

High-resolution HLA genotyping in inclusion body myositis refines 8.1 ancestral haplotype association to DRB1*03:01:01 and highlights pathogenic role of arginine-74 of DRβ1 chain

Nataliya Slater^a, Anuradha Sooda^a, Emily McLeish^a, Kelly Beer^{a,b}, Anna Bruschi^e, Rakesh Shakya^e, Christine Bundell^e, Ian James^{b,f}, Abha Chopra^f, Frank L. Mastaglia^{b,g}, Merrilee Needham^{a,b,c,d}, Jerome D. Coudert^{a,b,c,*}

^a Murdoch University, Centre for Molecular Medicine and Innovative Therapeutics, Murdoch, WA, Australia

^b Perron Institute for Neurological and Translational Science, Nedlands, WA, Australia

^c University of Notre Dame Australia, School of Medicine, Fremantle, WA, Australia

^d Fiona Stanley Hospital, Department of Neurology, Murdoch, WA, Australia

^e PathWest Laboratory Medicine, Dept of Clinical Immunology, QEII Medical Centre, Nedlands, WA, Australia

^f Murdoch University, Institute for Immunology and Infection Diseases, Murdoch, WA, Australia

^g University of Western Australia, Centre for Neuromuscular & Neurological Disorders, Crawley, WA, Australia

ARTICLE INFO

Handling Editor: Dr. M.E. Gershwin

Keywords:

Inclusion body myositis
Human leukocyte antigens
Autoimmunity genetic association
Next generation sequencing

ABSTRACT

Objectives: Inclusion body myositis (IBM) is a progressive inflammatory-degenerative muscle disease of older individuals, with some patients producing anti-cytosolic 5'-nucleotidase 1A (NT5C1A, aka cN1A) antibodies. Human Leukocyte Antigens (HLA) is the highest genetic risk factor for developing IBM. In this study, we aimed to further define the contribution of HLA alleles to IBM and the production of anti-cN1A antibodies.

Methods: We HLA haplotyped a Western Australian cohort of 113 Caucasian IBM patients and 112 ethnically matched controls using Illumina next-generation sequencing. Allele frequency analysis and amino acid alignments were performed using the Genentech/MiDAS bioinformatics package. Allele frequencies were compared using Fisher's exact test. Age at onset analysis was performed using the ggstatsplot package. All analysis was carried out in RStudio version 1.4.1717.

Results: Our findings validated the independent association of HLA-DRB1*03:01:01 with IBM and attributed the risk to an arginine residue in position 74 within the DRβ1 protein. Conversely, DRB4*01:01:01 and DQA1*01:02:01 were found to have protective effects; the carriers of DRB1*03:01:01 that did not possess these alleles had a fourteenfold increased risk of developing IBM over the general Caucasian population. Furthermore, patients with the abovementioned genotype developed symptoms on average five years earlier than patients without. We did not find any HLA associations with anti-cN1A antibody production.

Conclusions: High-resolution HLA sequencing more precisely characterised the alleles associated with IBM and defined a haplotype linked to earlier disease onset. Identification of the critical amino acid residue by advanced biostatistical analysis of immunogenetics data offers mechanistic insights and future directions into uncovering IBM aetiopathogenesis.

Financial support

This project was supported by a Spinnaker Health Research Foundation grant, a Biobank interim support program grant, a Research Excellence Award from the Western Australian Future Health Research & Innovation Fund, and a bequest donated by a late IBM patient "donor

name remains confidential". Nataliya Slater is the recipient of a Murdoch University Research Training Programme Scholarship and a Byron Kakulas Scholarship funded by the Perron Institute for Neurological and Translational Science.

The authors have no financial support or other benefits from commercial sources for the work reported in the manuscript to declare.

* Corresponding author. Murdoch University, CMMIT, 390 Discovery Way Murdoch, WA 6150 Australia.

E-mail address: jerome.coudert@murdoch.edu.au (J.D. Coudert).

<https://doi.org/10.1016/j.jaut.2023.103150>

Received 28 September 2023; Received in revised form 1 November 2023; Accepted 15 November 2023

Available online 2 December 2023

0896-8411/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Inclusion body myositis (IBM) is a chronic inflammatory disease of skeletal muscles affecting middle-aged and older adults. It has a distinctive pattern of muscle involvement, primarily impacting anterior limb muscles, including quadriceps femoris, tibialis anterior and finger flexors, leading to a gradual decline in strength and mobility. More than two-thirds of patients experience impaired swallowing due to pharyngeal muscle involvement, which when severe may result in nutritional deficiencies and recurring aspiration pneumonia [1]. A 2017 meta-analysis documented a worldwide prevalence of 24.8 cases per million individuals with a growing trend over time, possibly attributed to increasing disease awareness and improved diagnosis [2]. Despite the technological advances in diagnostic methods, IBM patients still experience, on average, a five-year delay between symptom onset and diagnosis [3,4]. Furthermore, there are currently no effective disease-modifying treatments for IBM, and misdiagnosis may result in the prescription of inadequate and potentially deleterious immunosuppressive therapies [1,5].

IBM pathology is complex and encompasses a combination of myofibre degeneration with disrupted proteostasis, mitochondrial abnormalities, and autoimmune manifestations. While the immune response is dominated by CD8⁺ T cells invading non-necrotic myofibers [6], infiltrates of CD4⁺ T cells and B cells have also been reported [7]. Additionally, the humoral autoimmune response features the production of antibodies against a muscle enzyme, cytosolic 5'-nucleotidase 1A (NT5C1A, aka cN1A), reported in 33–72 % of patients [8–12]. The interaction between T lymphocytes and antigenic peptides presented by major histocompatibility complex (MHC) molecules underpins adaptive cellular immune responses, including those against self. In humans, the MHC gene cluster is located on the short arm of chromosome 6 (6p21.3). It includes highly polymorphic alleles encoding Class I and Class II human leukocyte antigen (HLA) molecules that mediate the presentation of antigenic peptides to CD8⁺ and CD4⁺ T cells, respectively. HLA alleles are found at variable prevalence in different ethnicities. Indeed, some combinations of alleles are preferentially co-inherited and occur at non-random frequencies due to linkage disequilibrium (LD) [13]. The human MHC cluster exhibits a high degree of LD between HLA gene loci, resulting in extended lengths of conserved DNA sequences denoted as ancestral haplotypes (AH) [14].

