Article

Bioinformatics and connectivity map analysis suggest viral infection as a critical causative factor of Hashimoto's Thyroiditis

Dong-Woo Lim [1,2], Min-Seo Choi [1] and Seok-Mo Kim[3]*

[1] Department of Diagnostics, College of Korean Medicine, Dongguk University, Goyang, Republic of Korea, 10326; Lim, D.-W., greatwoodong@dongguk.edu; Choi, M.-S, msshys123@naver.com
[2] Institute of Korean Medicine, Dongguk University, Goyang, Republic of Korea, 10326;
[3] Gangnam Severance Hospital, Department of Surgery, Yonsei University College of Medicine, Seoul, Republic of Korea, 06273
* Correspondence: K, S.-M., seokmokim@yuhs.ac; Tel.: +82-2-2019-3370

Abstract: Hashimoto's thyroiditis (HT) is a common autoimmune disease, with its prevalence rapidly increasing. Both genetic and environmental risk factors contribute to the development of HT. Recently, viral infection has been suggested to act as a trigger of HT by eliciting the host immune response and subsequent autoreactivity. We analyzed features of HT through bioinformatics analysis so as to identify markers of HT development. We accessed public microarray data of HT patients from the Gene Expression Omnibus (GEO) and obtained differentially expressed genes (DEGs) under HT. Gene Ontology (GO) and KEGG pathway enrichment analyses were performed for functional clustering of our protein-protein interaction (PPI) network. Utilizing ranked gene lists, we performed Gene Set Enrichment Analysis (GSEA) using the clusterprofiler R package. By comparing the expression signatures of the huge perturbation database with the queried rank-ordered gene list, connectivity map (CMap) analysis was performed to screen potential therapeutic targets and agents. The gene expression profile of the HT group was in line with general characteristics of HT. Biological processes related to the immune response and viral infection pathways were obtained for the upregulated DEGs. GSEA results revealed activation of autoimmune disease-related pathways and several viral infection pathways. Autoimmune disease and viral infection pathways were highly interconnected by common genes, while the HLA genes which are shared by both were significantly upregulated. CMap analysis suggested that perturbagens, including SRRM1, NLK, and CCDC92, have the potential to reverse the HT expression profile. Several lines of evidence suggested that viral infection and the host immune response are activated during HT. Viral infection is suspected to act as a key trigger of HT by causing autoimmunity. SRRM1, an alternative splicing factor, which responds to viral activity, might serve as potential marker of HT.

Keywords: Hashimoto's thyroiditis, bioinformatics, viral infection, mRNA splicing, GSEA, autoimmune disease, CMap

1. Introduction

Hashimoto's thyroiditis (HT) is one of the most prevalent autoimmune diseases (AIDs) [1] and a common cause of hypothyroidism in developed countries [2]. Pathologic features of HT include extensive inflammation in thyroid tissue caused by infiltrating CD4+ T lymphocytes [3] and the presence of autoantibodies causing follicular cell damage and ultimately leading to disrupted thyroid function [4]. The clinical phenotype of HT can vary from asymptomatic to severe symptoms of hypothyroidism, including weight gain, menstrual disorders, and heat intolerance [5]. Although HT is organ-specific disease, yet the disease might be related with other AID or systemic autoimmune disorders in many cases [6]. Proposed risk factors for HT include genetic susceptibility, sex (female), senescence as well as environmental triggers such as iodine uptake, drugs, chem-

 doi:10.20944/preprints202212.0275.v1

icals, and viral infections [7]. However, the exact mechanism underlying thyroid autoimmunity in HT remains unclear.

AIDs lead to tissue injury caused by the autoreactivity of T- and B cell-mediated responses [8]. Mechanisms of immune tolerance generally prevent autoimmune reactions through the fine tuning process of positive and negative selection of potentially autoreactive lymphocytes [9]. The failure of these selection mechanisms gives rise to autoreactive T cells or antibodies, which may appear several years prior to the manifestation of AID [10,11]. Genetic susceptibility and environmental factors determine one's predisposition to AID [12].

Genetic predisposition and expression changes influence the autoimmune thyroid disease course by affecting host immunoreactivity or antigen presentation/recognition [12]. Polymorphisms in IL2 and CTLA4 or the upregulation of MHC class II molecules are implicated in the etiology of AID [10]. MHC, which is also known as human leukocyte antigen (HLA) in humans, plays a key role in AID by helping the immune system distinguish self from foreign, owing to its unlimited allelic diversity [13]. HLA alleles (HLA-DPB1) and their variants (HLA-DPB1*02:02 and HLA-DPB1*05:01) have been described as contributors to the early pathogenesis of autoimmune thyroiditis [14]. However, genetic susceptibility alone is often insufficient to give rise to autoimmunity [15].

Recently, viral infection has emerged as an attractive environmental trigger of autoimmunity, with multiple mechanisms described for different AIDs [16]. Molecular mimicry, bystander activation, and epitope spreading are the three major mechanisms underlying virus-induced autoimmunity [17]. Molecular mimicry is based on cross-reactivity due to the structural similarity between pathogen (viral particles) and self (self-antigens), providing a basis for virus-induced autoimmunity [18]. Subsequent tissue damage results in the release of damage-associated molecular patterns (DAMPs) that activate TLRs, leading to amplified immune activation [19]. Likewise, biomolecule based observation has been indicating viral infections as a key factor in induction and development of autoimmune diseases.

A case study reported HT onset after herpes simplex virus infection (3 to 6 months) in three patients, as indicated by the presence of IgM and IgG antibodies against the virus [20]. A study of 42 HT patients revealed high prevalence of Epstein-Barr virus infection in patient tissue (n=42) [21]. A larger study showed a clear association of hepatitis C virus (HCV) infection and thyroid autoimmunity [22]. A systematic review of 12 studies compared epidemiological differences in thyroid dysfunction between HCV-infected and non-HCV-infected, indicating an increased risk of thyroid dysfunction in the former [23].

Recent advances in microarray profiling and bioinformatics tools have enabled the analysis of massive transcriptome data, providing insight into the expression changes observed during various diseases. The Gene Expression Omnibus (GEO) database is a public repository with microarray, next-generation sequencing (NGS), and other genomics data, which provides access to large datasets submitted by various researchers [24]. Gene Set Enrichment Analysis (GSEA) is a powerful analytical method used for interpreting gene expression data via Gene Ontology (GO) terms or other gene set collections [25]. The Connectivity Map (CMap), a comprehensive, large-scale perturbation database containing 1.5 million gene expression profiles from cultured human cells, can be used to identify potential therapeutic targets or drugs for the submitted gene signature [26]. These bioinformatics tools can be utilized to address various biomedical issues by deciphering information hidden in large amount of biological datasets [27].

Investigation on biomarkers of HT should be conducted to improve diagnosis and provide feasible medical treatment. A previous study on HT utilized the GEO microarray database, yet only suggested hub genes deduced from protein-protein interaction (PPI) network of DEGs [28]. In this work, we extracted a list of DEGs from the GEO microarray

 doi:10.20944/preprints202212.0275.v1

database and divided them into functional clusters via PPI network construction and performed over-representation analysis ORA. Based on a ranked gene list, we conducted GSEA to scrutinize skewed distribution of genes related with specific BP terms and KEGG pathways in the HT group, visualizing these through various R packages. Finally, we identified promising targets and compounds for HT via CMap analysis.

2. Materials and Methods

Data resources and processing

The human microarray dataset of GSE138198 [29] was accessed via the GEO database (https://www.ncbi.nlm.nih.gov/geo/) at the National Center of Biotechnology Information (NCBI) [24]. The GSE138198 dataset comprises a total of 36 human thyroid tissue microarray samples, including 13 HT tissue samples and 3 normal thyroid (TN) tissue samples. The microarray dataset is based on the GPL6244 [HuGene-1_0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version].

Screening of DEGs

GEO2R, an interactive web tool for the analysis of GEO datasets, was used to screen DEGs between HT and TN (http://www.ncbi.nlm.nih.gov/geo/geo2r/). GEO2R identified DEGs via GEO queries and the Limma R package from R/bioconductor [30,31]. Genes with an adjusted p-value < 0.05 and $|\log_2(\text{Fold Change})| > 1$ were considered DEGs. Adjustment of the P-value was performed via the Benjamini & Hochberg method to reduce false discovery rate. DEGs were visualized with a color-differentiated (by significance and fold change) and labeled volcano plot using ggplot R packages. Clustered correlation heatmap of all samples was created using the 'pheatmap' package.

