International Journal of Celiac Disease, 2024, Vol. 12, No. 1, 33-40
Available online at http://pubs.sciepub.com/ijcd/12/1/7
Published by Science and Education Publishing
DOI:10.12691/ijcd-12-1-7

Relationship Between Genotype and Disease Phenotype for Gluten-Related Disorders

Michael F. Roberts[1,*], Nadine Grzeskowiak[2], Stephen E. Bricher[3]

[1]Department of Biology, Linfield University, McMinnville, Oregon 97128 USA
[2]Celiac Nurse Consulting, Salem, Oregon 97302 USA
[3]Department of Mathematics and Computer Science, Linfield University, McMinnville, Oregon 97128 USA
*Corresponding author: mrobert@linfield.edu

Received November 14, 2024; Revised December 16, 2024; Accepted December 22, 2024

Abstract We investigated gluten-related disorders (GRD) in intestinal biopsy-proven subjects to determine the relation between genotype (i.e., Human Leukocyte Antigen (HLA) DQ alleles) and phenotype (i.e., antibody levels and clearly defined symptoms). Subjects had known DQ genotypes, information on IgA anti-Endomysial antibodies (EMA) and/or anti-tissue Trans Glutaminase (tTG) antibodies (tTGA). Subjects also answered a survey characterizing gastrointestinal, skin, and neural symptoms. Though all were biopsy-proven, HLA genotypes included all combinations of DQ2.5, DQ2.2, DQ8, DQ7.5, DQ5, and DQ6. Subject antibody levels were proportional to the number of DQ2.5 haplotypes, and subjects with non-CD-associated haplotypes did not have positive antibody scores (despite positive biopsy). In addition, the total number of GRD-associated symptoms were approximately the same across all genotypes. The results suggest that gluten damage and symptoms are independent of genotype; i.e., not restricted to people with DQ2.5, DQ2.2, or DQ8 genotypes or to those with positive antibody scores. We propose a GRD classification in which clinicians consider GRD symptoms and signs even in people without CD-associated genotypes, as these likely make up the majority of those with gluten-related disorders.

Keywords: *gluten-related disorders, celiac disease, innate and acquired immune responses, IgA tTGA and EMA antibodies, gluten-related symptoms*

Cite This Article: Michael F. Roberts, Nadine Grzeskowiak, and Stephen E. Bricher, "Relationship Between Genotype and Disease Phenotype for Gluten-Related Disorders." *International Journal of Celiac Disease*, vol. 12, no. 1 (2024): 33-40. doi: 10.12691/ijcd-12-1-7.

1. Introduction

Gluten-related disorders (GRD) are inflammatory responses to proteins in wheat, barley, and rye. Though the proteins are not identical in the three grains, they form similar functional complexes that are referred to as *gluten*. The most relevant protein is the version found in wheat and is called *gliadin.* Humans ingest many dietary proteins, most of which (e.g., those from eggs or meats) can be completely digested by human intestinal enzymes. However, gliadin resists complete digestion because it contains high proportions of the amino acids glutamine and proline [1]. Fragments of gluten proteins thus remain in the intestinal lumen as large peptides and are not absorbed for nutritive purposes [2].

Gliadin causes intestinal epithelial cell tight junction remodeling and loss of barrier function [3]. The signs and symptoms of GRD are varied, but most (i.e., anemia, osteoporosis, and other nutrient malabsorption issues) occur because the tight junction modification allows gliadin entry into intestinal lamina propria, into blood, and across the brain-blood barrier [4]. In the intestinal wall, inflammation and infiltration of intraepithelial lymphocytes (IEL) [5,6] degrade epithelial cells enough to cause nutrient malabsorption [7]. In addition, other body regions are affected, as blistering skin conditions [8] and cerebellar damage [9] are common.

GRD can be classified into two separate disorders that Sapone et al [10] characterize as gluten sensitivity (GS) and celiac disease (CD). Celiac Disease appears to result from an acquired immune response, sometimes called "adaptive" [6] or "reactive" [11]. It thus is characterized by genotype-specific antibody production. GS is sometimes described as non-celiac gluten sensitivity (NCGS) or non-celiac wheat sensitivity (NCWS) [12]. We will use the term NCGS for clarity. Evidence suggests that NCGS does not invoke an acquired immune response [10]. However, it does initiate inflammation and epithelial damage [6,13]. In the Conclusions (Section 4), we will develop our connection between these two GRD.

The overall result of the partial digestion of dietary gliadin is a mix of peptides that can trigger three separate host responses that resemble responses to pathogenic invaders [14]. Alterations of the gastrointestinal tract, skin, and neural tissue have become well known as indicating possible presence of a GRD. At present the most widely used diagnostic approaches are biopsy for gastrointestinal changes and skin lesions, as well as antibody titers [15]

and genotype [1]. Other indicators are serological markers such as cytokines [5], and, in the case of people who implement a gluten free diet, resolution of nearly all GRD symptoms and signs [16].

Three levels of gliadin action can be distinguished, based on which peptide is involved:

1. A gliadin peptide of 20 amino acids has direct actions on intestinal epithelial cells; i.e., disruption or reorganization of tight junctions [3] that reduces intestinal barrier function and increases permeability of the epithelial layer.
2. Peptides with different sequences activate the innate immune system, leading to inflammatory responses that damage epithelial cells [6,17]; whether these peptides are gliadin fragments or other wheat proteins, is being investigated [6,12,18]
3. Another gliadin peptide of 33 amino acids activates acquired immune system responses [6] and production of antibodies.

These separate actions of gluten-containing grains have historically been combined into a single term "celiac disease," but the three different pathogenic mechanisms outlined above suggest the situation is more complex. These separate actions suggest there are three physiological responses to gliadin, which we now discuss.

1.1. Level 1: Tight Junction Remodeling

The direct tight junction effect of gliadin results in loss of epithelial integrity, allowing movement of large molecules (e.g., gliadin peptides) out of the lumen and into the intestinal extracellular space. The increased permeability is an early event in gluten exposure [3] and occurs in all people ingesting gluten [3,19]. It is therefore genotype-independent.

