

RESEARCH ARTICLE

Combined effect of glutamine at position 70 of HLA-DRB1 and alanine at position 57 of HLA-DQB1 in type 1 diabetes: An epitope analysis

Petroula Gerasimou^{1,2}, Vicky Nicolaidou³, Nicos Skordis⁴, Michalis Picosos⁵, Demetrios Monos⁶, Paul A. Costeas^{1*}

1 Karaiskakio Foundation, Nicosia, Cyprus, **2** University of Cyprus, Department of Biological Sciences, Nicosia, Cyprus, **3** Department of Life and Health Sciences, School of Sciences and Engineering, University of Nicosia, Nicosia Cyprus, **4** Division of Paediatric Endocrinology, Paedi Centre for Specialized Paediatrics, Nicosia, Cyprus, **5** Alithia Endocrinology Centre, Nicosia, Cyprus, **6** Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, United States of America

☯ These authors contributed equally to this work.

* paul.costeas@karaiskakio.org.cy



OPEN ACCESS

Citation: Gerasimou P, Nicolaidou V, Skordis N, Picosos M, Monos D, Costeas PA (2018) Combined effect of glutamine at position 70 of HLA-DRB1 and alanine at position 57 of HLA-DQB1 in type 1 diabetes: An epitope analysis. PLoS ONE 13(3): e0193684. <https://doi.org/10.1371/journal.pone.0193684>

Editor: Francesco Ria, Università Cattolica del Sacro Cuore, ITALY

Received: December 5, 2017

Accepted: February 15, 2018

Published: March 1, 2018

Copyright: © 2018 Gerasimou et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: The authors received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Abstract

The contribution of specific HLA Class II alleles in type 1 diabetes is determined by polymorphic amino acid epitopes that direct antigen binding therefore, along with conventional allele frequency analysis, epitope analysis can provide important insights into disease susceptibility. We analyzed the highly heterogeneous Cypriot population for the HLA class II loci of T1DM patients and controls and we report for the first time their allele frequencies. Within our patient cohort we identified a subgroup that did not carry the DRB1*03:01-DQA1*05:01-DQB1*02:01 and DRB1*04:xx-DQA1*03:01-DQB1*03:02 risk haplotypes but a novel recombinant one, DRB1*04:XX-DQA1*03:01-DQB1*02:01 designated DR4-DQ2.3. Through epitope analysis we identified established susceptibility (DQB1 A⁵⁷, DRB1 H¹³) and resistance (DQB1 D⁵⁷) residues as well as other novel susceptibility residues DRB1 Q⁷⁰, DQB1 L²⁶ and resistance residues DRB1 D⁷⁰, R⁷⁰ and DQB1 Y⁴⁷. Prevalence of susceptibility epitopes was higher in patients and was not exclusively a result of linkage disequilibrium. Residues DRB1 Q⁷⁰, DQB1 L²⁶ and A⁵⁷ and a 10 amino acid epitope of DQA1 were the most significant in discriminating risk alleles. An extended haplotype containing these epitopes was carried by 92% of our patient cohort. Sharing of susceptibility epitopes could also explain the absence of risk haplotypes in patients. Finally, many significantly associated epitopes were non-pocket residues suggesting that critical immune functions may exist spanning further from the binding pockets.

Introduction

Type 1 diabetes (T1D) is an organ-specific autoimmune disease affecting the insulin producing β cells of the pancreas, leading to absolute insulin deficiency [1]. It is one of the most common endocrine metabolic disorders in children and adolescents, however global incidence varies greatly between, and even within, countries and different ethnic populations. The lowest incidence is observed in China and the highest in Finland (0.1 v 40/ 100 000) [2] with a gradual decrease noted in countries located closer to the equator creating a north-south gradient [3]. In the Mediterranean region some areas however show notable increases such as Sardinia, which has a very high incidence similar to that of Finland, as opposed to other mainland Italian regions [4]. Great variability in T1D incidence has also been reported across different regions of Spain, where differences observed are more than threefold [5]. In the Cypriot population the incidence of T1D among children and adolescences shows an increasing trend; 10.76/100 000 between 1990–2000 [6] vs 14.9/100 000 within the subsequent five year period 2000–2004 [7]. Further studies spanning a 20-year period (1999–2009) confirm the overall increase in prevalence reporting also the largest increase in children less than five years of age [8].

Although close to 60 diabetes-associated genetic linkages have been reported to date [9], numerous studies show that the strongest association involves the inheritance of specific HLA alleles [10]. In particular, a number of HLA class II alleles namely DRB1*03:01, *04:01, *04:02, *04:05 and DQB1*02:01, *03:02 are strongly linked. A set of class II haplotypes DRB1*03:01-DQA1*05:01-DQB1*02:01 (known as DR3-DQ2.5) and DRB1*04:01/02/04/05/08-DQA1*03:01-DQB1*03:02 (known as DR4-DQ8), confer the highest risk with up to 50% of patients carrying both haplotypes, whilst haplotype DRB1*15:01-DQA1*01:02-DQB1*06:02 contributes to protection [11]. Independent associations with HLA class I alleles have also been suggested and shown to potentially affect age of onset [12–17]. More recent data have described a novel association of the less polymorphic non-classical HLA class I allele HLA-G [18–20].

HLA molecules function to present antigenic peptides to T cells and thus have a central role in immune cell activation and autoimmune disease. The peptide-binding grooves of HLA molecules are made of amino acid residues arranged in pockets; these amino acids are highly polymorphic and, either as single or groups of continuous or non-continuous residues, create millions of possible epitopes. These epitopes determine the repertoire of peptides a given HLA allele can present. Class I molecules (HLA-A, -B and -C) have binding grooves made of six pockets. Class II molecules (HLA-DR, -DQ, -DP) are heterodimers comprised of α and β chains creating binding grooves with four major pockets. In the HLA-DR heterodimer polymorphism is found only on the β chains whereas both α and β chains of the HLA-DQ and HLA-DP heterodimers are polymorphic. In addition, not all pairings between α and β chains are allowed explaining structural correlates and associations or non-associations with disease [21,22].

In addition to conventional allele frequency studies, the systematic analysis of HLA epitopes has recently been highlighted as a critical component in providing a better understanding of genetic susceptibility to T1D [23]. Indeed, even in the absence of statistically significant HLA allele association, disease susceptibility may be determined by the independent contribution of polymorphic residues participating in the formation of a functional arrangement within the binding cleft of an HLA molecule, a concept first proposed by Zerva et al. in 1996 [24]. Epitope analysis can also uncover allele associations that are missed due to their low frequency in the population or disparate alleles that share peptide-binding motifs, known as shared epitopes. In addition, single amino acid polymorphisms in the same allele have been shown to alter disease

susceptibility, for example aspartic acid in position 57 of the HLA-DQB1 is protective whereas substitution with alanine is associated with susceptibility [23,25]. Amino acid differences have been shown to discriminate even closely related alleles and alter the binding avidity for insulin peptides [26]. Finally, specific epitopes may explain disease susceptibility in patients that do not carry the established risk alleles.

In the current study we investigated the HLA frequencies in a cohort of Cypriot T1D patients, a population that has not been studied to date. We propose that studying genetically heterogeneous populations such as the Cypriot population can be critically informative in allowing further validation of established risk alleles and epitopes or uncovering new ones. We report here, for the first time, the allele frequencies in Cypriot T1DM patients and a new HLA Class II risk haplotype DR4-DQ2.3. Furthermore, we identified previously reported HLA class II susceptibility residues including DQB1 L²⁶ and A⁵⁷ but also susceptibility residue DRB1 Q⁷⁰ and alternative protective R⁷⁰ not previously reported to be associated with T1D. Finally, we show that an extended risk haplotype of HLA class II susceptibility epitopes identified in this study DRβQ⁷⁰-DQβL²⁶A⁵⁷DQαY¹¹R⁵²R⁵⁵F⁶¹T⁶⁴I⁶⁶L⁶⁹V/L⁷⁶H¹²⁹E/K¹⁷⁵ could account for 92% of our patient cohort.

