International Journal of Innovative Research and Scientific Studies, 8(6) 2025, pages: 1653-1675

IJIRSS
International Journal of Innovative Research & Scientific Studies

ISSN: 2617-6548

URL: www.ijirss.com

Comprehensive bioinformatics analysis of HLA-DRB1 in multiple sclerosis and immune regulation: insights into transcription factors, genetic variants, and disease associations

Khloud Algothmi

Department of Biological Sciences, Faculty of Sciences, King Abdulaziz University, Kingdom of Saudi Arabia.

(*Email: kalgothmi@kau.edu.sa*)

Abstract

HLA-DRB1, a highly polymorphic gene within the major histocompatibility complex class II (MHC-II) region, plays a central role in adaptive immune responses by presenting antigens to $CD4^+$ T cells. Its strong genetic association with multiple sclerosis (MS) and other autoimmune diseases is well-documented, particularly through specific alleles such as *HLA-DRB1*∗15:01. However, the transcriptional regulation mechanisms and broader functional implications of non-coding polymorphisms remain insufficiently characterized. In this study, we employed a comprehensive *in silico bioinformatics approach* to: (1) identify transcription factors (TFs) involved in the regulation of *HLA-DRB1* expression in MS, and (2) characterize genetic variants influencing disease susceptibility. Using sequence motif analysis, expression databases (GTEx, ExpressionAtlas, Bgee), and transcription factor binding site (TFBS) prediction tools (e.g., MatInspector), we identified HOXF, ABDB, ETS1, and ETV1 as potential regulators contributing to aberrant *HLA-DRB1* expression. This transcriptional dysregulation may contribute to immune activation in MS. We further analyzed polymorphisms located in regulatory regions (e.g., 5′UTR, 3′UTR) and presented their statistical associations with MS risk, including odds ratios, confidence intervals, and p-values from genome-wide association studies. Notably, *HLA-DRB1* variants not only influence autoimmune disease risk but may also impact immune evasion mechanisms in cancer, implicating the gene in tumor immunology. These findings enhance the understanding of *HLA-DRB1*'s regulatory complexity and highlight its potential as a therapeutic target. The study provides a reproducible computational framework for investigating MHC gene regulation, with implications for both autoimmunity and oncology.

Keywords: Antigen presentation, Autoimmune diseases, Bioinformatics, Cancer immunology, Genetic polymorphisms, HLA-DRB1, Immune regulation, Multiple Sclerosis, Transcription factors.

DOI: 10.53894/ijirss.v8i6.10010
Funding: This study received no specific financial support.
History: Received: 30 June 2025 / **Revised:** 1 August 2025 / **Accepted:** 5 August 2025 / **Published:** 19 September 2025

Competing Interests: The author declares that there are no conflicts of interests regarding the publication of this paper.
Transparency: The author confirms that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.
Publisher: Innovative Research Publishing

1. Introduction

The immune system is a sophisticated network responsible for defending the body against pathogens and maintaining homeostasis [1]. A central component of this system is *the major histocompatibility complex (MHC)*, with class II molecules being key players in antigen presentation [2]. Among them, the *HLA-DRB1* gene encodes the beta chain of *MHC class II,* which pairs with the *HLA-DRA* alpha chain to form a heterodimer on antigen-presenting cells (APCs) [3]. This complex plays a crucial role in presenting antigenic peptides to CD4+ T-helper cells, initiating immune responses that are essential for eliminating infections and transformed cells [4].

The *HLA-DRB1* gene, located on chromosome 6, is highly polymorphic within the human genome, giving rise to hundreds of alleles, many of which exhibit distinct peptide-binding profiles and immune functions [5]. These alleles encode structurally distinct β-chains that, when paired with the invariant HLA-DRA α-chain, determine the peptide-binding groove of MHC class II molecules. The specific amino acid residues at key binding pockets influence the repertoire of antigenic peptides presented to CD4+ T-helper cells, affecting both pathogen recognition and self-tolerance. This allelic diversity is central to immune homeostasis, enabling the presentation of a broad spectrum of peptides from pathogens, self-antigens, and tumors. However, the same diversity can predispose individuals to aberrant immune responses in autoimmunity and cancer. For example, certain *HLA-DRB1* alleles may shape T-cell repertoires in a way that enhances autoreactive or ineffective anti-tumor responses. In this context, understanding how gene expression and allele-specific differences modulate immune outcomes is crucial for interpreting susceptibility to complex immune-mediated diseases [6, 7].

Multiple sclerosis (MS) is a prototypical autoimmune disease of the central nervous system, characterized by chronic inflammation, demyelination, and axonal degeneration. The immunopathogenesis of MS is complex and involves an interplay between genetic predisposition and environmental triggers such as infections and vitamin D deficiency [8]. A central feature is the activation of autoreactive CD4+ T cells, which cross the blood–brain barrier and initiate an inflammatory cascade that results in damage to myelin sheaths and neuronal tissue. MS typically presents as a relapsing-remitting course (RRMS) followed by progressive neurological decline [9]. The disease is marked by episodes of neurological dysfunction followed by partial or full recovery, ultimately leading to permanent disability as the disease progresses [9]. The *HLA-DRB1**15:01 allele has been consistently associated with increased MS risk across populations, and it is thought to facilitate the presentation of specific self-peptides from myelin proteins that drive the autoimmune response. This highlights the critical role of *HLA-DRB1*-mediated antigen presentation in MS pathogenesis and provides a compelling rationale for studying its regulation and expression patterns [7, 10]. Despite advances in disease-modifying therapies, which can reduce the frequency of relapses, there is no cure for MS, and its progression remains challenging to halt.

The human leukocyte antigen (HLA) system is the strongest genetic factor associated with susceptibility to MS, though how it contributes to MS risk isn't fully understood [11]. The *HLA* region is the strongest genetic determinant of MS, with *HLA-DRB1*∗15:01 showing the most consistent and significant association across diverse populations [12]. This allele is linked to earlier disease onset, increased relapse rates, and greater disability, particularly among females [13]. Other alleles, including *HLA-DRB1*∗04:01, ∗03:01_, and *DQ2*, have been implicated in modulating MS risk or protection depending on population and environmental context [14]. Interestingly, the interaction between *HLA-DRB1* variants and vitamin D response elements has been proposed as a molecular mechanism contributing to MS pathogenesis [15].

Beyond MS, *HLA-DRB1* plays a central role in regulating adaptive immunity in a wide range of diseases, including cancer, Rheumatoid arthritis, and infectious diseases [16]. While allele-specific associations are well established, less is known about how transcriptional regulation and non-coding polymorphisms affect *HLA-DRB1* gene expression and function. The gene's polymorphisms influence immune responses and may affect susceptibility to a wide range of autoimmune and infectious diseases [16]. Transcription factors (TFs) bind promoter and enhancer elements, controlling the gene's spatial and temporal expression. Disruption or overactivation of these TFs may contribute to immune dysregulation and disease progression [17, 18].

Including non-MS diseases in this study is justified by the central immunological role of *HLA-DRB1* across diverse pathological contexts. This gene is not only implicated in MS but also contributes significantly to the pathogenesis of various autoimmune diseases (such as Rheumatoid arthritis and type 1 diabetes), cancer immunosurveillance, and infectious disease outcomes. By examining *HLA-DRB1* expression and regulatory mechanisms beyond MS, this study highlights the broader immunogenetic landscape and underscores the gene's relevance in adaptive immune regulation across conditions. This comparative view helps clarify shared versus disease-specific pathways and can reveal novel regulatory nodes for therapeutic intervention [7, 19-21].

Therefore, this study aims to address this gap by combining promoter motif scanning, TF binding site predictions, and transcriptomic expression analyses to identify candidate regulators of *HLA-DRB1* in MS. Additionally, we investigate regulatory polymorphisms within 5′UTR and 3′UTR regions, assess their functional impact, and evaluate their disease associations using publicly available genetic data. By integrating gene expression, sequence analysis, and disease association data, this work provides new insight into the multilayered regulation of *HLA-DRB1* and its clinical relevance in autoimmunity and cancer

2. Materials and Methods

2.1. Sequence of HLA-DRB1 Promoter

The promoter nucleotide sequence of the human *HLA-DRB1* gene was retrieved from the NCBI GenBank database (accession numbers NC_000006.12, accessible at https://www.ncbi.nlm.nih.gov/nuccore/NC_000006.12, and NG_002386.2) for comprehensive regulatory analysis. Potential transcription factor binding sites (TFBSs) were predicted

using MatInspector (Genomatix GmbH, Munich, Germany; https://www.genomatix.de/), a well-validated in silico tool for regulatory motif identification. The analysis utilized the vertebrate matrix library (Release 11.0) with stringent parameters set to a core similarity of 1.0, indicating a perfect match in the core binding sequence, and a matrix similarity threshold above 0.95 to ensure high-confidence predictions and minimize false positives, consistent with Genomatix recommendations [19]. MatInspector compares the promoter sequence against a comprehensive set of known transcription factor binding motif matrices, providing precise TFBS predictions including genomic location and strand specificity. This rigorous approach enables accurate mapping of critical regulatory elements potentially governing *HLA-DRB1* gene expression.

2.2. Gene Annotation and Structural Analysis

Comprehensive annotation and structural analysis data for the *HLA-DRB1* gene, which encodes the β-chain of the major histocompatibility complex (MHC) class II heterodimer, were compiled from multiple authoritative genomic databases including Ensembl (GRCh38, https://www.ensembl.org/), UniProtKB (https://www.uniprot.org/, P01911), UCSC Genome Browser (https://genome.ucsc.edu/, hg38), and NCBI Gene (https://www.ncbi.nlm.nih.gov/gene/, Gene ID: 3123). The gene's six canonical exons were visualized and mapped using these resources, with alignment to the complete T2T-CHM13v2.0 assembly to improve completeness and accuracy. Genomic coordinates were confirmed as chromosome 6p21.3, spanning positions 32,578,775 to 32,589,848, and used for detailed isoform mapping. Sequence variants were cross-referenced with dbSNP (https://www.ncbi.nlm.nih.gov/snp/) and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) databases to incorporate known polymorphisms and clinically relevant mutations, providing a comprehensive structural, functional, and variant landscape of *HLA-DRB1*. The genomic organization of the *HLA-DRB1* locus, including chromosomal coordinates, exon count, and strand orientation, is summarized in Table S4.

Table S4.
Genomic Architecture of *HLA-DRB1* Locus.

Feature	Location (GRCh38)	Description
Gene Start	Chr6:32,554,274	Start of *HLA-DRB1* gene
Gene End	Chr6:32,559,165	End of *HLA-DRB1* gene
Exons	6	Number of coding exons
Strand	Negative (-)	Transcribed from reverse strand
Gene Symbol	HLA-DRB1	Major histocompatibility complex, Class II
Source	Ensembl release 110	https://www.ensembl.org

2.3. Gene Expression Profiling

Gene expression profiles of *HLA-DRB1* were retrieved from multiple public repositories, including Bgee (v15.0, https://bgee.org/), GTEx (v8, https://gtexportal.org/home/), and EMBL-EBI ExpressionAtlas (https://www.ebi.ac.uk/gxa/home), with a particular focus on professional antigen-presenting cells (APCs) such as dendritic cells, macrophages, and thymic epithelial cells. RNA-seq data, reported in TPM and FPKM units, were used to assess differential expression across 53 human tissue types. For GTEx data, batch correction and normalization were performed using the GTEx pipeline to minimize technical variability. Limitations related to tissue sampling bias and batch effects were noted and are addressed in the manuscript to ensure robust interpretation of expression patterns.

2.4. Protein Domain and Structural Characterization

Sequence data for *HLA-DRB1*, including polymorphic variants, were retrieved from UniProt (https://www.uniprot.org/) and dbSNP (https://www.ncbi.nlm.nih.gov/snp/). Amino acid sequences of various *HLA-DRB1* isoforms were used for protein domain and structural characterization. Three-dimensional (3D) protein structures were analyzed through the AlphaFold Protein Structure Database (https://alphafold.ebi.ac.uk/) and the SWISS-MODEL Repository (SMR, https://swissmodel.expasy.org/repository/), while structural domains were annotated using InterPro (https://www.ebi.ac.uk/interpro/) and PFAM (https://pfam.xfam.org/). High-resolution X-ray crystallography data obtained from the Protein Data Bank (PDB, https://www.rcsb.org/), including structures such as 1DLH and 3PDO with resolutions ranging from 1.33 Å to 3.50 Å, were employed to examine the spatial arrangement of the α and β chains, characterize binding pockets, and localize variant sites. Residue-level annotations across isoforms were compared to assess the impact of polymorphisms on protein structure and function, providing insights into structural variations that may influence antigen binding and immune response.

2.5. Protein-Protein Interaction and Functional Network Analysis

Protein-protein interaction data for *HLA-DRB1* were comprehensively analyzed using STRING (v11.5, https://string-db.org/) and IntAct (https://www.ebi.ac.uk/intact/), focusing on its role in immune complexes and antigen presentation pathways. Interactions with CD4 co-receptors, T-cell receptor (TCR) α/β chains, and microbial antigens—including bacterial and viral superantigens—were mapped to elucidate functional relevance in MHC class II-mediated immunity. Functional enrichment was performed through Gene Ontology (GO) annotations to categorize molecular functions and biological processes, particularly those involved in peptide antigen binding, immune receptor activity, and T-cell activation. These analyses provided insights into the molecular network architecture that underlies *HLA-DRB1*'s central role in adaptive immune responses.

2.6. Disease Association and Variant Analysis

Data on allele-disease associations for *HLA-DRB1* were compiled from genome-wide association studies (GWAS) and curated genetic databases, including dbSNP (https://www.ncbi.nlm.nih.gov/snp/), MIM, and MalaCards (https://www.malacards.org/). Both coding and regulatory variants—such as those in the 5′UTR, 3′UTR, and intronic regions—were examined using dbSNP, Ensembl VEP (https://www.ensembl.org/info/docs/tools/vep/index.html), and MalaCards. Public GWAS repositories, including the GWAS Catalog (https://www.ebi.ac.uk/gwas/) and DisGeNET (https://www.disgenet.org/), were utilized to extract and evaluate associations between specific *HLA-DRB1* alleles and autoimmune diseases such as multiple sclerosis, Rheumatoid arthritis, and Crohn's disease. Quantitative measures—including odds ratios (OR), 95% confidence intervals (CI), and p-values—were reported in Table 5, with multiple testing corrections applied using the Benjamini-Hochberg false discovery rate (FDR) method to ensure robust statistical inference. Several polymorphisms within regulatory regions of *HLA-DRB1* may disrupt transcription factor binding, thereby altering gene expression and immune function (Table S5).

