Helsinki University Biomedical Dissertations No. 128

INSIGHTS INTO THE MOLECULAR GENETICS OF CELIAC DISEASE:

applying family- and population-based strategies

Lotta Koskinen

Department of Medical Genetics and
Research Program for Molecular Medicine
University of Helsinki
Helsinki, Finland

ACADEMIC DISSERTATION

To be presented, with the permission of the Faculty of Medicine, University of Helsinki, for public examination in Lecture Hall 3, Biomedicum Helsinki 1, Haartmaninkatu 8, Helsinki, on September 25th 2009, at 12 noon.

Helsinki 2009

Supervisors: Päivi Saavalainen, Ph.D.
Department of Medical Genetics, and
Research Program for Molecular Medicine
University of Helsinki, Helsinki, Finland

and

Juha Kere, M.D., Ph.D.
Department of Medical Genetics
University of Helsinki, Helsinki, Finland
and
Folkhälsan Institute of Genetics
Helsinki, Finland
and
Department of Biosciences and Nutrition at Novum
Karolinska Institutet, Huddinge, Sweden

Reviewers: Katarina Pelin, Ph.D.
Department of Biological and Environmental Sciences
Division of Genetics
University of Helsinki, Helsinki, Finland

and

Outi Vaarala, M.D., Ph.D.
Immune Response Unit
Department of Vaccination and Immune Protection
National Institute for Health and Welfare, Helsinki, Finland

Opponent: Jose Ramón Bilbao, Ph.D.
Immunogenetics Research Laboratory
Hospital de Cruces, Barakaldo, Spain
and
Department of Genetics, Physical Anthropology and Animal Physiology
University of the Basque Country, Spain

Yliopistopaino
Helsinki 2009

ISSN 1457-8433
ISBN 978-952-10-5701-4 (print)
ISBN 978-952-10-5702-1 (PDF)

TO MY FAMILY

Mikä onkaan suurempaa kuin löytää ratkaisematon arvoitus
kummallisine piirteineen?
Edith Södergran
(suom. Uuno Kailas)

What could be greater than the
strange features of a riddle unsolved?
Edith Södergran
(translated by Gounil Brown)

Table of contents

List of original publications

This Ph.D. thesis is based on the following publications:

I Koskinen LLE, Einarsdóttir E, Korponay-Szabó IR, Kurppa K, Kaukinen K, Sistonen P., Pocsai Z, Széles G, Ádány R, Mäki M, Kere J, Saavalainen P. Finemapping of the *CELIAC2* locus on chromosome 5q31-q33 in the Finnish and Hungarian populations. Tissue Antigens. *In press.*

II Koskinen LLE, Korponay-Szabó IR, Viiri K, Juuti-Uusitalo K, Kaukinen K, Lindfors K, Mustalahti K, Kurppa K, Ádány R, Pocsai Z, Széles G, Einarsdóttir E, Wijmenga C, Mäki M, Partanen J, Kere J, Saavalainen P. Myosin IXB gene region and gluten intolerance: linkage to coeliac disease and a putative dermatitis herpetiformis association. J Med Genet. 2008 45(4):222-227.

III Koskinen LLE*, Einarsdóttir E*, Dukes E, Heap GA, Dubois P, Korponay-Szabó IR, Kaukinen K, Kurppa K, Ziberna F, Vatta S, Not T, Ventura A, Sistonen P, Ádány R, Pocsai Z, Széles G, Mäki M, Kere J, Wijmenga C, van Heel DA, Saavalainen P. Association study of the *IL18RAP* locus in three European populations with coeliac disease. Hum Mol Genet. 2009 18(6):1148-1155.

IV Koskinen LLE*, Romanos J*, Kaukinen K, Mustalahti M, Korponay-Szabó IR, Barisani D, Bardella MT, Ziberna F, Vatta S, Széles G, Pocsai Z, Karell K, Haimila K, Ádány R, Not T, Ventura A, Mäki M, Partanen J, Wijmenga C, Saavalainen P. Cost-effective HLA typing with tagging SNPs predicts celiac disease risk haplotypes in the Finnish, Hungarian, and Italian populations. Immunogenetics 2009 61(4):247-256.

* Equal contribution

The publications are referred to in the text by their roman numerals.

Abbreviations

bp	base pair
CD4	cluster of differentiation 4, a glycoprotein expressed predominantly in helper T cells
CD8	cluster of differentiation 8, a glycoprotein expressed predominantly in cytotoxic T cells
C.I.	confidence interval
cM	centimorgan
DH	dermatitis herpetiformis
ELISA	enzyme-linked immunosorbent assay
EMA	endomysial antibody
EMSA	electophoretic mobility shift assay
ESPGAN	European Society for Paediatric Gastroenterology and Nutrition
HLA	human leukocyte antigen
HLA-DQ	a human leukocyte antigen heterodimer consisting of DQA1 and DQB1 molecules
IFNγ	interferon gamma
IgA	immunoglobulin A
IgG	immunoglobulin G
IL	interleukin
IL18R1	interleukin 18 receptor 1
IL18RAP	interleukin 18 receptor accessory protein, interleukin 18 receptor beta
kb	kilobase
kDa	kilodalton
KLRK1	killer cell lectin-like receptor subfamily K, member 1
LD	linkage disequilibrium
LOD	logarithm of odds
Mb	megabase
MHC	major histocompatibility complex
MICA	MHC class I polypeptide-related sequence A
MYO9B	myosin IXB
NPL	non-parametric linkage
OR	odds ratio
PBMC	peripheral blood mononuclear cell
SNP	single-nucleotide polymorphism
TDT	transmission-disequilibrium test
TGM2	transglutaminase 2 (C polypeptide, protein-glutamine-gamma-glutamyltransferase, tissue transglutaminase)
Th1	T helper cell type 1: a subset of CD4+ T cells
Th2	T helper cell type 2: a subset of CD4+ T cells
Th17	T helper cell type 17: a subset of CD4+ T cells

Abstract

Celiac disease, or gluten intolerance, is triggered by dietary glutens in genetically susceptible individuals and it affects approximately 1% of the Caucasian population. The best known genetic risk factors for celiac disease are HLA DQ2 and DQ8 heterodimers, which are necessary for the development of the disease. However, they alone are not sufficient for disease induction, other risk factors are required. This thesis investigated genetic factors for celiac disease, concentrating on susceptibility loci on chromosomes 5q31-q33, 19p13 and 2q12 previously reported in genome-wide linkage and association studies. In addition, a novel genotyping method for the detection of HLA DQ2 and DQ8 coding haplotypes was validated.

This study was conducted using Finnish and Hungarian family materials, and Finnish, Hungarian and Italian case-control materials. Genetic linkage and association were analysed in these materials using candidate gene and fine-mapping approaches.

The results confirmed linkage to celiac disease on the chromosomal regions 5q31-q33 and 19p13. Fine-mapping on chromosome 5q31-q33 revealed several modest associations in the region, and highlighted the need for further investigations to locate the causal risk variants. The *MYO9B* gene on chromosome 19p13 showed evidence for linkage and association particularly with dermatitis herpetiformis, the skin manifestation of celiac disease. This implies a potential difference in the genetic background of the intestinal and skin forms of the disease, although studies on larger samplesets are required.

The *IL18RAP* locus on chromosome 2q12, shown to be associated with celiac disease in a previous genome-wide association study and a subsequent follow-up, showed association in the Hungarian population in this study. The expression of *IL18RAP* was further investigated in small intestinal tissue and in peripheral blood mononuclear cells. The results showed that *IL18RAP* is expressed in the relevant tissues. Two putative isoforms of IL18RAP were detected by Western blot analysis, and the results suggested that the ratios and total levels of these isoforms may contribute to the aetiology of celiac disease.

A novel genotyping method for celiac disease-associated HLA haplotypes was also validated in this thesis. The method utilises single-nucleotide polymorphisms tagging these HLA haplotypes with high sensitivity and specificity. Our results suggest that this method is transferable between populations, and it is suitable for large-scale analysis.

In conclusion, this doctorate study provides an insight into the roles of the 5q31-q33, *MYO9B*, *IL18RAP* and *HLA* loci in the susceptibility to celiac disease in the Finnish, Hungarian and Italian populations, highlighting the need for further studies at these genetic loci and examination of the function of the candidate genes.

Introduction

Together with other common immune-mediated diseases, such as type 1 diabetes, rheumatoid arthritis and asthma, celiac disease falls into the category complex diseases because it appears to be causally influenced by multiple risk factors. These factors can be genetic or environmental, or may be comprised of an interplay between both. Complex diseases can present with a variety of symptoms and disease manifestations, and they do not conform to the classical Mendelian inheritance patterns, although familial clustering can often be seen.

As complex diseases affect a relatively large portion of the human population, and they are a major burden on health care systems, substantial resources are invested in understanding the aetiology of these diseases and developing better tools for disease diagnosis and individualised treatment. The challenge to research efforts is that the contribution of each risk factor to the disease is often very slight, and even the healthy population has varying degrees of predisposition in relation to the risk factors.

Celiac disease also belongs to a group of inflammatory diseases with autoimmune features. Its unique feature is that the major environmental risk factor for celiac disease is known: the pathology is triggered by exposure to dietary gluten (Dicke 1950). Other environmental risk factors for celiac disease may also be involved. It is well recognised that certain haplotypes within the major histocompatibility complex (MHC) region constitute the main genetic risk factors for celiac disease (Sollid *et al.* 1989, Karell *et al.* 2003). Their role in the pathogenesis of the disease has also been confirmed (Molberg *et al.* 1998), making celiac disease a relatively well-characterised disease when compared to several other disorders with complex aetiology. Nevertheless, our understanding of the genetic background and the pathogenesis of celiac disease remains incomplete, and is the focus of cutting-edge research. In particular, recent advances in the field of genetics have resulted in more powerful study designs to locate genetic risk factors for complex diseases such as celiac disease.

In this thesis, genetic susceptibility loci for celiac disease were studied using Finnish, Hungarian and Italian patient materials. The study was started in 2005 after the main susceptibility loci for celiac disease were identified through genome-wide linkage scans. Two of the candidate loci were investigated in this study (I, II). Novel risk genes for celiac disease were highlighted in the first genome-wide association scan in celiac disease published in 2007 and in its follow-up studies, and one of the newly identified genes was also analysed in this doctorate study (III). Lastly, a novel genotyping method to predict human leukocyte antigen (*HLA*) risk haplotypes for celiac disease was validated to be used in large scale research projects (IV).

Review of the literature

1. Celiac disease

Celiac disease (MIM #212750), also known as gluten intolerance, gluten sensitive enteropathy or celiac sprue, is a chronic inflammatory condition of the small intestinal mucosa with autoimmune features. It affects approximately 1% of the gluten-consuming population; although, similarly to atopic diseases and type 1 diabetes, the prevalence of celiac disease is increasing, and is already 2% in Finland (Lohi *et al.* 2007). Celiac disease is triggered by dietary glutens from wheat, barley and rye, which set off an inflammatory reaction in the small intestine, leading to villous atrophy and crypt hyperplasia. The villous atrophy classically leads to malabsorption of nutrients with various associated symptoms (Mäki and Collin 1997). Presently, the only cure for celiac disease is a strict gluten-free diet, which typically leads to a complete recovery of the small intestinal tissue.

The pathogenesis of celiac disease remains incompletely understood, but the immune system - possibly with interplay between the adaptive and innate immune responses - has a major role. Several twin and family studies have shown that the genetic component of celiac disease is strong and is the main factor contributing to the disease susceptibility. To date, the only known functionally-characterised genetic risk factors for celiac disease are certain alleles in *HLA* genes located on chromosome 6p21.3. However, these variants do not explain the whole genetic background of celiac disease, and there is evidence for several genetic non-*HLA* risk variants. The search for other susceptibility genes is ongoing, and along with a growing understanding of the normal variation in the human genome, the puzzle of the complex genetic background to celiac disease is slowly starting to unravel.

1.1. Clinical characteristics

The clinical picture of celiac disease was first described by Samuel Gee in 1888 (Gee 1888), although the adverse effects of ingested gluten were recognised over 60 years later (Dicke 1950). The various primary symptoms of celiac disease are caused by a flattening of the small intestinal mucosa, known as partial, subtotal or total villous atrophy (Figure 1). The villi are necessary for efficient absorption of ingested nutrients, and their absence leads to the classical symptoms of celiac disease, including diarrhoea, weight loss and fatigue. Severe malabsorption syndrome has also been common among patients affected with celiac disease. (Mäki and Collin 1997).

The clinical picture of celiac disease has changed over time, and the clinical signs of celiac disease are becoming milder than the symptoms that have previously been defined as classical symptoms. Along with the increasing prevalence of celiac disease, so-called atypical symptoms and manifestations of celiac disease have become common and the disease can even be apparently symptom-free. There are several atypical symptoms associated with celiac disease, such as various neurological symptoms, dental enamel defects, infertility, osteoporosis, aphthosis of the mouth mucosa, joint symptoms and elevated liver-enzyme concentrations. As untreated celiac disease is associated with osteoporosis and severe

(although rare) conditions such as malignant lymphoma and neural ataxias, asymptomatic patients are also recommended to follow a gluten-free diet. (Mäki and Collin 1997).

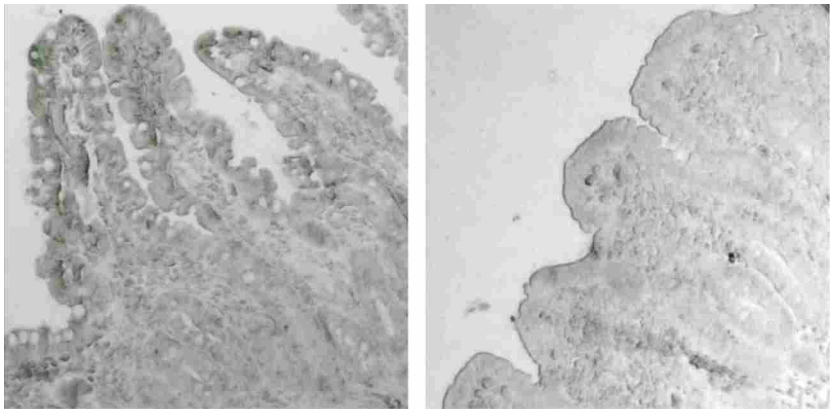

Figure 1 *Healthy small intestinal mucosa (on the left) is characterised by long villi extending into the small intestinal lumen. Partial, subtotal or total villous atrophy is typical in celiac disease (on the right).*

1.1.1. Dermatitis herpetiformis

Dermatitis herpetiformis (DH) (MIM 601230), the skin manifestation of celiac disease, affects approximately one-fourth of gluten intolerant patients (Collin and Reunala 2003). Its relationship with celiac disease was suggested by Samman in 1955 (Samman 1955) and established a decade later (Marks *et al.* 1966, Marks and Shuster 1968). DH is a blistering skin condition that can occur at any age (Reunala 2001). The majority of patients with DH are also affected with mostly asymptomatic intestinal manifestations of gluten intolerance (Gawkrodger *et al.* 1984, Savilahti *et al.* 1992). As with celiac disease, DH is cured by excluding gluten from the diet (Fry *et al.* 1973). Dapsone also efficiently controls the rash that is typical for DH (Reunala 2001).

Like celiac disease, DH has a complex aetiology, and appears to share the same genetic background as the intestinal form of gluten intolerance. For example, the *HLA* alleles known to be associated with celiac disease also predispose to DH (Spurkland *et al.* 1997). DH and celiac disease also co-occur in the same families (Hervonen *et al.* 2002), indicating that these two distinct phenotypes are different manifestations of the same disease. Further evidence supporting this idea comes from monozygotic twins, where one individual of the genetically identical twin pair can have DH and the other has celiac disease (Hervonen *et al.* 2000). There is also evidence for changing phenotype from intestinal celiac disease to DH after long-term gluten exposure (Kurppa *et al.* 2008). In conflict with these studies, there is evidence for distinct genetic susceptibility factors underlying the two conditions (Holopainen *et al.* 2001). To date, it is not known what triggers DH rather than celiac disease.

1.1.2. Refractory celiac disease

Refractory celiac disease is a rare condition which can be defined as persisting or recurring villous atrophy and crypt hyperplasia with increased intraepithelial lymphocytes despite strict

gluten-free diet for at least 12 months. Refractory celiac disease can also manifest with severe persisting symptoms and require interventions independent of the duration of gluten-free diet. Refractory celiac disease has two categories. Patients with type I refractory celiac disease do not have aberrant T cells in their small intestinal epithelium and seem to profit from immunosuppressive treatment. Patients with type II refractory celiac disease demonstrate an aberrant clonal intraepithelial T cell population, and have a high risk of developing overt lymphoma. Type II refractory celiac disease is usually resistant to immunosuppressive therapies: cytotoxic chemotherapy and autologous stem cell transplantation have been suggested for treatment. (Al-Toma *et al.* 2007).

1.1.3. Other associated disorders

Celiac disease co-occurs with several other disease conditions. As expected, other autoimmune diseases associated with HLA DQ2 and DQ8 are more common among patients with celiac disease than the rest of the population. Among these conditions are type 1 diabetes, autoimmune thyroid diseases, selective immunoglobulin A (IgA) deficiency, Sjögren's syndrome, alopecia areata and Addison's disease (Mäki and Collin 1997). Co-morbidities of celiac disease with schizophrenia (Eaton *et al.* 2006) and rheumatoid arthritis (Neuhausen *et al.* 2008) have also been reported.

1.2. Diagnostics

The diagnostic criteria for celiac disease, established by the European Society for Paediatric Gastroenterology and Nutrition (ESPGAN), are small bowel mucosal atrophy with improvement or normalisation on a gluten-free diet, and the presence of circulating antibodies specific for celiac disease (Working group of European Society of Paediatric Gastroenterology and Nutrition 1990). A requirement for all diagnostic testing of celiac disease is that the patient is on a gluten-containing diet. Typically, the first step in diagnosis is a serologic test, where celiac disease-specific antibodies (IgA specific for tissue transglutaminase [TGM2] and IgA endomysial antibody [EMA]) are measured from the peripheral blood of the patients. For IgA-deficient patients a serological test based on immunoglobulin G (IgG) antibodies can be performed. Small intestinal villous patterns from the biopsy specimens are assessed based on Marsh and Oberhuber's classification of duodenal histological lesions (Oberhuber *et al.* 1999).

Diagnosis of DH is based on the demonstration of granular IgA deposits in sub-epidermal skin (van der Meer 1969).

Traditionally, both positive serology and biopsy results are required for a presumptive diagnosis of celiac disease. However, as the clinical picture of celiac disease has been changing during the past years, gluten intolerance is no longer restricted only to villous atrophy. It is now recognised that individuals with normal mucosal villous structure can have celiac disease, with villous atrophy thought to develop gradually during the disease process (Mäki and Collin 1997). The question of whether the diagnostic criteria should be modified to be based less on the requirement of villous atrophy is subject to ongoing debate (Kaukinen *et al.* 2001, Kurppa *et al.* 2009). Given that celiac disease often presents with silent or atypical

symptoms, population screening for celiac disease using disease-specific autoantibodies has been suggested because early treatment of celiac disease may improve the quality of life for affected individuals (Viljamaa *et al.* 2005, Korponay-Szabó *et al.* 2007).

1.3. Epidemiology

The prevalence of celiac disease is around 1% in Western European populations, although there are differences between populations (Dube *et al.* 2005). A minority of the affected individuals are symptomatic: the majority of patients demonstrate silent or latent celiac disease, which can be detected by antibody screening. Therefore, the prevalence of symptomatic patients presents only the tip of the so-called celiac disease iceberg (Mäki and Collin 1997).

Like several other autoimmune and inflammatory diseases, the prevalence of celiac disease is increasing throughout the world. In Finland, the prevalence of both clinically diagnosed and previously unrecognised celiac disease has increased significantly during the past two decades. The total prevalence of celiac disease has recently been estimated to be already approximately 2% of the Finnish adult population; a two-fold increase compared to the prevalence 20 years earlier (Lohi *et al.* 2007). Such a rapid increase in prevalence may reflect changes in our environment and way of life.

The hygiene hypothesis has been proposed to play a major role in the relatively recent increase in the prevalence of autoimmune and allergic diseases. This hypothesis argues that improvements to sanitation and health care, which lead to the reduction of microbial contacts and incidence of infectious diseases, have increased the prevalence of allergies and diseases with autoimmune-features (Bach 2002, Romagnani 2004, Sewell *et al.* 2002). There is some epidemiological evidence supporting the hygiene hypothesis in relation to celiac disease, but more detailed studies are still required to evaluate its role. For example, low economic status and inferior hygienic living conditions have been shown to protect from celiac disease in a study where prevalence rates were compared between Russian Karelia and Finland (Kondrashova *et al.* 2008). The two investigated populations share partly the same ancestry and are equally exposed to grains, but their socio-economic conditions differ significantly. Kondrashova *et al.* (2008) demonstrated that the prevalence of autoantibodies specific for celiac disease was significantly lower among study subjects from Russian Karelia than Finland (Kondrashova *et al.* 2008). Some epidemiological studies have also suggested that the season of birth for a child with celiac disease differs significantly from the pattern in the general population (Ivarsson *et al.* 2003, Lewy *et al.* 2009). For example, Swedish children born in the summer had a higher risk of celiac disease when compared to children born in the winter (Ivarsson *et al.* 2003). This suggests that environmental risk factors with seasonal patterns, such as infections occurring most frequently in the winter, can play a role in celiac disease risk. It has also been observed that certain viral infections appear to increase the risk of celiac disease, such as adenovirus 12 and hepatitis C virus (Plot and Amital 2009). Thus, the role of infections and microbial contacts in the aetiology of celiac disease still requires further investigations.

As with most autoimmune conditions, females are more frequently affected with celiac disease. Two-thirds of patients with celiac disease are females, implying that hormones may affect disease susceptibility.

Dietary patterns, especially in childhood, may affect disease incidence or age of onset. The amount and timing of gluten introduction in diet together with the duration of breast feeding was reported to have a dramatic effect on the appearance of celiac disease in Sweden (Ivarsson *et al.* 2000). The annual incidence of celiac disease in children younger than 2 years had increased four-fold in two years (1985-1987) followed by a sharp decline to the previous levels ten years later. This epidemic, rare for common complex diseases, coincided with changes in infant feeding practices. Introduction of dietary gluten to children at a later age, increases in the amount of gluten consumed, and reduced breastfeeding after dietary gluten introduction were suggested to increase the incidence of celiac disease (Ivarsson *et al.* 2000).

2. Pathogenesis of celiac disease

Celiac disease is predominantly a T lymphocyte-mediated disorder, with activation of both cellular and humoral immune responses in the small intestinal mucosa. Both adaptive and innate immune responses are seen in active celiac disease. These immunological reactions are triggered by dietary gluten from wheat, barley or rye, and they can lead to destruction of the small intestinal villous structure. The pathogenesis of celiac disease remains incompletely understood, but is thought to consist of several intersecting molecular and cellular events in the small intestine. The intestinal lesion is characterised by infiltration of lymphocytes to the epithelium, increased density of leukocytes in the lamina propria, crypt cell hyperplasia and villous atrophy (Trier 1991, Sollid 2000).

2.1. Gluten as a triggering factor for celiac disease

Gluten is composed of storage proteins from the endosperm of wheat, barley or rye. It comprises over 80% of the protein in a wheat seed, and is an important source of nutrition worldwide. Wheat gluten is composed of two types of proteins: alcohol-soluble gliadins and alcohol-insoluble glutenins. Gliadins comprise a family of proteins with a high content of glutamine and proline residues. They are phylogenetically related to secalins of rye and hordeins of barley, and they are all capable of triggering celiac disease in genetically susceptible individuals (Mäki and Collin 1997). Oats are distantly related to wheat, barley and rye; therefore, the structure of its storage proteins (avenins) differs from the structure of gliadins, secalins and hordeins. Avenins appear to be incapable of triggering pathology in most patients with celiac disease, although some concerns remain (Lundin *et al.* 2003, Peräaho *et al.* 2004, Arentz-Hansen *et al.* 2004, Holm *et al.* 2006).

Due to their high content of proline and glutamine residues, gliadins, secalins and hordeins are relatively resistant to proteolytic digestion in the small intestine. This may explain the accumulation of large gliadin-derived polypeptides in the intestinal mucosa after gluten ingestion. Certain gliadin polypeptides have been observed to be more toxic than others for individuals with celiac disease: an immuno-dominant 33-mer gliadin peptide appears to be particularly good at stimulating T cell-mediated inflammatory responses seen in celiac small intestinal mucosa (Piper *et al.* 2002, Shan *et al.* 2002). Several other peptides have also been reported (Arentz-Hansen *et al.* 2002, Vader *et al.* 2002, Camarca *et al.* 2009). Other gliadin

epitopes are thought to be directly cytotoxic and have been shown to trigger innate immune responses (Giovannini *et al.* 2000, Hue *et al.* 2004, Barone *et al.* 2007).

2.2. Tissue transglutaminase and its roles in celiac disease

Tissue transglutaminase (TGM2) has been shown to be an important factor in the pathogenesis of celiac disease. It is a ubiquitously-expressed enzyme belonging to a family of calcium-dependent transamidating enzymes. It is usually active in the extracellular space and catalyses the covalent and irreversible cross-linking of a protein with a glutamine residue to a second protein with a lysine residue (Folk and Cole 1966, Folk and Chung 1985). With the formation of stable bonds between proteins, TGM2 is involved in stabilising tissue structures and is expressed at high levels during wound healing (Bowness *et al.* 1988) and angiogenesis (Haroon *et al.* 1999).

Gliadin peptides are ideal substrates for TGM2 because of their high glutamine content. These peptides thereby act as glutamine donors for TGM2. The reaction between gliadin peptides and TGM2 creates a covalent link between the molecules, leading to gliadin-TGM2 complexes (Szabolcs *et al.* 1987, Dieterich *et al.* 1997). Alternatively, in certain conditions, TGM2-catalysed deamidation of gliadins is favoured over cross-linking. When glutamine residues are deamidated they become negatively-charged glutamate residues. Deamidation is particularly favoured over protein cross-linking in low pH conditions. For example, the pH of the small intestine is slightly acidic, which may lead to TGM2 catalysing deamidation reactions (favourable for celiac disease) more frequently (Fleckenstein *et al.* 2002).

Deamidated gliadin peptides bind to HLA DQ2 molecules with higher affinity than unmodified peptides (van de Wal *et al.* 1998, Arentz-Hansen *et al.* 2000), which may be extremely important in the pathogenesis of celiac disease. Higher affinity interactions would enable efficient presentation of gluten-derived peptides to the immune system, inducing the adaptive immune reactions observed in celiac disease (Lundin *et al.* 1993). On the other hand, cross-linking of TGM2 with gliadin peptides enables the presentation of this autoantigen as a hapten-carrier complex to the B cells of the immune system, putatively explaining the production of autoantibodies specific for celiac disease (Sollid *et al.* 1997) as described in chapter 2.4.2.

2.3. Innate immunity and celiac disease

Innate immunity constitutes the front line of defence against pathogenic agents. It is activated within minutes or hours after encountering the pathogen. Innate immunity does not rely on adaptive mechanisms or antigen-specific lymphocytes, allowing it to respond to threat immediately. The innate immune system is able to discriminate between self antigens and non-self antigens and it recognises a broad variety of pathogens. (Janeway *et al.* 2001).

There is evidence that the innate immune system is involved in celiac disease. It is thought to be raised against gliadin peptides and it precedes and enables the initiation of adaptive immune responses (Maiuri *et al.* 2000). Gliadin peptides have been shown to induce rapid expression of interleukin (IL) 15 in duodenal biopsy samples from patients with untreated celiac disease (Maiuri *et al.* 2000, Mention *et al.* 2003). Hue *et al.* (2004) showed that the

toxic p31-49 gliadin peptide induces the expression of the non-classical HLA molecule MICA on the surface of enterocytes in active celiac disease. Gliadin peptides may induce innate immune responses through the increased expression of IL15 and upregulation of MICA by enterocytes. Expression levels of MICA have been found to correlate with the severity of the disease (Hue *et al.* 2004).

MICA is known to be a ligand for the killer-cell lectin-like receptor subfamily K, member 1 (KLRK1, NKG2D). KLRK1 is expressed by a subset of CD8+ T cells, T cells using a gamma-delta T cell receptor, and most natural killer cells (Bahram *et al.* 2005). Intraepithelial lymphocytes in the small intestine of patients with celiac disease also express KLRK1: such cells are capable of lysing epithelial cells through a KLRK1-MICA interaction (Hue *et al.* 2004, Meresse *et al.* 2004). Following these innate immune reactions, entry of gliadin peptides into the intestinal lamina propria is proposed to stimulate the action of the adaptive immune response.

It has also been reported that gliadin increases enterocyte apoptosis (Giovannini *et al.* 2000, Giovannini *et al.* 2003), suggesting a direct cytotoxic effect of gliadin that may contribute to the villous atrophy observed in celiac disease. However, contradictory effects of gliadin on intestinal enterocytes have been demonstrated, as gliadin has also been reported to rapidly induce actin rearrangements in intestinal epithelial cell lines and to enhance proliferation of epithelial cells (Barone *et al.* 2007). These effects may lead to the maintenance of the typical atrophic and proliferative alterations of the small intestine in celiac disease.

2.4. Adaptive immunity and celiac disease

Adaptive immunity is characterised by high specificity to antigens, development of immunological memory and usually requires a few days to activate after encountering a novel antigen. Adaptive immunity is comprised of T cell-mediated immunity and humoral immunity (Janeway *et al.* 2001), both of which are activated in celiac disease.

T cell-mediated immunity involves antigen presentation through HLA molecules by antigen presenting cells. The encounter between a naïve T cell and an antigen-HLA complex on antigen presenting cells stimulates T cells to become effector cells in the presence of sufficient co-stimulatory conditions. HLA class I molecules, expressed on virtually all nucleated cells, present intracellular pathogens to CD8+ T cells, which then differentiate and kill infected target cells. In contrast, HLA class II molecules are expressed almost exclusively on antigen presenting cells, such as dendritic cells, and they present antigens derived from extracellular or intravesicular pathogens to CD4+ T cells. These CD4+ T cells can then differentiate into three types of effector T cells: T helper type 1 (Th1), T helper type 2 (Th2) or T helper type 17 (Th17) cells. (Janeway *et al.* 2001, Dardalhon *et al.* 2008).

Th1, Th2 and Th17 cells activate different immune responses through the cytokines they produce. Th1 responses are characterised by the production of IL2, interferon-gamma (IFNγ) and tumour necrosis factor beta (TNFβ), and they are involved in activating CD8+ T cells and macrophages, and inducing B cells to produce IgG antibodies for opsonisation. Th2 responses are characterised by the production of IL4, IL5, IL13 and IL25, and they initiate the humoral immune response by activating naïve B cells to produce IgM antibodies followed by IgA, IgE, and IgG antibody isotypes. Th17 cells are a recently identified T helper cell population. These

cells are characterised by the production of IL17, IL22 and IL21, as well as the expression of the IL23 receptor. The immunological role of Th17 cells is not yet completely understood, but they appear to promote organ-specific inflammation. (Janeway *et al.* 2001, Dardalhon *et al.* 2008).

The normal consequence of activation of adaptive immune responses is clearance of a foreign antigen from the body. Although effector T cells have crucial roles in the elimination of infectious pathogens, they are also involved in certain disease conditions. Th2 cells are known to have a role in allergic reactions, whereas Th1 and also Th17 cells appear crucial in the induction and maintenance of chronic inflammatory processes such as autoimmunity. In autoimmune conditions, adaptive immune responses have developed against self antigens, a situation in which the immune system is incapable of eliminating the antigen. If these harmful immune responses are left unchecked, inflammatory injury to the body's tissues may ensue. (Janeway *et al.* 2001, Dardalhon *et al.* 2008)

Adaptive immune responses are active in celiac disease in the small intestinal mucosa. CD4+ T cell-mediated immune responses to gluten peptides have been demonstrated in patients, and the Th1 response appears to be particularly important for the disease pathology.

2.4.1. HLA molecules and T cell stimulation by antigen presentation

HLA DQ2 and DQ8 heterodimers are HLA class II molecules and are required for the development of celiac disease. These heterodimers consist of an alpha and a beta chain, which are encoded for by different genes, namely *DQA1* and *DQB1*, respectively. The pairing of these chains creates an antigen binding groove. The antigen binding grooves of different HLA heterodimers can have different properties due to their differing amino acid compositions. HLA DQ2 and DQ8 heterodimers have a particularly high preference for negatively-charged amino acids, such as glutamic acid residues in deamidated gliadin peptides (van de Wal *et al.* 1998, Arentz-Hansen *et al.* 2000). The DQ2 and DQ8 molecules mediate the immune responses triggered by dietary gluten. Stimulation with gliadin peptides leads to proliferation of T cells in biopsies from patients with DQ2 and DQ8 heterodimers (Lundin *et al.* 1993, Lundin *et al.* 1994). Thus, gliadin peptides stimulate T cells through presentation by these DQ heterodimers.

Recent evidence suggests that the preference for negatively-charged peptides by the DQ8 heterodimer is not the only mechanism by which this molecule is involved in the pathogenesis of celiac disease. A certain positively-charged amino acid within the binding groove of the DQ8 heterodimer appears to promote the recruitment of T cell receptors with a negative charge at the complementary region during the response against native (not deamidated) gluten peptides presented by the DQ8 heterodimers (Hovhannisyan *et al.* 2008). This model provides an explanation of how the inflammatory responses are initiated in the mucosa of patients with celiac disease, although it is not yet known whether the DQ2 heterodimer has similar functions.

The production of Th1 cytokines can lead to mucosal destruction and epithelial apoptosis in the small intestine. The cytokines produced by DQ2-restricted gliadin-specific T cell clones are dominated by IFNγ (Nilsen *et al.* 1998), and the histological damage to mucosa in celiac disease appears to be mediated through cytokines, especially IFNγ-driven pathways (Przemioslo *et al.* 1995, Kontakou *et al.* 1995b). However, IFNγ activation on its own does

not explain the intestinal damage seen in patients with celiac disease, as IFNγ can also be upregulated in treated patients without causing pathology (Forsberg *et al.* 2002). In addition to IFNγ, other cytokines show upregulation in an *in vivo* gluten challenge, such as tumour necrosis factor alpha (TNFα), IL2 and IL6 (Kontakou *et al.* 1995a). IL18, which maintains Th1 responses by supporting IFNγ production, is also expressed in samples from patients with celiac disease, but not in healthy controls (Salvati *et al.* 2002). Finally, there is emerging evidence about the role of Th17 cells in autoimmunity (Dardalhon *et al.* 2008), and the cytokines produced by these cells may play a role in the pathogenesis of celiac disease as well (Castellanos-Rubio *et al.* 2009).

2.4.2. Antibodies in celiac disease

Initiation of adaptive immune responses occurs when T cells are activated by signals from antigen presenting cells. Generally, B cells responsible for the antibody production require stimulatory signals from the effector T cells to be activated. When B cells are activated, they become antibody-secreting plasma cells, producing antibodies that can bind to antigens such as pathogens outside the cells. The antibodies can neutralise, opsonise or activate the complement system to eradicate the antigen. The effector functions of the antibody depend on its isotype. In autoimmune conditions, antibodies against self-antigens are often produced. (Janeway *et al.* 2001)

Patients with celiac disease have two types of antibodies, which are specific for the disease. They have antibodies against deamidated gliadin peptides (Ankelo *et al.* 2007, Kaukinen *et al.* 2007, Niveloni *et al.* 2007) and they also display autoantibodies of the IgA class against TGM2 (van de Wal *et al.* 1998). The anti-TGM2 antibody is proposed to be raised against TGM2 molecules within cross-linked TGM2-gliadin complexes. TGM2-specific B cells are hypothesised to endocytose the TGM2-gliadin complexes. These molecules are then further processed in the B cells and presented to T cells via HLA class II molecules on the B cell. The presentation of gliadin by HLA DQ2 molecules activates gliadin-specific T cells, which then help TGM2-specific B cells to start antibody production (Sollid *et al.* 1997). Dietary gluten is thus responsible for the celiac disease-specific antibodies: the antibodies disappear when gluten is excluded from the diet (Dieterich *et al.* 1997).

The antibodies against TGM2 are produced in the small intestinal mucosa of patients with celiac disease (Marzari *et al.* 2001). These antibodies are present at the site of their production and also found circulating in the blood. In untreated patients with celiac disease, the anti-TGM2 antibodies can also be found as deposits in the small intestinal mucosa on extracellular TGM2 below the epithelial layer and around blood vessels (Korponay-Szabó *et al.* 2004, Kaukinen *et al.* 2005). Even patients with early stage celiac disease who do not yet have villous atrophy have been shown to have these anti-TGM2 antibody deposits (Salmi *et al.* 2006), often before detectable levels of serum antibodies are seen.