Results

HLA class II allele frequencies in type 1 diabetes

We used a dataset of 170 T1D patients and 192 control subjects for whom high-resolution HLA genotyping was performed at 4 classical major histocompatibility complex class II loci DRB1, DQA1, DQB1 and DPB1. Allele frequency analysis (Fig 1 and S1–S4 Tables) showed that consistent with previous studies the alleles most significantly associated with disease were HLA-DRB1*03:01 (pcorr. = 1.06×10^{-12} , OR = 6.56) and *04:05 (pcorr. = 2×10^{-12} , OR = 7.28), HLA-DQB1*02:01 (pcorr. = 1.05×10^{-15} , OR = 6.4) and 03:02 (pcorr. = 1.19×10^{-12} and OR = 6.58), HLA-DQA1*03:01 (pcorr = 1.76×10^{-16} , OR = 6.84) and HLA-DPB1*03:01 (pcorr = 0.003, OR = 2.76). In contrast, the alleles most commonly found in the control population, and therefore deemed protective, were HLA-DRB1*14:01 (pcorr. = 4.29×10^{-8} , OR = 0.02), *11:04 (pcorr. = 7.26×10^{-5} , OR = 0.16), *10:01 (pcorr. = 5.06×10^{-4} , OR = 0.1) and *16:02 (pcorr. = 6.77×10^{-4} , OR = 0.07), HLA-DQB1*03:01 (2.85×10^{-13} , OR = 0.13) and *05:03 (pcorr. = 2.14×10^{-8} , OR = 0.02), HLA-DQA1*01:01 (pcorr. = 6.35×10^{-6} , OR = 0.31) and HLA-DPB1*04:02 (pcorr. = 4.62×10^{-12} , OR = 0.07).

HLA class II epitopes associated with susceptibility and resistance

The high-resolution HLA genotyping dataset of patients and control subjects was imported into the SKDM HLA Tool for analysis to identify epitopes associated with T1D. Both pocket and non-pocket residues were investigated for each allele and a number of them were found significantly associated. The susceptibility epitopes with the lowest p values ($p \leq 0.001$) and highest OR ($OR \geq 1.5$) and the resistance epitopes with the lowest p values ($p \leq 0.001$) and lowest OR ($OR \leq 1.5$) are summarized in Table 1. No statistically significant associations were found for HLA-DPB1 epitopes.

The strongest association was shown for the previously reported HLA-DQB1 residue A⁵⁷ ($p = 1.3 \times 10^{-27}$, OR 22.3). This is found in HLA-DQB1*02:01 and *03:02 alleles which were significantly more common in patients according to the allele frequency analysis (S2 Table). Aspartic acid in the same position (D⁵⁷) was identified as the residue most strongly associated with resistance ($p = 4.5 \times 10^{-25}$, OR 0.07). Two of the HLA-DQB1 alleles (*05:03, *03:01) sharing this residue were significantly more frequent in controls. Similarly, the presence of leucine (L) in position 26 was the second strongest residue associated with susceptibility ($p = 1.4 \times 10^{-22}$,

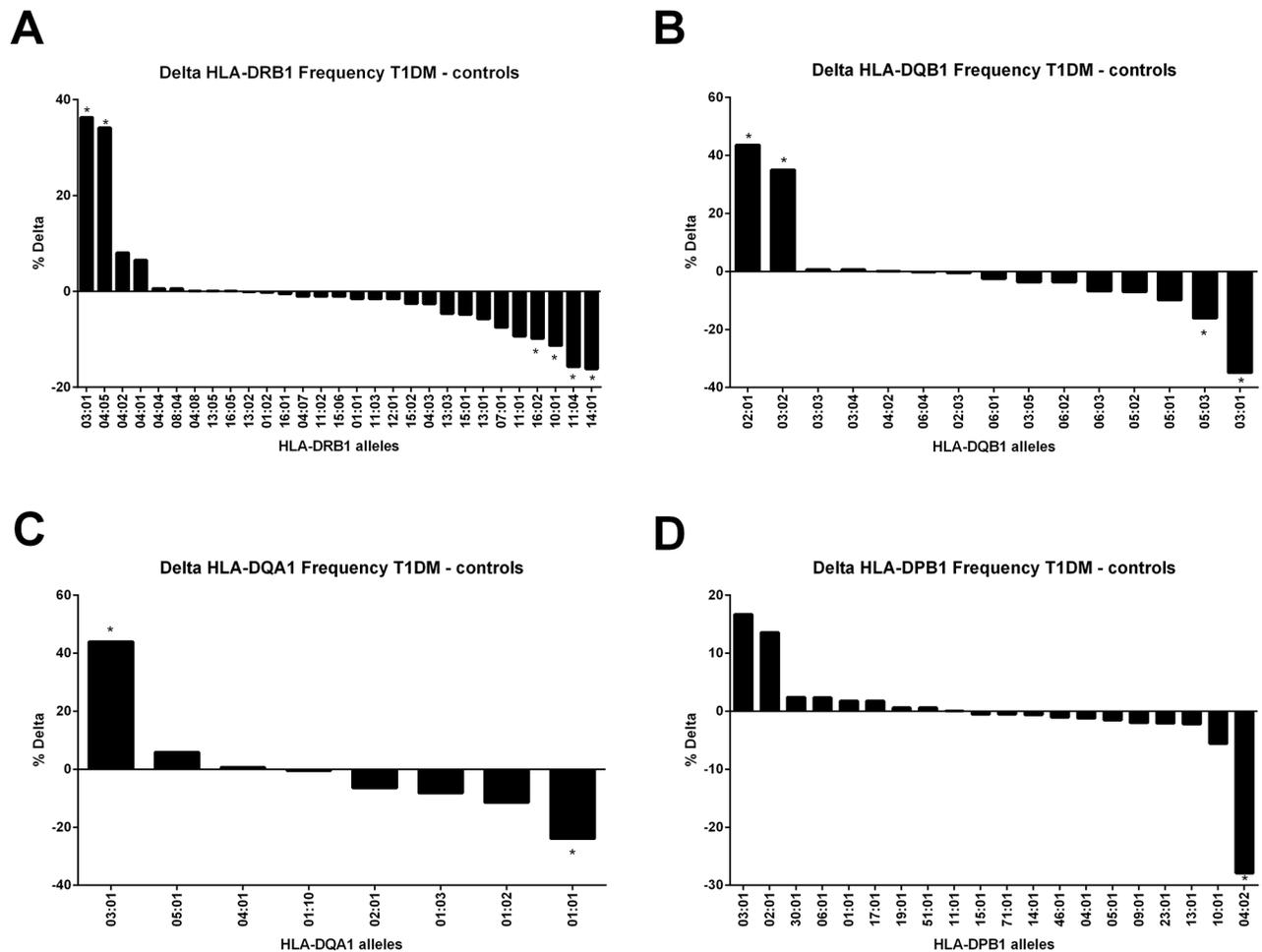


Fig 1. The population frequencies for HLA class II alleles. Population frequencies for HLA-DRB1 (A), HLA-DQB1 (B), HLA-DQA1 (C) and HLA-DPB1 (D) are shown as the delta between type 1 diabetes patients and controls. The population frequency counts only how many times an allele is present in the population so in case of homozygosity it is counted as 1. * $P_{\text{corr}} \leq 0.001$.

<https://doi.org/10.1371/journal.pone.0193684.g001>

OR 18.47), whilst tyrosine (Y) ($p = 1.1 \times 10^{-12}$, OR 0.14) or glycine (G) ($p = 1.9 \times 10^{-5}$, OR 0.32) at the same position confer resistance. L²⁶ is found in HLA-DQB1*02:01 and *03:02 alleles which are more common in patients. HLA-DQB1 susceptibility residues F⁴⁷ ($p = 1.6 \times 10^{-14}$, OR 6.2) and R⁷⁰ ($p = 1.6 \times 10^{-8}$, OR 12.1) were also found in the same position as resistance residues Y⁴⁷ ($p = 4 \times 10^{-6}$, OR 0.09) and G⁷⁰ ($p = 1.8 \times 10^{-7}$, OR 0.26). In contrast, resistance residues A¹³ ($p = 1.1 \times 10^{-12}$, OR 0.14), Y³⁷ ($p = 1.9 \times 10^{-6}$, OR 0.08), T²⁸ ($p = 4 \times 10^{-6}$, OR 0.09) were not in the same position as susceptibility ones.

The DRB1 locus contained the next most significant HLA residue associated with susceptibility to T1D; HLA-DRB1 Q⁷⁰ had the strongest association ($p = 3.6 \times 10^{-18}$, OR 12.42). This residue is shared between most high risk alleles including DRB1*03:01, *04:05 and *04:01, which were all significantly more common in patients (S1 Table). In addition, previously reported HLA-DRB1 residues H¹³ ($p = 8.8 \times 10^{-17}$, OR 7.57) and K⁷¹ ($p = 3. \times 10^{-11}$, OR 5.48) were also significantly associated with susceptibility. HLA-DRB1 H¹³ is shared by a number of HLA-DRB1 *04:xx alleles that confer high risk. HLA-DRB1 *04:01 which shares all three of these epitopes had the highest OR 10.07. HLA-DRB1 resistance residues R⁷⁰ ($p = 8.2 \times 10^{-12}$,

Table 1. HLA class II pocket epitopes associated with type 1 diabetes.

