



# Identification of Key Genes and Pathways in Peripheral Blood Mononuclear Cells of Type 1 Diabetes Mellitus by Integrated Bioinformatics Analysis

Xing Li<sup>1,2,3</sup>, Mingyu Liao<sup>3</sup>, Jiangheng Guan<sup>4</sup>, Ling Zhou<sup>3</sup>, Rufe Shen<sup>3</sup>, Min Long<sup>3</sup>, Jiaqing Shao<sup>1,2</sup>

<sup>1</sup>Department of Endocrinology, Jinling Hospital, Medical School of Nanjing University, Nanjing,

<sup>2</sup>Department of Endocrinology, Jinling Hospital, Nanjing Medical University, Nanjing,

<sup>3</sup>Department of Endocrinology, Translational Research Key Laboratory for Diabetes, Xinqiao Hospital, Army Medical University, Chongqing,

<sup>4</sup>Department of Neurosurgery, The General Hospital of Chinese PLA Central Theater Command, Wuhan, China

**Background:** The onset and progression of type 1 diabetes mellitus (T1DM) is closely related to autoimmunity. Effective monitoring of the immune system and developing targeted therapies are frontier fields in T1DM treatment. Currently, the most available tissue that reflects the immune system is peripheral blood mononuclear cells (PBMCs). Thus, the aim of this study was to identify key PBMC biomarkers of T1DM.

**Methods:** Common differentially expressed genes (DEGs) were screened from the Gene Expression Omnibus (GEO) datasets GSE9006, GSE72377, and GSE55098, and PBMC mRNA expression in T1DM patients was compared with that in healthy participants by GEO2R. Gene Ontology, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and protein-protein interaction (PPI) network analyses of DEGs were performed using the Cytoscape, DAVID, and STRING databases. The vital hub genes were validated by reverse transcription-polymerase chain reaction using clinical samples. The disease-gene-drug interaction network was built using the Comparative Toxicogenomics Database (CTD) and Drug Gene Interaction Database (DGIdb).

**Results:** We found that various biological functions or pathways related to the immune system and glucose metabolism changed in PBMCs from T1DM patients. In the PPI network, the DEGs of module 1 were significantly enriched in processes including inflammatory and immune responses and in pathways of proteoglycans in cancer. Moreover, we focused on four vital hub genes, namely, chitinase-3-like protein 1 (CHI3L1), C-X-C motif chemokine ligand 1 (CXCL1), matrix metalloproteinase 9 (MMP9), and granzyme B (GZMB), and confirmed them in clinical PBMC samples. Furthermore, the disease-gene-drug interaction network revealed the potential of key genes as reference markers in T1DM.

**Conclusion:** These results provide new insight into T1DM pathogenesis and novel biomarkers that could be widely representative reference indicators or potential therapeutic targets for clinical applications.

**Keywords:** Computational biology; Diabetes mellitus, type 1; Immune system; Leukocytes, mononuclear; Protein interaction maps

## INTRODUCTION

Type 1 diabetes mellitus (T1DM) is a chronic autoimmune

disease that leads to autoimmune destruction of insulin-secreting pancreatic  $\beta$ -cells and ultimately hyperglycemia, which seriously affects human health and living standards [1]. More

Correspond authors: Jiaqing Shao <https://orcid.org/0000-0002-9739-5410>  
Department of Endocrinology, Jinling Hospital, Medical School of Nanjing University,  
Nanjing Medical University, 305 Zhongshan E Rd, Nanjing 210016, China  
E-mail: shaojqiaqing@nju.edu.cn

Min Long <https://orcid.org/0000-0003-1071-8131>  
Department of Endocrinology, Translational Research Key Laboratory for Diabetes,  
Xinqiao Hospital, Army Medical University (Third Military Medical University),  
Chongqing 400037, China  
E-mail: longmin\_casper@163.com

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

than 13 million people suffer from T1DM [2], and the incidence increases by 3% to 4% per year globally [3-5]. T1DM was originally observed to be accompanied by pancreas inflammation and islet cell autoantibody elevation in patients [6]. This autoimmune reaction is triggered by both genetic susceptibility and environmental factors such as viral infection [6]. The key pathogenic process is antigen-driven T lymphocyte-mediated islet inflammation, which leads to the specific destruction of islet cells [7]. Importantly, autoimmune reactions often precede the onset of clinical diabetes and may occur many years in advance. However, changes in innate immune effector cells have been observed at the same or even earlier stage of the production of diabetes autoantibodies [8,9].

Peripheral blood mononuclear cells (PBMCs), also known as human mononuclear cells from peripheral blood, reflect immune responses and act as an important component of immune monitoring [10]. Due to their accessibility and effectiveness for batch testing, PBMCs are utilized as a useful tool for studying various aspects of pathology and biology in research and clinical applications. However, the initiating factors of T1DM and how innate immunity participates are still worth investigating. Early and efficient prevention and diagnosis are critical for higher-risk individuals with T1DM. Gene chips have been widely applied as a gene detection technology, and the corresponding data have been deposited in public databases. Integrating and reanalyzing these genomic data offer possibilities for identifying certain disease-related biomarkers. Recently, many studies have been carried out according to microarray data profiles to elucidate the role of PBMCs in T1DM [10]. However, the differentially expressed genes (DEGs) identified with microarrays are based on a single cohort in each study and have a limited sample scale and unknown confounding factors. To identify novel key DEGs that remained representative and accurate, we used integrated bioinformatics methods overlapping genes obtained from different microarrays.

In the current study, we downloaded three microarray datasets, namely, GSE9006 [10], GSE72377, and GSE55098 [11], from the NCBI Gene Expression Omnibus (GEO) database, including gene expression data for PBMC samples from 70 patients with T1DM and 54 healthy controls. We identified DEGs using the interactive web tool GEO2R. The functional and pathway analyses were performed via Gene Ontology (GO) and Database for Annotation, Visualization and Integrated Discovery (DAVID). Subsequently, we integrated the DEG protein-protein interaction (PPI) network with module screen-

ing to identify hub genes in PBMCs of T1DM. Identifying novel key DEGs and their enriched functions and signaling pathways might help provide a convenient and effective approach for clinical prevention, treatment and management.

## METHODS

### Microarray data information

NCBI-GEO is regarded as a free public database of microarray/gene profile and we obtained the gene expression profile of GSE9006, GSE72377, and GSE55098 analyzed by PBMCs from totally 70 T1DM patients and 54 healthy controls. GSE9006 was on account of GPL96 and GPL97 Platforms ([HG-U133A] Affymetrix Human Genome U133A Array and [HG-U133B] Affymetrix Human Genome U133B Array), including 43 samples of PBMCs from T1DM patients and 24 healthy controls. The 43 newly diagnosed T1DM patients (26 females and 17 males; age  $10.1 \pm 3.8$  years; race 28 Caucasian, seven Hispanic, three African-American, and four mixed ethnicities; glycosylated hemoglobin [HbA1c]  $11.8\% \pm 2.0\%$ ) was age-matched to healthy group, and diagnosed by American Diabetes Association (ADA) and World Health Organization criteria [12] with a include criteria that positive autoantibodies to insulin, IA-2, and GAD65 [10]. GSE72377 was on account of GPL10558 Platforms (Illumina HumanHT-12 V4.0 expression beadchip, Illumina, San Diego, CA, USA), including 15 samples of PBMCs from T1DM patients and 20 healthy controls. GSE55098 was on account of GPL570 Platforms ([HG-U133\_Plus\_2] Affymetrix Human Genome U133 Plus 2.0 Array) including 12 samples of PBMCs from T1DM patients and 10 healthy controls. The diagnosis of 12 newly diagnosed T1DM patients (five women and seven men; age  $17.50 \pm 3.68$  years; fasting blood glucose  $6.37 \pm 1.93$  mmol/L; HbA1c  $11.78\% \pm 3.63\%$ ; lower fasting C-peptide  $0.47 \pm 0.20$  ng/mL; markedly higher glutamic acid decarboxylase antibody [GADA] level 149.85 U/mL) [11] was made according to 2011 ADA criteria [13], and their age-matched healthy participants showed clinical parameters within normal range.