It is still unclear whether the celiac disease-specific autoantibodies play a role in the pathogenesis of the disease. However, there is mounting support of this idea (Lindfors *et al.* 2009). Celiac disease-specific IgA antibodies have been shown to disturb angiogenesis (Myrsky *et al.* 2008) and inhibit the differentiation of intestinal epithelial cells *in vitro* (Halttunen and Mäki 1999). A subset of anti-TGM2 IgA antibodies have also been reported to

be functionally active, increasing epithelial permeability in patients with untreated celiac disease (Smecuol *et al.* 1997, Zanoni *et al.* 2006). Furthermore, there is evidence that celiac disease specific autoantibodies are able to activate monocytes (Zanoni *et al.* 2006) and induce monocyte-mediated antibody-dependent cytotoxicity (Saalman *et al.* 1998). Thus, the activation of monocytes by the anti-TGM2 antibodies may participate in the pathogenesis of celiac disease.

The adaptive and innate immune responses in celiac disease are summarised in Figure 2.

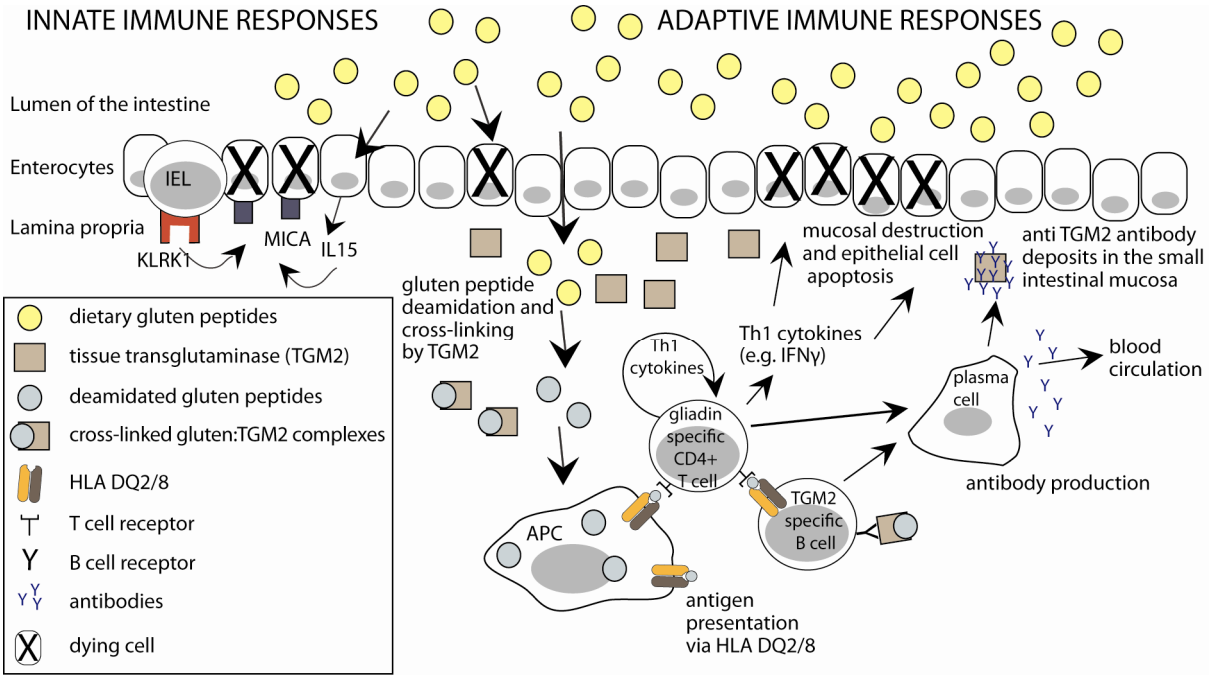

Figure 2 *A schematic picture of the innate (on the left) the adaptive (on the right) immune responses in celiac disease. Both responses are triggered by dietary gluten. In innate immune responses, gluten has direct cytotoxic effects on the enterocytes of the small intestinal epithelium. It also triggers the IL15, MICA and KLRK1 –mediated cell-lysis. The adaptive immune responses are driven by deamidated gluten peptides. TGM2 is responsible for the deamidation and is also able cross-link with gluten peptides. The deamidated gluten peptides are presented efficiently to the CD4+ T cells by HLA DQ2 and DQ8 molecules. The activation of T cell-mediated immune responses leads to inflammation in the small intestine. The cross-linking of TGM2 with gluten enables the activation of TGM2-specific B cells and the production of anti-TGM2 antibodies. IEL: intraepithelial lymphocyte, APC: antigen presenting cell.*

3. Genetics of celiac disease

The mode of inheritance of celiac disease is not known, although there is strong evidence for genes playing an important role in the disease susceptibility. Family and twin studies can be used in estimating the proportion of genetic and environmental risk factors in the disease prevalence. The studies have shown that genetic components play a major role in the induction and manifestation of celiac disease. Depending on the study, the proband-wise concordance for celiac disease is 75-86% between monozygotic twins (Hervonen *et al.* 2000, Walker-Smith 1973, Bardella *et al.* 2000, Greco *et al.* 2002, Nistico *et al.* 2006), whereas the proband-wise concordance of celiac disease between dizygotic twins has been estimated to be 16.7-20% (Greco *et al.* 2002, Nistico *et al.* 2006). The difference in concordances between monozygotic twins and dizygotic twins provides an estimate of the size of the genetic component in celiac disease, which appears higher than in many other complex immunological disorders. For example, in type 1 diabetes the proband-wise concordance is 23% for monozygotic twins and 5% for dizygotic twins (Kaprio *et al.* 1992). The concordance of celiac disease between ordinary siblings and dizygotic twins, both sharing approximately 50% of their genes, is at the same level, indicating that a shared environment (apart from gluten ingestion) may only have a minor effect on the concordance in dizygotic twins (Greco *et al.* 2002, Bevan *et al.* 1999). According to Nistico *et al.* (2006), the heritability of celiac disease is up to 87%.

Family studies have shown that the prevalence of celiac disease is 10% among first-degree relatives of gluten intolerant patients (Dube *et al.* 2005). The familiar clustering of celiac disease can be shown as a ratio of the prevalence of celiac disease between siblings of patients with celiac disease, and the population (λs). λs is estimated to be 30-60 in celiac disease, further confirming the existence of a large genetic component in celiac disease (Risch 1987, Petronzelli *et al.* 1997).

Certain common *HLA* alleles are known to be important genetic risk factors for celiac disease. Despite their crucial role in pathogenesis, their effect on the genetic susceptibility for celiac disease seems to be minor, indicating that other *HLA*-unlinked susceptibility genes for celiac disease exist (Bevan *et al.* 1999, Risch 1987, Petronzelli *et al.* 1997). *HLA* haplotype sharing studies in families have shown that the contribution of the *HLA* region to the development of celiac disease among siblings is less than 40% (Bevan *et al.* 1999, Petronzelli *et al.* 1997). These results indicate that the remaining genetic risk is likely to be conferred by non-*HLA* genes (Bevan *et al.* 1999, Petronzelli *et al.* 1997). There is abundant evidence for several *HLA*-unlinked susceptibility genes for celiac disease, with each gene conferring only a minor effect on the disease risk.

3.1. Molecular genetics

The genetic information of humans is stored in the cell nuclei in 23 chromosome pairs and also within the mitochondria. The genetic information is composed of deoxyribonucleic acid (DNA), which is an antiparallel double-helical macromolecule consisting of pairs of 4 nucleotides (adenine [A], cytosine [C], guanine [G], and thymine [T]) forming an organised linear sequence (Strachan and Read 2004). DNA contains the genetic instructions that are

used in the development and function of cells and organisms. Differences between individuals are largely due to different types of genetic polymorphisms, i.e. differences in the DNA. Most of these polymorphisms are part of normal genetic variation, whereas some of them are known to directly cause diseases, or they may increase the risk of being affected by certain disease conditions - such as celiac disease. Using molecular genetics tools, such as genetic linkage and association approaches, it is possible to locate and identify these variations.

3.1.1. HLA genes

The MHC region on human chromosome 6p21.3 is the most polymorphic region in the entire human genome. The allelic diversity of certain genes is extremely high, e.g. the HLA-B gene has more than 1100 known alleles and the *HLA-DRB1* gene has more than 600 known alleles (IMGT/HLA Database: http://www.ebi.ac.uk/imgt/hla/allele.html - 12.02.2009) (Robinson *et al.* 2003). The celiac disease-associated *DQB1* and *DQA1* genes have 96 and 35 alleles, respectively. The MHC region harbours several genes that have crucial roles in controlling and regulating immune responses. The *HLA* genes can be divided into three classes according to their main functions (Figure 3): *HLA* class I and II genes are responsible for presenting antigens; *HLA* class III genes comprise a more heterogenic group of various, mostly immune-related genes.

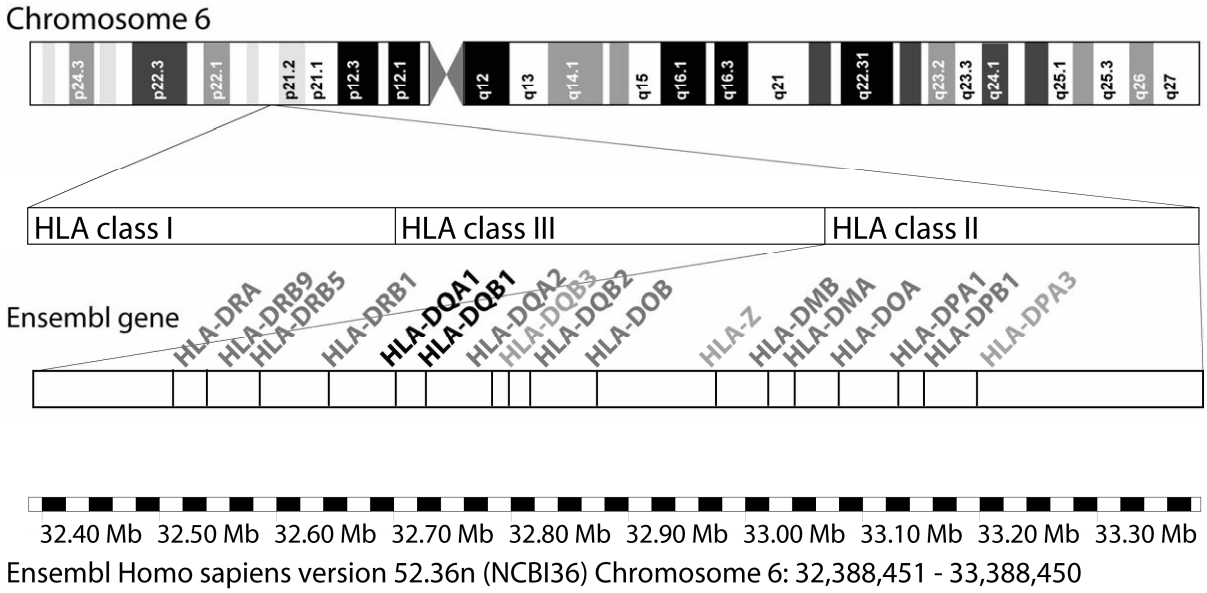

Figure 3 *The HLA region on chromosome 6 and HLA class II genes. The celiac disease-associated DQA1 and DQB1 genes are highlighted with black. HLA-DQB3, HLA-Z and HLA-DPA3 are pseudogenes.*

HLA genes are associated with several autoimmune and inflammatory disorders (McDevitt 1998, Reveille 2006) such as ankylosing spondylitis (with *B27*), narcolepsy (with *DQB1*0602*), type 1 diabetes (with *DRB1*0301-DQB1*0201* and *DRB1*04-DQB1*0302*), rheumatoid arthritis (with *DRB1*0401/0404/0405/0101*), systemic lupus erythematosus (with *DRB1*0301/*1501/*1503/*08*), Sjögren's syndrome (with *DRB1*0301*), systemic sclerosis (with *DRB1*11/*0301/*1502*), and autoimmune thyroid disease (Graves' disease and

autoimmune thyroiditis) (with *DRB1*0301-DQA1*0501-DQB1*0201* haplotype). Some of these associations are relatively weak, whereas for other diseases the particular *HLA* haplotype is always required for disease onset. Certain *HLA* haplotypes show a particularly strong association with celiac disease, which was demonstrated as early as 1972 by Falchuk and Strober, and by Stokes *et al*. The strongest association signal was subsequently refined to the *HLA* class II molecules DR3 and DQ2 (Keuning *et al.* 1976, Solheim *et al.* 1976, Tosi *et al.* 1983).

Approximately 90% of European Caucasians with celiac disease carry the HLA-DQ2 heterodimer coded by the *DQA1*05* and *DQB1*02* alleles (Sollid *et al.* 1989, Karell *et al.* 2003). A majority of the patients carry the DQ2.5 (or *DR3-DQ2*) haplotype, where the alpha and beta chains of the DQ2 heterodimer are encoded together *in cis* on a *DRB1*03* haplotype (including the *DRB1*03, DQA1*0501* and *DQB1*0201* alleles). The DQ2 heterodimer can also be encoded in *in trans* configuration by the DQ2.2 (*DR7-DQ2*) and DQ7 (*DR5/6-DQ7*) haplotypes, with the *DQA1*05* allele deriving from *DRB1*11, *12 or *13* haplotypes (*DRB1*11/12/13, DQA1*0505, DQB1*0301*) and the *DQB1*02* allele deriving from a *DRB1*07* haplotype (*DRB1*07, DQA1*02, DQB1*0202*) (Mazzilli *et al.* 1992, Sollid and Thorsby 1993).

The DQ8 heterodimer encoded by the DR4-DQ8 haplotype (*DRB1*04, DQA1*03, DQB1*0302*) is common in celiac patients who do not carry the DQ2 heterodimer (Spurkland *et al.* 1992). As explained in Chapter 2.4.1, both DQ2 and DQ8 heterodimers have a central role in the pathogenesis of celiac disease (Molberg *et al.* 1998). Only a small number of patients do not carry DQ2 or DQ8, and the vast majority of these patients carry one chain of the DQ2 heterodimer - i.e. encoded by either the DQ2.2 or DQ7 haplotype (Karell *et al.* 2003, Polvi *et al.* 1998). The *HLA* haplotypes and heterodimers conferring risk to celiac disease are illustrated in Figure 4.

The HLA DQ2 and DQ8 heterodimers are necessary but not sufficient, on their own, for the development of celiac disease, as they are common also in the healthy population (Sollid *et al.* 1989, Polvi *et al.* 1996). This indicates that there are other risk genes for celiac disease in the human genome. Several attempts have been made to identify the non-HLA risk genes and genetic variants for celiac disease. Along with increasing knowledge about variation in the human genome, as well as innovations to genotyping methodology, our understanding of the genetic factors predisposing to celiac disease is growing.

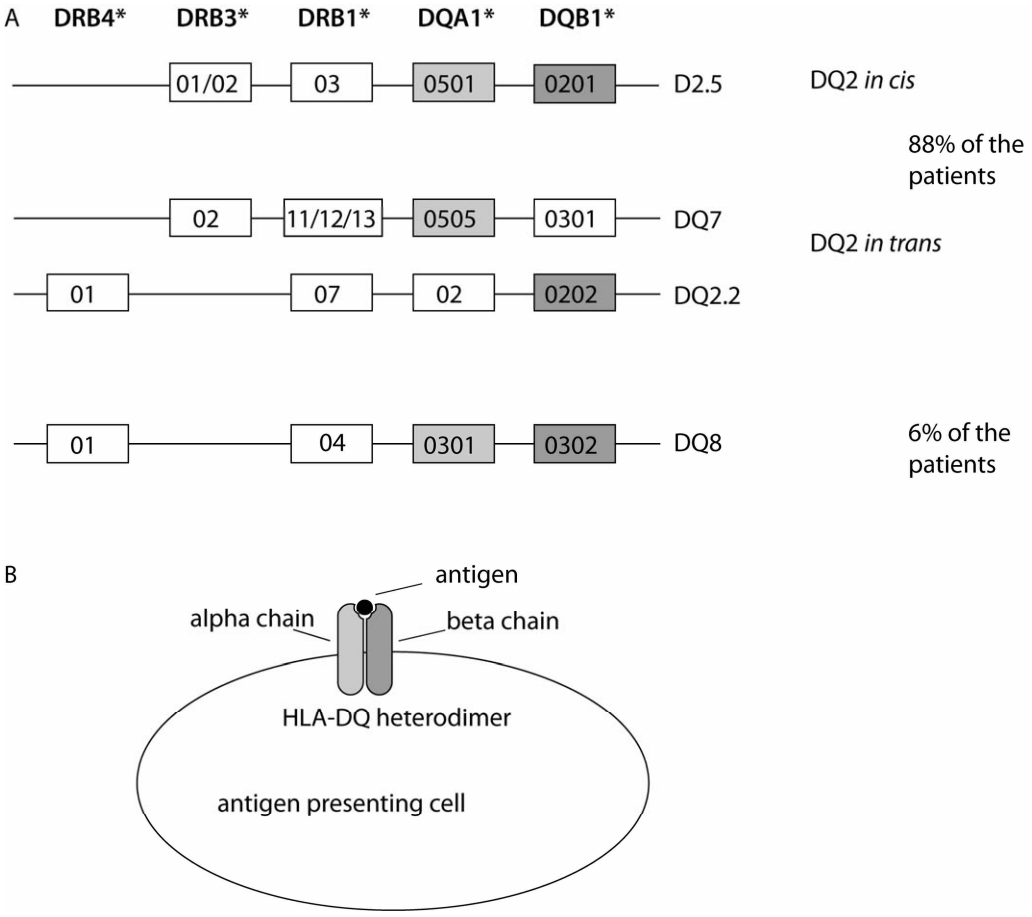

Figure 4 ***A.*** *HLA haplotypes conferring risk to celiac disease. The DQ2 heterodimer (coded by DQA1*05 and DQB1*02) can derive either in cis or in trans configuration. The frequencies shown are taken from Karell et al. (2003).* ***B.*** *The DQ heterodimer consists of an alpha chain and a beta chain encoded by the DQA1 and DQB1 genes, respectively.*

3.1.2. Genome-wide linkage studies and positional candidate genes

The genome-wide linkage study approach is a method to locate disease susceptibility loci, and this method has been highly successful in identifying risk loci for monogenic, Mendelian, diseases. The method employs families with at least two affected individuals, genetic markers (e.g. microsatellites or single-nucleotide polymorphisms [SNP]), and recombination fractions or shared chromosomal segments between the affected siblings or other family members. This method works best for rare disease alleles with relatively high penetrance, and it is robust against allelic heterogeneity. No prior hypothesis or knowledge is needed in relation to which chromosome the risk-conferring genes may be located because the whole genome is under inspection. Encouraged by findings for Mendelian disorders, the linkage method has been widely used in attempts to locate risk genes for diseases with complex aetiology.

The first genome-wide linkage scan in celiac disease was published in 1996 (Zhong *et al.* 1996). Since then, 12 other whole-genome linkage scans have been performed (Greco *et al.* 1998, King *et al.* 2000, Naluai *et al.* 2001, Liu *et al.* 2002, Woolley *et al.* 2002, Popat *et al.* 2002a, Neuhausen *et al.* 2002, Van Belzen *et al.* 2003, van Belzen *et al.* 2004, Rioux *et al.* 2004, Garner *et al.* 2007, Ding *et al.* 2008). With the exception of the MHC locus, the results

of the linkage scans have been somewhat contradictory. However, a number of chromosomal regions have been repeatedly highlighted, e.g. 5q31-33 (*CELIAC2*) (Greco *et al.* 1998, Naluai *et al.* 2001, Liu *et al.* 2002), 2q32 (*CELIAC3*) (King *et al.* 2000, Naluai *et al.* 2001, Rioux *et al.* 2004) and 19p13 (*CELIAC4*) (Popat *et al.* 2002a, Van Belzen *et al.* 2003, van Belzen *et al.* 2004) (Table 1).

***CELIAC2* locus on chromosome 5q31-q33**

The *CELIAC2* locus has been linked with celiac disease in genome-wide linkage scans in the Italian, Swedish-Norwegian and Finnish populations (Greco *et al.* 1998, Naluai *et al.* 2001, Liu *et al.* 2002). This region also showed linkage in a Finnish candidate locus study (Holopainen *et al.* 2001). A meta- and pooled-analysis performed in Italian, Finnish, Swedish, Norwegian, French and UK populations revealed strong linkage of this region with celiac disease, suggesting that the region is the strongest non-*HLA* susceptibility locus for celiac disease (Babron *et al.* 2003). The locus harbours several genes related to immune responses, such as the Th2 cytokine cluster and interleukin genes, making it highly interesting for celiac disease and other inflammatory diseases. Chromosome 5q31-q33 has also been associated with several other autoimmune and inflammatory disorders, such as Crohn's disease, type 1 diabetes, asthma, psoriasis and rheumatoid arthritis, further emphasising its role in inflammatory conditions (Rioux *et al.* 2001, Morahan *et al.* 2001, Cookson and Moffatt 2000, Samuelsson *et al.* 1999, Tokuhiro *et al.* 2003). Nevertheless, no consistent gene associations have been found in the region to date.

Two systematic fine-mapping studies of the *CELIAC2* locus have been published recently (Amundsen *et al.* 2007, Adamovic *et al.* 2008b). The aim of these studies was to identify genetic variants responsible for the previously identified linkage signal in the Scandinavian population (Naluai *et al.* 2001). Adamovic *et al.* (2008b) reported strong evidence for linkage of the locus with celiac disease, and modest association was also demonstrated. Amundsen *et al.* (2007) found several single- and multi-point associations albeit at modest significance levels. However, no risk variants were confirmed in these two partly overlapping studies, implying an incomplete coverage of the 5q31-q33 region and highlighting the need for studies with greater statistical power.

***CELIAC3* locus on chromosome 2q32**

The *CELIAC3* locus on chromosome 2q32 has shown linkage with celiac disease and several other autoimmune diseases (Gough *et al.* 2005). This region harbours the *CTLA4, ICOS* and *CD28* genes, all of which belong to the CD28 family of T cell co-stimulatory receptors expressed on T cells following activation. *CD28* is required for the activation of T cells (Yamada *et al.* 2002). In contrast, *CTLA4* is a negative regulator of T cell responses and may be required for the induction of tolerance (Yamada *et al.* 2002). It mediates the effects of regulatory T cells known to be important in suppressing immune responses (Rudd 2009). *ICOS* is necessary for the maintenance of humoral immunity, as it is involved in antibody isotype-switching and also induces the secretion of Th2 cytokines (Shilling *et al.* 2006).

A number of studies have suggested that *CTLA4* and *ICOS* are the strongest candidate genes in the region. Despite several fine-mapping attempts, the risk-conferring genetic factor still remains unknown. Association of the *CTLA4* region and celiac disease has been reported in several populations (Popat *et al.* 2002c, Hunt *et al.* 2005, Brophy *et al.* 2006). However, *ICOS* appears to give the strongest linkage and association signals in the Finnish population

(Holopainen *et al.* 2004, Haimila *et al.* 2004, Haimila *et al.* 2009). The linkage disequilibrium (LD) pattern in the region is complex, consisting of two main haplotype blocks: one covering *CTLA4* and the first exon of *ICOS*, and the other spanning the rest of *ICOS*. This makes it challenging to pinpoint the risk gene using genetic approaches. Population heterogeneity in the region may also explain the inconsistent results.

***CELIAC4* locus on chromosome 19p13**

The chromosomal region 19p13 (*CELIAC4*) has shown evidence for genetic linkage with celiac disease, with the strongest reported linkage signal coming from the Dutch population (Van Belzen *et al.* 2003, van Belzen *et al.* 2004). A subsequent fine-mapping study comprised of 463 cases and 686 controls from the Dutch population revealed association of five SNPs in the 3' end of the myosin IXB gene (*MYO9B*) with celiac disease (Monsuur *et al.* 2005). The function of *MYO9B*, also known as unconventional myosin, is relatively poorly understood. It is a single-headed processive motor that contains a Rho-GTPase-activating protein domain (O'Connell and Mooseker 2003) similar to genes involved in tight junction functions (Matter and Balda 2003, Bruewer *et al.* 2004). Thus, it is possible that variations in *MYO9B* play a role in the impairment of epithelial permeability of the small intestine (Matter and Balda 2003), leading to increasing amounts of gluten entering the lamina propria. The *MYO9B* association with celiac disease has also been replicated in the Spanish population (Sanchez *et al.* 2007), as well as in other pathologies, such as inflammatory bowel disease (van Bodegraven *et al.* 2006, Nunez *et al.* 2007a, Cooney *et al.* 2009), schizophrenia (Jungerius *et al.* 2008), refractory celiac disease (Wolters *et al.* 2007), systemic lupus erythematosus (Sanchez *et al.* 2007), rheumatoid arthritis (Sanchez *et al.* 2007) and type 1 diabetes (Santiago *et al.* 2008). Despite these numerous associations, *MYO9B* still remains a controversial candidate gene for celiac disease and inflammatory bowel disease because several replication studies in different populations have not found association (Hunt *et al.* 2006, Amundsen *et al.* 2006a, Amundsen *et al.* 2006b, Giordano *et al.* 2006, Nunez *et al.* 2006, Latiano *et al.* 2007, Cirillo *et al.* 2007). Such variable results could be explained by differences in allele frequencies in different populations, lower risk effect than estimated in the original association study, or differences in LD patterns within the gene region. In addition, the possibility of other risk gene(s) in the region cannot be excluded because *MYO9B* is in LD with genes located both upstream and downstream.

***CELIAC5* locus on chromosome 15q11-q13**

Given that the results of the genome-wide linkage scans had been somewhat contradictory, perhaps due to genetic heterogeneity between different populations, Woolley *et al.* (2002) performed a genome-wide linkage scan in a relatively homogenous subpopulation to decrease the effects of underlying allelic and locus heterogeneity. The subpopulation originated from North-Eastern Finland, from the Koilliskaira region. Genealogic studies had confirmed that the families selected for this study shared a common ancestor in the 16th century, in a region which belongs to the so-called late settlement region in the Finnish population history. Due to a small founder population settling in these regions, leading to genetic drift and bottlenecks in Eastern and Northern Finland, several rare disease genes have become enriched in the region. Likewise, many diseases that are common in other parts of Europe are absent there. The population structure in these regions, representing a relatively homogenous genetic background, has been ideal in identifying genes causing monogenic diseases. The genetic

background of homogenous regional isolates with long blocks of LD is also thought to be useful in dissecting diseases with a more complex genetic background (Peltonen *et al.* 1995). Indeed, there are some examples where this approach has been successful (Laitinen *et al.* 2004).

Woolley *et al.* (2002) identified a region on chromosome 15q11-q13 (*CELIAC5*) as a potential genetic susceptibility locus for celiac disease. This region still requires confirmation as a true susceptibility locus for celiac disease in larger populations.

In general, the findings of genome-wide linkage scans in celiac disease mirror the results from genome-wide linkage scans performed for other complex diseases. The linkage scans have highlighted several chromosomal regions of the human genome, but replication of the original results has often proved to be challenging and further fine-mapping studies have frequently not revealed genetic associations that completely explain the original observed linkage signals. (Risch 2000).

3.1.3. Functional candidate genes

Studies on functional candidate genes attempt to test whether a particular gene with a known function is genetically linked or associated with the disease of interest. The candidate gene can be selected based on existing knowledge of its possible role in disease pathogenesis. A gene can also be selected as a candidate gene based on known associations with similar or co-occurring diseases, which imply shared genetic risk factors. In addition, genes showing interesting expression patterns, e.g. in studies with whole-genome expression arrays, are often tested for genetic association.

Several candidate gene studies have been performed in celiac disease, with particular interest in genes having a known immunological function (Lopez-Vazquez *et al.* 2002, Rueda *et al.* 2004, Curley *et al.* 2006, Abel *et al.* 2006, Santin *et al.* 2007, Wapenaar *et al.* 2007, Nunez *et al.* 2007b, Castellanos-Rubio *et al.* 2008, Chernavsky *et al.* 2008). These studies include genes that are related to both innate and adaptive immunity. Combined positional and functional candidate gene studies have also been used in celiac disease, applying both knowledge relating to the function of the gene and its location within candidate susceptibility loci for the disease. Epithelial barrier integrity is impaired in the small intestine of patients with celiac disease, so genes known to relate to tight junction assemblies have also been tested as susceptibility candidates (Wapenaar *et al.* 2008). Similarly, genes that have shown association with other autoimmune or inflammatory disorders have been tested for association with celiac disease (Alizadeh *et al.* 2007, Santin *et al.* 2008, Nunez *et al.* 2008). Recent findings of celiac disease candidate gene studies are summarised in Table 2.

Many of the candidate gene studies do not confirm associations to the tested variants, or only weak evidence for association is found. The lack of confirmed association through this experimental approach may reflect the insufficient knowledge we have about gene functions and/or biochemical pathways involved in the pathogenesis of celiac disease. The inability to replicate findings may also be caused by overestimations of possible effect sizes, especially in studies conducted using relatively small samplesets. Replication of reported associations in different samplesets and populations are essential for validating novel findings.

Table 1 *Chromosomal regions highlighted in the genome-wide linkage scans and subsequent fine-mapping studies.*

Locus	Chromosome	Candidate genes highlighted in finemapping studies	OR (if calculated)	References	Other diseases
CELIAC2	5q31-q33	several candidate genes with immunological functions		*Linkage / association:* Holopainen *et al.* 2001, Greco *et al.* 1998, Naluai *et al.* 2001, Liu *et al.* 2002, Babron *et al.* 2003, Amundsen *et al.* 2007, Adamovic *et al.* 2008b	Crohn's disease: Rioux *et al.* 2001, T1D: Morahan *et al.* 2001, asthma: Cookson and Moffatt 2000, PSO: Samuelsson *et al.* 1999, RA: Tokuhiro *et al.* 2003
CELIAC3	2q32	*CTLA4*	1.41 (1.16-1.73) (Hunt *et al.* 2005)	*Linkage / association:* Popat *et al.* 2002c, Hunt *et al.* 2005, Brophy *et al.* 2006, Popat *et al.* 2002b	Autoimmunity: review in Gough *et al.* 2005, IgA deficiency: Haimila *et al.* 2009
		ICOS	1.98 (1.22-3.19) (Haimila *et al.* 2004)	*Linkage / association:* Holopainen *et al.* 2004, Haimila *et al.* 2004, Haimila *et al.* 2009	
CELIAC4	19p13.1	*MYO9B*	1.56 (1.27-1.93) (Monsuur *et al.* 2005)	*Linkage / association:* Van Belzen *et al.* 2003, van Belzen *et al.* 2004, Monsuur *et al.* 2005, Sanchez *et al.* 2007, Refractory CD: Wolters *et al.* 2007. *NO association:* Hunt *et al.* 2006, Amundsen *et al.* 2006a, Giordano *et al.* 2006, Latiano *et al.* 2007, Cirillo *et al.* 2007, Núñez *et al.* 2006	schizophrenia: Jungerius *et al.* 2008, IBD: van Bodegraven *et al.* 2006, Nunez *et al.* 2007a, Cooney *et al.* 2009, *NO association:* Amundsen *et al.* 2006b, SLE: Sanchez *et al.* 2007, RA: Sanchez *et al.* 2007, T1D: Santiago *et al.* 2008
CELIAC5	15q11-q13			*Linkage:* Woolley *et al.* 2002	

OR: odds ratio, T1D: type 1 diabetes, PSO: psoriasis, IBD: inflammatory bowel disease, SLE: systemic lupus erythematosus, RA: rheumatoid arthritis.

Table 2 *Recent candidate gene studies in celiac disease.*

Chromosome	Gene	Full gene name	OR (95% C.I.)	Number of patients/families	Reference	Replication studies in celiac disease	Other diseases showing association
1q21–q24	*FcgRIIa*	Fc fragment of IgG, low affinity IIa, receptor (CD32)	1.37 (1.18-1.58)	519 cases, 1359 controls	Alizadeh *et al.* 2007	*No association*: Adamovic *et al.* 2008a, Sareneva *et al.* 2009	SLE: Karassa *et al.* 2002, T1D: Alizadeh *et al.* 2007
1p13	*PTPN22*	protein tyrosine phosphatase, non-receptor type 22	1.82 (1.1-3.0)	262 cases, 214 controls	Santin *et al.* 2008		T1D: Bottini *et al.* 2004, Graves' disease: Velaga *et al.* 2004 , RA: Begovich *et al.* 2004, SLE: Kyogoku *et al.* 2004
1p31	*IL23R*	interleukin 23 receptor	1.57 (1.06-2.32)	598 cases, 546 controls	Nunez *et al.* 2008	*No association:* Weersma *et al.* 2008, *Linkage but no association: Einarsdottir et al. 2009*	*IBD:* Duerr *et al.* 2006, Weersma *et al.* 2008, Lappalainen *et al.* 2008, Einarsdottir *et al.* 2009, MS: Nunez *et al.* 2008
2q12-q22	*IL1B* (in interaction with *TNFA* on 6p21.3)	interleukin 1 beta (in interaction with tumor necrosis factor alpha)	29.59 (1.79-489.46)	228 cases, 244 controls	Chernavsky *et al.* 2008		
2q33	*SERPINE2*	serpin peptidase inhibitor, clade E	0.48 (0.34-0.67)	262 cases, 214 controls	Castellanos-Rubio *et al.* 2008	*No association:* Hunt *et al.* 2008	
5q31	*SLC22A5*	solute carrier family 22, member 5	1.39 (1.03-1.88)	309 cases, 358 controls	Curley *et al.* 2006		Crohn´s disease: Peltekova *et al.* 2004
6p12	*MICA*	MHC class I polypeptide-related sequence A	12.5 (4.66-33.11)	54 cases, 116 controls	Lopez-Vazquez *et al.* 2002	*Association:* Rueda *et al.* 2003	
7q21	*MAGI2*	membrane associated guanylate kinase, WW and PDZ domain containing 2	1.19 (1.08-1.32)	1215 cases, 2114 controls	Wapenaar *et al.* 2008		UC: Wapenaar *et al.* 2008

9q34	*PPP6C*	protein phosphatase 6, catalytic subunit	1.70 (1.31-2.20)	262 cases, 214 controls	Castellanos-Rubio *et al.* 2008	*No association:* Hunt *et al.* 2008	
	PBX3	pre-B-cell leukemia homeobox 3	2.60 (1.63-4.13)	262 cases, 214 controls	Castellanos-Rubio *et al.* 2008	*No association:* Hunt *et al.* 2008	
10p11	*PARD3*	par-3 partitioning defective 3 homolog (C. elegans)	1.23 (1.11-1.37)	1215 cases, 2114 controls	Wapenaar *et al.* 2008		
12q14	*IFNγ*	interferon, gamma	1.67 (1.13-2.47)	220 families, 330 cases, 499 controls	Rueda *et al.* 2004	*No association:* Wapenaar *et al.* 2004	
19p13	*CYP4F3*	cytochrome P450, family 4, subfamily F, polypeptide 3	1.77 (1.03-3.05)	309 cases, 358 controls	Curley *et al.* 2006		
	CYP4F2	cytochrome P450, family 4, subfamily F, polypeptide 2	1.33 (1.06-1.66)	309 cases, 358 controls	Curley *et al.* 2006		
	ICAM-1	intercellular adhesion molecule 1	4.2 (2.3-7.5)	180 cases, 212 controls	Abel *et al.* 2006		IBD: Papa *et al.* 2004
19q13	*KIR2DL5B*	killer cell immunoglobulin-like receptor, two domains, long cytoplasmic tail, 5B	2.60 (1.21-6.19)	343 cases, 160 controls	Santin *et al.* 2007		
22q11	*MIF*	macrophage migration inhibitory factor	1.33 (1.00-1.76)	531 cases, 887 controls	Nunez *et al.* 2007b		RA: Baugh *et al.* 2002, SLE: Donn *et al.* 2001, PSO: Donn *et al.* 2004

OR: odds ratio, 95% C.I.: 95% confidence interval, SLE: systemic lupus erythematosus, T1D: type 1 diabetes, RA: rheumatoid arthritis, IBD: inflammatory bowel disease, UC: ulcerative colitis, MS: multiple sclerosis, PSO: psoriasis

3.1.4. Genome-wide association studies

The sequencing of the entire human genome, completed at the beginning of the millennium, has opened up new possibilities for understanding the structure and variations of the human genetic material (Lander *et al.* 2001, Venter *et al.* 2001). This information can be used for studying susceptibility variants for human diseases and for asking more targeted questions in gene mapping. Another important approach for understanding the architecture of the human genome is the International HapMap Project, which aims to unravel the LD patterns and common variation of the human genome in different populations (International HapMap Consortium 2003, International HapMap Consortium *et al.* 2007). Information about conserved haplotype blocks in the genome can be used in capturing most of the common patterns of variation within the genome, providing an essential tool for the study of genetic susceptibility factors for diseases.

Together with increasing amounts of information on genomic variation, new genotyping methods have been developed that enable simultaneous and fast genotyping of hundreds of thousands of SNPs in a large number of samples. These innovations have given rise to genome-wide association studies, where up to one million genetic markers can be genotyped. Such an approach allows the whole human genome to be simultaneously tested as disease susceptibility candidates. Genome-wide association studies have opened up a new era in the genetic research of complex diseases, with novel findings that were not detected using the genome-wide linkage approach.