Since the 1990s, genetic association studies have consistently reported the association of HLA-DRB1*03:01 (originally reported as the HLA-DR3 serotype [15]) and the 8.1 ancestral haplotype (8.1 AH) with IBM in Caucasian individuals [16–20]. There have also been reports linking HLA-DRB1*01:01 [20,21], A*03 [22], and DQB1*05 [22] to the disease. Notably, prior studies were constrained by their use of serological or genotypic data with limited resolution, which provided only a broad characterisation of allele groups. More recently, Rothwell and colleagues applied GWAS analysis to Caucasian IBM patients sourced from 11 countries through the MYOGEN consortium; HLA allele imputation from SNP polymorphisms showed independent associations with three alleles: HLA-DRB1*03:01, *01:01 and *13:01 [23]. However, the identity of secondary risk alleles, and of protective alleles, has yet to be confirmed in other populations using higher resolution genotyping methods.

Our case-control study aimed to further refine IBM genetic associations within the 8.1 AH by applying high-resolution direct sequencing of HLA gene loci. We employed 3-field-resolution HLA genotyping enabled by Illumina next-generation sequencing, focusing specifically on individuals of Caucasian ethnic ancestry who constitute over 98 % of people with IBM in Western Australia.

2. Patients and methods

2.1. Participant recruitment

Study participants were recruited through the myositis clinics at Murdoch University and Perron Institute for Neurological and Translational Science in Perth, Western Australia. A total of one hundred and thirteen Caucasian IBM patients and one hundred and twelve ethnically matched controls donated blood or saliva samples for the study after providing informed consent. Participants' ethnicity as Caucasian was self-reported (N = 167) or reported by treating neurologist or nurse (N = 58). The term Caucasian relates to white individuals of European descent; however, the sub-populations making-up Caucasians vary significantly between geographical locations; therefore, we matched our Australian cohort of Caucasian IBM patients with a control Caucasian cohort from the same region in order to limit the possibility of genetic variations between the two groups due to a different composition of their ancestry. To our knowledge, none of the study participants were related by birth. All the patients were diagnosed by an experienced neuromuscular neurologist as having IBM based on history, clinical examination, and where available, serological findings. Many patients underwent a confirmatory biopsy, (62 % could be traced), and diagnosis confirmed in accordance with the ENMC criteria [23]. Some patients were unable or unwilling to have biopsies, and some were not available for review. Age at disease onset was self-reported; patients were asked to recollect their initial symptoms and the year in which they first became aware of them.

All participants provided informed consent prior to enrolment to the study and sample donation. The research protocol was evaluated and approved by the Murdoch University Human Research Ethics Committee (projects 2015/111 and 2020/188).

2.2. DNA extraction

DNA was extracted from either blood or saliva. Venous blood was collected into lithium heparin-coated tubes (BD Vacutainer, Becton Dickinson, Australia). The buffy coat was separated by centrifugation and stored at –80 °C until processing. DNA was extracted from the buffy coat using a QIAamp DNA Blood It (QIAGEN, Hilden, Germany) according to the manufacturer's protocol.

Saliva samples were self-collected using Oragene DNA collection kits (DNA Genotek, Ontario, Canada) and stored at ambient temperature until processing. DNA was extracted from saliva using PrepIT-L2P kit (DNA Genotek, Ontario, Canada) according to the manufacturer's protocol. The concentration and purity of the extracted DNA were measured using a Nanodrop spectrophotometer at 260/280 nm and by resolving DNA aliquots on a 1 % agarose gel.

2.3. HLA genotyping

HLA genotyping was conducted by Illumina next-generation sequencing at the Institute for Immunology and Infectious Diseases, Murdoch University, which is accredited by the American Society for Histocompatibility and Immunogenetics (ASHI) and the National Association of Testing Authorities (NATA). The validated sequencing protocol using locus-specific PCR amplification of genomic DNA has been described in Currenti et al. [24]. Briefly, polymorphic regions of HLA Class I A, B, C (exons 2, 3) and Class II DRB1, 3, 4, 5 (exons 2, 3), DQB1 (exons 2, 3), DQA1 (exon 2), and DPB1 (exon 2) were PCR amplified using sample-specific Molecular IndexED (MID) primers. Amplicons were quantified, normalised, and pooled in equimolar ratios. Sequencing libraries were created and quantified using the Jetseq qPCR Library Quantification Kit (Meridian Bioscience Inc., OH, USA). Samples were sequenced on an Illumina MiSeq platform using the MiSeq V3 600-cycle kit (2 × 300 base pair reads) (Illumina Inc., CA, USA). Reads were quality-filtered, separated by MID tags and passed through an

in-house accredited HLA allele caller software. Alleles were called according to the latest IPD-IMGT/HLA nomenclature [25]. Allele identity assignment was based on G groups that combine sub-alleles with identical nucleotide sequences across the sequenced exons but polymorphism across the other exons. Additionally, some DQB1 and DRB1, 3, 4, 5 alleles were assigned an i group. These alleles were identical across the sequenced exons 2 and 3 but did not have reference data in the IPD-IMGT database for exon 3. For ease of analysis and in-text reporting, the G or i group designations have been removed; however, the full names of the alleles identified in this study featuring the G and i designations are listed in [Supplementary Table 1](#), and the corresponding sub-allele list can be reconstructed using the HLA ambiguity tables located at <https://www.iiid.com.au/laboratory-testing> (v.3.32.0 to v.3.51.0). The data presented for the HLA allele sequences are deposited in the NCBI Sequence Read Archive (SRA) repository with the accession number PRJNA1042617. <http://www.ncbi.nlm.nih.gov/bioproject/1042617>.

