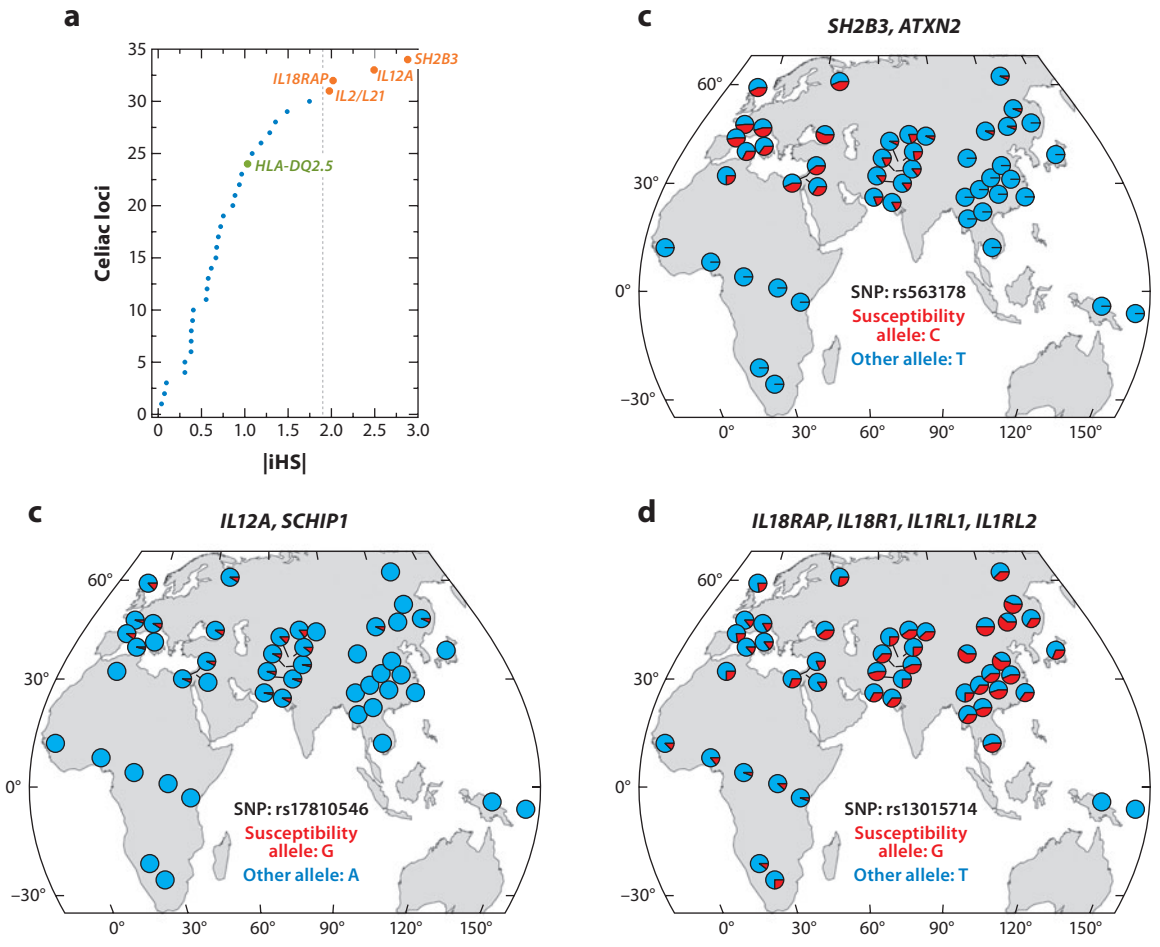


Figure 9

Signatures of positive selection on celiac disease (CD) susceptibility alleles. (*a*) |iHS| (absolute value of integrated haplotype similarity score) values for single-nucleotide polymorphisms (SNPs) tagging genomic regions associated with CD. For this analysis, we used the European HapMap phase II SNPs (the population consisted of individuals of European descent from Utah) because all genome-wide association studies for CD have been performed in populations of European descent. The dashed line represents the ninety-fifth percentile of the genome-wide |iHS| distribution for the European samples from HapMap. |iHS| above the ninety-fifth percentile (*red dots*) are therefore suggestive of positive selection. (*b*) Worldwide frequency distribution of the SNP linking the *SH2B3/ATXN2* locus with susceptibility to CD. (*c*) Worldwide frequency distribution of the SNP linking the *IL12A/SCHIP1* locus with susceptibility to CD. (*d*) Worldwide frequency distribution of the SNP linking the *IL18RAP/IL18R1/IL1RL1/IL1RL2* locus with susceptibility to CD. The red fraction of the pie charts in panels *b–d* represents, for the corresponding SNP, the frequency of the CD susceptibility allele, whereas the blue fraction represents the frequency of the protective allele in different areas of the world.

of positive selection is the one associated with increased susceptibility to CD.

The latter observation can be easily explained if having CD is associated with some sort of selective advantage, which would bypass the negative effects associated with the disease. For example, increased levels of IL-15 or IFN-α and/or the absence of villi in CD patients could confer protection against intestinal infections that lead to death in young children. Although this is an interesting possibility, our evolutionary results do not fully support these

hypotheses. Indeed, if having CD were an advantageous phenotype, the strongest signatures of selection would be associated with the DQ2.5 haplotype(s) because they explain most of the genetic variance associated with CD. However, our results do not provide strong evidence for the action of positive selection on the DQ2.5 haplotype(s). The lack of signal of selection could be due simply to the low power of the integrated haplotype similarity test (as all other neutrality tests) to detect selection in high-recombining regions such as the MHC region. Yet, the fact that we observed strong signatures of selection (using the same test) for the HLA-DRB1 haplotypes associated with RA, multiple sclerosis, and systemic lupus erythematosus appears to disfavor such a hypothesis. Moreover, the four loci identified as targeted by positive selection are not specific to CD. Indeed, these same loci also represent risk factors for other autoimmune and/or inflammatory disorders, such as T1D, ulcerative colitis, and Crohn's disease, among others (**Table 1**).

Thus, as previously suggested (135, 136), these CD-risk alleles were positively selected probably because they confer increased resistance to past or present infectious agents. Studies on the functional role of the *SH2B3* risk allele (an amino acid–altering mutation) strongly support this hypothesis (136). Indeed, stimulation of peripheral-blood mononuclear cells with MDP, a specific ligand of the pattern-recognition receptor NOD2, shows that cells isolated from individuals homozygous for the *SH2B3* CD-risk allele display an increased production of proinflammatory cytokines, such as IL-1β, IL-6, and IL-8, compared with homozygous or heterozygous individuals for the other, nonrisk allele (136). Thus, individuals homozygous for the *SH2B3* allele probably enjoy increased protection against certain infectious agents because they can induce stronger proinflammatory responses, but at the cost of increased susceptibility to autoimmune or inflammatory disorders.