Genome-wide association studies are an ideal way to study common risk variants for common diseases. Given that the effect size of identified risk variants has been relatively low in many of diseases, large sample materials are required to obtain sufficiently high statistical power for analysis. Typically, at least several hundred cases and controls are used in these studies. By the end of February 2009, 261 genome-wide association studies had been published, reporting associations to 1183 SNPs in 178 diseases and traits (Hindorff *et al.* 2009) (http://www.genome.gov/26525384). In most of the common diseases, genome-wide association studies have revealed several independent loci with low effect sizes, suggesting that genetic susceptibility for common diseases is a result of joint effects of various risk-conferring variants.

The first genome-wide association study in celiac disease was published in 2007. It was performed to identify novel risk factors in the British population and examined 778 unrelated cases and 1422 population controls (van Heel *et al.* 2007). The strongest findings detected in the scan at the non-*HLA* regions were replicated in the Dutch (508 cases, 929 controls) and Irish (486 cases, 560 controls) populations. This study identified and confirmed the *KIAA1109-TENR-IL2-IL21* gene cluster on chromosome 4q27 (*CELIAC6*) as a risk locus for celiac disease (van Heel *et al.* 2007). This association with celiac disease has recently been replicated in the Swedish-Norwegian (Adamovic *et al.* 2008a) and Italian populations (Romanos *et al.* 2009); the association has also been demonstrated in other autoimmune and inflammatory diseases, such as type 1 diabetes, rheumatoid arthritis and ulcerative colitis (Todd *et al.* 2007, Zhernakova *et al.* 2007, Albers *et al.* 2009, Festen *et al.* 2009). Interestingly, a previous meta-analysis combining genome-wide linkage studies from the Italian, British, Swedish-Norwegian and Finnish populations also showed evidence for linkage to this locus (Babron *et al.* 2003).

Due to the extensive LD in the 4q region, the risk-causing variant associated with celiac disease still remains to be determined. This LD block contains three known protein-coding genes: *TENR, IL2, IL21*, and a predicted gene *KIAA1109* with an unknown function. The *TENR* gene is expressed mainly in the testis (Schumacher *et al.* 1995), making it an unlikely risk gene for intestinal inflammatory diseases. In contrast, KIAA1109 is expressed in multiple tissues (NCBI UniGene EST Profile Viewer: http://www.ncbi.nlm.nih.gov/sites/entrez) and cannot be excluded as a candidate for celiac disease. IL2, formerly known as T cell growth factor, is secreted by antigen stimulated T cells and plays an important role in T cell and B cell activation and proliferation (Lowenthal *et al.* 1985, Smith 1988). IL21 has a role in enhancing the proliferation and IFNγ production of T cells, B cells and natural killer cells (Kasaian *et al.* 2002, Brandt *et al.* 2007, Ettinger *et al.* 2005), and it has also been implicated in the Th17 pathway (Korn *et al.* 2007). These important immunological functions make *IL2* and *IL21* highly interesting candidate genes for celiac disease, as well as other immunological disorders. However, further investigations are required to determine the functional risk variants within this LD block.

Intriguingly, the two strongest non-*HLA* susceptibility loci for celiac disease identified in linkage studies (*CELIAC2* and *CELIAC3*) showed only weak evidence for association in the 2007 genome-wide association study. The *MYO9B* region on chromosome 19p13 (*CELIAC4*) showed no association signal at all.

The genome-wide association study by van Heel *et al.* (2007) was followed-up by a study of >1000 of the most strongly associated non-*HLA* SNPs and was performed in the UK (719 cases, 1561 controls), Irish (416 cases, 957 controls) and Dutch (508 cases, 888 controls) populations. This study revealed several loci that are likely to be important for celiac disease susceptibility (Hunt *et al.* 2008). Seven previously unknown risk regions were identified (Table 3); six of them harbouring genes controlling immune responses (Hunt *et al.* 2008).

The most significantly associated region outside the *HLA* region and the previously identified *KIAA1109-TENR-IL2-IL21* region was located within an LD block containing the regulator of G-protein signalling 1 (*RGS1*) gene on chromosome 1q31 (CELIAC7) (Hunt *et al.* 2008). The *RGS1* association has recently been replicated in Italian patients with celiac disease (Romanos *et al.* 2009), and in British patients with type 1 diabetes (Smyth *et al.* 2008). Genes belonging to the RGS family are involved in attenuating the signalling activity of G proteins by acting as GTPase-activating proteins. RGS1 is known to regulate chemokine receptor signalling and it is involved in B cell activation and proliferation (Kehrl 1998). Hunt *et al.* (2008) detected *RGS1* expression in small intestinal biopsies, providing further evidence for the involvement of *RGS1* in celiac disease. Interestingly, *RGS1* is expressed at higher levels in intestinal intraepithelial lymphocytes than in systemic T cells (Shires *et al.* 2001). Intestinal epithelial lymphocytes appear to have a key role in epithelial cell death and the development of villous atrophy in celiac disease (Hue *et al.* 2004).

Another region showing strong association with celiac disease in the Hunt *et al.* study (2008) is located on chromosome 2q11-q12 (*CELIAC8*), harbouring four genes (*IL1RL1, IL18R1 IL18RAP, SLC9A4)* in strong LD with each other. Two intergenic SNPs, located within the same 400 kilobase (kb) LD block, showed strong evidence for association with celiac disease (Hunt *et al.* 2008). It is not yet clear how the risk variants at this locus contribute to celiac disease susceptibility. *IL1RL1* and *SLC9A4* are not expressed in small intestinal tissue; thus, they are unlikely to contribute to the aetiology of celiac disease but cannot be completely excluded (Hunt *et al.* 2008). *IL18R1* and *IL18RAP* form the receptor for

IL18, making both genes very good candidates for celiac disease susceptibility because Th1 inflammatory responses characterised by the expression of IFNγ and active IL18 cytokine are found in mucosal samples of untreated celiac patients (Salvati *et al.* 2002). Hunt *et al.* (2008) found that genotypes that increased disease risk correlated with different (lower) mRNA expression of *IL18RAP*, but not of *IL18R1*, pointing to a primary role for *IL18RAP*. Association to this locus was not replicated in Italian (Romanos *et al.* 2009) or in Spanish (Dema *et al.* 2009) patients with celiac disease, suggesting that this locus is not associated with celiac disease in the Southern European populations. Alternatively, the region may confer a weaker risk effect than estimated in the previous studies. Association of this locus with type 1 diabetes has been reported in British patients (Smyth *et al.* 2008) and also with inflammatory bowel disease in Dutch patients (Zhernakova *et al.* 2008).

Hunt *et al.* (2008) also identified a genetic variant within a large cluster of chemokine receptor genes on chromosome 3p21 (*CELIAC9*). The SNP providing the strongest association signal in this locus was located between the *CCR3* and *CCR2* genes, both of which encode receptors for chemokines. This locus is also associated with celiac disease in the Spanish population (Dema *et al.* 2009), but not in the Italian population (Romanos *et al.* 2009). The inability to replicate the finding may reflect population heterogeneity or a weaker effect size in the Italian population than previously estimated. This locus does not appear to be associated with type 1 diabetes, as was previously reported in the British population (Smyth *et al.* 2008).

A 70 kb LD block on chromosome 3q25-q26 (*CELIAC10*) showed strong association with celiac disease in the Hunt *et al.* study (2008), and subsequently also in the Italian population too (Romanos *et al.* 2009). This region is located immediately upstream of the *IL12A* gene. Together with IL12B, IL12A forms the heterodimeric IL12 molecule which has a broad range of biological activities on T cells and natural killer cells. Interestingly, IL12 is required for Th1 lineage differentiation and induces IFNγ secretion by T cells, which are both shown to be important in celiac disease.

Several SNPs within a 70 kb LD block on chromosome 3q28 (*CELIAC11*) showed association with celiac disease in the study by Hunt *et al.* (2008). The SNP showing strongest association in this region is located in an intron of the LIM domain containing preferred translocation partner in lipoma (*LPP*) gene. Not much is known about the *LPP* gene, but it is highly expressed in the small intestine (Hunt *et al.* 2008) and is also known to have an important role in cell adhesion (Petit *et al.* 2003). However, it is not known how *LPP* would contribute to the pathogenesis of celiac disease. The *LPP* region also showed modest association with celiac disease ain the Italian population (Romanos *et al.* 2009).

Another susceptibility locus for celiac disease is located on chromosome 6q24.1 (*CELIAC12*), harbouring the T cell activation RhoGTPase activating protein (*TAGAP*) gene (Hunt *et al.* 2008): modest association with celiac disease was also found in the Italian population (Romanos *et al.* 2009). *TAGAP* is associated with type 1 diabetes in the UK population (Smyth *et al.* 2008) and this locus has also shown evidence for linkage in a large Dutch pedigree (van Belzen *et al.* 2004). *TAGAP* is expressed during T cell activation (Mao *et al.* 2004) and presents with three different isoforms (NCBI Entrez Gene: http://www.ncbi.nlm.nih.gov/sites/entrez?db=gene), but little is known about its functions.

The seventh newly identified risk locus by Hunt *et al.* (2008) is located on chromosome 12q24 (*CELIAC13*). Genetic variants situated close to the SH2B adaptor protein 3 (*SH2B3*) gene, also known as lymphocyte adaptor protein (*LNK*), were found to increase the risk of

celiac disease. The association results have been replicated in the Italian population (Romanos *et al.* 2009) and strong association with type 1 diabetes has also been reported in this region (Todd *et al.* 2007). *SH2B3* is involved in mediating interactions between extracellular receptors and intracellular signalling pathways in T cell activation (Li *et al.* 2000, Takaki *et al.* 2000, Buza-Vidas *et al.* 2006). Hunt *et al.* (2008) demonstrated expression of *SH2B3* in the small intestine, with increased expression levels in untreated celiac biopsies that possibly reflects the leukocyte recruitment and activation observed in celiac disease (Hunt *et al.* 2008).

As the strongest association signals yielded in the genome-wide association study by van Heel *et al.* (2007) had been investigated further by Hunt *et al.* (2008), Trynka *et al.* (2009) examined the variants that showed modest association with celiac disease in the 2007 genome-wide association study. 458 SNPs were selected for genotyping in 1682 cases and 3258 controls from three populations (UK, Irish and Dutch). The results were combined with the initial UK data from the 2007 genome-wide association study, and showed association to eight SNPs. These SNPs were also genotyped in an independent cohort from Italy (538 cases and 593 controls). The combined analysis revealed the strongest association signals within the *OLIG3-TNFAIP3* locus on chromosome 6p23.3 and the *REL* locus on chromosome 2p16.1. The *OLIG3-TNFAIP3* locus is associated with rheumatoid arthritis (Plenge *et al.* 2007, Thomson *et al.* 2007), systemic lupus erythematosus (Musone *et al.* 2008, Graham *et al.* 2008), type 1 diabetes (Fung *et al.* 2009) and psoriasis (Nair *et al.* 2009). *REL* has demonstrated modest evidence for association with ulcerative colitis (Zhernakova *et al.* 2008). *OLIG3* is involved in the development of neurons of the dorsal horn of the spinal cord (Muller *et al.* 2005), making it an unlikely susceptibility gene for autoimmunity. Both *TNFAIP3* and *REL* are involved in the NF-kappa-B signalling pathway, which has a role in innate immunity (Belguise and Sonenshein 2007, Lee *et al.* 2000, Boone *et al.* 2004, Turer *et al.* 2008). However, no difference was observed in *TNFAIP3* or *REL* mRNA expression levels between untreated celiac patients and controls. Small intestinal biopsy samples and whole blood samples of different genotypes demonstrated no difference in mRNA expression (Trynka *et al.* 2009), highlighting the need for further investigation of these genes in relation to the pathogenesis of celiac disease.

The findings of the 2007 genome-wide association study, its follow up studies and replication efforts are summarised in Table 3.

As seen in other common complex diseases, the genome-wide association approach has been highly successful in finding new susceptibility genes in celiac disease as well. More novel risk loci have been found and confirmed in this single scan and its follow-up studies than with any other experimental approach used previously. Most importantly, these results appear to be replicable in other populations and similar diseases, further validating the initial findings. Nevertheless, further studies are required to pinpoint the risk genes within the LD blocks and identify the risk-causing variants, as well as to characterise the function of the variants in the pathogenesis of celiac disease.

Table 3 *Most promising susceptibility loci and genes for celiac disease identified in the genome-wide association study and its follow-up studies.*

Locus	Chromosome	Gene/s	Full gene name	OR (95% C.I.)	Population	References	Replications / other diseases
CELIAC7	1q31	*RGS1*	regulator of G-protein signalling 1	0.71 (0.63-0.80)	Meta-analysis (UK, Irish, Dutch)	van Heel *et al.* 2007, Hunt *et al.* 2008	T1D: Smyth *et al.* 2008, CD: Romanos *et al.* 2009
CELIAC8	2q11-q12	*IL1RL1, IL18R1, IL18RAP, SLC9A4*	interleukin 1 receptor-like 1, interleukin 18 receptor 1, interleukin 18 receptor accessory protein, solute carrier family 9	1.27 (1.15-1.40)	Meta-analysis (UK, Irish, Dutch)	van Heel *et al.* 2007, Hunt *et al.* 2008	IBD: Zhernakova *et al.* 2008, T1D: Smyth *et al.* 2008, *CD (no association):* Romanos *et al.* 2009, Dema *et al.* 2009.
	2p16	*REL*	v-rel reticuloendotheliosis viral incogene hololog	0.84 (0.78-0.90)	Meta-analysis (UK, Irish, Dutch, Italian)	van Heel *et al.* 2007, Trynka *et al.* 2009	UC: Zhernakova *et al.* 2008
CELIAC10	3q25-q26	*IL12A*	interleukin 12A	1.34 (1.19-1.51)	Meta-analysis (UK, Irish, Dutch)	van Heel *et al.* 2007, Hunt *et al.* 2008	CD: Romanos *et al.* 2009, T1D (*no association*): Smyth *et al.* 2008
CELIAC11	3q28	*LPP*	LIM domain containing preferred translocation partner in lipoma	1.21 (1.11-1.31)	Meta-analysis (UK, Irish, Dutch)	van Heel *et al.* 2007, Hunt *et al.* 2008	CD: Romanos *et al.* 2009, T1D (*no association*): Smyth *et al.* 2008
CELIAC9	3p21	*CCR2, CCR3*	chemokine receptors 2 and 3	1.21 (1.10-1.32)	Meta-analysis (UK, Irish, Dutch)	van Heel *et al.* 2007, Hunt *et al.* 2008	CD: Dema *et al.* 2009, *no association*: Romanos *et al.* 2009, T1D (*no association*): Smyth *et al.* 2008
CELIAC6	4q27	*KIAA1109, TENR, IL2, IL21*	KIAA1109, testis nuclear RNA-binding protein, interleukin 2, interleukin 21	0.63 (0.57-0.71)	Meta-analysis (UK, Irish, Dutch)	van Heel *et al.* 2007	T1D: Todd *et al.* 2007, Zhernakova *et al.* 2007, RA: Zhernakova *et al.* 2007, CD: Adamovic *et al.* 2008a, Romanos *et al.* 2009, UC: Festen *et al.* 2009

	6q23	*OLIG3, TNFAIP3*	oligodendrocyte lineage transcriptionfactor 3, tumor necrosis factor-alpha-induced protein 3	1.25 (1.15-1.34)	Meta-analysis (UK, Irish, Dutch, Italian)	van Heel *et al.* 2007, Trynka *et al.* 2009	SLE: Musone *et al.* 2008, Graham *et al.* 2008, RA: Plenge *et al.* 2007, Thomson *et al.* 2007, T1D: Fung *et al.* 2009, PSO: Nair *et al.* 2009
CELIAC12	6q25	*TAGAP*	T cell activation RhoGTP-ase activating protein	1.21 (1.11-1.31)	Meta-analysis (UK, Irish, Dutch)	van Heel *et al.* 2007, Hunt *et al.* 2008	T1D: Smyth *et al.* 2008, CD: Romanos *et al.* 2009
CELIAC13	12q24	*SH2B3*	SH2B adaptor protein 3	1.19 (1.10-1.30)	Meta-analysis (UK, Irish, Dutch)	van Heel *et al.* 2007, Hunt *et al.* 2008	T1D: Todd *et al.* 2007, Zhernakova *et al.* 2007, CD: Romanos *et al.* 2009

OR: odds ratio, 95% C.I.:95% confidence interval, CD: celiac disease, T1D: type 1 diabetes, RA: rheumatoid arthritis, IBD: inflammatory bowel disease, UC: ulcerative colitis, PSO: psoriasis.

Aims of the study

The broad aim of this study was to investigate genetic susceptibility loci for celiac disease in the Finnish, Hungarian and Italian populations. The specific aims were to:

1. Screen for previously reported genetic susceptibility loci for celiac disease in the Hungarian population (I).

2. Fine-map and use candidate gene analysis at the 5q31-q33 (*CELIAC2*) locus, with the ultimate goal of identifying risk genes and variants for celiac disease (I)

3. Evaluate the role of the 19p13 (*CELIAC4*) locus and the *MYO9B* gene in celiac disease in the Finnish and Hungarian populations (II).

4. Evaluate the role of the *IL18RAP* locus in celiac disease in the Finnish, Hungarian and Italian populations (III).

5. Validate a novel genotyping method for *HLA* risk haplotypes in celiac disease (IV).

Materials and methods

1. Study subjects

Study subjects were drawn from Finland, Hungary and Italy. The number of families and independent cases and controls from each population used in each project are presented in Table 4 and the study materials are described in detail below.

Table 4 *The number of Finnish, Hungarian and Italian samples used in each of the publications.*

	Finland		Hungary		Italy (Trieste)	Italy (Milan)
Publication	**Families**	**cases / controls**	**Families**	**cases / controls**	**cases / controls**	**cases / controls**
I	277	189* / 176*	374	258 / 403		
II	158 (45 with DH)		337 (39 with DH)	270 (62 with DH) / 270		
III	282#	844 / 698	342#	607 / 448	187 / 239	
IV	85	210 / 176		177 / 179	134 / 202	543 / 592

**Two SLC36A2 SNPs were genotyped for additional 375 Finnish cases and 310 population controls.*
#The family materials overlapped partially with the single cases.

1.1. Finnish population

The Finnish sample material consisted of two different family collections, independent single cases with celiac disease, and population controls.

The Finnish family and single patient cohorts were collected at the Paediatric Research Centre, University of Tampere Medical School and Tampere University Hospital. The first collection of Finnish celiac disease families (158 families in total) has been described earlier (Mustalahti *et al.* 2002b). The additional collection of family materials commenced in 2006 and is still ongoing. Both collections are based on screening family members of affected individuals for silent celiac disease. Earlier diagnoses were re-evaluated by scrutinising medical records and performing antibody screenings to reveal asymptomatic cases among family members. Diagnosis was based on initial small bowel biopsy (or skin biopsy for DH patients) and/or EMA positivity. Over 90% of patients fulfilled the strict diagnostic criteria.

The Finnish single case material consisted of patients with celiac disease or DH, and they were independent of the family materials. The diagnoses were made or re-evaluated using the same criteria as described for the families.

The Finnish controls were selected to be representative for the Finnish population and were part of four larger datasets. One dataset was chosen to be representative for the Finnish population density excluding Lapland and Northern Karelia. These samples were used in publications I, III and IV. The second set of control samples was collected at the University of Tampere, and they have been shown by antibody testing and biopsy to be negative for celiac disease and were used in publications I and IV. The third control set was a population-based sample of Finnish monozygotic twins described previously (Aulchenko *et al.* 2009), and the

fourth set was part of a national cohort collected to provide a comprehensive picture of the health state of the Finnish population. The details of this fourth cohort have also been previously described (Aromaa and Koskinen 2004, Stefansson *et al.* 2008). The third and fourth control sets were used in publication III. The controls matched the geographical spread of the patients, which represented the general (mixed) Finnish population with a slight bias to Southern Finland and Tampere region.

1.2. Hungarian population

The Hungarian celiac disease material consisted of pedigrees where at least one offspring was affected with celiac disease. In addition, a set of unrelated patients with celiac disease was also studied. The Hungarian patient material was collected at the Coeliac Disease Center, Pal Heim Children's Hospital, Budapest, and at the Department of Paediatrics, University of Debrecen, and new samples have been collected throughout this Ph.D. candidature. Diagnosis of the Hungarian patients was based on initial small bowel biopsy (or skin biopsy from DH patients) and/or EMA positivity. As in the Finnish patient collection, over 90% of the patients fulfilled the strict diagnostic criteria having both severe villous atrophy and EMA positivity; the remaining patients were diagnosed based on serology or histology. All DH patients were shown to have granular IgA deposition in their skin by direct immunofluorescence.

The Hungarian single case material consisted of patients with celiac disease or DH, and was independent of the family material. The diagnoses were made or re-evaluated using the same criteria as described for the families.

The Hungarian population control material consisted of samples representing the Hungarian population according to the population density. The details of the Hungarian control population have been described earlier (Central Statistical Office 2001, Széles *et al.* 2005).

1.3. Italian population

The Italian samples were collected from two distinct regions of Italy. The first set was from Trieste in North-Eastern Italy, the second set was from Milan in Northern Italy. All celiac patients were diagnosed in accordance with the ESPGAN criteria (Working group of European Society of Paediatric Gastroenterology and Nutrition 1990) and the intestinal biopsies were analysed using Oberhuber's classification (Oberhuber *et al.* 1999). In addition, the patients' serum samples tested positive for both anti-TGM2 antibody and anti-EMA.

The two sets of Italian population controls were selected to geographically match the Italian patient sets.

1.4. Ethical permits

The collection of patient and control materials was approved by the ethical committees of the Tampere and Helsinki University Hospitals, Finnish National Public Health Institute, Heim Pal Children's Hospital, Budapest, the University of Debrecen, the Independent Local Ethical

Committee of the Burlo Garofolo Children's Hospital in Trieste and the hospital Fondazione IRCCS Ospedale Maggiore Policlinico, Mangiagalli e Regina Elena, Milano, Italy. All enrolled participants were informed about the study according to the study protocol and gave written informed consent.

2. Methods

2.1. DNA extraction

Genomic DNA was extracted from whole blood samples by a standard non-enzymatic method, using the Flexigene DNA kit or the QIAamp DNA Blood Mini Kit (Qiagen GmbH, Hilden, Germany).

2.2. Marker selection

(I) Microsatellite markers and SNP markers were selected based on 11 previous genome-wide linkage scans and fine-mapping studies showing linkage with celiac disease in different populations (Zhong *et al.* 1996, Greco *et al.* 1998, King *et al.* 2000, Naluai *et al.* 2001, Liu *et al.* 2002, Woolley *et al.* 2002, Van Belzen *et al.* 2003, van Belzen *et al.* 2004, Rioux *et al.* 2004, King *et al.* 2001, Greco *et al.* 2001). The selected markers were used for a linkage screen in the Hungarian sib-pair families with celiac disease. The markers and their genetic positions are listed in Table 5.

The SNPs at the *CELIAC2* locus on 5q31-q33 were selected to tag immunologically interesting candidate genes as listed in Table 6. The tag-SNP selection was performed in the CEU population genotype data downloaded from www.hapmap.org with the Tagger program using aggressive tagging, implemented in Haploview (v. 3.3) (de Bakker *et al.* 2005, Barrett *et al.* 2005), using a minor allele frequency of 0.10, minimum haplotype frequency of 0.05, and r^2 of 0.80. The *FGF1*, *YIPF5* (*SMAP5*), *TCERG1*, *STK32A*, *JAKMIP2* and *SPINK1* genes showed strongest association results in recent articles by Amundsen *et al.* (2007), and Adamovic *et al.* (2008b) and thus were also genotyped in this study. The *SLC36A2* gene was chosen for typing because it was the only gene in the genome-wide association study (van Heel *et al.* 2007) that showed evidence for association at the *CELIAC2* locus.

Table 5 *The microsatellite markers and SNPs selected for the linkage screen in the Hungarian sib-pair families. Marshfield centimorgan (cM) map was used to determine the genetic positions of the markers.*

Chromosome	Marker	cM position	Chromosome	Marker	cM position
1	D1S243	0.00	5	D5S410	156.00
	D1S1612	16.22		D5S422	164.00
	D1S3669	37.05		D5S2069	182.00
2	D2S116	198.70	7	D7S821	109.12
	CTLA4-318C/T / rs5742909	199.11		D7S1799	113.92
	CTLA4+49A/G / rs231775	199.12	9	D9S925	32.00
	ICOS-693G/A / rs11883722	199.13		D9S1679	44.00
	ICOSivs+173T/C / rs10932029	199.14		D9S169	49.00
	ICOSc602A/C / rs10183087	199.15		D9S1817	59.00
	ICOSc1624C/T / rs10932037	199.16	11	D11S1760	12.00
	ICOSc2373G/C / rs4675379	199.17		D11S1999	17.00
	D2S1271	199.20		D11S4190	27.00
4	D4S412	5.00		D11S904	34.00
	D4S3009	23.00		D11S935	46.00
	D4S2960	29.00		D11S4142	111.00
	D4S2639	33.00		CD3D	114.00
5	IRF1p1	136.10		D11S4089	119.00
	CAh15a	136.20		D11S934	126.00
	CAh17a	136.30	15	D15S1002	15.00
	D5S816	139.00		D15S1031	22.00
	rs31036	146.64		D15S1232	31.00
	rs31037	146.65	16	D16S3253	72.00
	rs1424293	146.73		D16S3091	111.00
	D5S2033	149.00	19	D19S714	42.00
	D5S640	153.00		D19S407	48.00

(II) The four SNP markers within the *MYO9B* gene were selected based on the previously reported association results by Monsuur *et al.* (2005) (rs2305767, rs1457092, rs2305764) and Van Bodegraven *et al.* (2006) (rs1545620) (Figure 5).

(III) The two SNPs (rs13015714 and rs917997) located in the LD block containing the *IL18RAP* gene were selected for genotyping based in the previous association results by van Heel *et al.* (2007) and Hunt *et al.* (2008) (Figure 6).

(IV) The six HLA risk haplotype tagging SNPs (rs2187668, rs2395182, rs4713586, rs7775228, rs4639334, rs7454108) were selected for genotyping based on the publication by Monsuur et al. (2008), where these SNPs were shown to correlate with haplotypes known to be associated with celiac disease. See Table 7 and Figure 7 for details.

Table 6 *The genes and SNPs selected for the positional candidate gene analysis at the CELIAC2 locus on chromosome (chr) 5q31-q33. Mb: megabases*

Gene	Description	Tagging SNPs	Position on chr 5 (Mb)
IL3	interleukin 3	rs2073506	131.423
		rs40401	131.424
IRF1	interferon regulatory factor 1	rs10072700	131.845
		rs839	131.847
		rs2070729	131.848
		rs17622656	131.849
IL5	interleukin 5	rs2069812	131.908
IL13 & IL4	interleukin 13 & interleukin 4	rs2070874	132.038
		rs4426908	132.051
IL9	interleukin 9	rs2069885	135.256
		rs1859430	135.258
TGFBI	transforming growth factor, beta-induced	rs756462	135.396
		rs11738979	135.402
		rs4141306	135.403
CD14	CD14 antigen	rs4914	139.992
		rs2569190	139.993
IK	IK cytokine, down-regulator of HLA II	rs1583005	140.012
		rs4801	140.022
FGF1	fibroblast growth factor 1	rs4912868	141.953
		rs33999	141.954
		rs250092	141.958
		rs34019	141.981
		rs249916	141.989
		rs4912870	141.996
		rs1596776	142.014
		rs10067332	142.022
YIPF5 (SMAP5)	Yip1 domain family, member 5	rs5020191	143.518
		rs4289587	143.518
		rs6864640	143.520
TCERG1	transcription elongation regulator 1	rs2163770	145.810
		rs11743333	145.819
STK32A	serine/threonine kinase 32A	rs31036	146.595
		rs31037	146.596
		rs1424293	146.683
SPINK1	serine peptidase inhibitor, Kazal type 1	rs17107298	147.183
		rs3777125	147.184
IL17B	interleukin 17B	rs372402	148.732
		rs353268	148.737
CD74	CD74 antigen	rs17111145	149.762
		rs2288817	149.767
		rs7724855	149.772
SLC36A2	solute carrier family 36 (proton/amino acid symporter), member 2	rs13357969	150.732
		rs7708940	150.769

Figure 5 *The LD pattern in the MYO9B region on chromosome 19p13 in the CEU HapMap population (International HapMap Project: http://www.hapmap.org). The locations of the four genotyped SNPs (rs2305767, rs1545620, rs1457092, rs2305764) are marked with *.*

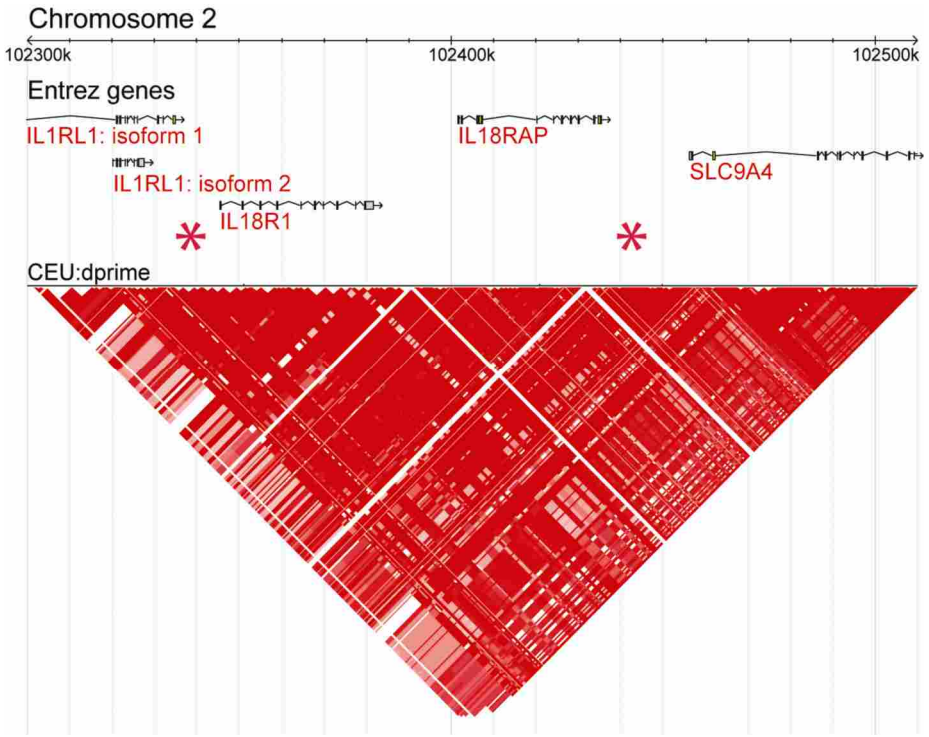

Figure 6 *The LD pattern in the IL18RAP region on chromosome 2q12 in the CEU HapMap population (International HapMap project: http://www.hapmap.org). The locations of the two genotyped SNPs (rs13015714 and rs917997) are marked with *.*

Table 7 *The HLA risk haplotype tag SNPs.*

SNP rs-number	Allele call		Applied Biosystems assay number	Basepair position	Tags DQ type	Positive predicting allele	Negative predicting allele
	VIC	FAM					
rs2395182	G	T	C__11409965_10	32521295	*DQ2.2*	T	
rs4639334	A	G	C__42975350_10	32710192	*DQ7*	A	
rs2187668	C	T	C__58662585_10	32713862	*DQ2.5*	T	
rs7775228	C	T	C__29315313_10	32766057	*DQ2.2*	C	
rs4713586	A	G	C__27950246_10	32767560	*DQ2.2*		G
rs7454108	C	T	C__29817179_10	32789461	*DQ8*	C	

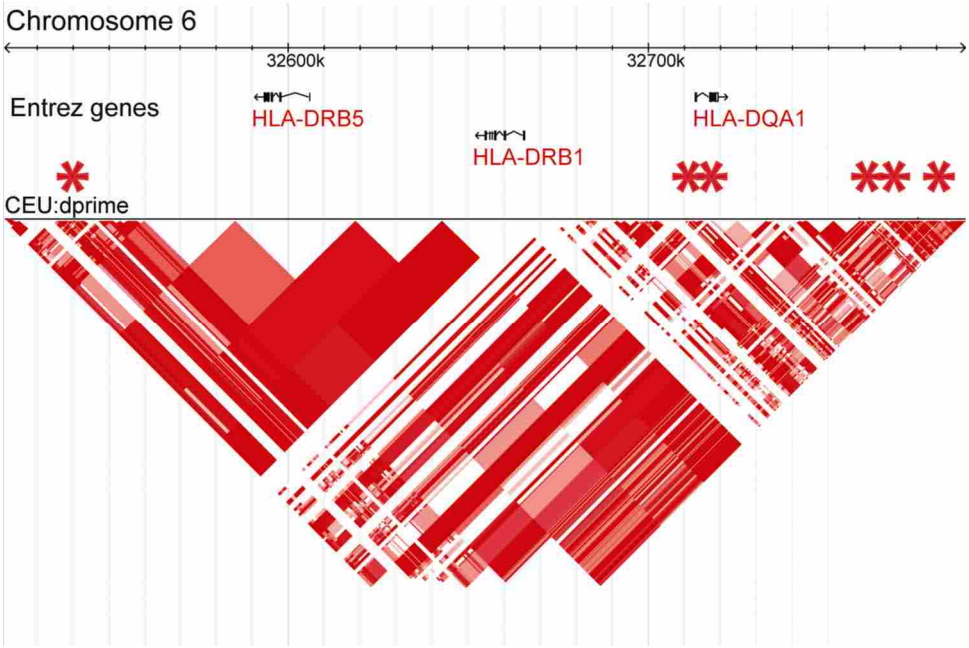

Figure 7 *The LD pattern in the HLA region on chromosome 6p21.3 in the CEU HapMap population. (International HapMap project:* http://www.hapmap.org*). The locations of the six HLA tag SNPs are marked with *.*

2.3. Power analysis (I, II, III)

Power analyses to detect associations to the Dutch *MYO9B* risk variants in the Monsuur *et al.* study (2005), the *IL18RAP* risk variants in the Hunt *et al.* study (2008) and the *SLC36A2* risk variants in the van Heel *et al.* study (2007) were performed assuming a disease prevalence of 1% and using risk parameters reported in the original studies, and significance level 0.05 was used. For the case-control datasets we used the Power for Association With Error program (http://linkage.rockefeller.edu/pawe/pawe.cgi) (Gordon *et al.* 2002, Gordon *et al.* 2003) or the Genetic Power Calculator program (http://pngu.mgh.harvard.edu/~purcell/gpc/) using the Case-control for discrete traits option. For the families, we used the Genetic Power Calculator program (http://pngu.mgh.harvard.edu/~purcell/gpc/) using the TDT for discrete traits option (Purcell *et al.* 2003).

The power analysis for the *MYO9B* variants showed 92% power for the Hungarian case-control set and 89% power for a TDT of Finnish and Hungarian trios. The power analysis for 85 trios with DH showed 47% power, and 56% power for 100 Hungarian DH cases and 270 controls.

Our power to detect association of *IL18RAP* to celiac disease was 66% in the Hungarian case-control dataset, 74% in the Finnish case-control dataset and 27% in the Italian case-control dataset.

The power to detect association to the SNPs in *SLC36A2* was analysed in the Finnish case-control materials. The power of an allelic test to detect association in the 189 Finnish cases and 176 controls was as low as 39%. Therefore, the number of cases and controls was increased (564 and 486, respectively) to yield power of >80%.

2.4. Genotyping

2.4.1. Microsatellite markers (I)

114 Hungarian affected sib-pair and multiplex families were genotyped for the microsatellite markers listed in Table 5 (I). The microsatellites were assayed using fluorescent-labelled primers and the ABI PRISM 3730 DNA analyser and analysed with ABI PRISM Genemapper (v.3) software (Applied Biosystems, CA, Foster City, USA) at the Department of Medical Genetics, University of Helsinki, Helsinki, Finland.

2.4.2. SNPs (I-IV)

Both the Sequenom (Sequenom Inc. San Diego, CA, USA) and TaqMan (Applied Biosystems, Foster City, CA, USA) genotyping platforms have been used for SNP genotyping in these doctorate studies. The Sequenom methodology is based on matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry (Sequenom Inc. San Diego, CA, USA) and up to 40 SNPs were multiplexed in one PCR reaction. Sequenom genotyping was performed either at the Finnish Genome Center, University of Helsinki,

Helsinki, Finland (I), or at the Mutation Analysis Facility, Karolinska Institute, Huddinge, Sweden (II).