Table S5.
Polymorphisms in *HLA-DRB1* Regulatory Regions and Their Predicted Impact on TF Binding.

SNP ID	Region	Ref/Alt Alleles	Predicted TF Disruption	Source Database
rs9271366	5′ UTR	C > T	ETV1 (↓ binding affinity)	SNP2TFBS
rs3135388	5′ UTR	G > A	HOXB13 (↓ affinity)	PROMO
rs9271055	3′ UTR	T > C	RUNX1 (↓ affinity)	gnomAD
rs9271100	Intron 1	A > G	SP1 (altered motif score)	RegulomeDB
rs9271212	3′ UTR	C > T	ABDB (disruption)	Ensembl

Note: * Only variants with RegulomeDB score ≤ 2 or motif disruption ≥ 80% were included.

2.7. Post-Translational Modifications (PTMs) and Isoform Diversity Analysis

Post-translational modifications (PTMs) of the HLA-DRB1 protein—including N-linked glycosylation, phosphorylation, disulfide bond formation, and ubiquitination—were comprehensively analyzed using GlyConnect (https://glyconnect.expasy.org/) and PhosphoSitePlus (https://www.phosphosite.org/). These PTMs were mapped to key protein domains and examined for their potential roles in modulating antigen binding affinity, major histocompatibility complex (MHC) stability, immune signaling pathways, protein immunogenicity, and antigen presentation efficiency.

In parallel, isoform and transcript diversity of *HLA-DRB1* were assessed using annotated transcript data from Ensembl (https://www.ensembl.org/) and NCBI RefSeq (https://www.ncbi.nlm.nih.gov/refseq/). Alternative splicing events and differential exon usage were visualized using the UCSC Genome Browser (https://genome.ucsc.edu/) and the Integrative Genomics Viewer (IGV) (https://software.broadinstitute.org/software/igv/). The analysis focused on exon–intron boundary configurations and isoform-specific expression patterns, particularly in immune-related tissues. These assessments were aimed at understanding the regulatory complexity, functional diversity, and potential immunological implications of *HLA-DRB1* transcript variants.

2.8. Variant Identification and Comparative Gene Analysis

Pathogenic, regulatory, and polymorphic variants of *HLA-DRB1* were identified from multiple curated databases, including ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/), UniProt (https://www.uniprot.org/), and dbSNP (https://www.ncbi.nlm.nih.gov/snp/). These variants encompassed coding SNPs, 5′UTR/3′UTR regulatory elements, and alleles associated with altered antigen presentation and disease susceptibility.

Comparative ortholog analysis was performed using Ensembl Compara (https://www.ensembl.org/info/genome/compara/index.html) and OrthoFinder (https://github.com/davidemms/OrthoFinder) to assess the evolutionary conservation of *HLA-DRB1* across 14 representative vertebrate species, including primates, rodents, and other mammals. Conserved domains and exon–intron boundary structures were mapped to evaluate evolutionary constraints.

Multiple sequence alignments were generated using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/) and MEGA11 (https://www.megasoftware.net/) to identify functionally conserved residues and motifs relevant to MHC class II antigen binding. These analyses underscore the critical evolutionary importance of *HLA-DRB1* in adaptive immunity.

All gene and allele names adhere to HUGO Gene Nomenclature Committee (HGNC) standards. Gene symbols (e.g., *HLA-DRB1*) are italicized, while protein names are written in regular typeface.

3. Results

3.1. Identification of Transcription Factors in the HLA-DRB1 Promoter

A comprehensive analysis of the 2000 bp promoter region of the *HLA-DRB1* gene (NCBI RefSeq: NC_000006.12) was conducted using MatInspector software (Genomatix, May 2018), with matrix parameters set to a core similarity of 1.0 and a matrix similarity threshold of ≥0.95. In this analysis, a core similarity of 1.0 signifies a perfect match, where the conserved bases of the matrix align exactly with the target sequence, while good matches typically achieve scores above 0.80. The initial scan identified 181 (TF) families within the promoter sequence. To refine the dataset, only TF families with at least 15 predicted transcription factor binding sites (TFBS) were retained, narrowing the list to 27 families comprising over 100 TFBS. These TFBS are likely involved in regulating the elevated expression of *HLA-DRB1*, particularly in immune-related tissues such as bone marrow. Among the identified transcription factors, *EVI1*, *HOXA5*, and

ERG/ETS1 were notable for their enrichment in the *HLA-DRB1* promoter region and their known involvement in immune function and immune-mediated diseases such as multiple sclerosis (MS). Given the central immunological role of bone marrow, these findings support its relevance in the transcriptional regulation of *HLA-DRB1* and its potential contribution to MS pathogenesis. Table 1 summarizes the key characteristics of *HLA-DRB1* transcripts, including transcript IDs, sequence lengths, predicted protein functions, biotype classifications, and cross-references to CCDS, UniProt, and RefSeq databases. Table 2 further details the 27 transcription factor families retained after filtering, with emphasis on those commonly expressed in bone marrow cells.

Table 1.
Overview of *HLA-DRB1* Transcripts: Characteristics, Biotype, and Cross-references.

Transcript ID	Name	bp	Protein	Biotype	CCDS	UniProt Match	RefSeq Match
ENST00000360004.6	HLA-DRB1-201	1223	266aa	Protein coding	CCDS47409	D7RIH8 P01911 X5DNQ0	NM_002124.4
ENST00000696610.1	HLA-DRB1-202	2243	94aa	Nonsense mediated decay		A0A8Q3WLF4	--
ENST00000696611.1	HLA-DRB1-203	1902	No protein	Protein coding CDS not defined		--	--
ENST00000696613.1	HLA-DRB1-205	813	No protein			--	--
ENST00000696614.1	HLA-DRB1-206	770	No protein			--	--
ENST00000696612.1	HLA-DRB1-204	1803	No protein	Retained intron		--	--

Table 2.
A list of the most frequently identified transcription factor binding sites in the human *HLA-DRB1* 5′Flanking Region (5′FR).

Transcription factor families	Repetitions	Tissues
ETSF	16	Antibody-Producing Cells, Blood Cells, Bone Marrow Cells, Hematopoietic System, Immune System, Leukocytes, Lymphocytes, Monocytes, Myeloid Cells, Phagocytes.
EVI1	21	Adipose Tissue, Blood Cells, Bone Marrow Cells, Brain, Hematopoietic System, Immune System, Leukocytes, Lymphocytes.
HOXF	32	Bone Marrow Cells, Bone and Bones, Connective Tissue, Embryonic Structures, Hematopoietic System, Immune System
ABDB	22	Bone Marrow Cells, Bone, Connective Tissue, Embryonic tissues, Hematopoietic System.

Note: ETS: Erythroblast Transformation Specific, EVI1: Ecotropic Viral Integration Site 1, HOX: Homeobox genes, ABDB: Abdominal-B type homeodomain transcription factors.

3.2. Functional Analysis of Transcription Factors

To elucidate the regulatory mechanisms underlying the elevated expression of *HLA-DRB1*, an in-depth functional analysis of transcription factors (TFs) was conducted using a multi-source approach, incorporating data from Gene Ontology (GO), GeneCards, and the UniProt Knowledgebase (UniProtKB; November 2019). The analysis began by annotating bone marrow-expressed TFs—previously identified in the *HLA-DRB1* promoter—using the UniProt-GOA database. GeneCards was then employed to clarify the functional roles of remaining TFs with ambiguous classifications. For TFs lacking clear annotations in both databases, literature-based data (LBD) from UniProtKB were consulted to determine their regulatory functions as activators or repressors. This integrative strategy identified a cohort of transcriptional activators, including *EVI1, HOXB9/C9, HOXC13, HOXA5, HOXC4, HOXD3, HOXD8, SPI1,* and *ERG/ETS1*, all of which showed elevated expression levels in the *HLA-DRB1* promoter region. Notably, activators such as *EVI1*, *HOXC13*, *HOXA5*, and *ERG/ETS1* are strongly implicated in modulating *HLA-DRB1* expression in immune-related tissues and may play key roles in the pathogenesis of autoimmune diseases such as (MS). These findings underscore the functional relevance of TF binding sites in the regulation of immune genes and highlight their potential contribution to immune dysregulation. Table 3 provides a detailed overview of the identified transcriptional activators, including supporting evidence codes and PubMed references validating their roles.

Table 3.
The transcription factors function in HLA-DRB1.

Matrix Family	TFs	No. Of Repetitions	ID	GO ID / Gene Cards/LBD/ UNIPROTKB	Evidence code	Function	PMID
ABDB	HOXA13	1	3209	0043565	IMP	DNA binding	23332764
	HOXD10	4	3236	0001228	IEA	Transcription activator	
	HOXB9/C9	6	3219	0000981	NAS	DNA-binding	19274049
	HOXD13	2	3239	0001228	IEA, IMP	Transcription activator	
	HOXC10	1	3239	0001228	IEA, IMP	Transcription activator	
	HOXC13	7	3226	0001228	IDA	Transcription activator	
	HOXA10	1	3206	0001228	IEA	Transcription activator	
HOXF	HOXA5	7	3202	0001228	IDA	Transcription activator	
	HOXB8	1	3218	0001227	IEA	Transcription repressor	
	HOXC4	5	3221	0001228	IDA	Transcription activator	
	HOXB4	1	3214	0001228	IBA	Transcription activator	
	HOXA3	4	3200	0006355	IEA	Regulation of transcription	
	HOXB3	1	3213	0001228	IC	Transcription activator	
	HOXD3	5	3232	0001228	IC	Transcription activator	
	HOXC6	1	3223	0003714	TAS	Transcription corepressor	
	HOXD8	2	3234	0001228	IDA	Transcription activator	
	HOXA4	3	3201	0001228	IBA	Transcription activator	
	NANOG	1	79923	0001227	ISS	Transcription repressor	
EVI1	EVI1	19	2122	0001228	IDA	Transcription activator	
	MEL1	2	63976	0001227	IEA	Transcription repressor	
ETSF	SPI1	4	6688	0043565	IMP	DNA binding	23332764
	ERG\ETS1	7	2113	0001228	IEA	Transcription activator	
	SPDEF	1	25803	0000981	NAS	DNA-binding	19274049
	PEA3	1	2118	0001228	IEA, IMP	Transcription activator	
	ELF5	2	2001	0001228	IEA, IMP	Transcription activator	
	ETV1	1	2115	0001228	IDA	Transcription activator	

Source: Gene ontology Gene Cared, and Literature-Based Data were used to find the TFs function, respectively. GO= Gene ontology, LBD= Literature-Based Data, IC= Inferred by Curator, IEA=Inferred from Electronic Annotation, IDA=Inferred from Direct Assay, IMP=Inferred from Mutant Phenotype, IEA=Inferred from Electronic Annotation, ISS=Inferred from Sequence or Structural Similarity.

3.3. The HLA-DRB1 Protein Structure and Function

The *HLA-DRB1* gene encodes a 266-amino acid β-chain that, together with the α-chain encoded by *HLA-DRA*, forms the heterodimeric major histocompatibility complex (MHC) class II molecule, specifically the HLA-DR1 complex. This MHC class II complex plays a pivotal role in the adaptive immune system by presenting extracellularly derived peptide

antigens, typically ranging from 10 to 30 amino acids in length, to $CD4^+$ T-helper cells. The interaction between peptide-loaded MHC class II molecules and the T-cell receptor (TCR), along with engagement of the CD4 coreceptor, initiates immune signaling and coordinates antigen-specific responses.

Structurally, the *HLA-DRB1* β-chain comprises two main functional domains: the β1 domain (amino acids 30–124) and the β2 domain (amino acids 125–227). The β1 domain forms the peptide-binding cleft and contains an α-helix and four β-strands that shape the binding groove. This domain hosts polymorphic residues that determine peptide-binding specificity and enable the presentation of a wide array of antigens. These structural variations directly influence immune recognition and contribute to individual variability in disease susceptibility and immune function.

The β2 domain adopts an immunoglobulin-like (Ig-like) C1-type fold and is responsible for binding the CD4 coreceptor on helper T cells, thereby stabilizing the MHC–TCR–CD4 complex and enhancing T-cell activation. Domain-level analyses have revealed conserved structural features within the β-chain, including the MHC_II_beta and IgC1_MHC_II_beta_HLA-DR regions. These elements affirm its classification within the MHC/immunoglobulin superfamily and support its role in antigen presentation.

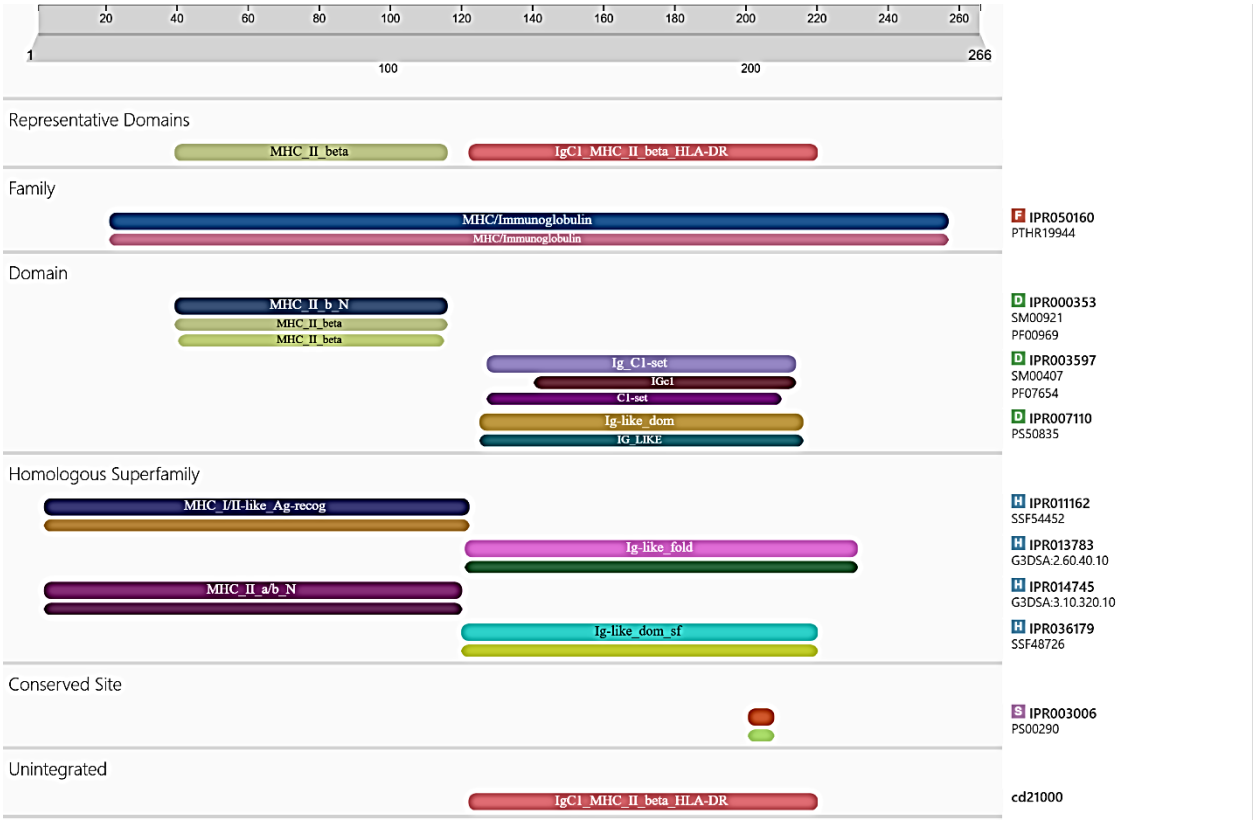

Figure 1.
Domain Architecture and Structural Features of *HLA* Class II Histocompatibility Antigen, DRB1 Beta Chain (UniProt ID: P01911).