1.2. Level 2: The Innate Immune Response

The next step following level 1 infiltration of gliadin peptides into the gastrointestinal intercellular tissue is the activation of innate immune system responses, including production of cytokines and access of cytotoxic IEL into the epithelial intestinal lining [2]. The buildup of cytotoxic IEL in the epithelial layer defines the Marsh 1 category of intestinal damage [20]. The relation between this damage and "villous atrophy" remains uncertain but it is proposed to be responsible for loss of epithelial cells [6,21]. Though some authors [18] state that in NCGS, gluten does not cause an inflammatory response, other authors [2,17] report that similar epithelial destruction can occur via innate immune mechanisms that *are not* genotype-specific [12,22]. This level 2 innate immune inflammatory response to gliadin could therefore be the process that is responsible for initial loss of epithelial cells. The inflammation also allows the next step in the gluten-response process, which *is* genotype-specific.

1.3. Level 3: The Acquired Immune Response

As a result of level 2 inflammation, a cellular enzyme, transglutaminase (TG), is released into extracellular space [23], modifying certain of the infiltrated gliadin sequences into different immunogenic peptides that promote an acquired immune antibody response that differs from the level 2 innate response [10,24]. The acquired immune response depends on the recognition of these peptides by proteins produced by genes of the HLA DQ system.

Human chromosome 6 contains the Human Leukocyte Antigen (HLA) genes involved in antigen surveillance. The DQ region of the chromosome has just two genes, A1 and B1. Each of these genes has more than ten alleles and the two can thus produce more than 100 A1-B1 allelic combinations in humans [25]. The two genes produce proteins that form a dimer (α–β protein pair) on the surface of immune surveillance cells (dendritic cells). This dimer presents a gliadin epitope to T helper cells (TH) for antibody production within the Level 3 acquired immune response [26]. The HLA haplotypes are thus the CD genotype. Because the two genes are adjacent on chromosome 6, the particular alleles on each parental chromosome are nearly always transmitted together to offspring.

The most relevant HLA haplotypes that produce gliadin-binding α–β proteins are DQ2.5, DQ2.2, and DQ8 [1,27]. Non CD-associated haplotypes are DQ5, DQ6, and in certain cases, DQ7.5. These we term "DQX.x."

1.4. Innate Mechanisms Causing GRD

Three lines of evidence suggest a primarily innate immune mechanism for initiation of villous atrophy: first, classical immunology [28] states that, in general, innate immune mechanisms are activated well prior to acquired immune mechanisms. Second, specific to the GRD, reports on cultured tissues [7] as well as studies on human subjects (next section) suggest that innate mechanisms cause villous atrophy in the absence of acquired immune signals. Third, the intestinal atrophic changes are reported to occur earlier than does the production of gliadin-specific antibodies in CD [2,14] and that innate immune mechanisms may trigger or enhance acquired immune responses [29]. Even so, some studies (e.g., [5,30]) suggest that acquired immune activation is needed for innate mechanisms to cause damage. We will return to this question in Section 4.3.

1.5. Diagnosis and Estimated Prevalence of GRD

Researchers [10,31] calculate that 10% of the U.S. population has one or another of the gluten disorders. Of these, CD itself is estimated to be about 3% of the U.S. population [31,32,33]. Even so, mis-diagnosis and under-diagnosis are common phenomena in clinical medicine and may result in serious delays in treatment [34,35]; errors in diagnosis are known to be significant in the GRD. Some reasons for under-estimation of GRD prevalence [1,26,36] may be: low specificity and sensitivity of antibody or biopsies; clinicians' greater interest in intestinal symptoms compared to skin or neural symptoms; and the tendency of people with a GRD to dismiss their symptoms [14,35]. In addition, several research groups [22,37,38,39,40,41] have reported cases that fit current criteria for CD (villous atrophy at Marsh 3 levels) but did not have HLA DQ2.5, DQ2.2, or DQ8. Such patients may

have NCGS and thus may be missed if clinicians focus only on treating those with the CD-associated genotypes.

Many people with CD-associated genotypes are told that because over 30% of all people have these genotypes but under 3% of the population has CD, there is less than a 10% chance that a person with the CD-associated genotype will actually get diagnosed with CD [42]. This is an under-estimate because researchers also acknowledge that current diagnostic practices are inadequate [1,43] and CD under-estimation may reach 90% [1,36]. If this estimate of under-diagnosis is correct, then a significant majority of those with CD-associated haplotypes already have CD at an undiagnosed degree of severity.

We therefore wished to determine the allele relationships of CD-associated haplotypes and whether it is possible to distinguish among the GRD via genotypes, symptom sets, antibody titers, or the three levels characterizing physiological responses to gluten. We studied people with biopsy-proven intestinal damage, most of whom also had known DQ haplotypes, clear symptoms involving intestinal, skin, and nervous system, and data for IgA antibodies (tTGA, EMA).

2. Methods

2.1. Study Population

We performed an analytical observational study [44] within a biopsy-proven cohort of people (age at positive biopsy between 2 and 81). These individuals had sought medical advice for abdominal, skin, and/or neural symptoms and had been biopsy-proven by endoscopy. The samples had been interpreted by a pathologist, who determined them to be at least Marsh 3 [20]. These people were invited to participate in the study. The final subject pool consisted of 111 people (23% male, 77% female, 92% Caucasian). All subjects filled out a symptom survey, 72 had HLA DQ genotypes determined, and 49 had data for both genotypes and IgA antibodies (tTGA/EMA). Subjects gave permission for use of their data and confidentiality was maintained throughout the study. Our procedure assured that individuals cannot be identified from their reported information. All clinical data had been obtained while subjects were consuming gluten.

