The role of HLA genotype in determining the phenotype of coeliac disease

PhD Thesis

Dr. Judit Bajor

Doctoral School of Pharmacological and Pharmaceutical Sciences

Leader of the Doctoral School: Erika Pintér M.D., Ph.D., D.Sc.

Program Leader: Péter Hegyi M.D., Ph.D., D.Sc.

Supervisor: Áron Vincze M.D., Ph.D.

University of Pécs, Medical School

First Department of Medicine

2020

1. INTRODUCTION

Coeliac disease (gluten-sensitive enteropathy, non-tropical sprue) is a systemic autoimmune disorder which affects genetically susceptible individuals and develops upon gluten intake in approximately 1% of the Hungarian population and worldwide. The immune response is triggered by the α-gliadin fraction of gluten in cereal grains (wheat, barley and rye) in the small intestine, leading to mucosal damage and the destruction of villous architecture. The immune response is not restricted to the small intestine but affects the whole body, therefore several extraintestinal symptoms can develop. Genetic predisposition plays an essential role in the disease pathogenesis: the prerequisite of lifelong gluten intolerance is the presence of HLA-DQ2 or HLA-DQ8. Gliadin peptides are recognized by T-cells only if they are presented by antigen-presenting cells expressing DQ2 or DQ8 heterodimers on the cellular surface. Thus carrying HLA-DQ2 or HLA-DQ8 haplotypes is technically a must to develop coeliac disease. 90-95% of patients carry HLA-DQ2 and the remaining 5-10% carry HLA-DQ8. DQ2 heterodimers have two forms: DQ2.5 (DQA1*0501/B1*0201) is associated with a high risk of coeliac disease, whereas DQ2.2 (DQA1*0201/B1*0202) imposes low risk only.

Coeliac disease manifests with diverse clinical picture. In spite of the high prevalence, the majority of patients remain unrecognized for a long time, because the classical malabsorptive symptoms can be identified only in the minority of the patients. More than 50% of the adult cases have no typical gastrointestinal symptoms. The clinical phenotype is classified according to the Oslo definitions: we differentiate classical and non-classical cases. The classical phenotype is associated with characteristic malabsorptive signs and symptoms: diarrhoea, weight loss, low albumin level, vitamin and trace element deficiencies, and in paediatric cases, failure to thrive. The non-classical phenotype lacks malabsorptive signs and symptoms; instead, we observe atypical abdominal complaints (e.g. bloating, pain) or extraintestinal signs and symptoms (e.g. iron deficiency anaemia, abnormal liver function tests, dermatitis herpetiformis, metabolic bone disease, type I diabetes mellitus, infertility or neurological symptoms).

Due to the genetic predisposition, first degree relatives of coeliac patients are at the highest risk of the disease. Trisomies (Down, Turner and Williams syndromes) are more frequently associated with coeliac disease. Another typical comorbid condition is IgA deficiency. Other comorbidities include immune-mediated diseases, dermatitis herpetiformis as a skin manifestation, dental enamel hypoplasia (DED) and recurrent oral

aphthous lesions (RAS) as oral manifestations. A gluten-dependent disease is the so-called gluten ataxia, which results from progressive loss of Purkinje-cells. The shared genetic background provides a basis for the co-occurrence of coealiac disease and other autoimmune conditions. These disorders often include type I diabetes mellitus in childhood and autoimmune thyroiditis in adulthood. Sjögren syndrome, SLE, Addison's disease, rheumatoid arthritis, autoimmune liver diseases or myastaenia gravis develop less often as well as inflammatory bowel disease. The refractory coeliac disease is a rare but severe complication of coeliac disease, in type 2 form, the majority of the patients develop ulcerative jejunitis, followed by the enteropathy-associated T-cell lymphoma (EATL), which occurs in coeliac patients almost exclusively.

2. HLA

2.1. HLA molecules in the pathogenesis of coeliac disease

HLA molecules predisposing to coeliac disease are HLA-DQ2.5, HLA-DQ2.2 and HLA-DQ8, their importance is highlighted by their key role in disease pathogenesis. Gluten peptide-binding properties of HLA and stabile complex formation are essential contributors to T-cell response against gluten and the development of coeliac disease. Gluten-triggered T-cell stimulation depends on the quality and quantity of the DQ2 molecules. DQ2.5 possesses the ability to form a stable bond with the proline-rich regions of gluten; therefore, it presents a large number of gluten peptides, resistant to degradation. DQ8 present the peptides less effectively, this complex is less resistant. Finally, DQ2.2 has only limited peptide-binding properties, the complex is less stable, T-lymphocyte activation is less effective. If one chromosome has DQ2.2 while the other has DQ7.5, DQ2.5 molecule (DQ2 trans) can be synthesized. Therefore, DQ2.2 is the risk factor of coeliac disease only if it co-occurs with DQ7.5. Besides the quality of the heterodimers on the surface of the antigen-presenting cells, their quantity matters as well: the more DQ-gluten complexes are formed, the stronger the T-cell response and the higher the risk of coeliac disease. In HLA-DQ2 homozygotes, a large number of DQ2 heterodimers is on the surface of the antigen-presenting cells, contrasting the lower number in heterozygotes: consequently, the risk of coeliac disease is the highest in homozygotes (Table 1).

Table 1: HLA molecules, alleles and haplotypes which are risk factors of coeliac disease

HLA molecule	HLA allele	HLA haplotype	Risk of coeliac disease
DQ2.5 cis	DQA1*0501-DQB1*0201	DR3-DQ2	high
DQ2.5 trans	DQA1*0505-DQB1*0301 DQA1*0201-DQB1*0202	DR5-DQ7 DR7-DQ2	high
DQ2.2	DQA1*0201-DQB1*0202	DR7-DQ2	low
DQ7.5	DQA1*0505-DQB1*0301	DR5-DQ7	very low
DQ8	DQA1*03-DQB1*0302	DR4-DQ8	low

2.2. The role and importance of HLA-typing in the diagnostics of coeliac disease

HLA-typing is not routinely recommended according to the current international guidelines. Being positive for HLA-DQ2 or -DQ8 does not confirm the diagnosis since one-third of the population test positively for either of these. However, the high sensitivity and negative predictive value of testing allow to exclude the disease based on HLA-typing: in fact, lack of HLA-DQ2 or -DQ8 excludes the diagnosis of coeliac disease. A great advantage of HLA-typing is its independency of gluten intake. If a patient has already started a gluten-free diet and rejects a gluten challenge, HLA-DQ should be determined.