	Location	EPITOPE	PATIENT (N = 170)	CONTROL (N = 192)	Pcorr. Value	OR	Associated alleles
DRB1	13	H	110	37	8.8x10 ⁻¹⁷	7.57	04:05, 04:02, 04:01, 04:04, 04:08, 04:07, 04:03
	70	Q	158	97	3.6x10 ⁻¹⁸	12.42	03:01, 04:05, 04:01, 04:04, 04:08, 01:02, 04:07, 15:06, 01:01, 15:02, 04:03, 15:01
	70	D	84	157	7.4x10 ⁻⁹	0.22	04:02, 08:04, 13:05, 16:05, 13:02, 16:01, 11:02, 11:03, 12:01, 13:03, 13:01, 07:01, 11:01, 16:02, 11:04
	70	R	2	52	8.2x10 ⁻¹²	0.04	10:01, 14:01
	71	K	91	33	3.0x10 ⁻¹¹	5.48	03:01, 04:01, 13:03
	71	R	132	185	5.12x10 ⁻⁶	0.14	04:05, 04:04, 08:04, 04:08, 13:05, 16:05, 01:02, 16:01, 04:07, 01:01, 12:01, 04:03, 07:01, 11:01, 16:02, 10:01, 11:04, 14:01
	74	R	106	24	4.2x10 ⁻¹²	6.6	03:01
	74	E	8	46	1.4x10 ⁻⁵	0.17	04:07, 04:03, 14:01
DQA1	11	Y	165	141	4.2x10 ⁻⁹	11	03:01, 05:01, 04:01, 02:01
	11	C	82	153	3.0x10 ⁻⁸	0.21	01:03, 01:02, 01:01
	55	R	165	141	4.2x10 ⁻⁹	11	03:01, 05:01, 04:01, 02:01
	55	G	82	153	3.0x10 ⁻⁸	0.24	01:03, 01:02, 01:01
	66	I	165	141	4.2x10 ⁻⁹	11	03:01, 05:01, 04:01, 02:01
	66	M	82	153	3.0x10 ⁻⁸	0.24	01:03, 01:02, 01:01
	69	L	165	141	4.2x10 ⁻⁹	11	03:01, 05:01, 02:01
	69	A	82	153	3.0x10 ⁻⁸	0.24	01:03, 01:02, 01:01
DQB1	26	L	161	92	1.4x10 ⁻²²	18.47	02:01, 03:02, 03:03, 06:04, 02:03, 06:02, 06:03
	26	Y	17	88	1.1x10 ⁻¹²	0.14	03:04, 06:01, 03:01
	26	G	78	140	1.9x10 ⁻⁵	0.32	04:02, 03:05, 05:02, 05:01, 05:03
	47	F	121	54	1.6x10 ⁻¹⁴	6.2	02:01, 02:03
	47	Y	140	189	4.0x10 ⁻⁶	0.09	03:02, 03:03, 03:04, 04:02, 06:04, 06:01, 03:05, 06:02, 06:03, 05:02, 05:01, 05:03, 03:01
	57	A	160	78	1.3x10 ⁻²⁷	22.3	02:01, 03:02, 03:04, 03:05
	57	D	17	121	4.5x10 ⁻²⁵	0.07	03:03, 04:02, 02:03, 06:01, 06:02, 06:03, 05:03, 03:01
	70	R	166	145	1.6x10 ⁻⁸	12.1	02:01, 03:02, 03:03, 03:04, 06:04, 02:03, 06:01, 03:05, 03:01
	70	G	77	146	1.8x10 ⁻⁷	0.26	06:02, 06:03, 05:02, 05:01, 05:03

<https://doi.org/10.1371/journal.pone.0193684.t001>

OR 0.04) or D⁷⁰ (p = 7.4x10⁻⁹, OR 0.22) and R⁷¹ (p = 5.12x10⁻⁶, OR 0.14) are at the same positions as susceptibility counterparts. Resistance residue R⁷⁰ in particular had the lowest OR and is shared among two protective alleles HLA-DRB1*10:01 and *14:01 which were significantly less common in patients. The presence of arginine (R) at position 74 is associated with susceptibility (p = 4.2x10⁻¹², OR 6.6), whereas glutamic acid (E) at the same position is significantly associated with resistance (p = 1.4x10⁻⁵, OR 0.17). Finally, residues R³⁰ and A³⁸ were also significantly associated with resistance and had the next lowest OR after R⁷⁰ (p = 0.0003, OR 0.07).

A number of DQA1 residues were strongly associated with T1D; Y¹¹, R⁵⁵, I⁶⁶ and L⁶⁹ (p = 4.2x10⁻⁹, OR 11) and the previously reported R⁵⁶ and V⁷⁶ (p = 1.6x10⁻¹⁵, OR 6.8 (S7 Table). All are present in DQA1*03:01 (R⁵⁶ and V⁷⁶ are only found in this allele) most frequently observed in patients (S3 Table). Furthermore, all resistance HLA-DQA1 residues G⁵⁵, C¹¹, M⁶⁶, A⁶⁹, G⁵⁶ and M⁷⁶ (p = 3x10⁻⁸, OR 0.24) are at the same positions as susceptibility counterparts and are shared between HLA-DQA1*01:03 and *01:01 alleles, which were significantly more frequent in controls.

HLA epitopes associated with susceptibility and resistance show gene-dose effect

The effect of homozygosity or heterozygosity of class II residues identified to be associated with susceptibility or resistance was analyzed. Homozygosity of susceptibility epitope DQB1 A⁵⁷ had the strongest association with T1D ($p = 1.25 \times 10^{-24}$, OR 100.6) (Table 2), whilst inheritance of two copies of the resistance residue at the same position, DQB1 D⁵⁷, had the strongest negative association ($p = 3.94 \times 10^{-9}$, OR 0.03). A gene dose effect was apparent since inheritance of one copy of A⁵⁷ decreased the probability of disease as shown by the lower OR ($p = 5.61 \times 10^{-23}$, OR 17.5), in the same way inheritance of one copy of the susceptibility residue D⁵⁷ ($p = 4.8 \times 10^{-22}$, OR 0.08) was not as protective as inheritance of two copies but still lowered probability of disease compared to inheritance of one copy of A⁵⁷. However, resistance appears to be dominant since inheritance of only one copy of D⁵⁷ still conferred a lower probability of diabetes (OR 0.08). A similar effect was observed for DQB1 L²⁶, inheritance of two copies is strongly associated with T1D ($p = 4.3 \times 10^{-21}$, OR 45.8), inheritance of one copy greatly decreased the probability of disease ($p = 3.9 \times 10^{-19}$, OR 14.9) and two copies of G²⁶ is protective ($p = 8.68 \times 10^{-6}$, OR 0.11). Inheritance of two copies of the HLA-DRB1 Q⁷⁰ susceptibility residue was also strongly associated with T1D ($p = 1.4 \times 10^{-17}$, OR 20.1) and probability decreased with inheritance of one copy ($p = 5 \times 10^{-15}$, OR 10.2). Among the two alternative resistance residues identified for this position, inheritance of R⁷⁰ appears to be more protective than inheritance of D⁷⁰; one copy of R⁷⁰ conferred lower probability of disease ($p = 9.44 \times 10^{-13}$, OR 0.04) than one ($p = 1.8 \times 10^{-7}$, OR 0.27) or two copies of D⁷⁰ ($p = 6.1 \times 10^{-10}$, OR 0.11). The probability

Table 2. Homozygous and heterozygous inheritance of shared HLA class II epitopes associated with type 1 diabetes.

