Immunogenetics of Type 1 diabetes and Celiac disease

Alshiekh, Shehab Abdulaziz

2020

Document Version:
Publisher's PDF, also known as Version of record

Received: 20 March 2017 | Revised: 2 May 2017 | Accepted: 12 May 2017

DOI: 10.1111/tan.13065

WILEY HLA Immune Response Genetics

ORIGINAL ARTICLE

Different *DRB1*03:01-DQB1*02:01* haplotypes confer different risk for celiac disease

S. Alshiekh[1,2] | L. P. Zhao[3] | Å. Lernmark[1] | D. E. Geraghty[4] | Å. T. Naluai[5] | D. Agardh[1]

[1]Department of Clinical Sciences, Lund University/CRC, Skåne University Hospital, Malmö, Sweden

[2]Department of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia

[3]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, Waltham

[4]Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, Waltham

[5]Department of Microbiology and Immunology, Institute of Biomedicine, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden

Correspondence
Daniel Agardh, Department of Clinical Sciences/Diabetes and Celiac Disease Unit, Lund University/Clinical Research Center (CRC), Jan Waldenströms gata 35; 60:11, Skåne University Hospital (SUS), SE-205 02, Malmö, Sweden.
Email: daniel.agardh@med.lu.se

Celiac disease is associated with the *HLA-DR3-DQA1*05:01-DQB1*02:01* and *DR4-DQA1*03:01-DQB1*03:02* haplotypes. In addition, there are currently over 40 non-HLA loci associated with celiac disease. This study extends previous analyses on different HLA haplotypes in celiac disease using next generation targeted sequencing.

Included were 143 patients with celiac disease and 135 non-celiac disease controls investigated at median 9.8 years (1.4-18.3 years). PCR-based amplification of HLA and sequencing with Illumina MiSeq technology were used for extended sequencing of the HLA class II haplotypes HLA-DRB1, DRB3, DRB4, DRB5, DQA1 and DQB1, respectively. Odds ratios were computed marginally for every allele and haplotype as the ratio of allelic frequency in patients and controls as ratio of exposure rates (RR), when comparing a null reference with equal exposure rates in cases and controls.

Among the extended HLA haplotypes, the strongest risk haplotype for celiac disease was shown for *DRB3*01:01:02* in linkage with *DQA1*05:01-DQB1*02:01* (RR = 6.34; *P*-value < .0001). In a subpopulation analysis, *DRB3*01:01:02-DQA1*05:01-DQB1*02:01* remained the most significant in patients with Scandinavian ethnicity (RR = 4.63; $P < .0001$) whereas *DRB1*07:01:01-DRB4*01:03:01-DQA1*02:01-DQB1*02:02:01* presented the highest risk of celiac disease among non-Scandinavians (RR = 7.94; $P = .011$).

The data also revealed 2 distinct celiac disease risk *DR3-DQA1*05:01-DQB*02:01* haplotypes distinguished by either the *DRB3*01:01:02* or *DRB3*02:02:01* alleles, indicating that different *DRB1*03:01-DQB1*02:01* haplotypes confer different risk for celiac disease. The associated risk of celiac disease for *DR3-DRB3*01:01:02-DQA1*05:01-DQB1*02:01* is predominant among patients of Scandinavian ethnicity.

KEYWORDS

celiac disease, genetic risk, HLA, massively parallel sequencing, next generation sequencing

1 | INTRODUCTION

Recent advance in immunogenetics opens the door to implement HLA typing in risk stratification of diseases in the clinic.[1] Development of human leukocyte antigen (HLA) genotyping by next generation targeted sequencing (NGTS) have led to greater allele resolution and less phase ambiguity at cheaper cost[2] in comparison to conventional methods such as Sanger sequencing or Sequence-Specific Oligonucleotides (SSO)/Sequence-Specific Primers (SSP).[3,4] In addition, NGTS has offered important insights in "single-molecule sequencing"[5] of growing number of polymorphic HLA alleles by typing all exons and introns through combining clonal (single-molecule) sequencing and

a high level of parallelism for a potential full understanding of regulation and expression of these genes.

Celiac disease is an autoimmune small bowel disorder strongly associated with the *HLA-DRB1*03:01:01 (DR3)-DQA1*05:01-DQB1*02:01* (DQ2.5-cis) and/or *DR4-DQA1*03:01-DQB1*03:02* haplotypes (DQ8).[6] Patients who do not carry either of these 2 haplotypes usually carry the *DQA1*02:01-DQB1*02:02* (DQ2.2) haplotype together with *DQA1*05:05-DQB1*03:01* haplotype (DQ7.5) on the opposite chromosome, also called DQ2.5-trans (*DQB1*02:02* from 1 chromosome and *DQA1*05:05* from the other chromosome). These patients are believed to be at equal risk as the DQ2.5-cis heterozygotes.[6]

Outside the HLA complex, genome-wide association studies (GWAS) have identified additional minor risk variants (single-nucleotide polymorphisms [SNP]) in over 40 non-HLA loci, however none of the identified SNPs being near the risk of that of the major known HLA high-risk haplotypes.[7,8] To date, the identified genetic variability is estimated to be responsible for approximately 54% of the heritability in celiac disease.[9,10] This suggests that additional genes, gene variants and/or environmental factors are involved in the development of celiac disease.