2.2. Data Analysis

HLA haplotypes. The study population haplotypes were determined by private laboratories; some laboratories reported alleles for all four A1 and B1 genes (i.e., A1*05 B1*02, A1*02 B1*02); others reported only those alleles considered by the laboratory to be CD-associated and stated the remainder as "X" or "not CD-associated." Study population haplotypes were quite varied and represented approximately 17 of the 36 DQ A1-B1 haplotypes identified among European Americans [25]. Only four of our subjects had non-CD-associated haplotypes. These four (all biopsy-proven) had various combinations of DQ5.1 and DQ6.1, which are generally considered not to be CD-associated. We will address these haplotypes in Section 3.

We quantified our subjects' HLA DQ haplotypes by characterizing the ability of their α–β protein pairs to display gliadin epitope and thus activate T cells and B cells to make tTGA and EMA; a similar approach is seen in [26]. Each dendritic cell is able to specify up to four different α–β protein pairs (2 genes times 2 alleles). Our approach is shown in Table 1.

Table 1. Scoring system for HLA DQ haplotypes

Haplotype	DQ5, 6, 7	DQ8	DQ2.2	DQ7.5[#]	DQ2.5
Score	0	0.5	0.5	1 or 0	1

[#] If DQ7.5 is present with DQ2.5 or DQ2.2, score is 1; otherwise, score is zero

The haplotypes DQ2.5, DQ2.2, and DQ8 activate T cells and we thus gave them scores from 0.5 to 1 per dimer. In contrast, certain haplotypes (e.g., DQ5 and DQ6) synthesize α–β pairs that appear not to bind epitope nor present it to T cells. To these we give a "zero" binding score. Our scoring system in Table 1 leads to our method of calculating haplotype binding scores (Figure 1).

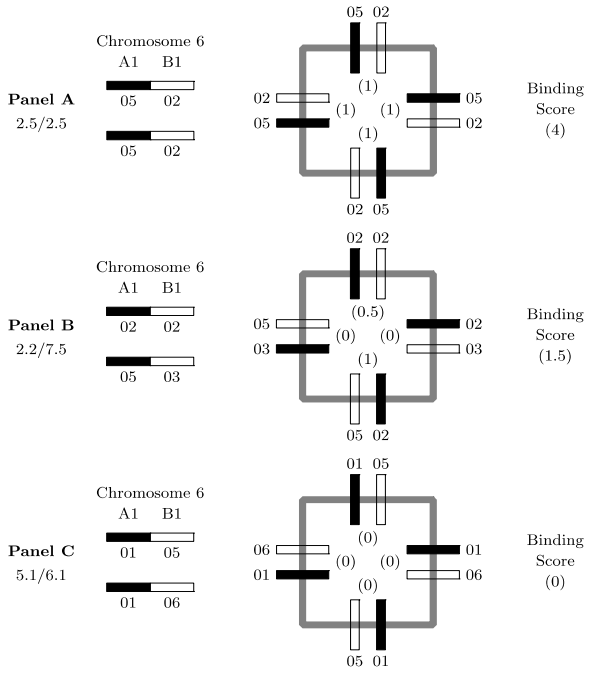

Figure 1. Haplotype-phenotype relation for T Cell activation. Three peoples' HLA diploid haplotypes are shown. At the left of each is the DQ haplotype of maternal and paternal chromosome 6, showing alleles for A1 and B1 genes (as shown, these genes are adjacent to each other on the chromosome). A dendritic cell (square outline), expresses the α–β protein pairs possible for each person (black and open rectangles match black and open portions of chromosome 6 at left). At the right of each is the predicted number of T cell activation units likely for each person, based on the idea that A1*05 B1*02 presents epitope strongly and activates T cells. Panel A: A1*05 B1*02 homozygote: all four α5 β2 heterodimers bind and present epitope and have score of 1. Panel B: A1*02 B1*02 presents with half the strength of A1*05 B1*02; the A1*05 B1*03 does not present epitope but allows the trans formation of α5 – β2 protein pair. Panel C: A1*01*B1*05 or A1*01 B1*06 do not present epitope at all and have score of 0. Thus, the DQ2.5/DQ2.5 homozygote (A) is predicted to be the strongest CD genotype, the DQ5/DQ6 genotype (C) is not predicted to be a CD genotype, and the DQ2.2/DQ7.5 (B) is predicted to be intermediate

For example, each DQ2.5 homozygote can make four identical α∗05 β∗02 proteins. As all the four possible α–β pairs are equally good at presenting gliadin, such homozygotes receive a haplotype binding score of 4 (Figure 1A). Because DQ7.5 has A1*05 and can double epitope binding when combined with DQ2.2 or DQ2.5, we gave DQ7.5 a binding score of 1 for these heterozygotes [45] so a DQ2.2/DQ7.5 heterozygote receives a score of 1.5 (Figure 1B). A DQ5 homozygote or a DQ5/DQ6 heterozygote receives a score of 0 (Figure 1C). Not included in Figure 1 are four additional haplotype binding scores: a DQ2.5/DQ2.2 heterozygote receives a score of 3; DQ8 homozygotes or DQ2.2 homozygotes receive a score of 2; DQ2.5/DQ5 heterozygotes receive a score of 1; and DQ2.2/DQ5 heterozygotes receive a score of 0.5. These seven haplotype binding scores quantify the genotype and are the basis for the genotype-phenotype analysis in Sections 3.3 and 3.4.

Mubarak et al. [46] introduced this approach by describing genotype action in terms of the HLA haplotypes synthesizing protein heterodimers that presented epitope to T cells for eventual antibody production. This allows quantification of the contributions of DQ2.5, 2.2, and 8 genotypes to antibody production. They proposed that antibody titers were proportional to the numbers of CD-associated heterodimers and also that CD is associated with antibody titers greater than 10 times reference values. However, though they illustrate the symptoms of subjects with different antibody concentration classes, they do not show the antibody values for individuals or groups with the same haplotype.

