Diabetes, Metabolic Syndrome and Obesity

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Identification of Immune Gene Signature Associated with T Cells and Natural Killer Cells in Type 1 Diabetes

Na Wang[1,2], Guofeng Wang[2], Xiuli Feng[2], Teng Yang[2]

[1]Department of Endocrinology, Lianyungang Clinical College of Nanjing Medical University, Lianyungang City, Jiangsu Province, 222000, People's Republic of China; [2]Department of Endocrinology, Jinzhou Medical University(The First People's Hospital of Lianyungang), Lianyungang City, Jiangsu Province, 222000, People's Republic of China

Correspondence: Guofeng Wang, Department of Endocrinology, Lianyungang Clinical College of Nanjing Medical University (The First People's hospital of Lianyungang), No. 6 Zhenhua East Road, Haizhou District, Lianyungang City, Jiangsu Province, 222000, People's Republic of China, Email wangguofengo@21cn.com

Purpose: This study aimed to investigate the abnormal infiltration of immune cells in type 1 diabetes mellitus (T1D) and elucidate their regulatory mechanisms.

Methods: Public T1D-related gene expression data were obtained from the Gene Expression Omnibus database.The GSE123658 dataset analyzed whole blood RNA-seq data from type 1 diabetic patients and healthy volunteers. The GSE110914 dataset analyzed neutrophils purified from peripheral blood of patients with symptomatic and pre-symptomatic type 1 diabetes (T1D), at risk of T1D, and healthy controls. Immune cell infiltration analysis was performed to identify abnormally infiltrating immune cells. Differentially expressed immune genes (DEIGs) in T1D samples were identified, followed by the construction of an immune gene signature (IGS) using a protein-protein interaction (PPI) network and Least absolute shrinkage and selection operator Cox regression analyses (LASSO Cox regression analyses). The regulatory mechanisms underlying IGS were explored using gene set enrichment analysis. Furthermore, expression validation, diagnostic efficacy evaluation, and upstream miRNA prediction of hub signature genes were performed. We verified the miRNA expression of the key gene colony stimulating factor 1 (CSF1) and microRNA-326 (miR-326) by reverse transcription-quantitative PCR (RT–qPCR).

Results: The proportion of infiltrating T and natural killer (NK) cells differed between the T1D and control samples, and 207 immune genes (IGs) related to these immune cells were extracted. After differential expression, PPI, and LASSO Cox regression analyses, four signature DEIGs were identified for IGS construction: notch receptor 1 (NOTCH1), Janus kinase 3 (JAK3), tumor necrosis factor receptor superfamily member 4(TNFRSF4), and CSF1. Key pathways such as the Toll-like receptor signaling pathway were significantly activated in the high-risk group. Moreover, the upregulation of CSF1 in T1D samples was confirmed using a validation dataset, and CSF1 showed high diagnostic efficacy for T1D. Furthermore, CSF1 was targeted by miR-326.We used validated key genes in T1D patients, several of which were confirmed by RT–qPCR.

Conclusion: In conclusion, the identified key IGs may play an important role in T1D. CSF1 can be developed as a novel diagnostic biomarker for T1D.

Keywords: Type 1 diabetes, immune cell infiltration, immune gene signature, diagnosis, CSF1

Introduction

Type 1 diabetes (T1D) is a major subtype of diabetes. It often occurs during childhood and adolescence, with a global prevalence of around 500,000 individuals.[1] Type 1 diabetes (T1D) is a chronic autoimmune disease characterized by the destruction of pancreatic β-cells, leading to insulin deficiency.[2] The prevalence of T1D is substantial and growing worldwide.[3] T1D often has preventable life-threatening complications, such as cardiovascular and nephropatic diabetic complications, and early diagnosis of T1D is associated with morbidity and quality of life for patients and their families.[4,5] Most screening tests for identifying individuals at risk for T1D target patients' relatives. However,

Received: 23 May 2024
Accepted: 23 July 2024
Published: 9 August 2024

approximately 90% of patients with T1D do not have a family history.[6] Therefore, there is an urgent need to elucidate the crucial mechanism of T1D and explore novel biomarkers for its early diagnosis.

Accumulating evidence has revealed that the complex pathogenesis of T1D is driven by multiple immune cell types, such as T lymphocytes, macrophages, and dendritic cells, which ultimately lead to the destruction of insulin-producing β cells.[7] Moreover, Studies have shown that T1DM development is caused by an imbalance between CD4 + effector cell (Teff) and Treg ratios. Teff can promote pancreatic islet β-cell death and immunoglobulin production. Treg cells inhibit initial T cell activation and inhibit functional of effector T cells by producing anti-inflammatory cytokines: IL-10 and transforming growth factor 1 (TGF-β1).[8] In addition, immune responses are observed during the development of T1D, which require the coordinated effects of multiple types of immune cells and a large number of immune-related genes (IGs).[9–11] Two hypotypes of $CD4^+$ T cells, Th1 cells and Tregs, promote the progression of T1D by secreting cytokines such as interleukin 1, which can destroy β cells.[12] Immune cells, such as natural killer (NK) cells, and IGs, such as Natural Cytotoxicity Triggering Receptor 3 (NCR3) and tumor necrosis factor (TNF), are key regulators of the occurrence and development of T1D.[13] However, the crucial immune cells and key mechanisms underlying aberrant immune cell infiltration in the development of T1D are not fully understood, which hinders the design of effective strategies to prevent and treat the disease.

In the present study, we downloaded publicly available T1D-related genes expression data and performed immune cell infiltration analysis to identify key immune cells and related IGs in T1D. The GSE123658 dataset analyzed whole blood RNA-seq data from type 1 diabetic patients and healthy volunteers. The GSE110914 dataset analyzed neutrophils purified from peripheral blood of patients with symptomatic and pre-symptomatic type 1 diabetes (T1D), at risk of T1D, and healthy controls. Immune cell infiltration analysis was performed to identify abnormally infiltrating immune cells.We screened key differentially expressed immune genes (DEIGs) in T1D samples and constructed an immune gene signature (IGS). The regulatory mechanisms underlying IGS were explored using gene set enrichment analysis (GSEA). Furthermore, hub signature genes were identified by expression validation using a validation dataset, followed by diagnostic efficacy evaluation and upstream miRNA prediction and were subsequently experimentally validated. This study provides novel insights into the design of efficient diagnostic biomarkers and therapeutic targets for T1D.

Materials and Methods

Patient Recruitment and Clinical Data Collection

During the experimental period, 60 T1DM patients were recruited for peripheral blood collection at the Lianyungang First People's Hospital. All patients underwent a definitive diagnosis. Inclusion criteria: (1) Meet WHO diagnostic criteria for diabetes; (2) Since the onset of the disease, it has always relied on insulin replacement therapy, with abdominal C-peptide<300pmol/L at the time of onset; (3) At least one antibody is positive for glutamic acid decarboxylase antibody (GADA), protein tyrosine phosphatase antibody (IA-2A), zinc transporter protein 8 antibody (ZnT8A), and pancreatic islet cell antibody (ICA); (4) Chinese Han population.In addition, we obtained peripheral blood from 22 healthy volunteers as controls. The study protocol was approved by the Ethics Committee of the Lianyungang First People's Hospital and complied with the Declaration of Helsinki. All the subjects were required to provide written informed consent.