Construction of PPI network and functional subcluster analysis

We used the Search Tool for the Retrieval of Interacting Genes (STRING, https://string-db.org/) online database to construct a PPI of up- or downregulated DEGs [32]. PPI networks were further imported to Cytoscape software (version 3.91) [33] for functional cluster analysis, which groups genes with similar functions, using the MCODE plugin to yield the top 3 clusters by score [34]. Sets of genes from these clusters were separately subjected to over-representation analysis.

GO and KEGG pathway enrichment analyses

GO and KEGG pathway enrichment analyses of DEGs were conducted using The Database for Annotation, Visualization and Integrated Discovery (DAVID database, Accessed on 31th July) [35]. Lists of clustered genes were uploaded with the identifier set to "official gene symbol". The data of enriched terms and KEGG pathways were processed and further visualized as a bubble plot created using the 'ggplot2' R package [36].

GSEA of GO and KEGG pathway

GSEA of all ranked genes was performed using the clusterprofiler R package and the genome-wide annotation package (OrgDb) of Bioconductor [37]. The package provides gseGO and gseKEGG functions for GSEA using GO (Biological process, Molecular function, Cellular component) and KEGG annotations. The list of significant gene set annotations was arranged in the order of Normalized enriched score (NES).

Argument parameters used in gseGO and gseKEGG were as follows; nPerm =10,000, minGSsize =3, maxGSSize = 800, pvalueCutoff = 0.05, pAdjustMethod = "none". Redundant GO terms were re-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1

moved via the simplify function [38]. Visualization of results obtained from ORA and GSEA analyses was performed using the "enrichplot" package [39].

Clinical HT biomarker identification via comparison of gene signatures in the CMap database

We selected the top 150 up- and downregulated DEGs of HT groups and submitted these to the Connectivity Map for analysis by querying the list on the CLUE web tool (https://clue.io/query/, version 1.1.1.43). The results present perturbagens, including KD (gene knock down), CP (compounds), OE (gene overexpression), and PCL (perturbagen lists), aligned by the calculated Connectivity score (median tau score), which varies from 100 to -100.

A negative connectivity score indicates an inversed gene expression signature between the query and perturbagen, which implies potential as a promising target or drug [40]. A list of rank-ordered perturbagens with a CMap connectivity score (tau) < -90 were selected and regarded as significant candidates.

3. Results

Validation of GEO data

Processed data from the GSE138198 human thyroid tissue microarray dataset was validated by checking distribution in a boxplot and heatmap clustering of samples. The boxplot revealed median-centered values, indicating that the data are well-normalized for all samples (Fig. 1A). A UMAP plot located each sample in a reduced dimension using Euclidean distance, concluding that samples within groups are in close proximity (HT vs TN), and, on the contrary, at long distance between the two groups (Fig. 1B). Likewise, the hierarchical correlation heatmap indicated that samples are clustered by groups (Fig. 1C).

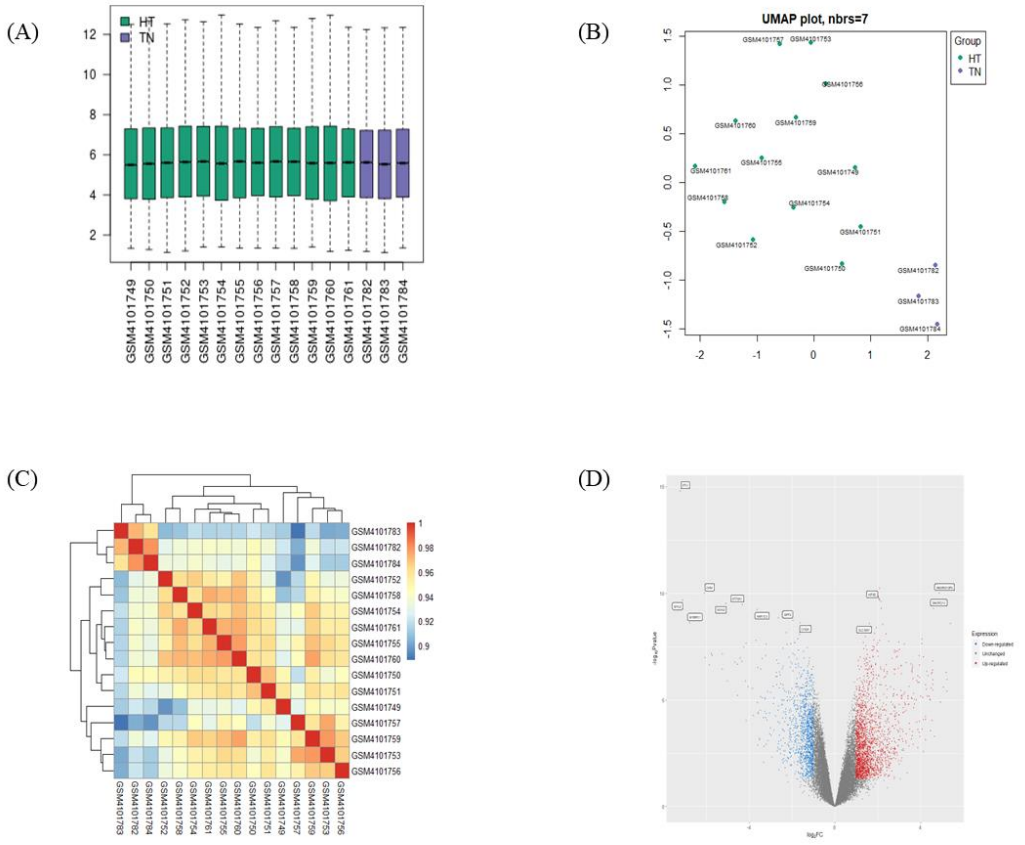


Fig. 1 Validation of selected gene expression profiles from GSE138198. A total 16 human samples were included, 13

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1

in the HT group and 3 in the TN group. (A) Box plot displays distributions of values of the selected samples from GSE138198 after normalization. (B) UMAP plot of 16 normalized samples embedded in an Euclidean space. (C) Clustered correlation heatmap showing correlation of microarray profile between 16 samples. (D) Volcano plot of significant differentially expressed genes (DEGs) in HT (adjusted p-value < 0.05 and $|\log_2 FC| > 1$). Plots were created using GEO2R or pheatmap R package.

Identification of DEGs extracted from HT

Volcano plot displayed the distribution of DEGs by their fold change and P-value distinguishable by color (Fig. 1D). After the validation of gene identifiers, a total of 838 upregulated DEGs and 595 downregulated DEGs were confirmed between the HT vs TN groups. Lists of the 15 most significant up- or downregulated genes are presented in Table 1. KIF5B and PTH were confirmed as the most significantly up- and downregulated DEGs, respectively, in the HT group as compared to TN group.

Table 1. List of Top 15 up- or downregulated Differentially Expressed genes (protein-coding genes only). Genes with |Log2(FC)|>1 and p-value<0.05 were considered as DEGs.

	Gene symbol	Gene name	$\log_2$(fold change)	$-\text{Log}_{10}$(P)
Up regulated	KIF5B	kinesin family member 5B	2.092	9.687
	SLC30A7	solute carrier family 30 member 7	1.613	8.584
	RHNO1	RAD9-HUS1-RAD1 interacting nuclear orphan 1	1.948	8.386
	PGK1	phosphoglycerate kinase 1	1.628	8.321
	ATP5EP2	ATP synthase, H+ transporting, mitochondrial F1 complex, epsilon subunit pseudogene 2	1.865	8.107
	PIGS	phosphatidylinositol glycan anchor biosynthesis class S	2.408	8.092
	CFL1	cofilin 1	1.441	7.967
	ACER3	alkaline ceramidase 3	1.833	7.798
	PARP3	poly(ADP-ribose) polymerase family member 3	2.09	7.783
	ATXN7L1	ataxin 7 like 1	2.526	7.756
	TMA7	translation machinery associated 7 homolog	2.999	7.713
	PTMA	prothymosin, alpha	1.341	7.685
	HLA-DMB	major histocompatibility complex, class II, DM beta	2.011	7.663
	SAMD9L	sterile alpha motif domain containing 9 like	3.262	7.617
	UNC93B1	unc-93 homolog B1 (C. elegans)	1.973	7.581
Downregulated	PTH	parathyroid hormone	-7.228	14.79
	CKM	creatine kinase, M-type	-6.094	10.002
	MYL2	myosin light chain 2	-7.113	9.68
	MYH2	myosin heavy chain 2	-5.091	9.511
	ATP2A1	ATPase sarcoplasmic/endoplasmic reticulum Ca2+ transporting 1	-4.317	9.467
	AKR1C3	aldo-keto reductase family 1, member C3	-3.636	9.222
	GPT2	glutamic-pyruvic transaminase 2	-1.961	8.731
	CHGA	chromogranin A	-1.615	8.614
	MYBPC1	myosin binding protein C, slow type	-6.775	8.608
	IGBP1	immunoglobulin (CD79A) binding protein 1	-1.63	8.189
	SAMD8	sterile alpha motif domain containing 8	-1.652	8.067