Antibodies. Though clinicians recognize several TG enzymes, current CD clinical antibody testing procedures are limited to the enzyme found in intestinal tissue ("tissue" transglutaminase -- tTG). The transglutaminase found in skin ("epithelial" or eTG) and the transglutaminase found in neural tissue ("neural" or nTG) are generally not tested in routine fashion. Our subjects thus had data only for tTGA/EMA. Antibodies are generally reported [47] in a qualitative fashion as binary, i.e., simply "negative" and "positive," or "low" and "high." Our attempt to develop a genotype-phenotype relationship required variable quantitative measures of antibody values that could be correlated with HLA haplotype, in agreement with recent recommendations [15] that give guidelines for use of different sources of antibody measurement.

Each subject's antibody information was provided by the laboratory making the analysis. Because each laboratory used different reference values, we quantified each subject's titer by dividing their experimental value by that lab's reference value (both measurements in mg/dl). This gave a non-dimensional antibody value that represented how much each subject's antibodies exceeded that lab's reference, sometimes called *Upper Limit of Normal [15]*. Thus, if a subject had tTGA of 100 mg/dl and the lab stated its reference value as 4 mg/dl, that subject's tTGA value was 25 "times reference" (Figure 3). For EMA, a person's serum titer is given as a dilution (e.g., 1:100) and the lab reference is given as another dilution, (e.g., 1:5). This person thus has an EMA of 20 times reference.

Of our 70 subjects with antibody data, 40 had both tTGA and EMA and 30 had only tTGA. We therefore needed a method that allowed us to use all subjects for which at least one of the antibody types was measured. Because EMA target the tTGA antibodies present [23] the two measure equivalent antibody responses and we were able to average the two to produce a single measure for this type of antibody, as validated with the Mann-Whitney U test [48], which requires the independence of the two samples. Indeed, we observed that whether a lab measured tTGA or EMA or both was random across our subjects and therefore concluded that the two measurements were independent. Thus, the subjects taking part in the genotype-phenotype analysis each had a diploid HLA DQ haplotype, EMA/tTGA titer, and symptom set.

A significant number of people with CD have a general IgA antibody deficiency [11]. Thus, we ensured that in our tabulations of low antibody titer subjects (in Figure 3, those with haplotype binding score of 0), all were known not to be IgA deficient.

Symptom categories. We developed a survey in which the subjects indicated symptom presence. We placed their data into categories for each of the three body regions affected by GRD: the gastrointestinal tract, skin, and nervous system. Our system for classifying symptoms is shown in Table 2.

Table 2. GRD Symptoms classified by body region

Gastrointestinal	Skin	Neural
Mouth: Aphthous ulcers Dental enamel defects --- Autoimmune: Thyroiditis Crohn's --- Abdominal: Pain Gas / bloating Diarrhea Vomiting --- Nutritional deficiencies Malabsorption Short stature Anemia Osteoporosis	Acne Itchy skin Hives Autoimmune: Vitiligo Alopecia areata RA Sjögren's Eczema Psoriasis Dermatitis Herpetiformis	"Brain Fog" Irritability Anxiety Depression Autoimmune: MS ALS Parkinson's Myasthenia gravis Neuropsychiatric Mood disorders Learning disabilities Schizophrenia Bipolar Dementia Gluten Ataxia Neuropathy

Though we organized symptoms in increasing order of severity, we were not able to correlate such an organization with genotype. However, we suggest that in a more general population, these groupings could be useful as a first step towards describing a relationship between symptom severity and HLA haplotype.

3. Results

3.1. Age at Diagnosis

The data for age at diagnosis (i.e., age at positive biopsy) are shown in Figure 2: panel A gives the proportions in each 10-year age interval and panel B gives the cumulative proportions over all ages.

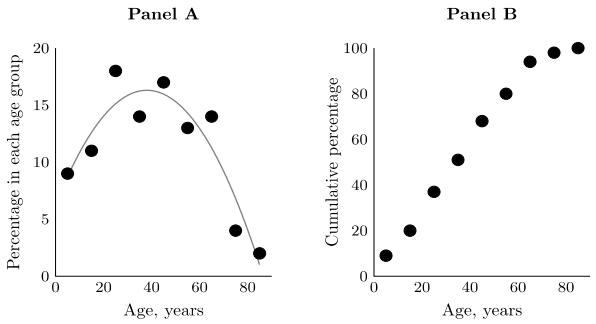

Figure 2. Distribution of the age at biopsy diagnosis for all subjects. Panel A: Percentage of subjects diagnosed in each 10-year age period. The curve is a best-fit quadratic function suggesting peak age of diagnosis is about 40 years of age. Panel B: cumulative proportions over all ages, suggesting that disease prevalence cannot be accurately estimated until people of almost all ages have been included

As other authors [49] have shown, most cases are diagnosed in middle age and GRD may appear at any time during a person's life. This suggests that GRD prevalence cannot be accurately estimated unless people of all ages are included; this can be referred to as the cumulative lifetime risk [35]. Because only a third of clinicians have ever diagnosed GRD or know that gluten-related symptoms may first present in adulthood [16], most people with GRD are not identified.

3.2. Distribution of Haplotypes

We compared the haplotype frequencies in our subjects with the data in Klitz et al. [25], who computed DQ haplotypes for 1899 people in the U.S. without considering whether they were associated with diseases. Table 3 shows the comparison between Klitz' study and ours.

Table 3. DQ haplotypes present in the U.S. population and the current study

			Percent	
A1	B1	DQ	In U.S. pop.	In this study
05	02	2.5	13	35
02	02	2.2	11	17
03	0302	8	7	14
x	X	X.x	69	34

At left are the HLA DQ A1 and B1 haplotypes. In the center is the DQ serotype. The first three rows indicate haplotypes that are CD-associated. In the bottom row, the x and X terms indicate haplotypes not-CD associated. The first three rows of data indicate haplotypes associated with CD; the bottom row indicates the non-CD-associated haplotypes. The "Freq. In US pop." column indicates the percentages of DQ haplotypes [25] that are in each of the above categories: 31% of Caucasian American haplotypes are those associated with CD (DQ2.5, DQ2.2, and DQ8) while 69% of haplotypes are not (e.g., DQ5, DQ6, DQ7). The final column represents the haplotype frequencies within our biopsy-proven subjects. The proportion of our subjects' haplotypes that are CD-associated is 66%, substantially higher than within the population as a whole. This difference is to be expected, as our subjects were not randomly chosen (they sought medical advice because they and/or their medical professionals were suspicious of their GRD symptoms).