Although celiac disease risk is strongly influenced by combinations of known DR3 and DR4 HLA haplotypes in linkage disequilibrium with DQ,[11] further exploration of the DRB3, DRB4, DRB5 and DRB1 haplotypes in relation to DQ remains mainly unresolved.[12] Many previous studies have tried to assess the risk associated with different HLA genotypes using statistical methods based on case-control studies.[13,14] However, analyzing ultra-polymorphic sequences in celiac disease are yet to be performed.

The impact of NGTS on genetic analysis in celiac disease has 2 main aspects. First, NGTS is able to target extended sequencing of the HLA class II haplotypes DRB1, DRB3, DRB4, DRB5, DQA1 and DQB1. Secondly, NGTS enables HLA haplotypes analysis and molecular dissection of celiac disease associations or identification of HLA-linked causative variants.[15] The aim of the present study was to use NGTA to study how HLA-DRB1, DRB3, DRB4 and DRB5 affect the risk of celiac disease inheritability in relation to the DQA1 and DQB1 haplotypes.

TABLE 1 Proportion of patients with celiac disease (cases) compared to non-celiac disease controls. Grading of intestinal biopsies were used according to the Marsh-Oberhuber classification and a Marsh score >1 was compliant with biopsy-proven celiac disease[27]

	Cases[a] $n = 143$	Controls $n = 135$	*P*-value
Gender			
Female	0.671	0.578	.137
Male	0.329	0.422	
Age (year)			
0-	0.063	0.393	<.0001
5-	0.147	0.356	
10-	0.371	0.156	
15-18	0.42	0.096	
Marsh score			
M0-M1	0.175[b]	0.993	<.0001
M2-M3	0.825	0.007	
Ethnicity			
Scandinavian	0.951	0.593	<.0001
Non-Scandinavian	0.049	0.407	
IgA-tTGA >4 U/mL			
Negative	0.078	0.911	<.0001
Positive	0.825	0.089	
IgG-tTGA >4 U/mL			
Negative	0.092	0.933	<.0001
Positive	0.908	0.067	

[a] One-hundred and twenty children had biopsy-proven active celiac disease.

[b] Four with treated celiac e disease had M0-M1, 21 had potential celiac disease.

2 | METHODOLOGY

2.1 | Study population

This case-control study comprised of 278 patients (174 females and 104 males) recruited as part of the GENEX study.[16] Study participants were investigated with an upper endoscopy with serial intestinal biopsies taken from the bulb and duodenum at median age 9.8 years (1.4-18.3 years) between 2010 and 2012 at the Department of Pediatrics, Skåne university hospital situated in Malmö, Sweden. At day of intestinal biopsy, serum samples from all patients were assessed for both IgA and IgG autoantibodies against tissue transglutaminase (tTGA) using radioligand binding assays previously described.[17] Among the 143 patients selected as cases, 118 had untreated celiac disease, 4 treated celiac disease and 21 were persistently tTGA positive and classified as having potential celiac disease (Table 1). For the 135 patients selected as controls, celiac disease was ruled out by the findings of intestinal biopsy and/or serology (Table 1). Study groups were selected to be of Scandinavian ethnicity and non-Scandinavian ethnicity based on parents' place of birth. Parents gave their informed written consent and the local ethical committee approved the study.

2.2 | Genetic sequencing

The NGTS HLA typing approach used PCR-based amplification of HLA and sequencing with Illumina MiSeq technology.[18,19] The laboratory steps comprise consecutive PCR reactions with bar coding incorporated into the PCRs for individual sample tracking followed by application to the MiSeq system using the ScisGo v4.0 genotyping system according to the manufacturer's instructions (Scisco Genetics Inc., Seattle Waltham). Robust assays for each target loci of all class II loci including HLA-DR alleles were used providing a depth of genotyping extending to HLA-DRB3,

DRB4 and DRB5 including exons 1 through 4 for all DR alleles. The analytical tools to define haplotypes and genotypes were developed in collaboration with Scisco Genetics Inc. (Seattle, Waltham). To date, these tools have been tested with 100% accuracy on 2000 control samples genotyped with the NGS approach used in this study.[18,19]

2.3 | Statistical analysis

A frequency matching strategy, matching cases and controls by gender was applied. Allelic association analysis of multiallelic genes of HLA-DRB1, DRB3, DRB4, DRB5, DQA1 and DQB1, were performed. Alleles of DRB3, DRB4 and DRB5 at the DR locus do not jointly reside on the same chromosome. Effectively, these 3 genes are designated as a "single gene" DRB345 for the analytic purpose. For each gene, allelic frequencies were computed among cases, controls, and combined cases and controls, under the Hardy-Weinberg equilibrium. To test if the allele is significantly associated with celiac disease, an allele-specific score test was applied.[20] The score test is to evaluate the allele-specific score under the null hypothesis with no allelic associations, and is referred here as H-score, because it was developed specifically for testing haplotype-association. H-score has an asymptotic normal distribution, which is then used to compute *P*-value under the null hypothesis. To retain meaningful interpretation on allele-specific odds ratios, it is assumed that under the null hypothesis for a reference allele, its allelic frequency in cases should be comparable to that in controls, and thus their ratio equals 1. In comparison with this assumed reference allele, odds ratios were computed marginally for every allele as the ratio of allelic frequency in cases over that in the controls, that is, ratio of exposure rates (RR). RR of less than 1 implies a protective allele, and RR of greater than 1 is a risk allele.