### Microarray data processing and DEGs identification

Data preprocessing included conversion from probes into gene symbols, data consolidation, and normalization. Probes without gene symbols or genes with more than one probe were removed or averaged, respectively. Volcano plot of each dataset was generated by ggpubr tool in R studio (R Foundation for

Statistical Computing, Vienna, Austria). DEGs between T1DM and healthy PBMC samples were identified via GEO2R online tools with  $|FC| > 1.1$  and  $P < 0.05$ . Then, the DEGs were checked in Venn software online to detect the commonly DEGs among three datasets. The DEGs with  $\log FC < 0$  was considered as down-regulated genes, while the DEGs with  $\log FC > 0$  was considered as an up-regulated gene. The heat map of top 10 DEGs were generated by Graphpad prism 8 (Graphpad, San Diego, CA, USA).

### Functional and pathway enrichment analysis

To understand the biological relevance of DEGs and the hub genes, we performed functional enrichment analysis using ClueGO. ClueGO facilitates the visualization of functionally related genes displayed as a clustered network and chart [14,15]. The statistical test used for the enrichment was based on two-sided hypergeometric test with a Bonferroni step down correction method and kappa score threshold of 0.4 ( $P < 0.05$ ). DAVID (<https://david-d.ncicrf.gov/>) was designed to identify a large number of genes or proteins function, which was used for Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, and GO biological functional enrichment analysis.

### PPI network and modular analysis

Interactions among the DEGs were explored with the Search Tool for the Retrieval of Interacting Genes/Proteins database (STRING, <https://string-db.org/>) [16]. Based on the STRING database, Cytoscape was applied to examine the potential correlation between these DEGs (confidence score  $\geq 0.4$ ) [17]. The number of nodes is 83, the number of edges is 105. The average node degree is 2.53, the average local clustering coefficient is 0.36, the expected number of edges is 54, and the PPI enrichment  $P$  value is  $4.79e-10$ . For checking modules of the PPI network, the MCODE app in Cytoscape was used (degree cutoff=2, node score cutoff=0.2; K-core=2, maximum depth from seed=100).

### Key gene identification

To find key hub genes in the above DEGs based on the PPI network, we used the plug—in cytoHubba of Cytoscape. The plug—in can predict and explore important nodes and subnetworks in a given network using several topological algorithms [18]. Maximal clique centrality (MCC) method was selected to explore hub genes. The heatmap of ranked 10 hub genes was

generated by Graphpad Prism 8. The network composed of hub genes and related biological functions was constructed by GlueGO.

### Clinical PBMCs samples collection

Ten patients new-diagnosed as T1DM and age-matched 10 healthy individual between December 2018 and November 2020 at Xinqiao Hospital (Second affiliated hospital of Army University/The Third Medical Military University) were enrolled and this study. The key clinical characteristics were shown in Supplementary Table 1. The study was approved by the ethical committee of Xinqiao Hospital (No.2016024), and written informed consent was obtained from all participated subjects. The diagnosis of T1DM was according to ADA criteria [13] briefly as follows: an onset of hyperglycemia symptoms, the absence of C-peptide and positivity for autoantibodies related to islet  $\beta$ -cell destruction. Patients with other acute or chronic diseases such as active or presumed infection, other autoimmune disease, were taking immune modulators, with organ dysfunction or pregnancy, were excluded. The blood collection was performed at 9:00 to 10:00 AM with the butterfly vein cannula drawn with a 10 mL syringe pre-filled with ethylenediaminetetraacetic acid (0.16%) at room temperature. PBMCs were immediately extracted from fresh blood within 2 hours after blood collection, using human PBMCs extracting kit (Solabio, Catalog: P8680, Beijing, China), according to manufacturer's protocol, then stored at  $-80^{\circ}\text{C}$  for RNA extraction.

### RNA extraction and hub gene validation

Total RNA was extracted from PBMCs, and 1  $\mu\text{g}$  RNA was reverse-transcribed into cDNA as described above. The reverse transcription-polymerase chain reaction (RT-PCR) was performed with SYBR Premix Ex Taq II (Takara, Terra Bella Ave, CA, USA) on a Bio-Applied Biosystems 7300 (Thermo Fisher Scientific, Waltham, MA, USA) using the following thermal cycling procedure:  $95^{\circ}\text{C}$  for 30 seconds, followed by 40 cycles of  $95^{\circ}\text{C}$  for 15 seconds,  $58^{\circ}\text{C}$  for 30 seconds, and  $72^{\circ}\text{C}$  for 20 seconds. Expression of the hub genes was normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) expression. The primer sequences for C-X-C motif chemokine ligand 1 (CXCL1), granzyme B (GZMB), matrix metalloproteinase-9 (MMP9), chitinase-3-like protein 1 (CHI3L1), and GAPDH were listed below: CXCL1 forward: AGCTTGCCCTCAATCCTGCATCC, reverse: TCCTTCAGGAACAGCCACCAGT;

GZMB forward: CGACAGTACCATTGAGTTGTGCG, reverse: TTCGTCCATAGGAGACAATGCCC; CHI3L1 forward: CCACAGTCCATAGAATCCTCGG, reverse: TGCCTGTCCTTCAGGTACTGCA; MMP9 forward: GCCACTACTGTGCCTTTGAGTC, reverse: CCCTCAGAGAATCGCCAGTACT.

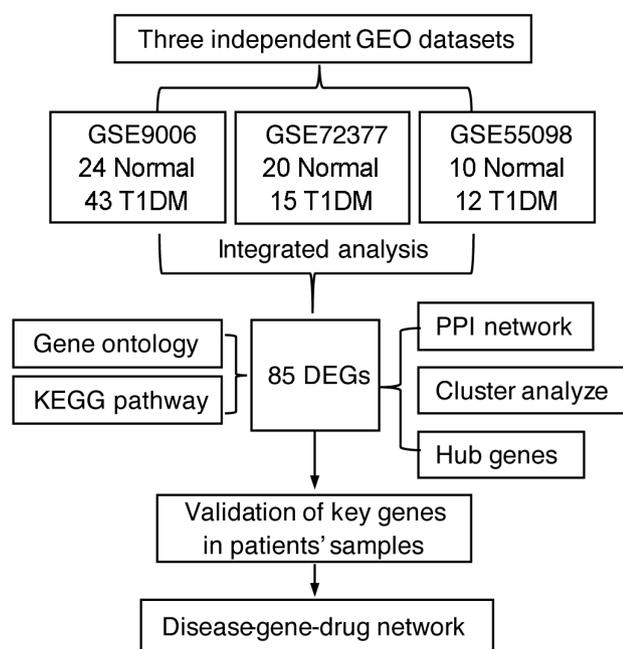
### Disease-gene-drug network-based analysis

The correlation among diseases, genes and drugs were analyzed via Comparative Toxicogenomics Database (CTD) database (<http://ctdbase.org/tools/batchQuery.go>) and the Drug Gene Interaction Database (DGIdb; <https://dgidb.org>). Endocrinology and immune-related diseases were selected. The disease-gene-drug interaction network was visualized using Cytoscape.