Locus	Genotype	P corr value	OR
HLA-DRB1	Q ⁷⁰⁺ /Q ⁷⁰⁺	1.4x10 ⁻¹⁷	20.1
	Q ⁷⁰⁺ /Q ⁷⁰⁻	5.0x10 ⁻¹⁵	10.2
	D ⁷⁰⁺ /D ⁷⁰⁻	1.8x10 ⁻⁷	0.27
	D ⁷⁰⁺ /D ⁷⁰⁺	6.1x10 ⁻¹⁰	0.11
	R ⁷⁰⁺ /R ⁷⁰⁻	9.44x10 ⁻¹³	0.04
	K ⁷¹⁺ /K ⁷¹⁻	2.6x10 ⁻¹²	5.5
	R ⁷¹⁺ /R ⁷¹⁻	4.7x10 ⁻⁵	0.2
	R ⁷¹⁺ /R ⁷¹⁺	3.9x10 ⁻⁹	0.1
	E ⁷⁴⁺ /E ⁷⁴⁻	3.4x10 ⁻⁷	0.16
	HLA-DQB1	L ²⁶⁺ /L ²⁶⁺	4.3x10 ⁻²¹
L ²⁶⁺ /L ²⁶⁻		3.9x10 ⁻¹⁹	14.9
G ²⁶⁺ /G ²⁶⁺		8.68x10 ⁻⁶	0.11
A ⁵⁷⁺ /A ⁵⁷⁺		1.25x10 ⁻²⁴	100.6
A ⁵⁷⁺ /A ⁵⁷⁻		5.61x10 ⁻²³	17.5
D ⁵⁷⁺ /D ⁵⁷⁻		4.8x10 ⁻²²	0.08
D ⁵⁷⁺ /D ⁵⁷⁺		3.94x10 ⁻⁹	0.03
R ⁷⁰⁺ /R ⁷⁰⁺		1.3x10 ⁻¹⁰	17.6
R ⁷⁰⁺ /R ⁷⁰⁻		2.8x10 ⁻⁸	10.4
G ⁷⁰⁺ /G ⁷⁰⁻		1.3x10 ⁻⁶	0.31
HLA-DQA1	G ⁷⁰⁺ /G ⁷⁰⁺	2.3x10 ⁻⁸	0.08
	L ⁶⁹⁺ /L ⁶⁹⁺	1.07x10 ⁻¹²	22.86
	L ⁶⁹⁺ /L ⁶⁹⁻	2.69x10 ⁻⁸	8.82

<https://doi.org/10.1371/journal.pone.0193684.t002>

of disease in the case of inheritance of two copies of R⁷⁰ was not calculated as this genotype was not present in the population.

Linkage disequilibrium between DQ and DR susceptibility epitopes

The prevalence of having both the highest susceptibility residues DRβ Q⁷⁰ and the DQβ A⁵⁷ was compared between patients and control subjects (Table 3). The majority of patients (90%) had both epitopes in contrast to 29.7% of the control subjects ($p \leq 0.0001$, OR 21.3). A very small percentage of patients had only the DRβ Q⁷⁰ epitope (2.9%) or only the DQβ A⁵⁷ epitope (4.1%) or none of the two epitopes (2.9%). Among control subjects the highest percentage (38.5%) had none of these susceptibility epitopes ($p \leq 0.0001$, OR 0.05). These findings suggest that our observations are not only due to linkage disequilibrium (LD) between the two loci. In addition, it was shown that these DR/DQ susceptibility epitopes were not in LD in patients ($p = 0.001$) but were in LD in controls ($p = 1.8 \times 10^{-6}$).

Shared susceptibility residues account for type 1 diabetes in patients lacking HLA class II risk haplotypes

Having identified a number of susceptibility and resistance residues associated with T1D we sought to further dissect our patient cohort with regards to their HLA genotypes aiming to explain disease susceptibility by the presence of shared susceptibility residues. More specifically, among our T1D patients the majority (135 of 170 or 79%) carried at least one or both of the susceptibility alleles DR3-DQ2.5 and DR4-DQ8 (Table 4). However, a significant number of patients (35 of 170 or 21%) did not carry any copies of either haplotype. We observed a new recombinant haplotype DRB1*04:XX-DQA1*03:01-DQB1*02:01, henceforth named DR4-DQ2.3, dominant within this subgroup; 22 (63%) of 35 patients carrying non-risk haplotypes and 13% of all patients carried one copy of the DR4-DQ2.3 haplotype in contrast to only 9 of 192 control subjects (4.7%). Interestingly, no DR4-DQ2.3 homozygous individuals were identified. The DR4-DQ2.3 haplotype contains the HLA-DQB1*02:01 allele instead of the HLA-DQB1*03:02, however both alleles share a number of DQB1 susceptibility residues such as L²⁶, R⁷⁰, L⁸⁵, E⁸⁶, T⁸⁹, but more importantly the most significant susceptibility residue A⁵⁷. Hence the presence of either HLA-DQB1*02:01 or HLA-DQB1*03:02 makes the patient homozygous for the A⁵⁷ residue, associated with the highest OR in the zygosity analysis ($p = 1.25 \times 10^{-24}$ OR 100.6).

Carrying at least one copy of any of the risk haplotypes DR3-DQ2.5, DR4-DQ8 or DR4-DQ2.3 could account for 92% (157 of 170) of patients (Table 4). However, an additional 8% of our patients (13 of 170) do not carry any of the risk haplotypes. We were able to verify that all these patients carried susceptibility associated residues identified in this study in the DRB1 locus and one or more in the DQA–DQB loci.

HLA epitopes associated with susceptibility and resistance and their potential function

To investigate whether epitopes differentiate associated alleles, allele sequences were retrieved and aligned using the IMGT/HLA database of the European Bioinformatics Institute. Within these sequences we noted all the susceptibility and resistance residues, both pocket and non-pocket, identified in this study and also found to have a proposed function according to literature [27–29]. The DQA, DQB and DRB domains show a considerable number of polymorphisms that are mainly involved in antigen binding by the anchoring pockets, the heterodimer

Table 3. Linkage disequilibrium of DR/DQ susceptibility epitopes in type 1 diabetes patients and control subjects.

HLA class II susceptibility epitopes	Patients (%) (N = 170)	Control subjects (%) (N = 192)	P Value	OR (95% CI)	LD in patients P value (OR)	LD in controls P value (OR)
DRβ Q ⁷⁰⁺ /DQβ A ⁵⁷⁺	153 (90)	57 (29.7)	≤ 0.0001	21.3 (11.8–38.4)	0.001 (20.5)	1.80x10 ⁻⁶ (4.9)
DRβ Q ⁷⁰⁺ /DQβ A ⁵⁷⁻	5 (2.9)	40 (20.8)	≤ 0.0001	0.1 (0.04–0.3)		
DRβ Q ⁷⁰⁻ /DQβ A ⁵⁷⁺	7 (4.1)	21 (10.9)	0.02	0.35 (0.14–0.84)		
DRβ Q ⁷⁰⁻ /DQβ A ⁵⁷⁻	5 (2.9)	74 (38.5)	≤ 0.0001	0.05 (0.02–0.12)		

n (%) p value Fisher exact test

<https://doi.org/10.1371/journal.pone.0193684.t003>

formation by salt bridges, T-cell receptor (TCR) or CD4 co-receptor binding and in the formation of the dimer of heterodimers.

HLA DRB1 E⁹, V¹¹, H¹³, Y²⁶, N³⁷ and R⁷⁴ that are associated with diabetes are amino acids that are part of binding pockets (Table 5). The residue at position 57 is involved in pocket 9 but also participates in hydrogen bond formation to the peptide. A serine in that position is highly associated to T1D while an alanine is associated to susceptibility. The residues at positions 67, 70 and 71 are also part of the pocket formation but are also sites for TCR contact; alternative residues showed either susceptibility or resistance. Lastly, position 112 has a potential function in the homodimer of heterodimers and the residue at position 140 is a potential contact side for the CD4 co-receptor. HLA DQB1 residues involved in the formation of the peptide pocket include positions 13, 26, 28, 30, 37, 47, 57, 67, 70, 71, 74, 85, 86, 89 and 90 (Table 6). Amino acids at position 30 and 57 are also involved in the formation of a hydrogen bond to the peptide while residues at position 67, 70 and 71 are also a potential TCR contact site. Amino acids 52, 53 and 55 act as a homodimerization patch in the dimer formation. Pocket residues of the HLA DQA1 molecule include positions 11, 52, 66, 69 and 76 (Table 7). Amino acids at positions 69 and 76 also form a hydrogen bond to the peptide. Important residues at positions 55–64 are potential TCR contact sites while position 129 is a potential CD4 contact site. The residue at position 175 upholds a function in the formation of the homodimer of heterodimers.

Table 4. HLA class II genotypes of type 1 diabetes patients and control subjects.

HLA class II Genotype	Patients (%) (N = 170)	Control subjects (%) (N = 192)	P value	OR
DR3—DQ2.5 / DRX	30 (17.65)	22 (11.46)	0.1	1.66
DR3—DQ2.5 / DR3—DQ2.5	18 (10.59)	0	<0.0001	46.7
DR3—DQ2.5 / DR4—DQ2.3	5 (2.94)	1 (0.52)	0.1	5.79
DR4—DQ8 / DRX	43 (25.29)	18 (9.38)	<0.0001	3.27
DR4—DQ8 / DR4—DQ8	5 (2.94)	1 (0.52)	0.1	5.79
DR4—DQ8 / DR4—DQ2.3	4 (2.35)	0	0.047	10.41
DR3—DQ2.5 / DR4—DQ8	30 (17.65)	1 (0.52)	<0.0001	40.93
DR4—DQ2.3 / DRX	22 (12.94)	8 (4.17)	0.004	3.42
DRX / DRX (non-risk)	13 (7.65)	141 (73.44)	<0.0001	0.03

DR3—DQ2.5 = DRB1*03:01-DQA1*05:01-DQB1*02:01, DR4—DQ8 = DRB1*04:XX-DQA1*03:01-DQB1*03:02, DR4—DQ2.3 = DRB1*04:XX-DQA1*03:01-DQB1*02:01

n (%) p value Fisher exact test

<https://doi.org/10.1371/journal.pone.0193684.t004>

Table 5. Type 1 diabetes associated polymorphic residues of DRB1 alleles with a proposed function.