2.4. Anti-cN1A ELISA

IBM patients' blood was collected by venepuncture into polymer-gel SST tubes (BD Vacutainer). The serum phase was separated by centrifugation and frozen at -80°C until further processing. Serum anti-cN1A antibodies were analysed by a semi-quantitative ELISA adapted from Bundell et al. [26]. 96-well plates (Maxisorp, Nunc, Roskilde, Denmark) were coated with 50 μL /well of full-length cN1A protein (GenScript, NJ, SA) diluted to 10 $\mu\text{g}/\text{mL}$ in carbonate-bicarbonate buffer (pH 9.6) for 2 h at ambient temperature. Wells were washed with PBS/0.1 % Tween (PBST) and saturated with blocking buffer (PBST/5 % skim milk powder) overnight at 4°C . After washing with PBST, 100 μL of a patient serum diluted to 1:1000 in blocking buffer was added in duplicate and incubated for 2 h at ambient temperature. Wells were washed with PBST before adding horseradish peroxidase (HRP)-conjugated anti-human secondary antibodies directed against pan IgG/M/A or IgG or IgM or IgA (Invitrogen, Rockford, IL, USA) and incubating for 1 h at ambient temperature. Wells were washed again before incubation with 50 μL per well TMB solution (ThermoFisher Scientific, Waltham, MA, USA) for 10 min. The reaction was stopped with 50 μL of 2 M H_2SO_4 solution. The absorbance at 450 nm was read using the FLUOstar Omega (BMG Labtech, Mornington, VIC, Australia) microplate reader. Each plate included a positive control of anti-cN1A antibodies purified from a seropositive patient's serum. Pooled sera of forty-three healthy volunteers diluted 1:1000 in blocking buffer was used as baseline control. Blank duplicates were obtained by performing all the steps in the absence of a sample. All absorbance values were adjusted by the average blank value. Absorbance obtained for patients' samples was recorded as a fold change relative to baseline control. Reference cut-off values corresponded to the 99th percentile of the healthy samples (Busselton Population Health Study participants, $n = 190$). Patients with increased value for at least one antibody isotype were considered seropositive.

2.5. Statistical analysis

HLA allele frequencies and amino acid position differences were analysed using Genentech/MiDAS package v.1.1.0 [27]. Genotype frequencies were compared using Fisher's exact test. P values were adjusted using the Bonferroni method (denominated p_a) and considered statistically significant at $p_a < 0.05$; odds ratios and 95 % confidence intervals (CI) were calculated. Independent associations were computed using stepwise conditional regression in MiDAS. Age at onset comparison was performed using ggstatsplot package v.0.8.0 [28]. Analysis was carried out in RStudio v.1.4.1717 [29].

3. Results

3.1. HLA-DRB1*03:01:01, DRB4*01:01:01 and DQA1*01:02:01 are independently associated with IBM

The MiDAS package supports the HLA allele frequency distribution analysis under multiple inheritance models. Within the "additive" model, each instance of an allele is counted as an individual event. Conversely, the "dominant" model treats both heterozygous and homozygous carriers as a single event. Since it is yet to be established whether the effects of HLA alleles in IBM are dose-dependent, meaning that the impact of an allele is more pronounced in homozygous than in heterozygous carriers, the "additive" inheritance model was applied for our initial analysis. Subsequently, the analysis was reiterated using the "dominant" inheritance model, and the outcomes were compared.

Comparison of the allele frequency distribution between IBM patients and controls under the "additive" inheritance model confirmed a strong positive association with the alleles of the 8.1 AH, specifically HLA-DRB1*03:01:01, DQB1*02:01:01, DQA1*05:01:01, DRB3*01:01:02, and B*08:01:01. Additionally, we identified some alleles with significantly decreased frequencies in IBM. They included DRB4*01:01:01, DQA1*02:01:01, DRB1*07:01:01, and DQA1*03:01:01. A complete table of results is shown in [Supplementary Table 1](#).

To account for the strong LD within the MHC region, we repeated the analysis while stepwise conditioning upon the most strongly associated variable or variables added to the model as covariates until the model fit could no longer be improved. As a result, only HLA-DRB1*03:01:01 ($p_a = 1.516 \times 10^{-8}$, OR = 6.187) of the 8.1 AH retained its significant positive association with the disease, while the association of the remaining alleles was no longer statistically significant. However, the independent negative association of DRB4*01:01:01 ($p_a = 2.371 \times 10^{-4}$, OR = 0.225) and DQA1*01:02:01 ($p_a = 3.463 \times 10^{-3}$, OR = 0.246) were confirmed ([Table 1](#)).

Allele frequency comparison was then carried out using the MiDAS "dominant" inheritance model combined with the conditional analysis. The results were comparable to those obtained under the "additive" model ([Supplementary Tables 2 and 3](#)).

3.2. Arginine-74 within the DR β 1 protein confers the associated risk, while glutamine at this position is protective

Next, we sought to locate the risk attributed to HLA-DRB1*03:01:01 at the level of a specific amino acid residue. To achieve this, we employed the MiDAS "hla_aa" function, which imports allele amino acid sequences from the IPD-IMGT/HLA database and aligns them, enabling the identification of variable residues. The effect of each residue is then estimated using the likelihood ratio test. [Supplementary Table 4](#) contains the complete list of the variable amino acid residues identified from our data set. As before, we subjected the findings to stepwise conditional analysis, identifying amino acid in position 74 encoded within the DRB1 locus as highly significant ($p_a = 1.08 \times 10^{-11}$) with a likelihood ratio of 69.40 ([Table 2](#)).

Given the significance of DRB1_74, we looked at the effect sizes of all residues in that position ([Table 3](#)). We found that the frequency of arginine (R) was significantly increased in IBM patients (38.05 % vs 13.84 %, $p_a = 1.547 \times 10^{-9}$, OR = 6.159), whereas the frequency of glutamine (Q) was significantly decreased (1.77 % vs 13.84 %, $p_a = 1.050 \times 10^{-4}$, OR = 0.096) compared to the control group.

Finally, we investigated which DRB1 alleles encoded the arginine and glutamine at position 74 of the β 1 chain. Using MiDAS alignment, we fine-mapped arginine-74 to HLA-DRB1*03:01:01 and DRB1*03:02:01 and glutamine-74 to DRB1*07:01:01. This finding was consistent with our earlier analysis, where we observed a significant association of both DRB1*03:01:01 and DRB1*07:01:01 with IBM, even though DRB1*07:01:01 association was lost in the conditional analysis ([Supplementary Table 1](#)). A significant role of DRB1*03:02:01 was

Table 1
HLA alleles independently associated with IBM after stepwise conditional analysis.

HLA allele	Allele Frequency (%)		P	P _a	OR [95 % CI]	Covariates
	IBM (n = 226)	Controls (n = 224)				
DRB1*03:01:01	37.61	13.84	3.30×10^{-10}	6.263×10^{-8}	6.187 [3.554–11.094]	DRB1*03:01:01 DRB1*03:01:01+DRB4*01:01:01
DRB4*01:01:01	7.96	32.14	5.27×10^{-6}	9.957×10^{-4}	0.225 [0.327–0.116]	
DQA1*01:02:01	9.73	21.88	7.87×10^{-5}	0.015	0.246 [0.119–0.482]	

Abbreviations: n – number of alleles, P – p value, P_a – adjusted p value (method = “Bonferroni”), OR – odds ratio, CI – confidence interval.

excluded as it was only present in one heterozygous patient.