## RESULTS

### Identification of DEGs in each GEO dataset

In this study, we performed a multistep analysis to explore key DEGs and their significant biological functions by integrated bioinformatics methods (Fig. 1). First, we selected and down-

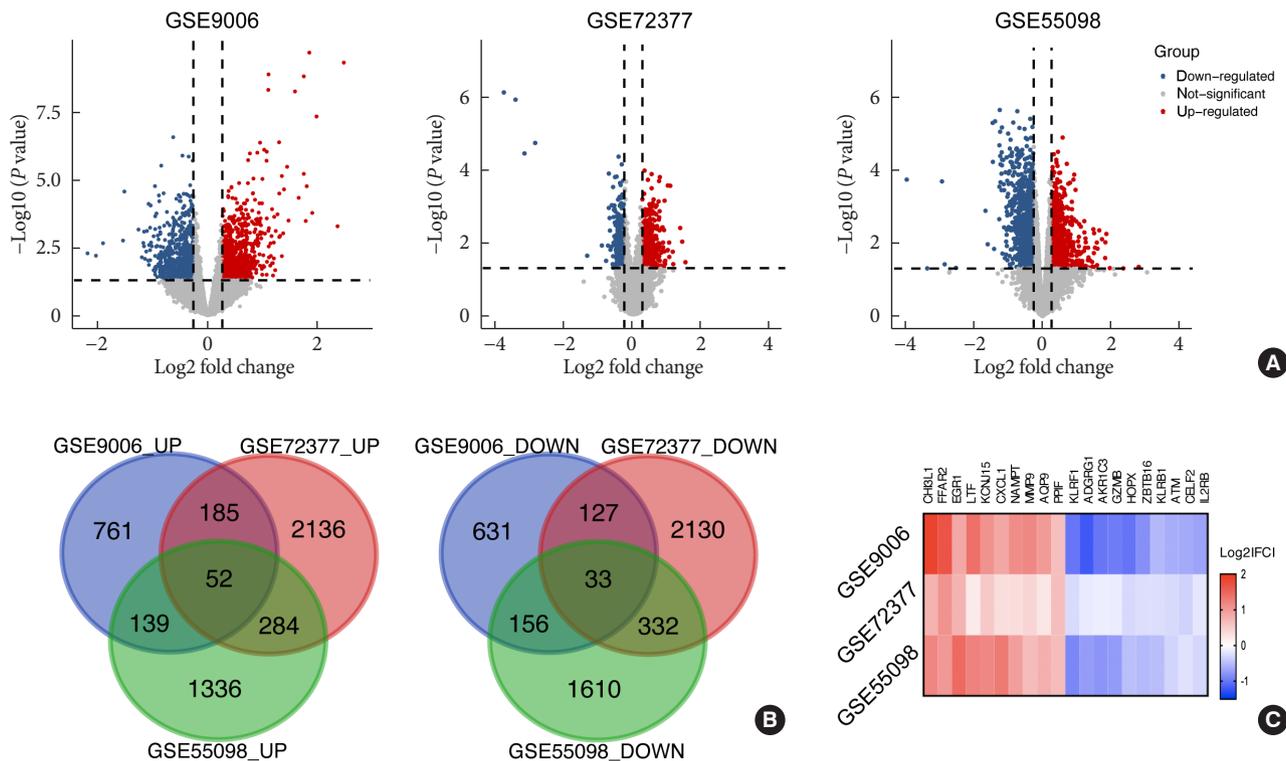


**Fig. 1.** Flow chart of bioinformatics and validation. GEO, Gene Expression Omnibus; T1DM, type 1 diabetes mellitus; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEG, differentially expressed genes; PPI, protein-protein interaction.

loaded three GEO datasets (GSE9006, GSE72377, and GSE55098) with gene expression profiles for PBMCs of T1DM patients and healthy controls. There were PBMC samples from 70 T1DM patients and 54 normal participants in our present study. On the basis of the cutoff criteria, DEGs in each GEO dataset were identified between PBMCs of the T1DM and normal groups (Supplementary Table 2). There were 2,084 DEGs, including 1,137 upregulated and 947 downregulated genes in GSE9006; 5,279 DEGs, including 2,657 upregulated and 2,622 downregulated genes in GSE72377; and 3,942 DEGs, including 1,811 upregulated and 2,131 downregulated genes in GSE55098. The volcano plots of the distribution of DEGs in each dataset are shown in Fig. 2A. Then, we used Venn diagram software to identify the common DEGs among datasets. The results showed that there were 85 common DEGs among the three datasets, including 52 upregulated genes and 33 downregulated genes (Fig. 2B), which are listed in Supplementary Table 3. Among them, the top 10 upregulated and downregulated DEGs were selected and are listed as a heatmap: the top 10 upregulated DEGs included CHI3L1, free fatty acid receptor 2 (FFAR2), early growth response 1 (EGR1), lactotransferrin (LTF), potassium inwardly rectifying channel subfamily J member 15 (KCNJ15), CXCL1, nicotinamide phosphoribosyltransferase (NAMPT), MMP9, aquaporin 9 (AQP9), and peptidylprolyl isomerase (PPIF), and the top 10 downregulated DEGs included killer cell lectin like receptor F1 (KLRF1), adhesion G protein-coupled receptor G1 (ADGRG1), aldo-keto reductase family 1 member C3 (AKR1C3), GZMB, HOP homeobox (HOPX), zinc finger and BTB domain containing 16 (ZBTB16), killer cell lectin like receptor B1 (KLRB1), ATM serine/threonine kinase (ATM), CUGBP elav-like family member 2 (CELF2), and interleukin 2 receptor subunit beta (IL2RB) (Fig. 2C).

### Biological functional enrichment analysis

To characterize the functional roles of the above DEGs, biological process enrichment analyses in upregulated or downregulated DEGs were conducted via ClueGO software. The results showed that upregulated DEGs were significantly enriched in response to amyloid-beta (A $\beta$ ) and cytokine receptor activity (Fig. 3A). The response to A $\beta$  involves various immune processes, including positive regulation of cytokine production involved in the inflammatory response, positive regulation of the reactive oxygen species biosynthetic process, regulation of the cytokine biosynthetic process, and interleukin-1 and inter-