Data Acquisition

The gene expression profiles GSE123658 (training dataset) and GSE110914 (validation dataset) were downloaded from the NCBI Gene Expression Omnibus (GEO) database.[14] The GSE123658 dataset includes 39 T1D and 43 control samples. After PCA, samples with poor clustering were eliminated and 27 T1D and 18 control samples were retained for analysis. The GSE110914 dataset included 5 T1D and 16 control samples.

Immune Cell Infiltration

Based on the GSE123658 dataset, the proportions of 22 types of immune cells in each sample were analyzed using the CIBERSORT algorithm.[15] Key immune cells with different infiltration proportions between T1D and control samples

were selected as the target immune cells. To explore the biological characteristics of target immune cells in T1D, items related to target immune cells were searched using the Molecular Signatures Database (MSigDB),[16] and IGs related to target immune cells were extracted.

GSEA of Pathways Associated with Target Immune Cells

To observe target immune cell-associated biological alterations in T1D, we subjected all DEGs to GSEA. We downloaded reference gene sets "c2.cgp.v2022.1.Hs.symbols.gmt" from database MSigDB.[17] Hs.symbols.gmt in the MSigDB database as the enrichment background, GSEA was performed using the R clusterProfiler package (version 4.4.4).[18] The significant enrichment threshold was set at $p<0.05$, and pathways associated with target immune cells were considered.

Differential Expression Analysis

Based on GSE123658 dataset, the differentially expressed genes (DEGs) between T1D and control samples were identified using R edgeR package.[19] The cutoff value for DEGs screening was adjusted to a p-value<0.05 and |log fold change (FC)| > 0.585.

DEIGs Identification

DEIGs were obtained by Venn diagram analysis of DEGs and IGs using the R Venn Detail package (version 1.2.0).[20]

Protein-Protein Interaction Network Analysis and Hub Gene Identification

Protein-protein interaction (PPI) relationships between DEIGs were analyzed using the STRING database (version 11.5).[21] The PPI network of the DEIGs was visualized using Cytoscape (version 3.9.2).[22] The interaction score was 0.4. The four topological properties of PPI nodes were analyzed using cytoHubba plug-in[23] in Cytoscape, including the maximal neighborhood component (MNC), maximal clique centrality (MCC), degree, and edge-percolated component (EPC). The overlapping genes of the top six DEIGs from the four algorithms were considered hub DEIGs.

LASSO Cox Regression Analysis

The expression of hub DEIGs in each sample was extracted and subjected to LASSO Cox regression analysis using the R glmnet package (version 4.0–2).[24] The optimal λ-value was obtained by ten-fold cross-validation to further screen for the optimized combination of DEIGs. The IG signature (IGS) was then constructed by these DEIGs, with the following formula: IGS score = $\sum$ βgene × Expgene, where βgene indicates the coefficient of each DEIG in LASSO Cox regression, and Expgene represents the expression value of the corresponding DEIG. The IGS score for each sample was calculated.

GSEA of IGS-Associated Pathways

The samples were divided into high-risk and low-risk groups based on the median IGS scores. To explore the potential biological mechanisms underlying IGS, we used c5.go.bp.v2022.1. Hs.symbols.gmt and c2.cp.kegg.v2022.1. Hs.symbols.gmt in the MSigDB database was used as the enriched background, and functional enrichment analysis was performed to determine the enriched GO terms and KEGG pathways using the R clusterProfiler package. The cutoff value was set at $p < 0.05$.

Differential Expression Validation

Based on the validation dataset GSE110914, the expression of signature genes in T1D and control samples was compared using a *t*-test. Genes with the same expression trend in the training dataset GSE123658 were selected as hub signature genes.

Diagnostic Efficacy Evaluation of Hub Signature Gene

Based on the expression data of the hub signature genes in the training and validation datasets, the diagnostic efficacy of the signature genes was evaluated by plotting the ROC curve using the R pROC package (version 1.18.0).[25] A higher

area under the curve (AUC) value indicated stronger diagnostic efficacy. Moreover, the relationship between signature genes and IGS score was analyzed using Spearman correlation analysis.

Prediction of Upstream miRNAs of Hub Signature Gene

T1D-related miRNAs were searched in the HMDD database.[26] The upstream miRNAs of the hub signature genes were predicted using the ENCORI database (https://starbase.sysu.edu.cn/index.php). T1D-related upstream miRNAs of hub signature genes were obtained.

Statistical Analysis

All statistical analyses were performed using R language (version 4.2.2), and $P < 0.05$ was considered statistically significant. All statistical analyses were performed using SPSS version 26.0 (SPSS Inc., Chicago, IL, USA), and graphs were created using GraphPad Prism 9 (GraphPad Software, La Jolla, California). Normally distributed measures were expressed as mean±SD, and comparisons between the two groups were made using the independent samples *t*-test; non-normally distributed measures were expressed as median (25th–75th interquartile range) and the Mann–Whitney *U*-test was used. Bivariate correlation analysis was performed between CSF1,miR-326 and clinical indicators. ROC curve was plotted and area under the curve (AUC) was calculated to analyze the diagnostic value of CSF1 and miR-326 in T1DM patients. All statistical analyses were two-sided and p values<0.05 were considered statistically significant.

RT-qPCR

Quantitative reverse transcription polymerase chain reaction (qRT-PCR) was used to detect CSF1 and miR-326 expression levels in PBMCs. 5mL of fasting venous blood was collected with EDTA dipotassium salt (EDTA-K2) anticoagulant tube, and Peripheral blood mononuclear cells (PBMCs) were isolated with peripheral blood lymphocyte isolation solution. Total RNA was extracted using Trizol reagent instructions. The concentration and purity of RNA were detected by ultraviolet spectrophotometer and the RNA were reverse-transcribed into cDNA according to reverse transcription kit. Finally, using cDNA as template, qRT-PCR was detected by Synergetic Binding Reagent (SYBR) method using Applied Biosystems (ABI) 7500 instrument of American Biosystems. Amplification procedure: 95°C 10min, 95°C 5s, 60°C 30s, 72°C 60s, a total of 40 cycles.Relative expression was calculated by the $2^{-\Delta\Delta Ct}$ method.

Primers for CSF1, miR-326 and internal reference Recombinant Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH),U6 and their corresponding kits were purchased from China's Ruibo Biotechnology Company.Primers were as described in Table 1. Sequences of h-miR-326 and U6 are confidential, High sensitivity small RNA quantitative detection method and kit, Patent license number: CA201410039162.6.