DRB1 allele	Amino Acid position													
						HB						HH		
								TCR				TCR		CD4
	P9	P6	P4	P4	P9	P9	P7	P7	P4/7	P4				
	9	11	13	26	37	57	67	70	71	74	77	112	140	
Susceptibility alleles														
03:01	E	S	S	Y	N	D	L	Q	K	R	N	H	T	
04:01	E	V	H	F	Y	D	L	Q	K	A	T	H	T	
04:05	E	V	H	F	Y	S	L	Q	R	A	T	H	T	
Protective alleles														
10:01	E	V	F	V	Y	D	L	R	R	A	T	H	T	
11:04	E	S	S	F	Y	D	F	D	R	A	T	H	T	
14:01	E	S	S	F	F	A	L	R	R	E	T	Y	T	
16:02	W	P	R	F	S	D	L	D	R	A	T	H	A	

Positions identified in bold show amino acids that are exclusive to the type 1 diabetes susceptibility or protective alleles. The function associated with each amino acid is depicted on top. P; pocket (peptide binding), HH; homodimer of heterodimer, HB; hydrogen bond to peptide, TCR; site of contact of HLA molecule to T-cell receptor. CD4; site of contact of HLA molecule to CD4.

<https://doi.org/10.1371/journal.pone.0193684.t005>

Overall, all risk associated HLA alleles contain more susceptibility residues and protective alleles contain more protective residues. Some residues, however, might not be as critical as others. For example, HLA-DRB1 susceptibility residue E⁹ and L⁶⁷ and resistance residue T⁷⁷ are found in both risk and protective HLA-DRB1 alleles. In contrast, all risk HLA-DRB1 alleles contain the susceptibility epitope Q⁷⁰ which is absent from protective alleles. Epitopes also alter the susceptibility of closely related alleles. For example, HLA-DQB1*03:01 and *03:02 have very similar amino acid sequences but differ at critical position 57, risk associated HLA-DQB1*03:02 contains alanine (A), whereas protective HLA-DRB1*03:01 contains aspartic acid (D) in the same position. In addition, risk associated HLA-DQB1*03:02 also contains the susceptibility epitope L²⁶. Finally, a sequence of 10 amino acid residues could differentiate

Table 6. Type 1 diabetes associated polymorphic residues of DQB1 alleles with a proposed function.

DQB1 allele	Amino Acid position																		
									HH									TCR	
				HB							SB								
	P4	P4	P4/7	P6	P9	P7					P9		P7	P4	P4/7	P4	P1	P1	P1
	13	26	28	30	37	47	52	53	55	57	66	67	70	71	74	85	86	89	90
Susceptibility alleles																			
02:01	G	L	S	S	I	F	L	L	L	A	D	I	R	K	A	L	E	T	T
03:02	G	L	T	Y	Y	Y	P	L	P	A	E	V	R	T	E	L	E	T	T
Protective alleles																			
03:01	A	Y	T	Y	Y	Y	P	L	L	D	E	V	R	T	E	L	E	T	T
05:03	G	G	T	H	Y	Y	P	Q	P	D	E	V	G	A	S	V	A	G	I

Positions identified in bold show amino acids that are exclusive to the type 1 diabetes positively or negatively associated alleles. The function associated with each amino acid is depicted on top. P; pocket (peptide binding), HH; homodimer of heterodimer, HB; hydrogen bond to peptide, SB; salt bridge, TCR; site of contact of HLA molecule to T-cell receptor.

<https://doi.org/10.1371/journal.pone.0193684.t006>

Table 7. Type 1 diabetes associated polymorphic residues of DQA1 alleles with a proposed function.

DQA1 allele	Amino Acid position									
				TCR		HB			CD4	
	P6	P1				P6	P6/9	P9		HH
	11	52	55	61	64	66	69	76	129	175
Susceptibility alleles										
03:01	Y	R	R	F	T	I	L	V	H	E
05:01	Y	R	R	F	T	I	L	L	H	K
Protective alleles										
01:01	C	S	G	G	R	M	A	M	Q	Q
01:02	C	S	G	G	R	M	A	M	Q	Q
01:03	C	S	G	G	R	M	A	M	H	Q

Positions identified in bold show amino acids that are exclusive to the type 1 diabetes susceptibility or protective alleles. The function associated with each amino acid is depicted on top. P; pocket (peptide binding), HH; homodimer of heterodimer, HB; hydrogen bond to peptide, TCR; T-cell receptor contact site to HLA molecule, CD4; site of contact of HLA molecule to CD4.

<https://doi.org/10.1371/journal.pone.0193684.t007>

risk and protective HLA-DQA1 alleles. Epitope Y¹¹R⁵²R⁵⁵F⁶¹T⁶⁴I⁶⁶L⁶⁹V/L⁷⁶H¹²⁹E/K¹⁷⁵ was exclusive to risk associated alleles and was not observed in any protective alleles.

Extended risk haplotype of HLA class II susceptibility epitopes accounts for 92% of type 1 patients

We counted the number of individuals carrying an extended haplotype containing all the HLA class II risk associated epitopes in our patient-control cohort (Table 8). The epitopes we included were the ones that differentiated risk and protective alleles as described above. The vast majority of patients (167, 92%) carried at least one copy of the DRβQ⁷⁰-DQB L²⁶A⁵⁷D-QαY¹¹R⁵²R⁵⁵F⁶¹T⁶⁴I⁶⁶L⁶⁹V/L⁷⁶H¹²⁹E/K¹⁷⁵ epitope haplotype compared with only 27% of the control subjects. A very small percentage of patients had only the DRβ Q⁷⁰ residue (2.9%) or only the DQB A⁵⁷ residue (4.1%) or none of the two (2.9%). Among control subjects the highest percentage (73%) had none of these susceptibility epitopes compared to only 8% of patients (p≤0.0001, OR 0.03).

Discussion

The association of class II alleles with T1D susceptibility is well documented even though the exact mechanism that confers the disease risk is yet to be fully understood. The allele frequencies of Cypriot T1D patients had not been previously reported. The vast majority of patients (79%) carried the established risk haplotypes DR3—DQ2.5 and DR4—DQ8 either in

Table 8. Extended haplotype of class II risk epitopes in type 1 diabetes patients and control subjects.

HLA class II Genotype	Patients (%) (N = 182)	Control subjects (%) (N = 192)	P value	OR	95% CI
DRB Q ⁷⁰ -DQB L ²⁶ A ⁵⁷ DQA Y ¹¹ R ⁵² R ⁵⁵ F ⁶¹ T ⁶⁴ I ⁶⁶ L ⁶⁹ V/L ⁷⁶ H ¹²⁹ E/K ¹⁷⁵ / X	103 (57%)	48 (25%)	<0.0001	3.8	2.5–5.9
DRB Q ⁷⁰ —DQB L ²⁶ A ⁵⁷ DQA Y ¹¹ R ⁵² R ⁵⁵ F ⁶¹ T ⁶⁴ I ⁶⁶ L ⁶⁹ V/L ⁷⁶ H ¹²⁹ E/K ¹⁷⁵ / DRB Q ⁷⁰ —DQB L ²⁶ A ⁵⁷ DQA Y ¹¹ R ⁵² R ⁵⁵ F ⁶¹ T ⁶⁴ I ⁶⁶ L ⁶⁹ V/L ⁷⁶ H ¹²⁹ E/K ¹⁷⁵	64 (35%)	3 (2%)	<0.0001	34.2	10.5–111.2
X / X (non-risk)	15 (8%)	141 (73%)	<0.0001	0.03	0.02–0.06

<https://doi.org/10.1371/journal.pone.0193684.t008>

heterozygous or homozygous, or carried both haplotypes, while only 20% of the control population were carriers ($p < 0.001$). A significant percentage (21%) of our diabetic cohort did not carry the risk haplotypes probably due to these not being present at high frequency in the Cypriot population, especially as compared to European Caucasians. In the latter, these haplotypes represent the first and second most common haplotypes respectively in contrast to ranking 87th and 91st in Cypriots. We thus believe that the Cypriot population represents an excellent study sample that can allow further dissection of T1D disease susceptibility. Using our highly diverse cohort of Cypriot patients, we were able to identify a new recombinant predisposing haplotype, DR4 –DQ2.3, carried by 13% of our patients, but only 4.7% of control subjects. This haplotype failed to reach significance in a large T1D Genetics Consortium investigating HLA-DR-DQ haplotypes in 607 Caucasian families and 38 Asian families [30]. A more recent study of the T1D Genetics Consortium analyzing more than 18,000 individuals of European descent also did not report this haplotype, with the authors discussing that the homogeneity of such populations may indeed limit the ability to interrogate rare alleles [31]. Further verification of the significance of the DR4 –DQ2.3 haplotype in disease susceptibility should be pursued in a larger cohort of Cypriot patients.