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1

	SOD1	superoxide dismutase 1, soluble	-1.722	7.862
	AKR1C1	aldo-keto reductase family 1 member C1	-3.674	7.796
	FKBP3	FK506 binding protein 3	-1.872	7.688
	TMEM159	transmembrane protein 159	-1.87	7.661

Functional clustering analysis of DEGs reveals the pathological characteristics of HT

The full PPI network of up- and downregulated DEGs is presented in Fig. S1. and Fig. S2. Further analysis of functional modules was conducted using the full PPI network with regard to GO (BP) and KEGG terms (Fig. 2, Fig. 3). PPI modules were in line with the general characteristics of HT. Upregulated DEGs were grouped in three notable clusters whose gene lists over-represented the immune response (Viral infection), immune system regulation (Autoimmune disease), and RNA translation (Ribosome), with scores of 13.467, 8.214, and 6.857, respectively (Fig. 2A, 2B, 2C). Meanwhile, downregulated DEGs were grouped into three remarkable clusters enriched in Muscle contraction (Cardiomyopathy), Translation (Ribosome), and Oxidative phosphorylation (Cellular respiration and its related diseases), with scores of 18.842, 11.091, and 7.60, respectively (Fig. 3A, 3B, 3C).

(A) Cluster 1
Score : 13.467
Immune response (to virus)
Viral infection

(B) Cluster 2
Score : 8.214
Immune system regulation
Autoimmune disease

(C) Cluster 3
Score : 6.857
RNA translation, mRNA
Ribosome

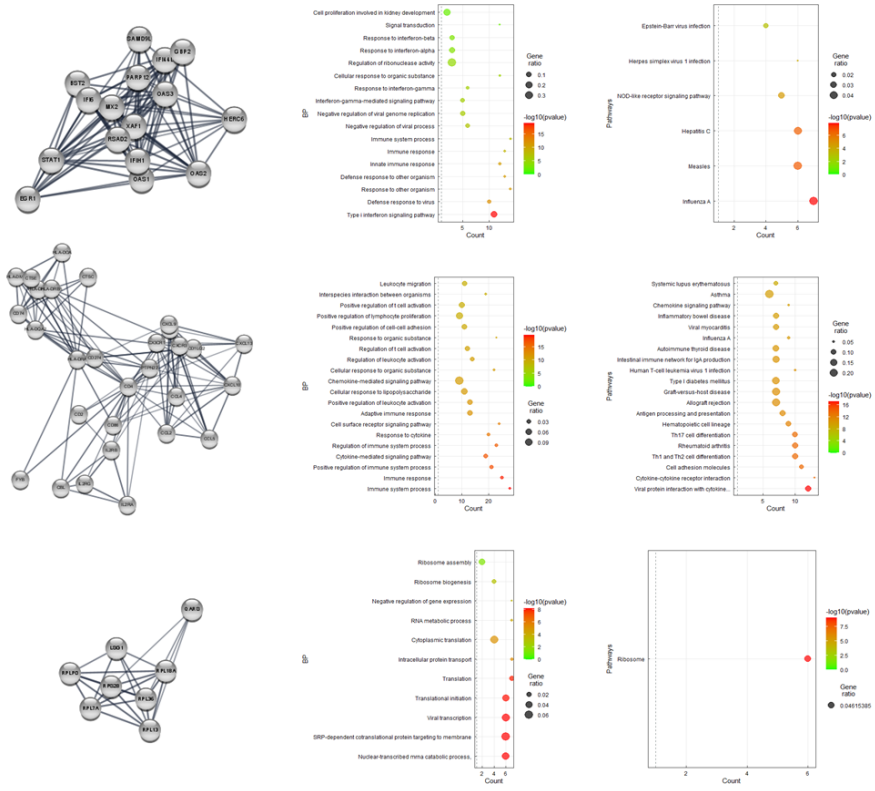


Fig. 2 Functional clustering of upregulated DEGs in HT samples using the MCODE algorithm. Results of BP and KEGG enrichment analyses presented as dot plots. Each cluster represents a set of highly connected genes. (A), (B), and (C) represent clusters 1-3, respectively, which have the highest clustering scores.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1

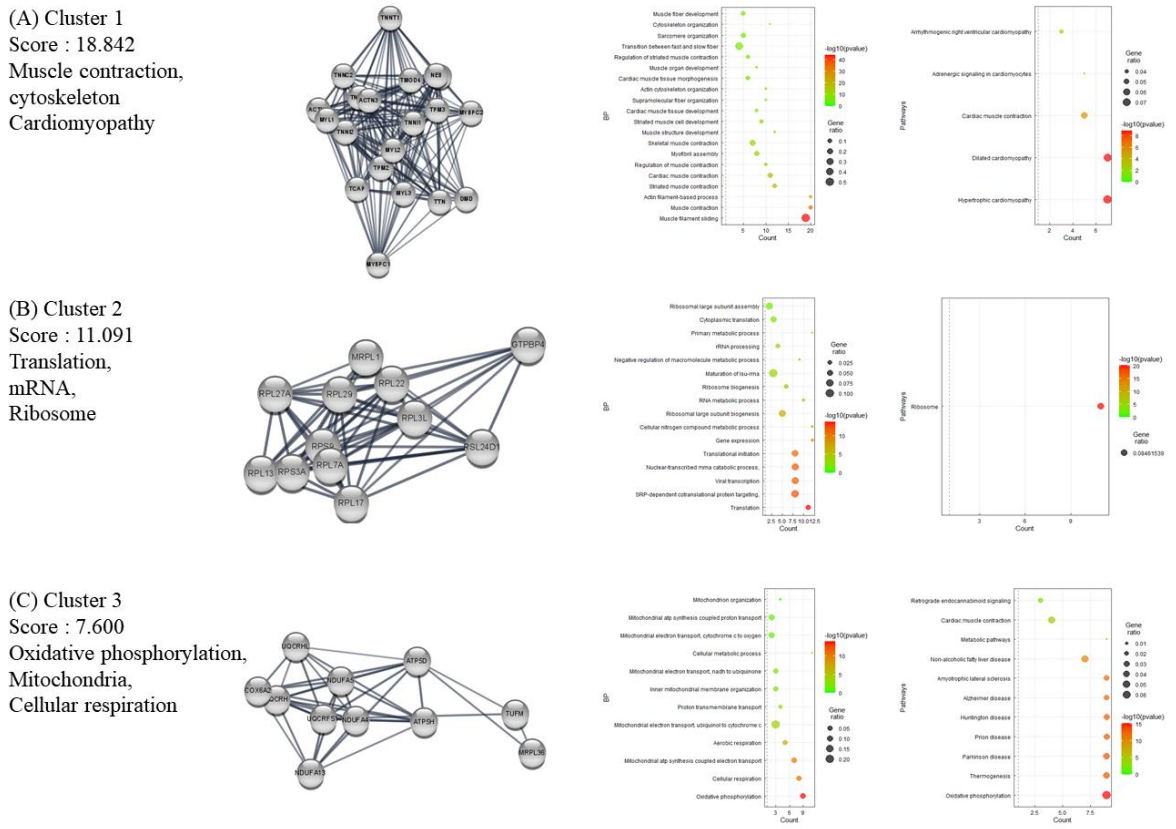


Fig. 3 Functional clustering of downregulated DEGs in HT samples using the MCODE algorithm. Results of BP and KEGG enrichment analyses presented as dot plots. Each cluster represents a set of highly connected genes. (A), (B), and (C) represents cluster 1-3, respectively, which had the highest clustering scores.