3.3. Carriage of protective alleles DRB4*01:01:01 and DQA1*01:02:01 negates the risk of HLA-DRB1*03:01:01

To investigate the potential risk-modifying effects of the protective alleles DRB4*01:01:01 and DQA1*01:02:01, we compared the frequency of HLA-DRB1*03:01:01 carriers who also possessed either one or none of these protective alleles by applying Fisher’s exact test. Our findings indicated that individuals carrying HLA-DRB1*03:01:01 but lacking both DRB4*01:01:01 and DQA1*01:02:01 faced a fourteen-fold increased risk of developing IBM over the general population (52.21 % of IBM vs 7.14 % of controls, $p_a = 2.021 \times 10^{-13}$, OR = 14.025) (Table 4). Moreover, carriage of at least one protective allele negated the risk effects of DRB1*03:01:01 (19.47 % of IBM vs 19.64 % of controls, $p = 1.00$, OR = 0.989). Additionally, only one patient and one control individual in our cohort possessed all three alleles, while most of the DRB1*03:01:01 carriers had either DRB4*01:01:01 or DQA1*01:02:01, suggesting that either one of these alleles is sufficient to exert the protective effect.

3.4. No HLA association with anti-cN1A serostatus

Our cohort comprised ninety-one IBM patients with known anti-cN1A serostatus, twenty-six (28.6 %) of whom were seropositive. Additionally, we tested thirty control donors among whom one individual (3.3 %) was determined to be seropositive. The comparison of HLA allele frequencies between the anti-cN1A seropositive and seronegative subgroups utilising both the “additive” and the “dominant” inheritance models did not identify statistically significant differences after P-value adjustment (Supplementary Tables 5 and 6).

3.5. Carriage of risk-associated HLA genotype correlates with earlier IBM onset

Ninety-four patients in our IBM cohort had age-at-onset data available (mean 63.7 ± 10.51 years). Welch’s t-test was employed to investigate if carriage of either the HLA-DRB1*03:01:01, DRB4*01:01:01, DQA1*01:02:01 or DRB1*03:01:01⁺/DRB4*01:01:01[−]/DQA1*01:02:01[−] genotype was associated with age at disease onset. In agreement with two previous studies [23,30], we found no association between carriage of individual HLA alleles and age at disease onset (Supplementary Fig. 1). Furthermore, we could not confirm the finding of earlier disease onset in carriers of the DRB1*03/*01 diplotype that had been previously reported in Australian IBM patients [30] (Supplementary Fig. 1). We did, however, discover that patients carrying DRB1*03:01:01⁺ but lacking both DRB4*01:01:01 and DQA1*01:02:01 were significantly more likely to develop disease symptoms earlier in life (62 vs 67 years, $p = 0.036$) (Fig. 1).

4. Discussion

Herein we present what we believe to be the first case-control study utilising high resolution next-generation sequencing to investigate HLA association with IBM in Caucasians to identify alleles associated with the

disease. NGS directly sequences the HLA genes, reducing the reliance on imputation methods which have been used in recent GWAS studies [23]; it provides higher genotyping resolution and prevents possible inaccurate calls for low-frequency and rare alleles [31]. In agreement with previous genomic analyses, we observed higher frequencies of the alleles linked within the 8.1 ancestral haplotype among IBM patients compared to the reference group of the same ethnicity. However, conditional statistical analysis revealed that only HLA-DRB1*03:01:01 independently conferred an increased risk of developing the disease. This finding corroborates conclusions from the previous multi-national GWAS study by Rothwell and colleagues, who noted the loss of significance within the 8.1 AH when conditioned upon the strong disease association of DRB1*03:01 [23]. However, we failed to confirm the association with DRB1*13:01 and *01:01 reported in their study, possibly because of the variable population mix of Caucasians from different European countries and the USA in their heterogeneous control group, compared with our Australian Caucasian control cohort. Applying “additive” and “dominant” inheritance models to our analysis led to comparable outcomes indicating the absence of a discernible dose-dependent impact of the reported alleles on IBM pathology.

It is important to note that the Illumina next-generation sequencing technology utilised in our study detects alleles based on their highly polymorphic regions that encode the extracellular antigen-presenting domains of MHC proteins and not on the complete genetic sequence. Thus, there is potential for ambiguity among alleles sharing identical sequences in exons 2 and 3 (exon 2 only for DQA1 and DPB1) while differing in the more conserved exons. For instance, the HLA-DRB1*03:01:01i group encompasses over fifty distinct alleles (for a comprehensive allele list, refer to <https://www.iid.com.au/laboratory-testing>). However, in our Caucasian cohort, the true HLA-DRB1*03:01:01 allele is the most probable due to its notably higher prevalence in this population. On this account, we investigated which amino acid residues contributed to the risk associated with the DRB1*03:01:01i alleles. Our analysis pinpointed the risk association to the presence of arginine in position 74. In contrast, glutamine in this position appears to protect against the disease as this residue is under-represented in IBM patients. Curiously, arginine-74 has also been implicated in the pathogenesis of myopathies other than IBM [19]. For instance, the GWAS study on the MYOGEN cohort has resulted in the identification of arginine-74, as well as asparagine-77, both located in the binding groove of the DRβ1 protein, as highly associated with idiopathic inflammatory myopathies (IIM) as a combined group of diseases [23]. Analysis of IIM subsets, however, reported more heterogeneous findings. In the same study, residue 74 of HLA-DRB1*03 was found associated with polymyositis but not with dermatomyositis (combined juvenile [JDM] and adult-onset [DM]), where an independent association of residue 57 in HLA-DQB1 was found [23]. In contrast, a more recent study focused on a cohort of DM patients, reported association of residue-74 with adult-onset DM while a stronger association of residue-37 and only a weak association of residue-74 with JDM [32]. The authors suggested that this difference may be driven by the association of this residue with anti-Jo-1, anti-PM/Scl and anti-cN1A antibodies which are more prevalent in adult patients [33]. Of note, we could not confirm the genetic association with anti-cN1A antibodies within our cohort, possibly due to the smaller size of our cohort resulting in diminished analysis sensitivity.

Table 2

Amino acid positions within the HLA molecules found independently associated with IBM after stepwise conditional analysis.