Results

Analysis of Immune Cell Infiltration

Based on the GSE123658 dataset, seven types of immune cells with significantly differential infiltration proportions between the T1D and control samples were identified, including naïve B and $CD4^+$ T cells, memory-activated $CD4^+$ T cells, Tregs, resting NK cells, activated NK cells, and resting dendritic cells (Figure 1A). Using the MSigDB database, the related items of T and NK cells were searched and 207 IGs were extracted from these related items.

Table 1 Primer Information

Target Name		Primer
h-CSF1	F	TCAGATGGAGACCTCGTGCC
	R	GGTGTTATCTCTGAAGCGCATG
GAPDH	F	GAACGGGAAGCTCACTGG
	R	GCCTGCTTCACCACCTTCT

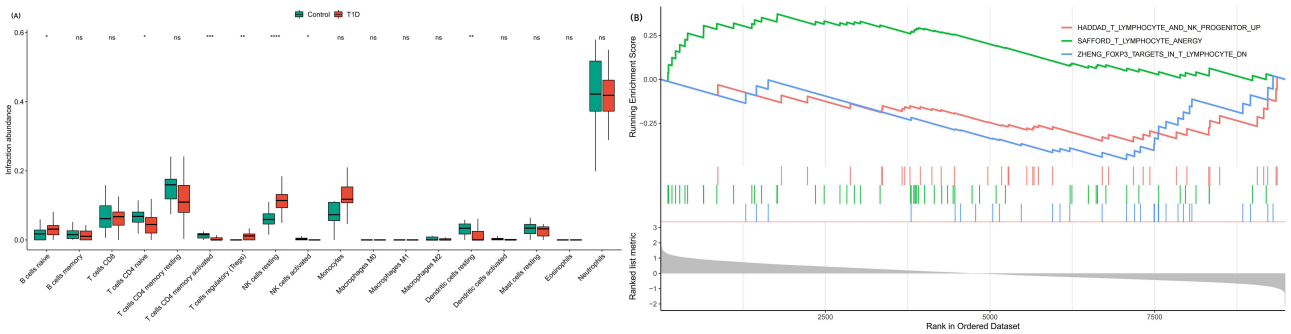

Figure 1 Results of immune cell infiltration and GSEA. (**A**): Immune cell infiltration analysis showed differences in proportion of infiltrating immune cells between T1D and control samples. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, and **** $p < 0.0001$ compared to control samples. (**B**) GSEA revealed differential pathways associated with T and NK cells between T1D and control samples.

GSEA of Pathways Associated with T and NK Cells

Based on the GSE123658 dataset, we analyzed the differential pathways between T1D and control samples and screened pathways associated with T and NK cells. The results showed that three pathways were enriched: HADDAD T LYMPHOCYTE AND NK PROGENITOR UP, SAFFORD T LYMPHOCYTE ANERGY, and ZHENG FOXP3 TARGETS IN T LYMPHOCYTE DN (Figure 1B).

DEG Screening

The procedure of this study was conducted based on the flow diagram (Figure 2). Based on the GSE123658 dataset, 1778 upregulated and 1044 downregulated DEGs were identified between the T1D and control samples (Figure 3A). Then, 32 DEIGs were obtained using Venn diagram analysis (Figure 3B).

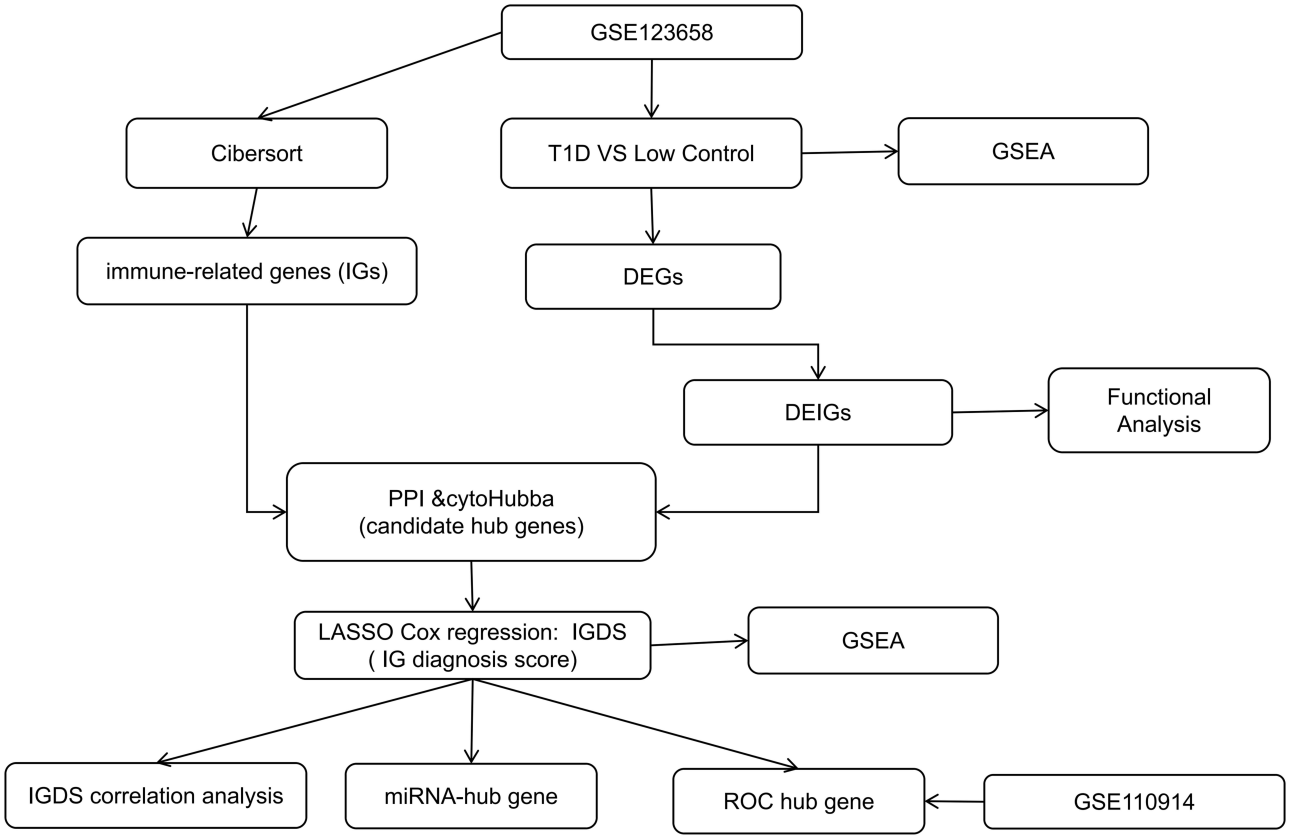

Figure 2 Flow diagram presenting the main plan and process of the study.