3. THE ROLE OF HLA-DQ2 GENE DOSE IN COELIAC DISEASE

3.1. Risk of coeliac disease by HLA haplotypes

Different HLA haplotypes impose different risks of developing coeliac disease; based on the literature, we distinguish high- and low-risk haplotypes. These are, however, not identically defined across studies. Most of the authors agree that HLA-DQ2.5 homozygotes are at high risk, heterozygotes are at moderate risk and HLA-DDQ2.2 positive individuals are at low risk of coeliac disease (Table 1).

3.2. Experimental data

Since the DQ2 molecule is the key in the pathogenesis, the number of HLA-DQB1*0201 copies is expected to have important consequences: all DQ molecule are DQ2 in homozygotes, whereas 4 different αβ-chain combinations can occur in heterozygotes and not all have the ability to present gliadin (only DQ2 and DQ8). Experimental data support this theory: the number of DQ2-peptide complexes significantly differ in homo- and heterozygotes. Antigen-presenting cells of HLA-DQ2.5 homozygotes can present more gluten peptides than those of HLA-DQ2.5 heterozygotes.

These data probably explain the fact that the risk of coeliac disease is 5-fold in HLA-DQ2.5 homozygotes compared to heterozygotes. The magnitude of gluten-triggered immune response is rather dependant on the strength of antigen presentation which depends on HLA-DQ2.5 gene dose: maximal T-cell activation and proinflammatory response are only possible in HLA-DQ2.5 homozygotes. The higher the number of high-affinity DQ2 molecules on the surface of antigen-presenting cells, the stronger the immune activation and the T-cell specific response, and the higher the IFN production. In this function, the importance of the α and β chains of DQ2.5 heterodimer might differ, the role of the β chain is rather decisive. Based on these in vitro experiments, we can assume that HLA-DQ2 gene dose influences the magnitude of the immune response, thereby affecting clinical phenotype – particularly disease severity – and the development of complications.

3.3. Clinical data

Based on the experimental data, a lot of observational studies, mostly retrospective and low-volume in nature, addressed the question of gene-dose effect. Until now, a consensus has not been reached: some publications support whereas others reject a significant gene-dose effect in coeliac disease.

3.4. The prognostic role of gene-dose effect

The clinical picture of coeliac disease is diverse, the scale of clinical presentation starts from fully asymptomatic cases and ends in cases with severe malabsorption. Factors determining the clinical picture are currently unknown nor is their prognostic role. It would be desirable to have markers predicting disease course and late complications (particularly malignant diseases). It is of similar importance to stratify the risk of patients who are known to be at risk of coeliac disease (first degree relatives, patients with autoimmune diseases, trisomies or IgA deficiency).

Taken together, the HLA status is a lifelong, gluten-independent marker, which has its firm role in the diagnostics of coeliac disease; however, the exploration of its prognostic utility is still awaiting clarification.

4. OBJECTIVES

We aimed to examine if HLA-DQ2 gene dose influences on the clinical phenotype of coeliac disease. If the answer is yes, we wonder which parameters are affected on which way. We analyzed the following variables:

i) clinical phenotype (classical vs non-classical),

ii) age at diagnosis,

iii) histology,

iv) anaemia,

v) metabolic bone disease,

vi) accompanying autoimmune diseases,

vii) development of complications (e.g. tumours).

To answer these questions, we initiated a meta-analysis and a multicentric retrospective cohort study.

5. THE STUDIES

5.1. Meta-analysis

5.1.1. Methods

To adapt the clinical question to meta-analysis, we used the PICO framework (P: population, I: intervention, C: control group, O: outcome). In our study, we compared multiple clinical outcomes (O) of patients carrying zero (I_1) or single dose of HLA-DQB1*02 allele (I_2) to those carrying a double dose of HLA-DQB1*02 allele (C) in coeliac disease (P). There has not been performed any meta-analysis in this topic yet. We report our work in accordance with the "Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) Statement".

5.1.1.1. Systematic literature search

We used 7 medical databases including MEDLINE (PubMed), Embase, CENTRAL, Web of Science, WHO Global Health Library, ClinicalTrials.gov and Scopus without any search restriction. We collected non-English-language papers as well.

5.1.1.2. Eligibility criteria, selection and data collection

Publications (including conference abstracts) were eligible if they analyzed and/or published clinical data about the investigated outcomes in relation to the HLA status (more specifically, HLA-DQB1*02 allele dose) determined with PCR in coeliac patients who were diagnosed according to the current guidelines. All records were combined in a reference manager software (EndNote X7.4, Clarivate Analytics, Philadelphia, PA, USA), which pool was subjected to duplicate removal and selection.

5.1.1.3 Data processing, analysis and risk of bias assessment

We divided the clinical phenotype at diagnosis into classical and non-classical as per the Oslo criteria; classical phenotype was defined based on the presence of

malabsorption, diarrhoea, weight loss, failure to thrive. Diarrhoea at diagnosis was further analyzed.

Age at diagnosis and age at first presentation of symptoms were analyzed separately. Diagnostic histology was described with two comparisons: on one hand, we compared cases with villous atrophy (Marsh 3) to those without villous atrophy (Marsh 1-2); on the other hand, we analyzed the severity of villous atrophy (Marsh 3c vs Marsh 3a-b). Regarding coeliac-specific serology, we assessed the titers of tissue transglutaminase (tTGA).

With respect to the frequency of accompanying diseases and disease complications, we assessed anaemia, metabolic bone disease, dermatitis herpetiformis and other autoimmune diseases (e.g. type I diabetes mellitus), dental complications (enamel defect, recurrent oral aphthous lesions) and malignant tumours (EATL, small bowel adenocarcinoma).

In the case of articles giving only HLA-DQ genotype, HLA-DQB1*02 gene dose was calculated, as follows: double-dose - HLA-DQ2.5 homozygotes (DQ2.5/DQ2.5) and compound heterozygotes (DQ2.5/DQ2.2); single-dose - HLA-DQ2.5 heterozygotes (DQ2.5/DQX) and HLA-DQ2 in trans (DQ2.2/DQ7); and zero-dose - HLA-DQ8/DQX and HLA-DQ2.2/DQX, where X represents any alleles except for DQ2.5.

All calculations were carried out with the Comprehensive Meta-Analysis software (Version 3, Biostat, Englewood, NJ) by the biostatistical group of Centre for Translational Medicine. For dichotomous variables (e.g. clinical phenotype and histology), we calculated odds ratio (OR) with 95% confidence interval (CI). For continuous variables (e.g. age), we calculated mean difference (MD) with CI. The level of significance was set to 0.05. During data aggregation in meta-analysis, we used the random-effect model with the DerSimonian-Laird estimation. In all comparisons, the reference group was the double dose of HLA-DQB1*02. To analyze age at diagnosis, we separated data of children and adults in two subgroups. Publication bias was assessed with funnel plot. Risk of bias was assessed by the topic-tailored items of the Newcastle-Ottawa Scale.