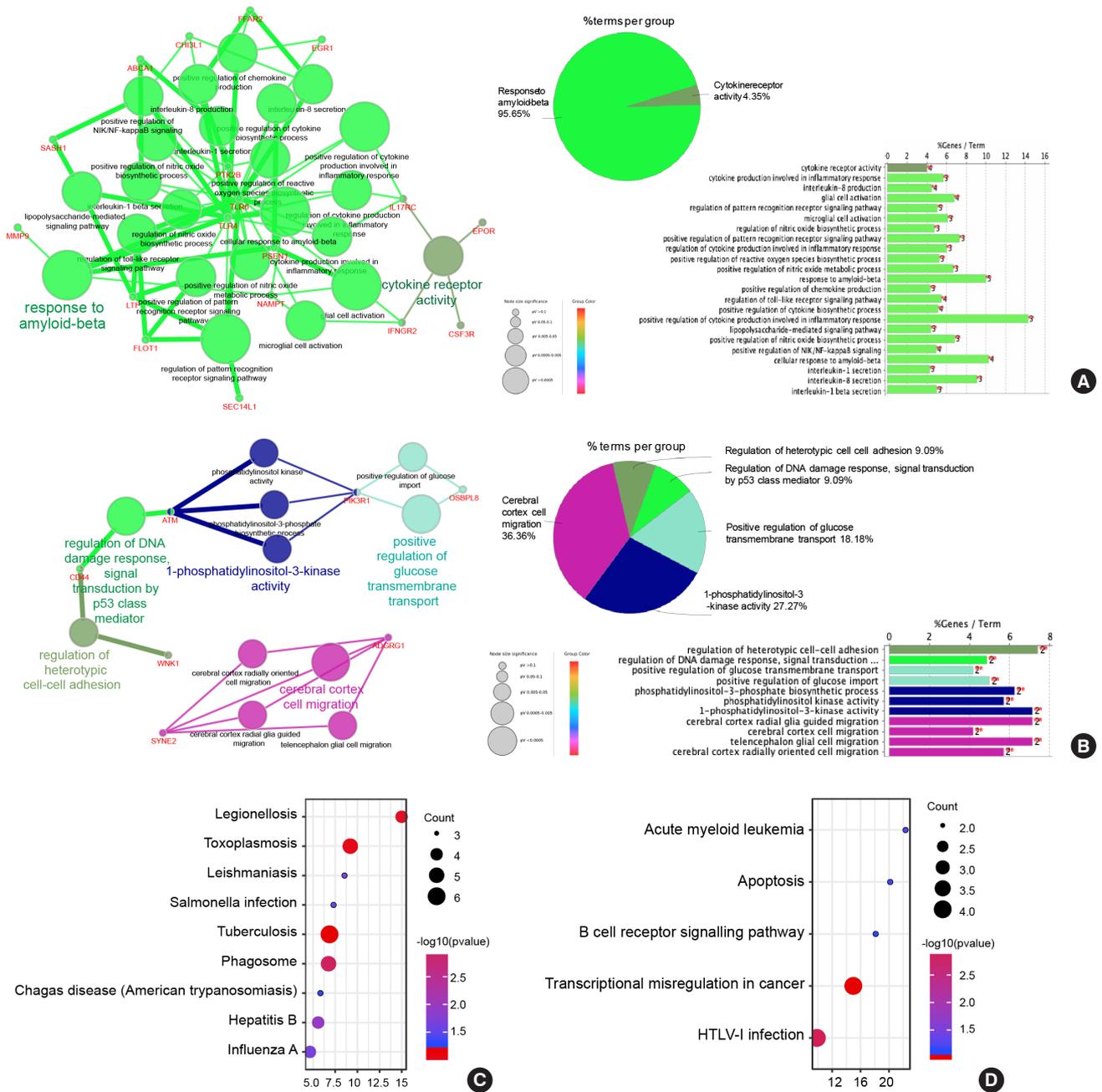
**Fig. 2.** Volcano plot of three datasets. (A) Volcano plots of the distribution of differentially expressed genes (DEGs) in each dataset, fold change  $> 1.2$ ,  $P < 0.05$ . (B) Venn plot of upregulated and downregulated DEGs in these datasets. (C) The expression heatmap of the top 10 upregulated and downregulated DEGs.

leukin-8 secretion. In addition, downregulated DEGs were significantly enriched in 1-phosphatidylinositol-3-kinase activity, cerebral cortex cell migration, positive regulation of glucose transmembrane transport, etc. (Fig. 3B). Interestingly, downregulated DEGs in PBMCs from T1DM patients were more enriched in glucose metabolism processes, not only positive regulation of glucose transmembrane transport but also positive regulation of glucose import. Moreover, by KEGG pathway enrichment, upregulated DEGs were enriched in infectious diseases, including tuberculosis, *Legionellosis*, hepatitis B, influenza A, and *Salmonella* infection, as well as parasitic diseases, such as toxoplasmosis, leishmaniasis, and Chagas disease (Fig. 3C). Downregulated DEGs were enriched in transcriptional misregulation in cancer, human T-lymphotropic virus type 1 (HTLV-1) infection, acute myeloid leukemia, apoptosis, and the B-cell receptor signaling pathway (Fig. 3D, Supplementary Table 4).

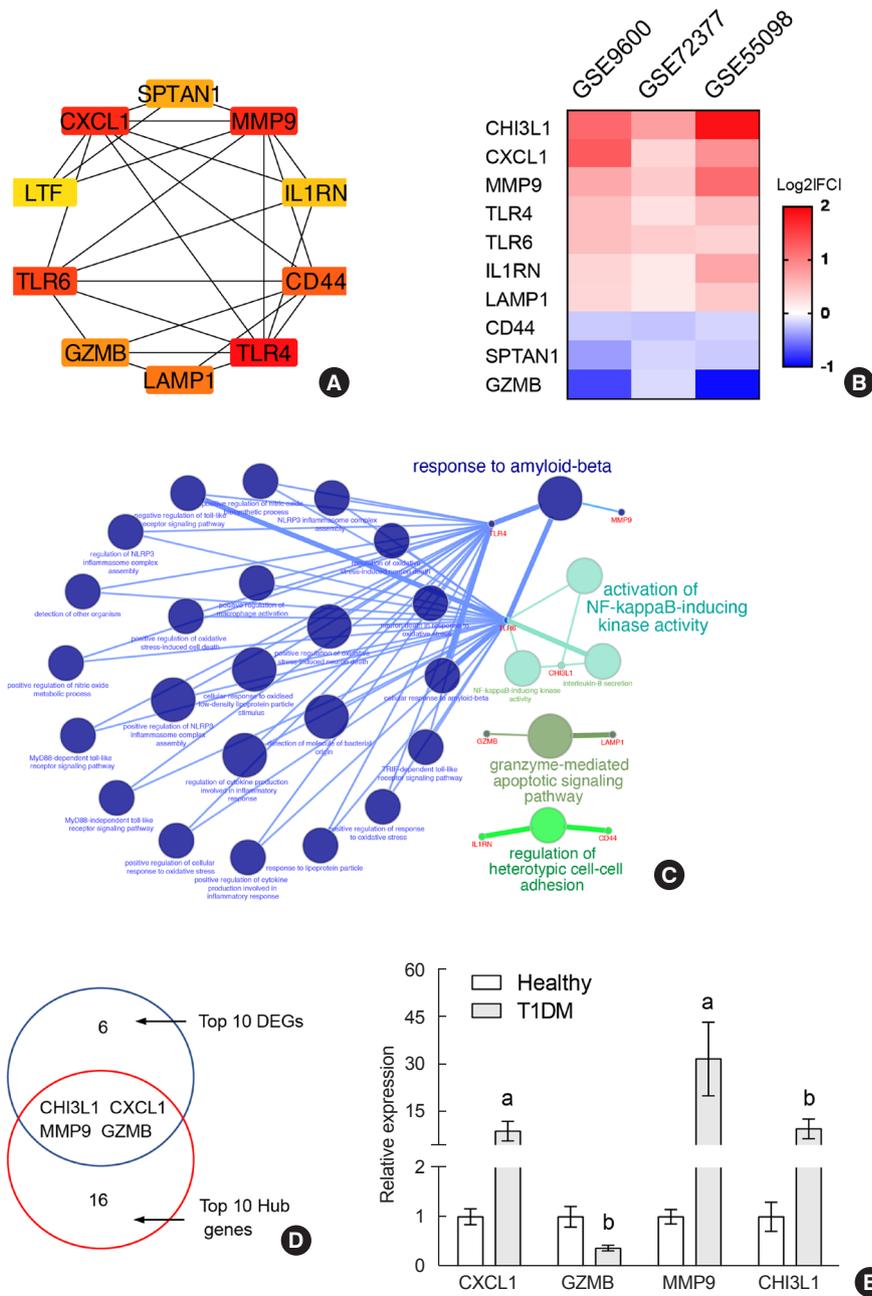
### PPI network and modular analysis

To establish the interactions of these DEGs in PBMCs of

T1DM, we used the STRING database to construct a PPI network. A total of 85 DEGs were imported into the DEG PPI network complex, which was improved by removing isolated genes via Cytoscape software (Supplementary Fig. 1A). On the basis of MCODE application in Cytoscape, we identified three modules in the whole network (Supplementary Fig. 1B-D). The genes of the top three modules are shown in Supplementary Table 5. Moreover, biological function and pathway enrichment analyses based on DAVID were performed for DEGs in the three modules (Supplementary Table 6, Supplementary Fig. 1E): the DEGs of module 1 were significantly enriched in inflammatory response, immune response, positive regulation of NLR family pyrin domain containing 3 (NLRP3) inflammasome complex assembly, etc.; the DEGs of module 2 were enriched in negative regulation of apoptotic process, inflammatory response, immune response, positive regulation of NLRP3 inflammasome complex assembly, etc.; the results of module 3 enrichment were similar to those of modules 1 and 2 (Supplementary Table 6, Supplementary Fig. 1). For the KEGG pathway, module 1 DEGs were significantly enriched in pathways