BP GSEA results revealed a strong connection between HT and immune regulation

BP GSEA results were presented in a dot plot, which clearly showed the significantly activated or inactivated biological processes in HT (Fig. 4A). As expected, immune response-related BPs were upregulated in HT patients. In contrast, muscular processes were distinctively downregulated in HT patients. Subsequently, Emapplot showed close interactions between the most significant BP terms based on overlapping genes (Fig. 4B). The BP term 'Cytokine production' had solid interactions with other terms, including 'positive regulation of immune system process', 'regulation of immune response', 'hemopoiesis', and 'cellular response to cytokine'. These five essential BP terms and enriched genes were further plotted using the Cnetplot function (Fig. 4C). Twenty-seven genes, including CASP8, IL6, TNF, IL1B, and SOCS1, were common between these five BPs, which play key roles in cytokine and inflammatory signaling (Supplementary file 1). Ridgeplot analysis illustrated the overall distribution of component genes consist of each BP (Fig. 4D). Similarly, BP terms related to various immune responses and cytokine production exhibited increased activity. The term 'Lymphocyte activation' showed predominant results in both p-value and NES, exhibiting enriched distributions of activated genes (Fig. 4E). Top BP GSEA results for the HT group are presented in Table 2. Figure S3 presents the top 5 BPs or pathways and the distributions of associated genes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1

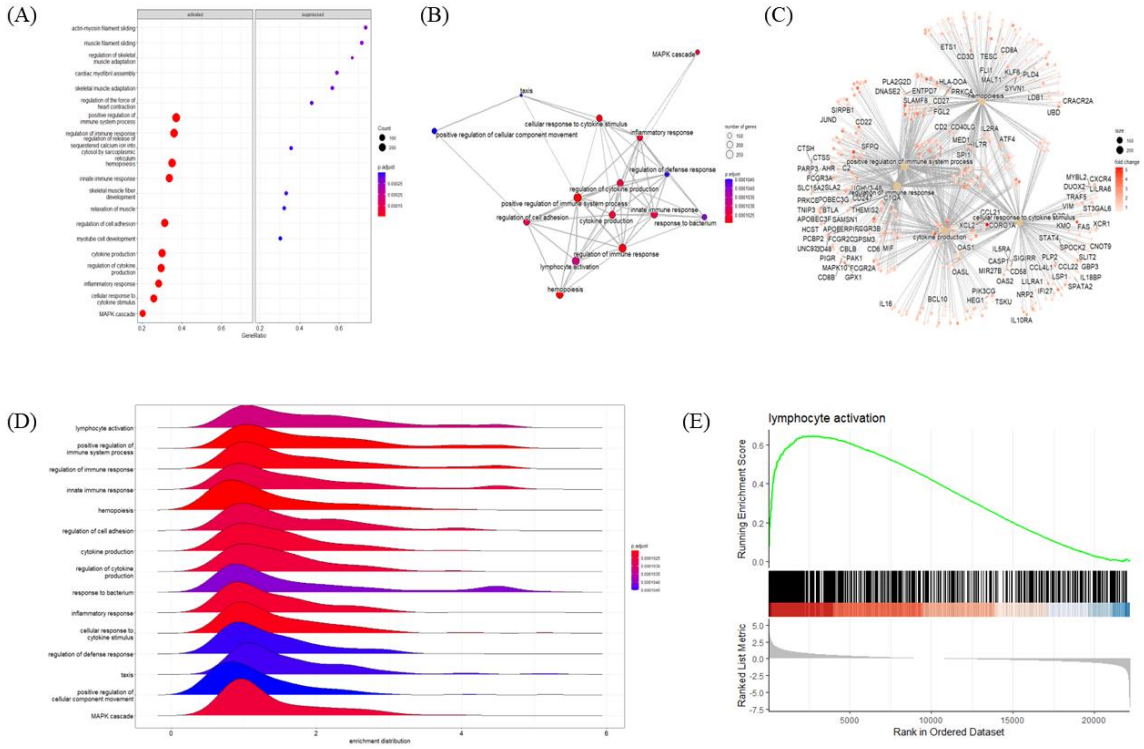


Fig. 4 Gene Set Enrichment Analysis (GSEA) plots of BP terms enriched in HT samples. (A) Bubble plot of BP terms, (B) Emapplot, (C) Cnet plot, (D) Ridge plot with ranked DEGs of the HT group from microarray data. (E) GSEA BP plot with ranked DEGs of the HT group from microarray data. All plots were created using clusterprofiler R package.

Table 2. Top 15 BP terms from GSEA of Hashimoto's thyroiditis samples, ranked by Normalized Enrichment Score (NES).

ID	Description	Enrichment Score	NES	p-value
GO:0002250	adaptive immune response	0.658	2.637	0.00011
GO:0050870	positive regulation of T cell activation	0.674	2.564	0.00012
GO:0046651	lymphocyte proliferation	0.641	2.502	0.00011
GO:0046649	lymphocyte activation	0.608	2.494	0.00010
GO:0007159	leukocyte cell-cell adhesion	0.625	2.483	0.00011
GO:0019724	B cell mediated immunity	0.684	2.475	0.00013
GO:1990868	response to chemokine	0.706	2.445	0.00013
GO:0002366	leukocyte activation involved in immune response	0.619	2.407	0.00011
GO:0002263	cell activation involved in immune response	0.615	2.389	0.00011
GO:0070661	leukocyte proliferation	0.606	2.387	0.00011
GO:0050865	regulation of cell activation	0.584	2.377	0.00011
GO:0002252	immune effector process	0.580	2.359	0.00011
GO:0002684	positive regulation of immune system process	0.570	2.347	0.00010
GO:0050776	regulation of immune response	0.559	2.304	0.00010

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1

GO:0019882	antigen processing and presentation	0.658	2.286	0.00013

KEGG GSEA results suggested a link between autoimmune thyroiditis and viral infections

The association between HT and other disease pathways was investigated via KEGG GSEA of HT transcriptional profiles from microarray data. The KEGG GSEA dot plot showed a predominance of viral infection-related pathways (including Human T-cell leukemia virus infection, Human cytomegalovirus infection, Epstein-Barr virus infection, Herpes simplex virus 1 infection, etc.) in the activated panel, with strong statistical significance (Fig. 5A). Emapplot displayed a number of interconnected viral disease pathways, while the 'Epstein-barr virus infection' pathway was a hub in the network (Fig. 5B). Five representative KEGG pathways with their connected genes were plotted, suggesting TNF and IL6 as common targets of all five pathways (Fig. 5C). Ridgeplot showed a distribution of gene expression (activated) in several viral-infectious diseases pathways such as 'Herpes simplex virus 1 infection' (Fig. 5D). Upon a closer look, the comparison of GSEA plots between 'Herpes simplex virus 1 infection' and 'Autoimmune thyroid disease' revealed that both had a skewed distribution (activated) of genes with a high normalized enrichment score (NES) (Fig. 5E). The top KEGG GSEA results of HT group are presented in Table 3.