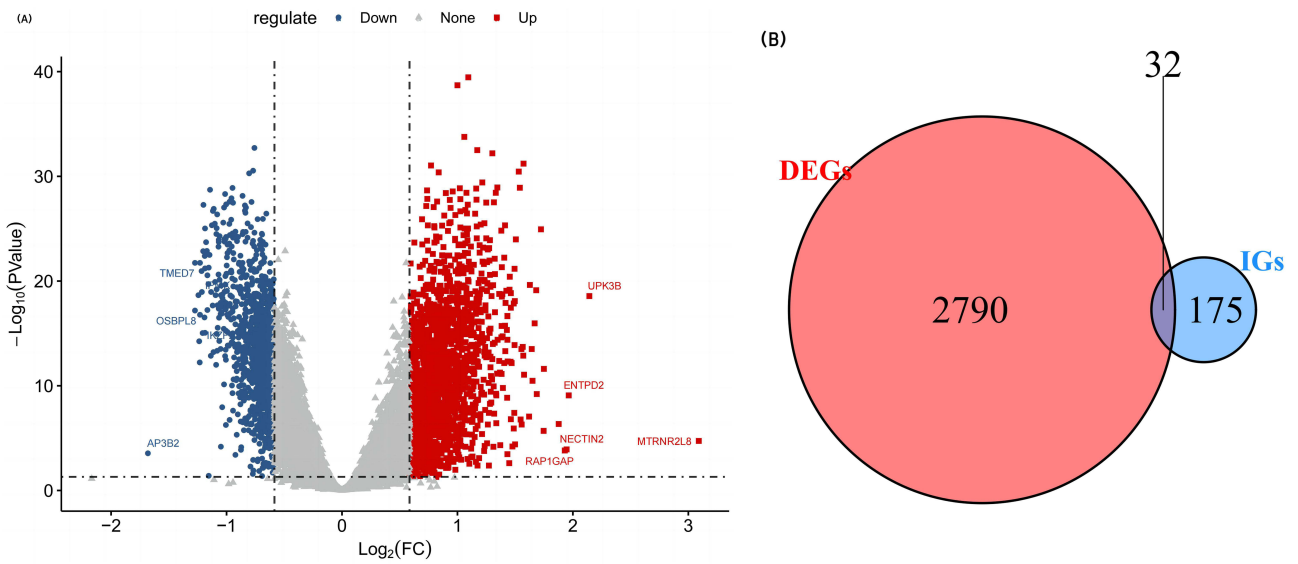

Figure 3 Identification of DEGs. (**A**) Volcano plot of DEGs between T1D and control samples. Red nodes indicate upregulated genes and green nodes indicate downregulated genes. (**B**) Venn diagram analysis of DEGs and IGs. T1D: type 1 diabetes.

PPI Network Analysis and Hub Gene Identification

Using the STRING database, a PPI network, including 16 DEIGs and 14 pairs, was constructed (Figure 4A). The top six genes were identified using four topological analysis algorithms (MNC, MCC, degree, and EPC). Through intersection analysis (Figure 4B), five overlapping genes were identified as hub DEIGs in the PPI network: NOTCH1, JAK3, TNFRSF4, CSF1, and nuclear factor of activated T cells 1 (NFATC1).

IGS Construction

Using LASSO Cox regression analysis, we further screened the optimized combination of DEIGs from the hub DEIGs in the PPI network to construct an IGS (Figure 5A and B). Four signature genes were identified for IGS construction: NOTCH1, JAK3, TNFRSF4, and CSF1. The IGS score was calculated for each sample.

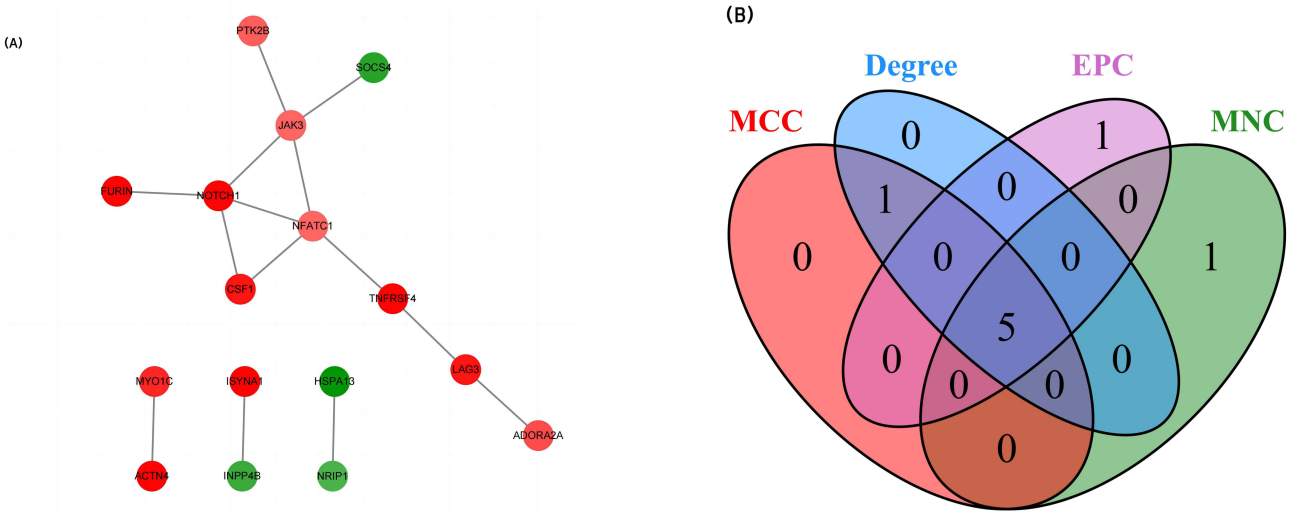

Figure 4 PPI analysis for DEIGs. (**A**) PPI network of DEIGs. Red nodes indicate upregulated genes and green nodes indicate downregulated genes. (**B**) Venn diagram analysis of top six genes screened from four topological analysis algorithms (MNC, MCC, degree, and EPC).