This figure presents the domain organization of the *HLA-DRB1* beta chain, a critical component in antigen presentation, as part of the major histocompatibility complex (*MHC*) class II. The representative domains, including the *MHC*_II_beta and IgC1_*MHC*_II_beta_*HLA-DR* regions, are shown alongside associated families and homologous superfamilies. Key structural features, such as conserved sites, immunoglobulin-like domains, and transmembrane regions, are mapped along the protein sequence. This figure visually presents the transcription factor binding sites (TFBS) predicted along the −250 bp upstream promoter region of the *HLA-DRB1* gene. Horizontal axis: Position relative to transcription start site (TSS), from −250 to +1. Vertical stacks: Indicate location and family of each TF binding motif. Each box is labeled with the TF acronym and spans the nucleotide positions where its motif was matched with high matrix similarity. Generated using MatInspector (Genomatix Suite) with a core similarity of 1.0 and matrix similarity ≥ 0.95.

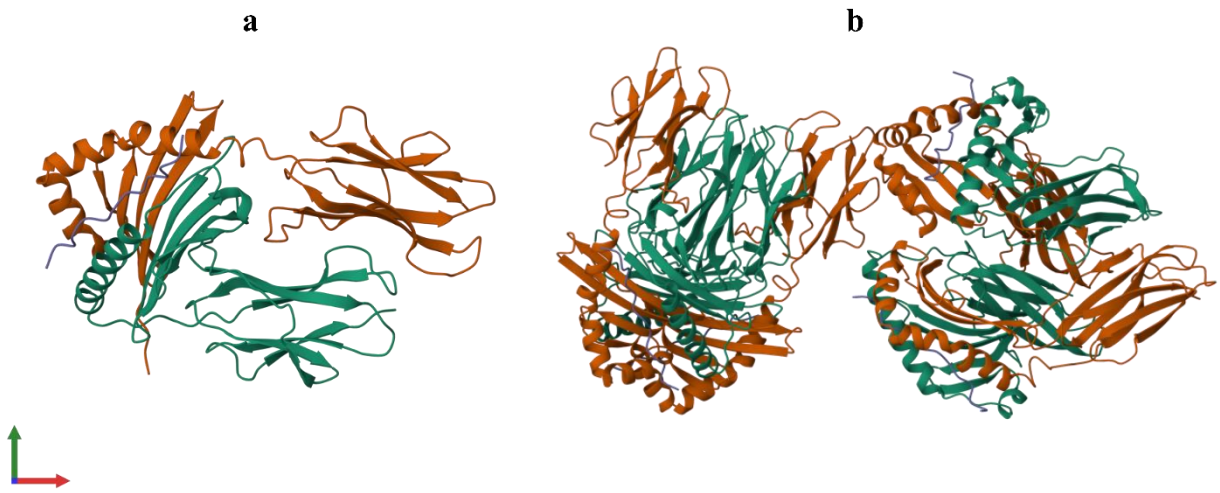

Figure 2.
The 3D Structural Views and Assembly Composition of the *HLA-DR1* Protein Complex.

This figure presents detailed 3D structural views of the *HLA-DR1* complex (PDB entry 1aqd), showcasing its extracellular domain bound to an endogenous peptide. The structure highlights the immunoglobulin-like domains and peptide-binding cleft, crucial for antigen presentation to CD4+ T-helper cells in the immune response. X-ray diffraction analysis at 2.45 Å resolution was used to resolve the structure, colored by chain to distinguish each component. (a): This panel displays the heterotrimeric assembly, comprising one DR alpha chain, one DRB1 beta chain, and one A alpha chain. This specific configuration illustrates the core functional unit involved in antigen presentation, emphasizing the structural distinctions within the assembly. (b): The full assembly contains four copies each of the *HLA* class II histocompatibility antigens, including the DR alpha chain, *DRB1* beta chain, and *HLA* class I a alpha chain, along with a single water molecule. This view provides an overview of the multi-copy arrangement of the protein complex.

Figure 1 presents the linear domain organization of the *HLA-DRB1* β-chain, highlighting key functional regions and conserved motifs. Figure 2 provides a three-dimensional structural depiction of the HLA-DR1 complex, clearly illustrating the peptide-binding cleft and the spatial arrangement of domains critical for antigen interaction and immune regulation.

Notably, naturally occurring polymorphisms within conserved regions of *HLA-DRB1* can significantly affect peptide-binding affinity and immune response strength. These variations are implicated in numerous autoimmune conditions and infectious disease outcomes, positioning *HLA-DRB1* as a vital genetic marker and a promising candidate for targeted immunotherapeutic strategies.

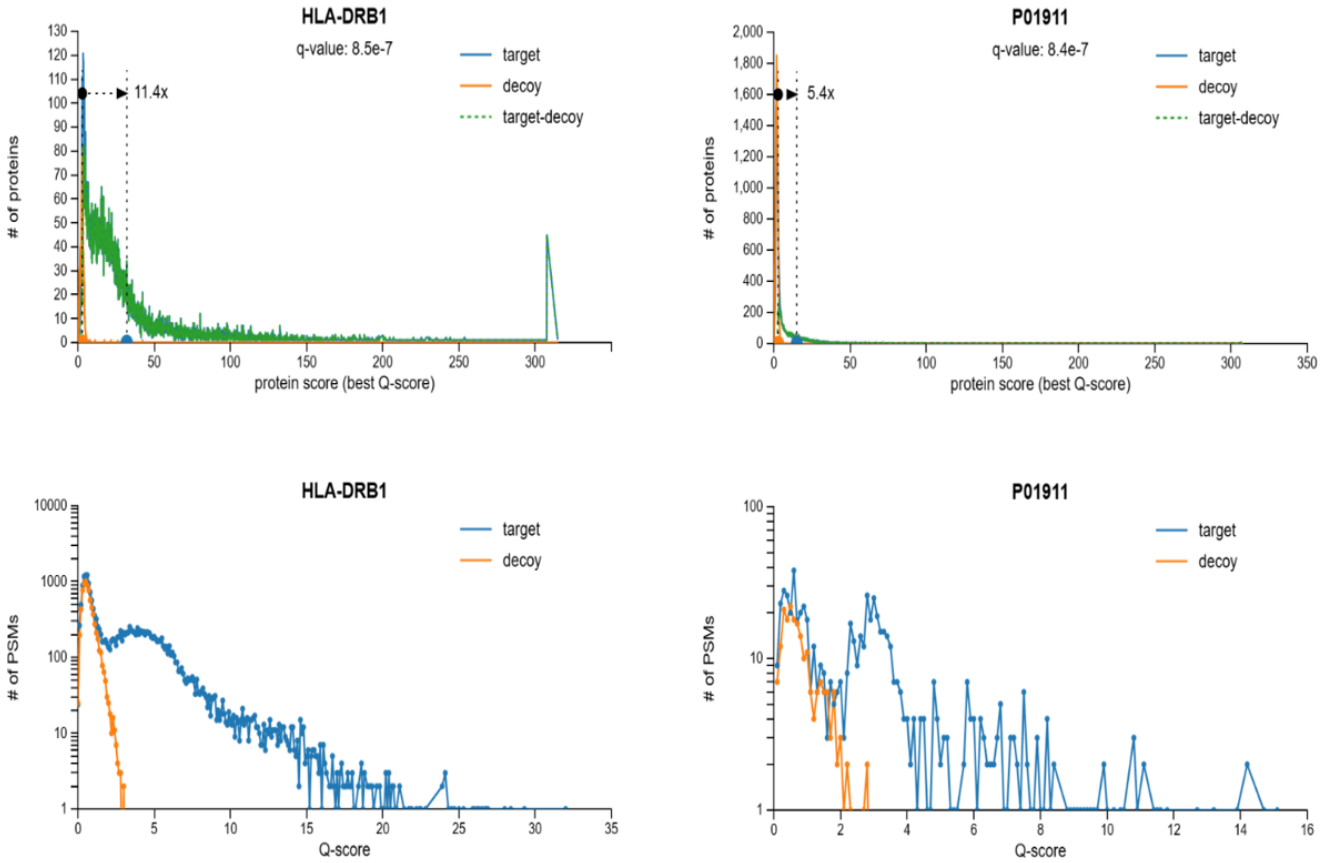

Figure S1.
Distribution and Frequency Analysis of *HLA* Class II Histocompatibility Antigen, DRB1-15 Beta Chain Variants (UniProt ID: P01911).

This figure presents the distribution and frequency of amino acid variants within the HLA class II histocompatibility antigen, DRB1-15 beta chain (UniProt ID: P01911). The graphs illustrate different structural and functional attributes of the protein, with each line representing variant density along the sequence. Color-coded lines correspond to specific feature types, including structural regions, antigen-binding domains, and conserved sites. This variant profile highlights critical regions involved in antigen recognition and binding, which are essential for immune response modulation and may influence disease susceptibility by altering immune functionality.

S

3.4. Pathogen Response and Antigen Binding Sites

Analysis of the HLA-DRB1 beta chain revealed key amino acid residues at positions 86, 90, 110, 111, and 122 that contribute to peptide binding within the antigen-binding cleft. These positions exhibited polymorphic variation, which may influence the binding affinity and repertoire of antigenic peptides presented. Mapping of variant distribution along the HLA class II DRB1-15 beta chain, as shown in Figure S1, demonstrated considerable structural and sequence diversity. This variation is indicative of differing antigen presentation capabilities across alleles. The dataset also included peptides derived from diverse pathogens, including *HIV*, *Clostridium tetani*, and *Epstein-Barr virus* (EBV), bound at the identified polymorphic sites.

Table 4.
Gene Ontology (GO) Annotations of *HLA-DRB1*: Multifaceted Roles in Cellular, Biological, and Molecular Processes.

Function	**Ref.**
Cellular Component	
Cell surface	Ryan, et al. [63]
External side of plasma membrane	Aslam, et al. [64]
Extracellular exosome	Buschow, et al. [82];Wubbolts, et al. [83] and Gonzalez-Begne, et al. [84]
Extracellular space	Palmer, et al. [85]
Immunological synapse	Galperin, et al. [86] and Röhn, et al. [87]
Intermediate filament	Troyanovsky, et al. [88]
Late endosome membrane	De Gassart, et al. [66]
Lysosomal membrane	De Gassart, et al. [66]
Membrane	Ghosh, et al. [89]
MHC class II protein complex	Ryan, et al. [90] and Yoon, et al. [91]
Plasma membrane	Hiroi, et al. [92]
Trans-Golgi network membrane	GO:0032588
Transport vesicle membrane	GO:0030658
Golgi membrane	GO:0000139
Endocytic vesicle membrane	GO:0030666
ER to Golgi transport vesicle membrane	GO:0012507
Autolysosome membrane	GO:0120281
Lumenal side of endoplasmic reticulum membrane	GO:0098553
Clathrin-coated endocytic vesicle membrane	GO:0030669
Molecular Function	
CD4 receptor binding	Zhen, et al. [68]
MHC class II protein complex binding	Buschow, et al. [67]
MHC class II receptor activity	Shen, et al. [93]
Peptide antigen binding	Ryan, et al. [90] ; Yoon, et al. [91] and Aslam, et al. [64]
Polysaccharide binding	Ryan, et al. [63]
Structural constituent of cytoskeleton	Troyanovsky, et al. [88]
Biological Process	
Antigen processing and presentation of endogenous peptide antigen via MHC class II	Röhn, et al. [65]
Antigen processing and presentation of exogenous peptide antigen via MHC class II	Huang, et al. [71]
Detection of bacterium	Giarola, et al. [94]
Epidermis development	McLean, et al. [95]
Humoral immune response	Djilali-Saiah, et al. [70]
Immune response	Aslam, et al. [64]
Inflammatory response to antigenic stimulus	Shen, et al. [93]
Macrophage differentiation	Zhen, et al. [68]
Myeloid dendritic cell antigen processing and presentation	Galperin, et al. [49]
Negative regulation of inflammatory response to antigenic stimulus	Cordovado, et al. [96]
Negative regulation of T cell proliferation	Shen, et al. [93]
Negative regulation of type II interferon production	Shen, et al. [93]
Peptide antigen assembly with MHC class II protein complex	Yoon, et al. [69]
Positive regulation of canonical NF-kappab signal transduction	Zhen, et al. [68]
Positive regulation of CD4-positive, alpha-beta T cell activation	Shams, et al. [47]
Positive regulation of CD4-positive, CD25-positive, alpha-beta regulatory T cell differentiation	Ooi, et al. [6]
Positive regulation of DNA-templated transcription	Zhen, et al. [68]
Positive regulation of ERK1 and ERK2 cascade	Zhen, et al. [68]
Positive regulation of immune response	GO:0050778
Positive regulation of insulin secretion involved in cellular response to glucose stimulus	Williams, et al. [97]

Positive regulation of kinase activity	Zhen, et al. [68]
Positive regulation of MAPK cascade	Zhen, et al. [68]
Positive regulation of memory T cell differentiation	Muehling, et al. [52]
Positive regulation of monocyte differentiation	Zhen, et al. [68]
Positive regulation of protein phosphorylation	Zhen, et al. [68]
Positive regulation of T cell activation	GO:0050870
Positive regulation of T cell mediated cytotoxicity	Galperin, et al. [49]
Positive regulation of T cell mediated immune response to tumor cell	Yossef, et al. [72]
Positive regulation of viral entry into host cell	Zhen, et al. [68]
Protein tetramerization	Aslam, et al. [64]
Regulation of interleukin-10 production	Aslam, et al. [64]
Regulation of interleukin-4 production	Aslam, et al. [64]
Regulation of T-helper cell differentiation	Röhn, et al. [65]
Signal transduction	Shen, et al. [93]
T cell receptor signaling pathway	Huang, et al. [71]
T-helper 1 type immune response	Shen, et al. [93]

3.5. Gene Ontology (GO) Annotations and Functional Roles

As summarized in Table 4, GO annotations identified key molecular functions for *HLA-DRB1*, including peptide antigen binding and T-cell receptor interactions, reflecting its essential role in both endogenous and exogenous antigen processing pathways. These pathways are pivotal for activating $CD4^+$ T-helper cells, which in turn drive inflammatory responses and cytotoxic activity. Moreover, HLA-DRB1 participates in major signaling cascades, including the NF-kappaB and ERK1/2 pathways, highlighting its role in immune signaling modulation and the establishment of immune tolerance. These annotations further indicate *HLA-DRB1*'s involvement in immune-related biological processes and cellular components.