Less intuitive, at least in the context of CD, is the functional role described for the positively selected risk allele in the *IL18RAP* locus. Indeed, carriers homozygous for the *IL18RAP* risk allele have a significantly lower level of IL-18RAP expression (at the messenger RNA level) (96, 136). This finding suggests that individuals at risk of CD show reduced signaling in response to IL-18 and that they generate less IFN-γ. This observation is surprising, given the well-described intestinal inflammation and high mucosal IFN-γ levels observed in CD (111). Interestingly, this same risk allele has also been associated with different isoforms of IL-18RAP. Individuals who are homozygous or heterozygous for the *IL18RAP* risk allele have increased amounts of a short form of IL-18RAP (37 kDa versus 70 kDa for the longer isoform) compared with individuals who are homozygous for the other allele (137). Although the function of this short isoform remains to be determined, it may increase IL-18-induced signaling, an explanation that would be more compatible with our current knowledge of the pathogenesis of CD.

Altogether, these data suggest that the high prevalence of CD in modern societies is at least partially the by-product of past selection for increased immune responses to combat pathogens. The massive increase in human population sizes and the exposure to new zoonoses after the development of agriculture and following the domestication of animals may have resulted in the spreading of new infectious diseases (138), which in turn may have promoted the selection of genetic polymorphisms that increase predisposition to CD. Interestingly, the susceptibility alleles targeted by positive selection are absent or are found at very low frequencies among African populations, in which agriculture was introduced more recently, whereas these alleles attain considerable frequencies in Europe and Asia (**Figure 9**). Although speculative, this observation might explain why the prevalence of CD is higher in European Americans than it is in African Americans (139, 140).

GENERAL PERSPECTIVES

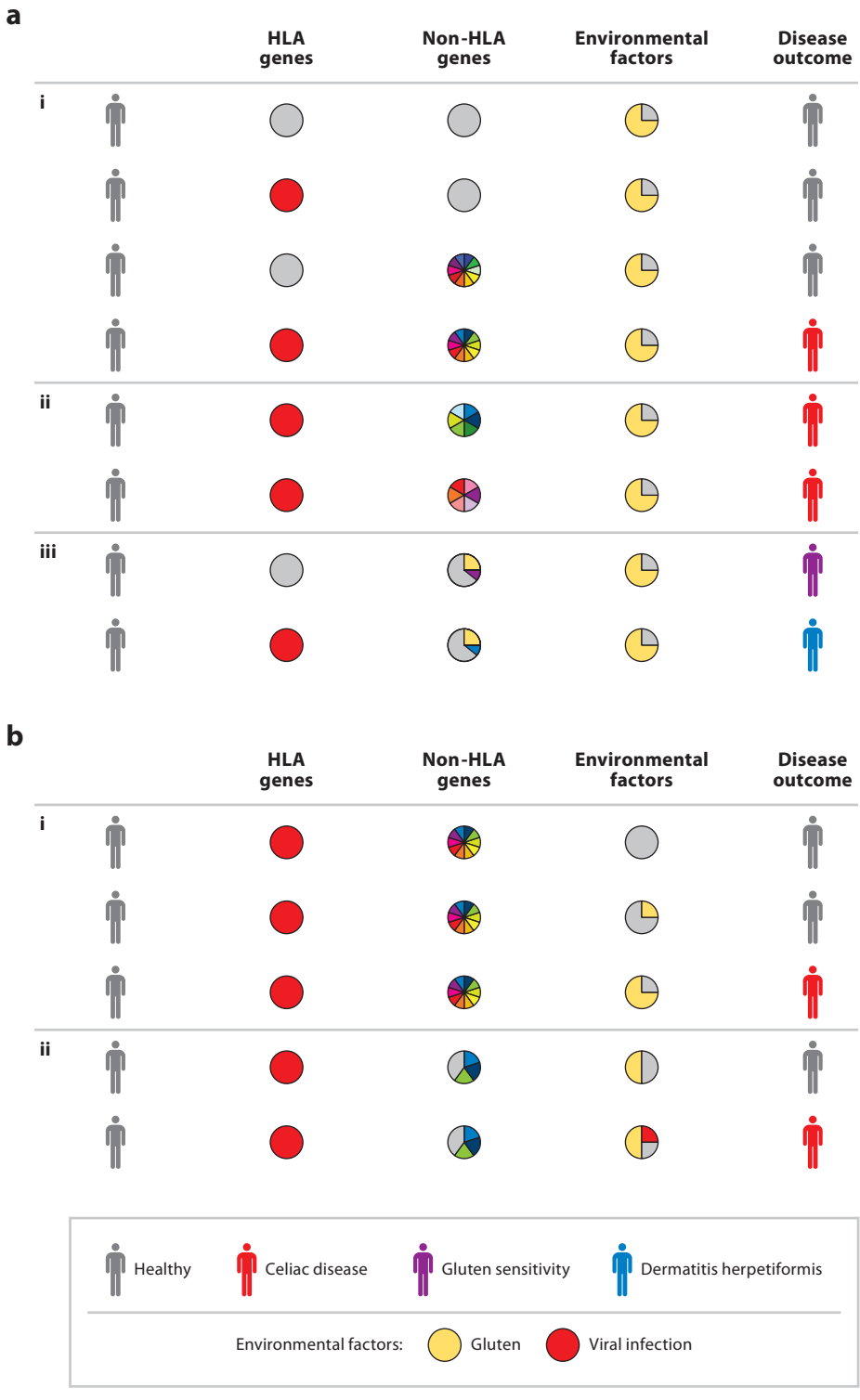
Data from genetic, immunological, and epidemiological studies converge to suggest that

CD is primarily a T cell–mediated immune disorder induced by dietary gluten, in which $CD4^+$ T cells and MHC class II molecules play a central role (**Figure 10*a*,*b***). In particular, typical CD, as defined by the presence of villous atrophy and anti-TG2 antibodies, is found only in patients with HLA-DQ2- or HLA-DQ8-restricted antigluten $CD4^+$ T cells (**Figure 10*a*,*b***). However, numerous observations in human and mouse also suggest that adaptive antigluten $CD4^+$ T cell immunity is not sufficient for the development of villous atrophy and that other cell types are required for the induction of tissue damage. For instance, the role of cytotoxic IELs (i.e., $CD8^+$ T cells) in CD is supported by the observation that their expansion and activation correlate with the presence of villous atrophy. In accordance with this hypothesis, polymorphisms in genes involved in the differentiation (*RUNX3*) and migration (*CCR9*) to the epithelium of cytotoxic $CD8^+$ T cells confer susceptibility to CD. In addition, immunological studies suggest that these intraepithelial cytotoxic $CD8^+$ T cells mediate the destruction of stressed epithelial cells by acquiring an NK-like phenotype. The role of NK cell–like–mediated responses in CD pathogenesis is further supported by GWAS showing that the genomic regions impacting susceptibility to CD have an approximately 40-fold enrichment for genes involved in NK cell activation.

Accumulating evidence from genetic and epidemiological studies suggests that viral infections might be an important triggering factor of CD. On one hand, high levels of IFN-α expression were reported in the intestinal mucosa of CD patients (100), and recurring rotavirus infections were found to increase the incidence of CD (141). On the other hand, GWAS identified viral response–associated genes such as *TLR7*, *TLR8*, and *IRF4* as risk factors for CD. Altogether, these observations suggest that repeated viral infections might constitute a risk factor for CD, particularly among patients with polymorphisms in viral response genes.