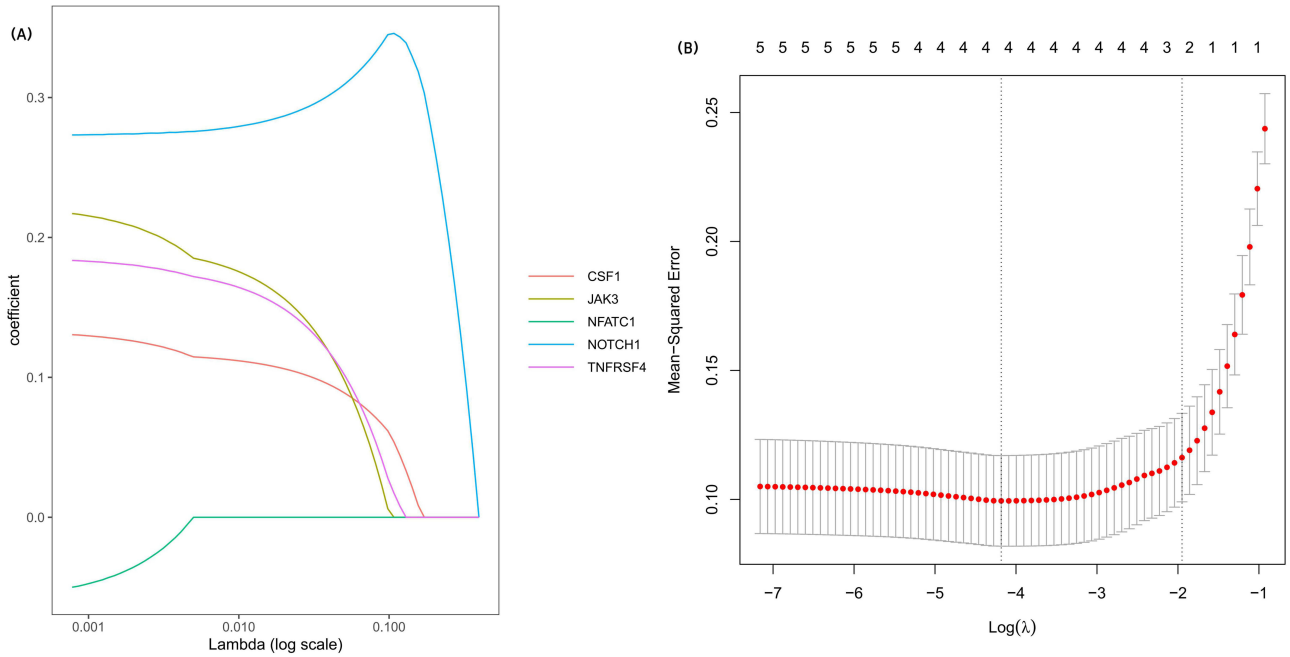

Figure 5 LASSO Cox regression analysis. (**A**) LASSO coefficient spectrum of hub DEIGs. (**B**) Selection of optimized lambda. LASSO.

GSEA of IGS-Associated Pathways

The samples were divided into high-risk and low-risk groups based on the median IGS score. GSEA showed the significantly enriched KEGG pathways (Figure 6A) and GO terms in the high-risk group (Figure 6B). Key KEGG pathways, such as the Toll-like receptor signaling pathway and ABC transporters, and GO terms, such as nervous system processes and inflammatory responses, were significantly activated in the high-risk group. In contrast, KEGG pathways such as ribosome and oxidative phosphorylation and GO terms such as mitochondrial electron transport of cytochrome C to oxygen and formation of the cytochrome translation initiation complex were remarkably suppressed in the high-risk group.

Validating the Differential Expression of Signature Genes

We first extracted the expression of the signature genes based on the training dataset GSE123658. Compared to the control samples, four signature genes, NOTCH1, JAK3, TNFRSF4, and CSF1, were all significantly upregulated in T1D samples ($p < 0.001$) (Figure 7A). We also analyzed the expression of these genes based on the validation dataset GSE110914. We found that only CSF1 expression was dramatically upregulated in T1D samples ($p < 0.001$) (Figure 7B), similar to the results obtained from the training dataset. Thus, CSF1 has been recognized as a hub signature gene in T1D.

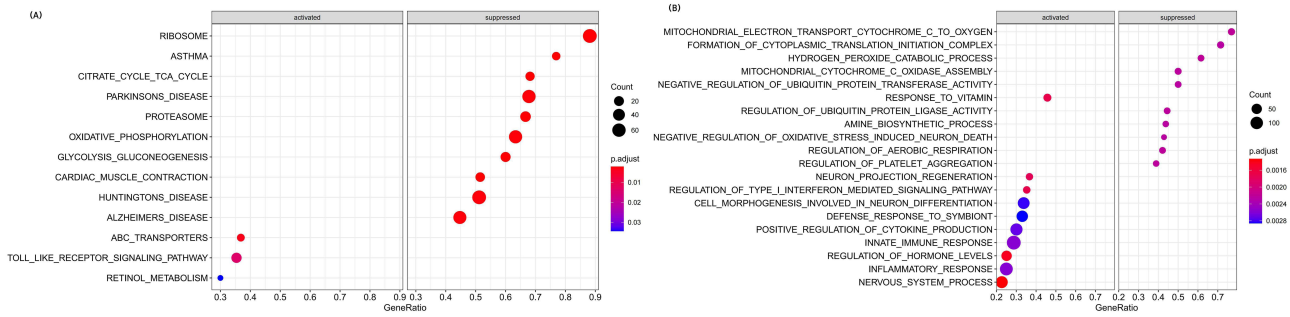

Figure 6 GSEA enrichment results. (**A**) KEGG pathway enrichment results. (**B**) GO enrichment results. Bubble size represents the number of genes enriched. Color, from blue to red, represents the reduction in p-value.

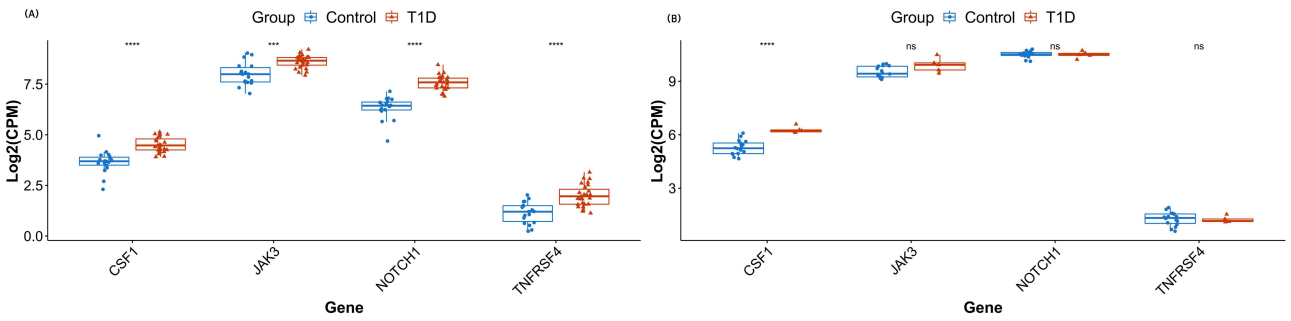

Figure 7 Validation of differential expression of signature genes. (**A**) Expression of signature genes based on training dataset GSE123658. (**B**) Expression of signature genes based on validation dataset GSE110914. *** p < 0.001 and **** p < 0.0001 compared to control. T1D: type 1 diabetes.

CSF1 Had High Diagnostic Efficacy for T1D

We further analyzed the diagnostic efficacy of CSF1 based on the gene expression data from the training and validation datasets. The results showed that the AUC value of CSF1 in the two datasets was > 0.7 (Figure 8A), indicating that CSF1 has a high diagnostic efficacy for T1D. Moreover, CSF1 expression significantly correlated with the IGS score (R = 0.82 and P = 4.1e-12) (Figure 8B).