The majority of our subjects had DQ2.5, 2.2, or DQ8 haplotypes. Some of our subjects were homozygous DQ2.5 or DQ8 but in most, the other haplotype was one not associated with CD (e.g., DQ5, DQ6, or DQ7). Indeed, the most frequent combination, DQ2.5/DQX.x heterozygotes, was seen in 25 subjects.

3.3. The Haplotype-antibody Relation

The relation between haplotype binding score and antibody values is presented in Figure 3, where haplotype binding scores for the seven categories of subjects are indicated on the horizontal axis and IgA tTGA/EMA antibody values (rescaled as "times reference" values – Section 2.2) for each haplotype binding score are given on the vertical axis (median ± SE, where SE denotes the standard error).

Use of the Mann-Whitney test showed that, for subjects in which both antibody measurements were made, the "times reference" values computed separately for tTGA and EMA were not significantly different ($p>0.05$). This is consistent with a report [50] that Caucasians have high degrees of tTGA and EMA concordance. Thus, we calculated a single mean value for those subjects who had measures of both. In addition, the independence of the two measurements meant that for each subject with only one antibody measurement, we could use that value as the subject's antibody titer.

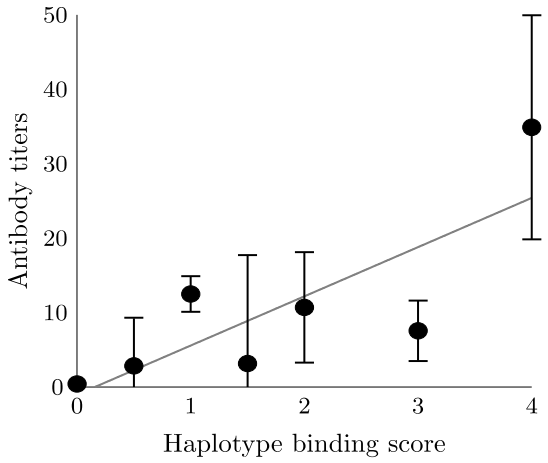

Figure 3. Relation between antibody titers (times reference) on the vertical axis and DQ haplotype binding scores (Figure 1) on the horizontal axis. The circles represent median ± SE of tTGA/EMA for each haplotype binding score. Thus, subjects with haplotype binding score of 4 are DQ2.5/2.5 homozygotes while subjects with score of 0 are DQX.x homozygotes lacking any CD-associated haplotypes. The slope of the linear regression is significantly different from zero ($p<0.05$) and suggests that antibody production is proportional to allele dose of the HLA DQ A1 and B1 genes

Figure 3 shows the relation between haplotype binding scores and IgA tTGA/EMA titers. The various values of haplotype binding score were obtained by following the procedure described in Section 2.2. These haplotype binding scores are based on our hypothesis that DQ2.2 and 8 have half the antibody-producing activity of DQ2.5 (Table 1; Figure 1). Though we arbitrarily set DQ 2.2 and DQ8 haplotype binding scores at 0.5, performing simulations with values other than 0.5 did not change the

proportionality between antibody titer and our measure of DQ haplotype binding score. It is possible, however, that better estimates of haplotype binding properties can be obtained in studies of data from more subjects with DQ2.2 and DQ8 haplotypes. We suggest that it is important to develop a widely applicable relation between CD genotype and phenotype, and that our approach is a useful first step in that direction.

3.4. Symptom-haplotype Relation

We determined the relation between symptoms and haplotype across all the symptom categories during the period subjects were eating gluten. Subsequently, nearly all our subjects went on a gluten free diet (GFD) once their clinical data were taken; all of those who did noted relief from symptoms. For the symptoms taken as a whole (for all the three body areas considered together), there was no tendency for the number of symptoms present to be affected by haplotype (Figure 4). Indeed, even those subjects with HLA DQ X.x haplotypes (zero haplotype binding score) had the same number of reported symptoms. This is likely because the gluten-caused physiological and innate immune changes were independent of their genotype. We are not aware of other studies classifying GRD symptoms as we have done (Table 2) and recommend this approach to other researchers.

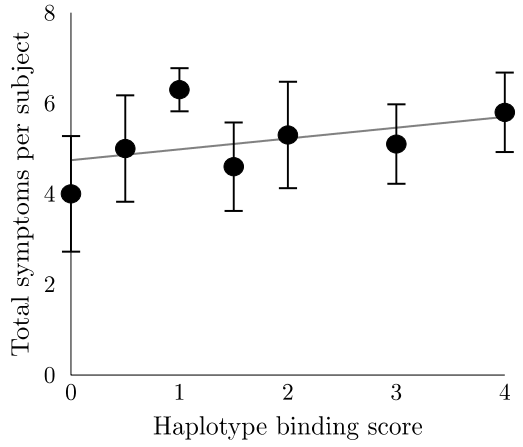

Figure 4. Relation of symptoms per subject (intestinal+skin+neural) to their haplotype binding score. The circles represent means ± SE for the number of reported symptoms. The slope of the linear regression is not significantly different from zero, indicating that symptom frequency is not associated with haplotype binding score. Note in particular that those with zero haplotype score also have the same number of symptoms as those with high binding scores

4. Conclusions

Our study shows a quantitative relationship between genotype (HLA DQ haplotype) and phenotype (quantitative measures of antibody production) for biopsy-proven subjects with CD (Figure 3). Our antibody data are thus compatible with a model of CD determined by three haplotypes, DQ2.5, DQ2.2, and DQ8. Other haplotypes (e.g., DQ5, DQ6, and DQ7.5) do not on their own contribute to significant antibody levels; this is seen in our subjects with DQ5/DQ6 haplotypes: their antibody titers are not above population reference values.