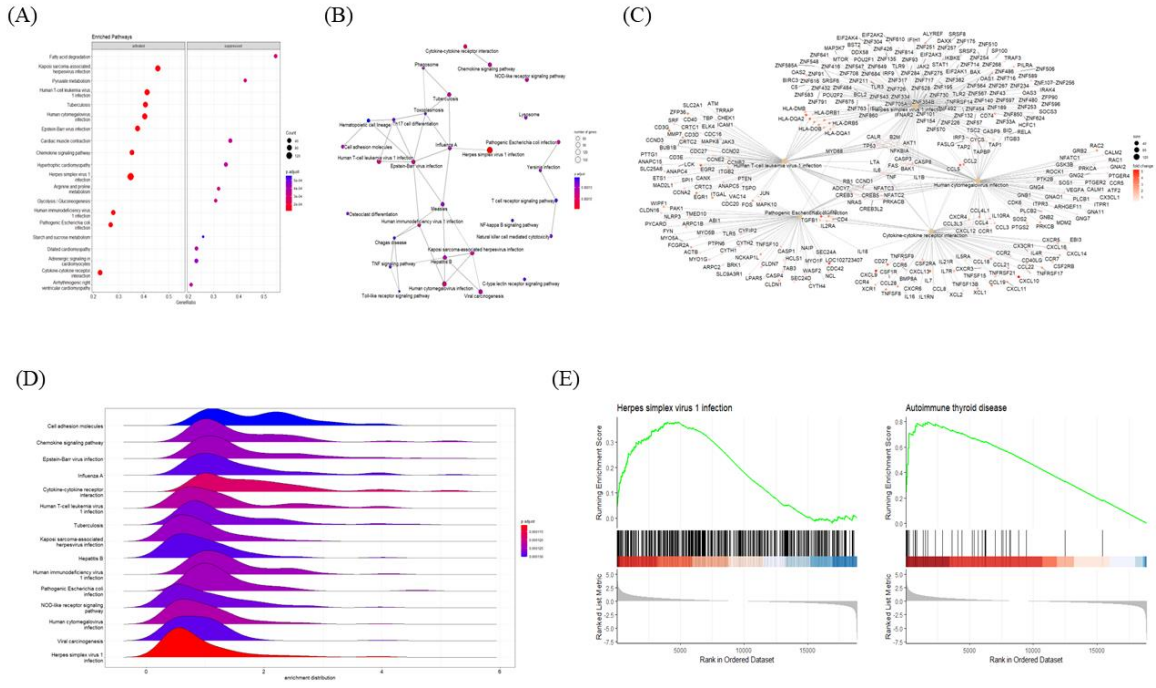


Fig. 5 Gene Set Enrichment Analysis (GSEA) plots of KEGG pathways enriched in HT samples. (A) Bubble plot of KEGG pathways, (B) Emapplot, (C) Cnet plot, and (D) Ridge plot analyzed with ranked DEGs of the HT group from microarray data. (E) GSEA KEGG plot with ranked DEGs of the HT group from microarray data. All plots were created using the clusterprofiler R package.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1

Table 3. Top 20 KEGG terms from GSEA of Hashimoto's thyroiditis samples, ranked by Normalized Enrichment Score (NES).

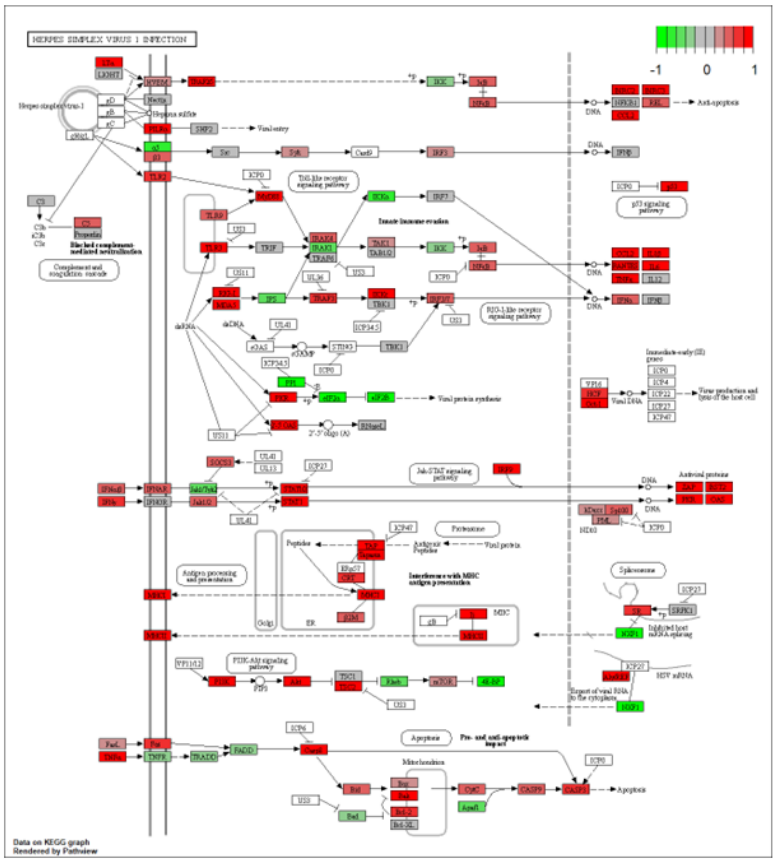
ID	Description	Enrichment Score	NES	p-value
hsa04061	Viral protein interaction with cytokine and cytokine receptor	0.753	2.683	0.00014
hsa04672	Intestinal immune network for IgA production	0.843	2.672	0.00015
hsa04640	Hematopoietic cell lineage	0.742	2.643	0.00014
hsa05322	Systemic lupus erythematosus	0.785	2.524	0.00015
hsa05320	Autoimmune thyroid disease	0.797	2.499	0.00015
hsa05340	Primary immunodeficiency	0.814	2.480	0.00016
hsa04514	Cell adhesion molecules	0.647	2.449	0.00013
hsa05330	Allograft rejection	0.821	2.427	0.00016
hsa05310	Asthma	0.837	2.426	0.00016
hsa04658	Th1 and Th2 cell differentiation	0.678	2.405	0.00014
hsa05332	Graft-versus-host disease	0.817	2.385	0.00016
hsa05140	Leishmaniasis	0.693	2.379	0.00014
hsa05150	Staphylococcus aureus infection	0.672	2.365	0.00014
hsa04662	B cell receptor signaling pathway	0.675	2.361	0.00014
hsa04062	Chemokine signaling pathway	0.607	2.360	0.00013
hsa04940	Type I diabetes mellitus	0.775	2.351	0.00016
hsa05323	Rheumatoid arthritis	0.663	2.341	0.00014
hsa04659	Th17 cell differentiation	0.637	2.313	0.00014
hsa05169	Epstein-Barr virus infection	0.587	2.291	0.00013
hsa04650	Natural killer cell-mediated cytotoxicity	0.622	2.274	0.00014

Common DEGs involved in two intuitive pathways provide insight into disease etiology

As our results repeatedly indicated a high correlation of HT with viral infectious disease, we mapped both pathways and colored DEGs (red to green) according to their relative expression levels using the pathview R package (Fig. 6A, 6B). Common genes between the two pathways were plotted using the cnetplot function with modified arguments to investigate only pathways of interest. Of note, two seemingly independent pathways shared many upregulated MHC class I and II genes (Fig. 7). These 14 HLA genes consisted of two MHC I class and 12 MHC II class molecules. Due to the unillustrated intracellular signaling pathway in the 'Autoimmune thyroid disease' KEGG map, other immune-related intersecting genes were not presented as results (Fig. 6B).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1

(A)



(B)

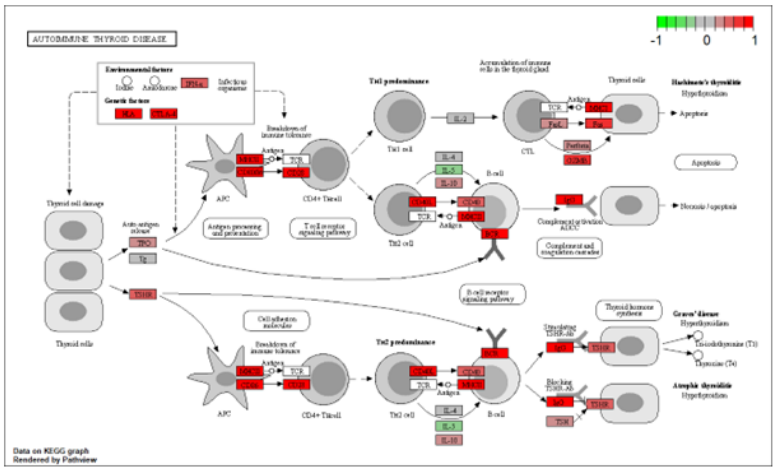
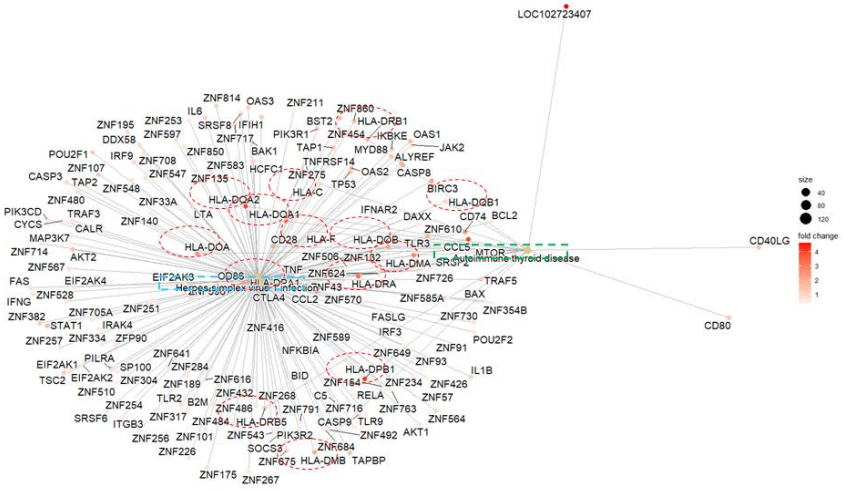