Analysis of Upstream miRNAs of CSF1

Using the HMDD database, 25 miRNAs associated with T1D were identified. In addition, 112 miRNAs upstream of CSF1 were identified using the ENCORI database. Notably, among the miRNAs upstream of CSF1, only hsa-miR-326 is associated with T1D.

Clinical and Laboratory Parameters of Study Participants

The clinical and laboratory parameters of 60 T1DM patients and 22 healthy patients were seen in Table 2. Compared with Control group alone, body mass index (BMI), triglycerides (TG), white blood cell (WBC), glycated hemoglobin A1c (HbA1c) and CSF1 in T1DM group were increased($p<0.05$), while albumin (ALB) and miR-326 were decreased ($p<0.05$). The results showed that the relative expression levels of CSF1 were consistent with the previous findings. There was no significant difference in other indexes between the two groups.

Bivariate Correlations Between CSF1 Levels and Other Variables

In all study participants,CSF1 levels were positively correlated with BMI (r=0.652, p<0.001; Table 3; Figure 9a), TG (r=0.411, p=0.024; Table 3; Figure 9b), WBC (r=0.64, p<0.001; Table 3; Figure 9c), HbA1c (r=0.463, p=0.01; Table 3; Figure 9d) and cardiovascular and cerebrovascular diseases (CCVD) (r=0.817, p<0.001; Table 3; Figure 9e).

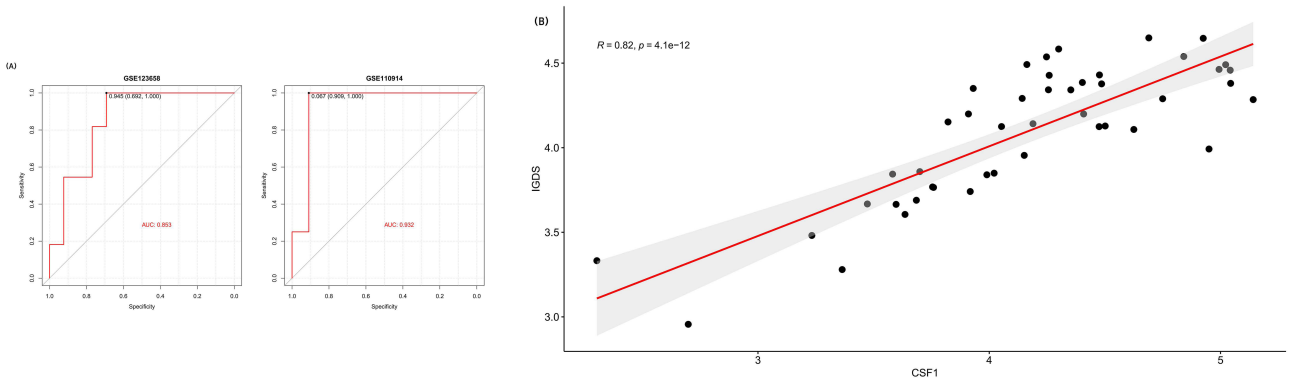

Figure 8 CSF1 had high diagnostic efficacy for T1D and was correlated with IGS score. (**A**) ROC curve showed diagnostic efficacy of CSF1 based on training dataset GSE123658 and validation dataset GSE110914. (**B**) Correlation between CSF1 expression and IGS score. T1D: type 1 diabetes.

Table 2 General Clinical and Laboratory Parameters of Study Participants

Variables	T1DM group	Control group	p value
n	60	22	
Age(years)[a]	55.07±16.06	54.27±8.17	0.877
T1D duration(years)[b]	16 (5,20)	0(0)	0.068
BMI (kg/m2)[a]	26.14±3.75	22.38±1.61	0.003
TG (mmol/L)[a]	4.60±0.83	1.10±0.33	<0.001
HDL-C(mmol/L)[a]	1.33±0.3	1.23±0.24	0.321
ALB(g/L)[b]	34.9(32.9,37.8)	40.7(39,41.7)	0.002
DBil(μmol/L)[b]	2.1(1.7,3)	3(2,3.3)	0.315
Scr(μmol/L)[b]	64.6(44.2,103.1)	62.8(59,73.1)	0.393
WBC (10^9/L)[b]	5.93(4.49,10.18)	4.71(4.26,5.1)	0.017
HbA1c(%)[a]	9.45±1.45	4.92±0.44	<0.001
CCVD	10(0.33)	0(0)	0.039
miR-326[b]	0.07(0.02,0.27)	0.58(0.31,1.31)	0.002
CSF1[b]	0.53(0.29,0.74)	0.21(0.14,0.26)	0.037

Notes: Enumeration data were compared using $\chi 2$ test. [a]Data normally distributed are shown as mean±SD. Independent sample *t*-test was performed. [b]Data with skewed distributions are shown as median (IQR, 25th–75th). Mann–Whitney *U*-test was performed.

Abbreviations: BMI, body mass index; TG, triglycerides; HDL-C, high-density lipoprotein cholesterol; ALB, albumin; DBil, Direct bilirubin; Scr, Serum creatinine; WBC, white blood cell; HbA1c, glycated hemoglobin A1c; CCVD, cardiovascular and cerebrovascular diseases; T1DM, type-1 diabetes mellitus.

Table 3 Bivariate Correlations Between CSF1 Levels and Other Variables

CSF1	*r*	*p* value
BMI	0.652**	<0.001
TG	0.411*	0.024
WBC	0.640**	<0.001
HbA1c	0.463*	0.01
CCVD	0.817**	<0.001

Notes: Spearman correlation analysis was used. *p* values <0.05 and <0.01 were considered significant. *p<0.05; **p<0.01.

Curve Area Under the Receiver Operating Characteristic Curve of CSF1 Levels in Patients with T1DM

In order to explore whether CSF1 with differential expression has diagnostic value in T1DM patients, we established the ROC curve. The results showed that CSF1 distinguished patients with T1DM with an AUC of 0.803 (95% confidence interval (CI) =0.646–0.959), sensitivity of 86.7%, and specificity of 81.8%. CSF1 levels were valuable in the diagnosis of T1DM patients (AUC,0.803; 95% CI, 0.646–0.959; $p < 0.001$; Figure 10).