The TaqMan chemistry is based on commercially available on-demand SNP genotyping assays with fluorescent-labelled allele-specific probes provided by Applied Biosystems (Applied Biosystems, Foster City, CA, USA). The PCR reactions and allelic discriminations were performed using ABI PRISM 7900 Sequence Detection System instrument (Applied Biosystems, Foster City, CA, USA). The TaqMan genotyping was performed in our laboratory at the Department of Medical Genetics, University of Helsinki, Helsinki, Finland (I, II, III, IV). Using TaqMan methodology it is only possible to genotype one SNP at a time: in contrast, several SNPs can be typed simultaneously using the Sequenom platform. Nevertheless, TaqMan is optimal for small projects involving genotyping of only a few SNPs, and the methodology is relatively fast and easy to perform.

In publication III, the SNP genotypes of a subset of Finnish controls were derived from two larger studies (Stefansson *et al.* 2008, Aulchenko *et al.* 2009). The genotyping was performed using the HumanHap300-Duo Genotyping BeadChips (Illumina Inc., CA, USA).

2.4.3. HLA typing (IV)

From the Finnish family cohort selected for publication IV, 212 individuals had previously been genotyped for the *DQB1* and *DRB1* genes at the Finnish Red Cross Blood Service. The genotyping of the *DQB1* polymorphisms was performed using the Olerup SSP *DQB1* low resolution kit (Olerup SSP AB, Saltsjöbaden, Sweden). *DRB1* genotypes were determined using *HLA*-linked microsatellite markers. This method has been described earlier by Karell *et al.* (2000). For 136 unrelated Finnish patients and 52 controls without celiac disease, the *HLA*-typing was performed using the DELFIA® Celiac Disease Hybridization Assay Kit (PerkinElmer Life and Analytical Sciences, Wallac Oy, Turku, Finland). This method detects the presence of the *DQA1*05*, *DQB1*02* and *DQB1*0302* alleles.

From the Hungarian population, 78 patients had previously been genotyped for the *DQB1* and *DRB1* polymorphisms using the Olerup SSP DQ low resolution and Olerup SSP DR low resolution kits (Olerup SSP AB, Saltsjöbaden, Sweden).

Among the Italian sets, 97 of the celiac cases from the Trieste region had been previously genotyped for the *DQB1* and *DQA1* genes using the low and high resolution Dynal Classic SSP DQ Kits (Dynal A.S., Oslo, Norway) based on PCR with allele-specific primers.

2.5. Data analysis

2.5.1. Data storage and management

The genotype data, pedigree structures and affection status information were imported into a BC|GENE LIMS database (Biocomputing Platforms, Espoo, Finland).

2.5.2. Quality control

In the family materials, genotyping results were tested for Mendelian errors using the PedCheck program (v. 1.1) (O'Connell and Weeks 1998). Mendelian inconsistency was solved by removing all genotypes for that marker in that family. Unlikely double or multiple recombinations were checked using the --best option of the Merlin program (v. 2 beta) (Abecasis *et al.* 2002). If recombinations occurred close to each other in an individual, the subject was omitted from the analysis. Furthermore, individual markers showing several recombinations were omitted from further analyses.

In unrelated case-control samplesets, possible deviations from Hardy-Weinberg Equilibrium were analysed separately in the cases and controls for each population using a p-value of 0.05 as cut-off. Hardy-Weinberg Equilibrium and genotyping call rates in each dataset were calculated using Haploview (de Bakker *et al.* 2005, Barrett *et al.* 2005).

2.5.3. Linkage analysis (I-III)

Multipoint non-parametric linkage (NPL) was tested in the Finnish and Hungarian families, both separately and in the combined material, using the Merlin program (v. 2 beta) (Abecasis *et al.* 2002). NPL analysis was chosen because the mode of inheritance of celiac disease is not known and unlikely to follow Mendelian patterns. Linkage scores (Zmean, non-parametric logarithm of odds [LOD]) were calculated using the NPL_{all} statistics. The genetic positions of the SNP markers were interpolated using the CARTOGR program (Knuuttila *et al.* 2007).

2.5.4. Transmission-disequilibrium test (TDT) (I-III)

TDT was performed using the Genehunter program (v.2) (II, III) for single SNPs and using a sliding window for haplotypes comprising of up to four markers (Kruglyak *et al.* 1996, Kruglyak and Lander 1998). If there was more than one affected offspring in a pedigree, only the index patient or the first diagnosed sibling was selected to the analysis. In publication I, TDT was performed using Unphased v. 3.0.13, assuming linkage (Dudbridge 2008) because all affected children in the pedigrees were included in the analysis. The TDT was analysed in both single SNPs and haplotypes in LD blocks. LD blocks were estimated using Haploview v. 4.1 (de Bakker *et al.* 2005, Barrett *et al.* 2005), and $D' \geq 0.8$ was considered as high LD and used when estimating haplotypes.

2.5.5. Association analysis in case-control materials (I-III)

(I) Allelic, haplotypic and genotypic associations of the markers at the *CELIAC2* locus were analysed. For the allelic and haplotypic analysis, Unphased v. 3.0.13 (Dudbridge 2008) was used and the haplotypes were estimated as in the family materials. The genotypic association was calculated using the χ^2 test in the R package v. 2.6.0 found at http://www.r-project.org/.

(II) The *MYO9B* allele and haplotype frequency calculations were performed using the Haploview program (v. 3.3) (Barrett *et al.* 2005) in the Hungarian case-control sampleset. The

association was calculated separately for patients with DH and patients with only the intestinal form of celiac disease. A combined association analysis was also performed. The index patients from each family with DH were added to the case-control comparison to increase the statistical power of the analysis. The haplotypes of the unrelated individuals were constructed with the SNPHAP program (v. 1.3) (Clayton 2002). Odds ratios (OR) were calculated using Fisher's exact test, and one-sided p-values were reported.

(III) Allele and haplotype frequency calculations, as well as association analyses, were performed using Haploview (v. 4.0) (Barrett *et al.* 2005). Two sided p-values are reported and OR were calculated. A haplotype cut-off of > 1% was used in the controls. Combined ORs from different populations were calculated using the Mantel-Haenszel meta-analysis approach. Forest plots were constructed using rmeta v2.14 on the R platform available from http://cran.r-project.org/web/packages/rmeta/index.html. The Breslow-Day test for rs917997 was significant (p-value 0.03) in a meta-analysis of our Finnish, Hungarian and Italian datasets combined with the previously published UK, Dutch, Irish (van Heel *et al.* 2007, Hunt *et al.* 2008) and Italian (Romanos *et al.* 2009) datasets, indicating heterogeneity of odds ratios. However, Cochran's Q statistic revealed no evidence of allelic heterogeneity in this data (p-value 0.5).

2.5.6. Validation of the HLA tag SNP genotyping assays (IV)

The results from the tag SNP genotypes were compared to previously established *HLA* genotypes by commercial *HLA* typing kits and *HLA*-linked microsatellite markers. The DQ-types were determined from the tag SNP results as described by Monsuur *et al.* (2008). If a mismatch between the two typing methods was found, the DQ type was verified in our laboratory using the Olerup SSP *DQB1* and *DRB1* low resolution kits (Olerup SSP AB, Saltsjöbaden, Sweden). DQ-typing results obtained by traditional typing methods and the tag SNP approach were compared to evaluate the sensitivity, specificity, positive predictive value and correlation (r^2) of the test using GraphPad Prism 4 (GraphPad Software, La Jolla, CA, USA). From family materials, all available family members with sufficient DR-DQ data were counted, but only one patient per family was selected for the haplotype frequency estimations.

2.6. Electrophoretic mobility shift assay (II)

The sequence spanning the SNP rs2305764 in the MYO9B gene was analysed with the Consite program (http://asp.ii.uib.no:8090/cgi-bin/CONSITE/consite), and a putative transcription factor binding site was observed. Double-stranded 41 base pair (bp) probes (CTGATGCCGAGGCATATACG[A/G]CCCCCTTTTTGTGTGAGTGT) spanning the rs2305764 SNP were thus designed for an electrophoretic mobility shift assay (EMSA). The probes were 32P radio-labelled and incubated with nuclear extracts of HeLa cells according to Murumägi *et al.* (2003). As a negative control, probes were incubated with nuclear extraction buffer alone. After 30 min incubation, the samples were separated using non-denaturing gel electrophoresis and visualised by autoradiography (Murumägi *et al.* 2003). To test the specificity of the bound nuclear protein(s) to the rs2305764 SNP sequence, EMSA was competed with the addition of excess unlabelled probes.

2.7. Western blot (II, III)

(II) Two commercial antibodies against MYO9B were tested for Western blotting using peripheral blood mononuclear cells (PBMC) and small intestinal biopsy samples. We were not able to produce reliable staining results, due to either non-specific or complete lack of staining.

(III) PBMCs were purified from buffy coats of anonymous blood donors by density gradient centrifugation and cultured for 72 hours in complete RPMI-1640 media. Western blot was performed on the samples according to standard protocols with reducing (SDS) conditions. Cell lysates of the collected samples (40µg) were treated with Complete Protease Inhibitor Cocktail (Roche Diagnostics GmbH, Germany) according to the manufacturer's instructions. Membranes were stained with 1 µg/ml mouse anti-human IL18RAP (clone 4G4, Abnova, Taiwan), and with anti-beta-actin antibody for loading control (clone 8226, Abcam, UK) after stripping. HRP-conjugated goat anti-mouse IgG (Abcam, UK) was used as secondary antibody for detection with Western Lightning Chemiluminescence Reagent Plus (PerkinElmer LAS, Inc., Boston, USA). Statistical comparison between carriers and non carriers of rs917997 allele A was performed using GraphPad Prism 4. Only one sample had the genotype AA, so it was grouped with the heterozygotes in the analysis.

2.8. Peripheral blood mononuclear cell cultures and enzyme-linked immunosorbent assay (III)

Cryopreserved PBMCs from 38 healthy donors were cultured for 22 h in X-VIVO15 media (Lonza, USA) with anti-CD3-anti-CD28 beads (Invitrogen, UK) at a ratio of 1:5 beads:cell with/without IL18 (R&D systems, UK) at doses 0 ng/ml, 1 ng/ml, 10 ng/ml and 100 ng/ml. All donor samples were cultured in triplicate in 96-well plates on the same day to eliminate experiment and batch variation. IFNγ was measured in the supernatants by enzyme-linked immunosorbent assay (ELISA) (E-Bioscience, CA, USA), and the mean of triplicate wells was analysed. Statistical comparison between the three stratified rs917997 genotype groups was performed using curve-fit regression analysis accounting for dose-response (GraphPad Prism 4).

2.9. Immunohistochemical analysis (II, III)

(II) Two commercial antibodies against MYO9B were tested for immunohistochemical staining of frozen and paraffin-embedded intestinal biopsies. Ten intestinal samples were analysed from the following subjects: 3 adults with untreated celiac disease who were positive for serum IgA class EMA, 4 adults with treated celiac disease (on a strict gluten-free diet for at least one year) who were EMA-negative, and 3 healthy control subjects who were EMA-negative and had normal villous structure. One of the treated celiac disease patients was known to be negative for the putative risk-conferring genotypes; one was known to be positive for the risk-increasing genotype. Unfortunately, reliable staining results were not produced (data not shown).

(III) IL18RAP protein expression was analysed by staining formalin-fixed, paraffin-embedded biopsy samples from 5 untreated (1 homozygous for the risk haplotype, 4 homozygous for the protective haplotype) and 9 treated (5 homozygous for the risk haplotype, 4 homozygous for the protective haplotype) patients with celiac disease, as well as 5 controls (2 homozygous for the risk haplotype, 3 homozygous for the protective haplotype), with 5 µg/µl mouse anti-human IL18RAP (clone 4G4, Abnova, Taiwan). Mouse IgG2a kappa (5 µg/µl; clone MG2a-52, Abcam, UK) was used as an isotype control on each slide. Epitope unmasking was performed by boiling at 95°C in 10 mmol/l citrate buffer, pH 6.0, for 1 min + 1 min, and immunohistochemistry was performed using Vectastain Elite ABC kit (Vector Laboratories, CA, USA).

Results

1. Linkage screen of 11 candidate loci in Hungarian families (I)

The 114 Hungarian affected sib-pair and multiplex families were analysed for genetic linkage using 40 microsatellite markers and 10 SNPs in 11 previously reported candidate loci for celiac disease. The results of the analysis are shown in Table 8. Markers on chromosomal regions 5q31-q33 and 19p showed the strongest evidence for genetic linkage, although none of the results reached statistical significance.

2. Fine-mapping of the CELIAC2 locus on chromosome 5q31-q33 (I)

2.1. Linkage in Finnish and Hungarian pedigrees

The *CELIAC2* locus on chromosome 5q31-33 showed evidence for linkage in the Hungarian pedigrees, and it has previously been shown to be linked with celiac disease in Finland and multiple other populations (Zhong *et al.* 1996, Greco *et al.* 1998, Naluai *et al.* 2001, Liu *et al.* 2002, Woolley *et al.* 2002, Greco *et al.* 2001). Thus, we fine-mapped this locus in large Finnish and Hungarian family materials, including both multiplex and trio families, partially overlapping with the families that were studied in the previous linkage scans in the Hungarian population (within this study) and also from the Finnish population (Holopainen *et al.* 2001, Liu *et al.* 2002).

On chromosome 5q31-q33, 48 SNPs tagging 17 genes were successfully genotyped in a total of 651 Finnish and Hungarian families. The 138 Finnish families that were informative for linkage analysis demonstrated some linkage to the whole region (maximum Zmean 2.66, p=0.004 at 140.02 – 140.04 cM) (Figure 8). The 134 Hungarian families informative for analysis showed a weaker linkage signal to the region (maximum Zmean 2.10, p=0.02 at 148.07 – 148.08 cM) (Figure 8). The combined Finnish and Hungarian family material gave a maximum Zmean of 3.22, (p=0.0006 at 149.51 cM) (Figure 8).

2.2. Association analysis in Finnish and Hungarian pedigrees and case-control materials

TDT was analysed in the Finnish and Hungarian families. The analysis revealed association of different genes with celiac disease in the two populations. The alleles and haplotypes demonstrating nominal association are shown in Tables 9 and 10, respectively.

Allelic, genotypic and haplotypic association in the cases and controls was also analysed. None of the associations demonstrated in the Finnish and Hungarian families were replicated in the case-control sets (Table 9).

Table 8 *Results of the linkage screen of candidate loci for celiac disease in 114 Hungarian affected sib-pair families. Mb: megabases*

Chromosomal region	Location (Mb)	Number of markers	Merlin Zmean (MAX)	p-value	Linkage previously shown by:
1p	2.1-17.7	3	-0.57	0.70	King *et al.* 2000, Liu *et al.* 2002
2q33 (*CELIAC3*)	201.4-204.7	9	0.44	0.30	Naluai *et al.* 2001, Rioux *et al.* 2004, Holopainen *et al.* 2004, Naluai *et al.* 2000
4p	3.4-18.5	4	0.76	0.20	Liu *et al.* 2002, Woolley *et al.* 2002
5q31-q33 (*CELIAC2*)	132-168.2	12	1.51	0.06	Holopainen *et al.* 2001, Greco *et al.* 1998, Naluai *et al.* 2001, Liu *et al.* 2002, Greco *et al.* 2001
7q	95.9-104.0	2	0.05	0.50	Zhong *et al.* 1996, Liu *et al.* 2002
9p	18.3-33.9	4	0.90	0.20	Naluai *et al.* 2001, Liu *et al.* 2002, van Belzen *et al.* 2004
11p-q	5.3-125.6	9	0.07	0.50	Zhong *et al.* 1996, Greco *et al.* 1998, King *et al.* 2000, Woolley *et al.* 2002, van Belzen *et al.* 2004, King *et al.* 2001
15q (*CELIAC3*)	25.5-32.8	3	0.36	0.40	Woolley *et al.* 2002
16q	71.8-81.5	2	0.91	0.20	King *et al.* 2000, Liu *et al.* 2002, van Belzen *et al.* 2004
19p (*CELIAC4*)	16.1-19.9	2	1.38	0.08	Van Belzen *et al.* 2003 and 2004

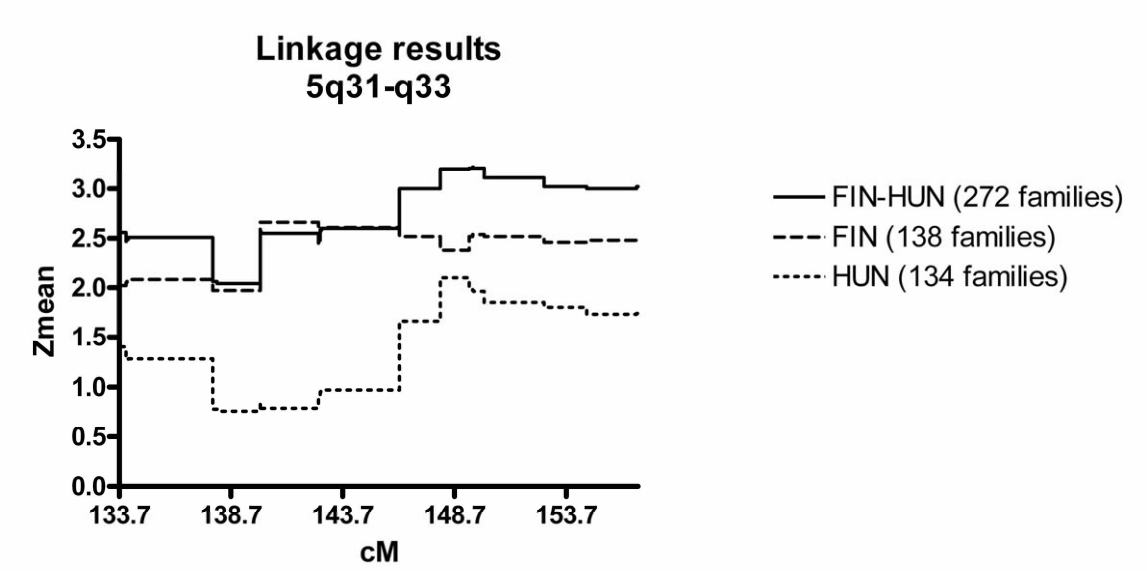

Figure 8 *Linkage results for a 20 centimorgan (cM) region on chromosome 5q31-q33 in the Finnish (FIN) and Hungarian (HUN) multiplex families.*

Table 9 *Allelic TDT in the Finnish and Hungarian families and the allelic and genotypic association results in the Finnish and Hungarian cases and controls. Only alleles and genotypes showing nominal association in at least one of the sample sets are shown. In the TDT, the case frequencies refer to the allele frequencies in the affected offspring and the control frequencies refer to the number of nontransmissions of the alleles, and also transmissions to unaffected offspring. P-values ≤ 0.05 were considered as significant and are highlighted with bold. Nominal p-values are shown.*

				TDT				Case-control					
				Finland		Hungary		Finland			Hungary		
Gene	Marker	Mb	Minor allele*	Freq. case, control	p-value	Freq. case, control	p-value	Freq. case, control	allelic test, p-value	genotypic test, p-value	Freq. case, control	allelic test, p-value	genotypic test, p-value
IL3	rs2073506	131.4	T	0.16, 0.12	0.11	0.06 0.13	**0.0003**	0.11, 0.12	0.54	0.42	0.11, 0.10	0.73	0.59
IL3	rs40401	131.4	T	0.37, 0.33	0.39	0.21, 0.28	**0.03**	0.31, 0.35	0.30	**0.001**	0.25, 0.26	0.77	0.05
IRF1	rs839	131.8	T	0.39, 0.35	0.12	0.30, 0.32	0.52	0.32, 0.34	0.45	0.59	0.28, 0.33	**0.04**	0.39
IL5	rs2069812	131.9	A	0.26, 0.31	0.17	0.23, 0.27	0.49	0.37, 0.31	0.07	0.27	0.26, 0.31	0.06	**0.05**
TGFBI	rs11738979	135.4	T	0.35, 0.39	0.29	0.41, 0.47	**0.03**	0.38, 0.38	0.99	0.76	0.49, 0.46	0.34	0.19
TGFBI	rs4141306	135.4	A	0.44, 0.48	0.33	0.47, 0.55	**0.005**	0.48, 0.43	0.24	0.68	0.57, 0.55	0.44	0.67
FGF1	rs33999	142.0	G	0.53, 0.45	**0.04**	0.21, 0.38	0.19	0.42, 0.42	1.00	0.74	0.39, 0.39	0.86	0.19
FGF1	rs250092	142.0	A	0.16, 0.18	0.35	0.14, 0.19	0.06	0.12, 0.17	**0.03**	**0.03**	0.15, 0.14	0.45	0.39
YIPF5	rs5020191	143.5	C	0.38, 0.35	0.48	0.39, 0.42	0.40	0.38, 0.30	**0.05**	0.13	0.35, 0.39	0.21	0.44
YIPF5	rs6864640	143.5	T	0.20, 0.19	0.47	0.20, 0.20	0.71	0.17, 0.21	0.18	0.29	0.25, 0.19	**0.02**	0.05
TCERG1	rs11743333	145.8	G	0.36, 0.35	0.95	0.34, 0.36	0.31	0.36, 0.38	0.59	0.71	0.31, 0.36	**0.05**	0.07
SPINK1	rs3777125	147.2	G	0.27, 0.29	0.66	0.38, 0.40	0.60	0.34, 0.25	**0.01**	0.11	0.41, 0.39	0.36	0.55
SPINK1	rs17717320	147.2	G	na	na	na	na	0.15, 0.19	0.17	**0.05**	0.17, 0.18	0.70	0.20
CD74	rs7724855	149.8	A	0.18, 0.17	0.60	0.17, 0.14	0.19	0.14, 0.14	0.97	0.83	0.18, 0.14	**0.05**	0.13
SLC36A2	rs7708940	150.8	G	0.48, 0.39	**0.01**	0.44, 0.46	0.42	0.46, 0.44	0.50	0.70	0.43, 0.43	0.95	0.99

*Mb: position at megabases on chromosome 5, na: not analysed *: Minor alleles in the Finnish family controls*

Table 10 *Haplotype TDT in Finnish and Hungarian families and in Finnish cases and controls. In the TDT, case frequencies refer to the allele frequencies in the affected offspring and the control frequencies refer to the number of non-transmissions of the alleles, and also transmissions to unaffected offspring. P-values ≤ 0.05 were considered significant; only significant results are shown.*

Finnish families			
SLC36A2 (2 SNPs) p-value = 0.002			
Haplotype	**Freq. cases, controls**	**OR**	**95% C.I.**
A-A	0.47, 0.55	reference	
G-A	0.04, 0.07	0.67	0.24 - 1.85
G-G	0.49, 0.35	1.59	1.06 – 2.39

Hungarian families			
IL3 (2 SNPs) p-value = 0.002			
Haplotype	**Freq. cases, controls**	**OR**	**95% C.I.**
C-C	0.79, 0.73	reference	
C-T	0.15, 0.14	0.96	0.67 - 1.38
T-T	0.06, 0.13	0.43	0.26 - 0.72

OR: odds ratio, C.I.: confidence interval

3. Linkage and association study of MYO9B, a positional candidate gene at the CELIAC4 locus (II)

Given that *MYO9B* demonstrated strong association with celiac disease in the Dutch population (Monsuur *et al.* 2005), and our evidence for linkage in this region with the Hungarian population (I), we examined *MYO9B* genetic variants in our sample collections. In total, 1265 Finnish and Hungarian patients with gluten intolerance were included in the study, of which 161 (13%) suffered from DH. The putative risk haplotype identified in the study by Monsuur *et al.* (2005) is referred to as A-G-A-A.

3.1 Linkage in Finnish and Hungarian pedigrees

All four markers showed significant linkage to celiac disease in the combined Finnish and Hungarian material of 240 multiplex families, with the maximum NPL LOD score of 3.76 (p=0.00002); Table 11.

The gut and skin manifestations of gluten intolerance were also analysed separately. That is, the families were divided in two groups according to presence (or absence) of any DH in the families. Linkage and association of the four *MYO9B* SNPs were studied in a total of 85 DH families. The *MYO9B* region showed linkage to DH in the combined material of 47 multiplex families (non-parametric LOD score of 1.13, p=0.01; Table 11). In the Finnish families, the region showed stronger linkage in DH families than in non-DH families; the 11 Hungarian affected sib-pair families did not have enough power to show linkage to DH. Although there were only 47 pedigrees with DH in this study (i.e. one-fifth of the pedigrees), their contribution to the linkage score was considerable (contributing to approximately one-third of the total LOD score).

Table 11 *Results of the NPL test in the Finnish and Hungarian pedigrees. LOD: non-parametric LOD score, N: number of informative families. The LOD scores with a p-value <0.05 are shown in bold. DH: families with dermatitis herpetiformis, NON-DH: families without any DH.*

	Hungarian and Finnish families			Hungarian families			Finnish families		
Diagnosis group	**LOD**	**p-value**	**N**	**LOD**	**p-value**	**N**	**LOD**	**p-value**	**N**
All (DH+NON-DH)	**3.76**	**0.00002**	240	**3.09**	**0.00008**	131	**0.83**	**0.03**	109
NON-DH	**2.57**	**0.0003**	193	**2.60**	**0.0003**	120	0.22	0.2	73
DH	**1.13**	**0.01**	47	0.35	0.10	11	**0.68**	**0.04**	36

3.2. No association in the large Finnish and Hungarian family cohorts

No association of the *MYO9B* marker alleles or haplotypes with celiac disease was found in the large Hungarian and Finnish family materials (Table 12).

Table 12 *Results of the allele and haplotype TDT in the Hungarian and Finnish pedigrees with only one affected child per family included in the analysis. The number of transmissions/non-transmissions (T/NT) is shown. Families with and without DH are included.*

	Hungarian and Finnish families			Hungarian families			Finnish families		
Marker*allele	**T/NT**	**χ^2**	**p-value**	**T/NT**	**χ^2**	**p-value**	**T/NT**	**χ^2**	**p-value**
rs2305767*A	140 / 129	0.45	0.50	101 / 92	0.42	0.52	39 / 37	0.05	0.82
rs1545620*G	124 / 138	0.75	0.387	86 / 98	0.78	0.38	38 / 40	0.05	0.82
rs1457092*A	124 / 137	0.65	0.42	88 / 99	0.65	0.42	36 / 38	0.05	0.82
rs2305764*A	134 / 140	0.13	0.72	98 / 102	0.08	0.78	36 / 38	0.05	0.82
AGAA haplotype	78 / 88	0.60	0.44	51 / 63	1.26	0.26	27 / 25	0.08	0.78
GTCG haplotype	87 / 88	0.01	0.94	58 / 58	0.00	1.00	29 / 30	0.02	0.90

3.3. Putative association in the Hungarian patients with dermatitis herpetiformis

The TDT of the combined Hungarian and Finnish DH family material revealed that none of the SNPs are significantly associated with DH (Table 13). However, an independent TDT analysis of the Hungarian DH families showed that three of the SNPs were associated with DH, the most significant being rs2305764 (p=0.003) (Table 13). The Dutch risk haplotype A-G-A-A was over-transmitted to the offspring in the Hungarian families with DH (12 transmitted vs. 1 non-transmitted, p=0.002). Moreover, the opposite haplotype G-T-C-G was under-transmitted to the offspring in these pedigrees (1 transmitted vs. 12 non-transmitted, p=0.002).

Table 13 *TDT results of the Finnish and Hungarian pedigrees with patients presenting DH, the skin manifestation of gluten intolerance. The number of transmissions/non-transmissions (T/NT) is shown. P-values ≤ 0.05 were considered significant and are shown in bold.*

	Hungarian and Finnish families			Hungarian families			Finnish families		
Marker*allele	**T/NT**	**χ^2**	**p-value**	**T/NT**	**χ^2**	**p-value**	**T/NT**	**χ^2**	**p-value**
rs2305767*A	22 / 13	2.31	0.13	16 / 4	7.20	**0.007**	6 / 9	0.60	0.44
rs1545620*G	18 / 11	1.69	0.19	10 / 3	3.77	**0.05**	8 / 8	0.00	1.0
rs1457092*A	19 / 11	2.13	0.14	13 / 3	6.25	**0.01**	6 / 8	0.29	0.59
rs2305764*A	24 / 13	3.27	0.07	18 / 4	8.91	**0.003**	6 / 9	0.60	0.44
AGAA haplotype	15 / 7	2.91	0.09	12 / 1	9.31	**0.002**	3 / 6	1.00	0.32
GTCG haplotype	7 / 15	2.91	0.09	1 / 12	9.31	**0.002**	6 / 3	1.00	0.32

Hungarian gluten-sensitive patients without DH showed the reverse trend in the TDT analysis: the A-G-A-A haplotype was under-transmitted to patients (p=0.02; Table 14). In Finnish families, gluten intolerance both with and without DH was not found to be associated with any marker or haplotype using TDT analysis (Tables 13 and 14).

Table 14 *TDT results of the Finnish and Hungarian pedigrees with only the intestinal form of gluten intolerance. The number of transmissions/non-transmissions (T/NT) is shown. P-values ≤ 0.05 were considered significant and are shown in bold.*

	Hungarian and Finnish families			Hungarian families			Finnish families		
Marker*allele	**T/NT**	**χ^2**	**p-value**	**T/NT**	**χ^2**	**p-value**	**T/NT**	**χ^2**	**p-value**
rs2305767*A	116 / 116	0.00	1.00	85 / 88	0.05	0.82	33 / 28	0.41	0.52
rs1545620*G	100 / 125	2.78	0.10	76 / 95	2.11	0.15	30 / 32	0.06	0.80
rs1457092*A	103 / 126	2.31	0.13	75 / 96	2.58	0.11	30 / 30	0.00	1.00
rs2305764*A	108 / 127	1.54	0.22	80 / 98	1.82	0.18	30 / 29	0.02	0.90
AGAA haplotype	61 / 81	2.82	0.09	39 / 62	5.24	**0.02**	24 / 19	0.58	0.45
GTCG haplotype	80 / 71	0.54	0.46	57 / 46	1.17	0.28	23 / 27	0.32	0.57

Association was examined further in the Hungarian population using an independent case-control sampleset (270 cases, 270 controls), which included 62 additional DH patients. Allele, genotype and haplotype frequencies of the Hungarian singleton patients were compared to respective frequencies in the Hungarian control population; 38 index patients from the Hungarian families with DH were also added to the comparison. Differences in frequencies were not statistically significance, but suggested that individuals homozygous for the Dutch risk-haplotype alleles to have a higher risk of DH than the heterozygotes or individuals homozygous for the alternative haplotype alleles. The A-G-A-A haplotype frequency for the DH cases was 41.5%, compared to 35.4% for the controls (OR 1.30, p=0.07, 95% CI 0.93 - 1.82). Conversely, the G-T-C-G haplotype frequency for the DH cases was 31.2% compared to 37.9% for controls (OR 0.74, p=0.05, 95% CI 0.52 - 1.05). Individuals homozygous for the A-G-A-A haplotype had an increased risk of DH (OR 1.74, p=0.06, 95% CI 0.92 - 3.29), whereas the G-T-C-G homozygotes had a lower risk of DH (OR 0.65, p=0.17, 95% CI 0.31 - 1.36).

3.4. Electrophoretic mobility shift assay results

The function of the highest risk-conferring SNP rs2305764 found in the Dutch population (Monsuur *et al.* 2005) was examined further using EMSA to determine its effect on gene regulation. The SNP is located in intron 28 (at 141 bp from exon 28 and 123 bp from exon 29) of the *MYO9B* gene, and is at a possible binding site for transcription factors. No difference in the binding of nuclear proteins was detected between the two alleles (Figure 9).

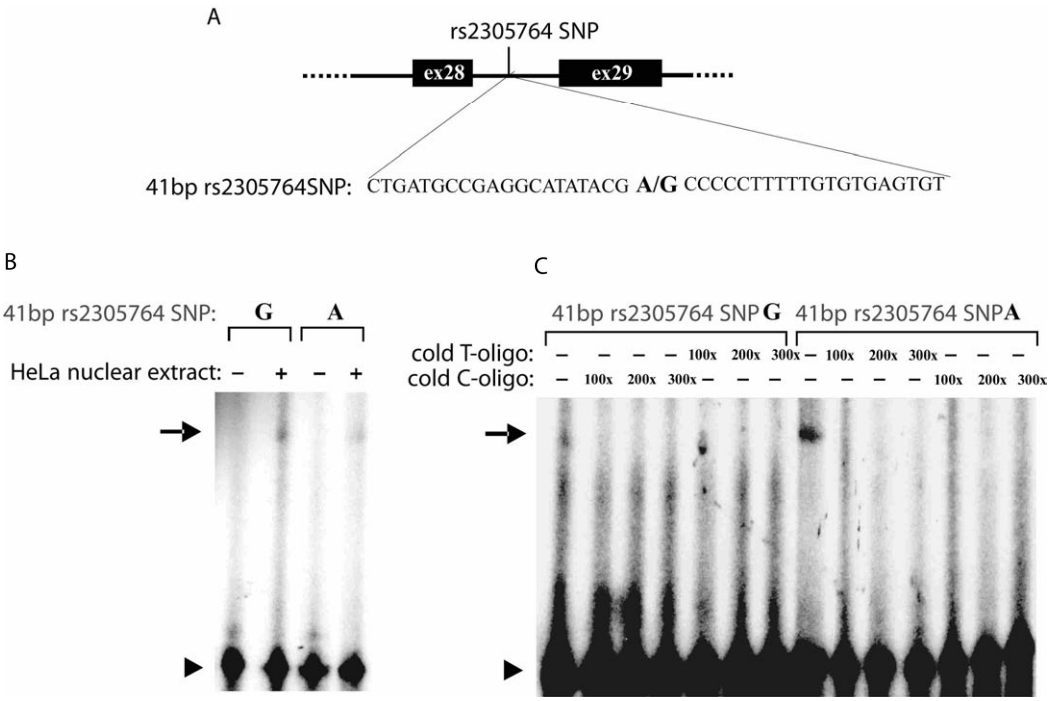

Figure 9 ***A****. Schematic view of the location of the SNP rs2305764 in the intronic region between exons 28 and 29 in the MYO9B gene. 41 base pair (bp) probes spanning the SNP rs2305764 were designed for electrophoretic mobility shift assay (EMSA).* ***B.*** *EMSA results show that the sequence around the rs2305764 SNP binds nuclear protein(s) from HeLa nuclear extract. Identical electrophoretic mobility shift patterns marked by arrows were produced with both A (risk) and G alleles.* ***C.*** *The specificity of the binding of nuclear protein(s) in the rs2305764 SNP sequence. Binding of factors can be competed out with excess addition of unlabelled ('cold') probes. In addition, EMSA can also be competed out with the reciprocal addition of cold oligos, suggesting that the same protein(s) are binding to the oligos bearing the rs2305764 SNP alleles G and A shown by arrow. Free labelled probes are marked by arrowheads.*

4. Association study of the IL18RAP locus in celiac disease (III)

The *IL18RAP* region was first implicated in celiac disease in a genome-wide association study by van Heel and colleagues in 2007, and subsequently confirmed as a novel risk locus by Hunt *et al.* in 2008. Since novel gene associations should be validated in multiple populations, we studied the risk-conferring variants within the *IL18RAP* region in our celiac disease materials.

4.1. No evidence for linkage or association in the Finnish or Hungarian families

Finnish and Hungarian family materials did not show genetic linkage or association with the two markers at the *IL18RAP* locus. Linkage analysis was conducted in 109 Finnish families, yielding a LOD score of 0.16 (p=0.2), and 107 Hungarian families (LOD 0.09, p=0.3). TDT analysis found no evidence for transmission disequilibrium in either population (Finnish: 45/49 p=0.7, Hungarian: 97/78, p=0.2). These results may be (partially) explained by the low information content from analysis of only two bi-allelic markers that are in strong LD.

4.2. Association in the case-control materials

Table 15 shows the association analysis results of the two SNPs at the *IL18RAP* locus and their haplotypes in Finnish, Hungarian and Italian case-control materials. Carriage of rs13015714*G and rs917997*A was significantly associated with increased risk for celiac disease in the Hungarian case-control material (p=0.001, OR 1.38, 95% CI 1.14-1.68 and p=0.001, OR 1.39, 95% CI 1.14-1.70, respectively). The corresponding GA haplotype demonstrated association (p=0.0001, OR 1.48, 95% CI 1.21-1.80) with disease-risk in Hungarian cases, whilst the TG-haplotype showed protection (p=0.0040, OR 0.75, 95% CI 0.62-0.91). Significant association to celiac disease was not found in the Finnish or Italian populations, although frequency differences between cases and controls showed the same tendency as in the Hungarian samples, as well as the previous UK and Dutch studies.

Meta-analysis of our Finnish, Hungarian and Italian datasets combined with previously published UK (van Heel *et al.* 2007, Hunt *et al.* 2008), Dutch (Hunt *et al.* 2008), Irish (Hunt *et al.* 2008) and Italian (Romanos *et al.* 2009) datasets demonstrated strong association of the *IL18RAP* locus SNPs with celiac disease (Figure 10). The OR for carriage of rs917997 was 1.25 ($p=1.89 \times 10^{-11}$, 95% CI 1.17-1.33). Meta analysis of our Finnish, Hungarian and Italian samplesets independently yielded an OR of 1.23 (p=0.001, 95% CI 1.09-1.39).