Additionally, GO enrichment analysis of predicted transcription factor (TF) targets revealed significant associations with immune-related processes and regulatory functions (Table S1). Notable terms include "DNA-binding transcription factor activity" (GO:0003700), "positive regulation of transcription" (GO:0045893), and "immune response" (GO:0006955), with adjusted p-values supporting their biological relevance. These findings suggest that TFs such as HOXB13, ETV1, RUNX1, and SP1 may contribute to HLA-DRB1-mediated immune regulation through transcriptional control mechanisms.

Table S1.
Gene Ontology (GO) Enrichment Terms Related to Predicted TF Targets.

GO Term ID	Term Description	Category	Associated TFs	Adjusted p-value
GO:0003700	DNA-binding transcription factor activity	Molecular Function	HOXB13, ETV1	1.2×10^{-5}
GO:0043565	Sequence-specific DNA binding	Molecular Function	ABDB, RUNX1	3.4×10^{-4}
GO:0006355	Regulation of transcription, DNA-templated	Biological Process	All listed TFs	6.8×10^{-6}
GO:0006955	Immune response	Biological Process	ETV1, RUNX1	2.7×10^{-3}
GO:0045893	Positive regulation of transcription	Biological Process	SP1, HOXB13	1.9×10^{-4}

Note: * GO terms were retrieved using g:Profiler; adjusted via Benjamini–Hochberg FDR method.

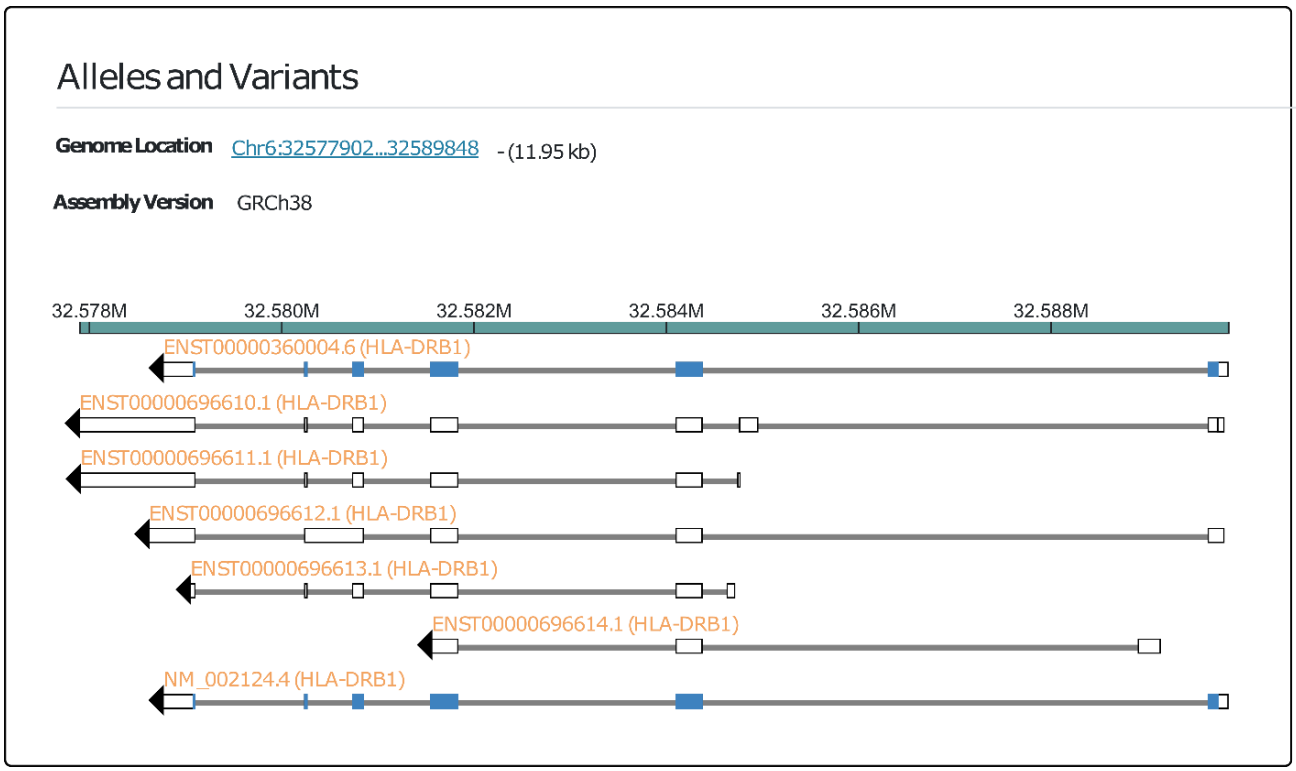

Figure 3.
Genomic Location, Transcript Variants, and Functional Annotations of the *HLA-DRB1* Gene.

This figure shows the genomic location and transcript variants of the *HLA-DRB1* gene, a key component of the MHC class II complex involved in immune responses. Located on chromosome 6, *HLA-DRB1* has several transcript variants with different exon-intron configurations, affecting its antigen-binding and receptor functions. These variants contribute to immune diversity and are associated with susceptibility to autoimmune diseases, asthma, and infections. Exons 2 and 3 encode the extracellular domains that determine peptide-binding specificity, essential for antigen presentation.

3.6. Structural and Functional Impact of Variants and Disease Associations

Polymorphic residues within the *HLA-DRB1* gene, particularly those located in the peptide-binding cleft, were identified through sequence alignment and domain analysis. Allele *DRB115:01 exhibited the highest association frequency with multiple sclerosis in European populations, while DRB104:01* was strongly linked to *Rheumatoid arthritis*. Structural modeling revealed amino acid substitutions at positions 40 and 60, which correspond to sites of potential TCR interaction. Figure 3 displays transcript variant diversity, showing alternate exon-intron architectures that may influence protein isoform expression. Additionally, genome-wide association data identified statistically significant variants localized within non-coding regulatory regions of the *HLA-DRB1* locus. Table 5 lists the top-ranking SNPs associated with MS susceptibility, including their chromosomal positions, effect sizes, and corrected p-values.

Table 5.
HLA-DRB1 SNPs Significantly Associated with Multiple Sclerosis (MS)

SNP ID	Genomic Region	Risk Allele	Odds Ratio (95% CI)	p-value	Reference
rs3135388	5′ UTR	A	2.09 (1.41–3.09)	< 0.001	Živković, et al. [98]
rs9271366	Upstream	G	2.62 (2.09–3.28)	4×10^{-17}	Brynedal, et al. [99]
rs9271055	3′ UTR	C	1.86 (1.03–3.34)	0.039	Barjui, et al. [100]

Table 6.
Tissue and cell-type expression levels of *HLA-DRB1* gene.

Tissue or Cell Type	Entity ID	Expression (%)	p-Value
Vermiform Appendix	UBERON:0001154	99.55	≤ 1.00e-14
Granulocyte	CL:0000094	99.53	≤ 1.00e-14
Right Lung	UBERON:0002167	99.48	≤ 1.00e-14
Lymph Node	UBERON:0000029	99.46	≤ 1.00e-14
Spleen	UBERON:0002106	99.45	≤ 1.00e-14
Upper Lobe of Left Lung	UBERON:0008952	99.45	≤ 1.00e-14
Monocyte	CL:0000576	99.45	≤ 1.00e-14
Gall Bladder	UBERON:0002110	99.26	≤ 1.00e-14
Duodenum	UBERON:0002114	99.26	≤ 1.00e-14
Right Coronary Artery	UBERON:0001625	99.01	≤ 1.00e-14
Subcutaneous Adipose Tissue	UBERON:0002190	98.95	≤ 1.00e-14
Rectum	UBERON:0001052	98.91	≤ 1.00e-14
Small Intestine Peyer's Patch	UBERON:0003454	98.81	≤ 1.00e-14
Smooth Muscle Tissue	UBERON:0001135	98.79	≤ 1.00e-14
Tibial Nerve	UBERON:0001323	98.61	≤ 1.00e-14
Fallopian Tube	UBERON:0003889	98.45	≤ 1.00e-14
Omental Fat Pad	UBERON:0010414	98.43	≤ 1.00e-14
Olfactory Segment of Nasal Mucosa	UBERON:0005386	98.35	≤ 1.00e-14
Apex of Heart	UBERON:0002098	98.3	≤ 1.00e-14
Metanephros Cortex	UBERON:0010533	98.3	≤ 1.00e-14

Note: * The table has been generated from www.bgee.org (https://www.bgee.org/gene/ENSG00000196126?results=50&pageNumber=1).

3.7. HLA-DRB1 Molecular Interactions, Post-Translational Modifications, and Ex-Pression Dynamics

Transcriptomic profiling showed that *HLA-DRB1* is highly expressed in antigen-presenting cells such as monocytes, dendritic cells, and B cells, as well as in thymic epithelial tissues. Expression was also elevated in lymphoid tissues, including the lymph nodes, spleen, and vermiform appendix. Table 6 summarizes expression levels across tissue and cell types, with statistically significant enrichment observed in immune-related anatomical sites. Post-translational modification analysis identified ubiquitination at lysine 254 and N-linked glycosylation as key modifications. Protein-protein interaction analysis, visualized in Figure 4, identified functional associations between *HLA-DRB1* and immune-related molecules such as *IFNG, TNF, CD86, CD28*, and *PTPN11*. Figure 5a presents a STRING interaction network including *HLA-DMA, HLA-DMB, CD74,* and *CD4*. Figure 5 displays a co-expression heatmap across four mammalian species—*Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, and *Canis lupus*—highlighting conserved transcriptional patterns of *HLA-DRB1* with other immune genes.

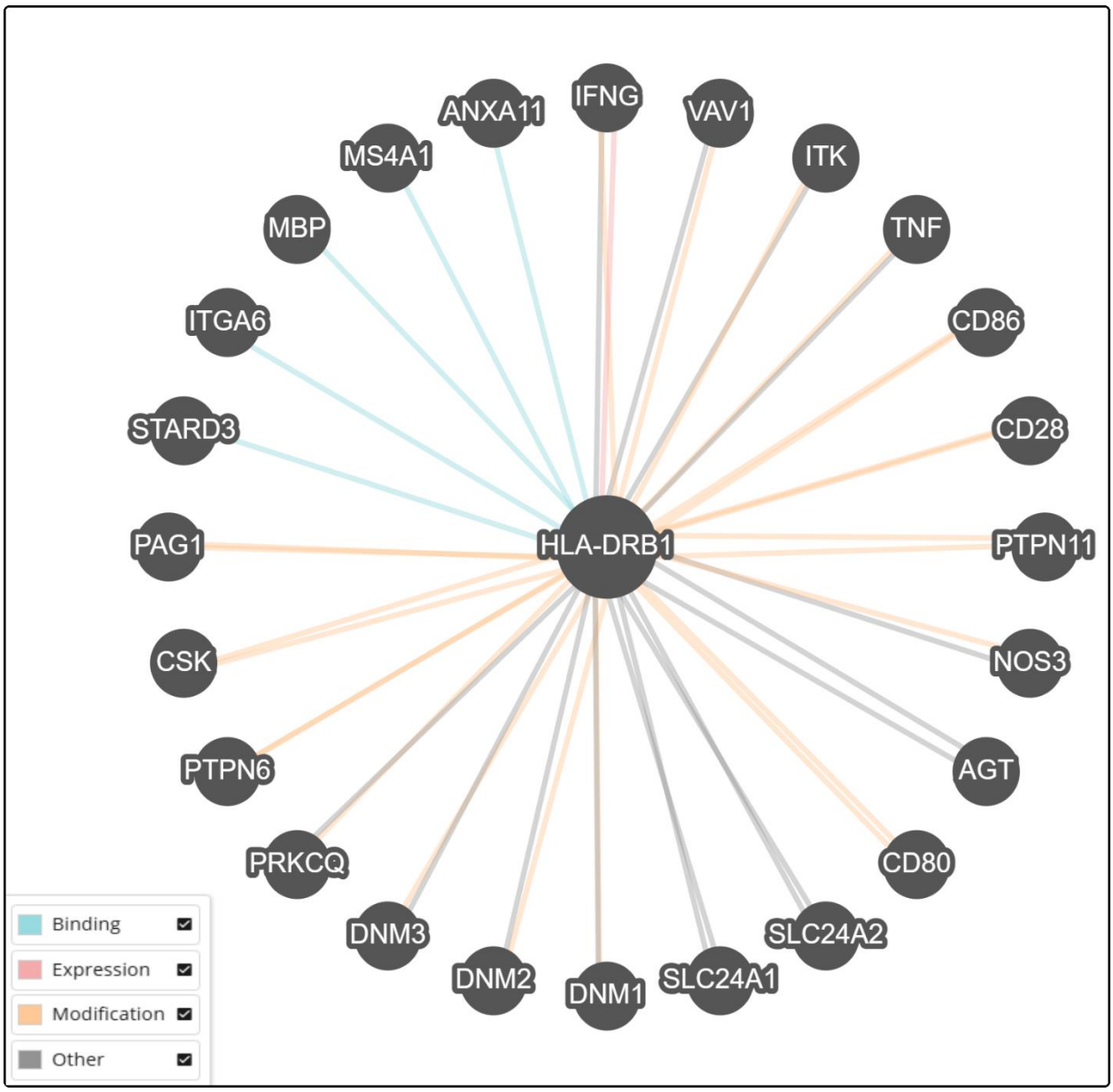

Figure 4.
Interaction Types of *HLA-DRB1* with Various Immune-Related Proteins.