Discussion

Immune cells and related IGs are essential regulators of T1D. However, the aberrant infiltration of immune cells and their regulatory mechanisms in T1D are yet to be completely understood. In the current study, we found differential infiltration proportions of T and NK cells between T1D and control samples and identified 207 IGs related to these immune cells. After differential expression, PPI, and LASSO Cox regression analyses, four signature DEIGs were identified for IGS

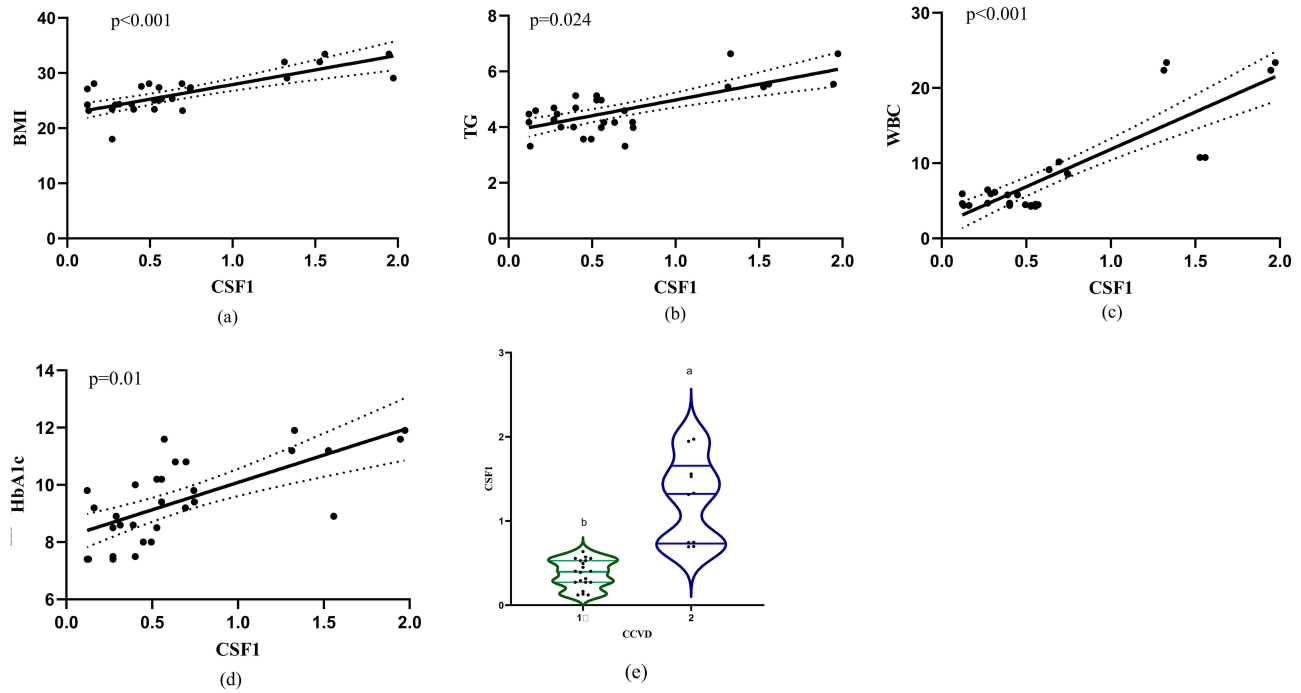

Figure 9 Bivariate correlations between CSF1 levels and other variables(**a**)correlation analysis of CSF1 with BMI; (**b**)correlation analysis of CSF1 with TG; (**c**)correlation analysis of CSF1 with WBC;(**d**)correlation analysis of CSF1 with HbA1c;(**e**)correlation analysis of CSF1 with CCVD, the 1 means no CCVD and the 2 represents CCVD, P<0.001.

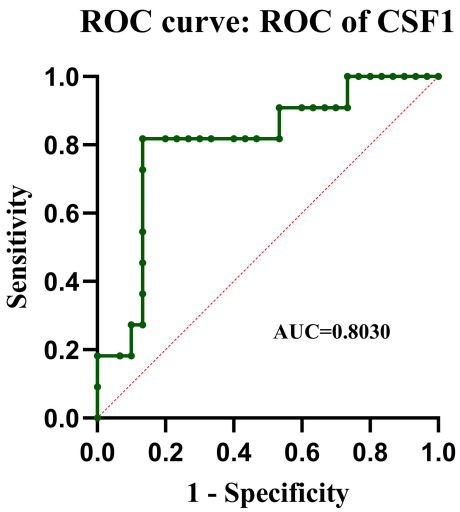

Figure 10 ROC curve of CSF1 levels in diagnosis of T1DM patients (AUC,0.803; 95% CI, 0.646–0.959; p < 0.001).

construction: NOTCH1, JAK3, TNFRSF4, and CSF1. Key pathways such as the Toll-like receptor signaling pathway were significantly activated in the high-risk group. Moreover, the upregulation of CSF1 in T1D samples was confirmed using the validation dataset GSE110914, and CSF1 showed high diagnostic efficacy for T1D. Furthermore, CSF1 is targeted by miR-326. These data reveal the molecular mechanisms underlying aberrant immune cell infiltration in T1D.

Increasing evidence has revealed that T cells play a key role in T1D induction.[27] $CD4^+$T cells are subdivided into Th1, Th2, Th17, and Tregs according to their cytokine secretion profiles. Th1-produced cytokines such as IL-2 and IFN-γ are key mediators of β cell autoreactivity; Th2-produced cytokines such as IL-4, IL-5, and IL-13 are stimulators of humoral immune responses; Th17 T cells contribute to the induction of autoimmune tissue injury; and Tregs are primary

controllers of effector T responses and peripheral tolerance. These $CD4^+$T cell subsets are involved in the T1D development.[28] Moreover, Tregs promote T1D progression in non-obese diabetic (NOD) mice by reinforcing NK cells.[29] NK cells have been implicated in the maintenance of immune responses and peripheral tolerance mechanisms, and their role in T1D pathogenesis has been demonstrated.[30,31] In the present study, we found that immune cells, such as naïve $CD4^+$T cells, activated CD4 memory activated, Tregs, resting NK cells, and activated NK cells, had significantly different infiltration proportions between T1D and control samples, suggesting a key role for T cells and NK cells in the pathogenesis of T1D.

To better understand the regulatory mechanisms of T and NK cells in T1D, we explored the related pathogenic genes by overlapping IGs and DEGs, resulting in 32 DEIGs. Using PPI and LASSO Cox regression analyses, four signature genes were identified for IGS construction: NOTCH1, JAK3, TNFRSF4, and CSF1. To further elucidate the potential mechanisms underlying IGS, we conducted GESA and found that key pathways, such as the toll-like receptor signaling pathway, were significantly activated in the high-risk group. The Toll-like receptor signaling pathway has been confirmed to participate in T1D development.[32,33] These data reveal that the Toll-like receptor signaling pathway may be a key mechanism mediating the role of IGS in T1D.