5.1.2. Results

5.1.2.1. Search and selection

Our search strategy yielded 6704 records (Pubmed [Medline]: 954, Embase: 2277, CENTRAL: 43, Web of Science: 925, WHO Global Health Library: 795, ClinicalTrials.gov: 6, Scopus: 1704), 59 of which were eligible for qualitative synthesis, 24 of which were eligible for quantitative synthesis. Results are summarized in Table 2.

Table 2: Results of the meta-analysis

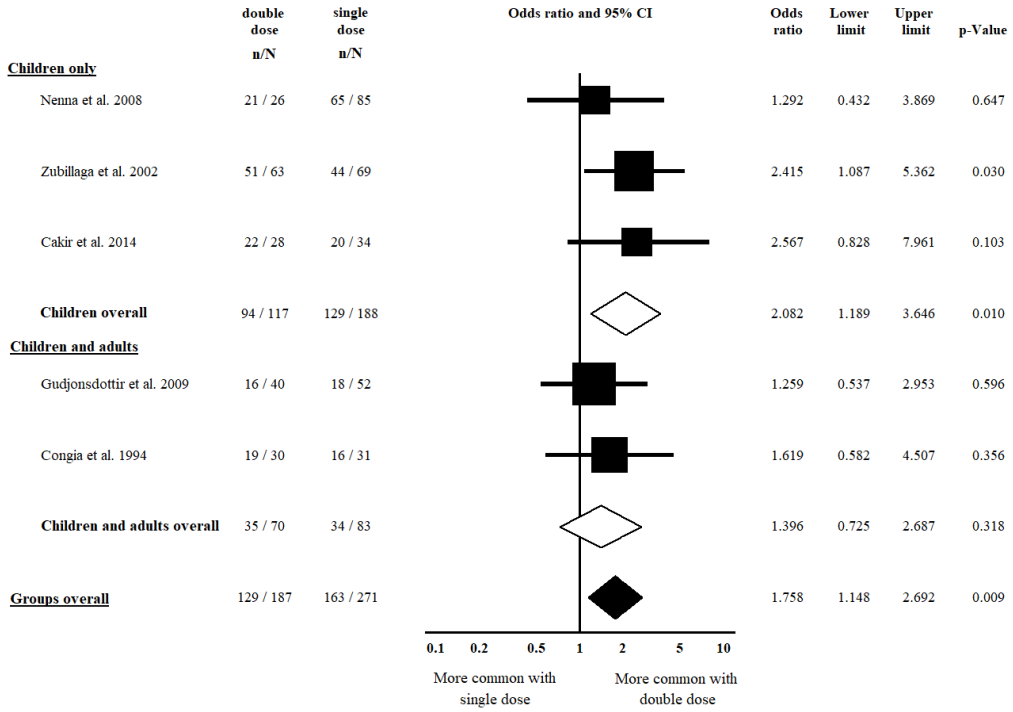
Outcomes, subgroups by age	**Double vs. single dose of HLA-DQB1*02**			**Double vs. zero dose of HLA-DQB1*02**		
	N^0 of patients	**OR (95% CI), *p*-value**	**Heterogeneity (I^2, chi^2)**	**N^0 of patients**	**OR (95% CI), *p*-value**	**Heterogeneity (I^2, chi^2)**
Atrophic vs. non-atrophic	722	0.991 (0.406-2.420), *p*=0.984	11.8%, *p*=0.338	430	2.626 (1.060-6.505), *p*=0.037*	21.3%, *p*=0.260
children	237	1.729 (0.319-9.370), *p*=0.525	71.6%, *p*=0.061*	159	1.757 (0.236-13096), *p*=0.583	0.0%, *p*=0.542
adults	379	0.537 (0.175-1.652), *p*=0.278	0.0%, *p*=0.682	200	2.534 (0.675-9.507), *p*=0.168	0.0%, *p*=0.945
Marsh 3c vs. Marsh 3a-b	862	0.870 (0.514-1.470), *p*=0.602	39.7%, *p*=0.127	418	0.822 (0.333-2.032), *p*=0.671	46.8%, *p*=0.068*
children	399	0.821 (0.401-1.681), *p*=0.590	0.0%, *p*=0.397	251	0.975 (0.296-3.208), *p*=0.967	65.2%, *p*=0.035*
adults	442	0.957 (0.420-2.184), *p*=0.918	82.4%, *p*=0.017*	147	0.753 (0.157-3.599), *p*=0.722	50.2%, *p*=0.134
Classical vs. non-classical	458	1.758 (1.148-2.692), *p*=0.009*	0.0%, *p*=0.744	221	1.701 (0.725-3.991), *p*=0.222	40.7%, *p*=0.168
children	305	2.082 (1.189-3.646), *p*=0.010*	0.0%, *p*=0.609	81	3.139 (1.142-8.630), *p*=0.027*	0.0%, *p*=0.747
Diarrhea vs. non-diarrhea	934	1.147 (0.863-1.523), *p*=0.345	0.0%, *p*=0.860	421	1.092 (0.655-1.818), *p*=0.337	0.0%, *p*=0.856
children	616	1.143 (0.819-1.593), *p*=0.432	0.0%, *p*=0.727	308	1.111 (0.569-2.170), *p*=0.758	0.0%, *p*=0.724
adults	318	1.158 (0.671-1.998), p=0.599	0.0%, *p*=1.000	113	1.065 (0.484-2.342), *p*=.0875	0.0%, *p*=1.000
Type 1 diabetes mellitus	840	0.914 (0.437-1.912), *p*=0.811	71.8%, *p*=0.006*	411	1.169 (0.410-3.331), *p*=0.770	89.8%, *p*<0.001*
children	766	0.597 (0.218-1.634), *p*=0.315	79.3%, *p*=0.008*	365	0.242 (0.045-1.312), *p*=0.100	79.4%, *p*=0.008*
		MD (95% CI), *p*-value	**Heterogeneity (I^2, chi^2)**		**MD (95% CI), *p*-value**	**Heterogeneity (I^2, chi^2)**
Age at diagnosis	512	-0.523 (-1.630 to 0.585), *p*=0.355	28.6%, *p*=0.231	147	-7.332 (-19.833 to 5.169), *p*=0.250	71.4%, *p*=0.015*
children	377	-0.303 (-1.156 to 0.551), *p*=0.487	5.5%, *p*=0.366	106	-2.026 (-5.824 to 1.771), *p*=0.296	48.7%, *p*=0.142
adults	133	-5.000 (-10.876 to 0.876), *p*=0.095	0.0%, *p*=1.000	41	-15.000 (-25.509 to -4.491), *p*=0.005*	0.0%, *p*=1.000

Asterisks indicate a p<0.05 for OR and MD, and a p<0.10 for heterogeneity tested with chi^2-test. CI: confidence interval; OR: odds ratio; MD: mean difference.