Although these findings illustrate how genetic risk factors and environmental factors can synergize to lead to increased development of disease, particular exogenous factors may promote CD by compensating for the lack of certain susceptibility genes (other than HLA) (**Figure 10*b***). As an extreme example, IFN-α treatment in hepatitis C patients induced CD in patients bearing the CD-associated HLA. Conversely, given the right genetic makeup, it may be possible to reach the same outcome without the need for additional environmental hits outside of gluten consumption (**Figure 10*a***). One can therefore imagine a spectrum of disease susceptibility: On one end are individuals with a large number of genetic susceptibility markers for CD and limited need for environmental hits, and on the other end are individuals with a limited number of genetic risk factors (e.g., the correct HLA genes but a limited number of non-HLA genes) who require multiple environmental hits to develop disease. The former group of patients may get CD as soon as gluten is introduced into the diet, whereas the latter group may never develop the disease or may develop it late in life. How gluten influences disease development can vary depending on the amount of gluten and when it is introduced into the infant diet (142), which further suggests a complex interplay between genes and environment.

Similar levels of genetic susceptibility to CD may be attributable to distinct sets of non-HLA genes, as different genetic pathways may lead to the same immunological outcome (**Figure 10*a***). This is well illustrated by the genetic network showing how different gene combinations lead to expression of IFN-α, IL-15, or both (**Figure 8**). Both cytokines play a critical role in the induction of inflammatory T cell responses, and both promote NK cell activity in cytotoxic $CD8^+$ T cells. Interestingly, some preliminary evidence suggests that CD patients could be subdivided into patients who express only one of the cytokines, patients who express both, and patients who express neither. Thus, similar effector immune responses and the same disease outcome can be achieved in many ways, which supports the idea of genetic heterogeneity among CD patients.



Furthermore, there may be an even larger genetic heterogeneity if one expands the classical definition of CD by including gluten-mediated intestinal disorders without intestinal damage and by including diseases that are associated only with extraintestinal manifestations. For example, cases of patients with dermatitis herpetiformis and no villous atrophy have been reported (**Figure 10*a***). Conversely, some patients who lack HLA-DQ2 or HLA-DQ8 molecule may still have an intestinal epithelial stress response leading to clinical symptoms associated with irritable bowel syndrome (**Figure 10*a***). Genetic studies are now required to unravel the genetic idiosyncrasies associated with these different manifestations of CD-like disorders.

CONCLUSION

GWAS have identified a fair number of genomic regions associated with CD, but much more work remains to be done before we know how genetic variation in these regions impacts immunological (or other) phenotypes. Resolving these issues will not be easy, given the limited possibilities for further genetic mapping in regions with strong LD as well as the tremendous challenge of linking mutations with altered function in complex biological systems. Despite these challenges, the general findings from GWAS studies are in exquisite agreement with existing immunological models. The remarkable concordance between genetic and immunological observations encourages further efforts to harness the presently established pathogenic players of CD to develop alternative therapies and effective prevention. On the genetic side, future studies should aim to identify rare gene mutations with high disease penetrance, which could point to novel molecular targets that would be particularly effective for therapeutic intervention. In addition, it will be of particular interest to unravel the function of present-day uncharacterized genes or gene-desert genomic regions that show consistent associations with CD. Such studies might provide us with important clues about unexpected biological pathways implicated in CD pathogenesis. Finally, we should take advantage of the recent development of several technologies (for example, expression microarrays, RNA sequencing, and mass spectrometry) that allow assessment of the levels of interindividual phenotypic variation at the genome-wide level. For example, it would be interesting to characterize genome-wide transcriptional signatures (i.e., expression levels) that are associated with different forms and stages of the disease. These molecular signatures could be used as prognostic tools, but they could also illuminate the specific immunological mechanisms associated with specific forms of the disease.

Figure 10

Role of genetic factors in celiac disease (CD) development. Various scenarios for the interplay between human leukocyte antigen (HLA) and non-HLA genes and environmental factors. (*a*) HLA and non-HLA genes contribute to CD development under similar environmental pressures. (*i*) HLA genes are necessary but not sufficient for the development of CD. (*ii*) Different combinations of non-HLA genetic risk factors can lead to CD in individuals who carry the predisposing HLA molecules. (*iii*) Nonclassical gluten-induced pathologies. Patients who lack the predisposing HLA molecules but carry particular non-HLA risk factors may develop irritable bowel syndrome–like disorders in response to gluten (*purple*). Conversely, patients with predisposing HLA molecules and other non-HLA risk factors may develop dermatitis herpetiformis (*blue*) in the absence of intestinal manifestations. (*b*) Impact of environmental factors on CD development in HLA-DQ2- or HLA-DQ8-carrying individuals. (*i*) Quantitative differences in gluten-intake influence on CD development. (*ii*) Viral infection and/or other environmental factors promote CD in individuals with low non-HLA risk factors who otherwise would not develop CD.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank CD patients and their family members, as well as the University of Chicago Celiac Disease Center, for their support of our research. We are grateful to F.G. Sperone for generating the geographic information system maps shown in **Figure 1**. We thank H. Fehlner-Peach for critical reading of the manuscript. We thank C. Catassi and D.D. Kasarda for helpful discussions. This work was supported by National Institutes of Health grants R01 67180 and R01 58727 to B.J. and by grants from the Research Council of Norway, the European Commission FP7 program (grant MRTN-CT-2006–036032), the South-Eastern Norway Regional Health Authority, the Juvenile Diabetes Research Foundation, and the Norwegian Foundation for Health and Rehabilitation to L.M.S. V.A. is supported by the Crohn's and Colitis Foundation. L.B.B. is supported by the Human Frontiers Science foundation (grant LT6572009-L).