Table 15 *IL18RAP allele and haplotype association results in Finnish, Hungarian and Italian case-control materials. P-values < 0.05 were considered significant and are shown in bold.*

	n cases	n controls	Marker	Allele/ haplotype	Frequency cases, controls	p-value	OR	95% CI
Finnish	844	698	rs13015714	G	0.21, 0.19	0.07	1.18	0.98-1.41
			rs917997	A	0.21, 0.19	0.09	1.17	0.98-1.40
				TG	0.78, 0.81	0.05	0.84	0.70-1.00
				GA	0.21, 0.19	0.14	1.14	0.96-1.37
Hungarian	607	448	rs13015714	G	0.31, 0.24	**0.001**	1.38	1.14-1.68
			rs917997	A	0.30, 0.24	**0.001**	1.39	1.14-1.70
				TG	0.68, 0.74	**0.004**	0.75	0.62-0.91
				GA	0.30, 0.22	**0.0001**	1.48	1.21-1.80
				GG	0.01, 0.02	**0.04**	0.46	0.21-0.97
				TA	0.01, 0.02	0.12	0.58	0.26-1.28
Italian	187	239	rs13015714	G	0.28, 0.26	0.47	1.14	0.80-1.61
			rs917997	A	0.28, 0.24	0.23	1.24	0.87-1.76
				TG	0.71, 0.74	0.35	0.85	0.60-1.20
				GA	0.28, 0.24	0.27	1.22	0.86-1.74
				GG	0.01, 0.02	0.27	0.41	0.08-1.97

n: number of analysed samples, OR: odds ratio, C.I.: confidence interval

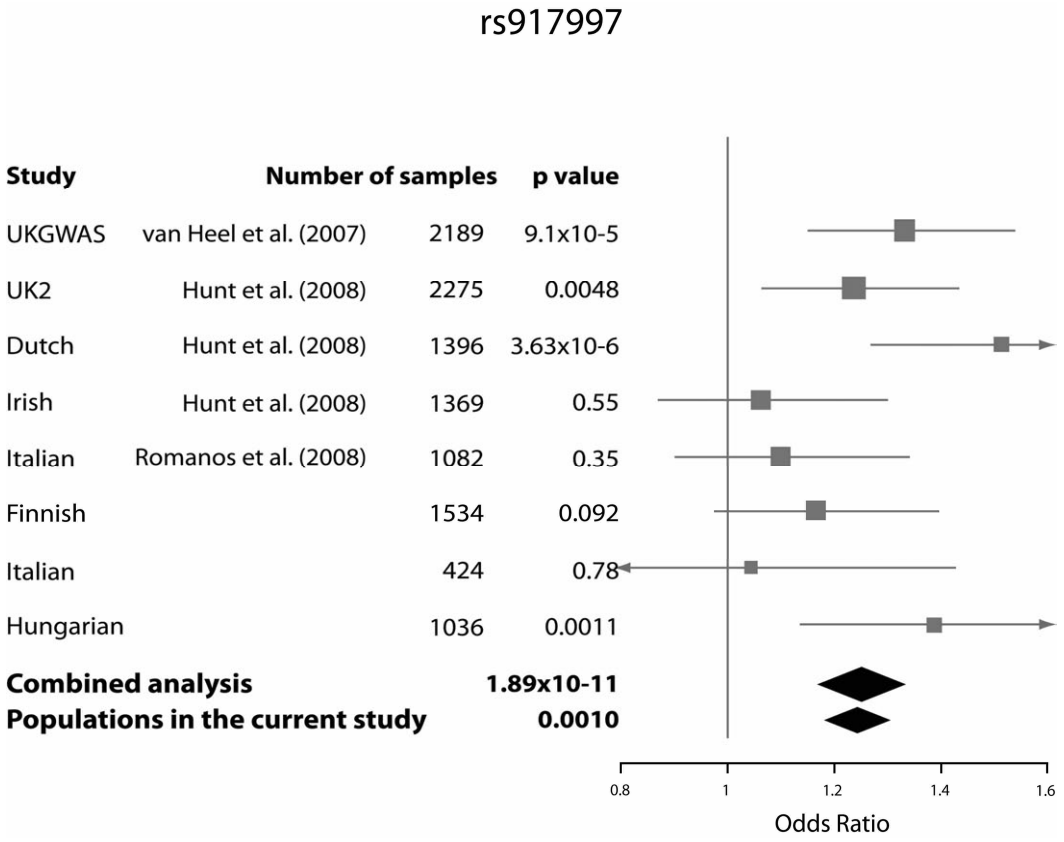

Figure 10 *Meta-analysis of our populations combined with those of van Heel et al. (2007), Hunt et al. (2008) and Romanos et al. (2009) for the IL18RAP SNPs rs917997. The number of the subjects in each population, association p-values and odds ratios (carriage of risk allele) are shown. Box sizes correspond to sample sizes of each population.*

4.3. IL18RAP expression and function in blood leucocytes

Expression levels of IL18RAP were studied in PBMCs from anonymous blood donors by Western blot. Figure 11A shows IL18RAP protein in PBMCs from 16 blood donors genotyped for rs917997. Two bands of differing size both appeared to be specific for IL18RAP. A 70-kilodalton (kDa) protein, which corresponds to the estimated full-length IL18RAP, was visible in 15 out of 16 donors. One sample demonstrated two proteins of approximately 70-kDa in size (lane 1). A protein of 37-kDa was also present in 14/16 PBMC samples. Carriers of the risk allele rs917997*A (AA and AG genotypes) demonstrated higher expression levels of the 37-kDa protein than non-carriers, but lower expression of the 70-kDa protein; however, these differences were not statistically significant (Figure 11B).

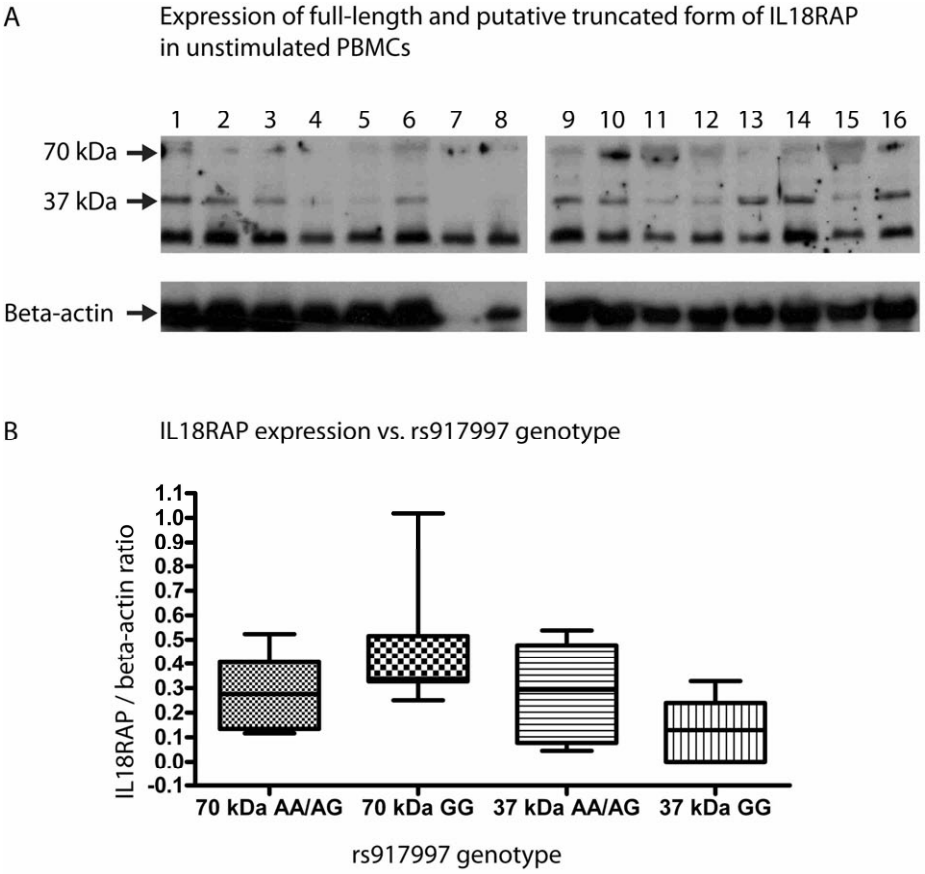

Figure 11 ***A.*** *IL18RAP expression in unstimulated peripheral blood mononuclear cells (PBMCs). The rs917997 genotypes for the PBMC samples are as follows: AA, lane 5; AG, lanes 1-4 and 13-16; GG, lanes 6-12. The lowest band at ~25 kDa is not specific for IL18RAP because it was also present upon staining with secondary antibody only.* ***B.*** *Densitometric analysis of the 37-kDa and 70-kDa bands standardised over beta-actin levels and grouped by rs917997 genotype.*

IL18 induced IFNγ secretion in a dose–dependent manner in polyclonally-stimulated (anti-CD3/CD28) PBMCs (Figure 12), but not in resting cells. However, the rs917997 *IL18RAP* genotypes showed no difference in PBMC responsiveness among 38 healthy donors.

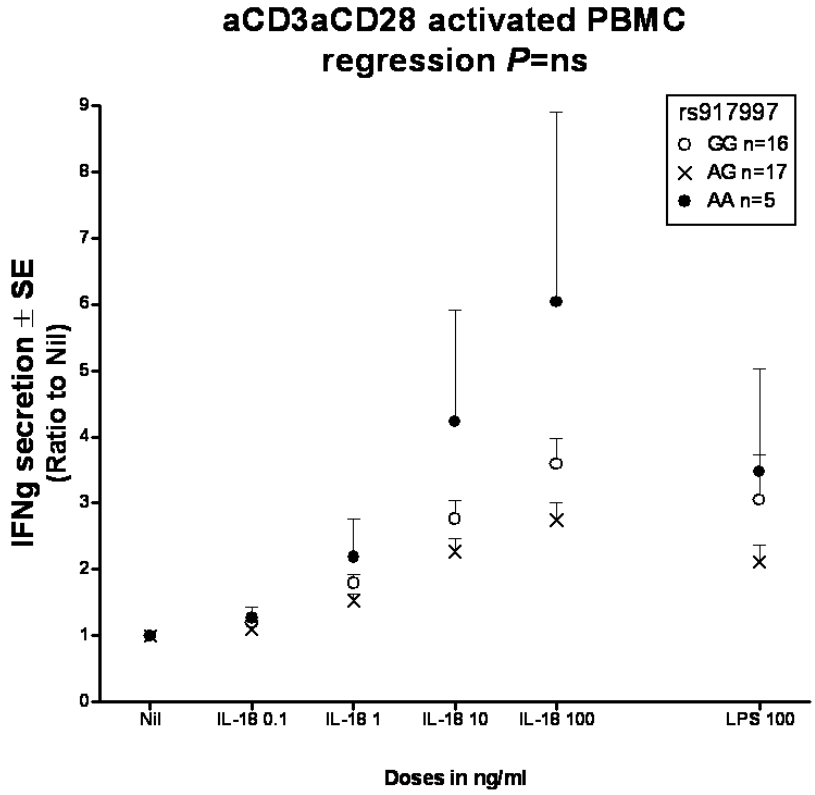

Figure 12 *IFNγ secretion peripheral blood mononuclear cells (PBMCs) stimulated with anti-CD3-anti-CD28 in the presence of increasing concentration of IL18. IL18 induces IFNγ secretion in a dose-dependent manner, but the rs917997 genotypes demonstrated no difference in IFNγ secretion.*

4.4. IL18RAP expression in small intestinal tissue

To examine the expression of *IL18RAP* in celiac disease, immunohistochemical analysis of small intestinal biopsy samples from untreated and treated celiac patients and controls was performed. Mononuclear cells in the lamina propria of two samples demonstrated staining with a monoclonal antibody specific for IL18RAP. This staining was also observed in entero-endocrine cells in the crypts of two samples, as well as in Paneth cells of one sample (Figure 13A, B, and C, respectively). No differences were observed between untreated and treated celiac patients, between patients and controls, or individuals with different rs917997 genotypes (data not shown).

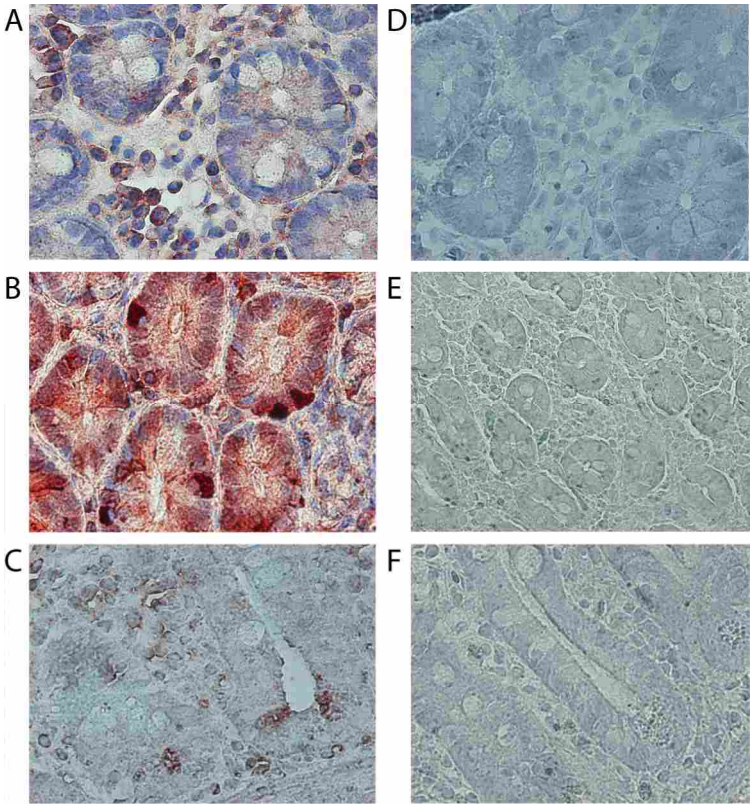

Figure 13 *IL18RAP expression in the small intestinal tissue of treated patients with celiac disease, homozygous for the risk haplotype. Staining with monoclonal antibody specific for IL18RAP (A-C) or isotype control (D-F).*

5. Validation of HLA tag SNPs in Finnish, Hungarian and Italian populations (IV)

5.1. Correlation of HLA results using old and new methodology

To predict the *DQ2.2, DQ2.5*, *DQ7* and *DQ8* risk haplotypes for celiac disease, six tagging SNPs were genotyped in 400 Finnish individuals (212 from a family cohort, 136 unrelated patients and 52 controls), 79 Hungarian patients and 97 Trieste-Italian cases from whom traditional *HLA* genotypes were available.

In the Finnish samples, validation of *DQ2.5*, *DQ2.2*, and *DQ7* haplotypes was only performed for the family cohort because the presence of these haplotypes cannot always be determined using the DELFIA assay when the parental haplotypes are unknown. The validation study results are presented in Table 16, showing sensitivities and specificities ranging from 0.95 to 1. *HLA* testing using the DELFIA kit found 7 individuals to be *DQA1*05* positive and *DQB1*02* negative, implying a serotype of DQ7. Five individuals were *DQB1*02* positive and *DQA1*05* negative, and are possibly DQ2.2. Tagging SNP analysis to determine presence of *DQ7* and *DQ2.2* matched these results. 108 individuals were positive for both *DQA1*05* and *DQB1*02*, and the DELFIA assay could not distinguish whether these subjects carried *DQ2.5* or both *DQ2.2* and *DQ7*. However, the tagging SNP

approach predicted 105 of these to carry the *DQ2.5* haplotype, and 3 subjects had the *DQ2.2*/*DQ7* genotype.

In the Hungarian and Trieste-Italian populations, the sensitivities and specificities ranged from 0.99 to 1 (Table 16).

The Finnish, Hungarian and Trieste-Italian materials were also analysed together for the correlation. The overall correlation was high, the sensitivities for *DQ2.2*, *DQ2.5*, *DQ7* and *DQ8* ranging from 0.97 to 1, and the specificities ranging from 0.996 to 1 (Table 16).

Table 16 *Validation of HLA tag SNPs in Finnish, Hungarian and Italian (Trieste) samples. All: Finland, Hungary and Italy combined, PPV: positive predictive value, r^2: correlation*

DQ2.2	**Finland**	**Hungary**	**Italy (Trieste)**	**All**
Number of chromosomes tested	354	152	166	672
Sensitivity	1	1	1	1
Specificity	1	1	1	1
PPV	1	1	1	1
r^2	1	1	1	1
conflicting results	0	0	0	0

DQ2.5	**Finland**	**Hungary**	**Italy (Trieste)**	**All**
Number of chromosomes tested	388	158	186	732
Sensitivity	0.98	1	1	0.99
Specificity	0.99	1	1	0.99
PPV	0.99	1	1	0.99
r^2	0.94	1	1	0.97
conflicting results	0.015	0	0	0.008

DQ7	**Finland**	**Hungary**	**Italy (Trieste)**	**All**
Number of chromosomes tested	422	158	186	766
Sensitivity	1	1	1	1
Specificity	1	0.99	0.99	0.997
PPV	1	0.95	0.97	0.97
r^2	1	0.94	0.96	0.97
conflicting results	0	0.006	0.0005	0.0026

DQ8	**Finland**	**Hungary**	**Italy (Trieste)**	**All**
Number of chromosomes tested	736	158	192	1086
Sensitivity	0.95	1	1	0.97
Specificity	0.999	0.99	0.99	0.996
PPV	0.98	0.88	0.83	0.93
r^2	0.93	0.87	0.82	0.81
conflicting results	0.0041	0.006	0.01	0.0055

Out of all the 576 tested individuals, 12 (2.1%) showed different results in the tagging SNP assay when compared to the traditional *HLA* typing results. Nine of them were celiac disease affected individuals from the Finnish family cohort, one was a Hungarian celiac patient and two were Italian (Trieste) celiac patients. In addition, four patients from the Italian (Milan) sampleset were predicted to have three *HLA* alleles by the tag-SNP method but when

genotyped for the *DQB1* and *DRB1* genes using the Olerup SSP low-resolution kits, two of the samples showed a rare *DR11-DQ2* haplotype.

5.2. HLA haplotype frequencies in the population materials

The six *HLA* risk haplotype tagging SNPs were also genotyped in an extended set of Finnish, Hungarian and Italian (Milan and Trieste) cases and controls. The *HLA* haplotype and genotype frequencies determined from the SNP allele frequencies are presented in Table 17. Excess of *DQ2.5* among patients was seen both in haplotype (Table 17A) and carrier (Table 17B) frequencies in all 3 populations. According to the SNP genotype results, 90.2% of the Finnish patients with celiac disease carried the DQ2 heterodimer (88.3% *DQ2.5* and 1.9% *DQ2.2*/*DQ7*), and 6.4% carried the *DQ8* haplotype without DQ2. Respectively, these frequencies among Hungarian patients were DQ2 97.2% (DQ2.5 87.5% and DQ2.2/DQ7 9.7%) and DQ8 2.3%, and among the Italian (Trieste) patients DQ2 89.6% (DQ2.5 74.7% and, DQ2.2/DQ7 14.9%), and DQ8 7.5%, and among Italian (Milan) patients DQ2 88.4% (DQ2.5 64.1% and DQ2.2/DQ7 24.3%) and DQ8 6.2%.

A dose effect of DQ2.5 homozygosity on celiac disease risk was observed in all three populations when comparing the genotypes of cases and population controls (Table 17B). Higher DQ2.2 haplotype and DQ2.5/DQ2.2 genotype frequencies were seen among the patients in Hungary and Italy, but to a lesser extent in Finland (Tables 17A and B). In addition, we studied the risk effect of the second haplotype in the presence of one DQ2.5 haplotype in all 4 population groups (Table 17C). The Italian-Milan population was the only dataset with a large enough sample size for such an analysis, thus it provides the most reliable results. The results show that homozygosity for DQ2.5 increases the risk of celiac disease 5.5-fold when compared to individuals with DQ2.5/DQX (where DQX implies haplotypes other than DQ2.5, DQ2.2, DQ7 or DQ8), while DQ2.2 increased the risk of celiac disease 3.1-fold. DQ7 or DQ8 in the presence of DQ2.5 did not confer additional risk to the disease.

Table 17 *A. HLA-haplotype frequencies (%) in Finnish, Hungarian and Italian controls and cases with celiac disease determined by tagging SNP analysis. The total number of chromosomes is shown in brackets.* ***B.*** *HLA-genotype frequencies (%) in Finnish, Hungarian and Italian controls and cases with celiac disease, determined by tagging SNP analysis. The total number of individuals is shown in brackets.* ***C.*** *DQ2.5 stratified analysis using logistic regression and DQ2.5/DQX as the reference group.*

A.	**Finland**		**Hungary**		**Italy (Trieste)**		**Italy(Milan)**		**CEU**
DQ type	controls (352), patients (530)		controls (358) patients (352)		controls (404) patients (268)		controls (1164) patients (1070)		
DQ2.2	6.0	5.8	8.9	20.1	11.9	17.2	10.8	27.1	13.3 (*DQA1*0201*)
DQ2.5	8.8	50.9	10.1	49.6	15.8	44.8	8.4	37.1	15.6 (*DQB1*0201*)
DQ7	6.8	4.0	21.5	13.0	29.5	16.0	27.8	17.9	16.7 (*DQB1*0301*)
DQ8	11.1	6.2	9.8	2.8	5.4	6.7	5.7	4.6	14.4 (*DQB1*0302*)
DQX	67.3	33.0	49.7	14.4	37.4	15.3	47.3	13.3	
Total	100	100	100	100	100	100	100	100	

DQX: not DQ2.2, DQ2.5, DQ7, DQ8, CEU: HLA allele frequencies of the CEU population from de Bakker et al. (2006).

B.	**Finland**		**Hungary**		**Italy (Trieste)**		**Italy (Milan)**	
Genotype	controls (176), patients (265)		controls (179) patients (176)		controls (202) patients (133)		controls (582) patients (535)	
All *DQ2+*	17.6	90.2	21.8	97.2	36.1	89.6	22.5	88.4
DQ2.5/DQX	13.6	58.5	10.1	28.4	13.4	23.1	6.9	18.3
DQ2.5/DQ2.5	0.6	13.6	2.2	11.9	3.0	14.9	0.7	10.1
DQ2.5/DQ2.2	0.0	6.4	1.7	29.5	3.0	15.7	2.9	23.9
DQ2.5/DQ7	1.1	4.9	3.4	14.2	7.9	15.7	4.1	9.3
DQ2.5/DQ8	1.7	4.9	0.6	3.4	1.5	5.2	1.5	2.4
DQ2.2/DQ7	0.6	1.9	3.9	9.7	7.4	14.9	6.4	24.3
All *DQ2-, DQ8+*	19.3	6.4	17.3	2.3	8.9	7.5	9.5	6.2
DQ8/DQX	15.9	4.2	10.1	0.0	5.0	3.7	5.0	2.6
DQ8/DQ8	1.1	0.8	1.7	0.0	0.0	0.7	0.3	0.6
DQ8/DQ2.2	1.7	1.1	2.2	1.1	1.0	1.5	0.7	2.1
DQ8/DQ7	0.6	0.4	3.4	1.1	3.0	1.5	3.4	0.9
All *DQ2-, DQ8-*	63.1	3.4	60.9	0.6	55.0	2.9	68.0	5.4
DQ2.2/DQX	9.7	1.5	10.1	0.0	9.4	2.2	10.3	1.7
DQ2.2/DQ2.2	0.0	0.4	0.0	0.0	1.5	0.0	0.7	1.1
DQ7/DQX	9.1	0.8	22.3	0.6	16.8	0.0	28.4	1.3
DQ7/DQ7	1.1	0.0	5.0	0.0	11.9	0.0	6.7	0.0
DQX/DQX	43.2	0.8	23.5	0.0	15.3	0.7	22.0	1.3

DQX: not DQ2.2, DQ2.5, DQ7, DQ8

C.				**95% C.I.**	
Population	***HLA* genotypes**	**p-value**	**OR**	**Lower**	**Upper**
Finnish	*DQ2.5/DQ2.5*	0.098	5.574	0.730	42.569
	DQ2.5/DQ2.2	0.998	NA	NA	NA
	DQ2.5/DQ7	0.994	1.006	0.214	4.740
	DQ2.5/DQ8	0.556	0.671	0.178	2.529
Hungarian	*DQ2.5/DQ2.5*	0.297	1.890	0.571	6.258
	DQ2.5/DQ2.2	0.005	6.240	1.731	22.497
	DQ2.5/DQ7	0.445	1.500	0.530	4.249
	DQ2.5/DQ8	0.490	2.160	0.243	19.194
Italian-Milan	*DQ2.5/DQ2.5*	0.002	5.510	1.871	16.228
	DQ2.5/DQ2.2	0.000	3.073	1.644	5.744
	DQ2.5/DQ7	0.602	0.850	0.462	1.565
	DQ2.5/DQ8	0.263	0.590	0.234	1.488
Italian-Trieste	*DQ2.5/DQ2.5*	0.046	2.903	1.018	8.281
	DQ2.5/DQ2.2	0.036	3.048	1.073	8.657
	DQ2.5/DQ7	0.752	1.143	0.498	2.622
	DQ2.5/DQ8	0.180	2.613	0.641	10.647

DQ2.5/DQX is the reference group; OR: odds ratio, C.I.: confidence interval, NA: not applicable

Discussion

The main focus of this Ph.D. thesis was the examination of linkage and association of the 5q31-q33, 19p13 and *IL18RAP* loci with genetic susceptibility to celiac disease. The study highlights the challenges of gene mapping in complex diseases. Common risk variants with modest risk effects and low penetrance show association when population-based strategies, employing large sample materials, are applied; however in the presence of allelic heterogeneity and perhaps rare risk variants, linkage mapping utilising family materials appears a more suitable approach. These issues are further discussed in the following pages. In addition, the novel HLA genotyping method that was developed for large-scale investigations is also evaluated.

1. Strong linkage but only weak evidence for association: susceptibility loci on chromosomes 5q31-q33 and 19p13

Previous genome-wide linkage scans have demonstrated that the chromosomal regions 5q31-q33 and 19p13 constitute important susceptibility loci for celiac disease (Greco *et al.* 1998, Naluai *et al.* 2001, Liu *et al.* 2002, Popat *et al.* 2002a, Van Belzen *et al.* 2003, van Belzen *et al.* 2004, Babron *et al.* 2003). Linkage signals in both these regions were also replicated in this study. Linkage of the 5q31-33 locus has previously been reported in the Finnish population (Holopainen *et al.* 2001, Liu *et al.* 2002), but this was the first study to report linkage in the Hungarian population. The linkage signal from the 5q31-q33 locus appeared slightly stronger in the Finnish families than in the Hungarian families, while the 19p13 locus demonstrated a stronger signal in Hungarian families. Despite this evidence of linkage of both loci with celiac disease, only modest association was detected for the regions. The 5q31-q33 region is known to harbour several genes with immunological functions, and is thus a very interesting candidate locus for celiac disease and other immunological-related disorders. However, pinpointing the associated gene(s) in this locus has been challenging. In contrast, chromosome 19p13 harbours a candidate gene, *MYO9B*, which has been shown to be associated with celiac disease (Monsuur *et al.* 2005). Our studies of the 19p13 region concentrated on replication of this previously reported *MYO9B* association.

We selected 17 candidate genes from the 5q31-q33 locus that were located under the linkage peak observed in the Finnish population, and performed association analyses in both Finnish and Hungarian family and case-control materials. The association analyses showed several moderate associations, and mainly for different genes in the different datasets. These results are in concordance with previous fine-mapping studies (Amundsen *et al.* 2007, Adamovic *et al.* 2008b) in which the complexity of the *CELIAC2* locus and evidence for multiple weak associations in the region were demonstrated. In these two previous studies, only moderate single- and multipoint associations were demonstrated, despite the usage of densely-spaced markers. Only one SNP, rs33999, located in *FGF1*, showed association in both Swedish-Norwegian families of the Adamovic *et al.* study (2008b) and the Finnish families used in our study. However, these results were not replicated in the Finnish cases and controls, which might also be an indication of lack of statistical power in that material. None of the genes and SNPs showing evidence for association in the Amundsen *et al.* study (2007)

demonstrated significant association in the Finnish or Hungarian materials. These two Scandinavian studies (Amundsen *et al.* 2007, Adamovic *et al.* 2008b), which applied overlapping materials but different fine-mapping strategies, were unable to provide more conclusive association results than our study. In these two studies and ours, the association signals appear insufficient to account for the linkage signal observed in the study populations. Therefore, even a substantial increase of markers may not significantly alter our results. The results from all three studies still point to an insufficient sample size or strong allelic heterogeneity in the region.

The *CELIAC2* region only demonstrated weak evidence for association in the genome-wide association study performed in the British population, despite being the strongest non-*HLA* candidate locus for celiac disease according to genome-wide linkage studies. The strongest association within 5q31-q33 was observed for two SNPs in LD with the *SLC36A2* gene (van Heel *et al.* 2007). However, this association did not gain genome-wide significance. A follow-up study by Trynka *et al.* (2009) did not show consistent associations within this region either. We demonstrated some evidence for association of the two *SLC36A2* SNPs with celiac disease in the Finnish family materials, but it was not replicated in the Finnish cases and controls. The Hungarian materials showed no evidence for association. The weak association findings and the failure to consistently replicate them may reflect genetic heterogeneity, differences between populations, or rare variants.

The 19p13 region has also been implicated in susceptibility to celiac disease previously, most notably in Dutch linkage scans (Van Belzen *et al.* 2003, van Belzen *et al.* 2004). Fine-mapping in the region showed association with the *MYO9B* gene in the Dutch population (Monsuur *et al.* 2005). *MYO9B* has subsequently been widely studied in European populations. *MYO9B* was associated with celiac disease, systemic lupus erythematosus, and rheumatoid arthritis in the Spanish population (Sanchez *et al.* 2007), and three studies have pointed to the role of this gene in inflammatory bowel disease (van Bodegraven *et al.* 2006, Nunez *et al.* 2007a, Cooney *et al.* 2009). However, all other recent attempts to replicate the findings in celiac disease have failed in the UK, Swedish-Norwegian, Italian and Spanish populations (Hunt *et al.* 2006, Amundsen *et al.* 2006a, Giordano *et al.* 2006, Nunez *et al.* 2006, Latiano *et al.* 2007, Cirillo *et al.* 2007).

In our study, the Dutch risk markers of within the *MYO9B* region were analysed in large Hungarian and Finnish family sets, as well as an independent Hungarian case-control sampleset, and demonstrated strong linkage to celiac disease, thereby confirming the region as a true susceptibility locus. However, neither a single allele nor any of the haplotypes showed association with celiac disease when all the patient materials were combined. Instead, when the family material was divided into two groups according to the occurrence of DH, the Dutch *MYO9B* risk SNP alleles were found to be significantly over-transmitted to the offspring of families with DH. This finding was strongest in the Hungarian families with DH, but the same trend was also seen in the case-control comparison. No association was detected in the Finnish population in any patient group, but genetic linkage was seen only in families with DH. Our findings suggest potential, although relatively weak, genetic differences in risk of the intestinal and skin manifestations of celiac disease, and emphasise the importance of more phenotypically detailed replication studies in larger cohorts of this small subset of patients with gluten intolerance.

To date, functional evidence for a role of *MYO9B* in celiac disease and other diseases is still lacking. But, interestingly, *MYO9B* is more highly expressed in the skin than in the

intestine (NCBI UniGene EST Profile Viewer http://www.ncbi.nlm.nih.gov/UniGene/). The function of the non-coding highest risk SNP rs2305764 reported in the Dutch population (Monsuur *et al.* 2005) was examined further in our study. An EMSA was used to study the effect of this SNP on gene regulation, but no differences in the binding of nuclear proteins were found between the two alleles. The SNP rs2305764 may also serve as a binding site for splice factors, which could be tested in future studies.

Our findings with the 19p13 region may be explained by a more heterogeneous genetic background amongst celiac families with only the intestinal form of the disease, compared to families also exhibiting DH, explaining why linkage and association were seen only in the latter subset. On the other hand, strong linkage was also demonstrated in the Hungarian families displaying only intestinal pathology. Thus, the weak DH association does not fully explain the linkage results. Rather, it suggests either allelic heterogeneity or completely different primary risk alleles or genes for gluten intolerance in the 19p13 region, or the presently studied haplotype block.

Indeed, one indication of this is that the Finnish families with DH did not show association with *MYO9B*, even though the families with DH showed stronger linkage to *MYO9B* than the families without DH. The existence of another risk gene or haplotype may also explain the inability to replicate association in other reported studies. In fact, this haplotype block shows LD with the downstream region of *MYO9B* and the LD actually covers up to four other genes, namely *MDS032*, *OCEL1*, *NR2F6* and *USHBP1* (Figure 5, page 43). Given that this region has not been genotyped systematically in previous studies by Monsuur *et al.* (2005) or Curley *et al.* (2006), or in any recent replication study, more thorough analysis of this haplotype block is needed.

Our results, particularly the linkage results from chromosomes 5q31-q33 and 19p13, strongly support the presence of genuine risk factors for celiac disease at these loci. It is possible that there are several variants conferring low risk or alternatively rare variants conferring high risk to celiac disease at these two loci, explaining the failure to find confirmed association in samplesets with relative small to medium sample size. In addition, the complexity of celiac disease; with its heterogeneous genetic background and different manifestations, may explain the previous failures to detect association in the 5q31-q33 locus or to replicate *MYO9B* associations in different populations. Larger samplesets and high-throughput sequencing in these loci are likely to be required for identification of the true risk genes and variants. Furthermore, studies with high density SNP-coverage in multiple populations might help to reveal the complex genetic patterns in the regions.

2. Association but not linkage: the IL18RAP locus on chromosome 2q12

The *IL18RAP* locus on chromosome 2q12 was shown to be associated with celiac disease in the genome-wide association scan by van Heel *et al.* (2007) and in the subsequent follow-up study by Hunt *et al.* (2008). Our study aimed to further investigate this association between celiac disease and polymorphisms at the *IL18RAP* locus. We genotyped two previously identified celiac disease risk SNPs at this locus in patients and controls from the Finnish, Hungarian and Italian populations. We replicated association of the *IL18RAP* locus with celiac disease in the Hungarian case-control dataset, but not in the Finnish or Italian datasets

alone. The results from the Finnish population, although not significant, suggest that variants at the *IL18RAP* locus may also be true celiac disease risk factors in the Finnish population. However, the risk effect is likely to be considerably smaller than in the Hungarian population. The frequencies of the previously reported risk alleles were higher in Italian cases than in controls, although the difference was not statistically significant. The low power of the Italian sampleset may explain our inability to replicate the association reported by Hunt *et al.* (2008) in this dataset. Alternatively, as reported by a recent study by Romanos *et al.* (2009), the *IL18RAP* locus may simply not be associated with celiac disease in Italian populations.

A meta-analysis of the datasets from our study, van Heel *et al.* (2007), Hunt *et al.* (2008) and Romanos *et al.* (2009) yielded highly significant association (OR 1.25, $p=1.89x10^{-11}$). This indicates that the *IL18RAP* locus is a true risk factor for celiac disease in multiple populations.

The GA risk haplotype showed stronger association than any single marker in the Hungarian population. In contrast, the rare GG and TA haplotypes showed no risk, or perhaps a protective effect. Due to the strong LD between the two markers, the additional information gained from haplotypic analysis is limited. Therefore, genotyping of other tagging markers is not likely to help in further identification of risk and non-risk haplotypes. Identification of the primary risk-conferring variant in this haplotype is likely to require both sequencing of this LD block and clear functional data.

Despite the association seen in case-control materials, Finnish and Hungarian family materials with celiac disease did not demonstrate linkage or association to the *IL18RAP* locus. This may be partially explained by the low information content resulting from the analysis of only two bi-allelic markers in nearly complete LD, and also the large number of non-informative homozygous parents in the dataset. Additionally, the result supports the prevailing dogma that family studies suffer a reduction in power when detecting disease variants with small risk effects (OR only 1.20-1.48 in this study) and low penetrance within families (Risch and Merikangas 1996). Case-control settings with large sample sizes are more powerful to detect such minor risk factors (Risch and Merikangas 1996).

Confirming previously reported risk factors in multiple populations increases confidence in novel findings, as demonstrated by the current and previous studies on the *IL18RAP* locus in celiac disease. The frequencies of the associated alleles may also vary between populations and affect the power of the study. The risk variants at the *IL18RAP* locus are relatively common in all currently studied populations, but do demonstrate a difference in Northern-Southern European variation: being the least frequent in Finland and the most frequent in Hungary and Italy.