Within the DR locus, alleles of DRB1 are in linkage disequilibrium with alleles of DRB345. Similarly, DQA1 and DQB1 are in high linkage disequilibrium with each other in the DQ locus. Further, DR and DQ loci are physically close to each other, and many of their alleles are highly associated with each other. Since phases of these individual genes are unresolved, statistically-driven haplotype analysis on DRB1-DRB345, DQA1-DQB1 or DR-DQ were performed. Without assigning haplotypes directly, the statistical approach is to numerate all possible haplotype configurations with empirically computed prior probabilities, and produces estimates of haplotype frequencies for all possible haplotypes, provided that expectation numbers of corresponding haplotypes are 10 or more copies in the combined cases and controls. For both allelic and haplotypic association analyses, the R function "haplo.cc" (http://cran.r-project.org/web/packages/haplo.stats/index.html) was used. Just as those statistics produced for allelic association analysis, haplotypic frequencies were computed among cases, controls and pooled case-control. Then, the function "haplo.cc" produces H-scores and *P*-values for all included haplotypes. Following the same principle of choosing the null reference haplotype, in other words, comparing haplotypic frequencies between cases and controls, the haplotype-specific RR was computed.

During the evaluation of allelic associations within individual genes, the bootstrap technique was used for random sampling with replacement to form a bootstrap sample to estimate allele-specific rate of ratios (allelic frequency) among cases over controls, which is equivalent to odds ratio given the assumption that 2 allelic frequencies are the same between cases and controls. By using 1000 bootstrap samples, we estimated associated standard errors and thus associated confidence intervals for individually estimated RRs.

3 | RESULTS

3.1 | Haplotypic associations of HLA-DR and DQ

The result from the complete haplotype analyses of the DR and DQ alleles is listed in Table S1 (Supporting information). *DRB1*03:01:01-DRB3*01:01:02-DQA1*05:01:01-DQB1*02:01:01*, *DRB1*04:01:01-DRB4*01:03:01-DQA1*01:03:01-DQB1*03:02:01*, *DRB1*04:04:01-DRB4*01:03:01-DQA1*03:01:01-DQB1*03:02:01* and *DRB1*11:01:01-DRB2*02:02:01-DQA1*05:05:01-DQB1*03:01:01* were all haplotypes associated with celiac disease (*P*-value < .01) (Table 2). In contrast, *DRB1*01:02:01-DRB3*01:01:02-DQA1*01:01:02-DQB1*05:01:01*, *DRB1*07:01:01-DRB4*01:01:01-DQA1*02:01-DQB1*02:02:01*, *DRB1*07:01:01-DRB4*01:03:01-DQA1*02:01-DQB1*02:02:01*, *DRB1*07:01:01-DRB4*01:03:01-DQA1*02:01-DQB1*03:03:02*, *DRB1*10:01:01-DQA1*01:05:01-DQB1*05:01:01*, *DRB1*11:01:01-DRB3*02:02:01-DQA1*05:05:01-DQB1*03:01:01*, *DRB1*13:01:01-DRB3*01:01:02-DQA1*01:03:01-DQB1*03:01:01*, *DRB1*15:02:01-DRB5*01:01:01-DQA1*01:02:01-DQB1*06:02:01* were protective for celiac disease (*P*-value <.05) (Table 2).

3.2 | Haplotype frequency stratified by ethnicity

The results from the stratified haplotype analyses of the DR and DQ alleles are given separately for Scandinavian (Table 3) and non-Scandinavian (Table 4) ethnicity, respectively. Ten patients carried a variant of the celiac risk haplotype DQ2.5 (*DRB1*03:01:01-DRB3*01:01:02-DQA1*05:01:01-DQB1*02:01:01*). Instead of *DRB3*01:01:02*, these patients carried *DRB3*02:02:01*. Seven of the 10 patients were of Scandinavian descent and out of these 7, 6 carried the *DRB3*01:01:02-DQA1*05:01-DQB*02:01*(DQ2.5) haplotype and 1 carried the *DRB4*01:03:01-DQA1*02:01-DQB1*02:02:01* (DQ2.2) risk haplotype on their opposing chromosome. Three patients were of non-Scandinavian descent and they carried either a *DQA1*02:01-DQB1*02:01*

TABLE 2 Frequency and ratio of exposure rates (RR) of extended HLA haplotypes in association with celiac disease in cases (n = 143) and controls (n = 135)