Recent studies have supported the significance of epitope analysis as an additional piece of the complex puzzle of deciphering autoimmune disease susceptibility [23,31,32], we thus attempted a similar analysis in our own cohort. The polymorphic residues of the HLA class II molecules are important not only for peptide binding but also interaction with the T cell receptor and CD4 as well as dimerization and stability of the heterodimer. Therefore, unlike previous studies, in our study we included residues outside of the binding pockets. Using the SKDM HLA Tool, an independent tool from ones used in previous studies, we were able to confirm the significance of a number of previously reported susceptibility and resistance epitopes, the vast majority found within the antigen-binding clefts of MHCII.

Susceptibility residue HLA-DQB A⁵⁷ and protective counterpart consisting of aspartic acid (D) in the same position were found to have the strongest association in agreement with previous reports [23,24,31]. This aspartic acid forms a salt bridge with a conserved arginine (R) at position 76 of HLA-DQA [28], and has been correlated with protection. In contrast, the presence of a non-charged amino acid at position 57, likely incapable of forming a salt bridge, predisposes to T1D [33,34]. In addition, absence of HLA-DQB D⁵⁷ in combination with HLA-DQA R⁵² has been associated with susceptibility [35]; the proximity of these residues to the interface of the dimer may affect the stability or structure of the dimer of heterodimers [28]. The most significantly associated HLA-DRB1 susceptibility residue identified in our study was Q⁷⁰ not previously associated with T1D. We were able to show that this residue alone discriminated between resistance and susceptibility HLA-DRB1 alleles. In addition to Q⁷⁰, we found residues HLA-DRβ V¹¹, H¹³ and L⁶⁷ that were previously reported to have the highest association with RA susceptibility, whereas D⁷⁰ strongly correlated with resistance [32]. The same study identified a two amino acid epitope QA^{70,74} associated with RA susceptibility. In addition to resistance residue HLA-DRβ D⁷⁰, we identified an alternative in the same position of the HLA-DRB1 allele, R⁷⁰, which was actually more protective. For the different epitopes a gene dose effect was also observed. In support of these findings, previously published work has shown a dose effect in the response of insulin B chain reactive T cells, with stronger responses shown in the presence of at least one non-risk DQB D⁵⁷ as compared to subjects lacking this epitope on both DQ alleles [36].

The importance of epitope analysis becomes apparent when considering closely related haplotypes or alleles with different risk determined by the presence of certain residues. For example, the closely related haplotypes DRB1*04:01-DQA1*03:01-DQB1*03:02 and DRB1*04:04-DQA1*03:01-DQB1*03:02 differ only at positions 71 (lysine vs. arginine) and 86

(glycine vs. valine) of DRB1; however, the former is highly predisposing whereas the latter haplotype is neutral [30]. Similarly, we showed that closely related alleles HLA-DQB1*03:02, which is risk associated, and HLA-DQB1*3:01, which is protective, differ at critical positions 26 and 57. In addition, sharing of epitopes by disparate alleles may explain disease association but also disease susceptibility in the absence of high risk alleles. For example, two distinct HLA molecules not closely related but both risk associated, HLA-DQA1 *03:01 and *05:01, shared an extended haplotype of 10 amino acid residues (DQA Y¹¹R⁵²R⁵⁵F⁶¹T⁶⁴I⁶⁶L⁶⁹V/L⁷⁶H¹²⁹E/K¹⁷⁵) all found to be significantly associated with disease susceptibility and all entirely different from all other DQA alleles suggesting that this constitutes a shared epitope for T1D. Finally, we observed that whilst the majority of our patient cohort carried the established DR3—DQ2.5 and DR4—DQ8 risk haplotypes, 13% carried one copy of the DR4—DQ2.3 haplotype and a small number of patients did not carry any of the known risk alleles. We were able to find that all patients however, even those that did not carry susceptibility haplotypes, carried identified susceptibility epitopes in their DRB1 locus and one or more in the DQA1 and DQB1 loci.

Our study reports for the first time a new haplotype, DR4 –DQ2.3 in T1D. In addition, our study lends further support to the significant role of certain HLA risk epitopes. HLA residues DRB Q⁷⁰, DQB L²⁶ and A⁵⁷ and a 10 amino acid epitope of DQA were identified to be the most significant in discriminating risk alleles. Of our patient cohort 92% were carriers of the DQA Y¹¹R⁵²R⁵⁵F⁶¹T⁶⁴I⁶⁶L⁶⁹V/L⁷⁶H¹²⁹E/K¹⁷⁵ with DRβ Q⁷⁰ and DQB L²⁶A⁵⁷, in contrast to 25% of our controls, suggesting that this extended HLA class II epitope haplotype is involved in the disease pathogenesis while other genetic factors may act as disease modifiers. Since these amino acids are implicated in functions other than antigen binding this may suggest contribution outside of the peptide groove and binding affinity to auto-antigens, to other allosteric sites that also hold important immune functions. Certainly, the contribution of either pocket or non-pocket residues to pathogenesis can only be proven by functional assays studying autoimmune TCR engagement with self-peptide/MHC. Most of the information currently available comes from studies of TCR interaction with foreign peptide however it is well-known that autoreactive T cell receptors engage peptide/MHC in configurations differently than those of pathogen responses [37,38]. Further structural data from self-reactive CD4 T cell receptors binding to self-peptide/MHC II complexes are needed to verify the implication of individual epitopes.

Materials and methods

Study population

A dataset of previously enrolled consented case (T1D patients) and control (individuals with no history of diabetes) subjects was used [19,39]. The case cohort consisted of 170 Greek Cypriot patients with a cut off age of disease onset set at 39 years (89 females, 81 males, mean age 11 years, age of onset: 0–8 years 59 patients, 9–13 years 60 patients, 14–39 51 patients). Patients over 40 years were considered as Latent Autoimmune Diabetes in Adults (LADA) and were not included in the final cohort. The inclusion criteria were: (i) the patient had to be of Greek-Cypriot origin, (ii) diagnosis of T1D was based on clinical (polyuria, polydipsia, weight loss) and laboratory findings (fasting blood sugar level >125mg/dl, glycosuria, ketonuria, frequent metabolic acidosis, absent C-peptide, elevated Hb A1c and in all cases GAD antibodies), (iii) insulin treatment at onset and thereafter, (iv) in case of affected siblings only one sibling was included in the study. Information regarding gender, date of birth, ethnicity, age of onset and family history of T1D were collected for each patient. The control group consisted of 192 healthy individuals of the same ethnic descent (78 females, 114 males). The study was reviewed

and approved by the Cyprus National Bioethics Committee. In the case of minors/children, a written informed consent was obtained from parents or legal guardians.

Genotyping of HLA class II loci

Genomic DNA was extracted from whole blood with the use of commercially available QIAGEN[®] genomic DNA extraction kit. High resolution HLA genotyping was performed by His-toGenetics at 4 classical major histocompatibility complex loci DRB1, DQA1, DQB1 and DPB1 using Next Generation Sequencing (NGS), as previously described [40].

HLA allele frequency and epitope analysis

For allele frequency and epitope analysis the SKDM HLA Tool beta was used [41], which can test for HLA allele differences between two populations and perform amino acid analysis by retrieving amino acid sequences. Highest polymorphism is found amongst residues lining the binding pockets of HLA molecules however in this analysis both pocket and non-pocket amino acid epitopes were investigated. Once primary associations are identified other parameters are determined such as zygosity, interaction and linkage disequilibrium among amino acid epitopes of the same HLA molecule or between HLA isotypes.

The SKDM output includes the difference (Delta) in frequency between case and control alleles for a particular locus. A corresponding odds ratio (OR) and a corrected p-value are also supplied. P-values are corrected by the number of distinct alleles present in cases and controls. A list of statistically significant residues as a table denoting the alleles (AAs) where a residue is present, its position (Pos) in the alignment, the single letter alias of the amino acid (AA), whether it is associated (Assoc) with cases (+) or controls (-), a p-value, a p-value corrected (p^{corr}) by the number of AA interrogated and an associated odds-ratio (OR). For each zygosity comparison, the OR is calculated by Haldane's modification of Woolf's method: $OR = [(a + \frac{1}{2})(d + \frac{1}{2})] / [(b + \frac{1}{2})(c + \frac{1}{2})]$, and the significance of its derivation from unity is estimated by Fisher's exact test. HLA epitopes with a corrected P value ≤ 0.001 were defined as statistically significant. In addition, odds ratio (OR) ≥ 1.5 determined susceptibility while OR ≤ 0.5 determined resistance.