LITERATURE CITED

1. Simoons FJ. 1981. Celiac disease as a geographic problem. In *Food, Nutrition and Evolution: Food as an Environmental Factor in the Genesis of Human Variability*, ed. DN Walcher, N Kretchmer, pp. 179–99. New York: Masson
2. Losowsky MS. 2008. A history of coeliac disease. *Dig. Dis.* 26:112–20
3. Adams F. 1956. *The Extant Works of Aretaeus, the Cappodocian*. London: Sydenham Soc.
4. Gee S. 1888. On the coeliac affection. *St. Bartholomew's Hosp. Rep.* 24:17–20
5. Dicke WK, Weijers HA, Van De Kamer JH. 1953. Coeliac disease. II. The presence in wheat of a factor having a deleterious effect in cases of coeliac disease. *Acta Paediatr.* 42:34–42
6. Paulley JW. 1954. Observation on the aetiology of idiopathic steatorrhoea: jejunal and lymph-node biopsies. *Br. Med. J.* 2:1318–21
7. Marsh MN. 1992. Mucosal pathology in gluten sensitivity. In *Coeliac Disease*, ed. M Marsh, pp. 136–91. Oxford, UK: Blackwell Sci.
8. Risch N. 1987. Assessing the role of HLA-linked and -unlinked determinants of disease. *Am. J. Hum. Genet.* 40:1–14
9. Petronzelli F, Bonamico M, Ferrante P, Grillo R, Mora B, et al. 1997. Genetic contribution of the HLA region to the familial clustering of coeliac disease. *Ann. Hum. Genet.* 61:307–17
10. Bevan S, Popat S, Braegger CP, Busch A, O'Donoghue D, et al. 1999. Contribution of the MHC region to the familial risk of coeliac disease. *J. Med. Genet.* 36:687–90
11. Greco L, Romino R, Coto I, Di Cosmo N, Percopo S, et al. 2002. The first large population based twin study of coeliac disease. *Gut* 50:624–28
12. Falchuk ZM, Rogentine GN, Strober W. 1972. Predominance of histocompatibility antigen HLA-8 in patients with gluten-sensitive enteropathy. *J. Clin. Investig.* 51:1602–5
13. Stokes PL, Asquith P, Holmes GK, Mackintosh P, Cooke WT. 1972. Histocompatibility antigens associated with adult coeliac disease. *Lancet* 2:162–64
14. Keuning JJ, Pena AS, van Leeuwen A, van Hooff JP, va Rood JJ. 1976. HLA-DW3 associated with coeliac disease. *Lancet* 1:506–8
15. Ek J, Albrechtsen D, Solheim BG, Thorsby E. 1978. Strong association between the HLA-Dw3-related B cell alloantigen -DRw3 and coeliac disease. *Scand. J. Gastroenterol.* 13:229–33
16. Tosi R, Vismara D, Tanigaki N, Ferrara GB, Cicimarra F, et al. 1983. Evidence that celiac disease is primarily associated with a DC locus allelic specificity. *Clin. Immunol. Immunopathol.* 28:395–404
17. Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E. 1989. Evidence for a primary association of celiac disease to a particular HLA-DQ α/β heterodimer. *J. Exp. Med.* 169:345–50

18. Spurkland A, Sollid LM, Polanco I, Vartdal F, Thorsby E. 1992. HLA-DR and -DQ genotypes of celiac disease patients serologically typed to be non-DR3 or non-DR5/7. *Hum. Immunol.* 35:188–92
19. van de Wal Y, Kooy YM, van Veelen PA, Pena SA, Mearin LM, et al. 1998. Small intestinal T cells of celiac disease patients recognize a natural pepsin fragment of gliadin. *Proc. Natl. Acad. Sci. USA* 95:10050–54
20. Sjöström H, Lundin KE, Molberg O, Korner R, McAdam SN, et al. 1998. Identification of a gliadin T-cell epitope in coeliac disease: general importance of gliadin deamidation for intestinal T-cell recognition. *Scand. J. Immunol.* 48:111–15
21. Arentz-Hansen H, Korner R, Molberg O, Quarsten H, Vader W, 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.* 191:603–12
22. Kim CY, Quarsten H, Bergseng E, Khosla C, Sollid LM. 2004. Structural basis for HLA-DQ2-mediated presentation of gluten epitopes in celiac disease. *Proc. Natl. Acad. Sci. USA* 101:4175–79
23. Henderson KN, Tye-Din JA, Reid HH, Chen Z, Borg NA, et al. 2007. A structural and immunological basis for the role of human leukocyte antigen DQ8 in celiac disease. *Immunity* 27:23–34
24. Hovhannisyan Z, Weiss A, Martin A, Wiesner M, Tollefsen S, et al. 2008. The role of HLA-DQ8 β57 polymorphism in the anti-gluten T-cell response in coeliac disease. *Nature* 456:534–38
25. Yoshida K, Corper AL, Herro R, Jabri B, Wilson IA, Teyton L. 2010. The diabetogenic mouse MHC class II molecule I-Ag7 is endowed with a switch that modulates TCR affinity. *J. Clin. Investig.* 120:1578–90
26. Tye-Din JA, Stewart JA, Dromey JA, Beissbarth T, van Heel DA, et al. 2010. Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci. Transl. Med.* 2:41ra51
27. Ferguson A, MacDonald TT, McClure JP, Holden RJ. 1975. Cell-mediated immunity to gliadin within the small-intestinal mucosa in coeliac disease. *Lancet* 1:895–97
28. MacDonald TT, Ferguson A. 1976. Hypersensitivity reactions in the small intestine. 2. Effects of allograft rejection on mucosal architecture and lymphoid cell infiltrate. *Gut* 17:81–91
29. Lundin KE, Scott H, Hansen T, Paulsen G, Halstensen TS, et al. 1993. Gliadin-specific, HLA-DQ (α1*0501, β1*0201) restricted T cells isolated from the small intestinal mucosa of celiac disease patients. *J. Exp. Med.* 178:187–96
30. Lundin KE, Scott H, Fausa O, Thorsby E, Sollid LM. 1994. T cells from the small intestinal mucosa of a DR4, DQ7/DR4, DQ8 celiac disease patient preferentially recognize gliadin when presented by DQ8. *Hum. Immunol.* 41:285–91
31. Molberg O, Kett K, Scott H, Thorsby E, Sollid LM, Lundin KE. 1997. Gliadin specific, HLA DQ2–restricted T cells are commonly found in small intestinal biopsies from coeliac disease patients, but not from controls. *Scand. J. Immunol.* 46:103–9
32. Molberg O, McAdam SN, Korner R, Quarsten H, Kristiansen C, et al. 1998. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. *Nat. Med.* 4:713–17
33. van de Wal Y, Kooy Y, van Veelen P, Pena S, Mearin L, et al. 1998. Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. *J. Immunol.* 161:1585–88
34. Dieterich W, Ehnis T, Bauer M, Donner P, Volta U, et al. 1997. Identification of tissue transglutaminase as the autoantigen of celiac disease. *Nat. Med.* 3:797–801
35. Baklien K, Brandtzaeg P, Fausa O. 1977. Immunoglobulins in jejunal mucosa and serum from patients with adult coeliac disease. *Scand. J. Gastroenterol.* 12:149–59
36. Ventura A, Magazzu G, Greco L. 1999. Duration of exposure to gluten and risk for autoimmune disorders in patients with celiac disease. SIGEP Study Group for Autoimmune Disorders in Celiac Disease. *Gastroenterology* 117:297–303
37. Ferguson A, Arranz E, O'Mahony S. 1993. Clinical and pathological spectrum of coeliac disease—active, silent, latent, potential. *Gut* 34:150–51
38. Louka AS, Sollid LM. 2003. HLA in coeliac disease: unravelling the complex genetics of a complex disorder. *Tissue Antigens* 61:105–17

