



REVIEW ARTICLE

HLA class II genes in precision-based care of childhood diseases: what we can learn from celiac disease

Giovanna Del Pozzo¹, Federica Farina¹, Stefania Picascia², Mariavittoria Laezza¹, Serena Vitale² and Carmen Gianfrani^{2,3}

Celiac disease (CeD) is a chronic immuno-mediated enteropathy caused by dietary gluten with marked autoimmunity traits. The human leukocyte antigen (HLA) class II heterodimers represent the main predisposing factor, although environmental agents, as viral infection, gut microbiota, and dietary regimen, also contribute to CeD risk. These molecules are involved in autoimmunity as they present self-antigens to autoreactive T cells that have escaped the thymic negative selection. In CeD, the HLA class II risk alleles, DQA1*05-DQB1*02 and DQA1*03-DQB1*03, encode for DQ2.5 and DQ8 heterodimers, and, furthermore, disease susceptibility was found strictly dependent on the dose of these genes. This finding questioned how the expression of HLA-DQ risk genes, and of relative surface protein on antigen-presenting cells, might be relevant for the magnitude of anti-gluten inflammatory response in CeD patients, and impact the natural history of disease, its pathomechanisms, and compliance to dietary treatment. In this scenario, new personalized medical approaches will be desirable to support an early, accurate, and non-invasive diagnosis, and to define genotype-guided preventive and therapeutic strategies for CeD. To reach this goal, a stratification of genetic risk, disease outcome, and therapy compliance based on HLA genotypes, DQ2.5/DQ8 expression measurement and magnitude of T cell response to gluten is mandatory.

Pediatric Research _____; <https://doi.org/10.1038/s41390-020-01217-4>

IMPACT:

- This article revises the current knowledge on how different HLA haplotypes, carrying the DQ2.5/DQ8 risk alleles, impact the onset of CeD.
- This review discusses how the expression of susceptibility HLA-DQ genes can determine the risk assessment, outcome, and prevention of CeD.
- The recent insights on the environmental factors contributing to CeD in childhood are reviewed.
- This review discusses the use of HLA risk gene expression as a tool to support medical precision approaches for an early and non-invasive diagnosis of CeD, and to define genotype-guided preventive and therapeutic strategies.

INTRODUCTION

In the past decades, advances in genetics knowledge and availability of big clinical and biological data have deeply changed the healthcare approach. A “precision medicine,” tailored on person’s genetic background, lifestyle, and environment, has gradually replaced the “one-size-fits-all” curative approach.¹ The successful potentiality of this emerging approach in medical practice for treatment and prevention of diseases is primarily dependent on individual genetic variability.

The major histocompatibility complex (MHC) or human leukocyte antigens (HLA) in humans is fundamental in the immune surveillance to microbial infections and neoplastic cell transformation. Owing to precise bindings of antigenic peptides, HLA molecules activate specific T cell-mediated response, with key role in disease prevention or outcome. However, certain HLA heterodimers of class II region are deleterious in autoimmunity, as they present self-antigens to autoreactive T cells that have escaped the thymic negative selection.² The HLA region includes a cluster of highly polymorphic genes that represents a sort of

person’s ID card. HLA class I and class II genes encode for heterodimeric proteins that, in specialized antigen-presenting cells (APCs), bind short peptides from a large diversity of immunogenic proteins. In autoimmune diseases (ADs), the HLA class II genes represent the main genetic predisposing factor, as the encoded molecules present self-antigenic peptides to T cells that survived the central immune tolerance.³ The activation of autoreactive T cells triggers an inflammatory cascade that leads to the destruction of own tissue.⁴

CeD, is an autoimmune-like food intolerance,⁵ occurring in subjects carrying DQA1*05-DQB1*02 (DR3/DQ2.5 haplotype) and DQA1*03-DQB1*03 (DR4/DQ8 haplotype) risk alleles. Very frequently, children with CeD are also affected by type 1 diabetes (T1D), due to the common HLA risk genes and pathomechanisms, although target organs and penetrance in the general population are different.⁶ Of note, children carrying both HLA-DQ2.5/DQ8 genes have a higher risk to have CeD and T1D comorbidities than the single disease alone.⁷ It has been demonstrated that predisposing DQ2.5 and DQ8 molecules present gluten- and

¹Institute of Genetics and Biophysics “Adriano Buzzati Traverso”, Italian National Council of Research(CNR), Naples, Italy; ²Institute of Biochemistry and Cell Biology, Italian National Council of Research(CNR), Naples, Italy and ³European Laboratory for the Investigation of Food Induced Diseases (ELFID), University Federico II, Naples, Italy
Correspondence: Giovanna Del Pozzo (giovanna.delpozzo@igb.cnr.it)

Received: 19 May 2020 Revised: 17 September 2020 Accepted: 25 September 2020

Published online: 29 October 2020

proinsulin-derived peptides more efficiently if post-translationally modified by an enzymatic reaction that deamidates glutamine in glutamic acid. The preferential binding of DQ2.5/DQ8 to negatively charged residues is a prerequisite to activate the autoantigen-reactive CD4⁺ T cells.⁸

This review discusses the current knowledge on the role of HLA class II genes in T cell response to gluten, and how the expression of HLA susceptibility genes can be used to better determine the risk assessment, outcome, and prevention of CeD.