5.1.2.2. The association between gene dose and clinical phenotype

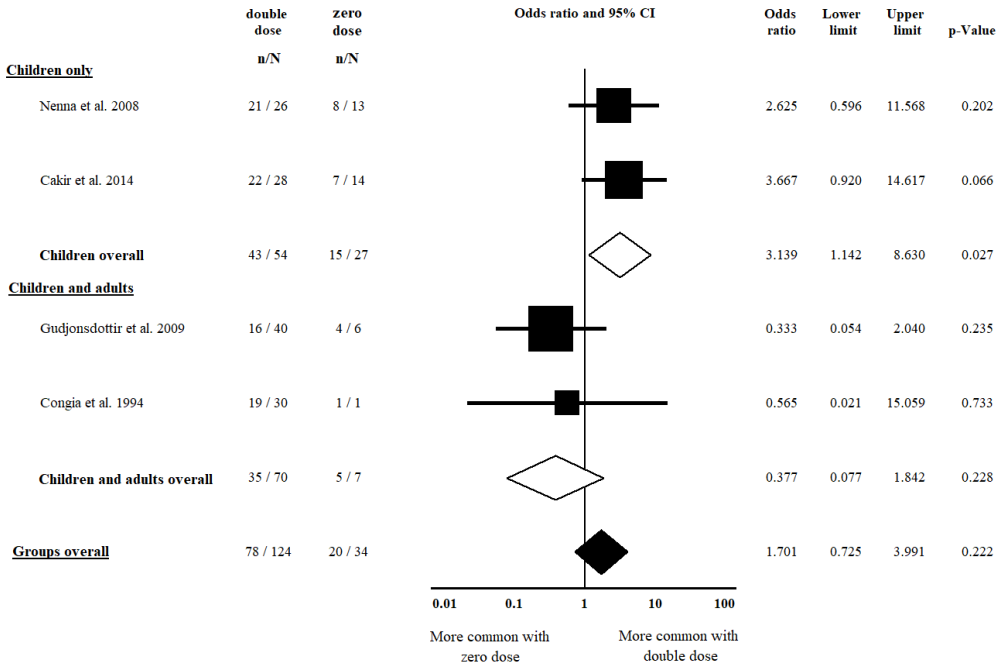
Patients with a double dose of HLA-DQB1*02 allele were more likely to develop classical clinical phenotype than those who carried a single dose of the allele (Figure 1; OR=1.758, CI:1.148-2.692, p=0.009) (I^2=0.0%, p=0.744, showing the homogeneity of the data).

*Figure 1: Classical phenotype with double vs simple HLA-DQB1*02 allele dose*



The difference was more prominent if we included only children in the analysis (Figure 1; OR=2.082, CI:1.189-3.646, p=0.010) (I^2=0.0%, p=0.609, showing the homogeneity of the data). In the comparison of double vs single dose, we observed a significant association in children exclusively (Figure 2; OR=3.139, CI:1.142-8.630, p=0.027) (I^2=0.0%, p=0.747, showing the homogeneity of the data). When we analyzed diarrhoea at diagnosis (and ignored other signs and symptoms of malabsorption), we failed to prove a significant association in the comparison of double vs single HLA-DQB1*02 allele dose, nor did for double vs zero dose or in the subgroup of children (Table 2).

*Figure 2: Classical phenotype with double vs zero HLA-DQB1*02 allele dose*



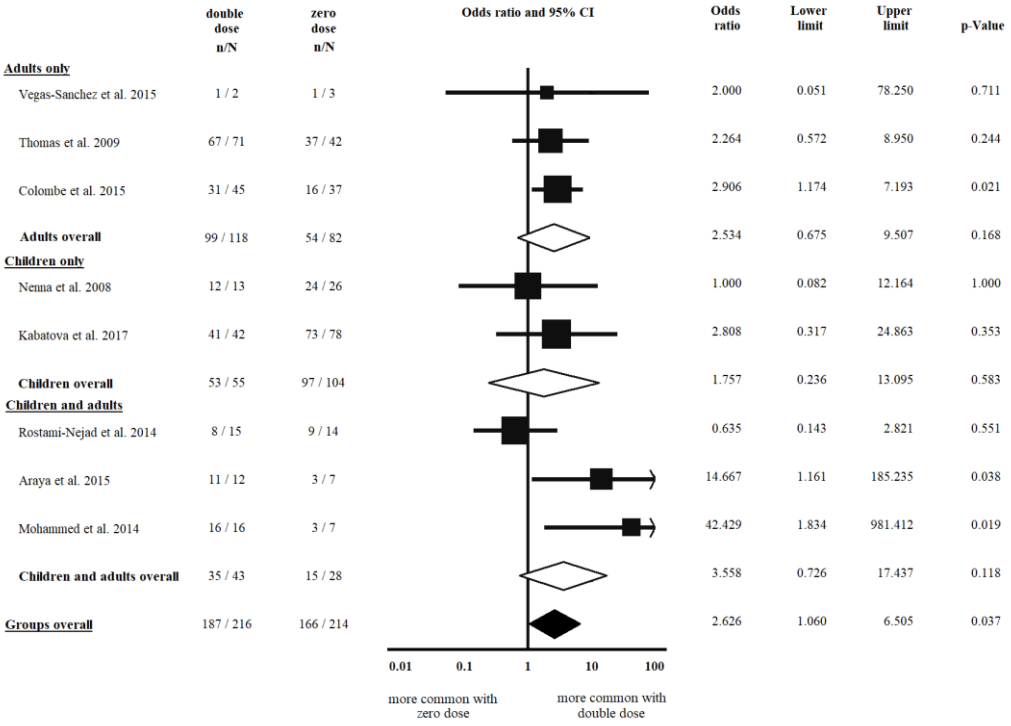
5.1.2.3. The association between gene dose and age at diagnosis

When we analyzed the association between HLA-DQB1*02 allele dose and age at diagnosis, we did not find a significant difference between groups; hence, a double dose of the allele was not accompanied with an earlier manifestation of the disease (MD: -0.523, CI:-1.630-+0.585, p=0.355; I^2=28.6% és MD:-7.332, CI:-19.833-+5.169, p=0.250; I^2=71.4%, showing considerable heterogeneity of the data). This proved to be true for children as well; however, statistical heterogeneity substantially reduced (Table 2).