39. Kutlu T, Brousse N, Rambaud C, Le Deist F, Schmitz J, Cerf-Bensussan N. 1993. Numbers of T cell receptor (TCR) $\alpha\beta^+$ but not of TCR $\gamma\delta^+$ intraepithelial lymphocytes correlate with the grade of villous atrophy in coeliac patients on a long term normal diet. *Gut* 34:208–14
40. Stein H, Dienemann D, Sperling M, Zeitz M, Riecken EO. 1988. Identification of a T cell lymphoma category derived from intestinal-mucosa-associated T cells. *Lancet* 2:1053–54
41. Spencer J, MacDonald TT, Diss TC, Walker-Smith JA, Ciclitira PJ, Isaacson PG. 1989. Changes in intraepithelial lymphocyte subpopulations in coeliac disease and enteropathy associated T cell lymphoma (malignant histiocytosis of the intestine). *Gut* 30:339–46
42. Jabri B, de Serre NP, Cellier C, Evans K, Gache C, et al. 2000. Selective expansion of intraepithelial lymphocytes expressing the HLA-E-specific natural killer receptor CD94 in celiac disease. *Gastroenterology* 118:867–79
43. Roberts AI, Lee L, Schwarz E, Groh V, Spies T, et al. 2001. NKG2D receptors induced by IL-15 costimulate CD28-negative effector CTL in the tissue microenvironment. *J. Immunol.* 167:5527–30
44. Meresse B, Chen Z, Ciszewski C, Tretiakova M, Bhagat G, et al. 2004. Coordinated induction by IL15 of a TCR-independent NKG2D signaling pathway converts CTL into lymphokine-activated killer cells in celiac disease. *Immunity* 21:357–66
45. Hue S, Mention JJ, Monteiro RC, Zhang S, Cellier C, et al. 2004. A direct role for NKG2D/MICA interaction in villous atrophy during celiac disease. *Immunity* 21:367–77
46. Meresse B, Curran SA, Ciszewski C, Orbelyan G, Setty M, et al. 2006. Reprogramming of CTLs into natural killer–like cells in celiac disease. *J. Exp. Med.* 203:1343–55
47. Green PH, Jabri B. 2003. Coeliac disease. *Lancet* 362:383–91
48. Polanco IBI, van Leeuwen A, Schreuder I, Khan PM, Guerrero J, et al. 1981. Gluten-sensitive enteropathy in Spain: genetic and environmental factors. In *The Genetics of Coeliac Disease*, ed. RB McConnell, pp. 211–34. Lancaster, UK: MTP
49. Green PH, Cellier C. 2007. Celiac disease. *N. Engl. J. Med.* 357:1731–43
50. Troncone R, Greco L, Mayer M, Mazzarella G, Maiuri L, et al. 1996. In siblings of celiac children, rectal gluten challenge reveals gluten sensitization not restricted to celiac HLA. *Gastroenterology* 111:318–24
51. Catassi C. 1996. Screening of coeliac disease. *Proc. Int. Symp. Coeliac Dis.* 7:23–33
52. Maki M. 1997. Changing features of coeliac disease. *Proc. Int. Symp. Coeliac Dis.* 7:1–7
53. Ascher H. 1997. The role of quantity and quality of gluten-containing cereals in the epidemiology of coeliac disease. *Proc. Int. Symp. Coeliac Dis.* 7:15–22
54. Corazza GR, Andreani ML, Biagi F, Corrao G, Pretolani S, et al. 1997. The smaller size of the "coeliac iceberg" in adults. *Scand. J. Gastroenterol.* 32:917–19
55. Feighery C. 1999. Fortnightly review: coeliac disease. *BMJ* 319:236–39
56. Jabri B, Sollid LM. 2006. Mechanisms of disease: immunopathogenesis of celiac disease. *Nat. Clin. Pract. Gastroenterol. Hepatol.* 3:516–25
57. Kaukinen K, Peraaho M, Collin P, Partanen J, Woolley N, et al. 2005. Small-bowel mucosal transglutaminase 2–specific IgA deposits in coeliac disease without villous atrophy: a prospective and randomized clinical study. *Scand. J. Gastroenterol.* 40:564–72
58. Shewry PR, Tatham AS, Kasarda DD. 1992. Cereal proteins and coeliac disease. In *Coeliac Disease*, ed. M Marsh, pp. 305–48. Oxford, UK: Blackwell
59. Cataldo F, Lio D, Simpore J, Musumeci S. 2002. Consumption of wheat foodstuffs not a risk for celiac disease occurrence in Burkina Faso. *J. Pediatr. Gastroenterol. Nutr.* 35:233–34
60. Arnaiz-Villena A, Benmamar D, Alvarez M, Diaz-Campos N, Varela P, et al. 1995. HLA allele and haplotype frequencies in Algerians. Relatedness to Spaniards and Basques. *Hum. Immunol.* 43:259–68
61. Djoulah S, Sanchez-Mazas A, Khalil I, Benhamamouch S, Degos L, et al. 1994. HLA-DRB1, DQA1 and DQB1 DNA polymorphisms in healthy Algerian and genetic relationships with other populations. *Tissue Antigens* 43:102–9
62. Ayed K, Ayed-Jendoubi S, Sfar I, Labonne MP, Gebuhrer L. 2004. HLA class-I and HLA class-II phenotypic, gene and haplotypic frequencies in Tunisians by using molecular typing data. *Tissue Antigens* 64:520–32
63. Catassi C, Ratsch IM, Gandolfi L, Pratesi R, Fabiani E, et al. 1999. Why is coeliac disease endemic in the people of the Sahara? *Lancet* 354:647–48