Supporting information

S1 Table. Allele frequency analysis for HLA-DRB1. The HLA-DRB1 typing of the patient and control populations. The table includes presence in the population and frequency, allele number and frequency, delta difference between the T1D and CTL population frequencies, a corrected P-value and the Odds Ratio (OR).

(DOCX)

S2 Table. Allele frequency analysis for HLA-DQB1. The HLA-DQB1 typing of the patient and control populations. The table includes presence in the population and frequency, allele number and frequency, delta difference between the T1D and CTL population frequencies, a corrected P-value and the Odds Ratio (OR).

(DOCX)

S3 Table. Allele frequency analysis for HLA-DQA1. The HLA-DQA1 typing of the patient and control populations. The table includes presence in the population and frequency, allele number and frequency, delta difference between the T1D and CTL population frequencies, a corrected P-value and the Odds Ratio (OR).

(DOCX)

S4 Table. Allele frequency analysis for HLA-DPB1. The HLA-DPB1 typing of the patient and control populations. The table includes presence in the population and frequency, allele number and frequency, delta difference between the T1D and CTL population frequencies, a corrected P-value and the Odds Ratio (OR).

(DOCX)

S5 Table. HLA-DRB1 Pocket epitopes.

(DOCX)

S6 Table. HLA-DQB1 pocket epitopes.

(DOCX)

S7 Table. HLA-DQA1 pocket epitopes.

(DOCX)

S8 Table. HLA-DRB1 non-pocket epitopes.

(DOCX)

S9 Table. HLA-DRB1 non-pocket zygosity.

(DOCX)

S10 Table. HLA-DQB1 non-pocket epitopes.

(DOCX)

S11 Table. HLA-DQB1 non-pocket zygosity.

(DOCX)

S12 Table. HLA-DQA1 non-pocket epitopes.

(DOCX)

S13 Table. HLA-DQA1 non-pocket zygosity.

(DOCX)

Acknowledgments

We would like to express our gratitude to the Cyprus Diabetes Association (CDA) and the patients that agreed to be part of our research.

Author Contributions

Conceptualization: Petroula Gerasimou, Paul A. Costeas.

Data curation: Petroula Gerasimou, Vicky Nicolaidou.

Formal analysis: Petroula Gerasimou, Vicky Nicolaidou, Nicos Skordis, Michalis Picolos, Demetrios Monos, Paul A. Costeas.

Funding acquisition: Paul A. Costeas.

Investigation: Petroula Gerasimou, Vicky Nicolaidou, Nicos Skordis, Michalis Picolos, Demetrios Monos, Paul A. Costeas.

Methodology: Petroula Gerasimou, Vicky Nicolaidou, Nicos Skordis, Michalis Picolos, Demetrios Monos, Paul A. Costeas.

Project administration: Petroula Gerasimou, Paul A. Costeas.

Resources: Nicos Skordis, Michalis Picolos, Demetrios Monos, Paul A. Costeas.

Software: Demetrios Monos.

Supervision: Paul A. Costeas.

Validation: Petroula Gerasimou, Vicky Nicolaidou.

Visualization: Vicky Nicolaidou, Demetrios Monos, Paul A. Costeas.

Writing – original draft: Petroula Gerasimou, Vicky Nicolaidou.

Writing – review & editing: Petroula Gerasimou, Vicky Nicolaidou, Paul A. Costeas.

References

1. Eisenbarth GS. Banting lecture 2009: An unfinished journey: Molecular pathogenesis to prevention of type 1A diabetes. *Diabetes*. 2010. pp. 759–774. <https://doi.org/10.2337/db09-1855> PMID: 20350969
2. Diamond T, Group P. Incidence and trends of childhood Type 1 diabetes worldwide 1990–1999. *Diabet Med*. 2006; 23: 857–66. <https://doi.org/10.1111/j.1464-5491.2006.01925.x> PMID: 16911623
3. Green A, Gale EAM, Patterson CC. Incidence of childhood-onset insulin-dependent diabetes mellitus: the EURODIAB ACE study. *Lancet*. 1992; 339: 905–909. [https://doi.org/10.1016/0140-6736\(92\)90938-Y](https://doi.org/10.1016/0140-6736(92)90938-Y) PMID: 1348306
4. Casu A, Pascutto C, Bernardinelli L, Songini M. Type 1 diabetes among Sardinian children is increasing: The Sardinian diabetes register for children aged 0–14 years (1989–1999). *Diabetes Care*. 2004; 27: 1623–1629. <https://doi.org/10.2337/diacare.27.7.1623> PMID: 15220238
5. Bahillo M, Hermoso F, Ochoa C, García-Fernández J, Rodrigo J, Marugán J, et al. Incidence and prevalence of type 1 diabetes in children aged $0-15\text{ yr}$ in Castilla-Leon (Spain). *Pediatr Diabetes*. Blackwell Publishing Ltd; 2007; 8: 369–373. <https://doi.org/10.1111/j.1399-5448.2007.00255.x> PMID: 18036062
6. Skordis N, Theodorou S, Apsiotou T, Stavrou S, Herakleous E, Savva SC. The incidence of type 1 diabetes mellitus in Greek-Cypriot children and adolescents in 1990–2000. *Pediatr Diabetes*. 2002; 3: 200–204. <https://doi.org/10.1034/j.1399-5448.2002.30406.x> PMID: 15016148
7. Toumba M, Savva SC, Bacopoulou I, Apsiotou T, Georgiou T, Stavrou S, et al. Rising incidence of type 1 diabetes mellitus in children and adolescents in Cyprus in 2000–2004. *Pediatr Diabetes*. 2007; 8: 374–376. <https://doi.org/10.1111/j.1399-5448.2007.00262.x> PMID: 18036063
8. Skordis N, Efstathiou E, Kyriakides TC, Savvidou A, Savva SC, Phylactou L a, et al. Epidemiology of type 1 diabetes mellitus in Cyprus: rising incidence at the dawn of the 21st century. *Hormones (Athens)*. 2012; 11: 86–93. Available: <http://www.ncbi.nlm.nih.gov/pubmed/22450348>
9. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, Erlich HA, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat Genet*. 2009; 41: 703–707. <https://doi.org/10.1038/ng.381> PMID: 19430480
10. Concannon P, Erlich HA, Julier C, Morahan G, Nerup J, Pociot F, et al. Type 1 diabetes: Evidence for susceptibility loci from four genome-wide linkage scans in 1,435 multiplex families. *Diabetes*. 2005; 54: 2995–3001. <https://doi.org/10.2337/diabetes.54.10.2995> PMID: 16186404
11. Noble JA, Valdes AM. Genetics of the HLA region in the prediction of type 1 diabetes. *Curr Diab Rep*. 2011; 11: 533–542. <https://doi.org/10.1007/s11892-011-0223-x> PMID: 21912932
12. Noble JA, Valdes AM, Varney MD, Carlson JA, Moonsamy P, Fear AL, et al. HLA class I and genetic susceptibility to type 1 diabetes: Results from the type 1 diabetes genetics consortium. *Diabetes*. 2010; 59: 2972–2979. <https://doi.org/10.2337/db10-0699> PMID: 20798335
13. Noble JA, Valdes AM, Bugawan TL, Apple RJ, Thomson G, Erlich HA. The HLA class I A locus affects susceptibility to type 1 diabetes. *Hum Immunol*. 2002; 63: 657–664. [https://doi.org/10.1016/S0198-8859\(02\)00421-4](https://doi.org/10.1016/S0198-8859(02)00421-4) PMID: 12121673
14. Tait BD, Colman PG, Morahan G, Marchinovska L, Dore E, Gellert S, et al. HLA genes associated with autoimmunity and progression to disease in type 1 diabetes. *Tissue Antigens*. 2003; 61: 146–153. <https://doi.org/10.1034/j.1399-0039.2003.00013.x> PMID: 12694582
15. Valdes AM, Erlich HA, Noble JA. Human leukocyte antigen class I B and C loci contribute to Type 1 Diabetes (T1D) susceptibility and age at T1D onset. *Hum Immunol*. Elsevier; 2005; 66: 301–313. <https://doi.org/10.1016/J.HUMIMM.2004.12.001> PMID: 15784469
16. Nejentsev S, Howson JMM, Walker NM, Szeszkó J, Field SF, Stevens HE, et al. Localization of type 1 diabetes susceptibility to the MHC class I genes HLA-B and HLA-A. *Nature*. 2007; 450: 887–892. <https://doi.org/10.1038/nature06406> PMID: 18004301