CORRELATION BETWEEN WORLDWIDE SPREADING OF CED INTOLERANCE AND HLA CLASS II GENES

CeD is a complex and multifactorial pathology in which both genetic and environmental factors contribute to the disease onset.⁹ The main genetic risk for CeD is represented by DQA1 and DQB1 genes at HLA class II locus, encoding, respectively, the α and β chain of the HLA-DQ heterodimers. In particular, the great majority of patients (90–95%) express DQ2.5 heterodimer, encoded by DQA1*05 and DQB1*02 alleles, either in *cis* or in *trans* configuration, and carried by DR3-DQ2.5 and DR5-DQ7/DR7-DQ2.2 haplotypes, respectively. The remaining 5–10% of patients express the DQ8 heterodimer, encoded by DQA1*03 and DQB1*03 alleles carried by DR4-DQ8 haplotype. However, recent genome-wide association studies (GWAS) have identified a number of loci, located outside the MHC region, as additional genetic factors associated with increased risk to have CeD.¹⁰

The HLA-DQ2.5/DQ8 risk alleles are in the genome of 30–35% of Caucasian populations; however, <2% of them develops CeD, indicating that the presence of these specific HLA genes is necessary but not sufficient for the onset of the disease. Various environmental agents have been hypothesized to act as predisposing factors, among them the increased consumption of gluten-containing cereals, such as wheat, barley, and rye, has been accounted as the main one.^{11,12} In the recent years, the CeD is markedly raised in Asia, especially in the Southeastern part of the continent, because of a progressive Westernization of the diet with the introduction of a large amount of gluten-rich food, such as pizza and pasta.¹³ The worldwide incidence of CeD significantly correlates with the frequency of HLA risk genes and wheat intake, which globally ranges between 21 and 564 g per person per day.¹² Notably, in the African continent, the Burkina Faso records an almost absence of CeD diagnosis, as well as a very low frequency of HLA-DQ2.5/DQ8 genes combined to low wheat consumption.¹⁴ Nevertheless, in some countries, this direct correlation between CeD prevalence and HLA/wheat risk factors was not observed. For example, the neighboring countries of Algeria and Tunisia consume an elevated quantity of wheat and barley, and present similar frequency of DR3-DQ2.5 and DR4-DQ8 haplotypes; nevertheless, their CeD prevalences are totally different (5.6% in Algeria and 0.28% in Tunisia).¹⁵ This co-occurrence of the two risk factors is known as the “evolutionary paradox of CeD” and is explained by the positive role of CeD-predisposing haplotypes in the populations that have based their diet on wheat.¹⁵ A positive selection of the HLA-DQ2 genotype has been proposed because it is associated with the protection against dental caries, a disease causing a tremendous negative impact on the health and reproductive fitness of ancient populations.¹²

In Europe, a large variability of CeD frequency is reported, ranging from 1.6 to 2.3% in Finland, Sweden, and Italy to 0.3% in Germany.^{16–18} The comparable levels of gluten intake and of HLA risk alleles frequency in these countries suggest that other genetic and/or environmental factors might contribute to CeD development.

However, a great geographical variability exists in the frequency of the different haplotypes carrying the DQ2 and DQ8 risk genes. In a study on a prospective cohort of pediatric CeD patients from Southern Europe, it has been reported that the frequency of DR7-

DQ2 haplotype is higher in Southern Europe than in other Caucasian populations, in accordance with what is reported in adults. Delgado et al.¹⁹ noted similar phenotypic profile of the disease, including clinical and histological characteristics, between children carrying HLA DR3-DQ2.5 (80.2% in their cohort of enrolled patients) and HLA DR7-DQ2.2 (9.9%) haplotypes in the absence of DR3-DQ2.5, highlighting that HLA DR7-DQ2.2 haplotype should be typed for a correct screening of risk in this geographic area.