64. Mankai A, Landolsi H, Chahed A, Gueddah L, Limem M, et al. 2006. Celiac disease in Tunisia: serological screening in healthy blood donors. *Pathol. Biol.* 54:10–13
65. Maki M, Mustalahti K, Kokkonen J, Kulmala P, Haapalahti M, et al. 2003. Prevalence of celiac disease among children in Finland. *N. Engl. J. Med.* 348:2517–24
66. Lohi S, Mustalahti K, Kaukinen K, Laurila K, Collin P, et al. 2007. Increasing prevalence of coeliac disease over time. *Aliment. Pharmacol. Ther.* 26:1217–25
67. Mustalahti KC, Catassi C, Reunanen C, Fabiani A, Heier M, et al. 2010. The prevalence of celiac disease in Europe: results of a centralized, international mass screening project. *Ann. Med.* 42:587–95
68. Kondrashova A, Mustalahti K, Kaukinen K, Viskari H, Volodicheva V, et al. 2008. Lower economic status and inferior hygienic environment may protect against celiac disease. *Ann. Med.* 40:223–31
69. Remes-Troche JM, Ramirez-Iglesias MT, Rubio-Tapia A, Alonso-Ramos A, Velazquez A, Uscanga LF. 2006. Celiac disease could be a frequent disease in Mexico: prevalence of tissue transglutaminase antibody in healthy blood donors. *J. Clin. Gastroenterol.* 40:697–700
70. Wu J, Xia B, von Blomberg BM, Zhao C, Yang XW, et al. 2010. Coeliac disease: emerging in China? *Gut* 59:418–19
71. van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, et al. 2007. A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. *Nat. Genet.* 39:827–29
72. Djilali-Saiah I, Caillat-Zucman S, Schmitz J, Chaves-Vieira ML, Bach JF. 1994. Polymorphism of antigen processing (TAP, LMP) and HLA class II genes in celiac disease. *Hum. Immunol.* 40:8–16
73. Sollid LM. 2002. Coeliac disease: dissecting a complex inflammatory disorder. *Nat. Rev. Immunol.* 2:647–55
74. Sollid LM, Thorsby E. 1993. HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. *Gastroenterology* 105:910–22
75. Ploski R, Ek J, Thorsby E, Sollid LM. 1993. On the HLA-DQ(α 1*0501, β 1*0201)-associated susceptibility in celiac disease: a possible gene dosage effect of DQB1*0201. *Tissue Antigens* 41:173–77
76. van Belzen MJ, Koeleman BP, Crusius JB, Meijer JW, Bardoel AF, et al. 2004. Defining the contribution of the HLA region to *cis* DQ2-positive coeliac disease patients. *Genes Immun.* 5:215–20
77. Vader W, Stepniak D, Kooy Y, Mearin L, Thompson A, et al. 2003. 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* 100:12390–95
78. Karell K, Louka AS, Moodie SJ, Ascher H, Clot F, et al. 2003. HLA types in celiac disease patients not carrying the DQA1*05-DQB1*02 (DQ2) heterodimer: results from the European Genetics Cluster on Celiac Disease. *Hum. Immunol.* 64:469–77
79. Al-Toma A, Goerres MS, Meijer JW, Pena AS, Crusius JB, Mulder CJ. 2006. Human leukocyte antigen–DQ2 homozygosity and the development of refractory celiac disease and enteropathy-associated T-cell lymphoma. *Clin. Gastroenterol. Hepatol.* 4:315–19
80. Vader LW, de Ru A, Van Der Wal Y, Kooy YM, Benckhuijsen W, et al. 2002. Specificity of tissue transglutaminase explains cereal toxicity in celiac disease. *J. Exp. Med.* 195:643–49
81. Fleckenstein B, Molberg O, Qiao SW, Schmid DG, von der Mulbe F, et al. 2002. Gliadin T cell epitope selection by tissue transglutaminase in celiac disease. Role of enzyme specificity and pH influence on the transamidation versus deamidation process. *J. Biol. Chem.* 277:34109–16
82. Shan L, Molberg O, Parrot I, Hausch F, Filiz F, et al. 2002. Structural basis for gluten intolerance in celiac sprue. *Science* 297:2275–79
83. Todd JA, Bell JI, McDevitt HO. 1988. HLA antigens and insulin-dependent diabetes. *Nature* 333:710
84. Klareskog L, Ronnelid J, Lundberg K, Padyukov L, Alfredsson L. 2008. Immunity to citrullinated proteins in rheumatoid arthritis. *Annu. Rev. Immunol.* 26:651–75
85. Dessen A, Lawrence CM, Cupo S, Zaller DM, Wiley DC. 1997. X-ray crystal structure of HLA-DR4 (DRA*0101, DRB1*0401) complexed with a peptide from human collagen II. *Immunity* 7:473–81
86. Hill JA, Southwood S, Sette A, Jevnikar AM, Bell DA, Cairns E. 2003. Cutting edge: the conversion of arginine to citrulline allows for a high-affinity peptide interaction with the rheumatoid arthritis–associated HLA-DRB1*0401 MHC class II molecule. *J. Immunol.* 171:538–41
87. Siegel M, Strnad P, Watts RE, Choi K, Jabri B, et al. 2008. Extracellular transglutaminase 2 is catalytically inactive, but is transiently activated upon tissue injury. *PLoS ONE* 3:e1861

88. Fallang LE, Bergseng E, Hotta K, Berg-Larsen A, Kim CY, Sollid LM. 2009. Differences in the risk of celiac disease associated with HLA-DQ2.5 or HLA-DQ2.2 are related to sustained gluten antigen presentation. *Nat. Immunol.* 10:1096–101
89. Makrygiannakis D, af Klint E, Lundberg IE, Lofberg R, Ulfgren AK, et al. 2006. Citrullination is an inflammation-dependent process. *Ann. Rheum. Dis.* 65:1219–22
90. Stadinski BD, Delong T, Reisdorph N, Reisdorph R, Powell RL, et al. 2010. Chromogranin A is an autoantigen in type 1 diabetes. *Nat. Immunol.* 11:225–31
91. Acha-Orbea H, McDevitt HO. 1987. The first external domain of the nonobese diabetic mouse class II I-A β chain is unique. *Proc. Natl. Acad. Sci. USA* 84:2435–39
92. Quartey-Papafio R, Lund T, Chandler P, Picard J, Ozegbe P, et al. 1995. Aspartate at position 57 of nonobese diabetic I-Ag7 β-chain diminishes the spontaneous incidence of insulin-dependent diabetes mellitus. *J. Immunol.* 154:5567–75
93. Nistico L, Fagnani C, Coto I, Percopo S, Cotichini R, et al. 2006. Concordance, disease progression, and heritability of coeliac disease in Italian twins. *Gut* 55:803–8
94. Gregersen PK, Olsson LM. 2009. Recent advances in the genetics of autoimmune disease. *Annu. Rev. Immunol.* 27:363–91
95. Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, et al. 2010. Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet.* 42:295–302
96. Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, et al. 2008. Newly identified genetic risk variants for celiac disease related to the immune response. *Nat. Genet.* 40:395–402
97. Chen Y, Kuchroo VK, Inobe J, Hafler DA, Weiner HL. 1994. Regulatory T cell clones induced by oral tolerance: suppression of autoimmune encephalomyelitis. *Science* 265:1237–40
98. Coombes JL, Siddiqui KR, Arancibia-Carcamo CV, Hall J, Sun CM, et al. 2007. A functionally specialized population of mucosal $CD103^+$ DCs induces $Foxp3^+$ regulatory T cells via a TGF-β and retinoic acid–dependent mechanism. *J. Exp. Med.* 204:1757–64
99. Mucida D, Park Y, Kim G, Turovskaya O, Scott I, et al. 2007. Reciprocal TH17 and regulatory T cell differentiation mediated by retinoic acid. *Science* 317:256–60
100. Monteleone G, Pender SL, Alstead E, Hauer AC, Lionetti P, et al. 2001. Role of interferon α in promoting T helper cell type 1 responses in the small intestine in coeliac disease. *Gut* 48:425–29
101. Cammarota G, Cuoco L, Cianci R, Pandolfi F, Gasbarrini G. 2000. Onset of coeliac disease during treatment with interferon for chronic hepatitis C. *Lancet* 356:1494–95
102. George EK, Mearin ML, Bouquet J, von Blomberg BM, Stapel SO, et al. 1996. High frequency of celiac disease in Down syndrome. *J. Pediatr.* 128:555–57
103. Gerdes AM, Horder M, Bonnevie-Nielsen V. 1993. Increased IFN-α-induced sensitivity but reduced reactivity of 2′,5′-oligoadenylate synthetase (2,5AS) in trisomy 21 blood lymphocytes. *Clin. Exp. Immunol.* 93:93–96
104. Sollid LM, Molberg O, McAdam S, Lundin KE. 1997. Autoantibodies in coeliac disease: tissue transglutaminase—guilt by association? *Gut* 41:851–52
105. Bjorck S, Brundin C, Lorinc E, Lynch KF, Agardh D. 2010. Screening detects a high proportion of celiac disease in young HLA-genotyped children. *J. Pediatr. Gastroenterol. Nutr.* 50:49–53
106. Sulkanen S, Halttunen T, Laurila K, Kolho KL, Korponay-Szabo IR, et al. 1998. Tissue transglutaminase autoantibody enzyme–linked immunosorbent assay in detecting celiac disease. *Gastroenterology* 115:1322–28
107. Dieterich W, Laag E, Schopper H, Volta U, Ferguson A, et al. 1998. Autoantibodies to tissue transglutaminase as predictors of celiac disease. *Gastroenterology* 115:1317–21
108. Matysiak-Budnik T, Moura IC, Arcos-Fajardo M, Lebreton C, Menard S, et al. 2008. Secretory IgA mediates retrotranscytosis of intact gliadin peptides via the transferrin receptor in celiac disease. *J. Exp. Med.* 205:143–54
109. MacDonald TT, Bajaj-Elliott M, Pender SL. 1999. T cells orchestrate intestinal mucosal shape and integrity. *Immunol. Today* 20:505–10
110. Tang F, Chen Z, Ciszewski C, Setty M, Solus J, et al. 2009. Cytosolic PLA2 is required for CTL-mediated immunopathology of celiac disease via NKG2D and IL-15. *J. Exp. Med.* 206:707–19