HLA-DR-DQ haplotype				Study population				
DRB1	**DRB345**	**DQA1**	**DQB1**	**Pool**	**Cases**	**Controls**	**RR**	***P*-value**
*DRB1*01:01:01*		*DQA1*01:01:01*	*DQB1*05:01:01*	0.032	0.017	0.048	0.345	.038
*DRB1*01:02:01*		*DQA1*01:01:02*	*DQB1*05:01:01*	0.013	0.003	0.022	0.136	.046
*DRB1*03:01:01*	*DRB3*01:01:02*	*DQA1*05:01:01*	*DQB1*02:01:01*	0.262	0.444	0.070	6.343	<.0001
*DRB1*03:01:01*	*DRB3*02:02:01*	*DQA1*05:01:01*	*DQB1*02:01:01*	0.035	0.035	0.033	1.061	.880
*DRB1*04:01:01*	*DRB4*01:03:01*	*DQA1*03:01:01*	*DQB1*03:02:01*	0.101	0.157	0.041	3.821	<.0001
*DRB1*04:01:01*	*DRB4*01:03:01*	*DQA1*03:03:01*	*DQB1*03:02:01*	0.011	0.010	0.011	0.979	.943
*DRB1*04:02:01*	*DRB4*01:03:01*	*DQA1*03:01:01*	*DQB1*03:02:01*	0.005	0.000	0.011	0.000	.166
*DRB1*04:04:01*	*DRB4*01:03:01*	*DQA1*03:01:01*	*DQB1*03:02:01*	0.065	0.094	0.033	2.848	.006
*DRB1*04:04:01*	*DRB4*01:03:02*	*DQA1*03:01:01*	*DQB1*03:02:01*	0.011	0.014	0.007	2.000	.451
*DRB1*04:07:01*	*DRB4*01:03:01*	*DQA1*03:03:01*	*DQB1*03:01:01*	0.005	0.003	0.007	0.429	.528
*DRB1*07:01:01*	*DRB4*01:01:01*	*DQA1*02:01*	*DQB1*02:02:01*	0.027	0.035	0.019	1.842	.225
*DRB1*07:01:01*	*DRB4*01:03:01*	*DQA1*02:01*	*DQB1*02:02:01*	0.025	0.028	0.022	1.273	.661
*DRB1*07:01:01*	*DRB4*01:03:01*	*DQA1*02:01*	*DQB1*03:03:02*	0.014	0.010	0.019	0.526	.423
*DRB1*08:01:03*		*DQA1*04:01:01*	*DQB1*04:02:01*	0.025	0.010	0.041	0.244	.021
*DRB1*08:03:02*		*DQA1*06:01:01*	*DQB1*03:01:01*	0.005	0.000	0.011	0.000	.073
*DRB1*10:01:01*		*DQA1*01:05:01*	*DQB1*05:01:01*	0.022	0.007	0.037	0.189	.013
*DRB1*11:01:01*	*DRB3*02:02:01*	*DQA1*05:05:01*	*DQB1*03:01:01*	0.038	0.010	0.067	0.149	.003
*DRB1*11:04:01*	*DRB3*02:02:01*	*DQA1*05:05:01*	*DQB1*03:01:01*	0.022	0.014	0.030	0.467	.200
*DRB1*12:01:01*	*DRB3*02:02:01*	*DQA1*05:05:01*	*DQB1*03:01:01*	0.009	0.003	0.015	0.200	.156
*DRB1*13:01:01*	*DRB3*01:01:02*	*DQA1*01:03:01*	*DQB1*06:03:01*	0.020	0.004	0.037	0.108	.005
*DRB1*13:01:01*	*DRB3*02:02:01*	*DQA1*01:03:01*	*DQB1*06:03:01*	0.021	0.010	0.033	0.303	.048
*DRB1*13:02:01*	*DRB3*03:01:01*	*DQA1*01:02:01*	*DQB1*06:04:01*	0.023	0.014	0.033	0.424	.156
*DRB1*13:03:01*	*DRB3*01:01:02*	*DQA1*05:05:01*	*DQB1*03:01:01*	0.011	0.000	0.022	0.000	.027
*DRB1*14:54:01*	*DRB3*02:02:01*	*DQA1*01:04:01*	*DQB1*05:03:01*	0.009	0.007	0.011	0.636	.606
*DRB1*15:01:01*	*DRB5*01:01:01*	*DQA1*01:02:01*	*DQB1*06:02:01*	0.083	0.035	0.133	0.263	<.0001
*DRB1*15:02:01*	*DRB5*01:02*	*DQA1*01:03:01*	*DQB1*06:01:01*	0.013	0.000	0.026	0.000	.005
*DRB1*16:01:01*	*DRB5*02:02*	*DQA1*01:02:02*	*DQB1*05:02:01*	0.007	0.000	0.015	0.000	.038

(DQ2.2) risk haplotype or a *DQA1*0505:DQB1*0301* (DQ7.5) risk haplotype on the opposing chromosome (Table S1). *DRB1*03:01:01-DRB3*01:01:02-DQA1*05:01:01-DQB1*02:01:01* showed the highest risk for developing celiac disease in relation with other haplotypes (*P*-value <.0001). Among the non-Scandinavians, *DRB1*07:01:01-DRB4*01:01:01-DQA1*02:01-DQB1*02:02:01* showed the highest risk of celiac disease in relation to the other 2 risk haplotypes, *DRB1*03:01:01 DRB3*02:02:01 DQA1*05:01:01 DQB1*02:01:01* and *DRB1*11:04:01-DRB3*02:02:01-DQA1*05:05:01-DQB1*03:01:01*.