5.1.2.4. The association between gene dose and small intestinal histology

Patients with a double dose of HLA-DQB1*02 were more likely to have villous atrophy at diagnosis than those with a zero dose of the allele (Figure 3; OR=2.626, CI: 1.060-6,505, p=0.037). However, comparing double dose to single dose we did not observe a significant difference (OR=0.991, CI: 0.406-2.420, p=0.984). These results were consistent in the subgroup of children (Table 2).

*Figure 3: Villous atrophy with double vs zero HLA-DQB1*02 allele dose*



The frequency of total villous atrophy (Marsh 3c) was independent on HLA gene dose (OR=0.870, CI: 0.514-1.470, p=0.602; I^2=39.7%, p=0.127 for double vs single and OR=0.822, CI: 0.333-2.032, p=0.671; I^2=46.8%, p=0.068 for double vs zero dose) (Table 2). Findings from subgroup analysis were in line with these.

5.1.2.5. The association between gene dose and complications of coeliac disease

We were lacking sufficient amount of high-quality data to carry out analysis on age at first presentation of symptoms, serology, accompanying diseases and complications (anaemia, metabolic bone disease, oral manifestations, refractory coeliac disease and tumours). We examined the frequency of type I diabetes mellitus, where gene dose effect did not manifest itself, which results remained constant with subgroup analyses.

5.1.2.6. Conclusions from the findings

Carrying a double dose of HLA-DQB1*02 allele predisposes to developing classical clinical phenotype and villous atrophy at diagnosis. Based on these findings, patients with high-risk HLA genotype (homozygotes) might benefit from a closer follow-up and more frequent check-ups because both villous atrophy and classical phenotype predispose to a more severe disease course where the risk of complications is increased. However, we failed to prove a significant gene dose effect regarding age at diagnosis, diarrhoea at diagnosis and the

development of type I diabetes mellitus. Further high-volume, prospective cohort studies are required to clarify the exact role of gene dose in the course of coeliac disease.

5.2. Cohort study

5.2.1. Methods

We collected data from three academic hospitals (Division of Gastroenterology, First Department of Medicine, University of Pécs, Pécs; Second Department of Internal Medicine, Semmelweis University, Budapest; Second Department of Internal Medicine, University of Debrecen, Debrecen). Scientific and Research Ethics Committee of the Medical Research Council has granted ethical approval of this research project (45098-2/2016/EKU).

To phrase the clinical questions, we used the PICO framework again. We compared multiple clinical outcomes (O) of patients carrying a low-risk zero dose of HLA-DQB1*02 allele (I_1), those carrying a moderate-risk single dose of HLA-DQB1*02 allele (I_2) to those carrying a high-risk double dose of HLA-DQB1*02 allele (C) in coeliac disease (P). To be eligible, a patient had to be subjected to HLA-typing and had to have the required clinical data available stored in medical files.

HLA-DQB1*02 allele dose was determined based on HLA-typing (Table 3). The HLA risk categories corresponded to HLA-DQ2 gene dose, as follows: (i) high-risk HLA-DQ2.5 homozygotes (DQ2.5/DQ2.5) and compound heterozygotes (DQ2.5/DQ2.2) with a double dose of DQB1*02 alleles; (ii) intermediate-risk HLA-DQ2.5 heterozygotes (DQ2.5/DQX) and HLA-DQ2 in trans (DQ2.2/DQ7) with a single dose of DQB1*02 allele; and (iii) low-risk HLA groups (HLA-DQ8/DQX, HLA-DQ2.2/DQX, X corresponded to any alleles except for DQ2.5) with zero doses of DQB1*02 allele.

The primary outcome was the clinical phenotype, which was defined similarly to our meta-analysis, as per the Oslo criteria. Small intestinal histology was assessed according to the Corazza-Villanacci classification. tTGA-IgA was measured with ELISA. If the titer increased at least 10-fold of the normal, we considered it high positive, while below that level but above the cut-off provided by the manufacturer, we considered it low positive. Data about accompanying autoimmune diseases, dermatitis herpetiformis and malignant tumours were extracted from electronic medical files.

5.2.1.1. Data processing and analysis

Pearson's Chi^2-test was used to test the association between HLA-risk and categorical variables, while one-way ANOVA was used to compare age across groups.

The level of significance was set to 0.05. The analysis was carried out with IBM SPSS Statistics v 20.0 software (IBM's Corporate, New York, USA).

5.2.2. Results

5.2.2.1. Characteristics of the patients included

A total of 727 patients were treated with coeliac disease in the three academic hospitals between Nov 1997 and May 2016, 105 of which (14.4%) were eligible for inclusion in the study (since HLA-typing is not mandatory to set up the diagnosis of coeliac disease, it is not routinely performed in all cases). Regarding HLA-risk based on HLA-typing, 35.3%, 52.3% and 12.3% of patients proved to be at high, moderate and low risk. We observed female predominance (73 females, 32 males) without a significant association between gene-dose and sex. Table 3 shows the clinical phenotype of patients by HLA gene dose.

*Table 3. The association between B1*02 allele dose and clinical phenotype*

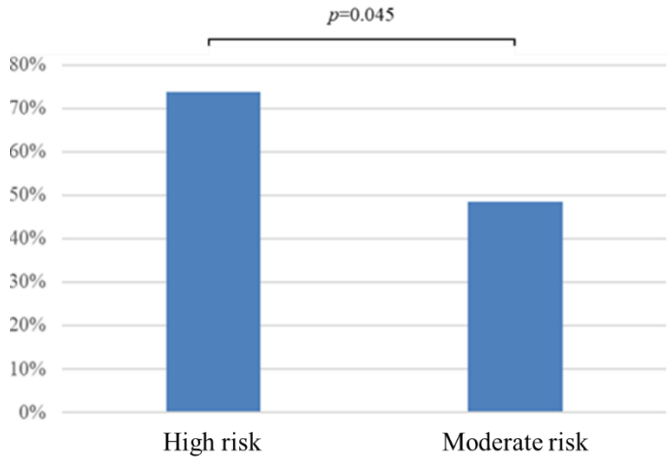
locus B1 PCR1	locus B1 PCR2	B1*02 allele dose	HLA genotype	Classical clinical phenotype (n)	Non-classical clinical phenotype (n)
High risk				**15**	**22**
B1*0201	B1*0201	double	DQ2.5/DQ2.5	8	16
B1*0201	B1*0202	double	DQ2.5/DQ2.2	7	6
Moderate risk				**25**	**30**
B1*0201	B1*05	single	DQ2.5/DQ5	5	9
B1*0201	B1*06	single	DQ2.5/DQ6	3	7
B1*0201	B1*0301	single	DQ2.5/DQ7	9	6
B1*0201	B1*0302	single	DQ2.5/DQ8	2	4
B1*0201	B1*0303	single	DQ2.5/DQ9	1	1
B1*0202	B1*0301	single	DQ2.2/DQ7	5	3
Low risk				**5**	**8**
B1*0202	B1*0202	zero	DQ2.2/DQ2.2	0	0
B1*0202	B1*04	zero	DQ2.2/DQ4	0	1
B1*0202	B1*06	zero	DQ2.2/DQ5	1	0
B1*0202	B1*0302	zero	DQ2.2/DQ8	2	1
B1*0302	B1*0302	zero	DQ8/DQ8	0	1
B1*0302	B1*04	zero	DQ8/DQ4	0	1
B1*0302	B1*06	zero	DQ8/DQ6	0	1
B1*0302	B1*0301	zero	DQ8/DQ7	2	3