The multinational study (TEDDY) assessed the CeD incidence on a large group of children at high genetic risk for both T1D and CeD, identified at birth on the basis of their DR3-DQ2.5 or DR4-DQ8 haplotypes. Teddy study has also evaluated the effects of sex, family history of CeD, and country of origin. The young patients that carry DR3-DQ2.5, especially homozygotes, were found to be at high CeD risk early in childhood. In particular, the frequency of DR3-DQ2.5 haplotype among children, who had a first-degree relative with CeD, was higher than the frequency among subjects without affected family members (86% vs. 62%).²⁰ Similar results on the highest risk carried by homozygous HLA DR3-DQ2.5 subjects were obtained by the PreventCeD study, which monitored newborns' first-degree relatives of celiac patients for CeD onset since birth.²¹

FROM DIAGNOSIS TO FOLLOW-UP

CeD diagnosis, both in children and adults, has been provided by the determination of IgA antibodies against type-2 (tissue) transglutaminase (anti-tTG) and IgA endomysial antibodies (EMA-IgA), HLA-DQ2, and/or DQ8 genotype and by the evaluation of the histological lesion through endoscopic biopsies taken from the second or third part of the duodenum. Recently, ESPGHAN (The European Society for Pediatric Gastroenterology Hepatology and Nutrition, <http://www.espghan.org/>) has issued new guidelines for the diagnosis of CeD in children, according to which biopsy is no longer necessary in pediatric patients, in the presence of symptoms, serological positivity (high-serum tTG-IgA, $\geq 10 \times$ ULN (upper limit of normal)) and HLA risk genes.^{22,23} Conversely, HLA-DQ2 and DQ8 typing is not required in patients with positive tTG-IgA and villous atrophy or with high tTG and EMA positivity. Therefore, the typing of HLA-DQ loci resulted to be particularly useful to refine and/or complete the diagnostic work-up for CeD in some complex clinical cases, where the absence or presence of DQ2/DQ8 genes can, respectively, rule out or support the diagnosis.

The only therapy currently available is the gluten-free diet (GFD), which consists of the total exclusion of gluten from the diet. If GFD is carefully respected, results in a full recovery of gut mucosa morphology and function, and positively influencing the clinical course of any ADs might be associated with CeD.^{24–26}

A follow-up within 6–12 months from diagnosis, and subsequently every 1–2 years, is necessary to verify the adherence to the GFD, the nutritional adequacy, and to improve the patient's quality of life. However, although GFD is advantageous for the restoration of the normal morphology and functionality of the intestinal mucosa, it does not guarantee the recovery of immune tolerance to gluten, even after many years of strict GFD.²⁶

CROSSTALK AMONG APC AND T CELLS IN THE INFLAMMATORY RESPONSE TO GLUTEN

Professional APCs are a heterogeneous population of cells, including dendritic cells (DCs), macrophages (M Φ), and B cells, specialized in antigen presentation, that express HLA molecules on cells' surface. The CeD-associated DQ2 and DQ8 heterodimers bind with high-affinity peptides containing negative charged amino acids at the HLA anchor residues (P4, P6, and P7 position for HLA-DQ2.5, at P1 and P9 for HLA-DQ8).²⁷ Gluten proteins have few negatively charged residues, although having many

Table 1. Risk classification according to genotype.

DR genotypes	DQ alleles	DQ heterodimers	Risk for celiac disease
DR3	DQA1*0501-DQB1*0201	DQ2.5	Very high
DR3	DQA1*0501-DQB1*0201	DQ2.5	DQ2.5 in <i>cis</i> with two copies of DQA1*05 and DQB1*02
DR3	DQA1*0501-DQB1*0201	DQ2.5	Very high
DR7	DQA1*0201-DQB1*0202	DQ2.2	DQ2.5 in <i>cis</i> with one copy of DQA1*05 and two copies of DQB1*02
DR5	DQA1*0505-DQB1*0301	DQ7	High
DR7	DQA1*0201-DQB1*0202	DQ2.2	DQ2.5 in <i>trans</i> with one copy of DQA1*05 and DQB1*02
DR3	DQA1*0501-DQB1*0201	DQ2.5	Moderate
DRX	DQA1*-DQB1*	DQX	DQ2.5 in <i>cis</i> with one copy of DQA1*05 and one copy of DQB1*02
DR3	DQA1*0501-DQB1*0201	DQ2.5	Moderate
DR5	DQA1*0505-DQB1*0301	DQ7	DQ2.5 in <i>cis</i> and <i>trans</i> with two copies of DQA1*05 and one copy of DQB1*02

glutamines that are a good substrate for the enzymatic activity of tTG. This ubiquitous enzyme is activated by the inflammatory conditions, such as those of celiac intestinal mucosa, and catalyses a covalent process converting glutamine (neutrally charged) in glutamic acid (negatively charged), within the consensus sequences QXP, particularly abundant in gluten proteins.²⁸ Thus, deamidation process provides negative charges necessary to strongly favor the binding of peptides to DQ2.5 and DQ8 molecule on the surface of APC. Among APCs, intestinal DCs are pivotal in activating celiac T cells. In CeD patients, the participation of diverse DC subsets in the pathological processes of CeD, as well as the involvement of tolerogenic DCs in Tregs' development, has been demonstrated.²⁹ According to these findings, Ráki et al.³⁰ observed an accumulation of CD11c⁺ DCs in celiac lesions, able to activate gluten-reactive T cells. Especially interferon- α (IFN α)-producing DCs contribute to the T-helper type 1 response in celiac disease.³¹