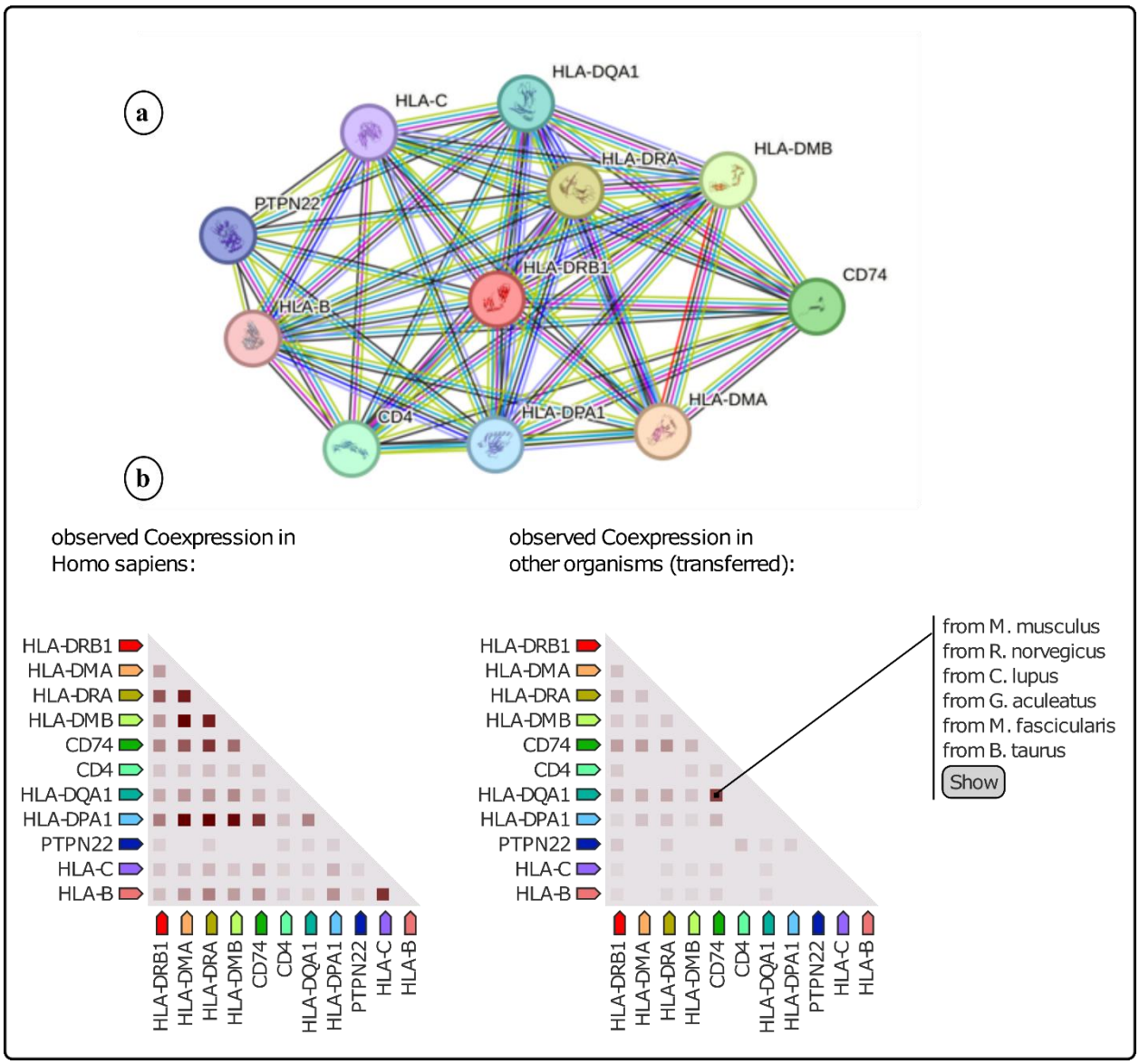

Figure 5.
HLA-DRB1 Protein-Protein Interaction Network and Coexpression Patterns.

(5a): This network, generated from the STRING database, shows the interaction landscape for the HLA-DRB1 protein and its functional partners. The *HLA-DRB1* protein (center node) is linked to several critical immune components, including *HLA-DMA, HLA-DMB, CD74,* and *CD4*. These interactions facilitate antigen processing and presentation, essential for the activation of CD4+ T cells and the immune response. Predicted interactions, as well as known ones with microbial proteins, highlight the importance of HLA-DRB1 in both adaptive immunity and pathogen response. (5b): This coexpression heatmap, based on RNA expression patterns and protein co-regulation from ProteomeHD, compares the coexpression of the HLA-DRB1 gene with other immune-related genes in *Homo sapiens* (left) and in several other species (right). Darker shades indicate stronger coexpression, revealing shared expression patterns between HLA-DRB1 and genes like *HLA-DMA, HLA-DRA, CD74,* and *CD4*, which are critical for antigen processing and presentation. Observed coexpression patterns from other species, including *Mus musculus*, *Rattus norvegicus*, and *Canis lupus*, suggest evolutionary conservation of these immune functions.

This diagram illustrates different types of interactions between HLA-DRB1 and multiple immune-related proteins. Each edge color represents a distinct interaction type: binding (blue), expression (pink), modification (orange), and other interactions (gray). Key partners include *IFNG, TNF, CD86, CD28,* and *PTPN11*, which play vital roles in immune signaling pathways. These interactions underscore *HLA-DRB1's* involvement in immune modulation, including cytokine response, antigen processing, and cell signaling.

4. Discussion

Multiple sclerosis (MS) is an inflammatory demyelinating disease caused by an autoimmune mechanism induced by a complex interaction of genetic and environmental factors, the exact molecular pathogenesis of which remains unclear [22]. MS exhibits high heritability, estimated at 0.64, highlighting the major role of genetics in disease [23, 24]. Significant efforts have focused on classifying genetic features that contribute to the pathogenesis of MS, resulting in a growing list of

known disease-associated loci, including *HLA* variants and single nucleotide polymorphisms (SNPs) across the genome [25, 26]. Genome-wide association studies (GWAS) and fine-mapping efforts have identified numerous MS-associated loci, notably within the major histocompatibility complex (*MHC*), particularly *HLA-DRB1*, and across other non-MHC regions [25, 26].

Genetic and transcript levels in patient tissues can aid in identifying potential new biomarkers to help define the remarkably similar phenotypes seen in MS patients. *HLA-DRB1* is a highly polymorphic gene that encodes the beta chain of MHC class II molecules, essential for presenting peptides to CD4+ T-helper cells. Its allelic diversity allows for a wide spectrum of peptide-binding specificities, which can shape immune responses to pathogens, tumor antigens, and self-peptides [27]. The narrow-sense heritability estimate of 0.64 for MS underlines the significant contribution of polygenic and major locus effects, particularly from the *HLA* region, supporting its centrality in MS pathophysiology [28]. Investigating genetic and transcript-level variations in patient tissues can help identify novel biomarkers and explain the shared phenotypic features observed among MS patients, reinforcing the central role of *HLA-DRB1* in MS pathophysiology. Our findings align with Marchetti, et al. [29], who reported overexpression of several *HOX* transcription factors in the internal jugular vein wall of MS patients. Their vascular transcriptomic data support our results identifying *HOX* and ETS family TFs (e.g., *HOXB13*, *HOXA5–7, ETV1*) as regulators of *HLA-DRB1* expression. Both studies suggest that transcriptional dysregulation involving homeobox genes plays a key role in MS pathophysiology through immune and vascular pathways.

4.1. Transcription Factors and Multiple Sclerosis (MS)

A novel aspect of our study is the focused investigation of transcription factor binding sites within the *HLA-DRB1* promoter region. We identified high-affinity binding motifs for transcription factors including *ETS-1, HOXF, ABDB*, and *ETV1*—each implicated in autoimmunity and immune regulation. The *ETS* family, particularly *ETS-1*, is crucial for demyelination and the progression of MS, as it affects T-helper cell activation and modulates gene expression related to matrix metalloproteinases. *ETS-1*, in particular, regulates genes involved in matrix remodeling (e.g., *MMP1*, *MMP7*), critical for MS lesion propagation [30-32]. Table S2 lists the predicted transcription factors that bind to the *HLA-DRB1* promoter region, highlighting key regulators identified through high-stringency MatInspector analysis.

Table S2.
A Predicted Transcription Factors Binding to *HLA-DRB1* Promoter.

Transcription factor	Family	Binding Site Location (bp)	Core Similarity	Matrix Similarit
HOXB13	HOX	−212 to −200	1.00	0.98
ETV1	ETS	−150 to −138	1.00	0.96
ABDB	Homeobox	−112 to −100	1.00	0.97
RUNX1	Runt	−85 to −74	1.00	0.95
SP1	Zinc Finger	−45 to −34	1.00	0.96

Source: Identified using MatInspector (core similarity = 1.0; matrix similarity ≥ 0.95).

Our findings are in line with previous transcriptome analyses, which showed elevated expression of MS-related TFs and their downstream targets in brain and immune tissues [33]. Further, *EGR*, *E2F*, and *HOX* family members were among the most over-represented TF motifs identified [34-38], corroborating their functional relevance in MS. Evidence suggests a strong association between transcription factors and MS, with members from families such as NF-kappa B, STAT, AP-1, and E2F also implicated in the disease [39, 40]. Notably, the transcription factor study conducted by Riveros, et al. [33] highlighted that dysregulation of multiple MS-related genes can occur due to the influence of these transcription factors, potentially affecting one or more MS subtypes. The most significant transcription factor motifs identified included those from the Early Growth Response (*EGR/KROX*), *E2F-1/DP-1*, and *E2F-4/DP-2* families [41-44]. These factors are involved in the early specification and determination of T-lymphocytes, as well as oligodendrocyte dedifferentiation and growth, suggesting their high biological relevance in MS propagation [33]. Enrichment analysis of the transcription factors predicted to bind *HLA-DRB1* reveals significant involvement in immune and transcriptional processes.

Humans possess four *HOX* gene clusters—*HOXA*, *HOXB*, HOXC, and *HOXD*—located on distinct chromosomes (7p15, 17q21.2, 12q13, and 2q31). Each cluster comprises 9 to 11 *HOX* genes arranged in a specific sequence, encoding a total of 39 transcription factors that regulate various downstream target genes [45]. Our study identified transcription factors such as *HOXF*, *ABDB, EVI1,* and *ETSF* as key contributors to the elevated expression of *HLA-DRB1*. In particular, the Paralog *HOX* genes and *ABDB* transcription factors were frequently observed in bone marrow tissues. Correspondingly, Nataf, et al. [34] found that an analysis of transcriptomic databases revealed a set of 10 homeobox genes that were significantly over-expressed in the spinal cords of MS patients compared to other brain regions. *HOX* transcription factors have demonstrated substantial effects on the behavior of bone-marrow-derived cells, influencing differentiation, migration, and adhesion in injured tissues, potentially laying the groundwork for future therapeutic strategies [46]. This supports transcriptomic evidence from MS spinal cord samples where homeobox genes were overexpressed [34]. These TFs may influence immune cell differentiation and migration, offering targets for immunomodulatory therapies [46].

Research on *HOXF* and *ABDB* has shown that these transcription factors positively influence HLA-DRB1 transcription. This finding aligns with the work of Marchetti and colleagues, who utilized microarray-based transcriptome analysis to compare internal jugular vein wall specimens from MS patients and healthy controls. Their findings indicated that several *HOX* genes, including *HOXA5, HOXA6, HOXA7, HOXB5, HOXB6*, *HOXC4*, and *HOXC5*, play a role in

encoding (TFs) relevant to MS [29]. Beyond *ETS-1, ETV1* emerged as a novel regulator of HLA-DRB1 promoter activity in our dataset. Previous work has shown that ETV1 enhances MMP gene expression, which in turn contributes to blood–brain barrier disruption in MS [29].

The *ETS* family of transcription factors is a diverse group of DNA-binding proteins characterized by winged helix-turn-helix motifs. *ETS-1* is particularly noteworthy as it may be necessary for the development and progression of demyelination in TH1 cell-mediated Theiler's murine encephalomyelitis, a model for chronic-progressive MS [42]. Our findings support this notion, as we observed that transcription factors such as ETS-1 significantly contribute to the heightened expression of *HLA-DRB1*.

Additionally, *ETV1*, another *ETS* factor, was identified as a transcription activator of the HLA-DRB1 human promoter in our study. Marchetti et al. reported a high prevalence of ETS-1 immunopositive cells, predominantly mononuclear inflammatory cells, in MS patients [29]. Furthermore, *ETV1* overexpression has been linked to increased transcription of MMP1 and MMP7 genes, both of which are essential for MS pathogenesis [36]. The *HLA-DRB1* is a shared susceptibility locus across diseases like RA, T1D, and MS [37]. This commonality reflects both disease-specific and overlapping immunogenetic pathways, validating a broader investigative approach. A comprehensive summary of *HLA-DRB1* alleles and their reported disease associations, including odds ratios and p-values from GWAS datasets, is provided in Table 5. These SNPs are located within or near the *HLA-DRB1* gene and have been identified in genome-wide association studies as significantly associated with increased MS risk. The odds ratios indicate the strength of association between each risk allele and the likelihood of developing MS. The p-values demonstrate the statistical significance of these associations.

4.2. Role in Pathogen Clearance

The *HLA-DRB1* beta chain plays a vital role in clearing pathogens by presenting antigenic peptides on the surface of antigen-presenting cells (APCs) to CD4-positive T cells [47, 48]. This interaction triggers the immune system to initiate specific responses, such as antibody production and macrophage activation, to eliminate infectious agents [47]. The ability of *HLA-DRB1* to present pathogen-derived peptides highlights its importance in clearing infections like HIV, *Bacillus anthracis*, and *Mycobacterium tuberculosis* [49-51]. Alleles such as *DRB101:01* and *DRB103:01* are particularly effective in this regard, enhancing immune responses and providing a robust defense against diverse viral and bacterial pathogens.

Notably, certain *HLA-DRB1* alleles efficiently present viral peptides, such as those from HIV and Rhinovirus, underscoring the gene's critical role in controlling viral load and clearing infections [52]. The *HLA-DRB1* molecule is key in clearing human rhinovirus (HRV) by presenting viral epitopes to T-helper 1 cells, activating immune responses. Various alleles, such as *DRB103:01, DRB104:04,* and *DRB1*15:01*, present epitopes from HRV-16 capsid proteins VP1 and VP2, including sequences like PRFSLPFLSIASAYYMFYDG and NEKQPSDDNWLNFDGTLLGN, to promote viral clearance [52]. Through this antigen presentation mechanism, *HLA-DRB1* is crucial in mounting a targeted immune defense to clear viral infections.