4 | DISCUSSION

The use of NGTS adds new insight into HLA risk alleles in celiac disease and highlighted by the finding that the haplotype with the highest risk of developing celiac disease DR3-DQ2.5 (*DRB1*03:01:01-DQA1*05:01:01-DQB1*02:01:01*) is differentiated at the DRB3 locus. The most frequent of these haplotype in the Scandinavian population is *DRB3*01:01:02* in linkage with *DQA1*05:01-DQB1*02:01* (hereafter called "8.1AH" for the extended ancestral haplotype: HLA-A1, B8, *DRB3*01:01*, *DRB1*03:01*, *DQB1*02:01*) and the less frequent version of this haplotype carries a *DRB3*02:02:01* in linkage with *DQA1*05:01-DQB1*02:01* (hereafter called 18.2AH for the extended ancestral haplotype: HLA-A30, B18, *DRB3*02:02*, *DRB1*03:01*, *DQB1*02:01*). Even though both of these haplotypes are associated with celiac disease, no patient carried the 18.2AH alone. Celiac disease patients who carried the 18.2AH haplotype on 1 chromosome, all carried another celiac disease risk haplotype [i.e. either the *DRB3*01-DQA1*05:01-DQB1*02:01* (DQ2.5), *DQA1*02:01-DQB1*02:01* (DQ2.2) or *DQA1*0505-DQB1*0301* (DQ7.5) haplotype] on the opposite chromosome (Table S1). The absence of patients carrying a single risk haplotype 18.2AH underscore that other HLA-linked loci are involved in the disease process. If only DQ alleles were involved, the 2 haplotypes A8.1AH and 18.2AH (which are identical at the DQA1 and DQB1 loci) should carry the same risk. This study as well as a previous Sardinian study[21] suggests that 18.2AH has a recessive pattern of inheritance. In

TABLE 3 Frequency and ratio of exposure rates (RR) of extended HLA haplotypes in association with celiac disease stratified by Scandinavian ethnicity

HLA-DR-DQ haplotype				Scandinavian ethnicity				
DRB1	**DRB345**	**DQA1**	**DQB1**	**Pool**	**Cases**	**Controls**	**RR**	***P*-value**
*DRB1*01:01:01*		*DQA1*01:01:01*	*DQB1*05:01:01*	0.039	0.018	0.075	0.240	.002
*DRB1*01:02:01*		*DQA1*01:01:02*	*DQB1*05:01:01*	0.007	0.004	0.012	0.333	.285
*DRB1*03:01:01*	*DRB3*01:01:02*	*DQA1*05:01:01*	*DQB1*02:01:01*	0.329	0.463	0.100	4.630	<.0001
*DRB1*03:01:01*	*DRB3*02:02:01*	*DQA1*05:01:01*	*DQB1*02:01:01*	0.016	0.026	0.000	>1	.038
*DRB1*04:01:01*	*DRB4*01:03:01*	*DQA1*03:01:01*	*DQB1*03:02:01*	0.127	0.165	0.062	2.661	.002
*DRB1*04:01:01*	*DRB4*01:03:01*	*DQA1*03:03:01*	*DQB1*03:02:01*	0.012	0.011	0.012	0.917	.890
*DRB1*04:02:01*	*DRB4*01:03:01*	*DQA1*03:01:01*	*DQB1*03:02:01*					
*DRB1*04:04:01*	*DRB4*01:03:01*	*DQA1*03:01:01*	*DQB1*03:02:01*	0.081	0.099	0.050	1.980	.092
*DRB1*04:04:01*	*DRB4*01:03:02*	*DQA1*03:01:01*	*DQB1*03:02:01*	0.014	0.015	0.012	1.250	.849
*DRB1*04:07:01*	*DRB4*01:03:01*	*DQA1*03:03:01*	*DQB1*03:01:01*					
*DRB1*07:01:01*	*DRB4*01:01:01*	*DQA1*02:01*	*DQB1*02:02:01*	0.030	0.033	0.025	1.320	.629
*DRB1*07:01:01*	*DRB4*01:03:01*	*DQA1*02:01*	*DQB1*02:02:01*	0.023	0.022	0.019	1.158	.843
*DRB1*07:01:01*	*DRB4*01:03:01*	*DQA1*02:01*	*DQB1*03:03:02*	0.016	0.011	0.031	0.355	.263
*DRB1*08:01:03*		*DQA1*04:01:01*	*DQB1*04:02:01*	0.028	0.011	0.056	0.196	.005
*DRB1*08:03:02*		*DQA1*06:01:01*	*DQB1*03:01:01*					
*DRB1*10:01:01*		*DQA1*01:05:01*	*DQB1*05:01:01*	0.014	0.007	0.025	0.280	.127
*DRB1*11:01:01*	*DRB3*02:02:01*	*DQA1*05:05:01*	*DQB1*03:01:01*	0.030	0.011	0.062	0.177	.002
*DRB1*11:04:01*	*DRB3*02:02:01*	*DQA1*05:05:01*	*DQB1*03:01:01*					
*DRB1*12:01:01*	*DRB3*02:02:01*	*DQA1*05:05:01*	*DQB1*03:01:01*	0.007	0.004	0.012	0.333	.285
*DRB1*13:01:01*	*DRB3*01:01:02*	*DQA1*01:03:01*	*DQB1*06:03:01*	0.021	0.004	0.050	0.080	.001
*DRB1*13:01:01*	*DRB3*02:02:01*	*DQA1*01:03:01*	*DQB1*06:03:01*	0.021	0.007	0.044	0.159	.009
*DRB1*13:02:01*	*DRB3*03:01:01*	*DQA1*01:02:01*	*DQB1*06:04:01*	0.019	0.015	0.025	0.600	.439
*DRB1*13:03:01*	*DRB3*01:01:02*	*DQA1*05:05:01*	*DQB1*03:01:01*					
*DRB1*14:54:01*	*DRB3*02:02:01*	*DQA1*01:04:01*	*DQB1*05:03:01*	0.009	0.007	0.012	0.583	.588
*DRB1*15:01:01*	*DRB5*01:01:01*	*DQA1*01:02:01*	*DQB1*06:02:01*	0.095	0.037	0.194	0.191	<.0001
*DRB1*15:02:01*	*DRB5*01:02*	*DQA1*01:03:01*	*DQB1*06:01:01*					
*DRB1*16:01:01*	*DRB5*02:02*	*DQA1*01:02:02*	*DQB1*05:02:01*					