Among APCs, M Φ from patients with CeD has shown greater antigen-presenting ability, which is exemplified by the upregulated expression of the co-stimulatory molecules CD80, CD86, and CD40, and a more highly activated state of T cells.³² Gliadin triggers a potent pro-inflammatory phenotype in human primary M Φ and a higher abundance of total macrophages from intestinal biopsies of CeD patients was confirmed when compared to healthy controls.³³ As a consequence of the environmental stimuli, a significant increase of several pro-inflammatory cytokines and chemokines in young CeD patients (median age 6.5 years) was found correlating with small bowel mucosa inflammation and damage.³⁴

B cells play a very important role in CeD considering that gluten-specific T cells induce B lymphocytes to produce antibodies directed against deamidated gluten peptides and tTG reflecting efficient T-B collaboration. A recent study³⁵ suggested that B cells produce antibodies targeting an N-terminal epitope of tTG and this event coincides with clinical onset of CeD, demonstrating that tTG-reactive B cells with this epitope specificity could be the main APC for autoantigen-specific T lymphocytes in CeD patients. A subset of plasma cells has also been identified from intestinal biopsies of patients expressing B cell receptors specific for gluten peptides or tTG. These plasma cells, expressing HLA class II genes, are the most abundant APC presenting the immunodominant gluten peptide DQ2.5-glia- α 1a in the tissues from these patients and contemporaneously promote and maintain intestinal inflammation.³⁶

Since gluten constitutes a very heterogeneous family of proteins, it is not surprising that many distinct T cell epitopes arise from its hydrolysis. The repertoire of gluten epitopes comprises a broad spectrum of peptides stimulating CD4⁺ T cells. Many distinct epitopes recognized by CeD-associated HLA molecules, especially HLA-DQ2.5, have been identified in all

wheat gluten family, as well as in all related prolamins of rye and barley, particularly, three of them from α -, ω -, and γ -gliadins were defined immunodominant. Several studies have addressed the repertoire of wheat gluten epitopes in adults and pediatric patients and hierarchy of dominant epitopes,^{21,37–39} some of them originated by tTG post-translational modification. A recent work demonstrated that children and adults with untreated CeD have T cell reactivity to a broad range of gluten peptides at the time of diagnosis, and that T cells from children equally react to both native and tTG deamidated gluten peptides.³⁰ Interestingly, the reactivity of T cells to gluten immunogenic peptides is highly variable, and a hierarchy of gluten epitopes has been reported both in children and in adults patients.^{38,39} These findings have an impact on the definition of the pool of immunodominant gluten epitopes, with implications for precision therapeutic approaches based on peptide-based immunotherapy.³⁰

FROM GENE DOSAGE TO EXPRESSION MODEL OF HLA RISK GENES IN CED

As described above, the adaptive CD4⁺ T cell response to wheat gluten is the key step in CeD pathogenesis. However, the anti-gluten T cell response must achieve a certain threshold in order to be pathogenic and to induce intestinal damage. To reach this threshold, several factors are important, that is, a number of HLA molecules and amount and stability of gliadin-HLA complex expressed on APC surface stimulating the CD4⁺ T cells.¹⁵ In this contest, a "gene dosage model," correlating the HLA-DQ antigen presentation of gluten epitopes to intestinal CD4⁺ T cell response, has been proposed, which is useful to explain the risk stratification associated with disease development. Accordingly, the strength of the pathogenetic CD4⁺ T cell response to gluten epitopes should be strictly dependent on the DQA1*05 and DQB1*02 number of alleles carried by APCs, which establish the amount of DQ2.5 heterodimers presenting antigenic peptides on APC surface. Individuals homozygous for the DR3-DQ2.5 haplotype, carrying two copies of DQA1*05:01 and DQB1*02:01 alleles, showed the highest risk disease (see Table 1) and induced the highest T cell response,⁴⁰ as well as subjects with DR3-DR7/DQ2.5-DQ2.2 genotype carrying one copy of DQA1*05:01 and two copies of DQB1*02 alleles. Nevertheless, DR5-DR7/DQ7-DQ2.2 individuals, with only one copy of DQA1*0505 and DQB1*0202 alleles in *trans* configurations, are in the high-risk class. Unexpectedly, DR3-DRX/DQ2.5-DQX subjects, also with only one copy of DQA1*0501 and DQB1*0201 alleles in *cis* configuration, as well as DR3-DR5/DQ2.2-DQ7 subjects, with two copies of DQA1*05 alleles (more specifically of DQA1*05:01 and DQA1*05:05) and one copy of DQB1*02 allele, are included in moderate risk class (see Table 1). We highlight that the DQA1*05:01 allele of the DR3-DQ2.5 haplotype was almost identical to the DQA1*05:05 allele of the