**Fig. 3.** Functional and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of common differentially expressed genes (DEGs). Biological functional enrichment of upregulated (A) and downregulated (B) DEGs via Cytoscape Gluego. Bubble plot of upregulated (C) and downregulated (D) DEG-enriched KEGG pathways via Database for Annotation, Visualization and Integrated Discovery (DAVID). FFAR2, free fatty acid receptor 2; CHI3L1, chitinase-3-like protein 1; EGR1, early growth response 1; ABCA1, ATP binding cassette subfamily A member 1; SASH1, SAM and SH3 domain containing 1; PTK2B, protein tyrosine kinase 2 beta; TLR, toll like receptor; MMP9, matrix metalloproteinase-9; IL17RC, interleukin 17 receptor C; PSEN1, presenilin 1; EPOR, erythropoietin receptor; LTP, lipid-transfer protein; NAMPT, nicotinamide phosphoribosyl transferase; IFNGR2, interferon gamma receptor 2; CSF3R, colony stimulating factor 3 receptor; FLOT1, flotillin 1; SEC14L1, SEC14 like lipid binding 1; ATM, ATM serine/threonine kinase; WNK1, WNK lysine deficient protein kinase 1; PIK3R1, phosphoinositide-3-kinase regulatory subunit 1; OSBPL8, oxysterol binding protein like 8; SYNE2, spectrin repeat containing nuclear envelope protein 2; ADGRG1, adhesion G protein-coupled receptor G1; HTLV-1, human T-lymphotropic virus type 1.



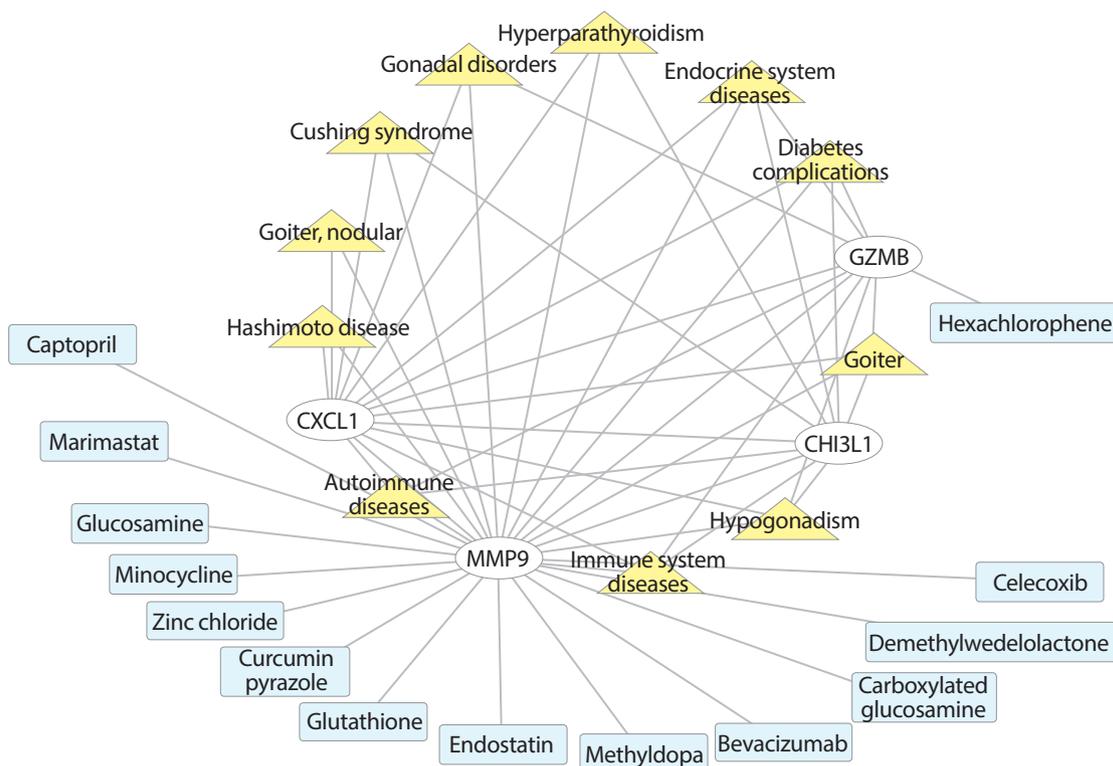
**Fig. 4.** Determination of vital hub genes and related biological processes. (A) Top 10 ranked hub gene networks generated by Cytoscape. (B) The expression heatmap of the top 10 ranked hub genes. (C) The network of hub genes and related biological processes constructed by ClueGO. (D) Venn plot of common genes between the top 10 regulated differentially expressed genes (DEGs) and hub genes: chitinase-3-like protein 1 (CHI3L1), chemokine C-X-C motif ligand 1 (CXCL1), matrix metalloproteinase-9 (MMP9), and granzyme B (GZMB). (E) Reverse transcription-polymerase chain reaction (RT-PCR) detection of CHI3L1, CXCL1, MMP9, and GZMB mRNA expression in peripheral blood mononuclear cell of healthy participants and type 1 diabetes mellitus (T1DM) patients ( $n=10$ ). SPTAN1, spectrin alpha non-erythrocytic 1; LTF, lactotransferrin; IL1RN, interleukin 1 receptor antagonist; TLR, toll like receptor; LAMP1, lysosomal associated membrane protein 1. <sup>a</sup> $P < 0.05$ , <sup>b</sup> $P < 0.01$ , unpaired  $t$ -test.

of proteoglycans in cancer, *Legionellosis* and *Salmonella* infection. Module 2 DEGs were also enriched in pathways of proteoglycans in cancer, transcriptional misregulation in cancer, and several infectious disease pathways, including hepatitis B, Chagas disease, and various immune processes, such as the Toll-like receptor signaling pathway. Module 3 pathways were similar to the module 1 and 2 pathways, which were related to infectious diseases, cancer and immune processes (Supplementary Table 7).