17. Howson JMM, Walker NM, Clayton D, Todd JA. Confirmation of HLA class II independent type 1 diabetes associations in the major histocompatibility complex including HLA-B and HLA-A. *Diabetes, Obes Metab.* 2009; 11: 31–45. <https://doi.org/10.1111/j.1463-1326.2008.01001.x> PMID: 19143813
18. Silva HP V, Ururahy MAG, Souza KSC, Loureiro MB, Oliveira YMC, Oliveira GHM, et al. The association between the HLA-G 14-bp insertion/deletion polymorphism and type 1 diabetes. *Genes Immun.* 2016; 17: 13–18. <https://doi.org/10.1038/gene.2015.45> PMID: 26492519
19. Gerasimou P, Skordis N, Picolos M, Spyridonidis A, Costeas P. HLA-G 14-bp polymorphism affects the age of onset in Type I Diabetes Mellitus. *Int J Immunogenet.* 2016; 43: 135–142. <https://doi.org/10.1111/iji.12259> PMID: 27080982
20. Eike MC, Becker T, Humphreys K, Olsson M, Lie BA. Conditional analyses on the T1DGC MHC dataset: Novel associations with type 1 diabetes around HLA-G and confirmation of HLA-B. *Genes Immun.* 2009; 10: 56–67. <https://doi.org/10.1038/gene.2008.74> PMID: 18830248
21. Kovats S, Nepom GT, Coleman M, Nepom B, Kwok WW, Blum JS. Deficient antigen-presenting cell function in multiple genetic complementation groups of type II bare lymphocyte syndrome. *J Clin Invest.* 1995; 96: 217–223. <https://doi.org/10.1172/JCI118023> PMID: 7615790
22. Tollefsen S, Hotta K, Chen X, Simonsen B, Swaminathan K, Mathews II, et al. Structural and Functional Studies of trans -Encoded HLA-DQ2.3 (DQA1*03:01/DQB1*02:01) Protein Molecule. *J Biol Chem.* 2012; 287: 13611–13619. <https://doi.org/10.1074/jbc.M111.320374> PMID: 22362761
23. Roark CL, Anderson KM, Simon LJ, Schuyler RP, Aubrey MT, Freed BM. Multiple HLA epitopes contribute to type 1 diabetes susceptibility. *Diabetes.* 2014; 63: 323–331. <https://doi.org/10.2337/db13-1153> PMID: 24357703
24. Zerva L, Cizman B, Mehra NK, Alahari SK, Murali R, Zmijewski CM, et al. Arginine at positions 13 or 70–71 in pocket 4 of HLA-DRB1 alleles is associated with susceptibility to tuberculoid leprosy. *JExpMed.* 1996; 183: 829–836. <https://doi.org/10.1084/jem.183.3.829>
25. Sato a K, Sturniolo T, Sinigaglia F, Stern LJ. Substitution of aspartic acid at beta57 with alanine alters MHC class II peptide binding activity but not protein stability: HLA-DQ (alpha1*0201, beta1*0302) and (alpha1*0201, beta1*0303). *Hum Immunol.* 1999; 60: 1227–36. Available: <http://www.ncbi.nlm.nih.gov/pubmed/10626736> PMID: 10626736
26. Ettinger R a, Papadopoulos GK, Moustakas AK, Nepom GT, Kwok WW. Allelic variation in key peptide-binding pockets discriminates between closely related diabetes-protective and diabetes-susceptible HLA-DQB1*06 alleles. *J Immunol.* 2006; 176: 1988–1998. PMID: 16424231
27. Schafer PH, Pierce SK, Jardetzky TS. The structure of mhc class ii: A role for dimer of dimers. *Semin Immunol.* 1995; 7: 389–398. <https://doi.org/10.1006/smim.1995.0043> PMID: 8775464
28. Brown JH, Jardetzky TS, Gorga JC, Stern LJ, Urban RG, Strominger JL, et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature.* 1993; 364: 33–39. <https://doi.org/10.1038/364033a0> PMID: 8316295
29. Bondinas GP, Moustakas AK, Papadopoulos GK. The spectrum of HLA-DQ and HLA-DR alleles, 2006: A listing correlating sequence and structure with function. *Immunogenetics.* 2007; 59: 539–553. <https://doi.org/10.1007/s00251-007-0224-8> PMID: 17497145
30. Erlich H, Valdes AM, Noble J, Carlson JA, Varney M, Concannon P, et al. HLA DR-DQ haplotypes and genotypes and type 1 diabetes risk analysis of the type 1 diabetes genetics consortium families. *Diabetes.* 2008; 57: 1084–1092. <https://doi.org/10.2337/db07-1331> PMID: 18252895
31. Hu X, Deutsch AJ, Lenz TL, Onengut-Gumuscu S, Han B, Chen WM, et al. Additive and interaction effects at three amino acid positions in HLA-DQ and HLA-DR molecules drive type 1 diabetes risk. *Nat Genet.* 2015; <https://doi.org/10.1038/ng.3353> PMID: 26168013
32. Freed BM, Schuyler RP, Aubrey MT. Association of the HLA-DRB1 epitope LA 67, 74 with rheumatoid arthritis and citrullinated vimentin binding. *Arthritis Rheum.* 2011; 63: 3733–3739. <https://doi.org/10.1002/art.30636> PMID: 22094856
33. TODD JA, BELL JI, McDEVITT HO. HLA antigens and insulin-dependent diabetes. *Nature.* Nature Publishing Group; 1988; 333: 710–710. <https://doi.org/10.1038/333710a0> PMID: 3386713
34. Horn GT, Bugawan TL, Long CM, Erlich HA. Allelic sequence variation of the HLA-DQ loci: relationship to serology and to insulin-dependent diabetes susceptibility. *Proc Natl Acad Sci U S A.* 1988; 85: 6012–6. Available: <http://www.ncbi.nlm.nih.gov/pubmed/2842756> PMID: 2842756
35. Khalil I, d'Auriol L, Gobet M, Morin L, Lepage V, Deschamps I, et al. A combination of HLA-DQ beta Asp57-negative and HLA DQ alpha Arg52 confers susceptibility to insulin-dependent diabetes mellitus. *J Clin Invest.* 1990; 85: 1315–9. <https://doi.org/10.1172/JCI114569> PMID: 2318983
36. Nakayama M, McDaniel K, Fitzgerald-Miller L, Kiekhäfer C, Snell-Bergeon JK, Davidson HW, et al. Regulatory vs. inflammatory cytokine T-cell responses to mutated insulin peptides in healthy and type 1

- diabetic subjects. *Proc Natl Acad Sci.* 2015; <https://doi.org/10.1073/pnas.1502967112> PMID: [25831495](https://pubmed.ncbi.nlm.nih.gov/25831495/)
37. Deng L, Mariuzza RA. Recognition of self-peptide-MHC complexes by autoimmune T-cell receptors. *Trends Biochem Sci.* NIH Public Access; 2007; 32: 500–8. <https://doi.org/10.1016/j.tibs.2007.08.007> PMID: [17950605](https://pubmed.ncbi.nlm.nih.gov/17950605/)
 38. Stadinski BD, Zhang L, Crawford F, Marrack P, Eisenbarth GS, Kappler JW. Diabetogenic T cells recognize insulin bound to IA2 in an unexpected, weakly binding register. *Proc Natl Acad Sci.* 2010; 107: 10978–10983. <https://doi.org/10.1073/pnas.1006545107> PMID: [20534455](https://pubmed.ncbi.nlm.nih.gov/20534455/)
 39. Castelli EC, Gerasimou P, Paz MA, Ramalho J, Porto IOP, Lima THA, et al. HLA-G variability and haplotypes detected by massively parallel sequencing procedures in the geographically distinct population samples of Brazil and Cyprus. *Mol Immunol.* 2017; 83: 115–126. <https://doi.org/10.1016/j.molimm.2017.01.020> PMID: [28135606](https://pubmed.ncbi.nlm.nih.gov/28135606/)
 40. Cereb N, Kim HR, Ryu J, Yang SY. Advances in DNA sequencing technologies for high resolution HLA typing. *Human Immunology.* 2015. pp. 923–927. <https://doi.org/10.1016/j.humimm.2015.09.015> PMID: [26423536](https://pubmed.ncbi.nlm.nih.gov/26423536/)
 41. Kanterakis S, Magira E, Rosenman KD, Rossman M, Talsania K, Monos DS. SKDM human leukocyte antigen (HLA) tool: A comprehensive HLA and disease associations analysis software. *Hum Immunol.* 2008; 69: 522–525. <https://doi.org/10.1016/j.humimm.2008.05.011> PMID: [18601964](https://pubmed.ncbi.nlm.nih.gov/18601964/)