DR5-DQ7 haplotype (only differing at one single residue in the leader sequence), and the DQB1*02:01 allele of the DR3-DQ2.5 was almost similar to the DQB1*02:02 allele of the DR7-DQ2.2 haplotype (only differing at one single residue at position 135 in the membrane-proximal domain).⁴¹ As a matter of fact, the risk classifications cannot be exhaustively explained by the “gene dosage model,” because many incongruities exist between the number of predisposing alleles and the different class of risk. One possible explanation invokes the contribution of other loci in linkage with HLA class II, including HLA class I genes. A GWAS identified as additional risk regions the class I, in particular with B*0801 allele, in strong linkage disequilibrium with DR3-DQ2.5 genes.¹⁰ Of note, a recent study from our group has demonstrated the presence in CeD patients of CD8⁺ T cells reactive to HLA class I-restricted gliadin peptides and releasing a high amount of IFN γ .^{42,43}

Recent research by our group addressed the question with a different approach. The RNA amount of DQA1*05 and DQB1*02 risk alleles was measured in APC with DQ2.5 genes either in *cis* (DR1/DR3)⁴⁴ or *trans* (DR5/DR7)⁴⁵ configurations. We demonstrated that the risk alleles, when in heterozygosis, are more expressed than non-CeD-associated allele on the other chromosome, in different haplotypes. In other words, one copy of risk allele is sufficient to synthesize high RNA amount, not statistically different with respect to the quantity produced by double copies of alleles. As a consequence, APC carrying homozygous genotype DR3-DQ2.5, heterozygous DR3-DR1/DQ2.5-DQ5, and DR5-DR7/DQ7-DQ2.2^{44,45} genotypes induce high expression of DQ α 1*05 and DQ β 1*02 protein chains. This results in a comparable DQ2.5 heterodimer density on APC either homozygous or heterozygous, generating a similar number of HLA–gliadin complexes. Gliadin-specific CD4⁺ T cell proliferation and IFN γ production does not change when these effector cells are stimulated by APC with different genotypes.

These findings provide a new model (see Fig. 1) in which the achievement of the threshold to activate pathogenic CD4⁺ T cell response is dependent on the expression of HLA risk genes and not on the gene dosage. One main consequence of this phenomenon is that the magnitude of the CD4⁺ T cell response depends on the antigen concentration and, therefore, by the amount of dietary gluten released in gut lumen upon the proteolytic digestive process.

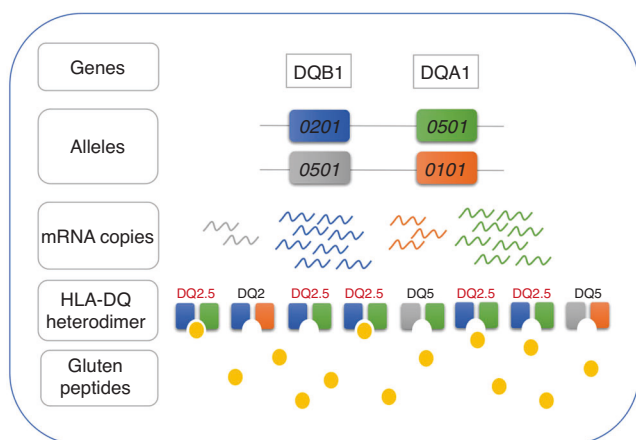


Fig. 1 The HLA haplotype carrying the DQ2.5 encoding genes, DQB1*02 and DQA1*05, is associated with a high/moderate risk to develop CeD. Interestingly, CeD-associated DQA1*05 and DQB1*02 mRNAs are more abundant than non-CeD-associated mRNA amount. The differential expression determines a high expression of HLA-DQ2.5 heterodimer on the surface of APC. As consequence, the expression of predisposing HLA genes became so influential along with the genetic configuration.

HLA-DQ RISK ALLELES' DIFFERENTIAL EXPRESSION

The differential expression of HLA class II risk alleles is a phenomenon that we also demonstrated in T1D and multiple sclerosis. These autoimmune diseases are associated with different DQ or DR predisposing alleles that, when in heterozygous, show the expression >50%, influencing the density of HLA molecules presenting self-antigens.^{46,47} Similarly, other groups demonstrated high expression of risk genes associated with different autoimmune pathologies, such as vitiligo^{48,49} and systemic lupus erythematosus.⁵⁰ In order to unravel the mechanism of differential expression, we investigated the gene regulation through the click-IT chemistry. We demonstrated that the great expression of risk alleles DQA1*05 and DQB1*02 is mainly determined by a high rate of transcription of risk alleles with respect to the non-CeD-associated alleles⁴⁵. This differential expression might be consequent to the presence of regulatory motives in the intergenic regions at the HLA locus.⁵¹ It is known that a regulative function has been assigned to XL9 site,⁵² located in the intergenic region between DRB1 and DQA1. This region shows high levels of acetylation, associated with more accessible chromatin, in which binding sites of many transcription factors have been identified.⁵⁰ As a consequence, polymorphism in this regulatory region impacts on the function of transcription factors and on the expression of HLA genes. In addition to specific single variants, also multiple haplotype changes⁵³ may determine increased gene expression. These intergenic regulatory variants, and/or haplotype polymorphisms, might explain the high mRNA level of DQA1*05 and DQB1*02 alleles in *cis* or *trans* configurations with respect to non-CeD-associated alleles.