Amino acid position	Residues	DF	LR	P	P _a	Covariates
DRB1_74	R, A, Q, L	4	69.40	3.03 x 10⁻¹⁴	1.08 x 10⁻¹¹	
DQB1_30	S, H	2	15.69	3.91 x 10 ⁻⁴	0.138	DRB1_74
DRB3_38	L, V, A	3	17.89	4.64 x 10 ⁻⁴	0.164	DRB1_74 + DQB1_30
C_113	Y	1	7.77	5.32 x 10 ⁻³	1.00	DRB1_74 + DQB1_30 + DRB3_38 + DPB1_76 + DRB1_86 + DRB1_-1
DPB1_76	M, V	2	10.03	6.63 x 10 ⁻³	1.00	DRB1_74 + DQB1_30 + DRB3_38
DRB1_-1	A, S	2	9.71	7.80 x 10 ⁻³	1.00	DRB1_74 + DQB1_30 + DRB3_38 + DPB1_76 + DRB1_86
B_65	Q, R	2	9.02	1.10 x 10 ⁻²	1.00	DRB1_74 + DQB1_30 + DRB3_38 + DPB1_76 + DRB1_86 + DRB1_-1 + C_113 + C_91
C_94	T	1	6.35	1.17 x 10 ⁻²	1.00	DRB1_74 + DQB1_30 + DRB3_38 + DPB1_76 + DRB1_86 + DRB1_-1 + C_113 + C_91 + B_65 + B_103
DRB1_86	V	1	6.33	1.19 x 10 ⁻²	1.00	DRB1_74 + DQB1_30 + DRB3_38 + DPB1_76
B_103	V	1	5.92	1.50 x 10 ⁻²	1.00	DRB1_74 + DQB1_30 + DRB3_38 + DPB1_76 + DRB1_86 + DRB1_-1 + C_113 + C_91 + B_65
DRB1_78	Y	1	4.77	2.90 x 10 ⁻²	1.00	DRB1_74 + DQB1_30 + DRB3_38 + DPB1_76 + DRB1_86 + DRB1_-1 + C_113 + C_91 + B_65 + B_103 + C_94
B_143	T	1	4.70	3.01 x 10 ⁻²	1.00	DRB1_74 + DQB1_30 + DRB3_38 + DPB1_76 + DRB1_86 + DRB1_-1 + C_113 + C_91 + B_65 + B_103 + C_94 + DRB1_78
C_91	G	1	4.30	3.81 x 10 ⁻²	1.00	DRB1_74 + DQB1_30 + DRB3_38 + DPB1_76 + DRB1_86 + DRB1_-1 + C_113

Abbreviations: DF – degrees of freedom, LR – likelihood ratio, P – p value, P_a – adjusted p value (method = “Bonferroni”).

Table 3

Comparison of amino acid frequencies at the DRB1_74 residue between IBM patients and control cases.

Amino acid residue	Residue frequency (%)		P	P _a	OR [95 % CI]
	IBM (n = 226)	Controls (n = 224)			
DRB1_74_R	38.05	13.84	3.095 x 10⁻¹⁰	1.547 x 10⁻⁹	6.159 [3.548–11.025]
DRB1_74_Q	1.77	13.84	2.100 x 10⁻⁵	1.050 x 10⁻⁴	0.096 [0.028–0.254]
DRB1_74_E	1.33	6.25	0.025	0.124	0.244 [0.056–0.721]
DRB1_74_L	2.21	4.46	0.184	0.920	0.472 [0.143–1.377]
DRB1_74_A	56.64	61.61	0.244	1.000	0.782 [0.515–1.180]

Abbreviations: n – number of alleles, P – p value, P_a – adjusted p value, OR – odds ratio, CI – confidence interval.

Table 4

Effect of protective alleles DRB4*01:01:01 and DQA1*01:02:01 in HLA-DRB1*03:01:01 carriers.

HLA genotype	Allele Carriers (%)		P	OR [95 % CI]
	IBM (n = 113)	Controls (n = 112)		
DRB1*03:01:01 without protective alleles	52.21	7.14	2.887 x 10 ⁻¹⁴	14.025 [6.105–36.501]
DRB1*03:01:01 with protective alleles	19.47	19.64	1	0.989 [0.484–2.019]

Abbreviations: n – number of carriers, P – p value, OR – odds ratio, CI – confidence interval.

Additionally, arginine-74 has been reported in association with various other autoimmune conditions, including rheumatoid arthritis [34], Grave’s disease [35,36], type I diabetes [36]. Extensive efforts have been dedicated to unravelling the mechanism by which it contributes to the risk of autoimmunity. Structural modelling of the DRβ1 chain by Miglioranza Scavuzzi and colleagues indicated that amino acid 74 resides within the peptide-binding pocket, and the substitution of a neutrally charged alanine or glycine with a positively charged arginine significantly altered the protein three-dimensional structure [37]. This alteration facilitated more efficient binding of self-peptides and their presentation to CD4⁺ T cells [38]. Building on those findings, Lee and colleagues screened a library of 150,000 small molecules to identify candidates that may disrupt the presentation by HLA-DRβ1-Arg74 of a self-antigen derived from thyroid stimulating hormone receptor (TSHR) that is responsible for autoimmune thyroid disease [38]. Intriguingly, the group identified one such molecule, Cepharanthine, which specifically out-competed the binding of a TSHR-derived self-peptide to

DRβ1-Arg74. In an experimental mouse model of Grave’s disease induced upon presentation of TSHR-derived peptide, Cepharanthine demonstrated therapeutic efficacy by reducing T cell activation and downstream production of pro-inflammatory cytokines [38].

Furthermore, an antigen-independent mechanism has been proposed in Systemic Lupus Erythematosus (SLE) where DRB1*03:01:01 is also a known risk driver [39]. The authors reported that a peptide that includes residues 65–79 of HLA-DRB1*03:01:01 could directly bind non-MHC receptors induced in the presence of IFN-γ in both human and mouse macrophages, initiating a signalling cascade, which resulted in mitochondrial perturbations, cell necrosis, protein unfolding and led to anti-dsDNA antibody production.