contrast, the 8.1AH haplotype is inherited to affected individuals with an additive effect, and often as a single risk haplotype where the opposite haplotype can be any haplotype. The 18.2AH haplotype confers perhaps a synergistic effect together with the 8.1AH haplotype increasing the risk of celiac disease in those individuals carrying both of these variants, while the 18.2AH haplotype does not seem to confer risk on its own. These analyses of the different extended DRB1-DRB345-DQ haplotypes suggest that the risk of celiac disease cannot be assigned to the DQ locus alone.

The population prevalence of celiac disease is about 1/91 and the heritability was estimated to 87%, the influence of common environmental factors to 12% and the contribution of the unshared environmental component of variance at the 1% level.[22] In a register based twin study of celiac disease cases, the heritability of diagnosed celiac disease was estimated at 75% (55% to 96%) and the non-HLA heritability contributed to 68% (40%-96%).[23,24] Surprisingly, the HLA loci only account for an additional 6% of the heritability of diagnosed celiac disease. A possible reason for the low estimate could be the fact that the HLA alleles were estimated from population frequencies and not actually genotyped.[25] Further exploring the unknown heritability of celiac disease and including extended HLA genotyping will perhaps reveal additional HLA alleles outside the DR-DQ region that are present in the population carrying genetic risk.

The translation product of the *DRB3*01* and *DRB3*02* alleles generate class II heterodimer molecules with the essentially non-polymorphic alpha (DRA) chain, which are important to the immune system by presenting peptides derived from extracellular proteins.[26] However, there is a paucity of information about the functional role of the heterodimers made up of *DRB3*01* and *DRB3*02*. The *DRB3*02* allele has been shown to be important in hepatitis virus clearance[25] and *DRB3*02* on the *DRB1*03:01* haplotype contributed to an increased risk for type 1 diabetes when compared to the *DRB1*03:01/*03:01* homozygotes carrying only the *DRB3*01* haplotype on both chromosomes.[27] We note that *DRB3*02* may confer a synergistic effect increasing the risk for type 1 diabetes as well as celiac disease in combination with another *DRB3*01* or *DRB3*02* allele, but a potentially protective effect for celiac disease, counteracting the *DQA1*05:01-DQB1*02:01* molecules, unless combined with other HLA risk haplotypes.

TABLE 4 Frequency and ratio of exposure rates (RR) of extended HLA haplotypes in association with celiac disease stratified by non-Scandinavian ethnicity

HLA-DR-DQ haplotype				Non-Scandinavian ethnicity				
DRB1	**DRB345**	**DQA1**	**DQB1**	**Pool**	**Cases**	**Controls**	**RR**	***P*-value**
*DRB1*01:01:01*		*DQA1*01:01:01*	*DQB1*05:01:01*					
*DRB1*01:02:01*		*DQA1*01:01:02*	*DQB1*05:01:01*	0.032	0.000	0.036	0.000	.461
*DRB1*03:01:01*	*DRB3*01:01:02*	*DQA1*05:01:01*	*DQB1*02:01:01*					
*DRB1*03:01:01*	*DRB3*02:02:01*	*DQA1*05:01:01*	*DQB1*02:01:01*	0.112	0.286	0.091	3.143	.034
*DRB1*04:01:01*	*DRB4*01:03:01*	*DQA1*03:01:01*	*DQB1*03:02:01*					
*DRB1*04:01:01*	*DRB4*01:03:01*	*DQA1*03:03:01*	*DQB1*03:02:01*					
*DRB1*04:02:01*	*DRB4*01:03:01*	*DQA1*03:01:01*	*DQB1*03:02:01*					
*DRB1*04:04:01*	*DRB4*01:03:01*	*DQA1*03:01:01*	*DQB1*03:02:01*					
*DRB1*04:04:01*	*DRB4*01:03:02*	*DQA1*03:01:01*	*DQB1*03:02:01*					
*DRB1*04:07:01*	*DRB4*01:03:01*	*DQA1*03:03:01*	*DQB1*03:01:01*					
*DRB1*07:01:01*	*DRB4*01:01:01*	*DQA1*02:01*	*DQB1*02:02:01*					
*DRB1*07:01:01*	*DRB4*01:03:01*	*DQA1*02:01*	*DQB1*02:02:01*	0.032	0.143	0.018	7.944	.011
*DRB1*07:01:01*	*DRB4*01:03:01*	*DQA1*02:01*	*DQB1*03:03:02*					
*DRB1*08:01:03*		*DQA1*04:01:01*	*DQB1*04:02:01*					
*DRB1*08:03:02*		*DQA1*06:01:01*	*DQB1*03:01:01*					
*DRB1*10:01:01*		*DQA1*01:05:01*	*DQB1*05:01:01*	0.048	0.000	0.055	0.000	.358
*DRB1*11:01:01*	*DRB3*02:02:01*	*DQA1*05:05:01*	*DQB1*03:01:01*	0.065	0.000	0.073	0.000	.280
*DRB1*11:04:01*	*DRB3*02:02:01*	*DQA1*05:05:01*	*DQB1*03:01:01*	0.081	0.286	0.055	5.200	.001
*DRB1*12:01:01*	*DRB3*02:02:01*	*DQA1*05:05:01*	*DQB1*03:01:01*					
*DRB1*13:01:01*	*DRB3*01:01:02*	*DQA1*01:03:01*	*DQB1*06:03:01*					
*DRB1*13:01:01*	*DRB3*02:02:01*	*DQA1*01:03:01*	*DQB1*06:03:01*					
*DRB1*13:02:01*	*DRB3*03:01:01*	*DQA1*01:02:01*	*DQB1*06:04:01*	0.040	0.000	0.045	0.000	.487
*DRB1*13:03:01*	*DRB3*01:01:02*	*DQA1*05:05:01*	*DQB1*03:01:01*	0.045	0.000	0.045	0.000	.487
*DRB1*14:54:01*	*DRB3*02:02:01*	*DQA1*01:04:01*	*DQB1*05:03:01*					
*DRB1*15:01:01*	*DRB5*01:01:01*	*DQA1*01:02:01*	*DQB1*06:02:01*	0.040	0.000	0.045	0.000	.405
*DRB1*15:02:01*	*DRB5*01:02*	*DQA1*01:03:01*	*DQB1*06:01:01*	0.040	0.000	0.045	0.000	.405
*DRB1*16:01:01*	*DRB5*02:02*	*DQA1*01:02:02*	*DQB1*05:02:01*	0.024	0.000	0.027	0.000	.526