FUTURE PERSPECTIVE FOR A PRECISION ASSESSMENT OF CED RISK BASED ON HLA GENE EXPRESSION

The strict correlation between the positivity for HLA-DQ2.5/DQ8 alleles and the risk to have CeD has been assessed for a long time. Furthermore, it is also known that risk contribution depends on the various HLA-DQ2.5/DQ8 carrying haplotypes.⁵⁴ However, the open question concerns the measurement and precise prevision to develop a disease that can affect people even late in life. It is also conceivable that a more precise evaluation of CeD risk may bring out the submerged part of the “celiac iceberg,” a still timely epidemiological model. According to this model, only a small fraction of cases are clinically diagnosed, whereas most part of cases remains underdiagnosed, having atypical and/or paucisymptomatic clinical presentations.^{55,56} A more precise risk assessment could improve the quality of life also for those children being first-degree relatives of CeD patients.

Several papers reported a risk gradient for CeD according to homozygous or heterozygous DQ2 genotype and associated with the presence of at least a copy of DQA1*05. Recent meta-analysis emphasized that also a single “dose” of HLA-DQB1*02 is associated with a relatively high risk for pediatric CeD.^{57–59} However, as discussed above, the gene dosage model is not satisfactory to explain the correlation between the number of predisposing genes and the various classes of risk to have CeD. Based on our recent findings, we propose the “gene expression model” that could fill the non-sufficient “gene dosage model.” We demonstrated that the mean percentages of expression of DQA1*05 and DQB1*02 risk alleles are, respectively, ~75–76% with respect to 24–25% of non-CeD-associated alleles.⁴⁵ More interesting, a significant difference in the risk allele expression was measured in APCs from celiacs with respect to healthy controls. Indeed, the differences corresponded to an increment of 7–8% DQ2 mRNA. These Δ value, if further confirmed in larger celiac cohorts, could represent a valid tool to have an early diagnosis for family members of affected subjects, or to estimate the times required for disease remission in patients on GFD. In addition, following the analysis of a large number of patients, our final goal

will be to use the Δ value for the assignment of subjects to different class of CeD risk.

CONCLUSION

Our findings demonstrated the relevance of *expression dosage* with respect to *gene dosage* for HLA class II predisposing alleles in the establishment of autoimmunity. Moreover, the assessment of expression could represent a promising marker to classify the level of CeD susceptibility, to support diagnosis for family members of pediatric patients and compliance to GFD therapy.

ACKNOWLEDGEMENTS

This work was supported with funds from the Italian Ministry of Health for the research grant RF-2016-02361372 to C.G., and from Italian Ministry of Research and DSB-CNR for Progetto Bandiera InterOmics 2017 to C.G. and G.D.P. We thank the Italian Celiac Disease Foundation (FC) for the Triennial Fellowship to S.V. (FC_009_2016) and to M.L. (Fellowship_6_FC_2019).

AUTHOR CONTRIBUTIONS

F.F., S.P., S.V., and M.L. made substantial contributions to the literature review and drafting the manuscript. G.D.P. and C.G. provided intellectual contributions and critically revised the manuscript. All authors read and approved the final manuscript for publication.

ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Budin-Ljøsne, I. & Harris, J. R. Ask not what personalized medicine can do for you - ask what you can do for personalized medicine. *Public Health Genomics* **18**, 131–138 (2015).
- Fernando, M. M. A. et al. Defining the role of the MHC in autoimmunity: a review and pooled analysis. *PLoS Genet.* **4**, e1000024 (2008).
- Takaba, H. & Takayanagi, H. The Mechanisms of T cell selection in the thymus. *Trends Immunol.* **38**, 805–816 (2017).
- Dendrou, C. A., Petersen, J., Rossjohn, J. & Fugger, L. HLA variation and disease. *Nat. Rev. Immunol.* **18**, 325–339 (2018).
- Lindfors, K. et al. Coeliac disease. *Nat. Rev. Dis. Primers* **5**, 3 (2019).
- Viken, M. K. et al. HLA class II alleles in Norwegian patients with coexisting type 1 diabetes and celiac disease. *HLA* **89**, 278–284 (2017).
- Smigoc Schweiger, D. et al. High-risk genotypes HLA-DR3-DQ2/DR3-DQ2 and DR3-DQ2/DR4-DQ8 in co-occurrence of type 1 diabetes and celiac disease. *Autoimmunity* **49**, 240–247 (2016).
- Van Lummel, M. et al. Type 1 diabetes-associated HLA-DQ8 transdimer accommodates a unique peptide repertoire. *J. Biol. Chem.* **287**, 9514–9524 (2012).
- Meijer, C., Shamir, R., Szajewska, H. & Mearin, L. Celiac disease prevention. *Front. Pediatr.* **6**, 368 (2018).
- Gutierrez-Achury, J. et al. Fine mapping in the MHC region accounts for 18% additional genetic risk for celiac disease. *Nat. Genet.* **47**, 577–578 (2015).
- de Punder, K. & Pruimboom, L. The dietary intake of wheat and other cereal grains and their role in inflammation. *Nutrients* **5**, 771–787 (2013).
- Lionetti, E. & Catassi, C. Co-localization of gluten consumption and HLA-DQ2 and -DQ8 genotypes, a clue to the history of celiac disease. *Dig. Liver Dis.* **46**, 1057–1063 (2014).
- Makharia, G. K. & Catassi, C. Celiac disease in Asia. *Gastroenterol. Clin. N. Am.* **48**, 101–113 (2019).
- Cataldo, F., Lio, D., Simpore, J. & Musumeci, S. Consumption of wheat foodstuffs: not a risk for celiac disease occurrence in Burkina Faso [3]. *J. Pediatr. Gastroenterol. Nutr.* **35**, 233–234 (2002).
- Abadie, V., Sollid, L. M., Barreiro, L. B. & Jabri, B. Integration of genetic and immunological insights into a model of celiac disease pathogenesis. *Annu. Rev. Immunol.* **29**, 493–525 (2011).
- Mustalahti, K. et al. The prevalence of celiac disease in Europe: results of a centralized, international mass screening project. *Ann. Med.* **42**, 587–595 (2010).