### Identification of key hub genes in PBMCs of T1DM and clinical validation

The top 10 hub genes were filtered out among DEGs using the cytoHubba application in Cytoscape. They were TLR4, CXCL1, MMP9, TLR6, CD44, LAMP1, GZMB, SPTAN1, IL1RN, and CHI3L1 ranked by the MCC method, and their full names and functions are listed in Supplementary Table 8. The PPI network of these hub genes is shown in Fig. 4A. The expression of upregulated hub genes (CHI3L1, CXCL1, MMP9, TLR4, TLR6, and IL1RN) and downregulated hub genes (LAMP1, CD44, SPTAN1, and GZMB) in the three datasets is

listed as a heatmap (Fig. 4B). Moreover, the biological functions of the 10 hubs were significantly enriched in response to A $\beta$  (TLR4, MMP9, and TLR6), activation of nuclear factor  $\kappa$ B (NF- $\kappa$ B)-inducing kinase activity (TLR6 and CHI3L1), granzyme-mediated apoptotic signaling pathway (GZMB and LAMP1), and regulation of heterotypic cell-cell adhesion (CD44 and IL1RN) (Fig. 4C). Interestingly, we found four common genes between the top 10 significantly changed DEGs and the top 10 hub genes: CXCL1, GZMB, MMP9, and CHI3L1 (Fig. 4D). We recruited age-matched healthy and new-onset T1DM participants. T1DM patients showed hyperglycemia, higher HbA1c levels, low C-peptide and positivity for GADA, while healthy participants displayed these indicators within the normal range (Supplementary Table 1). We detected transcriptional expression in PBMCs from 10 T1DM patients and 10 healthy individuals by RT-PCR and observed significantly and robustly increased CXCL, MMP9, and CHI3L1 and decreased GZMB in the T1DM group (Fig. 4E), which was consistent with the results from the three datasets, confirming that these four novel hub genes might be stable reflections of immune system conditions in T1DM patients.



**Fig. 5.** Disease-gene-drug network analysis centered on four vital hub genes. CXCL1, chemokine (C-X-C motif) ligand 1; MMP9, matrix metalloproteinase-9; CHI3L1, chitinase-3-like protein 1; GZMB, granzyme B.

### Disease-gene-drug network-based analysis

The four key hub genes (CXCL1, GZMB, MMP9, and CHI3L1) were screened for discovering disease-target-drug interactions using CTD (<http://ctdbase.org/tools/batchQuery.go>), DGIdb (<https://dgidb.org>), and Cytoscape (Fig. 5). The network showed that in addition to immune system and autoimmune diseases, CXCL1, GZMB, MMP9, and CHI3L1 were all associated with endocrinology system diseases, including diabetes complications, goiter, and hypogonadism. CXCL1, MMP9, and GZMB are related to gonadal disorders, and CXCL1, MMP9, and CHI3L1 are associated with Cushing syndrome and hyperparathyroidism. CXCL1 and MMP3 are both related to Hashimoto disease and nodular goiter. In the target-drug network, GZMB is a serine protease encoded by GZMB that can be inhibited by hexachlorophene (disinfectant), which helps in the development of small-molecule inhibitors of human Kaposi's sarcoma-associated herpesvirus (KSHV) [19]. The MMP9 gene represents MMP9 and may play an essential role in local proteolysis of the extracellular matrix and in leukocyte migration [20]. Among various drugs targeted to MMP9, MMP inhibitors and antagonists include captopril (angiotensin converting enzyme inhibitor) [21], marimastat (angiogenesis and metastasis inhibitor) [22], minocycline (tetracycline antibiotics) [23], carboxylated glucosamine [24], and glucosamine [25]. Zinc chloride is a binder for MMP9 [26]. Other clinical drugs also interact with MMP9, including glutathione, endostatin, methyl dopa, bevacizumab, and celecoxib. Glutathione maintains normal immune system function and has antioxidant and integrated detoxification effects. Endostatin is the most studied endogenous angiogenesis inhibitor. Methyl dopa is a centrally acting antiadrenergic drug. Bevacizumab is used to treat cancer of the membrane lining the internal organs in the abdomen and is usually given as part of a combination of cancer medicines. Celecoxib is a nonsteroidal anti-inflammatory drug (NSAID) for reducing pain or inflammation caused by many conditions, such as arthritis and ankylosing spondylitis.

## DISCUSSION

For more than 25 years, T1DM has been recognized as a chronic autoimmune disease characterized by progressive loss of beta cells in the pancreas that results in insulin dependence for a lifetime. The activation and dysfunction of innate immunity is an early event in T1DM that appears to be triggered by environmental factors on the appropriate genetic background

[27] and has recently aroused increasing attention. Innate immunity is the first line of defense against pathogens in the human body. Meanwhile, PBMCs, a group of peripheral blood cells with a nucleus such as lymphocytes (e.g., T-cells, B-cells, and natural killer cells) and monocytes, are vital components of innate immunity [28]. In the present study, we identified common DEGs, including 52 upregulated and 33 downregulated genes, on the basis of three GEO datasets using the Venn method. Seventy T1DM patients and 54 normal participants were enrolled in the present research. GO and KEGG pathway enrichment analyses showed that these DEGs were significantly enriched in pathways related to the immune system and glucose metabolic functions. The PPI network was constructed by DEGs. Module and hub gene detection and enrichment analysis were performed. Finally, four key hub genes were identified and validated via clinical samples. Furthermore, the network-based analysis evaluated the connection among the key hub genes, diseases and drugs, suggesting the multifunction of these genes, which might provide possible insight for future studies concerning disease pathogenesis, prevention, and management.

Previously, several bioinformatic studies investigated gene changes in important tissues/organs of diabetic patients, such as adipose tissue [29], muscle [30], pancreas [31], and kidney [32]. PBMCs are an important part of the immune system that participates in T1DM pathology, of which the previous gene sequencing results reported that the most overexpressed genes were IL1B, EGR2, EGR3, prostaglandin-endoperoxide synthase 2 (PTGS2), C-C motif chemokine receptor 1 (CCR1), and FosB proto-oncogene. The most significantly underexpressed genes included ras-related protein rab-12 (RAB12) and serine and arginine rich splicing factor 5 (SRSF5). Cellular functions were enriched in apoptosis of eukaryotic cells, proliferation of cells and development of lymphatic system cells [10].

In this study, we integrated three datasets with a larger sample size and found that DEGs were significantly centralized in various immune-related functions and pathways. GO biological functional enrichment showed that DEGs were particularly enriched in processes related to the immune system, such as upregulated response to A $\beta$  and cytokine receptor activity, as well as downregulated 1-phosphatidylinositol-3-kinase activity. There is extensive evidence that the accumulation of mononuclear phagocytes at sites of A $\beta$  deposition in the brain is an important pathological feature of Alzheimer's disease (AD) [33]. The enhanced biological process of response to A $\beta$  in

PBMCs of type 2 diabetes mellitus (T2DM) reflects the relevance between the immune state of diabetes and an early manifestation of AD [34]. Notably, downregulated processes included not only immune aspects but also glucose metabolism processes, such as positive regulation of glucose transmembrane transport and positive regulation of glucose import, revealing decreased glucose intake in PBMCs of T1DM patients, as well as the effect of hyperglycemia on the immune system. Meanwhile, upregulated pathways enriched by KEGG showed an obvious focus on infectious diseases (including tuberculosis, Legionellosis, and hepatitis B) and parasitic diseases (toxoplasmosis, leishmaniasis, and Chagas disease/American trypanosomiasis). T1DM increases the prevalence of infectious diseases [35]; moreover, considering the “hygiene hypothesis” that contact with microorganisms or parasites helps to prevent and treat T1DM [36,37], and based on genomic observation, our results provide a possible explanation for the relationship between T1DM and infectious/parasitic diseases.