5.2.2.2. The association of gene dose with age at diagnosis, clinical phenotype, serology and small intestinal histology

Age at diagnosis was 31.2 years (SD: 15,747 years, range: 0,5-78 years). Age at diagnosis was not different across HLA-risk groups (p=0.549). 42.9% and 57.1% of

patients had classical and non-classical clinical phenotype at diagnosis, without significant difference across HLA-risk groups (p=0.846). Regarding coeliac-specific serology, 70 patients had available tTG testing and 35 had EMA testing at diagnosis, 9 of the 70 patients with tTG were seronegative (3 cases with a double dose, 2 cases with a single dose and 4 cases with zero doses of HLA-DQB1*02 allele). Patients in the high-risk group carrying a double dose of HLA-DQB1*02 allele had significantly more often high tTG titers at least 10-fold of the normal at diagnosis. (Figure 4, p=0.045). However, we failed to observe a significant gene-dose effect regarding the severity of intestinal histological alterations (p=0.318).

Figure 4: Proportion of patients with high tTG titer at diagnosis in high- and moderate- HLA-risk groups



5.2.2.3. The association of gene dose with anaemia and metabolic bone disease at diagnosis

44.0% of patients had anaemia at diagnosis, which seemed not to be affected by HLA-DQB1*02 allele dose (p=0.611). DEXA was available in 62 patients, 61,3% of which had metabolic bone disease (osteopaenia or osteoporosis) independently of gene dose (p=0.374).

5.2.2.4. The association of gene dose with autoimmune diseases

11.4% of the patients had dermatitis herpetiformis, these cases were in the high- and moderate-risk groups exclusively. Of note, probably due to the low case number, we failed to prove a significant gene-dose effect regarding the frequency of dermatitis herpetiformis (p=0.381).

26.7% of patients had an accompanying autoimmune disease, with prominent female predominance (3:1). The most common disorder was autoimmune thyroiditis. There were 7 cases of inflammatory bowel disease, 2 cases of alopecia areata and sporadic cases of rheumatoid arthritis, myasthaenia gravis, lichen ruber planus, sarcoidosis,

psoriasis, sacroileitis and autoimmune liver disease, independently of HLA status (p=0.837). Refractory coeliac disease did not occur in our study population. Malignant tumours were diagnosed in 3 cases, all were in the moderate- and high-risk HLA groups but we did not perform further analysis due to the low case number.

5.2.3. Conclusions from the findings

We found significant associations between HLA-DQB1*02 allele dose and tTG titers at diagnosis but failed to confirm gene-dose effect regarding the other clinical parameters.

6. DISCUSSION

In vitro studies suggest that HLA-DQB1*02 allele dose exerts a significant effect on the clinical phenotype of coeliac disease. This question was addressed by several authors but, until now, no consensus has been achieved: findings from some studies support while others reject a gene-dose effect. In this work, we aimed to examine whether HLA-DQ2 gene dose influences the clinical phenotype of coeliac disease and if yes, which parameters are affected on what way. The findings of both our meta-analysis and the cohort study suggest the role of a HLA-DQB1*02 gene dose effect. Our data confirm the observation that in homozygotes (with a double allele dose), the inflammatory response is more intense (supported by the findings of the meta-analysis on classical clinical phenotype, malabsorptive signs and symptoms and villous atrophy, and by the findings of the cohort study on tTGA titer). Since the HLA status influences the clinical course of coeliac disease, it might be a potential candidate as a prognostic marker.

The importance of HLA-based risk stratification roots in the prediction of the development of accompanying autoimmune diseases (e.g. type I. diabetes mellitus) and the malignant complications of coeliac disease because these are the two most important determinants of morbidity and mortality. Among the autoimmune diseases, the co-occurrence of coeliac disease and type I diabetes mellitus is the most investigated. Both are HLA-associated diseases; in addition to genetic predisposition, immune dysregulation and environmental factors (e.g. gluten, viral pathogens, microbiome) are important contributors to the pathogenesis. It has long been known that children with type I diabetes mellitus are at an increased risk of coeliac disease, which results from the fact that a high proportion of diabetic patients has coeliac-related HLA genotypes (HLA-DQ2/DQ8), often leading to the development of coeliac disease years after the diagnosis of type I diabetes mellitus (within 5 years in average). Researchers think that the co-occurrence of

the diseases develops only in HLA-DQ2/DQ8 heterozygotes; moreover, others think that even HLA-DQ2 homozygosity is a risk factor.

Recent, large-volume prospective studies support the role of HLA-based risk stratification regarding the selection of patients who are at high risk of autoimmunity. In the USA, an analysis of more than 10000 subjects showed that HLA-DQ2 homozygotes are at higher risk of being EMA positive (OR: 3.94). In the population of the TEDDY Group, 6403 children were followed-up for 60 months and DR3-DQ2 positive children, especially the homozygotes, were at higher risk of developing coeliac disease at an early age. The CELIPREV Study concluded similarly, where the only significant predictor of the development of coeliac autoimmunity was HLA-DQ2 heterozygosity. Autoimmune diseases were not more common in another study but lymphatic complications occurred more often in coeliac patients carrying a double dose of HLA-DQ2.

Most importantly, HLA status should be highlighted in the prediction of the most feared complications, malignant tumours. Although evidence is limited, it supports that patients having a double HLA-DQB1*02 allele dose (homozygotes) might be exposed to a higher risk of malignant tumours. Al-Toma et al. found an association between HLA-DQ2.5 homozygosity and the most severe complications of coeliac disease, type II refractory coeliac disease (RCD2) and EATL. The observations of Biagi et al. confirm this association: in their study, severe complications including RCD1, RCD2, EATL, and small bowel adenocarcinoma developed more often in HLA-DQ2 homozygotes. Unfortunately, due to the limited number of cases, neither our meta-analysis nor our cohort study gave an answer to this question.