- Gatti, S. et al. Increased prevalence of celiac disease in school-age children in Italy. *Clin. Gastroenterol. Hepatol.* **18**, 596–603 (2020).
- Scherf, K. A. et al. Recent progress and recommendations on celiac disease from the working group on prolamin analysis and toxicity. *Front. Nutr.* **7**, 29 (2020).
- Delgado, J. F. et al. Paediatric celiac patients carrying the HLA-DR7-DQ2 and HLA-DR3-DQ2 haplotypes display small clinical differences. *Acta Paediatr.* **103**, 238–242 (2014).
- Liu, E. et al. Risk of pediatric celiac disease according to HLA haplotype and country. *N. Engl. J. Med.* **371**, 42–49 (2014).
- Camarca, A. et al. Gliadin-reactive T cells in Italian children from preventCD cohort at high risk of celiac disease. *Pediatr. Allergy Immunol.* **28**, 362–369 (2017).
- Husby, S. et al. European Society Paediatric Gastroenterology, Hepatology and Nutrition Guidelines for Diagnosing Coeliac Disease 2020. *J. Pediatr. Gastroenterol. Nutr.* **70**, 141–156 (2020).
- Werkstetter, K. J. et al. Accuracy in diagnosis of celiac disease without biopsies in clinical practice. *Gastroenterology* **153**, 924–935 (2017).
- Lauret, E. & Rodrigo, L. Celiac disease and autoimmune-associated conditions. *BioMed. Res. Int.* **2013**, 127589 (2013).
- Hogen Esch, C. E. et al. The PreventCD Study design: towards new strategies for the prevention of coeliac disease. *Eur. J. Gastroenterol. Hepatol.* **22**, 1424–1430 (2010).
- Lionetti, E. et al. Introduction of gluten, HLA status, and the risk of celiac disease in children. *N. Engl. J. Med.* **371**, 1295–1303 (2014).
- Sollid, L. M. The roles of MHC class II genes and post-translational modification in celiac disease. *Immunogenetics* **69**, 605–616 (2017).
- Sollid, L. M. & Jabri, B. Celiac disease and transglutaminase 2: a model for post-translational modification of antigens and HLA association in the pathogenesis of autoimmune disorders. *Curr. Opin. Immunol.* **23**, 732–738 (2011).
- Vorobjova, T., Uibo, O., Heilman, K. & Uibo, R. Increased density of tolerogenic dendritic cells in the small bowel mucosa of celiac patients. *World J. Gastroenterol.* **21**, 439–452 (2015).
- Ráki, M. et al. Similar responses of intestinal T cells from untreated children and adults with celiac disease to deamidated gluten epitopes. *Gastroenterology* **153**, e4 (2017).
- Di Sabatino, A. et al. Evidence for the role of interferon- α production by dendritic cells in the Th1 response in celiac disease. *Gastroenterology* **133**, 1175–1187 (2007).
- Ma, W. T., Gao, F., Gu, K. & Chen, D. K. The role of monocytes and macrophages in autoimmune diseases: a comprehensive review. *Front. Immunol.* **10**, 1140 (2019).
- Serena, G. et al. Intestinal epithelium modulates macrophage response to gliadin in celiac disease. *Front. Nutr.* **6**, 167 (2019).
- Vorobjova, T. et al. Celiac disease in children, particularly with accompanying type 1 diabetes, is characterized by substantial changes in the blood cytokine balance, which may reflect inflammatory processes in the small intestinal mucosa. *J. Immunol. Res.* **2019**, 6179243 (2019).
- Iversen, R. et al. Efficient T cell-B cell collaboration guides autoantibody epitope bias and onset of celiac disease. *Proc. Natl Acad. Sci. USA* **116**, 15134–15139 (2019).
- Høydaal, L. S. et al. Plasma cells are the most abundant gluten peptide MHC-expressing cells in inflamed intestinal tissues from patients with celiac disease. *Gastroenterology* **156**, 1428–1439.e10 (2019).
- Camarca, A. et al. 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).
- Hardy, M. Y. et al. Consistency in polyclonal T-cell responses to gluten between children and adults with celiac disease. *Gastroenterology* **149**, 1541–1552.e2 (2015).
- Tye-Din, J. A. et al. Comprehensive, quantitative mapping of T cell epitopes in gluten in celiac disease. *Sci. Transl. Med.* **2**, 41ra51 (2010).
- Vader, W. et al. The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. *Proc. Natl Acad. Sci. USA* **100**, 12390–12395 (2003).
- Sollid, L. M. The roles of MHC class II genes and post-translational modification in celiac disease. *Immunogenetics* **69**, 605–616 (2017).
- Picascia, S. et al. Gliadin-Specific CD8 + T cell responses restricted by HLA class I A*0101 and B*0801 molecules in celiac disease patients. *J. Immunol.* **198**, 1838–1845 (2017).
- Mazzarella, G. et al. Gliadin activates HLA class I-restricted CD8+ T cells in celiac disease intestinal mucosa and induces the enterocyte apoptosis. *Gastroenterology* **134**, 1017–1027 (2008).
- Pisapia, L. et al. HLA-DQ2.5 genes associated with celiac disease risk are preferentially expressed with respect to non-predisposing HLA genes: Implication for anti-gluten T cell response. *J. Autoimmun.* **70**, 63–72 (2016).
- Pisapia, L. et al. Differential expression of predisposing HLA-DQ2.5 alleles in DR5/DR7 celiac disease patients affects the pathological immune response to gluten. *Scientific Rep.* **10**, 17227 (2020).