4.3. Tumor Immunity

HLA-DRB1 alleles, particularly *DRB1*04:05*, contribute significantly to tumor immunity by presenting immunogenic epitopes to T-helper 1 cells. These epitopes, derived from the tumor-associated antigen WT1, activate antigen-specific CD4-positive T cells, promoting a robust anti-timor immune response. For example, epitopes such as KRYFKLSHLQMHSRKH and MTEYKLVVVGAVGVGKSALTIQLI are recognized in a broad range of solid and hematological malignancies, potentially enhancing tumor immunesurveillance and aiding in the elimination of cancerous cells [53]. In the tumor microenvironment, tumor-resident antigen-presenting cells (APCs) may process and present peptides generated through mechanisms like phagocytosis of apoptotic tumor cells, thus further enhancing the immune system's ability to target and destroy malignant cells [54]. This interaction highlights the critical role of *MHC* class II molecules, which, in complex with *HLA-DRA*, display tumor antigens for recognition by the immune system, triggering T-helper effector functions that play a vital role in both tumor regression and long-term cancer control. This highlights *HLA-DRB1*'s therapeutic relevance in cancer immunotherapy.

4.4. Disease Susceptibility and Autoimmunity

HLA-DRB1 alleles are central to both immune defense and immune tolerance, playing a dual role in pathogen recognition and preventing autoimmune reactions. These alleles present both foreign and self-peptides to CD4-positive T cells, guiding immune responses that can either combat infections or maintain tolerance to the body's own tissues. For example, certain alleles, such as *DRB1*04:02*, present citrullinated self-peptides from proteins like vimentin, which are linked to autoimmune diseases such as *Rheumatoid arthritis* [55]. Similarly, *DRB1*15:01* presents self-peptides derived from myelin basic protein, contributing to the autoimmune response in multiple sclerosis [56, 57].

Key binding sites on the *HLA-DRB1* protein, such as those at positions 86, 90, 110, 111, and 122, are essential for its function in antigen recognition and presentation. These sites enable the MHC-II complex to bind peptide fragments, including those from pathogens and self-proteins, leading to the activation of T-helper cells and immune responses. This balance is critical for ensuring that while *HLA-DRB1* alleles trigger immune responses against foreign pathogens, they also maintain immune tolerance to prevent self-reactivity and autoimmune disease [6, 58].

HLA-DRB1 is crucial for autoimmune disease susceptibility, primarily due to its role in presenting peptides that can mimic self-antigens. The molecular mimicry theory suggests that microbial peptides resembling self-antigens can trigger autoimmunity, especially in genetically predisposed individuals. This is particularly evident in diseases like (MS) and *R. arthritis* (RA). For instance, the allele *DRB1*15:01* has a strong association with MS, where viral or bacterial peptides may

activate autoreactive T cells, leading to immune attacks on the myelin sheath [59, 60]. Similarly, alleles such as *DRB1*04:01* and *DRB1*04:05* increase susceptibility to RA, with variations in peptide-binding specificity contributing to anti-CCP-positive *R. arthritis* [55].

Moreover, specific *HLA-DRB1* alleles like *DRB1*01:03* are linked to heightened risks for Crohn's disease and *Ulcerative colitis*, with the recessive nature of this allele reflected in decreased heterozygosity among affected individuals [60]. For sarcoidosis, the allele *DRB1*04:02* is associated with specific phenotypes, such as ocular involvement [58]. Additionally, *DRB1*15:01* contributes to Goodpasture syndrome by presenting self-peptides derived from *COL4A3*, triggering pathogenic immune responses [60].

Accordingly, *HLA-DRB1* plays a crucial role in disease susceptibility, with specific alleles linked to autoimmune disorders, infections, and cancers. Variants like those associated with Rheumatoid arthritis and multiple sclerosis may increase disease risk by presenting self-peptides that trigger improper immune responses [55]. In cancer, *HLA-DRB1's* ability to present tumor-associated antigens is vital for tumor surveillance and therapy. It can activate T-helper cells, enhancing cytotoxic T cell responses against tumors, making *HLA-DRB1* a key target in immunotherapy, vaccine development, and adoptive T cell therapies [59, 60].

These findings emphasize *HLA-DRB1's* dual role in immune defense and tolerance to prevent autoimmunity. Its polymorphic nature enhances peptide presentation, raising autoimmune disease risk. Identifying specific alleles enables genetic screening for early detection and personalized treatments. Targeting allele-specific pathways could restore immune regulation and prevent disease onset in high-risk individuals. Notably, several SNPs within the HLA-DRB1 gene have been significantly associated with increased MS susceptibility, as detailed in Table 5.

4.5. Structural Impact on Function

Structural polymorphisms within the *HLA-DRB1* gene—particularly those located in the peptide-binding groove—have profound implications for antigen presentation and immune system modulation. The β1 domain of the *HLA-DRB1* protein, which forms part of the peptide-binding cleft, plays a central role in anchoring antigenic peptides through conserved hydrogen-bond interactions [57]. These interactions enable the molecule to present a wide variety of peptide antigens to CD4+ T cells, thereby supporting immune surveillance and adaptive responses. In parallel, the β2 domain, with its immunoglobulin-like fold, contributes to the interaction with the CD4 coreceptor, facilitating T-cell activation while preserving immune tolerance [61]. Key domains such as *MHC_II_beta* and *IgC1_MHC_II_beta_HLA-DR* further enhance the molecule's structural capacity to accommodate diverse antigenic peptides, which is essential for mounting effective immune responses against pathogens and aberrant cells, while limiting autoreactivity [6].

Polymorphisms in these domains are critical for shaping immune recognition and tolerance. Polymorphic residues across both conserved and variable regions of the β1 and β2 domains are critical in determining peptide-binding specificity and T-cell receptor (TCR) interactions. These structural variations shape the functional versatility of *HLA-DRB1* but also increase susceptibility to autoimmune diseases like MS and *R. arthritis* [6, 60]. Alleles such as *HLA-DRB1 15:01* have been strongly associated with multiple sclerosis (MS), particularly in European populations, while *HLA-DRB1 04:01* shows a well-established link to *R. arthritis*. Specific amino acid substitutions—especially at positions 40 and 60—have been implicated in impaired TCR engagement, potentially triggering abnormal immune activation in these disorders (see Figure 3).

Functionally, *HLA-DRB1* typically presents exogenous peptides derived from lysosomal degradation of extracellular proteins, generally 10–30 amino acids in length. In cancerous tissues, it also displays tumor-associated antigens originating from apoptotic cells or tumor-secreted proteins. Intriguingly, *HLA-DRB1* can also present endogenous peptides, which play a critical role in T-cell selection and the maintenance of central tolerance. The mechanism of antigen processing varies between foreign and self-peptides: exogenous peptides are usually loaded onto MHC class II molecules before trimming, whereas self-peptides undergo trimming prior to binding. Effective binding to the MHC II groove often requires the presence of a bulky hydrophobic residue at the N-terminal end of the peptide, which stabilizes the complex [58, 62].

The *HLA-DRB1* gene exhibits notable transcriptional diversity through multiple splice variants with distinct exon-intron configurations, contributing to functional heterogeneity. These transcript variants influence receptor dynamics, antigen presentation efficiency, and ultimately immune competence. Structural variability in the extracellular domains of *HLA-DRB1* alters the conformation of the peptide-binding cleft, thereby broadening its peptide repertoire and enhancing its immunological utility [58, 62].

Collectively, these findings emphasize the central role of structural and regulatory polymorphisms in shaping *HLA-DRB1*'s functional profile. Understanding how these variants affect antigen recognition, immune activation, and disease predisposition provides a promising foundation for the development of precision immunotherapies. Future studies targeting allele-specific variants may pave the way for more effective interventions in autoimmune disorders and cancer.

4.6. GO Annotations, Immune Function, and Mechanism of HLA-DRB1

Gene Ontology (GO) annotations highlight the multifaceted role of *HLA-DRB1* in cellular, biological, and molecular processes (Table 4). In terms of cellular components, *HLA-DRB1* is associated with various membranes, including the autolysosome membrane, cell surface, clathrin-coated endocytic vesicle membrane, and the extracellular exosome [49, 63, 64]. Its presence is also noted in structures like the Golgi membrane, lysosomal membrane, and the immunological synapse [49, 65, 66]. Functionally, *HLA-DRB1* engages in critical molecular interactions, such as CD4 receptor binding and *MHC* class II protein complex binding [67, 68]. It plays a vital role in antigen processing and presentation, specifically for endogenous and exogenous peptide antigens [63, 64, 69].

Biologically, *HLA-DRB1* is involved in essential processes like immune response [64], humoral immune response [70], and macrophage differentiation [68]. It regulates T cell activation and differentiation, demonstrating its influence on adaptive immunity [6, 47]. The diverse functionalities of *HLA-DRB1* are underscored by its association with various biological processes, including positive regulation of cytokine production and signal transduction [68, 71].

In terms of enzymatic and pathway interactions, *HLA-DRB1* is implicated in several key pathways related to T cell receptor signaling and *MHC* class II antigen presentation, as highlighted in resources like PathwayCommons and Reactome (Figure 4). These annotations reflect *HLA-DRB1's* critical role in maintaining immune homeostasis while participating in the immune defense against pathogens. Understanding these functions paves the way for genetic screening and personalized therapeutic strategies targeting specific immune pathways linked to *HLA-DRB1*.

The primary function of the HLA-DRB1 protein is its role as a beta chain within the antigen-presenting *MHC* class II molecule [65, 71]. In conjunction with the alpha chain *HLA-DRA, HLA-DRB1* is essential for displaying antigenic peptides to CD4+ T-helper cells, which is crucial for activating the immune response [49, 52]. When an antigen is processed by professional antigen-presenting cells (APCs), the peptide fragments generated—typically 10 to 30 amino acids in length—are loaded onto the HLA-DRB1 molecule within the intracellular compartment [71]. This process involves the proteolysis of the endocytosed antigens in lysosomes, followed by the presentation of these peptides at the cell surface for recognition by T cell receptors (TCRs) on CD4+ T cells [47, 66, 72].

HLA-DRB1 not only presents extracellular-derived antigens but also plays a significant role in presenting peptides from intracellular proteins [71]. This occurs through mechanisms such as macroautophagy, where proteins from within the cell are delivered to autolysosomes for processing [68]. The binding and presentation of these self-peptides are crucial for thymic selection and maintaining central immune tolerance [55, 56]. The processing pathways utilized by *HLA-DRB1* can differ based on the source of the peptide; pathogen-derived antigens typically follow a "bind first, cut/trim later" approach, while autoantigens often undergo a "cut first, bind later" process [73].

For the subcellular localization, the *HLA-DRB1*, a single-pass type I membrane protein, is located at the cell membrane, endoplasmic reticulum, lysosomes, late endosomes, and autolysosomes [49, 66, 74]. It transits through the endocytic pathway to the cell membrane for antigen presentation [62, 66] and functions in immunological synapses between T cells and antigen-presenting cells [49].

These GO enrichment results (see Table S1) underscore the regulatory network surrounding *HLA-DRB1* transcription, particularly through TFs like *RUNX1* and *SP1*, which are already known to modulate immune-related gene clusters. Their involvement in immune response pathways and transcriptional activation further supports a model where epigenetic and transcriptional control converge on the *HLA-DRB1* locus to shape disease susceptibility and immune function variability.

Overall, the study of *HLA-DRB1* highlights its essential function within the *MHC* class II complex, reinforcing its role in maintaining immune system integrity and providing a molecular basis for understanding autoimmune diseases and immune response deficiencies [71]. It interacts with CD4 and TCRs to initiate adaptive immunity, processes peptides via *MHCII* pathways, and plays central roles in immune homeostasis. The multifunctionality of *HLA-DRB1* in immune regulation and tissue-specific immunology is illustrated by its diverse cellular roles, as summarized in Table S3.

Table S3.
Functional Roles of *HLA-DRB1* in Cellular Processes.

Process	Description	Evidence Source
Antigen processing and presentation	Presents exogenous peptide antigens via MHC class II to CD4+ T cells	Gene Ontology
T-helper cell activation	Central to initiating Th1/Th17-mediated immune responses	KEGG Pathway
Cytokine signaling modulation	Influences downstream signaling cascades via T-cell receptor engagement	Reactome Database
Apoptosis regulation	Indirectly affects apoptosis via immune-mediated inflammation	Abelin, et al. [54]
Tissue-specific immune surveillance	Variable expression influences tissue-specific immune reactivity	ExpressionAtlas

4.7. Role of Post-Translational Modifications

Post-translational modifications (PTMs) like ubiquitination, glycosylation, and disulfide bond formation play a critical role in regulating the expression and function of *HLA-DRB1* [55]. These modifications influence antigen presentation and immune regulation by affecting the stability, folding, and trafficking of the *HLA-DRB1* protein [55]. For instance, glycosylation at asparagine 48 and disulfide bonds at positions 44-108 and 146-202 help stabilize the protein structure [55]. Ubiquitination at lysine 254, mediated by *MARCHF1* and *MARCHF8*, directs *HLA-DRB1* into the endosome system for down-regulation, impacting immune responses [66]. Alterations in these PTMs can affect immune function, potentially contributing to autoimmune disease development. The (PTMs) like glycosylation, disulfide bonds, and ubiquitination are critical for *HLA-DRB1* folding, stability, and immune presentation.

4.8. Gene expression and Protein-Protein Interactions

The *HLA-DRB1* gene is expressed primarily in professional antigen-presenting cells (APCs) such as monocytes/macrophages, dendritic cells, and B cells, as well as in thymic epithelial cells. This gene shows enhanced

expression in lung and lymphoid tissues and is detected across a broad range of other tissues, including the vermiform appendix [54, 58]. At the protein level, *HLA-DRB1* forms a heterotrimer consisting of an alpha chain (*HLA-DRA*), a beta chain (*HLA-DRB1*), and a peptide bound within the *MHCII* complex. In the endoplasmic reticulum (ER), the alpha and beta chains associate with the invariant chain (CD74/Ii), which blocks the peptide-binding cleft and prevents premature peptide binding [75]. This complex is transported from the ER through the Golgi apparatus to lysosomes, where it encounters antigenic peptides derived from endocytosed proteins. Here, enzymes like cathepsins degrade the invariant chain, leaving a segment called CLIP in the peptide-binding cleft, which is later displaced by the *HLA-DM* heterodimer to allow high-affinity peptide binding [76-78]. The *MHCII* complex can also interact with CD4 coreceptors and the T-cell receptor (TCR) to facilitate immune recognition [79]. Additionally, during microbial infections, *MHCII* can bind certain toxins from *Staphylococcus aureus* and viral proteins from Epstein-Barr virus, interacting outside the peptide-binding region [80].