The HLA-based risk stratification in symptomless patiens raises ethical considerations. It is important to highlight that the predictive models are not suitable for foretelling that a patient will unavoidably develop a condition or a complication, these models are rather to operate with probability and offer an opportunity for prevention and early intervention. The question about whom and when to screen for has remained open: testing is recommended only for coeliac patients at diagnosis or for their family members as soon as possible. HLA typing might be beneficial to be performed prior to the first gluten intake or in infancy when the antibody-based diagnosis is uncertain. There is an urgent need for the introduction of an easily accessible screening strategy for first degree relatives of coeliac patients (women and HLA-DQ2 homozygotes are at the highest risk). A few authors recommend that serological testing should be done in every 2-5 years in high-risk individuals but HLA-typing makes repeated testing and follow-up unnecessary

in many cases. Early recognition of coeliac disease is particularly important in homozygotes to avoid failure to thrive due to severe malabsorption.

Based on our results, we agree with the statement saying that the high-risk individuals might benefit from HLA-based risk stratification: with early recognition of coeliac disease, early introduction of a gluten-free diet and with closer monitoring and follow-up one could more effectively prevent or recognize complications (particularly autoimmune and malignant diseases). This "personalized" risk stratification might inspire high-risk patients to adhere to the gluten-free diet more strictly because good dietary adherence can be scarce on the long-term, circa 17-80%. Of note, gluten was detected in the stool of patients who reported excellent dietary adherence, which finding confirms the need for better awareness and the introduction of ancillary therapies, especially in high-risk individuals. The HLA status might have therapeutic consequences in the future: blocking the binding sites of HLA can give the basis for targeted immunotherapy in high-risk individuals.

7. SUMMARY OF FINDINGS AND CONCLUSIONS

- Our meta-analysis proved that carrying a double dose of HLA-DQB1*02 allele predisposes to developing classical clinical phenotype and intestinal villous atrophy at diagnosis.

- Our meta-analysis did not prove a significant association between HLA-DQB1*02 gene dose and age at diagnosis, diarrhoea at diagnosis, the severity of villous atrophy and the development of type I diabetes mellitus.

- Our retrospective cohort study confirmed a significant association between HLA-DQB1*02 gene dose and tTGA levels at diagnosis.

- Our retrospective cohort study did not confirm the association between HLA-DQB1*02 gene dose and other clinical parameters including clinical phenotype, age at diagnosis, histological severity, anaemia, metabolic bone disease and the prevalence of accompanying autoimmune diseases.

- The findings of our meta-analysis and the cohort study imply a significant gene dose effect in coeliac disease. The observations support the theory that the inflammatory response is escalated in homozygotes (with a double dose of the allele), which is responsible for the increased frequency of classical clinical phenotype, malabsorptive symptoms and villous atrophy at diagnosis observed in the meta-analysis, and the higher titers of antibodies in our study population.

- HLA remains unchanged during the lifetime and is independent of gluten intake, which features make it adequate for risk stratification of coeliac patients. The early identification of patients carrying high-risk alleles offers a great opportunity for setting up individual therapeutic and follow-up plans and preventive strategies: cases of severe complications might be avoided with a strict gluten-free diet and more frequent medical check-ups.

8. PUBLICATIONS AND PRESENTATIONS RELATED TO THE TOPIC OF THE PHD THESIS

Publications serving the basis for the PhD dissertation (IF:4.838)

1. Bajor J, Szakács Zs, Farkas N, Hegyi P, Illés A, Solymár M, Pétervári E, Balaskó M, Pár G, Sarlós P, Szűcs Á, Czimmer J, Szemes K, Huszár O, Varjú P, Vincze Á. Classical coeliac disease is more frequent with a double dose of HLA-DQB1*02: A systematic review with meta-analysis. PLoS ONE, 2019, 14(2): e0212329. **IF: 2.776 (Q1)**

2. Bajor J, Szakács Zs, Juhász M, Papp M, Kocsis D, Szegedi É, Földi I, Farkas N, Hegyi P, Vincze Á: HLA-DQ2 homozygosis increases tTGA leves at diagnosis but does not influence the clinical phenotype of celiac disease: a multicentre study. Int J Immunogenet, 2019, 46(2): 74-81. IF: 1.031 (Q3)

3. Bajor J, Szakács Zs, Vincze Á: Response to Letter to the Editor: Relevance of HLA-DQB1*02 allele in predisposing to Celiac Disease. Int J Immunogenet, 2019, 46(4): 276-277. **IF: 1.031 (Q3)**

Other publications related to the topic of the PhD dissertation

1. Szakács Zs, Gede N, Gyöngyi Z, Solymár M, Csupor D, Erőss B, Vincze Á, Mikó A, Vasas A, Szapáry L, Dobszai D, Balikó V, Hágendorn R, Hegyi P, Bajor J: A call for research on the prognostic role of follow-up histology in celiac disease: a systematic review. Front Physiol, 2019 Nov 19; 10:1408. **IF: 3.201 (Q2)**

2. Szakács Zs, Csiszár B, Kenyeres P, Sarlós P, Erőss B, Hussain A, Nagy A, Kőszegi B, Veczák I, Farkas N, Bodis E, Márta K, Szentesi A, Tőkes-Fuzesi M, Berki T, Vincze Á, Toth K, Hegyi P, Bajor J: Haemorheological and haemostatic alterations in coeliac disease and inflammatory bowel disease in comparison with non-coeliac, non-IBD subjects (HERMES): a case–control study protocol. BMJ Open 2019;9:e026315. **IF: 2.376 (Q1)**

3. Bajor J: A malasszimiláció okai, diagnosztikája és kezelése. Orvostovábbképző Szemle, 2018, 25 (11): 17-18.

4. Szakács Zs, Mátrai P, Hegyi P, Szabó I, Vincze Á, Balaskó M, Mosdósi B, Sarlós P, Simon M, Márta K, Mikó A, Pécsi D, Demcsák A, Bajor J: Younger age at diagnosis predisposes to mucosal recovery in celiac disease on a gluten-free diet: a meta-analysis. Plos One, 2017, 12: (11) e0187526 **IF: 2.806 (Q1)**

5. Bajor J: Coeliakia – 2016, MBA, 69:321-327.

6. Bajor J: Gluténmentes diéta – az új divat? MBA, 2015, 68, 279-284.

7. Bajor J, Kocsis D, Juhász M: Nem-coeliakiás glutén-szenzitivitás – új glutén- asszociált entitás? CEUJGH, 2015, 1 (1): 14-18

8. Bajor J: Glutén szenzitivitás, az új glutén asszociált kórkép, MBA, 2014, 67, 93-96.

9. Bajor J: Irritábilis bél szindróma vagy glutén szenzitivitás, egy új entitás a glutén asszociált betegségek körében, HTSZ, 2014,19: 2-5.