However, all our subjects, including those who did not have DQ2.5, DQ2.2, or DQ8 haplotypes, had biopsy-proven Marsh 3 intestinal damage and improved on a GFD. Thus, the data emphasize a more important idea: that gluten-caused intestinal damage may occur in people without the genotypes generally associated with CD. These results are consistent with the immunological analysis discussed earlier (Sections 1.1, 1.2, and 1.3), that gluten-caused intestinal damage may result from levels 1 and 2, not just level 3.

Our results are consistent with earlier research indicating that people can have gluten-induced and biopsy-proven intestinal damage without possessing DQ2.5, DQ2.2, or DQ8 haplotypes. The existence of biopsy-proven intestinal damage in subjects not carrying CD-associated HLA DQ alleles may indicate that gluten damage leading to a “disorder” is started by innate immune actions that are genotype-independent (levels 1 and 2) and thus occurs prior to the acquired immune response. The fact that subjects’ symptoms improved when on GFD, plus the fact that the subjects were also known to be giardia-negative and olmesartan-negative, indicates that the source of the original damage is dietary gluten. Therefore, all our subjects had GRD, either in the form of CD or NCGS though CD is present only in those with DQ2.5, DQ2.2, and DQ8 haplotypes.

4.1. Haplotype Dominance Relationship in CD

Celiac Disease is a recognizable genetic entity. We thus agree with Sollid and Thorsby [51] that CD should be classified as a dominant disorder of the HLA DQ A1 and B1 loci. Greenberg et al. [52] originally proposed a recessive action in CD, but recessive alleles synthesize *non-functional* proteins and therefore primarily cause diseases in which a function is missing. Dominant action is indicated for CD [35,51] by the mechanism of HLA gene action: that the responsible HLA alleles (e.g., DQ2.5) make a functional protein (the α–β heterodimer) that *actively* generates an acquired immune system reaction to gliadin. Because of the dominant action of the HLA haplotypes, it should be more generally admitted by clinicians and researchers that the prevalence of CD may be roughly equal to the prevalence of the disease-causing haplotypes, and thus that genotype is both “necessary and sufficient” to predict CD [35]. Therefore, counting all the GRD, it is likely that a major part of the world population (over 30%) is affected by gluten.

4.2. Genotype-phenotype Relationship

Regarding HLA DQ haplotypes, our data (Figure 3) are consistent with the following view of A1 and B1 allele action in CD:

1. The A1*05 and B1*02 alleles (DQ2.5) serve as the strongest activators of antibody formation in CD. Lower antibody amounts are produced by A1*03 B1*0302 (DQ8) and A1*02 B1*02 (DQ2.2) haplotypes.

2. CD also depends on homozygosity for the above haplotypes: a person who is DQ2.5/DQ2.5 activates maximal antibody production.
3. The remaining genotypes (DQX.x/DQX.x) in homozygous form (i.e., not combined with DQ2.5, DQ2.2, or DQ8) did not produce antibodies at greater than reference values. However, the fact that these subjects had positive biopsies and symptoms indicates that intestinal damage can occur in a wider range of people than is currently thought (i.e., that epithelial damage is not restricted to CD and may occur as well in NCGS).

Previous researchers have developed ways to quantify the relation between HLA genotype and CD. However, because none of these studies quantified both genotype and phenotype, we consider our analysis (Figure 3) to be unique and quantitatively useful.

4.3. Summary

Subjects with haplotype binding scores of zero showed lack of elevated antibodies (Figure 3). This is the basis for our suggestion that HLA haplotype and antibody production may be unrelated to intestinal epithelial damage, which was present in all subjects. In agreement with this dissociation, the number of symptoms was consistent across haplotypes (Figure 4). Thus, the signs and symptoms appear likely to relate in large part to innate immune mechanisms that are genotype-independent (levels 1 and 2) and not solely to level 3 acquired immune mechanisms, which depend on HLA genotype.

Our data thus help reconcile the apparent contradictions between NCGS and CD. In our multilevel hypothesis (Sections 1.1, 1.2, and 1.3), level 2 innate immune responses initiate epithelial cell damage. This idea is consistent with the suggestion [30] that level 2 innate mechanisms of inflammation (in their term “epithelial stress”) can be a feature of GRD that occurs independently of the level 3 acquired immune response. In agreement with this idea, our data described above (Sections 2 and 3) suggest that biopsy-proven villous atrophy can be present in people who do not have HLA DQ2.5, DQ2.2, or DQ8 genotypes, meaning they do not activate acquired immune responses.

However, people with CD-associated genotypes activate acquired immune responses in which antibodies (level 3) are produced. The acquired response may magnify the damage initiated by level 2 via a mechanism suggested by Setty et al. [30], in which the cytotoxic profile of IEL is fully acquired by activation of acquired immune processes. Therefore, a person with CD may have more epithelial damage than a person with NCGS [53].

To summarize the relation among the three levels of gliadin action, we propose that direct gliadin action (level 1 – tight junction remodeling) allows infiltration of gliadin peptides that (level 2) allow innate immune processes to initiate damage of epithelial cells. This is followed (only in people with CD-associated genotypes) by activation of acquired immune responses (level 3) that magnify earlier innate cytotoxic effects.

The two forms of GRD can be characterized in terms of the three physiological responses to gluten discussed earlier (Sections 1.1, 1.2, and 1.3): someone with CD will experience levels 1, 2, and 3, while someone with NCGS will experience levels 1 and 2 but not 3. Though our subjects who were DQX.x homozygotes did not produce tTGA/EMA, they had biopsy-proven intestinal damage and their symptoms improved on a GFD. Levels 1 and 2 should be considered evidence for GRD aside from CD, implying that anyone who eats gluten is susceptible to inflammation and epithelial stress.