111. Nilsen EM, Jahnsen FL, Lundin KE, Johansen FE, Fausa O, et al. 1998. Gluten induces an intestinal cytokine response strongly dominated by interferon γ in patients with celiac disease. *Gastroenterology* 115:551–63
112. Perera L, Shao L, Patel A, Evans K, Meresse B, et al. 2007. Expression of nonclassical class I molecules by intestinal epithelial cells. *Inflamm. Bowel Dis.* 13:298–307
113. Kasaian MT, Whitters MJ, Carter LL, Lowe LD, Jussif JM, et al. 2002. IL-21 limits NK cell responses and promotes antigen-specific T cell activation: a mediator of the transition from innate to adaptive immunity. *Immunity* 16:559–69
114. Leonard WJ, Spolski R. 2005. Interleukin-21: a modulator of lymphoid proliferation, apoptosis and differentiation. *Nat. Rev. Immunol.* 5:688–98
115. Cellier C, Patey N, Mauvieux L, Jabri B, Delabesse E, et al. 1998. Abnormal intestinal intraepithelial lymphocytes in refractory sprue. *Gastroenterology* 114:471–81
116. Daum S, Cellier C, Mulder CJ. 2005. Refractory coeliac disease. *Best Pract. Res. Clin. Gastroenterol.* 19:413–24
117. Mention JJ, Ben Ahmed M, Begue B, Barbe U, Verkarre V, et al. 2003. Interleukin 15: a key to disrupted intraepithelial lymphocyte homeostasis and lymphomagenesis in celiac disease. *Gastroenterology* 125:730–45
118. Malamut G, El Machhour R, Montcuquet N, Martin-Lanneree S, Dusanter-Fourt I, et al. 2010. IL-15 triggers an antiapoptotic pathway in human intraepithelial lymphocytes that is a potential new target in celiac disease–associated inflammation and lymphomagenesis. *J. Clin. Investig.* 120:2131–43
119. Fehniger TA, Suzuki K, VanDeusen JB, Cooper MA, Freud AG, Caligiuri MA. 2001. Fatal leukemia in interleukin-15 transgenic mice. *Blood Cells Mol. Dis.* 27:223–30
120. Ruprecht CR, Gattorno M, Ferlito F, Gregorio A, Martini A, et al. 2005. Coexpression of CD25 and CD27 identifies FoxP3$^+$ regulatory T cells in inflamed synovia. *J. Exp. Med.* 201:1793–803
121. Ben Ahmed M, Belhadj Hmida N, Moes N, Buyse S, Abdeladhim M, et al. 2009. IL-15 renders conventional lymphocytes resistant to suppressive functions of regulatory T cells through activation of the phosphatidylinositol 3-kinase pathway. *J. Immunol.* 182:6763–70
122. Maiuri L, Ciacci C, Ricciardelli I, Vacca L, Raia V, et al. 2003. Association between innate response to gliadin and activation of pathogenic T cells in coeliac disease. *Lancet* 362:30–37
123. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, et al. 2009. STRING 8—a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res.* 37:D412–16
124. Azimi N, Brown K, Bamford RN, Tagaya Y, Siebenlist U, Waldmann TA. 1998. Human T cell lymphotropic virus type I Tax protein trans-activates interleukin 15 gene transcription through an NF-κB site. *Proc. Natl. Acad. Sci. USA* 95:2452–57
125. Washizu J, Nishimura H, Nakamura N, Nimura Y, Yoshikai Y. 1998. The NF-κB binding site is essential for transcriptional activation of the *IL-15* gene. *Immunogenetics* 48:1–7
126. Collin P, Kaukinen K, Valimaki M, Salmi J. 2002. Endocrinological disorders and celiac disease. *Endocr. Rev.* 23:464–83
127. Liu E, Eisenbarth GS. 2002. Type 1A diabetes mellitus–associated autoimmunity. *Endocrinol. Metab. Clin. North Am.* 31:391–410
128. Casella G, D'Inca R, Oliva L, Daperno M, Saladino V, et al. 2010. Prevalence of celiac disease in inflammatory bowel diseases: an IG-IBD multicentre study. *Dig. Liver Dis.* 42:175–78
129. Leeds JS, Horoldt BS, Sidhu R, Hopper AD, Robinson K, et al. 2007. Is there an association between coeliac disease and inflammatory bowel diseases? A study of relative prevalence in comparison with population controls. *Scand. J. Gastroenterol.* 42:1214–20
130. Jabri B, Sollid LM. 2009. Tissue-mediated control of immunopathology in coeliac disease. *Nat. Rev. Immunol.* 9:858–70
131. Martinon F, Mayor A, Tschopp J. 2009. The inflammasomes: guardians of the body. *Annu. Rev. Immunol.* 27:229–65
132. Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, et al. 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419:832–37
133. Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, et al. 2006. Positive natural selection in the human lineage. *Science* 312:1614–20