It is puzzling that HLA class II alleles are uncovered in association with IBM, a disease characterised by increased circulating and muscle-infiltrating cytotoxic CD8⁺ T cells [6,40,41]. While the current paradigm places the role of HLA class II on presenting exogenous antigens to CD4⁺ T cells [42], there is emerging evidence that class II molecules can

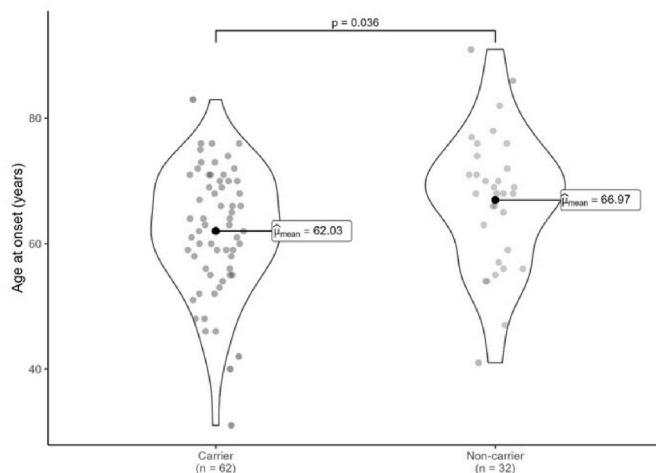


Fig. 1. Comparison of patients' age at IBM symptoms onset based on carriage of risk-associated DRB1*03:01:01/DRB4*01:01:01-/DQA1*01:02:01- genotype. The analysis was performed using Welch's *t*-test. The plot was generated using the ggstatsplot R package version 0.8.0.

play alternative roles in the inflammatory response, such as presenting viral antigens to CD8⁺ T cells [43,44] and activating NK cells via the Nkp44 receptor [45].

Our study also identified DRB4*01:01:01 and DQA1*01:02:01 as independently under-represented in the patient group and, thereby, presumably protective in IBM. Previous studies have reported an inverse correlation between the occurrence of the disease and the carriage of DRB4 alleles [46]. Our study provides a more precise understanding of this association by demonstrating that individuals carrying the DRB1*03:01:01 but lacking the DRB4*01:01:01 and DQA1*01:02:01 alleles have a fourteen-fold higher risk of developing IBM compared to the general population. Furthermore, in this patient cohort IBM patients carrying the abovementioned genotype developed disease symptoms on average five years earlier than non-carriers. In contrast, the presence of a single gene copy of either DRB4*01:01:01 or DQA1*01:02:01 alleles entirely negated the risk driven by DRB1*03:01:01. Although the mechanism by which these alleles exert their protective effect is yet to be determined, we can propose some potential avenues. For instance, HLA class II molecules have been found to interact with lymphocyte activation gene-3 (LAG-3; CD223), an inhibitory receptor that is induced on activated T cells (both CD4⁺ and CD8⁺) and NK cells [47]. By competing with self-reactive TCR for binding to HLA class II loaded with high-affinity self-antigens, LAG-3 may hinder the activity of autoreactive T cells that have escaped thymic negative selection, thereby controlling autoimmunity [48]. Additionally, polymorphism within HLA non-coding gene sequences can impact the level of cellular expression of the HLA molecule and, consequently, the extent of the resulting inflammatory response, as in the case of specific HLA-DP variants associated with the recovery from chronic HBV infection [49].

It is essential to interpret our findings within our study limitations. Specifically, the statistical power of our study is limited by the number of available participants, which is inherent to research in rare conditions. This constraint became particularly detrimental when assessing HLA alleles associated with anti-cN1A antibodies, as the smaller comparison groups led to diminished sensitivity in the analysis. This point is exemplified by a previous study that reported a strong correlation between anti-cN1A antibodies and HLA-DRB1*03:01 in a larger cohort of 313 Caucasian adult- and juvenile-onset myositis patients [33]; a finding we could not replicate in our smaller cohort. Finally, our observation of the correlation between the presence of the IBM-associated HLA genotype and the age at symptom onset requires replication in a different patient group because capturing symptom onset data relies on patients self-reporting, which introduces a subjective bias to the analysis.

Currently, no objective approach is available to accurately determine the timing of initial disease symptoms, especially in patients with slow disease progression.

In conclusion, this study utilised a high-resolution next-generation sequencing workflow to gain deeper insights into the role of HLA in IBM in a Caucasian patient cohort. It determined that carriage of HLA-DRB1*03:01:01 of the 8.1 AH cluster was the principal allelic component responsible for the increased disease risk and identified the combination of alleles that confers the highest risk profile for IBM as well as impact on the age of symptom onset. We anticipate that in the future, international endeavours focused on gathering and sharing high-resolution genetic data of this nature will enhance the precision and reliability of our understanding of the genomic landscape in rare diseases such as IBM, ultimately facilitating advancements in diagnosis and treatment. Additionally, studies investigating HLA – antigen – T cell receptor interactions and alternate mechanisms through which HLA molecules can facilitate IBM pathological processes could identify novel targets for therapeutic intervention.

Acknowledgements

We thank the patients of our IBM cohort and the control donors who consented to participate in this study. We are grateful to Ms Deborah Robertson and Ms Annik Panicker for liaising with the patients and collecting biological samples, Ms Linda Choo for biological samples registration, and the staff of the IIID sequencing platform at Murdoch University for their outstanding professionalism.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jaut.2023.103150>.