A major implication for clinicians is that regardless of genotype, for all people with signs or symptoms that suggest GRD, clinicians should follow up with additional diagnostic tests and GFD. These people should not be ignored just because they do not have DQ2.5, DQ2.2, or DQ8 HLA genotypes.

ACKNOWLEDGEMENTS

We received valuable comments on the manuscript from Lisa Shaver, ND, LSOM, Lac; and from William Messer, MD, PhD, Department of Molecular Biology and Immunology, Oregon Health and Sciences University School of Medicine. Preliminary versions were presented at meetings of the International Celiac Disease Symposia in Sorrento, Italy (2022) and Sheffield, England (2024).

The authors have no competing interests.

References

[1] Kagnoff, MF. Celiac disease: pathogenesis of a model immunogenetic disease. (2007). J. Clin. Invest. 2007; 117: 41–49.

[2] Jabri, B. and Sollid, L. Tissue-mediated control of immunopathology in coeliac disease. Nature Rev. Immunol. 2009; 9: 858–870.

[3] Drago, S. El Asmar, R., Di Pierro, M, et al. Gliadin, zonulin and gut permeability: Effects on celiac and non-celiac intestinal mucosa and intestinal cell lines. Scand. J. Gastroenterol. 2006; 41: 408–419.

[4] Lionetti, E., Leonardi, S., Franzonello, C., Mancardi, M., Ruggieri, M., and Catassi, C. Gluten Psychosis: Confirmation of a New Clinical Entity. Nutrients 2015; 7: 5532–5539.

[5] Abadie, V., Kim, SM, LeJeune, T., et al. IL-15, gluten and HLA-DQ8 drive tissue destruction in coeliac disease. Nature. 2020; 578: 600–604.

[6] Jabri, B., Kasarda, D., and Green, PH. Innate and adaptive immunity: the Yin and Yang of celiac disease. Immunol. Rev. 2005; 206: 219–231.

[7] Ciccocioppo, R., Di Sabatino, A. and Corazza, GR. The immune recognition of gluten in coeliac disease. Clin. Exper. Immunol. 2005; 140: 408–416.

[8] Graziano, M. and Rossi, M. An update on the cutaneous manifestations of coeliac disease and non-coeliac gluten sensitivity. Int. Rev. Immunol. 2018; 37(6): 291-300.

[9] Hadjivassiliou, M. Neurological Dysfunction in Coeliac Disease and Non-Coeliac Gluten Sensitivity. Am J Gastroenterol. 2016; 111: 561–567.

[10] Sapone, A., Lammers, KM, Mazzarella, G., et al. Differential Mucosal IL-17 Expression in Two Gliadin-Induced Disorders: Gluten Sensitivity and the Autoimmune Enteropathy Celiac Disease. Int Arch Allergy Immunol 2010; 152: 75–80.

[11] Lerner, A., Jeremias, P., and Matthias, T. Outside of Normal Limits: False Positive/Negative Anti TG2 Autoantibodies. International Journal of Celiac Disease 2015; 3: 87–90.

[12] Schuppan, D. Pickert, G., Ashfaq-Khan, M., and Zevallos, V. Non-celiac wheat sensitivity: differential diagnosis, triggers and implications. Best Practice & Research Clinical Gastroenterology 2015; 29: 469–476.

[13] Pena, AS. Immunogenetics of non celiac gluten sensitivity. Gastroenterol. Hepatol. Bed Bench 2014; 7(1): 1–5.

[14] Caio, G., Volta, U., Sapone, A., Leffler, DA, De Giorgio, R., Catassi, C. and Fasano, A. Celiac disease: a comprehensive current review. BMC Medicine 2019; 17: 142.

[15] Husby, S., Koletzko, S., Korponay-Szabo, I.R., et al. European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. J. Pediatr. Gastroenterol. Nutr. 2012; 54: 136–160.

[16] Daulatzai, M. Non-celiac gluten sensitivity triggers gut dysbiosis, neuroinflammation, gut-brain axis dysfunction, and vulnerability for dementia. CNS & Neurological Disorders - Drug Targets 2015; 14: 110–131.

[17] Hue, S., Mention, JJ, Monteiro, RC, et al. A Direct Role for NKG2D/MICA Interaction in Villous Atrophy during Celiac Disease. Immunity 2004; 21: 367–377.

[18] Dale, H.F., Biesiekierski, JR, and Lied, GA. Non-coeliac gluten sensitivity and the spectrum of gluten-related disorders: an updated overview. Nutrition Research Reviews 2019; 32: 28–37.

[19] Hollon, J., Puppa, E.L., Greenwald, B., Goldberg, E. Guerrerio, A., and Fasano. Effect of Gliadin on Permeability of Intestinal Biopsy Explants from Celiac Disease Patients and Patients with Non-Celiac Gluten Sensitivity. Nutrients 2015; 7, 1565–1576.

[20] Marsh, M.N. Gluten, Major Histocompatibility Complex, and the Small Intestine: A Molecular and Immunobiologic Approach to the Spectrum of Gluten Sensitivity (‘Celiac Sprue’). Gastroenterol. 1992; 102: 330–354.

[21] Jabri, B. and Sollid, LM. T cells in celiac disease. J Immunol. 2017; 198: 3005–3014.

[22] Samasca, G. Iancu, M., and Butnariu, A. Controversies in the laboratory diagnosis of celiac disease in children; new haplotypes discovered. Roum. Arch. Microbiol. Immunol. 2010; 69: 119–124.

[23] Dieterich, W., Trapp, D., Esslinger, B., et al. Identification of tissue transglutaminase as the autoantigen of celiac disease. Nature Medicine 1997; 3: 797–801.

[24] Jabri, B. and Abadie, V. IL-15 functions as a danger signal to regulate tissue-resident T cells and tissue destruction. Nat. Rev. Immunol. 2015; 15: 771–783.