Fig. 6 KEGG pathway mapping of (A) "Herpes simplex virus 1 infection" pathway and (B) "Autoimmune thyroid disease" based on microarray data of the HT group. The plot was created using Pathview R package. The red to green color indicates the relative gene expression.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1



14 common genes	HLA-DQA2 (Class II)	HLA-DPB1 (Class II)	HLA-DRA (Class II)	HLA-DMA (Class II)	HLA-DOA (Class II)	HLA-DPA1 (Class II)	HLA-DOB (Class II)	HLA-DQA1 (Class II)	HLA-DMB (Class II)	HLA-DRB1 (Class II)	HLA-DRB5 (Class II)	HLA-F (Class I)	HLA-DQB1 (Class II)	HLA-C (Class I)
log_2FC	3.965	3.889	3.221	2.784	2.551	2.511	2.197	2.153	2.011	1.349	1.268	1.004	0.935	0.877
Adj.pval	9.01E-07	7.41E-07	5.49E-06	3.69E-07	1.15E-03	2.36E-04	1.04E-02	8.66E-03	2.17E-08	6.34E-02	6.14E-03	2.42E-02	1.88E-01	2.98E-03

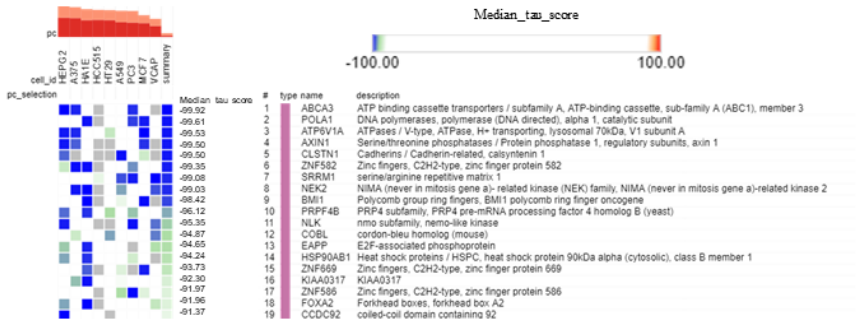
Fig. 7 Distribution of overlapping DEGs between the 'Herpes simplex virus infection' and 'Autoimmune thyroid disease' KEGG pathways of the HT group. Green and blue square boxes indicate pathway names, and the 14 red circles denote overlapping DEGs between the two pathways.

CMap analysis revealed potential markers and drugs based on the HT gene signature

The CLUE webtool deduced potential target genes and compounds for HT by comparing the gene signature against a gene expression database based on perturbations. Arranged by their median CMAP score, 19 KD gene perturbations were assumed to restore gene expression profiles of the queried signature (HT), with CMap score criteria below -90 (Fig. 8A). Six compound perturbations were expected to reverse the gene signature with the same CMap score as criteria (Fig. 8B). Among the 19 KD gene perturbagen lists, three genes (SRRM1, NLK, CCDC92) actually exhibited a significantly increased expression ($p < 0.05$) in the HT group (Fig. 8C). SRRM1 expression fold change, p-value, and transcriptional activity score (TAS) were most prominent among the three genes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 December 2022 doi:10.20944/preprints202212.0275.v1

(A)



(B)

(C)

Genes	CMap analysis		Gene expressions	
	C-score result	TAS (percentile)	Fold change (log2x)	Pvalue (-logP)
CCDC92	-91.37	0.06 (11%)	1.07	2.765
NLK	-95.35	0.12 (28%)	1.101	4.437
SRRM1	-99.08	0.49 (98%)	1.424	5.391

Fig. 8 Connectivity Map (CMap) analysis results showing (A) significant perturbation genes with the least CMap connectivity score, (B) significant perturbation compounds, and (C) three significant genes presented with their CMap score, fold-change, P-value, and TAS (transcriptional activity score) are presented.

3. Discussion

A US population-based review study reported that AID prevalence is 5-7% and is gradually increasing [10,41]. It has been suggested that genetic both susceptibility and environmental factor act as triggers for the breakdown of tolerance and progress of the disease [6]. Genetic predisposition, usually associated with HLA class II alleles and other immune mediators, is not sufficient on its own for the increasing prevalence of AIDs [42]. Rather, environmental factors, including viral infections, are regarded as major contributors to AID incidence [21].

Accumulating evidence indicates such a connection between viral infections and the development of AIDs, with an example being the increased prevalence of AIDs following an influenza pandemic reported in a population-based observational study [16,43]. A recent review on SARS-CoV-2 suggested that this virus could also trigger autoimmune responses through molecular mimicry, working in similar ways with other viruses [44]. Most prevalent autoimmune diseases/conditions involved with COVID-19 infection are Guillain-Barre syndrome, immune thrombocytopenic purpura, Kawasaki disease and autoimmune thyroid disease [45]. Investigations on the involvement of SARS-CoV-2 infection in autoimmune thyroid disease development are currently underway [46]. A recent systematic review revealed general characteristics of COVID-19-induced subacute thyroiditis patients [47]. The authors found subacute thyroiditis as the most common clinical thy-

 doi:10.20944/preprints202212.0275.v1

roidal syndrome associated COVID-19, indicating the direct effect of viral infection on autoimmune disease [47]. In other systematic review, authors pointed out that multifaceted effects of SARS-CoV-2 infection on thyroid functions are variable (thyrotoxicosis, hypothyroidism and non-thyroidal illness syndrome) and difficult to predict [48]. In this situation, it is of great importance to identify biomarkers to elucidate link between viral infections and development of HT.

Our analysis of microarray data revealed characteristic features of HT. Functional cluster analysis of upregulated DEGs provided strong evidence of an association between viral infection and the pathophysiological traits of AID through the activation of innate immune responses (Fig. 2) [49]. GSEA results indicated immune activation in the HT group, with an upregulation of genes implicated in lymphocyte activation, regulatory immune responses, cytokine production, and other related pathways (Fig. 4, 5). Decisively, the GSEA of the 'Herpes simplex virus 1 infection' and 'Autoimmune thyroid disease' pathways showed similar plots and presented high ranked significance (each 1st and 44th rank), with KEGG analysis of DEGs demonstrating an intimate relationship of the two pathways in HT (Fig. 5E, 6A, 6B).

In our data, of all HLA class I and II molecules investigated, the expression of 14 genes was significantly upregulated (Fig. 7). Aberrant HLA II expression in non-antigen presenting cells, such as epithelial cells and cultured thyrocytes, has been described in a number of AIDs [50]. Similarly, overexpression of HLA class I during the antiviral response was observed in tissue obtained via core needle biopsy from 46 HT patients [51]. Therefore, the increased expression of both HLA classes might be a general indicator of HT induced by the antiviral immune response.

Interestingly, ribosome-enriched clusters were noted among both up- or downregulated DEGs (Fig. 3). As cellular machinery that is hijacked by viruses, host ribosome factors, including ribosomal proteins (RPLs), play critical roles during viral infection [52]. While detailed functions of the various RPLs are under investigation, these are expected to impact viral replication and gene expression through interactions with viral proteins [52]. The result leads to the speculation of mixed reaction from host cellular defense mechanisms and viral activity. In line with our data, Wu et al. reported similar results from analysis of blood samples from patients with systemic lupus erythematosus, another AID [53]. When they analyzed PPI network consist of DEGs, several RPL genes were implicated as key genes in a module, and authors explained it as the outcome of host immune response to viral infection [53].

SRRM1 is an RNA-binding protein and splicing factor participating in alternative RNA splicing process [54]. It was suggested that host SRRM proteins are modulated by HIV-1 to facilitate its replication and release [55]. Further, observed changes in alternative splicing could be either a direct consequence of viral manipulation, the innate immune response, or cellular damage [56]. Dysregulation of post-transcription processes during the antiviral immune response, which is modulated by Type I and III interferons, may lead to autoimmunity [57]. In our study, SRRM1 was suggested as a critical marker of HT, as indicated by strictly negative CMap score (-99.08) and upregulated expression ($log_2FC = 1.424$) in the HT group (Fig. 8C).