It is not yet clear how *IL18RAP* locus risk variants would contribute to celiac disease susceptibility. *IL18RAP* is contained within a 400 kb LD block that also contains three other genes (*IL1RL1, IL18R1, SLC9A4*) (Figure 6, page 43). *IL1RL1* and *SLC9A4* are not expressed in small intestinal tissue and therefore were not considered a high priority in this study, but their role in pathogenesis cannot be excluded and they should be investigated further in future studies. IL18R1 and IL18RAP come together to form the receptor for IL18, making both of them very good candidate genes for celiac disease susceptibility. Further work is needed to confirm *IL18RAP* as the primary risk gene and its role in pathogenesis, and also to exclude *IL18R1*. It is also possible that these two neighbouring genes are co-regulated. However, in the study by Hunt *et al.* (2008), the disease associated genotypes correlated with lower mRNA

expression of *IL18RAP*, but no effect was observed for *IL18R1*: this points to a primary role for *IL18RAP*.

The mRNA correlation shown by Hunt *et al.* (2008) suggests that individuals with the risk allele have reduced *IL18RAP* expression, possibly leading to lower IFNγ secretion. This appears contradictory to the strong IFNγ response seen in active celiac disease. We attempted to directly correlate function with genotype by measuring IFNγ levels from activated PBMCs cultured in the presence of IL18. Although we saw a dose-dependent increase of IFNγ production in the presence of IL18, no significant differences were detected between groups stratified by rs917997 genotype. This experiment may have been limited by small sample size, the heterogenous nature of PBMCs, and the strong T cell activation induced by anti-CD3-anti-CD28 stimulation.

Since IL18RAP has been relatively poorly characterised to date, we studied its protein expression in PBMCs. In addition to the expected 70-kDa protein corresponding to the estimated full-length IL18RAP, we saw expression of an approximately 37-kDa protein, suggesting that an alternative isoform of IL18RAP exists. Although not statistically significant, these results suggest a correlation of the risk allele with higher expression levels of the 37-kDa variant, and lower expression levels of the 70-kDa variant. Hunt *et al.* (2008) showed that the risk allele correlates with lower levels of mRNA expression, consistent with our results for the full-length (70-kDa) protein. This may be a result of the design of the mRNA assay by Hunt *et al.* (2008), which identified only the full-length transcript.

Given that little is known about the IL18RAP protein and its alternative splicing or functional isoforms, future studies should concentrate on their systematic characterisation in celiac disease-specific tissues, cells and inflammatory conditions. Indeed, alternative splicing of IL18RAP mRNA has been described in human and rat, coding for putative truncated proteins and soluble forms of the receptor (Andre *et al.* 2003, Fiszer *et al.* 2007). It would be interesting to examine whether the 37-kDa band is a biologically active isoform encoded by an alternatively spliced mRNA. This is of particular interest because its higher expression in the carriers of the risk allele is in line with the observed IFNγ-mediated Th1 response in celiac disease, which is not the case for the 70-kDa band nor the mRNA expression reported by Hunt *et al.* (2008). However, we cannot exclude novel alternative roles of IL18 signalling or putative coupling of IL18RAP with other cytokine receptor chains, thus forming new receptor complexes with unknown or even opposite functions to the IL18 pathway. Our knowledge about the pathology of IFNγ and Th1 responses in celiac disease may also need a revision, as there is increasing evidence of Th17 cells involved in autoimmunity (Dardalhon *et al.* 2008).

Immunohistochemical analysis of IL18RAP showed a wide variety of expression patterns and revealed no consistent differences between the untreated and treated celiac patients and the controls. Furthermore, no differences were observed between different *IL18RAP* genotypes. However, this analysis confirms that *IL18RAP* is expressed in the gut inflammatory mononuclear cells, thereby further supporting the involvement of this protein in celiac disease pathogenesis.

Our study independently confirms the association results of Hunt *et al.* (2008) and supports *IL18RAP* as a novel celiac disease risk gene in the Hungarian population and possibly also in the Finnish and Italian populations. The unclear nature of the risk genotype effects on protein levels highlights the need for future studies examining alternative splicing and receptor isoforms, as well as novel protein interactions or functions that will increase our understanding of this poorly studied gene in celiac disease pathogenesis.

3. Towards large-scale screening of risk-conferring HLA haplotypes for celiac disease

Given that celiac disease is an important health problem because of its high prevalence, specific and non-specific morbidity and long-term complications (Mearin *et al.* 2005, Romanos *et al.* 2008), early diagnosis is important, especially in high-risk groups such as first-degree relatives and individuals with type 1 diabetes, iron-deficiency anaemia or Down syndrome. Although the presence of HLA DQ2 or DQ8 is not on its own sufficient for diagnosis, it can reveal the need for further serology tests and later a biopsy sampling in the risk groups. The absence of these molecules reduces the risk for the disease substantially.

Testing for *HLA* risk haplotypes is routinely performed using methods which require several reactions, multiple steps such as amplification and hybridisation to a membrane, special software or expertise in analysing the results, and significant financial cost. To make *HLA*-typing more automated and less expensive, Monsuur *et al.* (2008) established a new approach using six tagging SNPs to predict whether an individual is heterozygous or homozygous for the *DQ2.5*, *DQ2.2*, *DQ7* and *DQ8* haplotypes. The tag SNP selection was based on genotype data collected in the classical *HLA* genes and more than 7500 common SNPs and insertion-deletion polymorphisms across the human MHC region (de Bakker *et al.* 2006). Genotyping six SNPs is simple, fast and cost-effective compared to more classical techniques. In the Dutch, Spanish and Italian (Naples) populations, the sensitivity of this test was reported to be >0.991 and the specificity >0.996 (Monsuur *et al.* 2008). Since LD patterns within the *HLA* region may differ between populations, our aim was to validate this method in four new populations.

We genotyped the six *HLA*-tagging SNPs in the Finnish, Hungarian and Northern-Italian (Trieste) populations and compared the genotype with results from traditional *HLA* typing. Our findings showed that the sensitivity and specificity to detect the celiac disease risk alleles in the Finnish, Hungarian and Italian populations ranged from 95 to 100%. These results imply that this method can detect the haplotypes coding for DQ2 and DQ8 with high accuracy. In addition, the method is transferable to other Caucasian populations because it has now been validated in six different populations: Dutch, UK, Spanish, Finnish, Hungarian and Italian.

Moreover, additional sets of controls from each population were also genotyped, as well as a case-control sampleset from Milan. Although no previous *HLA*-typing results were available from these samples, the *HLA* allele frequencies determined by the SNP method followed the known *HLA* allele frequencies in these populations. For instance, the haplotype frequencies of the Italian (Milan) set are comparable to the haplotype frequencies reported by Margaritte-Jeannin *et al.* (2004). This genotyping method enables the identification of individuals being homozygous or heterozygous for the risk haplotypes, allowing us to study the risk of celiac disease conferred by different *HLA* genotypes. Our data supports several previous findings of the dose effect of functional DQ2 heterodimers in risk of celiac disease (Ploski *et al.* 1993, Louka *et al.* 2002, Vader *et al.* 2003, Margaritte-Jeannin *et al.* 2004).

Recently, the tagging SNP method was also shown to be useful for *HLA* testing in type 1 diabetes and it should be useful in other HLA DR3-DQ2 or DR4-DQ8 associated diseases, such as systemic lupus erythematosus and rheumatoid arthritis (Barker *et al.* 2008). The method can also be used to exclude diagnosis of celiac disease in the absence of HLA DQ2 or DQ8 when screening high-risk groups (e.g. relatives of patients with celiac disease, and

patients with Down syndrome), which reduces the cost of follow-up serology screening (Csizmadia *et al.* 2000, Kaukinen *et al.* 2002, Mustalahti *et al.* 2002a, Mearin *et al.* 2005, Romanos *et al.* 2008).

The main benefits of the SNP typing method are significant reductions in cost and time, as well as its potential as a high-throughput assay. The TaqMan method used in this study is most cost-reducing when applied to the 96 or 384 sample formats, making this method highly suitable for large sample cohorts, typically those used in research and population screening studies.

Conclusions and future challenges

In addition to the HLA DQ2 and DQ8 coding haplotypes, 14 susceptibility loci for celiac disease have now been identified through genome-wide linkage and association studies and their fine-mapping and follow-up studies. Many of these loci harbour more than one possible susceptibility gene, and only the role of DQ2 and DQ8 have been characterised in the pathogenesis of celiac disease. Several additional susceptibility genes have also been suggested by various candidate gene studies. The majority of the risk loci harbour genes with immunological functions, and many of these loci have also been associated with other autoimmune and inflammatory conditions. In particular, there appears to be multiple shared genetic risk factors for celiac disease, type 1 diabetes and inflammatory bowel disease, such as the *IL2/21* and *IL18RAP* loci. There are also several reports of shared risk genes between celiac disease, rheumatoid arthritis and systemic lupus erythematosus, such as the *TNFAIP3* locus.

The loci found by genome-wide linkage scans constitute large genetic regions, and the results of the fine-mapping studies have been somewhat contradictory. For example, *MYO9B* still constitutes a controversial candidate gene for celiac disease, although the original Dutch findings have been replicated in some other populations and similar pathologies. The lack of association with celiac disease in most replication studies may suggest a weaker effect size than estimated in the original Dutch study: a phenomenon often seen in genetic studies and referred to as "a winner's curse" (Ioannidis *et al.* 2001). More powerful study designs will probably be required to pinpoint the risk-conferring genes and variants in the *CELIAC2*, *CELIAC3*, and *CELIAC4* regions. Allelic heterogeneity and differences in the allele frequencies and LD patterns between populations are also possible explanations for the weak associations found and the discrepancies between different studies. The lack of association to these regions in the whole-genome association study by van Heel *et al.* (2007) might reflect insufficient power to detect associations in these regions, which may also have been a problem in previous fine-mapping studies.

Genome-wide association studies have been successful in detecting common risk variants (MAF > 1%) with modest effect sizes (OR 1.1 – 1.5). Even though several novel risk genes for complex diseases have been found through these scans, the effect of luck to detect risk variants with small effect sizes in previous genome-wide association studies should perhaps not be underestimated. It has been estimated that with 2000 cases and 3000 controls, the power to detect disease risk variants with minor allele frequencies above 5% is only 43% for alleles with a relative risk of 1.3 for a p-value threshold of 5×10^{-7} in a genome-wide association study published by the Wellcome Trust Case Control Consortium (2007). This implies that increasing sample sizes in genome-wide association scans should lead to substantial increases in the number of identified associations. This kind of studies with high statistical power to detect risk genes with small effect sizes will also be needed in the future for celiac disease, and they will most likely require multiple populations. There are already some examples of such meta-analyses of genome-wide association data, utilising thousands of cases and controls from several populations, leading to novel risk gene findings (Zeggini *et al.* 2008, Houlston *et al.* 2008, Cooper *et al.* 2008, Aulchenko *et al.* 2009).

The findings from genome-wide association studies have provided clues to understanding the architecture of the genetic risk of complex diseases in general. In addition to the identification of numerous novel susceptibility loci and new biological associations, these

studies have revealed future challenges. A substantial proportion of the detected risk variants are located in non-coding and intergenic regions (Hindorff *et al.* 2009) (http://www.genome.gov/26525384), calling for studies investigating the functions of non-coding variants. The associated SNPs are perhaps rarely the causal risk variants and techniques such as sequencing may be needed to identify the true functional risk variants. In celiac disease, all risk loci highlighted in the genome-wide association study and its follow-ups are located in intronic and intergenic regions within blocks of strong LD. The *IL18RAP* locus is an example of a locus where strong association is found and confirmed in multiple populations, but confirming the functions of the genes in pathogenesis requires further efforts.

To date, variants identified by genome-wide association studies only explain a small fraction of the overall genetic risk for disease, and a large portion of the genetic variance still remains to be found (Altshuler *et al.* 2008). For example, it has been estimated in celiac disease that the non-*HLA* risk SNPs found in the genome-wide association study and its follow-up studies account for only 3-4% of the heritability; the HLA region accounts for an additional 30% (Hunt and van Heel 2009). There is also evidence that the universe of rare structural changes contributing to each disease may be as extensive and diverse as that of common SNPs (Altshuler *et al.* 2008). Such variants are beyond the reach of current genome-wide association technologies. In addition, gene-gene interactions should be investigated further because they may have substantial effects on disease risk, despite the risk conferred by individual genes being small.

Functional candidate gene studies have suggested a number of susceptibility genes for celiac disease, but their role in disease risk still remains to be confirmed. Multiple studies have shown modest associations, often implying an under-powered study or perhaps even false-positive or false-negative findings. The findings of these studies often fail to be consistently replicated (Hirschhorn *et al.* 2002, Gorroochurn *et al.* 2007). Indeed, as we have learnt from genome-wide association scans, the effect sizes of the common risk-conferring variants for complex diseases are modest. Perhaps the risk effect has been estimated to be larger than it really is, leading to ambiguous and non-replicated/non-replicable association results. Genome-wide association studies may be making candidate gene studies a less popular approach to study disease-gene correlations, but there is still a niche for well-designed and well-powered candidate gene studies as well, if an interesting gene is proposed. Some loci or haplotypes may be missed in genome-wide studies, if none of the variants are in LD with the SNPs used in the genotyping arrays, or else because of allelic heterogeneity (Altshuler *et al.* 2008).

Finally, one may ask whether the high number of susceptibility genes, each only marginally increasing disease risk, will have clinical consequences. The novel association findings give us valuable information on the biology and molecular pathways underlying disease pathogenesis, thereby providing us with tools to create novel diagnostics and therapeutics in the future. Predictive gene testing could also be performed if modest risk effects can help guide clinical decision-making and increase knowledge about the efficiency of clinical intervention, and if there are means for preventive interventions.

To date, the *HLA* locus remains the strongest single risk locus for celiac disease, with the other identified loci having risk effects of less than 2-fold. Only genotyping of *HLA DQ* haplotypes is currently used to exclude the diagnosis of celiac disease, and a new tool for screening populations or other large samplesets has now been developed (Monsuur *et al.* 2008). We are beginning to understand more about the genetics of celiac disease, as well as

the nature and function of susceptibility genes and their interactions with each other and their environment. This new information will aid our understanding of the pathogenesis of celiac disease, hopefully leading to better patient outcomes. New advancements in the field of complex disease genetics can be expected in the near future as large-scale genome-wide association scans and high-throughput sequencing will take place. Such approaches are expected to shed further light on genetic susceptibility factors of common complex diseases.

Acknowledgements

This Ph.D. study was done at the Department of Medical Genetics at the University of Helsinki, Finland, during years 2005-2009. The former, and the current heads of the department, Professors Leena Peltonen-Palotie and Päivi Peltomäki, are acknowledged for providing excellent research facilities.

This work has been financially supported by the EU Commission Marie Curie Excellence Grant, the Academy of Finland, the University of Helsinki Funds, Biocentrum Helsinki, the Foundation of Pediatric Research, the Sigrid Juselius Foundation, the Finnish Cultural Foundation, the Maud Kuistila Memorial Foundation, and the Finnish Society for Gastroenterological Research. Personal support was provided by the Finnish Celiac Disease Society and University of Helsinki Funds. All sources of funding are gratefully acknowledged.

I wish to dedicate my deepest gratitude to all the voluntary patients and families who have participated in this study. Without their valuable contribution this work would not have been possible.

I have been privileged in having a large network of wonderful colleagues and collaborators who have contributed to my thesis project in many ways, and I wish to thank all of them, especially:

I am most grateful to my supervisor, Docent Päivi Saavalainen, for these years as her Ph.D. student: during these years I have learned and experienced much more than I could ever imagine. Päivi, thank you for sharing your vast knowledge on celiac disease and genetics, and for giving me a perfect combination of independence and guidance in my work. I truly admire your enthusiasm and ambition towards science, and the positive touch you have in everything you do. You are always full of new ideas and you have the ability to inspire people around you. Thank you for all your trust and support!

I am sincerely grateful to Professor Juha Kere for supervision and for all the support he has given to me and to the celiac disease genetics project during these years. I admire Juha's experience, positivity and dynamic drive: I have learnt incredibly much from him and his research group.

My cordial thanks to Docent Jukka Partanen for all the support he has given to me, especially in my very early steps towards Ph.D. I am grateful to Jukka for introducing me to the world of HLA and the celiac disease genetics project.

I wish to thank Professors Katarina Pelin and Outi Vaarala for the official review of my Ph.D. thesis. I am grateful for their excellent, constructive comments and suggestions which helped me to improve my thesis.

I am grateful to the members of my thesis committee, Docents Päivi Onkamo and Tiinamaija Tuomi, for all valuable concern, comments and suggestions they have given on my thesis project. I think our yearly meetings have been helpful and inspiring.

I am grateful to Professor Markku Mäki and his research group for fantastic collaboration, not only for providing us with patient samples, but also for always giving a helping hand when needed. I would like to thank especially Docent Katri Kaukinen and Doctors Kalle Kurppa, Katri Lindfors, Kati Juuti-Uusitalo and Keijo Viiri. I would also like to express my sincere gratitude to Docent Ilma Korponay-Szabó for the great collaboration and for her valuable time explaining and teaching me about the clinical side of celiac disease.

I am grateful to Professor Alessandro Ventura and Doctors Fabiana Ziberna, Tarcisio Not and Serena Vatta for their collaboration. I would also like to acknowledge Professor Róza Ádány and her research group, as well as Docent Pertti Sistonen for their valuable contributions to our studies.

I am highly thankful to Professor Cisca Wijmenga for extremely fruitful collaboration, which has truly been essential and important for my Ph.D. project. I would also like to thank Jihane Romanos from Cisca Wijmenga's group: it was nice to work with you! I would like to thank Professor David van Heel and his research group for being great collaborators, for sharing data and their expertise.

I am extremely thankful to Elísabet Einarsdóttir for sharing with me the research projects and the everyday life in the lab, in the office and while teaching: it has been a pleasure to work with such an excellent and enthusiastic geneticist. Elísabet is also thanked for giving great comments concerning my thesis manuscript. I would like to thank Emma Dukes for being extremely helpful in the lab in planning and setting up experiments. I am very grateful to Andrea de Kauwe and Emma Dukes for being my English consultants and for always helping me whenever I ran into problems with understanding immunology. Andrea de Kauwe is also greatly acknowledged for carefully revising the English in this thesis. I feel privileged having had such skilled post-docs working with me!

I am most grateful to Hanne Ahola for the laboratory assistance she has given me during the past years. It has been great to have such a skilled person taking care of the samples and everything in the lab!

During my years as a Ph.D. student I have learnt to know a number of wonderful people. I would like to thank all the current and previous members of Päivi's team for their cheerful company, various science-related and -unrelated discussions and all the fun and unforgettable moments we have spent in and outside work and at conference trips. All the members of Juha's group are warmly thanked for the friendly working atmosphere and joyful moments at lunch breaks, group trips, jogging events and singing rehearsals! I would also like to thank Markku's and Jukka's groups for the great company at various meetings, celebrations and conferences.

Dear friends outside work, I want to thank all of you. You keep reminding me there is life outside Biomedicum too, and I am grateful for your company at different activities and travelling to some of the most amazing places in the world. Thank you for all the unforgettable moments we have passed together: your friendship is important to me.

I would like to owe my warmest thanks to my family members for all the support they have given to me along the years. I am grateful to my parents Leena and Kari for encouraging and supporting me in my life and studies. My brother Ilkka and my sister-in-law Hanna are warmly thanked for their friendship. Many thanks to Emil and Anton, my lovely nephews, for all the lively and happy moments together.

I am grateful to Ville for his presence and love, and for being there for me.

Satta Helsinki, summer 2009

References

Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. Nat.Genet. 30:97-101. 2002.

Abel M, Cellier C, Kumar N, Cerf-Bensussan N, Schmitz J, Caillat-Zucman S. Adulthood-onset celiac disease is associated with intercellular adhesion molecule-1 (ICAM-1) gene polymorphism. Hum.Immunol. 67:612-617. 2006.

Adamovic S, Amundsen SS, Lie BA, Gudjonsdottir AH, Ascher H, Ek J, van Heel DA, Nilsson S, Sollid LM, Torinsson Naluai A. Association study of IL2/IL21 and FcgRIIa: significant association with the IL2/IL21 region in Scandinavian coeliac disease families. Genes Immun. 9:364-367. 2008a.

Adamovic S, Amundsen SS, Lie BA, Hellqvist A, Gudjonsdottir AH, Ek J, Nilsson S, Wahlstrom J, Ascher H, Sollid LM, Naluai AT. Fine mapping study in Scandinavian families suggests association between coeliac disease and haplotypes in chromosome region 5q32. Tissue Antigens 71:27-34. 2008b.

Albers HM, Kurreeman FA, Stoeken-Rijsbergen G, Brinkman DM, Kamphuis SS, van Rossum MA, Girschick HJ, Wouters C, Saurenmann RK, Hoppenreijs E, Slagboom P, Houwing-Duistermaat JJ, Verduijn W, Huizinga TW, Ten Cate R, Toes RE, Schilham MW. Association of the autoimmunity locus 4q27 with juvenile idiopathic arthritis. Arthritis Rheum. 60:901-904. 2009.

Alizadeh BZ, Valdigem G, Coenen MJ, Zhernakova A, Franke B, Monsuur A, van Riel PL, Barrera P, Radstake TR, Roep BO, Wijmenga C, Koeleman BP. Association analysis of functional variants of the FcgRIIa and FcgRIIIa genes with type 1 diabetes, celiac disease and rheumatoid arthritis. Hum.Mol.Genet. 16:2552-2559. 2007.

Al-Toma A, Verbeek WH, Mulder CJ. Update on the management of refractory coeliac disease. J.Gastrointestin Liver Dis. 16:57-63. 2007.

Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science 322:881-888. 2008.

Amundsen SS, Adamovic S, Hellqvist A, Nilsson S, Gudjonsdottir AH, Ascher H, Ek J, Larsson K, Wahlstrom J, Lie BA, Sollid LM, Naluai AT. A comprehensive screen for SNP associations on chromosome region 5q31-33 in Swedish/Norwegian celiac disease families. Eur.J.Hum.Genet. 15:980-987. 2007.

Amundsen SS, Monsuur AJ, Wapenaar MC, Lie BA, Ek J, Gudjonsdottir AH, Ascher H, Wijmenga C, Sollid LM. Association analysis of MYO9B gene polymorphisms with celiac disease in a Swedish/Norwegian cohort. Hum.Immunol. 67:341-345. 2006a.

Amundsen SS, Vatn M, IBSEN study group, Wijmenga C, Sollid LM, Lie BA. Association analysis of MYO9B gene polymorphisms and inflammatory bowel disease in a Norwegian cohort. Tissue Antigens 68:249-252. 2006b.

Andre R, Wheeler RD, Collins PD, Luheshi GN, Pickering-Brown S, Kimber I, Rothwell NJ, Pinteaux E. Identification of a truncated IL-18R beta mRNA: a putative regulator of IL-18 expressed in rat brain. J.Neuroimmunol. 145:40-45. 2003.

Ankelo M, Kleimola V, Simell S, Simell O, Knip M, Jokisalo E, Tarkia M, Westerlund A, He Q, Viander M, Ilonen J, Hinkkanen AE. Antibody responses to deamidated gliadin peptide show high specificity and parallel antibodies to tissue transglutaminase in developing coeliac disease. Clin.Exp.Immunol. 150:285-293. 2007.

Arentz-Hansen H, Fleckenstein B, Molberg O, Scott H, Koning F, Jung G, Roepstorff P, Lundin KE, Sollid LM. The molecular basis for oat intolerance in patients with celiac disease. PLoS Med. 1:e1. 2004.

Arentz-Hansen H, Korner R, Molberg O, Quarsten H, Vader W, Kooy YM, Lundin KE, Koning F, Roepstorff P, Sollid LM, McAdam SN. The intestinal T cell response to alpha-gliadin in adult celiac disease is focused on a single deamidated glutamine targeted by tissue transglutaminase. J.Exp.Med. 191:603-612. 2000.

Arentz-Hansen H, McAdam SN, Molberg O, Fleckenstein B, Lundin KE, Jorgensen TJ, Jung G, Roepstorff P, Sollid LM. Celiac lesion T cells recognize epitopes that cluster in regions of gliadins rich in proline residues. Gastroenterology 123:803-809. 2002.

Aromaa A, Koskinen S. Health and functional capacity in Finland. Baseline results of the Health 2000 Health Examination Survey. Publications of the National Public Health Institute, Helsinki, Finland. 2004. 171pp.

Aulchenko YS, Ripatti S, Lindqvist I, Boomsma D, Heid IM, Pramstaller PP, Penninx BW, Janssens AC, Wilson JF, Spector T, Martin NG, Pedersen NL, Kyvik KO, Kaprio J, Hofman A, Freimer NB, Jarvelin MR, Gyllensten U, Campbell H, Rudan I, Johansson A, Marroni F, Hayward C, Vitart V, Jonasson I, Pattaro C, Wright A, Hastie N, Pichler I, Hicks AA, Falchi M, Willemsen G, Hottenga JJ, de Geus EJ, Montgomery GW, Whitfield J, Magnusson P, Saharinen J, Perola M, Silander K, Isaacs A, Sijbrands EJ, Uitterlinden AG, Witteman JC, Oostra BA, Elliott P, Ruokonen A, Sabatti C, Gieger C, Meitinger T, Kronenberg F, Doring A, Wichmann HE, Smit JH, McCarthy MI, van Duijn CM, Peltonen L, ENGAGE Consortium. Loci influencing lipid levels and coronary heart disease risk in 16 European population cohorts. Nat.Genet. 41:47-55. 2009.

Babron MC, Nilsson S, Adamovic S, Naluai AT, Wahlstrom J, Ascher H, Ciclitira PJ, Sollid LM, Partanen J, Greco L, Clerget-Darpoux F, European Genetics Cluster on Coeliac Disease. Meta and pooled analysis of European coeliac disease data. Eur.J.Hum.Genet. 11:828-834. 2003.

Bach JF. The effect of infections on susceptibility to autoimmune and allergic diseases. N.Engl.J.Med. 347:911-920. 2002.

Bahram S, Inoko H, Shiina T, Radosavljevic M. MIC and other NKG2D ligands: from none to too many. Curr.Opin.Immunol. 17:505-509. 2005.

Bardella MT, Fredella C, Prampolini L, Marino R, Conte D, Giunta AM. Gluten sensitivity in monozygous twins: a long-term follow-up of five pairs. Am.J.Gastroenterol. 95:1503-1505. 2000.

Barker JM, Triolo TM, Aly TA, Baschal EE, Babu SR, Kretowski A, Rewers MJ, Eisenbarth GS. Two single nucleotide polymorphisms identify the highest-risk diabetes HLA genotype: potential for rapid screening. Diabetes 57:3152-3155. 2008.

Barone MV, Gimigliano A, Castoria G, Paolella G, Maurano F, Paparo F, Maglio M, Mineo A, Miele E, Nanayakkara M, Troncone R, Auricchio S. Growth factor-like activity of gliadin, an alimentary protein: implications for coeliac disease. Gut 56:480-488. 2007.

Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21:263-265. 2005.

Baugh JA, Chitnis S, Donnelly SC, Monteiro J, Lin X, Plant BJ, Wolfe F, Gregersen PK, Bucala R. A functional promoter polymorphism in the macrophage migration inhibitory factor (MIF) gene associated with disease severity in rheumatoid arthritis. Genes Immun. 3:170-176. 2002.

Begovich AB, Carlton VE, Honigberg LA, Schrodi SJ, Chokkalingam AP, Alexander HC, Ardlie KG, Huang Q, Smith AM, Spoerke JM, Conn MT, Chang M, Chang SY, Saiki RK, Catanese JJ, Leong DU, Garcia VE, McAllister LB, Jeffery DA, Lee AT, Batliwalla F, Remmers E, Criswell LA, Seldin MF, Kastner DL, Amos CI, Sninsky JJ, Gregersen PK. A missense single-nucleotide polymorphism in a gene encoding a protein tyrosine phosphatase (PTPN22) is associated with rheumatoid arthritis. Am.J.Hum.Genet. 75:330-337. 2004.

Belguise K, Sonenshein GE. PKCtheta promotes c-Rel-driven mammary tumorigenesis in mice and humans by repressing estrogen receptor alpha synthesis. J.Clin.Invest. 117:4009-4021. 2007.

Bevan S, Popat S, Braegger CP, Busch A, O'Donoghue D, Falth-Magnusson K, Ferguson A, Godkin A, Hogberg L, Holmes G, Hosie KB, Howdle PD, Jenkins H, Jewell D, Johnston S, Kennedy NP, Kerr G, Kumar P, Logan RF, Love AH, Marsh M, Mulder CJ, Sjoberg K, Stenhammer L, Walker-Smith J, Marossy AM, Houlston RS. Contribution of the MHC region to the familial risk of coeliac disease. J.Med.Genet. 36:687-690. 1999.

Boone DL, Turer EE, Lee EG, Ahmad RC, Wheeler MT, Tsui C, Hurley P, Chien M, Chai S, Hitotsumatsu O, McNally E, Pickart C, Ma A. The ubiquitin-modifying enzyme A20 is required for termination of Toll-like receptor responses. Nat.Immunol. 5:1052-1060. 2004.

Bottini N, Musumeci L, Alonso A, Rahmouni S, Nika K, Rostamkhani M, MacMurray J, Meloni GF, Lucarelli P, Pellecchia M, Eisenbarth GS, Comings D, Mustelin T. A functional variant of lymphoid tyrosine phosphatase is associated with type I diabetes. Nat.Genet. 36:337-338. 2004.

Bowness JM, Tarr AH, Wong T. Increased transglutaminase activity during skin wound healing in rats. Biochim.Biophys.Acta 967:234-240. 1988.

Brandt K, Singh PB, Bulfone-Paus S, Ruckert R. Interleukin-21: a new modulator of immunity, infection, and cancer. Cytokine Growth Factor Rev. 18:223-232. 2007.

Brophy K, Ryan AW, Thornton JM, Abuzakouk M, Fitzgerald AP, McLoughlin RM, O'morain C, Kennedy NP, Stevens FM, Feighery C, Kelleher D, McManus R. Haplotypes in the CTLA4 region are associated with coeliac disease in the Irish population. Genes Immun. 7:19-26. 2006.

Bruewer M, Hopkins AM, Hobert ME, Nusrat A, Madara JL. RhoA, Rac1, and Cdc42 exert distinct effects on epithelial barrier via selective structural and biochemical modulation of junctional proteins and F-actin. Am.J.Physiol.Cell.Physiol. 287:C327-35. 2004.

Buza-Vidas N, Antonchuk J, Qian H, Mansson R, Luc S, Zandi S, Anderson K, Takaki S, Nygren JM, Jensen CT, Jacobsen SE. Cytokines regulate postnatal hematopoietic stem cell expansion: opposing roles of thrombopoietin and LNK. Genes Dev. 20:2018-2023. 2006.

Camarca A, Anderson RP, Mamone G, Fierro O, Facchiano A, Costantini S, Zanzi D, Sidney J, Auricchio S, Sette A, Troncone R, Gianfrani C. Intestinal T cell responses to gluten peptides are largely heterogeneous: implications for a peptide-based therapy in celiac disease. J.Immunol. 182:4158-4166. 2009.

Castellanos-Rubio A, Martin-Pagola A, Santin I, Hualde I, Aransay AM, Castano L, Vitoria JC, Bilbao JR. Combined functional and positional gene information for the identification of susceptibility variants in celiac disease. Gastroenterology 134:738-746. 2008.

Castellanos-Rubio A, Santin I, Irastorza I, Castano L, Carlos Vitoria J, Ramon Bilbao J. TH17 (and TH1) signatures of intestinal biopsies of CD patients in response to gliadin. Autoimmunity 42:69-73. 2009.

Central Statistical Office. Demographic Yearbook of Hungary. Budapest, Hungary. 2001.

Chernavsky AC, Paez MC, Periolo N, Correa P, Guillen L, Niveloni SI, Maurino E, Bai JC, Anaya JM. The simultaneous presence of IL-1B and TNFA two-positions risk haplotypes enhances the susceptibility for celiac disease. Cytokine 42:48-54. 2008.

Cirillo G, Di Domenico MR, Corsi I, Gagliardo T, Del Giudice EM, Perrone L, Tolone C. Do MYO9B genetic variants predispose to coeliac disease? An association study in a cohort of South Italian children. Dig.Liver Dis. 39:228-231. 2007.

Clayton D. SNPHAP - A program for estimating frequencies of large haplotypes of SNPs. v.1.3: 2002.

Collin P, Reunala T. Recognition and management of the cutaneous manifestations of celiac disease: a guide for dermatologists. Am.J.Clin.Dermatol. 4:13-20. 2003.

Cookson WO, Moffatt MF. Genetics of asthma and allergic disease. Hum.Mol.Genet. 9:2359-2364. 2000.

Cooney R, Cummings JR, Pathan S, Beckly J, Geremia A, Hancock L, Guo C, Morris A, Jewell DP. Association between genetic variants in myosin IXB and Crohn's disease. Inflamm.Bowel Dis. 2009.

Cooper JD, Smyth DJ, Smiles AM, Plagnol V, Walker NM, Allen JE, Downes K, Barrett JC, Healy BC, Mychaleckyj JC, Warram JH, Todd JA. Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. Nat.Genet. 40:1399-1401. 2008.

Csizmadia CG, Mearin ML, Oren A, Kromhout A, Crusius JB, von Blomberg BM, Pena AS, Wiggers MN, Vandenbroucke JP. Accuracy and cost-effectiveness of a new strategy to screen for celiac disease in children with Down syndrome. J.Pediatr. 137:756-761. 2000.

Curley CR, Monsuur AJ, Wapenaar MC, Rioux JD, Wijmenga C. A functional candidate screen for coeliac disease genes. Eur.J.Hum.Genet. 14:1215-1222. 2006.

Dardalhon V, Korn T, Kuchroo VK, Anderson AC. Role of Th1 and Th17 cells in organ-specific autoimmunity. J.Autoimmun. 31:252-256. 2008.

de Bakker PI, McVean G, Sabeti PC, Miretti MM, Green T, Marchini J, Ke X, Monsuur AJ, Whittaker P, Delgado M, Morrison J, Richardson A, Walsh EC, Gao X, Galver L, Hart J, Hafler DA, Pericak-Vance M, Todd JA, Daly MJ, Trowsdale J, Wijmenga C, Vyse TJ, Beck S, Murray SS, Carrington M, Gregory S, Deloukas P, Rioux JD. A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. Nat.Genet. 38:1166-1172. 2006.

de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat.Genet. 37:1217-1223. 2005.

Dema B, Martinez A, Fernandez-Arquero M, Maluenda C, Polanco I, G de la Concha E, Urcelay E, Nunez C. Association of IL18RAP and CCR3 with celiac disease in the Spanish population. J.Med.Genet. 2009.

Dicke WK. Coeliakie. MD Thesis. Utrecht University, Utrecht, Netherlands. 1950.

Dieterich W, Ehnis T, Bauer M, Donner P, Volta U, Riecken EO, Schuppan D. Identification of tissue transglutaminase as the autoantigen of celiac disease. Nat.Med. 3:797-801. 1997.

Ding YC, Weizman Z, Yerushalmi B, Elbedour K, Garner CP, Neuhausen SL. An autosomal genome-wide screen for celiac disease in Bedouin families. Genes Immun. 9:81-86. 2008.

Donn RP, Plant D, Jury F, Richards HL, Worthington J, Ray DW, Griffiths CE. Macrophage migration inhibitory factor gene polymorphism is associated with psoriasis. J.Invest.Dermatol. 123:484-487. 2004.

Donn RP, Shelley E, Ollier WE, Thomson W, British Paediatric Rheumatology Study Group. A novel 5'-flanking region polymorphism of macrophage migration inhibitory factor is associated with systemic-onset juvenile idiopathic arthritis. Arthritis Rheum. 44:1782-1785. 2001.

Dube C, Rostom A, Sy R, Cranney A, Saloojee N, Garritty C, Sampson M, Zhang L, Yazdi F, Mamaladze V, Pan I, Macneil J, Mack D, Patel D, Moher D. The prevalence of celiac disease in average-risk and at-risk Western European populations: a systematic review. Gastroenterology 128:S57-67. 2005.

Dudbridge F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. Hum.Hered. 66:87-98. 2008.

Duerr RH, Taylor KD, Brant SR, Rioux JD, Silverberg MS, Daly MJ, Steinhart AH, Abraham C, Regueiro M, Griffiths A, Dassopoulos T, Bitton A, Yang H, Targan S, Datta LW, Kistner EO, Schumm LP, Lee AT, Gregersen PK, Barmada MM, Rotter JI, Nicolae DL, Cho JH. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science 314:1461-1463. 2006.

Eaton WW, Byrne M, Ewald H, Mors O, Chen CY, Agerbo E, Mortensen PB. Association of schizophrenia and autoimmune diseases: linkage of Danish national registers. Am.J.Psychiatry 163:521-528. 2006.