[25] Klitz, WM, Maiers, M., Spellman, S. et al. New HLA haplotype frequency reference standards: High-resolution and large sample typing of HLA DR-DQ haplotypes in a sample of European Americans. Tissue Antigens 2003; 62: 296–307.

[26] Stepniak, D. and Koning, F. Celiac Disease—Sandwiched between Innate and Adaptive Immunity. Human Immunology 2006; 67: 460–468.

[27] Fallang, L-E, Bergseng, E., Hotta, K., et al. Differences in the risk of celiac disease associated with HLA-DQ2.5 or HLA-DQ2.2 are related to sustained gluten antigen presentation Nature Immunol. 2009; 10: 1096–1101.

[28] Murphy, K. and Weaver, C. Janeway’s Immunobiology, 9th ed. New York, Garland Science. 2017; (pages 11, 37, 635).

[29] Woodward, J. Coeliac disease. Medicine 2015; 43: 234–238.

[30] Setty, M., Discepolo, V., Abadie, V., et al. Distinct and Synergistic Contributions of Epithelial Stress and Adaptive Immunity to Functions of Intraepithelial Killer Cells and Active Celiac Disease. Gastroenterol. 2015; 149: 681–691.

[31] Maki, M. and Caporale, D. HLA-DQ1 Alpha and Beta Genotypes Associated with Non-Celiac Gluten Sensitivity. The FASEB Journal, 2018; 31(1).

[32] Mooney, PD, Leeds, JS, Libzo, N. et al. Case-finding for coeliac disease in secondary care: A prospective multicentre UK study. Dig. Liver Dis. 2014; 46: 32–35.

[33] Liu, E., Dong, F., Barón, A. et al. High incidence of Celiac Disease in a long-term study of adolescents with susceptibility genotypes. Gastroenterol. 2017; 152: 1329–1337.

[34] Camillo, CA. Addressing the ethical problem of underdiagnosis in the post-pandemic Canadian healthcare system. Healthcare Management Forum 2023; 36: 420–423.

[35] Roberts, MF and Bricher, SE. Theoretical Framework for the Study of Genetic Diseases caused by Dominant Alleles. 2023; Life 13: 733.

[36] Fasano, A. and Catassi, C. Current Approaches to Diagnosis and Treatment of Celiac Disease: An Evolving Spectrum. Gastroenterol. 2001; 120: 636–651.

[37] Murray, JA. Celiac disease in patients with an affected member, Type 1 Diabetes, iron-deficiency, or osteoporosis? Gastroenterol. 2005; 128: S52–S56.

[38] Rostami-Nejad, M. Romanos, J, Rostami,K et al. Allele and haplotype frequencies for HLA-DQ in Iranian celiac disease patients. World J Gastroenterol 2014; 20: 6302–6308.

[39] Grillo, R. Petronzelli F., Ferrante, P. et al. Unusual HLA typing in celiac disease. Disease Markers 1996; 13: 61–64.

[40] Congia, M., Cucca, F., Frau, F., et al. A gene dosage effect of the DQA1*05/DQB1*0201 allelic combination influences the clinical heterogeneity of celiac disease. 1994; Human Immunol. 40: 138–142.

[41] Karell, K. Louka, AS, Moodie, SJ, et al. HLA Types in Celiac Disease Patients not Carrying the *DQA1*05-DQB1*02* (DQ2) Heterodimer: Results From the European Genetics Cluster on Celiac Disease. Human Immunology 2003; 64: 469–477.

[42] Schuppan, D., Esslinger, B., and Dieterich, W. Innate immunity and coeliac disease. Lancet 2003; 362: 3–4.

[43] Mearin, M, Ivarsson, A., and Dickey, W. Coeliac disease: is it time for mass screening? Best Practice & Research: Clin. Gastroenterol.; Kidlington 2005; 19: 441–452.

[44] Ranganathan P. and Aggarwal R. Study designs: Part 3 - Analytical observational studies. Perspect. Clin. Res. 2019; 10:91–4.

[45] Ploski, R., Thorsby, E., and Sollid, M. On the HLA-DQ (A1*0501 B1*0201) - associated susceptibility in celiac disease: A possible gene dosage effect of DOBI*0201. 1993; Tissue Antigens 41: 173–177.

[46] Mubarak, A. Spierings, E., Wolters, VM et al. Children with celiac disease and high tTGA are genetically and phenotypically different. World J Gastroenterol 2013; 19(41): 7114–7120.

[47] Kurppa, K. Räsänen, T., Collin, P. et al. Endomysial antibodies predict celiac disease irrespective of the titers or clinical presentation. World J Gastroenterol 2012; 18: 2511–2516.

[48] Ross, SM. Introduction to Probability and Statistics for Engineers and Scientists, 5th ed.; Academic Press: San Diego, CA, USA. 2014; (pages 527–531).

[49] Green, PHR, Stavropoulos, SN, Panagi, SG, et al. Characteristics of Adult Celiac Disease in the USA: Results of a National Survey. Am. J. Gastroenterol. 2001; 96: 126–131.

[50] Lebwohl, B. Celiac Disease and the Forgotten 10 %: The ‘‘Silent Minority’’. Digest. Diseases and Sciences 2015; 60: 1517–1518.

[51] Sollid, L. and Thorsby, E. HLA Susceptibility Genes in Celiac Disease: Genetic Mapping and Role in Pathogenesis. Gastroenterol. 1993; 105: 910–922.

[52] Greenberg, DA, Hodge, SE, and Rotter, JI, Evidence for Recessive and against Dominant Inheritance at the HLA-"Linked" Locus in Coeliac Disease. Am J. Hum. Genet. 1982; 34: 263–277.

[53] Uhde, M., Ajamian, M., Caio, G., et al. Intestinal cell damage and systemic immune activation in individuals reporting sensitivity to wheat in the absence of coeliac disease. Gut 2016; 65: 1930–1937.