References

- [1] M. Needham, F.L. Mastaglia, Inclusion body myositis: current pathogenetic concepts and diagnostic and therapeutic approaches, *Lancet Neurol.* 6 (7) (2007) 620–631.
- [2] A. Callan, G. Capkun, V. Vasantharasad, R. Freitas, M. Needham, A systematic review and meta-analysis of prevalence studies of sporadic inclusion body myositis, *J. Neuromuscul. Dis.* 4 (2) (2017) 127–137.
- [3] M. Needham, I. James, A. Corbett, T. Day, F. Christiansen, B. Phillips, et al., Sporadic inclusion body myositis: phenotypic variability and influence of HLA-DR3 in a cohort of 57 Australian cases, *J. Neurol. Neurosurg. Psychiatr.* 79 (9) (2008) 1056–1060.
- [4] Molberg Ø, C. Dobloug, Epidemiology of sporadic inclusion body myositis, *Curr. Opin. Rheumatol.* 28 (6) (2016) 657–660.
- [5] M. Needham, F.L. Mastaglia, Immunotherapies for immune-mediated myopathies: a current perspective, *Neurotherapeutics* 13 (1) (2016) 132–146.
- [6] A.G. Engel, K. Arahata, Monoclonal antibody analysis of mononuclear cells in myopathies. II: phenotypes of autoimmune cells in polymyositis and inclusion body myositis, *Ann. Neurol.* 16 (2) (1984) 209–215.
- [7] O. Benveniste, W. Stenzel, D. Hilton-Jones, M. Sani, O. Boyer, Engelen Bgm, Amyloid deposits and inflammatory infiltrates in sporadic inclusion body myositis: the inflammatory egg comes before the degenerative chicken, *Acta Neuropathol.* 129 (5) (2015) 611–624.
- [8] A. Amlani, M.Y. Choi, M. Tarnopolsky, L. Brady, A.E. Clarke, I. Garcia-De La Torre, et al., Anti-NT5c1A autoantibodies as biomarkers in inclusion body myositis, *Front. Immunol.* 10 (2019) 745.
- [9] N.A. Goyal, T.M. Cash, U. Alam, S. Enam, P. Tierney, N. Araujo, et al., Seropositivity for NT5c1A antibody in sporadic inclusion body myositis predicts more severe motor, bulbar and respiratory involvement, *J. Neurol. Neurosurg. Psychiatry* 87 (4) (2016) 373–378.
- [10] S.A. Greenberg, Cytoplasmic 5'-nucleotidase autoantibodies in inclusion body myositis: isotypes and diagnostic utility, *Muscle Nerve* 50 (4) (2014) 488–492.
- [11] M.K. Herbert, J. Stammen-Vogelzangs, M.M. Verbeek, A. Rietveld, I.E. Lundberg, H. Chinoy, et al., Disease specificity of autoantibodies to cytosolic 5'-nucleotidase 1A in sporadic inclusion body myositis versus known autoimmune diseases, *Ann. Rheum. Dis.* 75 (4) (2016) 696–701.
- [12] H.B. Larman, M. Salajegheh, R. Nazareno, T. Lam, J. Sauld, H. Steen, et al., Cytosolic 5'-nucleotidase 1A autoimmunity in sporadic inclusion body myositis, *Ann. Neurol.* 73 (3) (2013) 408–418.
- [13] D. Botstein, R.L. White, M. Skolnick, R.W. Davis, Construction of a genetic linkage map in man using restriction fragment length polymorphisms, *Am. J. Hum. Genet.* 32 (3) (1980) 314.