While the clinical phenotype of HT differs per case and is difficult to predict, it progresses slowly over months to years [58]. It is highly suggested that patients with autoimmune thyroid disorders should be monitored for the thyroid functions in consideration of the relationship with other systemic autoimmune diseases as well [6]. The predictive Thyroid Events Amsterdam (THEA) score takes into account the TSH, TPOAb levels, and familial background information to estimate the 5-year risk of overt hypothyroidism [59]. Based on our current findings, the evaluation of viral activity, as a key environmental factor in the context of HT, may improve prediction of further development of the disease. To this end, future studies should elucidate the relationship between SRRM1 (or the other alternative splicing factors) and the widely employed clinical HT biomarkers.

5. Conclusions

In conclusion, through the use of bioinformatics approaches for the analysis of microarray data, we provide evidence for the association between viral infection and immune activation in HT. We identified SRRM1, an mRNA splicing factor, as a key player in this association. Clinical validation of our current results remains to be performed. The history of viral infection should be seriously taken into account by clinicians during HT diagnosis. The host immune response to viral infection, especially with alternative mRNA splicing may provide us helpful indicators of HT development and could be harnessed as diagnostic or therapeutic purposes. Further elucidating the relationship between viral infection and HT may improve therapeutic approaches against this disease.

Supplementary Materials: The following supporting information can be downloaded online. Figure S1: PPI network of upregulated DEGs in the HT group obtained from thyroid tissue microarray data; Figure S2: Total PPI network of downregulated DEGs in the HT group obtained from thyroid tissue microarray data; Figure S3: Venn diagram showing gene distribution for the top five significant BPs and KEGG pathways. Supplementary datafile 1: constituent genes in top 5 BPs. Supplementary datafile 2: constituent genes in top 5 KEGG pathways.

Author Contributions: D.W.L.: data acquisition, data process, data analysis, article drafting. M.S.C.: figure design, manuscript revision, data discussion. S.M.K.: study design, conception, study review, manu-script revision. All authors reviewed and agreed on the final manuscript..

Funding: No funding was received for this article.

Institutional Review Board Statement: Not applicable

Data Availability Statement: The datasets generated during and/or analysed during the current study are available in the Gene Expression Omnibus (GEO) repository, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=gse138198.

Conflicts of Interest: The authors have no conflict of interest to declare..

References

1. Caturegli, P.; De Remigis, A.; Rose, N.R. Hashimoto thyroiditis: clinical and diagnostic criteria. *Autoimmun Rev* **2014**, *13*, 391-397, doi:10.1016/j.autrev.2014.01.007.
2. Abbasgholizadeh, P.; Naseri, A.; Nasiri, E.; Sadra, V. Is Hashimoto thyroiditis associated with increasing risk of thyroid malignancies? A systematic review and meta-analysis. *Thyroid research* **2021**, *14*, 1-11.
3. Kim, H.G.; Kim, E.K.; Han, K.H.; Kim, H.; Kwak, J.Y. Pathologic spectrum of lymphocytic infiltration and recurrence of papillary thyroid carcinoma. *Yonsei Med J* **2014**, *55*, 879-885, doi:10.3349/ymj.2014.55.4.879.
4. Bogusławska, J.; Godlewska, M.; Gajda, E.; Piekiełko-Witkowska, A. Cellular and molecular basis of thyroid autoimmunity. *European Thyroid Journal* **2022**, *11*.
5. Saraf, S.R.; Gadgil, N.M.; Yadav, S.; Kalgutkar, A.D. Importance of combined approach of investigations for detection of asymptomatic Hashimoto Thyroiditis in early stage. *J Lab Physicians* **2018**, *10*, 294-298, doi:10.4103/JLP.JLP_72_17.
6. Antonelli, A.; Ferrari, S.M.; Corrado, A.; Di Domenicantonio, A.; Fallahi, P. Autoimmune thyroid disorders. *Autoimmunity reviews* **2015**, *14*, 174-180.
7. Zaletel, K.; Gaberscek, S. Hashimoto's Thyroiditis: From Genes to the Disease. *Curr Genomics* **2011**, *12*, 576-588, doi:10.2174/138920211798120763.
8. Davidson, A.; Diamond, B. Autoimmune diseases. *New England Journal of Medicine* **2001**, *345*, 340-350.
9. Goodnow, C.C. Balancing immunity and tolerance: deleting and tuning lymphocyte repertoires. *Proceedings of the National Academy of Sciences* **1996**, *93*, 2264-2271.
10. Davidson, A.; Diamond, B. General features of autoimmune disease. In *The autoimmune diseases*, Elsevier: 2020; pp. 17-44.

11. Arbuckle, M.R.; McClain, M.T.; Rubertone, M.V.; Scofield, R.H.; Dennis, G.J.; James, J.A.; Harley, J.B. Development of autoantibodies before the clinical onset of systemic lupus erythematosus. *New England Journal of Medicine* **2003**, *349*, 1526-1533.
12. Marrack, P.; Kappler, J.; Kotzin, B.L. Autoimmune disease: why and where it occurs. *Nat Med* **2001**, *7*, 899-905, doi:10.1038/90935.
13. Gough, S.C.; Simmonds, M.J. The HLA Region and Autoimmune Disease: Associations and Mechanisms of Action. *Curr Genomics* **2007**, *8*, 453-465, doi:10.2174/138920207783591690.
14. Shin, D.-H.; Baek, I.-C.; Kim, H.J.; Choi, E.-J.; Ahn, M.; Jung, M.H.; Suh, B.-K.; Cho, W.K.; Kim, T.-G. HLA alleles, especially amino-acid signatures of HLA-DPB1, might contribute to the molecular pathogenesis of early-onset autoimmune thyroid disease. *PloS one* **2019**, *14*, e0216941.
15. Hewagama, A.; Richardson, B. The genetics and epigenetics of autoimmune diseases. *Journal of autoimmunity* **2009**, *33*, 3-11.
16. Smatti, M.K.; Cyprian, F.S.; Nasrallah, G.K.; Al Thani, A.A.; Almishal, R.O.; Yassine, H.M. Viruses and Autoimmunity: A Review on the Potential Interaction and Molecular Mechanisms. *Viruses* **2019**, *11*, doi:10.3390/v11080762.
17. Hussein, H.M.; Rahal, E.A. The role of viral infections in the development of autoimmune diseases. *Crit Rev Microbiol* **2019**, *45*, 394-412, doi:10.1080/1040841X.2019.1614904.
18. Cusick, M.F.; Libbey, J.E.; Fujinami, R.S. Molecular mimicry as a mechanism of autoimmune disease. *Clinical reviews in allergy & immunology* **2012**, *42*, 102-111.
19. Piccinini, A.; Midwood, K. DAMPening inflammation by modulating TLR signalling. *Mediators of inflammation* **2010**, *2010*.
20. Di Crescenzo, V.; D'Antonio, A.; Tonacchera, M.; Carlomagno, C.; Vitale, M. Human herpes virus associated with Hashimoto's thyroiditis. *Le Infezioni in Medicina* **2013**, *21*, 224-228.
21. Janegova, A.; Janega, P.; Rychly, B.; Kuracinova, K.; Babal, P. The role of Epstein-Barr virus infection in the development of autoimmune thyroid diseases. *Endokrynol Pol* **2015**, *66*, 132-136, doi:10.5603/EP.2015.0020.
22. Antonelli, A.; Ferri, C.; Pampana, A.; Fallahi, P.; Nesti, C.; Pasquini, M.; Marchi, S.; Ferrannini, E. Thyroid disorders in chronic hepatitis C. *The American journal of medicine* **2004**, *117*, 10-13.
23. Shen, Y.; Wang, X.L.; Xie, J.P.; Shao, J.G.; Lu, Y.H.; Zhang, S.; Qin, G. Thyroid Disturbance in Patients with Chronic Hepatitis C Infection: A Systematic Review and Meta-analysis. *J Gastrointestin Liver Dis* **2016**, *25*, 227-234, doi:10.15403/jgld.2014.1121.252.chc.
24. Edgar, R.; Domrachev, M.; Lash, A.E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic acids research* **2002**, *30*, 207-210.
25. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **2005**, *102*, 15545-15550.
26. Lamb, J.; Crawford, E.D.; Peck, D.; Modell, J.W.; Blat, I.C.; Wrobel, M.J.; Lerner, J.; Brunet, J.-P.; Subramanian, A.; Ross, K.N. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *science* **2006**, *313*, 1929-1935.
27. Trevino, V.; Falciani, F.; Barrera-Saldana, H.A. DNA microarrays: a powerful genomic tool for biomedical and clinical research. *Mol Med* **2007**, *13*, 527-541, doi:10.2119/2006-00107.Trevino.
28. Qiu, K.; Li, K.; Zeng, T.; Liao, Y.; Min, J.; Zhang, N.; Peng, M.; Kong, W.; Chen, L.L. Integrative Analyses of Genes Associated with Hashimoto's Thyroiditis. *J Immunol Res* **2021**, *2021*, 8263829, doi:10.1155/2021/8263829.