46. Farina, F. et al. HLA-DQA1 and HLA-DQB1 alleles, conferring susceptibility to celiac disease and type 1 diabetes, are more expressed than non-predisposing alleles and are coordinately regulated. *Cells* **8**, 751 (2019).
47. Pisapia, L. et al. The HLA-DRB1 risk alleles for multiple sclerosis are differentially expressed in blood cells of patients from Southern Italy. *Int. J. Immunogenet.* **46**, 479–484 (2019).
48. Cavalli, G. et al. MHC class II super-enhancer increases surface expression of HLA-DR and HLA-DQ and affects cytokine production in autoimmune vitiligo. *Proc. Natl Acad. Sci. USA* **113**, 1363–1368 (2016).
49. Hayashi, M. et al. Autoimmune vitiligo is associated with gain-of-function by a transcriptional regulator that elevates expression of HLA-A * 02: 01 in vivo. *Proc. Natl Acad. Sci. USA* **13**, 1357–1362 (2016).
50. Raj, P. et al. Regulatory polymorphisms modulate the expression of HLA class II molecules and promote autoimmunity. *eLife* **5**, 1–52 (2016).
51. Gianfrani, C. et al. Expression level of risk genes of MHC class II is a susceptibility factor for autoimmunity: new insights. *J. Autoimmun.* **89**, 1–10 (2018).
52. Majumder, P., Gomez, J. A. & Boss, J. M. The human major histocompatibility complex class II HLA-DRB1 and HLA-DQA1 genes are separated by a CTCF-binding enhancer-blocking element. *J. Biol. Chem.* **281**, 18435–18443 (2006).
53. Lam, T. H. et al. Unique allelic eQTL clusters in human MHC haplotypes. *G3* **7**, 2595–2604 (2017).
54. Margaritte-Jeannin, P. et al. 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).
55. Lebwohl, B., Sanders, D. S. & Green, P. H. R. Coeliac disease. *Lancet* **391**, 70–81 (2018).
56. Caio, G. et al. Celiac disease: a comprehensive current review. *BMC Med.* **17**, 142 (2019).
57. Poddighe, D., Rebuffi, C., De Silvestri, A. & Capittini, C. Carrier frequency of HLA-DQB1*02 allele in patients affected with celiac disease: a systematic review assessing the potential rationale of a targeted allelic genotyping as a first-line screening. *World J. Gastroenterol.* **26**, 1365–1381 (2020).
58. Capittini, C. et al. Relevance of HLA-DQB1*02 allele in the genetic predisposition of children with celiac disease: additional cues from a meta-analysis. *Medicina* **55**, 190 (2019).
59. De Silvestri, A. et al. HLA-DQ genetics in children with celiac disease: a meta-analysis suggesting a two-step genetic screening procedure starting with HLA-DQ β chains. *Pediatr. Res.* **83**, 564–572 (2018).