134. Voight BF, Kudaravalli S, Wen X, Pritchard JK. 2006. A map of recent positive selection in the human genome. *PLoS Biol.* 4:e72
135. Barreiro LB, Quintana-Murci L. 2010. From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nat. Rev. Genet.* 11:17–30
136. Zhernakova A, Elbers CC, Ferwerda B, Romanos J, Trynka G, et al. 2010. Evolutionary and functional analysis of celiac risk loci reveals SH2B3 as a protective factor against bacterial infection. *Am. J. Hum. Genet.* 86:970–77
137. Koskinen LL, Einarsdottir E, Dukes E, Heap GA, Dubois P, et al. 2009. Association study of the IL18RAP locus in three European populations with coeliac disease. *Hum. Mol. Genet.* 18:1148–55
138. Wolfe ND, Dunavan CP, Diamond J. 2007. Origins of major human infectious diseases. *Nature* 447:279–83
139. Not T, Horvath K, Hill ID, Partanen J, Hammed A, et al. 1998. Celiac disease risk in the USA: high prevalence of antiendomysium antibodies in healthy blood donors. *Scand. J. Gastroenterol.* 33:494–98
140. Kaistha A, Castells S. 2008. Celiac disease in African American children with type 1 diabetes mellitus in inner city Brooklyn. *Pediatr. Endocrinol. Rev.* 5(Suppl. 4):994–98
141. Stene LC, Honeyman MC, Hoffenberg EJ, Haas JE, Sokol RJ, et al. 2006. Rotavirus infection frequency and risk of celiac disease autoimmunity in early childhood: a longitudinal study. *Am. J. Gastroenterol.* 101:2333–40
142. Ivarsson A, Persson LA, Nystrom L, Ascher H, Cavell B, et al. 2000. Epidemic of coeliac disease in Swedish children. *Acta Paediatr.* 89:165–71
143. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57:289–300
144. Raychaudhuri S, Remmers EF, Lee AT, Hackett R, Guiducci C, et al. 2008. Common variants at CD40 and other loci confer risk of rheumatoid arthritis. *Nat. Genet.* 40:1216–23
145. Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, et al. 2009. Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. *Nat. Genet.* 41:1234–37
146. Smyth DJ, Plagnol V, Walker NM, Cooper JD, Downes K, et al. 2008. Shared and distinct genetic variants in type 1 diabetes and celiac disease. *N. Engl. J. Med.* 359:2767–77
147. Gregersen PK, Amos CI, Lee AT, Lu Y, Remmers EF, et al. 2009. REL, encoding a member of the NF-κB family of transcription factors, is a newly defined risk locus for rheumatoid arthritis. *Nat. Genet.* 41:820–23
148. McGovern DP, Gardet A, Torkvist L, Goyette P, Essers J, et al. 2010. Genome-wide association identifies multiple ulcerative colitis susceptibility loci. *Nat. Genet.* 42:332–37
149. Zhernakova A, Festen EM, Franke L, Trynka G, van Diemen CC, et al. 2008. Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring *CARD9* and *IL18RAP*. *Am. J. Hum. Genet.* 82:1202–10
150. Raychaudhuri S, Thomson BP, Remmers EF, Eyre S, Hinks A, et al. 2009. Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk. *Nat. Genet.* 41:1313–18
151. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, et al. 2009. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat. Genet.* 41:703–7
152. Barreto M, Santos E, Ferreira R, Fesel C, Fontes MF, et al. 2004. Evidence for *CTLA4* as a susceptibility gene for systemic lupus erythematosus. *Eur. J. Hum. Genet.* 12:620–26
153. Torres B, Aguilar F, Franco E, Sanchez E, Sanchez-Roman J, et al. 2004. Association of the CT60 marker of the *CTLA4* gene with systemic lupus erythematosus. *Arthritis Rheum.* 50:2211–15
154. De Jager PL, Jia X, Wang J, de Bakker PI, Ottoboni L, et al. 2009. Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci. *Nat. Genet.* 41:776–82
155. Coenen MJ, Trynka G, Heskamp S, Franke B, van Diemen CC, et al. 2009. Common and different genetic background for rheumatoid arthritis and coeliac disease. *Hum. Mol. Genet.* 18:4195–203
156. Zhernakova A, Alizadeh BZ, Bevova M, van Leeuwen MA, Coenen MJ, et al. 2007. Novel association in chromosome 4q27 region with rheumatoid arthritis and confirmation of type 1 diabetes point to a general risk locus for autoimmune diseases. *Am. J. Hum. Genet.* 81:1284–88

157. Julia A, Ballina J, Canete JD, Balsa A, Tornero-Molina J, et al. 2008. Genome-wide association study of rheumatoid arthritis in the Spanish population: KLF12 as a risk locus for rheumatoid arthritis susceptibility. *Arthritis Rheum.* 58:2275–86
158. Wellcome Trust Case Control Consort. 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 447:661–78
159. Cooper JD, Smyth DJ, Smiles AM, Plagnol V, Walker NM, et al. 2008. Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. *Nat. Genet.* 40:1399–401
160. Hakonarson H, Grant SF, Bradfield JP, Marchand L, Kim CE, et al. 2007. A genome-wide association study identifies *KIAA0350* as a type 1 diabetes gene. *Nature* 448:591–94
161. Hom G, Graham RR, Modrek B, Taylor KE, Ortmann W, et al. 2008. Association of systemic lupus erythematosus with C8orf13-BLK and ITGAM-ITGAX. *N. Engl. J. Med.* 358:900–9
162. Aust. N.Z. Mult. Scler. Genet. Consort. (ANZgene). 2009. Genome-wide association study identifies new multiple sclerosis susceptibility loci on chromosomes 12 and 20. *Nat. Genet.* 41:824–28
163. Comabella M, Craig DW, Camina-Tato M, Morcillo C, Lopez C, et al. 2008. Identification of a novel risk locus for multiple sclerosis at 13q31.3 by a pooled genome-wide scan of 500,000 single nucleotide polymorphisms. *PLoS ONE* 3:e3490
164. Silverberg MS, Cho JH, Rioux JD, McGovern DP, Wu J, et al. 2009. Ulcerative colitis–risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat. Genet.* 41:216–20
165. Grant SF, Qu HQ, Bradfield JP, Marchand L, Kim CE, et al. 2009. Follow-up analysis of genome-wide association data identifies novel loci for type 1 diabetes. *Diabetes* 58:290–95
166. Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, et al. 2007. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. *Nat. Genet.* 39:1477–82
167. Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, et al. 2008. Genetic variants near *TNFAIP3* on 6q23 are associated with systemic lupus erythematosus. *Nat. Genet.* 40:1059–61
168. Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, et al. 2009. Genome-wide scan reveals association of psoriasis with IL-23 and NF-κB pathways. *Nat. Genet.* 41:199–204
169. Yang W, Shen N, Ye DQ, Liu Q, Zhang Y, et al. 2010. Genome-wide association study in Asian populations identifies variants in *ETS1* and *WDFY4* associated with systemic lupus erythematosus. *PLoS Genet.* 6:e1000841
170. Bronson PG, Caillier S, Ramsay PP, McCauley JL, Zuvich RL, et al. 2010. *CIITA* variation in the presence of HLA-DRB1*1501 increases risk for multiple sclerosis. *Hum. Mol. Genet.* 19:2331–40
171. Zuvich RL, McCauley JL, Oksenberg JR, Sawcer SJ, De Jager PL, et al. 2010. Genetic variation in the IL7RA/IL7 pathway increases multiple sclerosis susceptibility. *Hum. Genet.* 127:525–35
172. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, et al. 2008. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat. Genet.* 40:955–62
173. Parkes M, Barrett JC, Prescott NJ, Tremelling M, Anderson CA, et al. 2007. Sequence variants in the autophagy gene *IRGM* and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat. Genet.* 39:830–32
174. Cooper JD, Walker NM, Smyth DJ, Downes K, Healy BC, Todd JA. 2009. Follow-up of 1,715 SNPs from the Wellcome Trust Case Control Consortium genome-wide association study in type I diabetes families. *Genes Immun.* 10(Suppl. 1):85–94S

Annual Review of Immunology
Volume 29, 2011

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Immunology* articles may be found at http://immunol.annualreviews.org/errata.shtml