In this study, the polymorphism of HLA-DRB1, DRB3, DRB4, DRB5, DQA1 and DQB1 genes is examined in a sample of patients with celiac residing in south of Sweden, in order to contribute new data on the association between HLA and celiac disease. The results confirm that the celiac risk haplotype known as DQ2.5 can be divided into 2 distinct haplotypes; *DRB3*01-DQA1*05:01-DQB*02:01* (also referred to as 8.1AH) and *DRB3*02-DQA1*05:01-DQB*02:01* (also referred to as 18.2AH), with a seemingly different risk associated with each haplotype. Unlike 8.1AH, the 18.2AH haplotype is likely to confer only a small risk of celiac disease on its own, despite being identical at the DQA1 and DQB1 loci.

The predictive strength of NGTS for assessing HLA-DR and DQ has provided a total of 15 new haplotypes found to have meaningful haplotypic frequencies in the general Swedish population and allowed for profound dissection using this "HLA-omic" analysis.[28] In addition, due to reduced costs of PCR, novel instrumentations and approaches in bioinformatics, make NGTS HLA typing affordable and precise. To our knowledge, this is the first study that implement NGTS in HLA haplotype analysis in celiac disease. Although this study showed differences between *DRB3*01-DQA1*05:01-DQB1*02:01* and *DRB3*02-DQA1*05:01-DQB1*02:01* on the disease risk, there are some potential limitations with this study. Firstly, the sample size was rather small and heterogeneous, which can affect the full spectrum of studied as well as other unstudied variants in the population of interest. The number of informative haplotypes (*DRB3*02-DQA1*05:01-DQB1*02:01*) in our population was limited; therefore, the statistical power to detect the differential effect of *DRB3*02-DQA1*05:01-DQB1*02:01* and *DRB3*01-DQA1*05:01-DQB1*02:01* was modest. Secondly, the study included study participants from a single clinical site in Sweden which can contain selection probability bias. Sequencing of the HLA coding region alone will be insufficient for a complete understanding of celiac disease. Further analyses are required to determine also the transcription of the fundamental genes involved in the HLA functional pathway, along with physically interacting targets and/or further investigation of regulatory regions such as those containing transcription factor-binding sites.

5 | CONCLUSION

In conclusion, extended HLA genotyping by NGTS found a differentiated effect of the presence of either *DRB3*01:01:02* and *DRB3*02:02:01* in linkage with *DQA1*05:01-DQB1*02:01* on the risk of celiac disease. The associated disease risk of *DR3-DRB3*01:01:02-DQA1*05:01-DQB1*02:01* seems more predominant among patients of Scandinavian ethnicity. This finding needs to be replicated in larger populations before these risk estimations of individuals for celiac disease can be further implemented in the clinical setting.

Conflict of interest

The authors have declared no conflicting interests.