Interactions between *HLA-DRB1* and *HLA-DM* are crucial for the immune system's ability to present antigens efficiently. This interaction ensures that only peptides with high affinity are presented to T cells, a critical process in immune surveillance [78]. These interactions facilitate peptide loading and presentation, while also revealing potential mechanisms of immune evasion by pathogens. Disruptions in this mechanism due to certain variants may lead to the improper activation of T cells, contributing to autoimmune diseases. Additionally, interactions with microbial antigens suggest that pathogens may exploit *HLA-DRB1* to evade immune detection or trigger autoimmune reactions through molecular mimicry [81].

While our findings offer valuable insights into the structural and regulatory mechanisms of *HLA-DRB1* in autoimmune pathogenesis, certain limitations should be acknowledged. The transcription factor predictions were conducted in silico using tools like MatInspector and remain computationally inferred without experimental validation. Additionally, allele-disease associations are largely based on existing databases, and population-level inferences may be restricted by ethnic and geographic variations. Future studies incorporating functional assays, diverse population cohorts, and epigenetic profiling are warranted to verify these associations and enhance translational applicability.

5. Conclusions

A comprehensive understanding of *HLA-DRB1*'s molecular mechanisms, particularly its transcriptional regulation by factors such as *ETS1*, *ETV1*, and *HOX* family members—is essential for developing more effective therapies for autoimmune diseases like multiple sclerosis (MS) and Rheumatoid arthritis, where current treatments often fail to provide lasting relief and may cause severe side effects. Our findings underscore the importance of allele-specific expression, especially of *HLA-DRB1 15:01*, which is strongly associated with MS susceptibility in diverse populations. Regulatory SNPs located in the 5′UTR and 3′UTR regions appear to influence transcription factor binding and alternative splicing patterns, offering mechanistic insight into differential gene expressions. By integrating genomic, transcriptomic, and structural data, we demonstrate how *HLA-DRB1* polymorphisms contribute to a wide range of immune functions—from pathogen clearance and tumor immunity to the development of autoimmunity. This system-level bioinformatics approach reveals that variations in gene regulation and expression, rather than sequence alone, may underlie disease susceptibility and progression. Future research focused on these regulatory networks could inform the design of personalized, targeted therapies. Moreover, leveraging bioinformatics tools to map transcriptional pathways presents promising opportunities to identify novel therapeutic targets, ultimately supporting the translational potential of modulating *HLA-DRB1* expression as a strategy to improve outcomes in autoimmune disorders, infections, and cancer

References

[1] A. K. Abbas, A. H. Lichtman, and S. Pillai, *Cellular and Molecular Immunology E-Book: Cellular and Molecular Immunology E-Book*. Elsevier Health Sciences, 2014.

[2] M. Wieczorek *et al.*, "Major histocompatibility complex (MHC) class I and MHC class II proteins: conformational plasticity in antigen presentation," *Frontiers in immunology,* vol. 8, p. 292, 2017.

[3] M. N. Lee and M. Meyerson, "Antigen identification for HLA class I–and HLA class II–restricted T cell receptors using cytokine-capturing antigen-presenting cells," *Science immunology,* vol. 6, no. 55, p. eabf4001, 2021.

[4] P. Selvaraj, D. N. Rajeswari, M. Jawahar, and P. Narayanan, "Influence of HLA-DRB1 alleles on Th1 and Th2 cytokine response to Mycobacterium tuberculosis antigens in pulmonary tuberculosis," *Tuberculosis,* vol. 87, no. 6, pp. 544-550, 2007.

[5] T. Wysocki, M. Olesińska, and A. Paradowska-Gorycka, "Current understanding of an emerging role of HLA-DRB1 gene in rheumatoid arthritis–from research to clinical practice," *Cells,* vol. 9, no. 5, p. 1127, 2020. https://doi.org/10.3390/cells9051127

[6] J. D. Ooi *et al.*, "Dominant protection from HLA-linked autoimmunity by antigen-specific regulatory T cells," (in eng), *Nature,* vol. 545, no. 7653, pp. 243-247, 2017.

[7] V. Matzaraki, V. Kumar, C. Wijmenga, and A. Zhernakova, "The MHC locus and genetic susceptibility to autoimmune and infectious diseases," *Genome Biology,* vol. 18, no. 1, p. 76, 2017. https://doi.org/10.1186/s13059-017-1207-1

[8] K. M. M. Koriem, "Multiple sclerosis: New insights and trends," *Asian Pacific Journal of Tropical Biomedicine,* vol. 6, no. 5, pp. 429-440, 2016.

[9] M. Kheradmand *et al.*, "Prevalence of subtypes of multiple sclerosis and the most common clinical symptoms in Iranian patients: A meta-analysis," *Clinical and Experimental Neuroimmunology,* vol. 10, no. 1, pp. 33-40, 2019.

[10] C. A. Dendrou, L. Fugger, and M. A. Friese, "Immunopathology of multiple sclerosis," *Nature Reviews Immunology,* vol. 15, no. 9, pp. 545-558, 2015. https://doi.org/10.1038/nri3871

[11] L. Handunnetthi, S. V. Ramagopalan, G. C. Ebers, and J. C. Knight, "Regulation of major histocompatibility complex class II gene expression, genetic variation and disease," *Genes & Immunity,* vol. 11, no. 2, pp. 99-112, 2010.

[12] L. Romero-Pinel *et al.*, "HLA-DRB1: genetic susceptibility and disability progression in a Spanish multiple sclerosis population," *European Journal of Neurology,* vol. 18, no. 2, pp. 337-342, 2011.
[13] D. Luckey, D. Bastakoty, and A. K. Mangalam, "Role of HLA class II genes in susceptibility and resistance to multiple sclerosis: Studies using HLA transgenic mice," *Journal of Autoimmunity,* vol. 37, no. 2, pp. 122-128, 2011. https://doi.org/10.1016/j.jaut.2011.05.001
[14] D. A. Dyment *et al.*, "Complex interactions among MHC haplotypes in multiple sclerosis: susceptibility and resistance," *Human Molecular Genetics,* vol. 14, no. 14, pp. 2019-2026, 2005.
[15] S. Katsavos and M. Anagnostouli, "Biomarkers in Multiple Sclerosis: An Up-to-Date Overview," (in eng), *Mult Scler Int,* vol. 2013, pp. 340508-340508, 2013.
[16] M.-T. Arango *et al.*, "HLA-DRB1 the notorious gene in the mosaic of autoimmunity," *Immunologic research,* vol. 65, pp. 82-98, 2017.
[17] C. Cheng *et al.*, "Understanding transcriptional regulation by integrative analysis of transcription factor binding data," (in eng), *Genome Res,* vol. 22, no. 9, pp. 1658-1667, 2012.
[18] J. P. Mpindi, "Bioinformatic tools for analysis, mining and modelling large-scale gene expression and drug testing datasets," 2016.
[19] W. Reith and B. Mach, "The bare lymphocyte syndrome and the regulation of MHC expression," *Annual Review of Immunology,* vol. 19, no. 1, pp. 331-373, 2001.
[20] R. Ferreira *et al.*, "High-Density SNP Mapping of the HLA Region Identifies Multiple Independent Susceptibility Loci Associated with Selective IgA Deficiency," *PLoS genetics,* vol. 8, p. e1002476, 2012.
[21] K. Singh *et al.*, "Deciphering the genetic landscape: Exploring the relationship between HLA-DQA1, HLA-DQB1, and HLA-DRB1 genes in diabetes mellitus," *Current Pharmacogenomics and Personalized Medicine,* vol. 21, no. 3, pp. 125-135, 2024.
[22] J. i. Satoh, "Bioinformatics approach to identifying molecular biomarkers and networks in multiple sclerosis," *Clinical and Experimental Neuroimmunology,* vol. 1, no. 3, pp. 127-140, 2010.
[23] T. James *et al.*, "Impact of genetic risk loci for multiple sclerosis on expression of proximal genes in patients," *Human Molecular Genetics,* vol. 27, no. 5, pp. 912-928, 2018. https://doi.org/10.1093/hmg/ddy001
[24] H. Westerlind *et al.*, "Modest familial risks for multiple sclerosis: a registry-based study of the population of Sweden," (in eng), *Brain,* vol. 137, no. Pt 3, pp. 770-778, 2014.
[25] M. Swami, "Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis," *Nature Medicine,* vol. 17, no. 9, pp. 1059-1059, 2011.
[26] L. Moutsianas *et al.*, "Class II HLA interactions modulate genetic risk for multiple sclerosis," (in eng), *Nat Genet,* vol. 47, no. 10, pp. 1107-1113, 2015.
[27] C. A. Dendrou, J. Petersen, J. Rossjohn, and L. Fugger, "HLA variation and disease," *Nature Reviews Immunology,* vol. 18, no. 5, pp. 325-339, 2018.
[28] A. Nova, T. Fazia, V. Saddi, M. Piras, and L. Bernardinelli, "Multiple sclerosis heritability estimation on sardinian ascertained extended families using Bayesian liability threshold model," *Genes,* vol. 14, no. 8, p. 1579, 2023.
[29] G. Marchetti *et al.*, "Changes in expression profiles of internal jugular vein wall and plasma protein levels in multiple sclerosis," (in eng), *Mol Med,* vol. 24, no. 1, pp. 42-42, 2018.
[30] D. M. Kofler, A. Farkas, M. v. Bergwelt-Baildon, and D. A. Hafler, "The link between CD6 and autoimmunity: genetic and cellular associations," *Current drug targets,* vol. 17, no. 6, pp. 651-665, 2016.
[31] R. Kozuka *et al.*, "Association between HLA-DQA1/DRB1 polymorphism and development of hepatocellular carcinoma during entecavir treatment," *Journal of Gastroenterology and Hepatology,* vol. 34, no. 5, pp. 937-946, 2019.
[32] S. V. Ramagopalan *et al.*, "Expression of themultiple sclerosis-associated MHC class II allele HLA-DRB1*1501 Is regulated by vitamin D," *PLOS Genetics,* vol. 5, no. 2, p. e1000369, 2009. https://doi.org/10.1371/journal.pgen.1000369
[33] C. Riveros *et al.*, "A transcription factor map as revealed by a genome-wide gene expression analysis of whole-blood mRNA transcriptome in multiple sclerosis," (in eng), *PLoS One,* vol. 5, no. 12, pp. e14176-e14176, 2010.
[34] S. Nataf, M. Guillen, and L. Pays, "TGFB1-mediated gliosis in multiple sclerosis spinal cords is favored by the regionalized expression of HOXA5 and the age-dependent decline in androgen receptor ligands," *International Journal of Molecular Sciences,* vol. 20, no. 23, p. 5934, 2019. https://doi.org/10.3390/ijms20235934
[35] E. Mahdipour and K. A. Mace, "Hox transcription factor regulation of adult bone-marrow-derived cell behaviour during tissue repair and regeneration," *Expert Opin Biol Ther,* vol. 11, no. 8, pp. 1079-90, 2011.
[36] M. Kouwenhoven *et al.*, "Multiple sclerosis: elevated expression of matrix metalloproteinases in bloodmonocytes," *Journal of Autoimmunity,* vol. 16, no. 4, pp. 463-470, 2001. https://doi.org/10.1006/jaut.2001.0505
[37] A. Suzuki, Y. Kochi, Y. Okada, and K. Yamamoto, "Insight from genome-wide association studies in rheumatoid arthritis and multiple sclerosis," *FEBS Letters,* vol. 585, no. 23, pp. 3627-3632, 2011.
[38] H. Shams *et al.*, "Characterization of a mycobacterium tuberculosis peptide that is recognized by human CD4+ and CD8+ T cells in the context of multiple HLA alleles," *The Journal of Immunology,* vol. 173, no. 3, pp. 1966-1977, 2004.
[39] M. Eggert, A. Kluter, U. K. Zettl, and G. Neeck, "Transcription Factors in Autoimmune Diseases," *Frontiers in Medicinal Chemistry,* vol. 383, no. 397, pp. 383-397, 2006.
[40] T. Zeis, U. Graumann, R. Reynolds, and N. Schaeren-Wiemers, "Normal-appearing white matter in multiple sclerosis is in a subtle balance between inflammation and neuroprotection," *Brain,* vol. 131, no. 1, pp. 288-303, 2008.
[41] G. Kristjansdottir *et al.*, "Interferon regulatory factor 5 (IRF5) gene variants are associated with multiple sclerosis in three distinct populations," (in eng), *J Med Genet,* vol. 45, no. 6, pp. 362-369, 2008. 10.1136/jmg.2007.055012
[42] I. Gerhauser, S. Alldinger, and W. Baumgärtner, "Ets-1 represents a pivotal transcription factor for viral clearance, inflammation, and demyelination in a mouse model of multiple sclerosis," *Journal of Neuroimmunology,* vol. 188, no. 1-2, pp. 86-94, 2007.
[43] G. Frisullo *et al.*, "Regulatory T cells fail to suppress CD4T+-bet+ T cells in relapsing multiple sclerosis patients," (in eng), *Immunology,* vol. 127, no. 3, pp. 418-428, 2009.
[44] K. Venken *et al.*, "Secondary progressive in contrast to relapsing-remitting multiple sclerosis patients show a normal CD4⁺CD25⁺ regulatory T-cell function and FOXP3 expression," *Journal of Neuroscience Research,* vol. 83, no. 8, pp. 1432-1446, 2006.

[45] B. Li, Q. Huang, and G.-H. Wei, "The role of HOX transcription factors in cancer predisposition and progression," *Cancers,* vol. 11, no. 4, p. 528, 2019. https://doi.org/10.3390/cancers11040528

[46] E. Mahdipour and K. A. Mace, "Hox transcription factor regulation of adult bone-marrow-derived cell behaviour during tissue repair and regeneration," *Expert Opinion on Biological Therapy,* vol. 11, no. 8, pp. 1079-1090, 2011.

[47] H. Shams *et al.*, "Characterization of a Mycobacterium tuberculosis peptide that is recognized by human CD4+ and CD8+ T cells in the context of multiple HLA alleles," (in eng), *Journal of immunology (Baltimore, Md. : 1950),* vol. 173, no. 3, pp. 1966-77, 2004.