Einarsdottir E, Koskinen LL, Dukes E, Kainu K, Suomela S, Lappalainen M, Ziberna F, Korponay-Szabó IR, Kurppa K, Kaukinen K, Ádány R, Pocsai Z, Széles G, Farkkila M, Turunen U, Halme L, Paavola-Sakki P, Not T, Vatta S, Ventura A, Lofberg R, Torkvist L, Bresso F, Halfvarson J, Mäki M, Kontula K, Saarialho-Kere U, Kere J, D'Amato M, Saavalainen P. IL23R in the Swedish, Finnish, Hungarian and Italian populations: association with IBD and psoriasis, and linkage to celiac disease. BMC Med.Genet. 10:8. 2009.

Ettinger R, Sims GP, Fairhurst AM, Robbins R, da Silva YS, Spolski R, Leonard WJ, Lipsky PE. IL-21 induces differentiation of human naive and memory B cells into antibody-secreting plasma cells. J.Immunol. 175:7867-7879. 2005.

Falchuk ZM, Strober W. HL-A antigens and adult coeliac disease. Lancet 2:1310. 1972.

Festen EA, Goyette P, Scott R, Annese V, Zhernakova A, Brant SR, Cho JH, Silverberg MS, Taylor KD, de Jong DJ, Stokkers PC, McGovern D, Palmieri O, Achkar JP, Xavier RJ, Duerr RH, Daly MJ, Wijmenga C, Weersma RK, Rioux JD. Genetic Variants in the Region Harbouring IL2/IL21 Associated to Ulcerative Colitis. Gut 2009.

Fiszer D, Rozwadowska N, Rychlewski L, Kosicki W, Kurpisz M. Identification of IL-18RAP mRNA truncated splice variants in human testis and the other human tissues. Cytokine 39:178-183. 2007.

Fleckenstein B, Molberg O, Qiao SW, Schmid DG, von der Mulbe F, Elgstoen K, Jung G, Sollid LM. 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. 277:34109-34116. 2002.

Folk JE, Chung SI. Transglutaminases. Methods Enzymol. 113:358-375. 1985.

Folk JE, Cole PW. Transglutaminase: mechanistic features of the active site as determined by kinetic and inhibitor studies. Biochim.Biophys.Acta 122:244-264. 1966.

Forsberg G, Hernell O, Melgar S, Israelsson A, Hammarstrom S, Hammarstrom ML. Paradoxical coexpression of proinflammatory and down-regulatory cytokines in intestinal T cells in childhood celiac disease. Gastroenterology 123:667-678. 2002.

Fry L, Seah PP, Riches DJ, Hoffbrand AV. Clearance of skin lesions in dermatitis herpetiformis after gluten withdrawal. Lancet 1:288-291. 1973.

Fung EY, Smyth DJ, Howson JM, Cooper JD, Walker NM, Stevens H, Wicker LS, Todd JA. Analysis of 17 autoimmune disease-associated variants in type 1 diabetes identifies 6q23/TNFAIP3 as a susceptibility locus. Genes Immun. 10:188-191. 2009.

Garner CP, Ding YC, Steele L, Book L, Leiferman K, Zone JJ, Neuhausen SL. Genome-wide linkage analysis of 160 North American families with celiac disease. Genes Immun. 8:108-114. 2007.

Gawkrodger DJ, Blackwell JN, Gilmour HM, Rifkind EA, Heading RC, Barnetson RS. Dermatitis herpetiformis: diagnosis, diet and demography. Gut 25:151-157. 1984.

Gee S. On the coeliac disease. St. Bart Hosp Rep 24:17-20. 1888.

Giordano M, Marano C, Mellai M, Limongelli MG, Bolognesi E, Clerget-Darpoux F, Momigliano-Richiardi P, Greco L. A family-based study does not confirm the association of MYO9B with celiac disease in the Italian population. Genes Immun. 7:606-608. 2006.

Giovannini C, Matarrese P, Scazzocchio B, Vari R, D'Archivio M, Straface E, Masella R, Malorni W, De Vincenzi M. Wheat gliadin induces apoptosis of intestinal cells via an autocrine mechanism involving Fas-Fas ligand pathway. FEBS Lett. 540:117-124. 2003.

Giovannini C, Sanchez M, Straface E, Scazzocchio B, Silano M, De Vincenzi M. Induction of apoptosis in caco-2 cells by wheat gliadin peptides. Toxicology 145:63-71. 2000.

Gordon D, Finch SJ, Nothnagel M, Ott J. Power and sample size calculations for case-control genetic association tests when errors are present: application to single nucleotide polymorphisms. Hum.Hered. 54:22-33. 2002.

Gordon D, Levenstien MA, Finch SJ, Ott J. Errors and linkage disequilibrium interact multiplicatively when computing sample sizes for genetic case-control association studies. Pac.Symp.Biocomput. 490-501. 2003.

Gorroochurn P, Hodge SE, Heiman GA, Durner M, Greenberg DA. Non-replication of association studies: "pseudo-failures" to replicate? Genet.Med. 9:325-331. 2007.

Gough SC, Walker LS, Sansom DM. CTLA4 gene polymorphism and autoimmunity. Immunol.Rev. 204:102-115. 2005.

Graham RR, Cotsapas C, Davies L, Hackett R, Lessard CJ, Leon JM, Burtt NP, Guiducci C, Parkin M, Gates C, Plenge RM, Behrens TW, Wither JE, Rioux JD, Fortin PR, Graham DC, Wong AK, Vyse TJ, Daly MJ, Altshuler D, Moser KL, Gaffney PM. Genetic variants near TNFAIP3 on 6q23 are associated with systemic lupus erythematosus. Nat.Genet. 40:1059-1061. 2008.

Greco L, Babron MC, Corazza GR, Percopo S, Sica R, Clot F, Fulchignoni-Lataud MC, Zavattari P, Momigliano-Richiardi P, Casari G, Gasparini P, Tosi R, Mantovani V, De Virgiliis S, Iacono G, D'Alfonso A, Selinger-Leneman H, Lemainque A, Serre JL, Clerget-Darpoux F. Existence of a genetic risk factor on chromosome 5q in Italian coeliac disease families. Ann.Hum.Genet. 65:35-41. 2001.

Greco L, Corazza G, Babron MC, Clot F, Fulchignoni-Lataud MC, Percopo S, Zavattari P, Bouguerra F, Dib C, Tosi R, Troncone R, Ventura A, Mantavoni W, Magazzu G, Gatti R, Lazzari R, Giunta A, Perri F, Iacono G, Cardi E, de Virgiliis S, Cataldo F, De Angelis G, Musumeci S, Clerget-Darpoux F. Genome search in celiac disease. Am.J.Hum.Genet. 62:669-675. 1998.

Greco L, Romino R, Coto I, Di Cosmo N, Percopo S, Maglio M, Paparo F, Gasperi V, Limongelli MG, Cotichini R, D'Agate C, Tinto N, Sacchetti L, Tosi R, Stazi MA. The first large population based twin study of coeliac disease. Gut 50:624-628. 2002.

Haimila K, Einarsdottir E, de Kauwe A, Koskinen LL, Pan-Hammarstrom Q, Kaartinen T, Kurppa K, Ziberna F, Not T, Vatta S, Ventura A, Korponay-Szabó IR, Ádány R, Pocsai Z, Széles G, Dukes E, Kaukinen K, Mäki M, Koskinen S, Partanen J, Hammarstrom L, Saavalainen P. The shared CTLA4-ICOS risk locus in celiac disease, IgA deficiency and common variable immunodeficiency. Genes Immun. 10:151-161. 2009.

Haimila K, Smedberg T, Mustalahti K, Mäki M, Partanen J, Holopainen P. Genetic association of coeliac disease susceptibility to polymorphisms in the ICOS gene on chromosome 2q33. Genes Immun. 5:85-92. 2004.

Halttunen T, Mäki M. Serum immunoglobulin A from patients with celiac disease inhibits human T84 intestinal crypt epithelial cell differentiation. Gastroenterology 116:566-572. 1999.

Haroon ZA, Hettasch JM, Lai TS, Dewhirst MW, Greenberg CS. Tissue transglutaminase is expressed, active, and directly involved in rat dermal wound healing and angiogenesis. FASEB J. 13:1787-1795. 1999.

Hervonen K, Hakanen M, Kaukinen K, Collin P, Reunala T. First-degree relatives are frequently affected in coeliac disease and dermatitis herpetiformis. Scand.J.Gastroenterol. 37:51-55. 2002.

Hervonen K, Karell K, Holopainen P, Collin P, Partanen J, Reunala T. Concordance of dermatitis herpetiformis and celiac disease in monozygous twins. J.Invest.Dermatol. 115:990-993. 2000.

Hindorff LA, Junkins HA, Mehta JP and Manolio TA (2009) A Catalog of Published Genome-Wide Association Studies. 2009:02/27 http://www.genome.gov/26525384

Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. A comprehensive review of genetic association studies. Genet.Med. 4:45-61. 2002.

Holm K, Mäki M, Vuolteenaho N, Mustalahti K, Ashorn M, Ruuska T, Kaukinen K. Oats in the treatment of childhood coeliac disease: a 2-year controlled trial and a long-term clinical follow-up study. Aliment.Pharmacol.Ther. 23:1463-1472. 2006.

Holopainen P, Mustalahti K, Uimari P, Collin P, Mäki M, Partanen J. Candidate gene regions and genetic heterogeneity in gluten sensitivity. Gut 48:696-701. 2001.

Holopainen P, Naluai AT, Moodie S, Percopo S, Coto I, Clot F, Ascher H, Sollid L, Ciclitira P, Greco L, Clerget-Darpoux F, Partanen J, Members of the European Genetics Cluster on Coeliac Disease. Candidate gene region 2q33 in European families with coeliac disease. Tissue Antigens 63:212-222. 2004.

Houlston RS, Webb E, Broderick P, Pittman AM, Di Bernardo MC, Lubbe S, Chandler I, Vijayakrishnan J, Sullivan K, Penegar S, Colorectal Cancer Association Study Consortium, Carvajal-Carmona L, Howarth K, Jaeger E, Spain SL, Walther A, Barclay E, Martin L, Gorman M, Domingo E, Teixeira AS, CoRGI Consortium, Kerr D, Cazier JB, Niittymäki I, Tuupanen S, Karhu A, Aaltonen LA, Tomlinson IP, Farrington SM, Tenesa A, Prendergast JG, Barnetson RA, Cetnarskyj R, Porteous ME, Pharoah PD, Koessler T, Hampe J, Buch S, Schafmayer C, Tepel J, Schreiber S, Volzke H, Chang-Claude J, Hoffmeister M, Brenner H, Zanke BW, Montpetit A, Hudson TJ, Gallinger S, International Colorectal Cancer Genetic Association Consortium, Campbell H, Dunlop MG. Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. Nat.Genet. 40:1426-1435. 2008.

Hovhannisyan Z, Weiss A, Martin A, Wiesner M, Tollefsen S, Yoshida K, Ciszewski C, Curran SA, Murray JA, David CS, Sollid LM, Koning F, Teyton L, Jabri B. The role of HLA-DQ8 beta57 polymorphism in the anti-gluten T-cell response in coeliac disease. Nature 456:534-538. 2008.

Hue S, Mention JJ, Monteiro RC, Zhang S, Cellier C, Schmitz J, Verkarre V, Fodil N, Bahram S, Cerf-Bensussan N, Caillat-Zucman S. A direct role for NKG2D/MICA interaction in villous atrophy during celiac disease. Immunity 21:367-377. 2004.

Hunt KA, Franke L, Deloukas P, Wijmenga C, van Heel DA. No evidence in a large UK collection for celiac disease risk variants reported by a Spanish study. Gastroenterology 134:1629-30; author reply 1630-1. 2008.

Hunt KA, McGovern DP, Kumar PJ, Ghosh S, Travis SP, Walters JR, Jewell DP, Playford RJ, van Heel DA. A common CTLA4 haplotype associated with coeliac disease. Eur.J.Hum.Genet. 13:440-444. 2005.

Hunt KA, Monsuur AJ, McArdle WL, Kumar PJ, Travis SP, Walters JR, Jewell DP, Strachan DP, Playford RJ, Wijmenga C, van Heel DA. Lack of association of MYO9B genetic variants with coeliac disease in a British cohort. Gut 55:969-972. 2006.

Hunt KA, van Heel DA. Recent advances in coeliac disease genetics. Gut 58:473-476. 2009.

Hunt KA, Zhernakova A, Turner G, Heap GA, Franke L, Bruinenberg M, Romanos J, Dinesen LC, Ryan AW, Panesar D, Gwilliam R, Takeuchi F, McLaren WM, Holmes GK, Howdle PD, Walters JR, Sanders DS, Playford RJ, Trynka G, Mulder CJ, Mearin ML, Verbeek WH, Trimble V, Stevens FM, O'Morain C, Kennedy NP, Kelleher D, Pennington DJ, Strachan DP, McArdle WL, Mein CA, Wapenaar MC, Deloukas P, McGinnis R, McManus R, Wijmenga C, van Heel DA. Newly identified genetic risk variants for celiac disease related to the immune response. Nat.Genet. 40:395-402. 2008.

International HapMap Consortium. The International HapMap Project. Nature 426:789-96. 2003.

International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, Belmont JW, Boudreau A, Hardenbol P, Leal SM, Pasternak S, Wheeler DA, Willis TD, Yu F, Yang H, Zeng C, Gao Y, Hu H, Hu W, Li C, Lin W, Liu S, Pan H, Tang X, Wang J, Wang W, Yu J, Zhang B, Zhang Q, Zhao H, Zhao H, Zhou J, Gabriel SB, Barry R, Blumenstiel B, Camargo A, Defelice M, Faggart M, Goyette M, Gupta S, Moore J, Nguyen H, Onofrio RC, Parkin M, Roy J, Stahl E, Winchester E, Ziaugra L, Altshuler D, Shen Y, Yao Z, Huang W, Chu X, He Y, Jin L, Liu Y, Shen Y, Sun W, Wang H, Wang Y, Wang Y, Xiong X, Xu L, Waye MM, Tsui SK, Xue H, Wong JT, Galver LM, Fan JB, Gunderson K, Murray SS, Oliphant AR, Chee MS, Montpetit A, Chagnon F, Ferretti V, Leboeuf M, Olivier JF, Phillips MS, Roumy S, Sallee C, Verner A, Hudson TJ, Kwok PY, Cai D, Koboldt DC, Miller RD, Pawlikowska L, Taillon-Miller P, Xiao M, Tsui LC, Mak W, Song YQ, Tam PK, Nakamura Y, Kawaguchi T, Kitamoto T, Morizono T, Nagashima A, Ohnishi Y, Sekine A, Tanaka T, Tsunoda T, Deloukas P, Bird CP, Delgado M, Dermitzakis ET, Gwilliam R, Hunt S, Morrison J, Powell D, Stranger BE, Whittaker P, Bentley DR, Daly MJ, de Bakker PI, Barrett J, Chretien YR, Maller J, McCarroll S, Patterson N, Pe'er I, Price A, Purcell S, Richter DJ, Sabeti P, Saxena R, Schaffner SF, Sham PC, Varilly P, Altshuler D, Stein LD, Krishnan L, Smith AV, Tello-Ruiz MK, Thorisson GA, Chakravarti A, Chen PE, Cutler DJ, Kashuk CS, Lin S, Abecasis GR, Guan W, Li Y, Munro HM, Qin ZS, Thomas DJ, McVean G, Auton A, Bottolo L, Cardin N, Eyheramendy S, Freeman C, Marchini J, Myers S, Spencer C, Stephens M, Donnelly P, Cardon LR, Clarke G, Evans DM, Morris AP, Weir BS, Tsunoda T, Mullikin JC, Sherry ST, Feolo M, Skol A, Zhang H, Zeng C, Zhao H, Matsuda I, Fukushima Y, Macer DR, Suda E, Rotimi CN, Adebamowo CA, Ajayi I, Aniagwu T, Marshall PA, Nkwodimmah C, Royal CD, Leppert MF, Dixon M, Peiffer A, Qiu R, Kent A, Kato K, Niikawa N, Adewole IF, Knoppers BM, Foster MW, Clayton EW, Watkin J, Gibbs RA, Belmont JW, Muzny D, Nazareth L, Sodergren E, Weinstock GM, Wheeler DA, Yakub I, Gabriel SB, Onofrio RC, Richter DJ, Ziaugra L, Birren BW, Daly MJ, Altshuler D, Wilson RK, Fulton LL, Rogers J, Burton J, Carter NP, Clee CM, Griffiths M, Jones MC, McLay K, Plumb RW, Ross MT, Sims SK, Willey DL, Chen Z, Han H, Kang L, Godbout M, Wallenburg JC, L'Archeveque P, Bellemare G, Saeki K, Wang H, An D, Fu H, Li Q, Wang Z, Wang R, Holden AL, Brooks LD, McEwen JE, Guyer MS, Wang VO, Peterson JL, Shi M, Spiegel J, Sung LM, Zacharia LF, Collins FS, Kennedy K, Jamieson R, Stewart J. A second generation human haplotype map of over 3.1 million SNPs. Nature 449:851-861. 2007.

Ioannidis JP, Ntzani EE, Trikalinos TA, Contopoulos-Ioannidis DG. Replication validity of genetic association studies. Nat.Genet. 29:306-309. 2001.

Ivarsson A, Hernell O, Nystrom L, Persson LA. Children born in the summer have increased risk for coeliac disease. J.Epidemiol.Community Health 57:36-39. 2003.

Ivarsson A, Persson LA, Nystrom L, Ascher H, Cavell B, Danielsson L, Dannaeus A, Lindberg T, Lindquist B, Stenhammar L, Hernell O. Epidemic of coeliac disease in Swedish children. Acta Paediatr. 89:165-171. 2000.

Janeway CA, Travers P, Walport M and Shlomchik MJ Immunobiology: the immune system in health and disease. 5th edition. Garland Publishing, New York, NY, USA. 2001. 732pp.

Jungerius BJ, Bakker SC, Monsuur AJ, Sinke RJ, Kahn RS, Wijmenga C. Is MYO9B the missing link between schizophrenia and celiac disease? Am.J.Med.Genet.B.Neuropsychiatr.Genet. 147:351-355. 2008.

Kaprio J, Tuomilehto J, Koskenvuo M, Romanov K, Reunanen A, Eriksson J, Stengard J, Kesaniemi YA. Concordance for type 1 (insulin-dependent) and type 2 (non-insulin-dependent) diabetes mellitus in a population-based cohort of twins in Finland. Diabetologia 35:1060-1067. 1992.

Karassa FB, Trikalinos TA, Ioannidis JP, FcgammaRIIa-SLE Meta-Analysis Investigators. Role of the Fcgamma receptor IIa polymorphism in susceptibility to systemic lupus erythematosus and lupus nephritis: a meta-analysis. Arthritis Rheum. 46:1563-1571. 2002.

Karell K, Klinger N, Holopainen P, Levo A, Partanen J. Major histocompatibility complex (MHC)-linked microsatellite markers in a founder population. Tissue Antigens 56:45-51. 2000.

Karell K, Louka AS, Moodie SJ, Ascher H, Clot F, Greco L, Ciclitira PJ, Sollid LM, Partanen J, European Genetics Cluster on Celiac Disease. 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. 64:469-477. 2003.

Kasaian MT, Whitters MJ, Carter LL, Lowe LD, Jussif JM, Deng B, Johnson KA, Witek JS, Senices M, Konz RF, Wurster AL, Donaldson DD, Collins M, Young DA, Grusby MJ. IL-21 limits NK cell responses and promotes antigen-specific T cell activation: a mediator of the transition from innate to adaptive immunity. Immunity 16:559-569. 2002.

Kaukinen K, Collin P, Laurila K, Kaartinen T, Partanen J, Mäki M. Resurrection of gliadin antibodies in coeliac disease. Deamidated gliadin peptide antibody test provides additional diagnostic benefit. Scand.J.Gastroenterol. 42:1428-1433. 2007.

Kaukinen K, Mäki M, Partanen J, Sievanen H, Collin P. Celiac disease without villous atrophy: revision of criteria called for. Dig.Dis.Sci. 46:879-887. 2001.

Kaukinen K, Partanen J, Mäki M, Collin P. HLA-DQ typing in the diagnosis of celiac disease. Am.J.Gastroenterol. 97:695-699. 2002.

Kaukinen K, Peraaho M, Collin P, Partanen J, Woolley N, Kaartinen T, Nuutinen T, Halttunen T, Mäki M, Korponay-Szabó I. Small-bowel mucosal transglutaminase 2-specific IgA deposits in coeliac disease without villous atrophy: a prospective and randomized clinical study. Scand.J.Gastroenterol. 40:564-572. 2005.

Kehrl JH. Heterotrimeric G protein signaling: roles in immune function and fine-tuning by RGS proteins. Immunity 8:1-10. 1998.

Keuning JJ, Pena AS, van Leeuwen A, van Hooff JP, va Rood JJ. HLA-DW3 associated with coeliac disease. Lancet 1:506-508. 1976.

King AL, Fraser JS, Moodie SJ, Curtis D, Dearlove AM, Ellis HJ, Rosen-Bronson S, Ciclitira PJ. Coeliac disease: follow-up linkage study provides further support for existence of a susceptibility locus on chromosome 11p11. Ann.Hum.Genet. 65:377-386. 2001.

King AL, Yiannakou JY, Brett PM, Curtis D, Morris MA, Dearlove AM, Rhodes M, Rosen-Bronson S, Mathew C, Ellis HJ, Ciclitira PJ. A genome-wide family-based linkage study of coeliac disease. Ann.Hum.Genet. 64:479-490. 2000.

Knuuttila J, Metzidis A, Sammalisto S, Peltonen L, Perola M, Saharinen J. CARTOGR: A tool to generate genetic marker maps for linkage analysis based on markers' physical location and genetic distances. Ver 2.2.0. 2007.

Kondrashova A, Mustalahti K, Kaukinen K, Viskari H, Volodicheva V, Haapala AM, Ilonen J, Knip M, Mäki M, Hyoty H, Epivir Study Group. Lower economic status and inferior hygienic environment may protect against celiac disease. Ann.Med. 40:223-231. 2008.

Kontakou M, Przemioslo RT, Sturgess RP, Limb AG, Ciclitira PJ. Expression of tumour necrosis factor-alpha, interleukin-6, and interleukin-2 mRNA in the jejunum of patients with coeliac disease. Scand.J.Gastroenterol. 30:456-463. 1995a.

Kontakou M, Przemioslo RT, Sturgess RP, Limb GA, Ellis HJ, Day P, Ciclitira PJ. Cytokine mRNA expression in the mucosa of treated coeliac patients after wheat peptide challenge. Gut 37:52-57. 1995b.

Korn T, Bettelli E, Gao W, Awasthi A, Jager A, Strom TB, Oukka M, Kuchroo VK. IL-21 initiates an alternative pathway to induce proinflammatory T(H)17 cells. Nature 448:484-487. 2007.

Korponay-Szabó IR, Halttunen T, Szalai Z, Laurila K, Kiraly R, Kovacs JB, Fesus L, Mäki M. In vivo targeting of intestinal and extraintestinal transglutaminase 2 by coeliac autoantibodies. Gut 53:641-648. 2004.

Korponay-Szabó IR, Szabados K, Pusztai J, Uhrin K, Ludmany E, Nemes E, Kaukinen K, Kapitany A, Koskinen L, Sipka S, Imre A, Mäki M. Population screening for coeliac disease in primary care by district nurses using a rapid antibody test: diagnostic accuracy and feasibility study. BMJ 335:1244-1247. 2007.

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am.J.Hum.Genet. 58:1347-1363. 1996.

Kruglyak L, Lander ES. Faster multipoint linkage analysis using Fourier transforms. J.Comput.Biol. 5:1-7. 1998.

Kurppa K, Collin P, Viljamaa M, Haimila K, Saavalainen P, Partanen J, Laurila K, Huhtala H, Paasikivi K, Mäki M, Kaukinen K. Diagnosing mild enteropathy celiac disease: a randomized, controlled clinical study. Gastroenterology 136:816-823. 2009.

Kurppa K, Koskinen O, Collin P, Mäki M, Reunala T, Kaukinen K. Changing phenotype of celiac disease after long-term gluten exposure. J.Pediatr.Gastroenterol.Nutr. 47:500-503. 2008.

Kyogoku C, Langefeld CD, Ortmann WA, Lee A, Selby S, Carlton VE, Chang M, Ramos P, Baechler EC, Batliwalla FM, Novitzke J, Williams AH, Gillett C, Rodine P, Graham RR, Ardlie KG, Gaffney PM, Moser KL, Petri M, Begovich AB, Gregersen PK, Behrens TW. Genetic association of the R620W polymorphism of protein tyrosine phosphatase PTPN22 with human SLE. Am.J.Hum.Genet. 75:504-507. 2004.

Laitinen T, Polvi A, Rydman P, Vendelin J, Pulkkinen V, Salmikangas P, Makela S, Rehn M, Pirskanen A, Rautanen A, Zucchelli M, Gullsten H, Leino M, Alenius H, Petays T, Haahtela T, Laitinen A, Laprise C, Hudson TJ, Laitinen LA, Kere J. Characterization of a common susceptibility locus for asthma-related traits. Science 304:300-304. 2004.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ, International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. Nature 409:860-921. 2001.

Lappalainen M, Halme L, Turunen U, Saavalainen P, Einarsdottir E, Farkkila M, Kontula K, Paavola-Sakki P. Association of IL23R, TNFRSF1A, and HLA-DRB1*0103 allele variants with inflammatory bowel disease phenotypes in the Finnish population. Inflamm.Bowel Dis. 14:1118-1124. 2008.

Latiano A, Mora B, Bonamico M, Megiorni F, Mazzilli MC, Cucchiara S, Palmieri O, Valvano MR, Annese V. Analysis of candidate genes on chromosomes 5q and 19p in celiac disease. J.Pediatr.Gastroenterol.Nutr. 45:180-186. 2007.

Lee EG, Boone DL, Chai S, Libby SL, Chien M, Lodolce JP, Ma A. Failure to regulate TNF-induced NF-kappaB and cell death responses in A20-deficient mice. Science 289:2350-2354. 2000.

Lewy H, Meirson H, Laron Z. Seasonality of birth month of children with celiac disease differs from that in the general population and between sexes and is linked to family history and environmental factors. J.Pediatr.Gastroenterol.Nutr. 48:181-185. 2009.

Li Y, He X, Schembri-King J, Jakes S, Hayashi J. Cloning and characterization of human Lnk, an adaptor protein with pleckstrin homology and Src homology 2 domains that can inhibit T cell activation. J.Immunol. 164:5199-5206. 2000.

Lindfors K, Kaukinen K, Mäki M. A role for anti-transglutaminase 2 autoantibodies in the pathogenesis of coeliac disease? Amino Acids 36:685-691. 2009.

Liu J, Juo SH, Holopainen P, Terwilliger J, Tong X, Grunn A, Brito M, Green P, Mustalahti K, Mäki M, Gilliam TC, Partanen J. Genomewide linkage analysis of celiac disease in Finnish families. Am.J.Hum.Genet. 70:51-59. 2002.

Lohi S, Mustalahti K, Kaukinen K, Laurila K, Collin P, Rissanen H, Lohi O, Bravi E, Gasparin M, Reunanen A, Mäki M. Increasing prevalence of coeliac disease over time. Aliment.Pharmacol.Ther. 26:1217-1225. 2007.

Lopez-Vazquez A, Rodrigo L, Fuentes D, Riestra S, Bousono C, Garcia-Fernandez S, Martinez-Borra J, Gonzalez S, Lopez-Larrea C. MHC class I chain related gene A (MICA) modulates the development of coeliac disease in patients with the high risk heterodimer DQA1*0501/DQB1*0201. Gut 50:336-340. 2002.

Louka AS, Nilsson S, Olsson M, Talseth B, Lie BA, Ek J, Gudjonsdottir AH, Ascher H, Sollid LM. HLA in coeliac disease families: a novel test of risk modification by the 'other' haplotype when at least one DQA1*05-DQB1*02 haplotype is carried. Tissue Antigens 60:147-154. 2002.

Lowenthal JW, Zubler RH, Nabholz M, MacDonald HR. Similarities between interleukin-2 receptor number and affinity on activated B and T lymphocytes. Nature 315:669-672. 1985.

Lundin KE, Nilsen EM, Scott HG, Loberg EM, Gjoen A, Bratlie J, Skar V, Mendez E, Lovik A, Kett K. Oats induced villous atrophy in coeliac disease. Gut 52:1649-1652. 2003.

Lundin KE, Scott H, Fausa O, Thorsby E, Sollid LM. 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. 41:285-291. 1994.

Lundin KE, Scott H, Hansen T, Paulsen G, Halstensen TS, Fausa O, Thorsby E, Sollid LM. Gliadin-specific, HLA-DQ(alpha 1*0501,beta 1*0201) restricted T cells isolated from the small intestinal mucosa of celiac disease patients. J.Exp.Med. 178:187-196. 1993.

Maiuri L, Ciacci C, Auricchio S, Brown V, Quaratino S, Londei M. Interleukin 15 mediates epithelial changes in celiac disease. Gastroenterology 119:996-1006. 2000.

Mao M, Biery MC, Kobayashi SV, Ward T, Schimmack G, Burchard J, Schelter JM, Dai H, He YD, Linsley PS. T lymphocyte activation gene identification by coregulated expression on DNA microarrays. Genomics 83:989-999. 2004.

Margaritte-Jeannin P, Babron MC, Bourgey M, Louka AS, Clot F, Percopo S, Coto I, Hugot JP, Ascher H, Sollid LM, Greco L, Clerget-Darpoux F. HLA-DQ relative risks for coeliac disease in European populations: a study of the European Genetics Cluster on Coeliac Disease. Tissue Antigens 63:562-567. 2004.

Marks J, Shuster S. Malabsorption and the skin. Br.Med.J. 4:455. 1968.

Marks J, Shuster S, Watson AJ. Small-bowel changes in dermatitis herpetiformis. Lancet 2:1280-1282. 1966.

Marzari R, Sblattero D, Florian F, Tongiorgi E, Not T, Tommasini A, Ventura A, Bradbury A. Molecular dissection of the tissue transglutaminase autoantibody response in celiac disease. J.Immunol. 166:4170-4176. 2001.

Matter K, Balda MS. Signalling to and from tight junctions. Nat.Rev.Mol.Cell Biol. 4:225-236. 2003.

Mazzilli MC, Ferrante P, Mariani P, Martone E, Petronzelli F, Triglione P, Bonamico M. A study of Italian pediatric celiac disease patients confirms that the primary HLA association is to the DQ(alpha 1*0501, beta 1*0201) heterodimer. Hum.Immunol. 33:133-139. 1992.

McDevitt HO. The role of MHC class II molecules in susceptibility and resistance to autoimmunity. Curr.Opin.Immunol. 10:677-681. 1998.

Mearin ML, Ivarsson A, Dickey W. Coeliac disease: is it time for mass screening? Best Pract.Res.Clin.Gastroenterol. 19:441-452. 2005.

Mention JJ, Ben Ahmed M, Begue B, Barbe U, Verkarre V, Asnafi V, Colombel JF, Cugnenc PH, Ruemmele FM, McIntyre E, Brousse N, Cellier C, Cerf-Bensussan N. Interleukin 15: a key to disrupted intraepithelial lymphocyte homeostasis and lymphomagenesis in celiac disease. Gastroenterology 125:730-745. 2003.

Meresse B, Chen Z, Ciszewski C, Tretiakova M, Bhagat G, Krausz TN, Raulet DH, Lanier LL, Groh V, Spies T, Ebert EC, Green PH, Jabri B. Coordinated induction by IL15 of a TCR-independent NKG2D signaling pathway converts CTL into lymphokine-activated killer cells in celiac disease. Immunity 21:357-366. 2004.

Molberg O, Mcadam SN, Korner R, Quarsten H, Kristiansen C, Madsen L, Fugger L, Scott H, Noren O, Roepstorff P, Lundin KE, Sjostrom H, Sollid LM. Tissue transglutaminase selectively modifies gliadin peptides that are recognized by gut-derived T cells in celiac disease. Nat.Med. 4:713-717. 1998.

Monsuur AJ, de Bakker PI, Alizadeh BZ, Zhernakova A, Bevova MR, Strengman E, Franke L, van't Slot R, van Belzen MJ, Lavrijsen IC, Diosdado B, Daly MJ, Mulder CJ, Mearin ML, Meijer JW, Meijer GA, van Oort E, Wapenaar MC, Koeleman BP, Wijmenga C. Myosin IXB variant increases the risk of celiac disease and points toward a primary intestinal barrier defect. Nat.Genet. 37:1341-1344. 2005.

Monsuur AJ, de Bakker PI, Zhernakova A, Pinto D, Verduijn W, Romanos J, Auricchio R, Lopez A, van Heel DA, Crusius JB, Wijmenga C. Effective detection of human leukocyte antigen risk alleles in celiac disease using tag single nucleotide polymorphisms. PLoS ONE 3:e2270. 2008.

Morahan G, Huang D, Ymer SI, Cancilla MR, Stephen K, Dabadghao P, Werther G, Tait BD, Harrison LC, Colman PG. Linkage disequilibrium of a type 1 diabetes susceptibility locus with a regulatory IL12B allele. Nat.Genet. 27:218-221. 2001.

Muller T, Anlag K, Wildner H, Britsch S, Treier M, Birchmeier C. The bHLH factor Olig3 coordinates the specification of dorsal neurons in the spinal cord. Genes Dev. 19:733-743. 2005.

Murumägi A, Vahamurto P, Peterson P. Characterization of regulatory elements and methylation pattern of the autoimmune regulator (AIRE) promoter. J.Biol.Chem. 278:19784-19790. 2003.

Musone SL, Taylor KE, Lu TT, Nititham J, Ferreira RC, Ortmann W, Shifrin N, Petri MA, Kamboh MI, Manzi S, Seldin MF, Gregersen PK, Behrens TW, Ma A, Kwok PY, Criswell LA. Multiple polymorphisms in the TNFAIP3 region are independently associated with systemic lupus erythematosus. Nat.Genet. 40:1062-1064. 2008.

Mustalahti K, Lohiniemi S, Collin P, Vuolteenaho N, Laippala P, Mäki M. Gluten-free diet and quality of life in patients with screen-detected celiac disease. Eff.Clin.Pract. 5:105-113. 2002a.

Mustalahti K, Sulkanen S, Holopainen P, Laurila K, Collin P, Partanen J, Mäki M. Coeliac disease among healthy members of multiple case coeliac disease families. Scand.J.Gastroenterol. 37:161-165. 2002b.

Myrsky E, Kaukinen K, Syrjanen M, Korponay-Szabó IR, Mäki M, Lindfors K. Coeliac disease-specific autoantibodies targeted against transglutaminase 2 disturb angiogenesis. Clin.Exp.Immunol. 152:111-119. 2008.

Mäki M, Collin P. Coeliac disease. Lancet 349:1755-1759. 1997.

Nair RP, Duffin KC, Helms C, Ding J, Stuart PE, Goldgar D, Gudjonsson JE, Li Y, Tejasvi T, Feng BJ, Ruether A, Schreiber S, Weichenthal M, Gladman D, Rahman P, Schrodi SJ, Prahalad S, Guthery SL, Fischer J, Liao W, Kwok PY, Menter A, Lathrop GM, Wise CA, Begovich AB, Voorhees JJ, Elder JT, Krueger GG, Bowcock AM, Abecasis GR, Collaborative Association Study of Psoriasis. Genome-wide scan reveals association of psoriasis with IL-23 and NF-kappaB pathways. Nat.Genet. 41:199-204. 2009.

Naluai AT, Nilsson S, Gudjonsdottir AH, Louka AS, Ascher H, Ek J, Hallberg B, Samuelsson L, Kristiansson B, Martinsson T, Nerman O, Sollid LM, Wahlstrom J. Genome-wide linkage analysis of Scandinavian affected sib-pairs supports presence of susceptibility loci for celiac disease on chromosomes 5 and 11. Eur.J.Hum.Genet. 9:938-944. 2001.

Naluai AT, Nilsson S, Samuelsson L, Gudjonsdottir AH, Ascher H, Ek J, Hallberg B, Kristiansson B, Martinsson T, Nerman O, Sollid LM, Wahlstrom J. The CTLA4/CD28 gene region on chromosome 2q33 confers susceptibility to celiac disease in a way possibly distinct from that of type 1 diabetes and other chronic inflammatory disorders. Tissue Antigens 56:350-355. 2000.

Neuhausen SL, Feolo M, Camp NJ, Farnham J, Book L, Zone JJ. Genome-wide linkage analysis for celiac disease in North American families. Am.J.Med.Genet. 111:1-9. 2002.

Neuhausen SL, Steele L, Ryan S, Mousavi M, Pinto M, Osann KE, Flodman P, Zone JJ. Co-occurrence of celiac disease and other autoimmune diseases in celiacs and their first-degree relatives. J.Autoimmun. 31:160-165. 2008.

Nilsen EM, Jahnsen FL, Lundin KE, Johansen FE, Fausa O, Sollid LM, Jahnsen J, Scott H, Brandtzaeg P. Gluten induces an intestinal cytokine response strongly dominated by interferon gamma in patients with celiac disease. Gastroenterology 115:551-563. 1998.