10. Kocsis D, Bajor J, Papp M, Miheller P, Herszényi L, Tulassay Zs, Juhász M: Questionnaire survey of celiac disease and non-celiac gluten sensitivity in outpatients' cohorts of three Hungarian gastroenterology referral centres. Zeitschrift für Gastroenterologie 05/2015; 53(05). **IF:0.291**

11. Bajor J: Coeliakia. Orvosi Hetilap. 2007, 148(15): 709-711.

12. Lomb Z, Bajor J, Garamszegi M, Grexa E, Bogner B, Tóvári L, Beró T: Peumatosis cystoides intestinalis. Orvosi Hetilap, 2005, 146 (8): 369-374.

13. Bajor J, Lomb Z, Anga B, Beró T: Szülést követően manifesztálódó coeliakia. Orvosi Hetilap, 2003, 144 (52): 2665-2669.

Posters presented in international congresses and related to the topic of the PhD dissertation

1. Szakács Zs, Nagy M, Csiszár B, Kenyeres P, Sarlós P, Erőss B, Hussain A, Nagy Á, Kőszegi B, Veczák I, Farkas N, Bódis E, Márta K, Szentesi A, Tőkés-Füzesi M, Berki T, Vincze Á, Tóth K, Hegyi P, Bajor J: Hemorheological and

hemostatic alterations in celiac disease and inflammatory bowel disease in comparison with non-celiac, non-IBD subjects (HERMES): Preliminary results of a case-control study. UEGW, 2019, október 19-23, Barcelona, Spanyolország.
2. Bajor J, Szakács Zs, Farkas N, Hegyi P, Vincze Á: Non-classical clinical presentation of celiac disease is more prevalent in the 21st century. A single center experience. UEGW, 2018, október 20-24. Bécs, Ausztria.

3. Korponay-Szabó IR, Gyimesi J, Barkaszi-Szabó Z, Riznik P, Balogh M, Tokodi I, Pollák É, Guthy I, Horváth Á, Werkstetter K, Koletzko S, Bódi P, Kis I, Bajor J, Csoszanszki N, Nemes É, Kadenczki O, B. Kovács J, Dolinsek J, Dolinsek J: Utilisation of diagnostic tools for coeliac disease in children in 2016-2017 - A nation-wide survey. UEGW, 2018, október 20-24. Bécs, Ausztria.

4. Bajor J, Zsolt Szakács Zs, Farkas N, Juhász M, Papp M, Kocsis D, Földi I, Szegedi É, Hegyi P, Vincze Á: HLA-DQ2 homozygosis increases tTGA levels at diagnosis but does not influence the clinical phenotype of celiac disease: a multicenter study. Tampere Coeliac Disease symposium, 2018, szept. 14-15. Tampere, Finland.

5. Szakács Zs, Csiszár B, Kenyeres P, Sarlós P, Erőss B, Hussain A, Nagy Á, Kőszegi B, Veczák I, Farkas N, Bódis E, Márta K, Szentesi A, Tőkés-Füzesi M, Berki T, Vincze Á, Tóth K, Hegyi P, Bajor J: Hemorheological and hemostatic alterations in celiac disease and inflammatory bowel disease in comparison with non-celiac non-IBD subjects (HERMES): preliminary results of a case-control study. Tampere Coeliac Disease symposium, 2018, szept. 14-15. Tampere, Finland.

6. Szakács Zs, Gede N, Gyöngyi Z, Solymár M, Csupor D, Erőss B, Vincze Á, Mikó A, Vasas A, Szapáry L, Dobszai D, Balikó V, Hágendorn R, Hegyi P, Bajor J: Follow-up histology has a limited prognostic role in coeliac disease: a systematic review and meta-analysis. Tampere Coeliac Disease symposium, 2018, szept. 14-15. Tampere, Finland.

7. Szakács Zs, Bajor J: Follow-up histology – what does it tell us in celiac disease - a meta-analysis and systematic review. 11th European Workshop on Immune Mediated Inflammatory Diseases, 2017, December 14-16. Utrecht, Netherlands.

8. Bajor J, Beró T: Rare co-existence of coeliac disease with Turner syndrome, myastenia gravis and chronic myeloid leukaemia, 164. Falk Symposium, Budapest, 2008. 05.02-03.

9. Bajor J, Beró T: Celiac disease associated rare complications. Report of two cases of liver cirrhosis and a case of lichen oris. 160. Falk Symposium. Portoroz, Szlovénia. 2007. 06.14-16.

10. Bajor J, Gódi Sz, Beró T: Coeliac disease during pregnancy and puerperium. 4. Central European Gastroenterology Meeting (CERUGEM), Visegrád, 2006.06.29-07.02.

11. Beró T, Bajor J, Gódi Sz, Anga B: Result of our coeliac health survey. 4. Central European Gastroenterology Meeting (CERUGEM), Visegrád, 2006.06.29-07.02.

Book chapters related to the topic of the PhD dissertation

1. Bajor J: Coeliakia (Klinikai belgyógyászat, Szerk: Tulassay Zs, 2017), 323-330.

2. Bajor J: A coeliakia és a háziorvos (Gasztroenterológia a háziorvosi gyakorlatban. Szerk: Magyar A, Bajor J, 2019)

Scientific metrics

Total number of publications: 183
Total number of citations: 392
Number of independent citations: 340
Cumulative impact factor: 102,787
Hirsch index: 9

9. ACKNOWLEDGEMENTS

I would like to thank dr. Tamás Beró who trusted me with the care of coeliac patients and convincingly conveyed the spirit adventure in regard to research.

I would like to express my gratitude to my supervisor, dr. Áron Vincze, who was always ready to give useful advice and supportive comments. His great precision set an excellent example for me.

I thank prof. dr. Péter Hegyi, who helped me to work in a truly supporting team. Without his guidance and persistent encouraging, the meta-analysis would not have been possible. His unshakeable faith in my success helped me through the most difficult challenges.

I would like to express the deepest appreciation to PhD student dr. Zsolt Szakács, who was my true companion in all scientific works. His methodological knowledge and systematic thinking made him an excellent contributor in completion of the papers and this dissertation. He became an exemplary colleague as well as a good friend during work.

Most importantly, I would say thank to all relatives and friends for the endless patience, support and understanding spirit.