REFERENCES

1. Duke JL, Lind C, Mackiewicz K, et al. Determining performance characteristics of an NGS-based HLA typing method for clinical applications. *HLA*. 2016;87:141-52.
2. Erlich HA. HLA typing using next generation sequencing: an overview. *Hum Immunol*. 2015;76:887-90.
3. Profaizer T, Lázár-Molnár E, Close DW, Delgado JC, Kumánovics A. HLA genotyping in the clinical laboratory: comparison of next-generation sequencing methods. *HLA*. 2016;88:14-24.
4. Monos D, Maiers MJ. Progressing towards the complete and thorough characterization of the HLA genes by NGS (or single-molecule DNA sequencing): consequences, opportunities and challenges. *Hum Immunol*. 2015;76:883-6.
5. Shiina T, Suzuki S, Ozaki Y, et al. Super high resolution for single molecule-sequence-based typing of classical HLA loci at the 8-digit level using next generation sequencers. *Tissue Antigens*. 2012;80:305-16.
6. Sollid LM, Markussen G, Ek J, Gjerde H, Vartdal F, Thorsby E. Evidence for a primary association of celiac disease to a particular HLA-DQ alpha/beta heterodimer. *J Exp Med*. 1989;169:345-50.
7. Sharma A, Liu X, Hadley D, et al. Identification of non-HLA genes associated with celiac disease and country-specific differences in a large, International Pediatric Cohort. *PLoS One*. 2016;11:e0152476.
8. Trynka G, Hunt KA, Bockett NA, et al. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet*. 2011;43:1193-201.
9. Romanos J, Rosen A, Kumar V, et al. Improving coeliac disease risk prediction by testing non-HLA variants additional to HLA variants. *Gut*. 2014;63:415-22.
10. Hunt KA, Zhernakova A, Turner G, et al. Newly identified genetic risk variants for celiac disease related to the immune response. *Nat Genet*. 2008;40:395-402.
11. Gutierrez-Achury J, Zhernakova A, Pulit SL, et al. Fine mapping in the MHC region accounts for 18% additional genetic risk for celiac disease. *Nat Genet*. 2015;47:577-8.
12. Djilali-Saiah I, Caillat-Zucman S, Schmitz J, Chaves-Vieira ML, Bach JF. Polymorphism of antigen processing (TAP, LMP) and HLA class II genes in celiac disease. *Hum Immunol*. 1994;40:8-16.
13. Hadley D, Hagopian W, Liu E, et al. HLA-DPB1*04:01 protects genetically susceptible children from celiac disease autoimmunity in the TEDDY Study. *Am J Gastroenterol*. 2015;110:915-20.
14. Paunić V, Gragert L, Madbouly A, Freeman J, Maiers M. Measuring ambiguity in HLA typing methods. *PLoS One*. 2012;7:e43585.
15. Carapito R, Radosavljevic M, Bahram S. Next-Generation Sequencing of the HLA locus: methods and impacts on HLA typing, population genetics and disease association studies. *Hum Immunol*. 2016;77:1016-1023.
16. Monten C, Bjelkenkrantz K, Gudjonsdottir AH, et al. Validity of histology for the diagnosis of paediatric coeliac disease: a Swedish multicentre study. *Scand J Gastroenterol*. 2016;51:427-33.
17. Grubin CE, Daniels T, Toivola B, et al. A novel radioligand binding assay to determine diagnostic accuracy of isoform-specific glutamic acid decarboxylase antibodies in childhood IDDM. *Diabetologia*. 1994;37:344-50.
18. Nelson WC, Pyo CW, Vogan D, et al. An integrated genotyping approach for HLA and other complex genetic systems. *Hum Immunol*. 2015;76:928-38.
19. Smith AG, Pyo CW, Nelson W, et al. Next generation sequencing to determine HLA class II genotypes in a cohort of hematopoietic cell transplant patients and donors. *Hum Immunol*. 2014;75:1040-6.
20. Schaid DJ, Rowland CM, Tines DE, Jacobson RM, Poland GA. Score tests for associations between traits and haplotypes when linkage phase is ambiguous. *Am J Hum Genet*. 2002;70:425-34.
21. Congia M, Frau F, Lampis R, et al. A high frequency of the A30, B18, DR3, DRw52, DQw2 extended haplotype in Sardinian celiac disease patients: further evidence that disease susceptibility is conferred by DQ A1*0501, B1*0201. *Tissue Antigens*. 1992;39:78-83.
22. Nisticò L, Fagnani C, Coto I, et al. Concordance, disease progression, and heritability of coeliac disease in Italian twins. *Gut*. 2006;55:803-8.
23. Ludvigsson JF, Brandt L, Montgomery SM, Granath F, Ekbom A. Validation study of villous atrophy and small intestinal inflammation in Swedish biopsy registers. *BMC Gastroenterol*. 2009;9:19.
24. Kuja-Halkola R, Lebwohl B, Halfvarson J, Wijmenga C, Magnusson PK, Ludvigsson JF. Heritability of non-HLA genetics in coeliac disease: a population-based study in 107 000 twins. *Gut*. 2016;65:1793-1798.
25. Harris RA, Sugimoto K, Kaplan DE, Ikeda F, Kamoun M, Chang KM. Human leukocyte antigen class II associations with hepatitis C virus clearance and virus-specific CD4 T cell response among Caucasians and African Americans. *Hepatology*. 2008;48:70-9.
26. Ting JP, Trowsdale J. Genetic control of MHC class II expression. *Cell*. 2002;109(Suppl):S21-33.
27. Erlich HA, Valdes AM, McDevitt SL, et al. Next generation sequencing reveals the association of DRB3*02:02 with type 1 diabetes. *Diabetes*. 2013;62:2618-22.
28. Zhao LP, Alshiekh S, Zhao M, et al. Next-generation sequencing reveals that HLA-DRB3, −DRB4, and -DRB5 may be associated with islet autoantibodies and risk for childhood type 1 diabetes. *Diabetes*. 2016;65:710-8.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Alshiekh S, Zhao LP, Lernmark Å, Geraghty DE, Naluai ÅT and Agardh D. Different *DRB1*03:01-DQB1*02:01* haplotypes confer different risk for celiac disease. *HLA*. 2017;90:95–101. https://doi.org/10.1111/tan.13065