- [14] M.A. Degli-Esposti, A.L. Leaver, F.T. Christiansen, C.S. Witt, L.J. Abraham, R. L. Dawkins, Ancestral haplotypes: conserved population MHC haplotypes, *Hum. Immunol.* 34 (4) (1992) 242–252.
- [15] M.J. Garlepp, B. Laing, P.J. Zilko, W. Ollier, F.L. Mastaglia, HLA associations with inclusion body myositis, *Clin. Exp. Immunol.* 98 (1) (1994) 40.
- [16] P. Price, L. Santoso, F. Mastaglia, M. Garlepp, C.C. Kok, R. Alcock, et al., Two major histocompatibility complex haplotypes influence susceptibility to sporadic inclusion body myositis: critical evaluation of an association with HLA-DR3, *Tissue Antigens* 64 (5) (2004) 575–580.
- [17] T.P. O'Hanlon, D.M. Carrick, F.C. Arnett, J.D. Reveille, M. Carrington, X. Gao, et al., Immunogenetic risk and protective factors for the idiopathic inflammatory myopathies: distinct HLA-A, -B, -Cw, -DRB1 and -DQA1 allelic profiles and motifs define clinicopathologic groups in caucasians, *Medicine (Baltim.)* 84 (6) (2005) 338–349.
- [18] A.P. Scott, N.G. Laing, F. Mastaglia, M. Needham, M.C. Walter, M.C. Dalakas, et al., Recombination mapping of the susceptibility region for sporadic inclusion body myositis within the major histocompatibility complex, *J. Neuroimmunol.* 235 (1) (2011) 77–83.
- [19] S. Rothwell, R.G. Cooper, I.E. Lundberg, F.W. Miller, P.K. Gregersen, J. Bowes, et al., Dense genotyping of immune-related loci in idiopathic inflammatory myopathies confirms HLA alleles as the strongest genetic risk factor and suggests different genetic background for major clinical subgroups, *Ann. Rheum. Dis.* 75 (8) (2016) 1558–1566.
- [20] B.M. Koffman, K. Sivakumar, T. Simonis, D. Stroncek, M.C. Dalakas, HLA allele distribution distinguishes sporadic inclusion body myositis from hereditary inclusion body myopathies, *J. Neuroimmunol.* 84 (2) (1998) 139–142.
- [21] A. Rojana-udomsart, I. James, A. Castley, M. Needham, A. Scott, T. Day, et al., High-resolution HLA-DRB1 genotyping in an Australian inclusion body myositis (s-IBM) cohort: an analysis of disease-associated alleles and diplotypes, *J. Neuroimmunol.* 250 (1) (2012) 77–82.
- [22] J.B. Lampe, G. Gossrau, A. Kempe, M. Füssel, K. Schwurack, R. Schröder, et al., Analysis of HLA class I and II alleles in sporadic inclusion-body myositis, *J. Neurol.* 250 (11) (2003) 1313–1317.
- [23] S. Rothwell, R.G. Cooper, I.E. Lundberg, P.K. Gregersen, M.G. Hanna, P. M. Machado, et al., Immune-array analysis in sporadic inclusion body myositis reveals HLA-DRB1 amino acid heterogeneity across the myositis spectrum, *Arthritis Rheumatol.* 69 (5) (2017) 1090–1099.
- [24] J. Currenti, A. Chopra, M. John, S. Leary, E. McKinnon, E. Alves, et al., Deep sequence analysis of HIV adaptation following vertical transmission reveals the impact of immune pressure on the evolution of HIV, *PLoS Pathog.* 15 (12) (2019), e1008177-e.
- [25] J. Robinson, D. Barker, X. Georgiou, M. Cooper, P. Flicek, S.G.E. Marsh, IPD-IMGT/HLA database, *Nucleic Acids Res.* 48 (2020) D948–D955.
- [26] C. Bundell, R. Shakra, A. Brusch, A. McLean Tookey, P. Hollingsworth, M. Needham, Diagnostic utility of cytoplasmic 5'-nucleotidase autoantibodies to identify inclusion body myositis patients, *Pathology* 49 (2017) S112–S113.
- [27] M. Migdal, D.F. Ruan, W.F. Forrest, A. Horowitz, C. Hammer, MiDAS-meaningful immunogenetic data at scale, *PLoS Comput. Biol.* 17 (7) (2021), e1009131-e.
- [28] I. Patil, Visualizations with statistical details: the 'ggstatsplot' approach, *J. Open Source Softw.* 6 (61) (2021) 3167.
- [29] R.C.R. Team, A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2021.
- [30] F.L. Mastaglia, M. Needham, A. Scott, I. James, P. Zilko, T. Day, et al., Sporadic inclusion body myositis: HLA-DRB1 allele interactions influence disease risk and clinical phenotype, *Neuromuscul. Disord.* 19 (11) (2009) 763–765.
- [31] V. Tam, N. Patel, M. Turcotte, Y. Bossé, G. Paré, D. Meyre, Benefits and limitations of genome-wide association studies, *Nat. Rev. Genet.* 20 (8) (2019) 467–484.
- [32] C.T. Deakin, J. Bowes, L.G. Rider, F.W. Miller, L.M. Pachman, H. Sanner, et al., Association with HLA-DRβ1 position 37 distinguishes juvenile dermatomyositis from adult-onset myositis, *Hum. Mol. Genet.* 31 (14) (2022) 2471–2481.
- [33] S. Rothwell, H. Chinoy, J.A. Lamb, F.W. Miller, L.G. Rider, L.R. Wedderburn, et al., Focused HLA analysis in Caucasians with myositis identifies significant associations with autoantibody subgroups, *Ann. Rheum. Dis.* 78 (7) (2019) 996–1002.
- [34] S. Raychaudhuri, C. Sandor, E.A. Stahl, J. Freudenberg, H.-S. Lee, X. Jia, et al., Five amino acids in three HLA proteins explain most of the association between MHC and seropositive rheumatoid arthritis, *Nat. Genet.* 44 (3) (2012) 291–296.
- [35] Y. Ban, T.F. Davies, D.A. Greenberg, E.S. Concepcion, R. Osman, T. Oashi, et al., Arginine at position 74 of the HLA-DR [beta]1 chain is associated with Graves' disease, *Gene Immun.* 5 (3) (2004) 203–208.
- [36] F. Menconi, R. Osman, M.C. Monti, D.A. Greenberg, E.S. Concepcion, Y. Tomer, Shared molecular amino acid signature in the HLA-DR peptide binding pocket predisposes to both autoimmune diabetes and thyroiditis, *Proceedings of the National Academy of Sciences - PNAS* 107 (39) (2010) 16899–16903.
- [37] Y. Ban, T.F. Davies, D.A. Greenberg, E.S. Concepcion, R. Osman, T. Oashi, et al., Arginine at position 74 of the HLA-DR beta1 chain is associated with Graves' disease, *Gene Immun.* 5 (3) (2004) 203–208.
- [38] H.J. Lee, M. Stefan-Lifshitz, C.W. Li, Y. Tomer, Genetics and epigenetics of autoimmune thyroid diseases: Translational implications, *Best Pract. Res. Clin. Endocrinol. Metabol.* 37 (2) (2023), 101661.
- [39] B. Miglioranza Scavuzzi, V. van Drongelen, B. Kaur, J.C. Fox, J. Liu, R.A. Mesquita-Ferrari, et al., The lupus susceptibility allele DRB1*03:01 encodes a disease-driving epitope, *Commun. Biol.* 5 (1) (2022) 751.
- [40] E. McLeish, A. Sooda, N. Slater, B. Kachigunda, K. Beer, S. Paramalingam, et al., Uncovering the significance of expanded CD8 + large granular lymphocytes in inclusion body myositis: insights into T cell phenotype and functional alterations, and disease severity, *Front. Immunol.* (2023) 14.
- [41] S.A. Greenberg, J.L. Pinkus, S.W. Kong, C. Baecher-Allan, A.A. Amato, D. M. Dorfman, Highly differentiated cytotoxic T cells in inclusion body myositis, *Brain* 142 (9) (2019) 2590–2604.
- [42] J. Neefjes, M.L.M. Jongsma, P. Paul, O. Bakke, Towards a systems understanding of MHC class I and MHC class II antigen presentation, *Nat. Rev. Immunol.* 11 (12) (2011) 823–836.
- [43] S. Ranasinghe, A. Lamothe Pedro, Z. Soghoian Damien, W. Kazer Samuel, B. Cole Michael, K. Shalek Alex, et al., Antiviral CD8+ T cells restricted by human leukocyte antigen class II exist during natural HIV infection and exhibit clonal expansion, *Immunity* 45 (4) (2016) 917–930.
- [44] T.E. Nyanhete, A.L. Frisbee, T. Bradley, W.J. Faison, E. Robins, T. Payne, et al., HLA class II-Restricted CD8+ T cells in HIV-1 Virus Controllers, *Sci. Rep.* 9 (1) (2019), 10165-17.
- [45] A. Niehrs, W.F. Garcia-Beltran, P.J. Norman, G.M. Watson, A. Hölzemer, A. Chapel, et al., A subset of HLA-DP molecules serve as ligands for the natural cytotoxicity receptor NKp44, *Nat. Immunol.* 20 (9) (2019) 1129–1137.
- [46] A. Rojana-udomsart, C. Mitropant, I. James, C. Witt, M. Needham, T. Day, et al., Analysis of HLA-DRB3 alleles and supertypal genotypes in the MHC Class II region in sporadic inclusion body myositis, *J. Neuroimmunol.* 254 (1–2) (2013) 174–177.
- [47] E. Baixeras, B. Huard, C. Miossec, S. Jitsukawa, M. Martin, T. Hercend, et al., Characterization of the lymphocyte activation gene 3-encoded protein. A new ligand for human leukocyte antigen class II antigens, *J. Exp. Med.* 176 (2) (1992) 327–337.
- [48] T. Maruhashi, I.-M. Okazaki, D. Sugiura, S. Takahashi, T.K. Maeda, K. Shimizu, et al., LAG-3 inhibits the activation of CD4 + T cells that recognize stable pMHCII through its conformation-dependent recognition of pMHCII, *Nat. Immunol.* 19 (12) (2018) 1415–1426.
- [49] R. Thomas, C.L. Thio, G.D. Kirk, W.K. Hoots, S. Wolinsky, M. Carrington, et al., A novel variant marking HLA-DP expression levels predicts recovery from hepatitis B virus infection, *J. Virol.* 86 (12) (2012) 6979–6985.