29. Subhi, O.; Schulten, H.-J.; Bagatian, N.; Al-Dayini, R.a.; Karim, S.; Bakhashab, S.; Alotibi, R.; Al-Ahmadi, A.; Ata, M.; Elaimi, A. Genetic relationship between Hashimotos thyroiditis and papillary thyroid carcinoma with coexisting Hashimotos thyroiditis. *Plos one* **2020**, *15*, e0234566.
30. Smyth, G.K. Limma: linear models for microarray data. In *Bioinformatics and computational biology solutions using R and Bioconductor*, Springer: 2005; pp. 397-420.
31. Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K.H.; Sherman, P.M.; Holko, M. NCBI GEO: archive for functional genomics data sets—update. *Nucleic acids research* **2012**, *41*, D991-D995.
32. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research* **2019**, *47*, D607-D613.
33. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **2003**, *13*, 2498-2504.
34. Bader, G.D.; Hogue, C.W. An automated method for finding molecular complexes in large protein interaction networks. *BMC bioinformatics* **2003**, *4*, 1-27.
35. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* **2009**, *4*, 44-57.
36. Bonnot, T.; Gillard, M.B.; Nagel, D.H. A simple protocol for informative visualization of enriched gene ontology terms. *Bio-protocol* **2019**, e3429-e3429.
37. Yu, G.; Wang, L.-G.; Han, Y.; He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology* **2012**, *16*, 284-287.
38. Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L., et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2021**, *2*, 100141, doi:10.1016/j.xinn.2021.100141.
39. Yu, G. Enrichplot: visualization of functional enrichment result. *R package version* **2019**, *1*.
40. Musa, A.; Ghoraie, L.S.; Zhang, S.D.; Glazko, G.; Yli-Harja, O.; Dehmer, M.; Haibe-Kains, B.; Emmert-Streib, F. A review of connectivity map and computational approaches in pharmacogenomics. *Brief Bioinform* **2018**, *19*, 506-523, doi:10.1093/bib/bbw112.
41. Jacobson, D.L.; Gange, S.J.; Rose, N.R.; Graham, N.M. Epidemiology and estimated population burden of selected autoimmune diseases in the United States. *Clinical immunology and immunopathology* **1997**, *84*, 223-243.
42. Munz, C.; Lunemann, J.D.; Getts, M.T.; Miller, S.D. Antiviral immune responses: triggers of or triggered by autoimmunity? *Nat Rev Immunol* **2009**, *9*, 246-258, doi:10.1038/nri2527.
43. Han, F.; Lin, L.; Warby, S.C.; Faraco, J.; Li, J.; Dong, S.X.; An, P.; Zhao, L.; Wang, L.H.; Li, Q.Y. Narcolepsy onset is seasonal and increased following the 2009 H1N1 pandemic in China. *Annals of neurology* **2011**, *70*, 410-417.
44. Liu, Y.; Sawalha, A.H.; Lu, Q. COVID-19 and autoimmune diseases. *Curr Opin Rheumatol* **2021**, *33*, 155-162, doi:10.1097/BOR.0000000000000776.
45. Yazdanpanah, N.; Rezaei, N. Autoimmune complications of COVID-19. *J Med Virol* **2022**, *94*, 54-62, doi:10.1002/jmv.27292.
46. Duntas, L.H.; Jonklaas, J. COVID-19 and Thyroid Diseases: A Bidirectional Impact. *J Endocr Soc* **2021**, *5*, bvab076, doi:10.1210/jendso/bvab076.

47. Christensen, J.; O'Callaghan, K.; Sinclair, H.; Hawke, K.; Love, A.; Hajkowicz, K.; Stewart, A.G. Risk factors, treatment and outcomes of subacute thyroiditis secondary to COVID-19: a systematic review. *Intern Med J* **2022**, *52*, 522-529, doi:10.1111/imj.15432.
48. Aemaz Ur Rehman, M.; Farooq, H.; Ali, M.M.; Ebaad Ur Rehman, M.; Dar, Q.A.; Hussain, A. The association of subacute thyroiditis with COVID-19: a systematic review. *SN Comprehensive Clinical Medicine* **2021**, *3*, 1515-1527.
49. Niewold, T.B. Interferon alpha as a primary pathogenic factor in human lupus. *Journal of interferon & cytokine research* **2011**, *31*, 887-892.
50. Buscema, M.; Todd, I.; Deuss, U.; Hammond, L.; Mirakian, R.; Pujol-Borrell, R.; Bottazzo, G.F. Influence of tumor necrosis factor-α on the modulation by interferon-γ of HLA class II molecules in human thyroid cells and its effect on interferon-γ binding. *The Journal of Clinical Endocrinology & Metabolism* **1989**, *69*, 433-439.
51. Weider, T.; Richardson, S.J.; Morgan, N.G.; Paulsen, T.H.; Dahl-Jørgensen, K.; Hammerstad, S.S. Upregulation of HLA class I and antiviral tissue responses in Hashimoto's thyroiditis. *Thyroid* **2020**, *30*, 432-442.
52. Li, S. Regulation of Ribosomal Proteins on Viral Infection. *Cells* **2019**, *8*, doi:10.3390/cells8050508.
53. Wu, C.; Zhao, Y.; Lin, Y.; Yang, X.; Yan, M.; Min, Y.; Pan, Z.; Xia, S.; Shao, Q. Bioinformatics analysis of differentially expressed gene profiles associated with systemic lupus erythematosus. *Molecular medicine reports* **2018**, *17*, 3591-3598.
54. Jiménez-Vacas, J.M.; Herrero-Aguayo, V.; Montero-Hidalgo, A.J.; Gómez-Gómez, E.; Fuentes-Fayos, A.C.; León-González, A.J.; Sáez-Martínez, P.; Alors-Pérez, E.; Pedraza-Arévalo, S.; González-Serrano, T. Dysregulation of the splicing machinery is directly associated to aggressiveness of prostate cancer. *EBioMedicine* **2020**, *51*, 102547.
55. Wojcechowskyj, J.A.; Didigu, C.A.; Lee, J.Y.; Parrish, N.F.; Sinha, R.; Hahn, B.H.; Bushman, F.D.; Jensen, S.T.; Seeholzer, S.H.; Doms, R.W. Quantitative phosphoproteomics reveals extensive cellular reprogramming during HIV-1 entry. *Cell Host Microbe* **2013**, *13*, 613-623, doi:10.1016/j.chom.2013.04.011.
56. Ashraf, U.; Benoit-Pilven, C.; Lacroix, V.; Navratil, V.; Naffakh, N. Advances in Analyzing Virus-Induced Alterations of Host Cell Splicing. *Trends Microbiol* **2019**, *27*, 268-281, doi:10.1016/j.tim.2018.11.004.
57. Liao, K.C.; Garcia-Blanco, M.A. Role of Alternative Splicing in Regulating Host Response to Viral Infection. *Cells* **2021**, *10*, doi:10.3390/cells10071720.
58. Akamizu, T.; Amino, N. Hashimoto's thyroiditis. *Endotext [Internet]* **2017**.
59. Strieder, T.G.; Tijssen, J.G.; Wenzel, B.E.; Endert, E.; Wiersinga, W.M. Prediction of progression to overt hypothyroidism or hyperthyroidism in female relatives of patients with autoimmune thyroid disease using the Thyroid Events Amsterdam (THEA) score. *Arch Intern Med* **2008**, *168*, 1657-1663, doi:10.1001/archinte.168.15.1657.