[48] J. Schulze zur Wiesch *et al.*, "Broad repertoire of the CD4+ Th cell response in spontaneously controlled hepatitis C virus infection includes dominant and highly promiscuous epitopes," (in eng), *Journal of immunology (Baltimore, Md. : 1950),* vol. 175, no. 6, pp. 3603-13, 2005.

[49] M. Galperin *et al.*, "CD4(+) T cell-mediated HLA class II cross-restriction in HIV controllers," (in eng), *Sci Immunol,* vol. 3, no. 24, 2018.

[50] W. W. Kwok *et al.*, "Frequency of epitope-specific naive CD4(+) T cells correlates with immunodominance in the human memory repertoire," (in eng), *Journal of immunology (Baltimore, Md. : 1950),* vol. 188, no. 6, pp. 2537-44, 2012.

[51] Y. Wang *et al.*, "Systemic immune dysregulation in severe tuberculosis patients revealed by a single-cell transcriptome atlas," *Journal of Infection,* vol. 86, no. 5, pp. 421-438, 2023.

[52] L. M. Muehling, D. T. Mai, W. W. Kwok, P. W. Heymann, A. Pomés, and J. A. Woodfolk, "Circulating Memory CD4+ T Cells Target Conserved Epitopes of Rhinovirus Capsid Proteins and Respond Rapidly to Experimental Infection in Humans," (in eng), *Journal of immunology (Baltimore, Md. : 1950),* vol. 197, no. 8, pp. 3214-3224, 2016.

[53] S. Anguille *et al.*, "Identification of a Wilms' tumor 1-derived immunogenic CD4(+) T-cell epitope that is recognized in the context of common Caucasian HLA-DR haplotypes," (in eng), *Leukemia,* vol. 27, no. 3, pp. 748-750, 2013.

[54] J. G. Abelin *et al.*, "Defining HLA-II Ligand Processing and Binding Rules with Mass Spectrometry Enhances Cancer Epitope Prediction," (in eng), *Immunity,* vol. 51, no. 4, pp. 766-779.e17, 2019.

[55] S. W. Scally *et al.*, "A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis," (in eng), *The Journal of experimental medicine,* vol. 210, no. 12, pp. 2569-82, 2013

[56] A. Kim *et al.*, "Divergent paths for the selection of immunodominant epitopes from distinct antigenic sources," (in eng), *Nature communications,* vol. 5, p. 5369, 2014.

[57] K. J. Smith, J. Pyrdol, L. Gauthier, D. C. Wiley, and K. W. Wucherpfennig, "Crystal structure of HLA-DR2 (DRA*0101, DRB1*1501) complexed with a peptide from human myelin basic protein," (in eng), *The Journal of experimental medicine,* vol. 188, no. 8, pp. 1511-20, 1998.

[58] E. Adamopoulou *et al.*, "Exploring the MHC-peptide matrix of central tolerance in the human thymus," (in eng), *Nature communications,* vol. 4, p. 2039, 2013.

[59] S. Sawcer *et al.*, "Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis," (in eng), *Nature,* vol. 476, no. 7359, pp. 214-9, 2011.

[60] M. Harkiolaki *et al.*, "T cell-mediated autoimmune disease due to low-affinity crossreactivity to common microbial peptides," (in eng), *Immunity,* vol. 30, no. 3, pp. 348-57, 2009.

[61] P. Jönsson *et al.*, "Remarkably low affinity of CD4/peptide-major histocompatibility complex class II protein interactions," (in eng), *Proceedings of the National Academy of Sciences of the United States of America,* vol. 113, no. 20, pp. 5682-7, 2016.

[62] D. Schmid, M. Pypaert, and C. Münz, "Antigen-loading compartments for major histocompatibility complex class II molecules continuously receive input from autophagosomes," (in eng), *Immunity,* vol. 26, no. 1, pp. 79-92, 2007.

[63] S. O. Ryan, J. A. Bonomo, F. Zhao, and B. A. Cobb, "MHCII glycosylation modulates Bacteroides fragilis carbohydrate antigen presentation," (in eng), *The Journal of experimental medicine,* vol. 208, no. 5, pp. 1041-53, 2011.

[64] A. Aslam, H. Chan, D. A. Warrell, S. Misbah, and G. S. Ogg, "Tracking antigen-specific T-cells during clinical tolerance induction in humans," (in eng), *PLoS One,* vol. 5, no. 6, p. e11028, 2010.

[65] T. A. Röhn *et al.*, "Upregulation of the CLIP self peptide on mature dendritic cells antagonizes T helper type 1 polarization," (in eng), *Nature immunology,* vol. 5, no. 9, pp. 909-18, 2004.

[66] A. De Gassart *et al.*, "MHC class II stabilization at the surface of human dendritic cells is the result of maturation-dependent MARCH I down-regulation," (in eng), *Proceedings of the National Academy of Sciences of the United States of America,* vol. 105, no. 9, pp. 3491-6, 2008.

[67] S. I. Buschow, B. W. van Balkom, M. Aalberts, A. J. Heck, M. Wauben, and W. Stoorvogel, "MHC class II-associated proteins in B-cell exosomes and potential functional implications for exosome biogenesis," (in eng), *Immunology and cell biology,* vol. 88, no. 8, pp. 851-6, 2010.

[68] A. Zhen, S. R. Krutzik, B. R. Levin, S. Kasparian, J. A. Zack, and S. G. Kitchen, "CD4 ligation on human blood monocytes triggers macrophage differentiation and enhances HIV infection," (in eng), *Journal of virology,* vol. 88, no. 17, pp. 9934-46, 2014.

[69] T. Yoon *et al.*, "Mapping the HLA-DO/HLA-DM complex by FRET and mutagenesis," (in eng), *Proceedings of the National Academy of Sciences of the United States of America,* vol. 109, no. 28, pp. 11276-81, 2012.

[70] I. Djilali-Saiah, A. Fakhfakh, H. Louafi, S. Caillat-Zucman, D. Debray, and F. Alvarez, "HLA class II influences humoral autoimmunity in patients with type 2 autoimmune hepatitis," (in eng), *Journal of hepatology,* vol. 45, no. 6, pp. 844-50, 2006.

[71] H. Huang, C. Wang, F. Rubelt, T. J. Scriba, and M. M. Davis, "Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening," (in eng), *Nature biotechnology,* vol. 38, no. 10, pp. 1194-1202, 2020.

[72] R. Yossef *et al.*, "Enhanced detection of neoantigen-reactive T cells targeting unique and shared oncogenes for personalized cancer immunotherapy," (in eng), *JCI insight,* vol. 3, no. 19, 2018.

[73] M. G. Gutierrez-Hoffmann, R. N. O'Meally, R. N. Cole, E. Tiniakou, E. Darrah, and M. J. Soloski, "Borrelia burgdorferi-Induced Changes in the Class II Self-Immunopeptidome Displayed on HLA-DR Molecules Expressed by Dendritic Cells," (in eng), *Frontiers in medicine,* vol. 7, p. 568, 2020.

[74] R. Faner, E. James, L. Huston, R. Pujol-Borrel, W. W. Kwok, and M. Juan, "Reassessing the role of HLA-DRB3 T-cell responses: evidence for significant expression and complementary antigen presentation," (in eng), *European journal of immunology,* vol. 40, no. 1, pp. 91-102, Jan 2010.

[75] S. J. Park, S. Sadegh-Nasseri, and D. C. Wiley, "Invariant chain made in Escherichia coli has an exposed N-terminal segment that blocks antigen binding to HLA-DR1 and a trimeric C-terminal segment that binds empty HLA-DR1," (in eng), *Proceedings of the National Academy of Sciences of the United States of America,* vol. 92, no. 24, pp. 11289-93, 1995.

[76] H. Kropshofer, S. O. Arndt, G. Moldenhauer, G. J. Hämmerling, and A. B. Vogt, "HLA-DM acts as a molecular chaperone and rescues empty HLA-DR molecules at lysosomal pH," (in eng), *Immunity,* vol. 6, no. 3, pp. 293-302, 1997.

[77] P. Ghosh, M. Amaya, E. Mellins, and D. C. Wiley, "The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3," (in eng), *Nature,* vol. 378, no. 6556, pp. 457-62,1995.

[78] W. Pos *et al.*, "Crystal structure of the HLA-DM-HLA-DR1 complex defines mechanisms for rapid peptide selection," (in eng), *Cell,* vol. 151, no. 7, pp. 1557-68, 2012.

[79] X. X. Wang *et al.*, "Affinity maturation of human CD4 by yeast surface display and crystal structure of a CD4-HLA-DR1 complex," (in eng), *Proceedings of the National Academy of Sciences of the United States of America,* vol. 108, no. 38, pp. 15960-5, 2011.

[80] K. Petersson *et al.*, "Crystal structure of a superantigen bound to MHC class II displays zinc and peptide dependence," (in eng), *The EMBO journal,* vol. 20, no. 13, pp. 3306-12, 2001.

[81] S. H. Krovi and V. K. Kuchroo, "Activation pathways that drive CD4+ T cells to break tolerance in autoimmune diseases," *Immunological Reviews,* vol. 307, no. 1, pp. 161-190, 2022.

[82] S. I. Buschow, B. W. Van Balkom, M. Aalberts, A. J. Heck, M. Wauben, and W. Stoorvogel, "MHC class II-associated proteins in B-cell exosomes and potential functional implications for exosome biogenesis," *Immunology and Cell Biology,* vol. 88, no. 8, pp. 851-856, 2010.

[83] R. Wubbolts *et al.*, "proteomic and biochemical analyses of human b cell-derived exosomes: Potential implications for their function and multivesicular body formation " *Journal of Biological Chemistry,* vol. 278, no. 13, pp. 10963-10972, 2003. https://doi.org/10.1074/jbc.M207550200

[84] M. Gonzalez-Begne *et al.*, "Proteomic analysis of human parotid gland exosomes by multidimensional protein identification technology (MudPIT)," *Journal of Proteome Research,* vol. 8, no. 3, pp. 1304-1314, 2009.

[85] D. J. Palmer, V. C. Kelly, A. M. Smit, S. Kuy, C. G. Knight, and G. J. Cooper, "Human colostrum: identification of minor proteins in the aqueous phase by proteomics," (in eng), *Proteomics,* vol. 6, no. 7, pp. 2208-16, Apr 2006.

[86] M. Galperin *et al.*, "CD4⁺ T cell–mediated HLA class II cross-restriction in HIV controllers," *Science Immunology,* vol. 3, no. 24, p. eaat0687, 2018. https://doi.org/10.1126/sciimmunol.aat0687

[87] T. A. Röhn *et al.*, "Upregulation of the CLIP self peptide on mature dendritic cells antagonizes T helper type 1 polarization," *Nature Immunology,* vol. 5, no. 9, pp. 909-918, 2004. https://doi.org/10.1038/ni1108

[88] S. M. Troyanovsky, R. E. Leube, and W. W. Franke, "Characterization of the human gene encoding cytokeratin 17 and its expression pattern," (in eng), *European journal of cell biology,* vol. 59, no. 1, pp. 127-37, 1992.

[89] D. Ghosh, D. Lippert, O. Krokhin, J. P. Cortens, and J. A. Wilkins, "Defining the membrane proteome of NK cells," (in eng), *Journal of mass spectrometry : JMS,* vol. 45, no. 1, pp. 1-25, 2010.

[90] S. O. Ryan, J. A. Bonomo, F. Zhao, and B. A. Cobb, "MHCII glycosylation modulates bacteroides fragilis carbohydrate antigen presentation," *Journal of Experimental Medicine,* vol. 208, no. 5, pp. 1041-1053, 2011.

[91] T. Yoon *et al.*, "Mapping the HLA-DO/HLA-DM complex by FRET and mutagenesis," *Proceedings of the National Academy of Sciences,* vol. 109, no. 28, pp. 11276-11281, 2012. https://doi.org/10.1073/pnas.1113966109

[92] S. Hiroi, H. Harada, H. Nishi, M. Satoh, R. Nagai, and A. Kimura, "Polymorphisms in the SOD2 and HLA-DRB1 genes are associated with nonfamilial idiopathic dilated cardiomyopathy in Japanese," (in eng), *Biochemical and biophysical research communications,* vol. 261, no. 2, pp. 332-9,1999.

[93] Z. Shen, G. Wang, J. Y. Fan, W. Li, and Y. F. Liu, "HLA DR B1*04, *07-restricted epitopes on Keratin 17forautoreactive T cells in psoriasis," (in eng), *Journal of dermatological science,* vol. 38, no. 1, pp. 25-39, Apr 2005.

[94] L. B. Giarola, R. R. Dos Santos, J. Bedendo, W. V. da Silva Júnior, and S. D. Borelli, "HLA molecules and nasal carriage of Staphylococcus aureus isolated from dialysis and kidney transplant patients at a hospital in Southern Brazil," (in eng), *BMC research notes,* vol. 5, p. 90, 2012.

[95] W. H. McLean *et al.*, "Keratin 16 and keratin 17 mutations cause pachyonychia congenita," (in eng), *Nat Genet,* vol. 9, no. 3, pp. 273-8, 1995.

[96] S. K. Cordovado *et al.*, "Nephropathy in type 1 diabetes is diminished in carriers of HLA-DRB1*04: the genetics of kidneys in diabetes (GoKinD) study," (in eng), *Diabetes,* vol. 57, no. 2, pp. 518-22, 2008.

[97] R. C. Williams *et al.*, "HLA-DRB1 reduces the risk of type 2 diabetes mellitus by increased insulin secretion," (in eng), *Diabetologia,* vol. 54, no. 7, pp. 1684-92, 2011.

[98] M. Živković *et al.*, "The tag SNP for HLA-DRB1∗1501, rs3135388, is significantly associated with multiple sclerosis susceptibility: Cost-effective high-throughput detection by real-time PCR," *Clinica Chimica Acta,* vol. 406, no. 1, pp. 27-30, 2009. https://doi.org/10.1016/j.cca.2009.05.004

[99] B. Brynedal *et al.*, "HLA-A confers an HLA-DRB1 independent influence on the risk of multiple sclerosis," *PLoS One,* vol. 2, no. 7, p. e664, 2007.

[100] S. P. Barjui, Z. F. Shahrakei, and N. Jivad, "Association of HLA rs3135388 and rs9271366 single-nucleotide polymorphisms with multiple sclerosis in the chaharmahal and bakhtiari province, Iran," *Epidemiology and Health System Journal,* vol. 10, no. 4, pp. 170-174, 2023.