Based on the module analysis of the PPI network and hub gene identification, we obtained three modules and four key genes in the whole network. As the module contained a cluster of the most vital genes, representative module 1 displayed enrichment in inflammatory response, immune response, positive regulation of NLRP3 inflammasome complex assembly, etc. Interestingly, NLRP3 is a critical component of innate immunity and contributes to the pathology of various metabolic diseases, including diabetes and insulin resistance [38,39]. Previous studies reported that NLRP3 expression was higher in PBMCs of diabetes patients, while NLRP3 expression was considerably decreased after medication treatment and lifestyle interventions [40]. In addition, NLRP3 deficiency protects against T1DM through the regulation of chemotaxis into pancreatic islets [41]. We also found an upregulated response to the A $\beta$  signaling pathway. Islet amyloid polypeptide, a protein that forms amyloid deposits in the pancreas during T2DM, is a known activator of the NLRP3 inflammasome [42]. Thus, the amyloid-NLRP3 pathway might be a potential target.

Moreover, four vital hub genes, namely, CHI3L1, CXCL1, MMP9, and GZMB, were identified as both top-ranked DEGs and hub genes. First, CHI3L1 is expressed by macrophages [43] and is elevated in patients with obesity, cardiovascular disease, and T2DM [44]. The connection between CHI3L1 and NF- $\kappa$ B helped to counteract tumor necrosis factor  $\alpha$ -mediated inflammation, and CHI3L1 normalized impaired insulin action and insulin-stimulated glucose uptake [45]. Thus, the ele-

vated CHI3L1 in PBMCs might be a protective response in T1DM. Second, CXCL1 is a small peptide belonging to the CXC chemokine family that acts as a chemoattractant for several immune cells, especially neutrophils, to the site of injury or infection [46,47]. CXCL1 is elevated during inflammatory responses [48]; thus, the upregulation of CXCL1 in our study might reflect arousal of the immune response against hyperglycemia. Third, MMP9 belongs to the zinc metalloproteinase family and plays roles in neutrophil action, such as degrading extracellular matrix, activating IL-1 $\beta$ , and cleaving several chemokines [49]. Although MMP9 participates in various diseases and acts as an important drug target in the disease-gene-drug network, the association between MMP9 and PBMCs from diabetes patients, as well as its diagnostic and therapeutic potential in T1DM, are still worth further clarification. Finally, serine protease GZMB is a marker of the cytotoxic potential of cytotoxic T lymphocyte (CTL)-mediated  $\beta$ -cell destruction/apoptosis in T1DM, and its level increases when CTLs enter the islets [50]. We concluded that there is a reduced GZMB level in PBMCs of T1DM, suggesting a decreased peripheral CTL, which was inconsistent with previous studies that in T1DM, GZMB was required for efficient early activation of CTL and maintained homeostasis of peripheral CD8+ T-cells for efficient proliferation in the pancreatic lymph nodes [51]. The reason might be that although the GZMB gene was expressed in PBMCs, it does not represent the GZMB protease concentration level, especially at its site of action, and its apoptotic contribution to pancreatic  $\beta$ -cells.

However, there were several limitations of the present study that require acknowledgment. First, our study utilized statistics and bioinformatics, and further studies with much larger sample sizes and detailed cell types are needed. In addition, the predicted results still require confirmation by well-designed intervention experiments *in vivo* and *in vitro*. Whether the hub gene in PBMCs could be a gold standard predictor or therapeutic target for T1DM still needs to be further investigated.

Taken together, our integrated bioinformatics analysis study identified significantly changed functions and pathways in PBMCs of T1DM, as well as key hub genes that may be considered major and potential PBMC biomarkers of T1DM. An expression validation experiment using clinical samples was finally used to test the robustness of the results. Therefore, this study provided reliable and readily accessible indicators for the immune system in T1DM patients, which would be helpful for

convenient genetic evaluation, aiming to guide prevention, diagnosis, and even targeted therapy in T1DM.

## SUPPLEMENTARY MATERIALS

Supplementary materials related to this article can be found online at <https://doi.org/10.4093/dmj.2021.0018>.

## CONFLICTS OF INTEREST

No potential conflict of interest relevant to this article was reported.

## AUTHOR CONTRIBUTIONS

Conception or design: X.L., J.S.

Acquisition, analysis, or interpretation of data: X.L., M.L., J.G., L.Z., R.S.

Drafting the work or revising: X.L., M.L., J.S.

Final approval of the manuscript: X.L., M.L., J.G., L.Z., R.S., M.L., J.S.

## ORCID

Xing Li <https://orcid.org/0000-0001-6469-4839>

Min Long <https://orcid.org/0000-0003-1071-8131>

Jiaqing Shao <https://orcid.org/0000-0002-9739-5410>

## FUNDING

This work was supported by National Natural Science Foundation of China (No. 82000795) and 69th Postdoctoral Science Foundation of China (No. 2021M693952) to Xing Li; National Natural Science Foundation of China (No. 81873174) to Jiaqing Shao; Chongqing Natural Science Foundation (Outstanding Youth Foundation) (No. cstc2020jcyj-jqx0017) to Min Long.

## ACKNOWLEDGMENTS

None

## REFERENCES

- Bluestone JA, Herold K, Eisenbarth G. Genetics, pathogenesis and clinical interventions in type 1 diabetes. *Nature* 2010;464:1293-300.
- Chiang JL, Kirkman MS, Laffel LM, Peters AL; Type 1 Diabetes Sourcebook Authors. Type 1 diabetes through the life span: a position statement of the American Diabetes Association. *Diabetes Care* 2014;37:2034-54.
- GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018;392:1789-858.
- McKnight JA, Wild SH, Lamb MJ, Cooper MN, Jones TW, Davis EA, et al. Glycaemic control of type 1 diabetes in clinical practice early in the 21st century: an international comparison. *Diabet Med* 2015;32:1036-50.
- Mayer-Davis EJ, Lawrence JM, Dabelea D, Divers J, Isom S, Dolan L, et al. Incidence trends of type 1 and type 2 diabetes among youths, 2002-2012. *N Engl J Med* 2017;376:1419-29.
- Herold KC, Vignali DA, Cooke A, Bluestone JA. Type 1 diabetes: translating mechanistic observations into effective clinical outcomes. *Nat Rev Immunol* 2013;13:243-56.
- Bulek AM, Cole DK, Skowera A, Dolton G, Gras S, Madura F, et al. Structural basis for the killing of human beta cells by CD8(+) T cells in type 1 diabetes. *Nat Immunol* 2012;13:283-9.
- Ferreira RC, Guo H, Coulson RM, Smyth DJ, Pekalski ML, Burren OS, et al. A type I interferon transcriptional signature precedes autoimmunity in children genetically at risk for type 1 diabetes. *Diabetes* 2014;63:2538-50.
- Beyan H, Drexhage RC, van der Heul Nieuwenhuijsen L, de Wit H, Padmos RC, Schloot NC, et al. Monocyte gene-expression profiles associated with childhood-onset type 1 diabetes and disease risk: a study of identical twins. *Diabetes* 2010;59:1751-5.
- Kaizer EC, Glaser CL, Chaussabel D, Banchereau J, Pascual V, White PC. Gene expression in peripheral blood mononuclear cells from children with diabetes. *J Clin Endocrinol Metab* 2007;92:3705-11.
- Yang M, Ye L, Wang B, Gao J, Liu R, Hong J, et al. Decreased miR-146 expression in peripheral blood mononuclear cells is correlated with ongoing islet autoimmunity in type 1 diabetes patients 1miR-146. *J Diabetes* 2015;7:158-65.
- American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2007;30 Suppl 1:S42-7.
- American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care* 2011;34 Suppl 1:S62-9.

14. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 2009;25:1091-3.
15. Bindea G, Galon J, Mlecnik B. CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics* 2013;29:661-3.
16. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 2015;43(Database issue):D447-52.
17. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13:2498-504.
18. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. CytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 2014;8 Suppl 4:S11.
19. Shahian T, Lee GM, Lazic A, Arnold LA, Velusamy P, Roels CM, et al. Inhibition of a viral enzyme by a small-molecule dimer disruptor. *Nat Chem Biol* 2009;5:640-6.
20. Tschesche H, Knauper V, Kramer S, Michaelis J, Oberhoff R, Reinke H. Latent collagenase and gelatinase from human neutrophils and their activation. *Matrix Suppl* 1992;1:245-55.
21. Williams RN, Parsons SL, Morris TM, Rowlands BJ, Watson SA. Inhibition of matrix metalloproteinase activity and growth of gastric adenocarcinoma cells by an angiotensin converting enzyme inhibitor in in vitro and murine models. *Eur J Surg Oncol* 2005;31:1042-50.
22. Chen X, Ji ZL, Chen YZ. TTD: Therapeutic Target Database. *Nucleic Acids Res* 2002;30:412-5.
23. Lee CZ, Yao JS, Huang Y, Zhai W, Liu W, Guglielmo BJ, et al. Dose-response effect of tetracyclines on cerebral matrix metalloproteinase-9 after vascular endothelial growth factor hyperstimulation. *J Cereb Blood Flow Metab* 2006;26:1157-64.
24. Mendis E, Kim MM, Rajapakse N, Kim SK. Carboxy derivatized glucosamine is a potent inhibitor of matrix metalloproteinase-9 in HT1080 cells. *Bioorg Med Chem Lett* 2006;16:3105-10.
25. Rajapakse N, Mendis E, Kim MM, Kim SK. Sulfated glucosamine inhibits MMP-2 and MMP-9 expressions in human fibrosarcoma cells. *Bioorg Med Chem* 2007;15:4891-6.
26. Martens E, Leyssen A, Van Aelst I, Fiten P, Piccard H, Hu J, et al. A monoclonal antibody inhibits gelatinase B/MMP-9 by selective binding to part of the catalytic domain and not to the fibronectin or zinc binding domains. *Biochim Biophys Acta* 2007;1770:178-86.
27. Ilonen J, Lempainen J, Veijola R. The heterogeneous pathogenesis of type 1 diabetes mellitus. *Nat Rev Endocrinol* 2019;15:635-50.
28. Delves PJ, Roitt IM. *Roitt's essential immunology*. 12th ed. Chichester: Wiley-Blackwell; 2011. p546.
29. Doumatey AP, Xu H, Huang H, Trivedi NS, Lei L, Elkahoulou A, et al. Global gene expression profiling in omental adipose tissue of morbidly obese diabetic African Americans. *J Endocrinol Metab* 2015;5:199-210.
30. Shao K, Shen LS, Li HH, Huang S, Zhang Y. Systematic-analysis of mRNA expression profiles in skeletal muscle of patients with type II diabetes: the glucocorticoid was central in pathogenesis. *J Cell Physiol* 2018;233:4068-76.
31. Zhong M, Wu Y, Ou W, Huang L, Yang L. Identification of key genes involved in type 2 diabetic islet dysfunction: a bioinformatics study. *Biosci Rep* 2019;39:BSR20182172.
32. Van JA, Scholey JW, Konvalinka A. Insights into diabetic kidney disease using urinary proteomics and bioinformatics. *J Am Soc Nephrol* 2017;28:1050-61.
33. Gold M, El Khoury J.  $\beta$ -Amyloid, microglia, and the inflammasome in Alzheimer's disease. *Semin Immunopathol* 2015;37:607-11.
34. Morgese MG, Schiavone S, Trabace L. Emerging role of amyloid beta in stress response: implication for depression and diabetes. *Eur J Pharmacol* 2017;817:22-9.
35. Carey IM, Critchley JA, DeWilde S, Harris T, Hosking FJ, Cook DG. Risk of infection in type 1 and type 2 diabetes compared with the general population: a matched cohort study. *Diabetes Care* 2018;41:513-21.
36. Mendez-Samperio P, de-la-Rosa-Arana JL. Prevention of type 1 diabetes through parasite infection. *Immunotherapy* 2015;7:595-8.
37. Kuhlreiber WM, Faustman DL. BCG therapy for type 1 diabetes: restoration of balanced immunity and metabolism. *Trends Endocrinol Metab* 2019;30:80-92.
38. Stienstra R, van Diepen JA, Tack CJ, Zaki MH, van de Veerdonk FL, Perera D, et al. Inflammasome is a central player in the induction of obesity and insulin resistance. *Proc Natl Acad Sci U S A* 2011;108:15324-9.
39. He Y, Hara H, Nunez G. Mechanism and regulation of NLRP3 inflammasome activation. *Trends Biochem Sci* 2016;41:1012-21.
40. Wan Z, Fan Y, Liu X, Xue J, Han Z, Zhu C, et al. NLRP3 inflam-

- masome promotes diabetes-induced endothelial inflammation and atherosclerosis. *Diabetes Metab Syndr Obes* 2019;12:1931-42.
41. Wu KK. Deep fibromatosis of the foot with metatarsal involvement. *J Foot Surg* 1989;28:586-90.
  42. Masters SL, Dunne A, Subramanian SL, Hull RL, Tannahill GM, Sharp FA, et al. Activation of the NLRP3 inflammasome by islet amyloid polypeptide provides a mechanism for enhanced IL-1 $\beta$  in type 2 diabetes. *Nat Immunol* 2010;11:897-904.
  43. Rehli M, Krause SW, Andreesen R. Molecular characterization of the gene for human cartilage gp-39 (CHI3L1), a member of the chitinase protein family and marker for late stages of macrophage differentiation. *Genomics* 1997;43:221-5.
  44. Batinic K, Hobaus C, Grujicic M, Steffan A, Jelic F, Lorant D, et al. YKL-40 is elevated in patients with peripheral arterial disease and diabetes or pre-diabetes. *Atherosclerosis* 2012;222:557-63.
  45. Gorgens SW, Eckardt K, Elsen M, Tennagels N, Eckel J. Chitinase-3-like protein 1 protects skeletal muscle from TNF $\alpha$ -induced inflammation and insulin resistance. *Biochem J* 2014;459:479-88.
  46. Moser B, Clark-Lewis I, Zwahlen R, Baggiolini M. Neutrophil-activating properties of the melanoma growth-stimulatory activity. *J Exp Med* 1990;171:1797-802.
  47. Schumacher C, Clark-Lewis I, Baggiolini M, Moser B. High- and low-affinity binding of GRO alpha and neutrophil-activating peptide 2 to interleukin 8 receptors on human neutrophils. *Proc Natl Acad Sci U S A* 1992;89:10542-6.
  48. Silva RL, Lopes AH, Guimaraes RM, Cunha TM. CXCL1/CXCR2 signaling in pathological pain: role in peripheral and central sensitization. *Neurobiol Dis* 2017;105:109-16.
  49. Opdenakker G, Van den Steen PE, Dubois B, Nelissen I, Van Coillie E, Masure S, et al. Gelatinase B functions as regulator and effector in leukocyte biology. *J Leukoc Biol* 2001;69:851-9.
  50. Graham KL, Krishnamurthy B, Fynch S, Mollah ZU, Slattery R, Santamaria P, et al. Autoreactive cytotoxic T lymphocytes acquire higher expression of cytotoxic effector markers in the islets of NOD mice after priming in pancreatic lymph nodes. *Am J Pathol* 2011;178:2716-25.
  51. Kobayashi M, Kaneko-Koike C, Abiru N, Uchida T, Akazawa S, Nakamura K, et al. Genetic deletion of granzyme B does not confer resistance to the development of spontaneous diabetes in non-obese diabetic mice. *Clin Exp Immunol* 2013;173:411-8.