Nistico L, Fagnani C, Coto I, Percopo S, Cotichini R, Limongelli MG, Paparo F, D'Alfonso S, Giordano M, Sferlazzas C, Magazzu G, Momigliano-Richiardi P, Greco L, Stazi MA. Concordance, disease progression, and heritability of coeliac disease in Italian twins. Gut 55:803-808. 2006.

Niveloni S, Sugai E, Cabanne A, Vazquez H, Argonz J, Smecuol E, Moreno ML, Nachman F, Mazure R, Kogan Z, Gomez JC, Maurino E, Bai JC. Antibodies against synthetic deamidated gliadin peptides as predictors of celiac disease: prospective assessment in an adult population with a high pretest probability of disease. Clin.Chem. 53:2186-2192. 2007.

Núñez C, Márquez A, Varadé J, Martínez A, Polanco I, Maluenda C, Fernández-Arquero M, de la Concha EG, Urcelay E. No evidence of association of the *MYO9B* polymorphisms with celiac disease in the Spanish population Tissue antigens 17-Oct-2006 published online: 2006.

Nunez C, Dema B, Cenit MC, Polanco I, Maluenda C, Arroyo R, de las Heras V, Bartolome M, de la Concha EG, Urcelay E, Martinez A. IL23R: a susceptibility locus for celiac disease and multiple sclerosis? Genes Immun. 9:289-293. 2008.

Nunez C, Marquez A, Varade J, Martinez A, Polanco I, Maluenda C, Fernandez-Arquero M, de la Concha EG, Urcelay E. No evidence of association of the MYO9B polymorphisms with celiac disease in the Spanish population. Tissue Antigens 68:489-492. 2006.

Nunez C, Oliver J, Mendoza JL, Gomez-Garcia M, Pinero A, Taxonera C, Diaz-Rubio M, Lopez-Nevot MA, de la Concha EG, Nieto A, Urcelay E, Martinez A, Martin J. MYO9B polymorphisms in patients with inflammatory bowel disease. Gut 56:1321-1322. 2007a.

Nunez C, Rueda B, Martinez A, Lopez-Nevot MA, Fernandez-Arquero M, de la Concha EG, Martin J, Urcelay E. Involvement of macrophage migration inhibitory factor gene in celiac disease susceptibility. Genes Immun. 8:168-170. 2007b.

Oberhuber G, Granditsch G, Vogelsang H. The histopathology of coeliac disease: time for a standardized report scheme for pathologists. Eur.J.Gastroenterol.Hepatol. 11:1185-1194. 1999.

O'Connell CB, Mooseker MS. Native Myosin-IXb is a plus-, not a minus-end-directed motor. Nat.Cell Biol. 5:171-172. 2003.

O'Connell JR, Weeks DE. PedCheck: a program for identification of genotype incompatibilities in linkage analysis. Am.J.Hum.Genet. 63:259-266. 1998.

Papa A, Danese S, Urgesi R, Grillo A, Guglielmo S, Roberto I, Semeraro S, Scaldaferri F, Pola R, Flex A, Fedeli G, Gasbarrini G, Pola P, Gasbarrini A. Intercellular adhesion molecule 1 gene polymorphisms in inflammatory bowel disease. Eur Rev Med Pharmacol Sci. 8:187-91. 2004

Peltekova VD, Wintle RF, Rubin LA, Amos CI, Huang Q, Gu X, Newman B, Van Oene M, Cescon D, Greenberg G, Griffiths AM, St George-Hyslop PH, Siminovitch KA. Functional variants of OCTN cation transporter genes are associated with Crohn disease. Nat Genet. 36:471-475. 2004.

Peltonen L, Pekkarinen P, Aaltonen J. Messages from an isolate: lessons from the Finnish gene pool. Biol.Chem.Hoppe Seyler 376:697-704. 1995.

Peräaho M, Collin P, Kaukinen K, Kekkonen L, Miettinen S, Mäki M. Oats can diversify a gluten-free diet in celiac disease and dermatitis herpetiformis. J.Am.Diet.Assoc. 104:1148-1150. 2004.

Petit MM, Meulemans SM, Van de Ven WJ. The focal adhesion and nuclear targeting capacity of the LIM-containing lipoma-preferred partner (LPP) protein. J.Biol.Chem. 278:2157-2168. 2003.

Petronzelli F, Bonamico M, Ferrante P, Grillo R, Mora B, Mariani P, Apollonio I, Gemme G, Mazzilli MC. Genetic contribution of the HLA region to the familial clustering of coeliac disease. Ann.Hum.Genet. 61:307-317. 1997.

Piper JL, Gray GM, Khosla C. High selectivity of human tissue transglutaminase for immunoactive gliadin peptides: implications for celiac sprue. Biochemistry 41:386-393. 2002.

Plenge RM, Cotsapas C, Davies L, Price AL, de Bakker PI, Maller J, Pe'er I, Burtt NP, Blumenstiel B, DeFelice M, Parkin M, Barry R, Winslow W, Healy C, Graham RR, Neale BM, Izmailova E, Roubenoff R, Parker AN, Glass R, Karlson EW, Maher N, Hafler DA, Lee DM, Seldin MF, Remmers EF, Lee AT, Padyukov L, Alfredsson L, Coblyn J, Weinblatt ME, Gabriel SB, Purcell S, Klareskog L, Gregersen PK, Shadick NA, Daly MJ, Altshuler D. Two independent alleles at 6q23 associated with risk of rheumatoid arthritis. Nat.Genet. 39:1477-1482. 2007.

Ploski R, Ek J, Thorsby E, Sollid LM. On the HLA-DQ(alpha 1*0501, beta 1*0201)-associated susceptibility in celiac disease: a possible gene dosage effect of DQB1*0201. Tissue Antigens 41:173-177. 1993.

Plot L, Amital H. Infectious associations of Celiac disease. Autoimmun.Rev. 8:316-319. 2009.

Polvi A, Arranz E, Fernandez-Arquero M, Collin P, Mäki M, Sanz A, Calvo C, Maluenda C, Westman P, de la Concha EG, Partanen J. HLA-DQ2-negative celiac disease in Finland and Spain. Hum.Immunol. 59:169-175. 1998.

Polvi A, Eland C, Koskimies S, Mäki M, Partanen J. HLA DQ and DP in Finnish families with celiac disease. Eur.J.Immunogenet. 23:221-234. 1996.

Popat S, Bevan S, Braegger CP, Busch A, O'Donoghue D, Falth-Magnusson K, Godkin A, Hogberg L, Holmes G, Hosie KB, Howdle PD, Jenkins H, Jewell D, Johnston S, Kennedy NP, Kumar P, Logan RF, Love AH, Marsh MN, Mulder CJ, Sjoberg K, Stenhammar L, Walker-Smith J, Houlston RS. Genome screening of coeliac disease. J.Med.Genet. 39:328-331. 2002a.

Popat S, Hearle N, Hogberg L, Braegger CP, O'Donoghue D, Falth-Magnusson K, Holmes GK, Howdle PD, Jenkins H, Johnston S, Kennedy NP, Kumar PJ, Logan RF, Marsh MN, Mulder CJ, Torinsson Naluai A, Sjoberg K, Stenhammar L, Walters JR, Jewell DP, Houlston RS. Variation in the CTLA4/CD28 gene region confers an increased risk of coeliac disease. Ann.Hum.Genet. 66:125-137. 2002b.

Popat S, Hearle N, Wixey J, Hogberg L, Bevan S, Lim W, Stenhammar L, Houlston RS. Analysis of the CTLA4 gene in Swedish coeliac disease patients. Scand.J.Gastroenterol. 37:28-31. 2002c.

Przemioslo RT, Lundin KE, Sollid LM, Nelufer J, Ciclitira PJ. Histological changes in small bowel mucosa induced by gliadin sensitive T lymphocytes can be blocked by anti-interferon gamma antibody. Gut 36:874-879. 1995.

Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19:149-150. 2003.

Reunala TL. Dermatitis herpetiformis. Clin.Dermatol. 19:728-736. 2001.

Reveille JD. The genetic basis of autoantibody production. Autoimmun.Rev. 5:389-398. 2006.

Rioux JD, Daly MJ, Silverberg MS, Lindblad K, Steinhart H, Cohen Z, Delmonte T, Kocher K, Miller K, Guschwan S, Kulbokas EJ, O'Leary S, Winchester E, Dewar K, Green T, Stone V, Chow C, Cohen A, Langelier D, Lapointe G, Gaudet D, Faith J, Branco N, Bull SB, McLeod RS, Griffiths AM, Bitton A, Greenberg GR, Lander ES, Siminovitch KA, Hudson TJ. Genetic variation in the 5q31 cytokine gene cluster confers susceptibility to Crohn disease. Nat.Genet. 29:223-228. 2001.

Rioux JD, Karinen H, Kocher K, McMahon SG, Karkkainen P, Janatuinen E, Heikkinen M, Julkunen R, Pihlajamäki J, Naukkarinen A, Kosma VM, Daly MJ, Lander ES, Laakso M. Genomewide search

and association studies in a Finnish celiac disease population: Identification of a novel locus and replication of the HLA and CTLA4 loci. Am.J.Med.Genet.A. 130:345-350. 2004.

Risch N. Assessing the role of HLA-linked and unlinked determinants of disease. Am.J.Hum.Genet. 40:1-14. 1987.

Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 273:1516-1517. 1996.

Risch NJ. Searching for genetic determinants in the new millennium. Nature 405:847-856. 2000.

Robinson J, Waller MJ, Parham P, de Groot N, Bontrop R, Kennedy LJ, Stoehr P, Marsh SG. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. Nucleic Acids Res. 31:311-314. 2003.

Romagnani S. The increased prevalence of allergy and the hygiene hypothesis: missing immune deviation, reduced immune suppression, or both? Immunology 112:352-363. 2004.

Romanos J, Barisani D, Trynka G, Zhernakova A, Bardella MT, Wijmenga C. Six new coeliac disease loci replicated in an Italian population confirm association with coeliac disease. J.Med.Genet. 46:60-63. 2009.

Romanos J, Rybak A, Wijmenga C, Wapenaar MC. Molecular diagnosis of celiac disease: are we there yet? Expert Opin. Med. Diagn. 2:399-416. 2008.

Rudd CE. CTLA-4 co-receptor impacts on the function of Treg and CD8+ T-cell subsets. Eur.J.Immunol. 39:687-690. 2009.

Rueda B, Martinez A, Lopez-Nevot MA, Mas-Fontao A, Paco L, Ortega E, Fernandez-Arquero M, Urcelay E, Gomez de la Concha E, Martin J. A functional variant of IFNgamma gene is associated with coeliac disease. Genes Immun. 5:517-519. 2004.

Rueda B, Pascual M, Lopez-Nevot MA, Koeleman BP, Ortega E, Maldonado J, Lopez M, Martin J. Association of MICA-A5.1 allele with susceptibility to celiac disease in a family study. Am.J.Gastroenterol. 98:359-362. 2003.

Saalman R, Wold AE, Dahlgren UI, Fallstrom SP, Hanson LA, Ahlstedt S. Antibody-dependent cell-mediated cytotoxicity to gliadin-coated cells with sera from children with coeliac disease. Scand.J.Immunol. 47:37-42. 1998.

Salmi TT, Collin P, Jarvinen O, Haimila K, Partanen J, Laurila K, Korponay-Szabó IR, Huhtala H, Reunala T, Mäki M, Kaukinen K. Immunoglobulin A autoantibodies against transglutaminase 2 in the small intestinal mucosa predict forthcoming coeliac disease. Aliment.Pharmacol.Ther. 24:541-552. 2006.

Salvati VM, MacDonald TT, Bajaj-Elliott M, Borrelli M, Staiano A, Auricchio S, Troncone R, Monteleone G. Interleukin 18 and associated markers of T helper cell type 1 activity in coeliac disease. Gut 50:186-190. 2002.

Samman PD. Idiopathic steatorrhea. Proc.R.Soc.Med. 48:774-775. 1955.

Samuelsson L, Enlund F, Torinsson A, Yhr M, Inerot A, Enerback C, Wahlstrom J, Swanbeck G, Martinsson T. A genome-wide search for genes predisposing to familial psoriasis by using a stratification approach. Hum.Genet. 105:523-529. 1999.

Sanchez E, Alizadeh BZ, Valdigem G, Ortego-Centeno N, Jimenez-Alonso J, de Ramon E, Garcia A, Lopez-Nevot MA, Wijmenga C, Martin J, Koeleman BP. MYO9B gene polymorphisms are associated with autoimmune diseases in Spanish population. Hum.Immunol. 68:610-615. 2007.

Santiago JL, Martinez A, Nunez C, de la Calle H, Fernandez-Arquero M, de la Concha EG, Urcelay E. Association of MYO9B haplotype with type 1 diabetes. Hum.Immunol. 69:112-115. 2008.

Santin I, Castellanos-Rubio A, Aransay AM, Castano L, Vitoria JC, Bilbao JR. The functional R620W variant of the PTPN22 gene is associated with celiac disease. Tissue Antigens 71:247-249. 2008.

Santin I, Castellanos-Rubio A, Perez de Nanclares G, Vitoria JC, Castano L, Bilbao JR. Association of KIR2DL5B gene with celiac disease supports the susceptibility locus on 19q13.4. Genes Immun. 8:171-176. 2007.

Sareneva I, Koskinen LL, Korponay-Szabó IR, Kaukinen K, Kurppa K, Ziberna F, Vatta S, Not T, Ventura A, Ádány R, Pocsai Z, Széles G, Mäki M, Saavalainen P, Einarsdottir E. Linkage and association study of FcgammaR polymorphisms in celiac disease. Tissue Antigens 73:54-58. 2009.

Savilahti E, Reunala T, Mäki M. Increase of lymphocytes bearing the gamma/delta T cell receptor in the jejunum of patients with dermatitis herpetiformis. Gut 33:206-211. 1992.

Schumacher JM, Lee K, Edelhoff S, Braun RE. Distribution of Tenr, an RNA-binding protein, in a lattice-like network within the spermatid nucleus in the mouse. Biol.Reprod. 52:1274-1283. 1995.

Sewell DL, Reinke EK, Hogan LH, Sandor M, Fabry Z. Immunoregulation of CNS autoimmunity by helminth and mycobacterial infections. Immunol.Lett. 82:101-110. 2002.

Shan L, Molberg O, Parrot I, Hausch F, Filiz F, Gray GM, Sollid LM, Khosla C. Structural basis for gluten intolerance in celiac sprue. Science 297:2275-2279. 2002.

Shilling RA, Bandukwala HS, Sperling AI. Regulation of T:B cell interactions by the inducible costimulator molecule: does ICOS "induce" disease? Clin.Immunol. 121:13-18. 2006.

Shires J, Theodoridis E, Hayday AC. Biological insights into TCRgammadelta+ and TCRalphabeta+ intraepithelial lymphocytes provided by serial analysis of gene expression (SAGE). Immunity 15:419-434. 2001.

Smecuol E, Bai JC, Vazquez H, Kogan Z, Cabanne A, Niveloni S, Pedreira S, Boerr L, Maurino E, Meddings JB. Gastrointestinal permeability in celiac disease. Gastroenterology 112:1129-1136. 1997.

Smith KA. Interleukin-2: inception, impact, and implications. Science 240:1169-1176. 1988.

Smyth DJ, Plagnol V, Walker NM, Cooper JD, Downes K, Yang JH, Howson JM, Stevens H, McManus R, Wijmenga C, Heap GA, Dubois PC, Clayton DG, Hunt KA, van Heel DA, Todd JA. Shared and distinct genetic variants in type 1 diabetes and celiac disease. N.Engl.J.Med. 359:2767-2777. 2008.

Solheim BG, Ek J, Thune PO, Baklien K, Bratlie A, Rankin B, Thoresen AB, Thorsby E. HLA antigens in dermatitis herpetiformis and coeliac disease. Tissue Antigens 7:57-59. 1976.

Sollid LM. Molecular basis of celiac disease. Annu.Rev.Immunol. 18:53-81. 2000.

Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E. Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. J.Exp.Med. 169:345-350. 1989.

Sollid LM, Molberg O, McAdam S, Lundin KE. Autoantibodies in coeliac disease: tissue transglutaminase--guilt by association? Gut 41:851-852. 1997.

Sollid LM, Thorsby E. HLA susceptibility genes in celiac disease: genetic mapping and role in pathogenesis. Gastroenterology 105:910-922. 1993.

Spurkland A, Ingvarsson G, Falk ES, Knutsen I, Sollid LM, Thorsby E. Dermatitis herpetiformis and celiac disease are both primarily associated with the HLA-DQ (alpha 1*0501, beta 1*02) or the HLA-DQ (alpha 1*03, beta 1*0302) heterodimers. Tissue Antigens 49:29-34. 1997.

Spurkland A, Sollid LM, Polanco I, Vartdal F, Thorsby E. HLA-DR and -DQ genotypes of celiac disease patients serologically typed to be non-DR3 or non-DR5/7. Hum.Immunol. 35:188-192. 1992.

Stefansson H, Rujescu D, Cichon S, Pietilainen OP, Ingason A, Steinberg S, Fossdal R, Sigurdsson E, Sigmundsson T, Buizer-Voskamp JE, Hansen T, Jakobsen KD, Muglia P, Francks C, Matthews PM, Gylfason A, Halldorsson BV, Gudbjartsson D, Thorgeirsson TE, Sigurdsson A, Jonasdottir A, Jonasdottir A, Bjornsson A, Mattiasdottir S, Blondal T, Haraldsson M, Magnusdottir BB, Giegling I, Moller HJ, Hartmann A, Shianna KV, Ge D, Need AC, Crombie C, Fraser G, Walker N, Lonnqvist J, Suvisaari J, Tuulio-Henriksson A, Paunio T, Toulopoulou T, Bramon E, Di Forti M, Murray R, Ruggeri M, Vassos E, Tosato S, Walshe M, Li T, Vasilescu C, Muhleisen TW, Wang AG, Ullum H, Djurovic S, Melle I, Olesen J, Kiemeney LA, Franke B, GROUP, Sabatti C, Freimer NB, Gulcher JR, Thorsteinsdottir U, Kong A, Andreassen OA, Ophoff RA, Georgi A, Rietschel M, Werge T, Petursson H, Goldstein DB, Nothen MM, Peltonen L, Collier DA, St Clair D, Stefansson K. Large recurrent microdeletions associated with schizophrenia. Nature 455:232-236. 2008.

Stokes PL, Asquith P, Holmes GK, Mackintosh P, Cooke WT. Histocompatibility antigens associated with adult coeliac disease. Lancet 2:162-164. 1972.

Strachan T and Read AP Human molecular genetics. 3th edition. Garland Publishing, New York, NY, USA and London, UK. 2004. 674pp.

Szabolcs M, Sipka S, Csorba S. In vitro cross-linking of gluten into high-molecular-weight polymers with transglutaminase. Acta Paediatr.Hung. 28:215-227. 1987.

Széles G, Voko Z, Jenei T, Kardos L, Pocsai Z, Bajtay A, Papp E, Pasti G, Kosa Z, Molnar I, Lun K, Ádány R. A preliminary evaluation of a health monitoring programme in Hungary. Eur.J.Public Health 15:26-32. 2005.

Takaki S, Sauer K, Iritani BM, Chien S, Ebihara Y, Tsuji K, Takatsu K, Perlmutter RM. Control of B cell production by the adaptor protein lnk. Definition Of a conserved family of signal-modulating proteins. Immunity 13:599-609. 2000.

Thomson W, Barton A, Ke X, Eyre S, Hinks A, Bowes J, Donn R, Symmons D, Hider S, Bruce IN, Wellcome Trust Case Control Consortium, Wilson AG, Marinou I, Morgan A, Emery P, YEAR Consortium, Carter A, Steer S, Hocking L, Reid DM, Wordsworth P, Harrison P, Strachan D, Worthington J. Rheumatoid arthritis association at 6q23. Nat.Genet. 39:1431-1433. 2007.

Todd JA, Walker NM, Cooper JD, Smyth DJ, Downes K, Plagnol V, Bailey R, Nejentsev S, Field SF, Payne F, Lowe CE, Szeszko JS, Hafler JP, Zeitels L, Yang JH, Vella A, Nutland S, Stevens HE, Schuilenburg H, Coleman G, Maisuria M, Meadows W, Smink LJ, Healy B, Burren OS, Lam AA, Ovington NR, Allen J, Adlem E, Leung HT, Wallace C, Howson JM, Guja C, Ionescu-Tirgoviste C, Genetics of Type 1 Diabetes in Finland, Simmonds MJ, Heward JM, Gough SC, Wellcome Trust Case Control Consortium, Dunger DB, Wicker LS, Clayton DG. Robust associations of four new chromosome regions from genome-wide analyses of type 1 diabetes. Nat.Genet. 39:857-864. 2007.

Tokuhiro S, Yamada R, Chang X, Suzuki A, Kochi Y, Sawada T, Suzuki M, Nagasaki M, Ohtsuki M, Ono M, Furukawa H, Nagashima M, Yoshino S, Mabuchi A, Sekine A, Saito S, Takahashi A, Tsunoda T, Nakamura Y, Yamamoto K. An intronic SNP in a RUNX1 binding site of SLC22A4, encoding an organic cation transporter, is associated with rheumatoid arthritis. Nat.Genet. 35:341-348. 2003.

Tosi R, Vismara D, Tanigaki N, Ferrara GB, Cicimarra F, Buffolano W, Follo D, Auricchio S. Evidence that celiac disease is primarily associated with a DC locus allelic specificity. Clin.Immunol.Immunopathol. 28:395-404. 1983.

Trier JS. Celiac sprue. N.Engl.J.Med. 325:1709-1719. 1991.

Trynka G, Zhernakova A, Romanos J, Franke L, Hunt K, Turner G, Platteel M, Ryan AW, de Kovel C, Barisani D, Bardella MT, McManus R, Van Heel DA, Wijmenga C. Coeliac disease associated risk variants in TNFAIP3 and REL implicate altered NF-{kappa}B signalling. Gut 58:1078-1083. 2009.

Turer EE, Tavares RM, Mortier E, Hitotsumatsu O, Advincula R, Lee B, Shifrin N, Malynn BA, Ma A. Homeostatic MyD88-dependent signals cause lethal inflamMation in the absence of A20. J.Exp.Med. 205:451-464. 2008.

Vader W, Kooy Y, Van Veelen P, De Ru A, Harris D, Benckhuijsen W, Pena S, Mearin L, Drijfhout JW, Koning F. The gluten response in children with celiac disease is directed toward multiple gliadin and glutenin peptides. Gastroenterology 122:1729-1737. 2002.

Vader W, Stepniak D, Kooy Y, Mearin L, Thompson A, van Rood JJ, Spaenij L, Koning F. 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.U.S.A. 100:12390-12395. 2003.

Van Belzen MJ, Meijer JW, Sandkuijl LA, Bardoel AF, Mulder CJ, Pearson PL, Houwen RH, Wijmenga C. A major non-HLA locus in celiac disease maps to chromosome 19. Gastroenterology 125:1032-1041. 2003.

van Belzen MJ, Vrolijk MM, Meijer JW, Crusius JB, Pearson PL, Sandkuijl LA, Houwen RH, Wijmenga C. A genomewide screen in a four-generation Dutch family with celiac disease: evidence for linkage to chromosomes 6 and 9. Am.J.Gastroenterol. 99:466-471. 2004.

van Bodegraven AA, Curley CR, Hunt KA, Monsuur AJ, Linskens RK, Onnie CM, Crusius JB, Annese V, Latiano A, Silverberg MS, Bitton A, Fisher SA, Steinhart AH, Forbes A, Sanderson J, Prescott NJ, Strachan DP, Playford RJ, Mathew CG, Wijmenga C, Daly MJ, Rioux JD, van Heel DA. Genetic variation in myosin IXB is associated with ulcerative colitis. Gastroenterology 131:1768-1774. 2006.

van de Wal Y, Kooy Y, van Veelen P, Pena S, Mearin L, Papadopoulos G, Koning F. Selective deamidation by tissue transglutaminase strongly enhances gliadin-specific T cell reactivity. J.Immunol. 161:1585-1588. 1998.

van der Meer JB. Granular deposits of immunoglobulins in the skin of patients with dermatitis herpetiformis. An immunofluorescent study. Br.J.Dermatol. 81:493-503. 1969.

van Heel DA, Franke L, Hunt KA, Gwilliam R, Zhernakova A, Inouye M, Wapenaar MC, Barnardo MC, Bethel G, Holmes GK, Feighery C, Jewell D, Kelleher D, Kumar P, Travis S, Walters JR, Sanders DS, Howdle P, Swift J, Playford RJ, McLaren WM, Mearin ML, Mulder CJ, McManus R, McGinnis R, Cardon LR, Deloukas P, Wijmenga C. A genome-wide association study for celiac disease identifies risk variants in the region harboring IL2 and IL21. Nat.Genet. 39:827-829. 2007.

Velaga MR, Wilson V, Jennings CE, Owen CJ, Herington S, Donaldson PT, Ball SG, James RA, Quinton R, Perros P, Pearce SH. The codon 620 tryptophan allele of the lymphoid tyrosine phosphatase (LYP) gene is a major determinant of Graves' disease. J.Clin.Endocrinol.Metab. 89:5862-5865. 2004.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. The sequence of the human genome. Science 291:1304-1351. 2001.

Viljamaa M, Collin P, Huhtala H, Sievanen H, Mäki M, Kaukinen K. Is coeliac disease screening in risk groups justified? A fourteen-year follow-up with special focus on compliance and quality of life. Aliment.Pharmacol.Ther. 22:317-324. 2005.

Walker-Smith JA. Discordance for childhood coeliac disease in monozygotic twins. Gut 14:374-375. 1973.

Wapenaar MC, Monsuur AJ, Poell J, van 't Slot R, Meijer JW, Meijer GA, Mulder CJ, Mearin ML, Wijmenga C. The SPINK gene family and celiac disease susceptibility. Immunogenetics 59:349-357. 2007.

Wapenaar MC, Monsuur AJ, van Bodegraven AA, Weersma RK, Bevova MR, Linskens RK, Howdle P, Holmes G, Mulder CJ, Dijkstra G, van Heel DA, Wijmenga C. Associations with tight junction

genes PARD3 and MAGI2 in Dutch patients point to a common barrier defect for coeliac disease and ulcerative colitis. Gut 57:463-467. 2008.

Wapenaar MC, van Belzen MJ, Fransen JH, Sarasqueta AF, Houwen RH, Meijer JW, Mulder CJ, Wijmenga C. The interferon gamma gene in celiac disease: augmented expression correlates with tissue damage but no evidence for genetic susceptibility. J.Autoimmun. 23:183-190. 2004.

Weersma RK, Zhernakova A, Nolte IM, Lefebvre C, Rioux JD, Mulder F, van Dullemen HM, Kleibeuker JH, Wijmenga C, Dijkstra G. ATG16L1 and IL23R are associated with inflammatory bowel diseases but not with celiac disease in the Netherlands. Am.J.Gastroenterol. 103:621-627. 2008.

Wellcome Trust Case Control Consortium, Burton PR, Clayton DG, Cardon LR, Craddock N, Deloukas P, Duncanson A, Kwiatkowski DP, McCarthy MI, Ouwehand WH, Samani NJ, Todd JA, Donnelly Chair P, Analysis Committee:, Barrett JC, Burton PR, Davison D, Donnelly P, Easton D, Evans DM, Leung HT, Marchini JL, Morris AP, Spencer CC, Tobin MD, Cardon LR, Clayton DG, UK Blood Services & University of Cambridge Controls:, Attwood AP, Boorman JP, Cant B, Everson U, Hussey JM, Jolley JD, Knight AS, Koch K, Meech E, Nutland S, Prowse CV, Stevens HE, Taylor NC, Walters GR, Walker NM, Watkins NA, Winzer T, Todd JA, Ouwehand WH, 1958 Birth Cohort Controls:, Jones RW, McArdle WL, Ring SM, Strachan DP, Pembrey M, Bipolar Disorder (Aberdeen):, Breen G, Clair DS, (Birmingham):, Caesar S, Gordon-Smith K, Jones L, (Cardiff):, Fraser C, Green EK, Grozeva D, Hamshere ML, Holmans PA, Jones IR, Kirov G, Moskivina V, Nikolov I, O'Donovan MC, Owen MJ, Craddock N, (London):, Collier DA, Elkin A, Farmer A, Williamson R, McGuffin P, (Newcastle):, Young AH, Ferrier IN, Coronary Artery Disease (Leeds):, Ball SG, Balmforth AJ, Barrett JH, Bishop TD, Iles MM, Maqbool A, Yuldasheva N, Hall AS, (Leicester):, Braund PS, Burton PR, Dixon RJ, Mangino M, Stevens S, Tobin MD, Thompson JR, Samani NJ, Crohn's Disease (Cambridge):, Bredin F, Tremelling M, Parkes M, (Edinburgh):, Drummond H, Lees CW, Nimmo ER, Satsangi J, (London):, Fisher SA, Forbes A, Lewis CM, Onnie CM, Prescott NJ, Sanderson J, Matthew CG, (Newcastle):, Barbour J, Mohiuddin MK, Todhunter CE, Mansfield JC, (Oxford):, Ahmad T, Cummings FR, Jewell DP, Hypertension (Aberdeen):, Webster J, (Cambridge):, Brown MJ, Clayton DG, (Evry F, Lathrop MG, (Glasgow):, Connell J, Dominiczak A, (Leicester):, Samani NJ, (London):, Marcano CA, Burke B, Dobson R, Gungadoo J, Lee KL, Munroe PB, Newhouse SJ, Onipinla A, Wallace C, Xue M, Caulfield M, (Oxford):, Farrall M, Rheumatoid Arthritis:, Barton A, The Biologics in RA Genetics and Genomics Study Syndicate (BRAGGS) Steering Committee*, Bruce IN, Donovan H, Eyre S, Gilbert PD, Hilder SL, Hinks AM, John SL, Potter C, Silman AJ, Symmons DP, Thomson W, Worthington J, Type 1 Diabetes:, Clayton DG, Dunger DB, Nutland S, Stevens HE, Walker NM, Widmer B, Todd JA, Type 2 Diabetes (Exeter):, Frayling TM, Freathy RM, Lango H, Perry JR, Shields BM, Weedon MN, Hattersley AT, (London):, Hitman GA, (Newcastle):, Walker M, (Oxford):, Elliott KS, Groves CJ, Lindgren CM, Rayner NW, Timpson NJ, Zeggini E, McCarthy MI, Tuberculosis (Gambia):, Newport M, Sirugo G, (Oxford):, Lyons E, Vannberg F, Hill AV, Ankylosing Spondylitis:, Bradbury LA, Farrar C, Pointon JJ, Wordsworth P, Brown MA, Autoimmune Thyroid Disease:, Franklyn JA, Heward JM, Simmonds MJ, Gough SC, Breast Cancer:, Seal S, Breast Cancer Susceptibility Collaboration (UK)*, Stratton MR, Rahman N, Multiple Sclerosis:, Ban M, Goris A, Sawcer SJ, Compston A, Gambian Controls (Gambia):, Conway D, Jallow M, Newport M, Sirugo G, (Oxford):, Rockett KA, Kwiatkowski DP, DNA, Genotyping,Data QC and Informatics (Wellcome Trust Sanger Institute, Hinxton):, Bumpstead SJ, Chaney A, Downes K, Ghori MJ, Gwilliam R, Hunt SE, Inouye M, Keniry A, King E, McGinnis R, Potter S, Ravindrarajah R, Whittaker P, Widden C, Withers D, Deloukas P, (Cambridge):, Leung HT, Nutland S, Stevens HE, Walker NM, Todd JA, Statistics (Cambridge):, Easton D, Clayton DG, (Leicester):, Burton PR, Tobin MD, (Oxford):, Barrett JC, Evans DM, Morris AP, Cardon LR, (Oxford):, Cardin NJ, Davison D, Ferreira T, Pereira-Gale J, Hallgrimsdottir IB, Howie BN, Marchini JL, Spencer CC, Su Z, Teo YY, Vukcevic D, Donnelly P, Principal Investigators:, Bentley D, Brown MA, Cardon LR, Caulfield M, Clayton DG, Compston A, Craddock N, Deloukas P, Donnelly P, Farrall M, Gough SC, Hall AS, Hattersley AT, Hill AV, Kwiatkowski DP, Matthew CG, McCarthy MI, Ouwehand WH, Parkes M, Pembrey M, Rahman N, Samani NJ, Stratton MR, Todd JA, Worthington J, AITD Replication Group:, Mitchell SL, Newby PR, Brand OJ, Carr-Smith J, Pearce SH, Gough SC, IL23R replication:, McGinnis R, Keniry A, Deloukas P, The Australo-Anglo-American Spondylitis Consortium (TASC), Reveille JD, Zhou X, Bradbury LA, Sims AM, Dowling A, Taylor J, Doan T, Cardon LR, Davis JC, Pointon JJ, Savage L, Ward MM, Learch TL, Weisman MH, Wordsworth P, Brown MA. Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat.Genet. 39:1329-1337. 2007.

Wolters VM, Verbeek WH, Zhernakova A, Onland-Moret C, Schreurs MW, Monsuur AJ, Verduijn W, Wijmenga C, Mulder CJ. The MYO9B gene is a strong risk factor for developing refractory celiac disease. Clin.Gastroenterol.Hepatol. 5:1399-405, 1405.e1-2. 2007.

Woolley N, Holopainen P, Ollikainen V, Mustalahti K, Mäki M, Kere J, Partanen J. A new locus for coeliac disease mapped to chromosome 15 in a population isolate. Hum.Genet. 111:40-45. 2002.

Working Group of European Society of Paediatric Gastroenterology and Nutrition.Revised criteria for diagnosis of coeliac disease. Arch.Dis.Child. 65:909-911. 1990.

Yamada A, Salama AD, Sayegh MH. The role of novel T cell costimulatory pathways in autoimmunity and transplantation. J.Am.Soc.Nephrol. 13:559-575. 2002.

Zanoni G, Navone R, Lunardi C, Tridente G, Bason C, Sivori S, Beri R, Dolcino M, Valletta E, Corrocher R, Puccetti A. In celiac disease, a subset of autoantibodies against transglutaminase binds toll-like receptor 4 and induces activation of monocytes. PLoS Med. 3:e358. 2006.

Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, Hu T, de Bakker PI, Abecasis GR, Almgren P, Andersen G, Ardlie K, Bostrom KB, Bergman RN, Bonnycastle LL, Borch-Johnsen K, Burtt NP, Chen H, Chines PS, Daly MJ, Deodhar P, Ding CJ, Doney AS, Duren WL, Elliott KS, Erdos MR, Frayling TM, Freathy RM, Gianniny L, Grallert H, Grarup N, Groves CJ, Guiducci C, Hansen T, Herder C, Hitman GA, Hughes TE, Isomaa B, Jackson AU, Jorgensen T, Kong A, Kubalanza K, Kuruvilla FG, Kuusisto J, Langenberg C, Lango H, Lauritzen T, Li Y, Lindgren CM, Lyssenko V, Marvelle AF, Meisinger C, Midthjell K, Mohlke KL, Morken MA, Morris AD, Narisu N, Nilsson P, Owen KR, Palmer CN, Payne F, Perry JR, Pettersen E, Platou C, Prokopenko I, Qi L, Qin L, Rayner NW, Rees M, Roix JJ, Sandbaek A, Shields B, Sjogren M, Steinthorsdottir V, Stringham HM, Swift AJ, Thorleifsson G, Thorsteinsdottir U, Timpson NJ, Tuomi T, Tuomilehto J, Walker M, Watanabe RM, Weedon MN, Willer CJ, Wellcome Trust Case Control Consortium, Illig T, Hveem K, Hu FB, Laakso M, Stefansson K, Pedersen O, Wareham NJ, Barroso I, Hattersley AT, Collins FS, Groop L, McCarthy MI, Boehnke M, Altshuler D. Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. Nat.Genet. 40:638-645. 2008.

Zhernakova A, Alizadeh BZ, Bevova M, van Leeuwen MA, Coenen MJ, Franke B, Franke L, Posthumus MD, van Heel DA, van der Steege G, Radstake TR, Barrera P, Roep BO, Koeleman BP, Wijmenga C. 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. 81:1284-1288. 2007.

Zhernakova A, Festen EM, Franke L, Trynka G, van Diemen CC, Monsuur AJ, Bevova M, Nijmeijer RM, van 't Slot R, Heijmans R, Boezen HM, van Heel DA, van Bodegraven AA, Stokkers PC, Wijmenga C, Crusius JB, Weersma RK. Genetic analysis of innate immunity in Crohn's disease and ulcerative colitis identifies two susceptibility loci harboring CARD9 and IL18RAP. Am.J.Hum.Genet. 82:1202-1210. 2008.

Zhong F, McCombs CC, Olson JM, Elston RC, Stevens FM, McCarthy CF, Michalski JP. An autosomal screen for genes that predispose to celiac disease in the western counties of Ireland. Nat.Genet. 14:329-333. 1996.