NOTCH1 is a member of the canonical Notch signaling pathway that regulates adult β cell proliferation and maturity.[34] Eom et al demonstrated that inhibition of NOTCH1 suppressed insulin secretion and reduced islet and β cell mass.[35] Wang et al revealed that NOTCH1 deficiency impairs cognitive function in T1D mice,[36] suggesting a potential role of NOTCH1 in T1D-related complications. JAK3 is a Janus kinase involved in the initiation of cytokine-triggered signaling events via the activation of STAT proteins. JAK3 has been reported to be a promising target for the treatment of T1D, and the JAK3 inhibitor JANEX-1 can prevent T1D in NOD mice.[37] TNFRSF4 (also known as OX40 or CD134) belongs to the TNF receptor superfamily and is expressed on various immune cells including activated T and NK cells. High TNFRSF4 expression is observed in $CD4^+CD25^{high}$ cells in children with T1D.[38] Bresson et al demonstrated that treatment with a TNFRSF4 agonist can ameliorate the antigen-specific prevention of T1D in NOD mice.[39] Considering the key functions of NOTCH1, JAK3, and TNFRSF4 in T1D development, we speculated that these genes may be vital mediators of T and NK cells in T1D. However, the expression of these genes did not show significant differences between T1D and control samples based on the validation dataset, which may be due to the small sample size (only five T1D samples in this dataset). Therefore, further studies with larger sample sizes are required to validate the differential expression and function of these genes in patients with T1D. In addition, CSF-1 is a cytokine that regulates the functions of multiple immune cells such as monocytes and macrophages.[40] CSF-1 expression is related to the inflammatory status of periodontal tissues and type 2 diabetes.[41] However, the role of CSF-1 in T1D has not been reported. Despite these, CSF-1 was found to be upregulated in T1D samples based on both training and validation datasets and had a high diagnostic efficacy for T1D. Thus, we speculated that CSF-1 may also be a key mediator of T and NK cells in T1D, and has the potential as a diagnostic biomarker for T1D.

To further elucidate the regulatory mechanism of CSF1 in T1D, we predicted the miRNAs upstream of CSF1. Notably, among the miRNAs upstream of CSF1, only hsa-miR-326 is associated with T1D. This study suggested that CSF1 and miR-326 play an important role in T1DM patients and CSF1 level may be a novel biomarker to distinguish T1DM patients.

Previous study highlights the increased prevalence of overweight and obesity in adults with T1DM.[42] In a US-based multicentre study, HbA1c was higher in obese adults when compared with normal BMI adults.[43] Overweight and obese T1DM patients had an increased prevalence of complications when compared to that of normal BMI T1DM patients. The prevalence of hypertension was 59% among T1DM patients with overweight/obesity whereas 27.5% in normal-weight T1DM patients.[42] Hypertension is considered to be an important risk for the development of microvascular and macrovascular complications.[42–44]Hematopoietic CSF1 deficiency causes smaller atherosclerotic lesions.[45] Consistent with previous studies, our study showed that BMI was significantly higher in the T1DM group compared with the control group ($p < 0.05$). HbA 1 c and CCVD were significantly higher in overweight and obese T1DM patients than in controls ($p < 0.05$).Macrophage colony-stimulating factor (CSF1) is the primary growth factor required for the control of monocyte and macrophage differentiation, survival, proliferation and renewal.[46] Highly significant correlations were observed between the expression of CSF1 mRNA in the omentum (OM) in overweight T1DM patients and the number of

infiltrating macrophages. Obese people have higher levels of macrophages, mainly in OM.OM macrophage counts correlated with BMI and waist circumference and tended to correlate with the number of features comprising the metabolic syndrome.[47] Our findings showed that CSF1 was positively correlated with BMI, TG, WBC in T1DM patients, which was consistent with previous studies.

Limitations of the study include that are triggered by different factors may not be identifiable due to the constraints of our relatively small sample size. Another restricting factor was the selection criteria used in this study, which led to an unequal distribution of participants in the study groups, as both the cause and effect were measured at a single point in time. Since this study is cross-sectional, it is challenging to establish a clear understanding of temporality and biological plausibility. In light of these limitations, further studies are needed to confirm the miR-326-CSF1 relationship and clarify its role in T1D development. Consequently, we recommend that future research endeavors more comprehensive analytical studies with larger sample sizes to address these limitations.

In conclusion, abnormal infiltration of T and NK cells may play a key role in T1D development by targeting key IGs, such as NOTCH1, JAK3, TNFRSF4, and CSF1. Thus, CSF1 can be developed as a novel diagnostic biomarker for T1D. Our findings provide novel insights into immunoregulatory therapy for T1D.

Data Sharing Statement

The datasets generated and analyzed during the present study are available from the corresponding author on reasonable request.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by the 2021 Lianyungang Aging Health Research Project (No. L202105), the 2022 Jiangsu Aging Health Research Project (No. LKM2022065), the 2022 Lianyungang Health Science and Technology Project (No. 202206),the 2023 Lianyungang City sixth “521” project research funding project (No.LYG06521202372), and 2023 soft research project (No. Lkxyb23130).

Disclosure

The authors declare that they have no competing interests in this work.

References

1. Katsarou A, Gudbjörnsdottir S, Rawshani A. et al. Type 1 diabetes mellitus. *Nature Reviews Disease Primers*. 2017;3(1):1–17. doi:10.1038/nrdp.2017.16
2. Cano-Cano F, Gómez-Jaramillo L, Ramos-García P, et al. IL-1β implications in type 1 diabetes mellitus progression: systematic review and meta-analysis. *J Clin Med*. 2022;11(5):1303. doi:10.3390/jcm11051303
3. Gregory GA, Robinson TIG, Linklater SE, et al. Global incidence, prevalence, and mortality of type 1 diabetes in 2021 with projection to 2040: a modelling study. *Lancet Diabetes Endocrinol*. 2022;10(10):741–760. doi:10.1016/S2213-8587(22)00218-2
4. Rawshani A, Sattar N, Franzén S, et al. Excess mortality and cardiovascular disease in young adults with type 1 diabetes in relation to age at onset: a nationwide, register-based cohort study. *Lancet*. 2018;392(10146):477–486. doi:10.1016/S0140-6736(18)31506-X
5. Foster NC, Beck RW, Miller KM, et al. State of type 1 diabetes management and outcomes from the T1D exchange in 2016–2018. *Diabetes Technol Ther*. 2019;21(2):66–72. doi:10.1089/dia.2018.0384
6. Sims EK, Besser REJ, Dayan C, et al. Screening for type 1 diabetes in the general population: a status report and perspective. *Diabetes*. 2022;71(4):610–623. doi:10.2337/dbi20-0054
7. Scherm MG, et al. Beta cell and immune cell interactions in autoimmune type 1 diabetes: how they meet and talk to each other. *Mol Metabol*. 2022;1:64.
8. Zhang H, Kong H, Zeng X, et al. Subsets of regulatory T cells and their roles in allergy. *J Transl Med*. 2014;12(1):125. doi:10.1186/1479-5876-12-125

9. Vallianou NG, Kopchick JJ, et al. Diabetes type 1: can it be treated as an autoimmune disorder? *Rev Endocr Metab Disord*. 2021;22(1):1–18. doi:10.1007/s11154-021-09629-1
10. Marek-Trzonkowska N, Myśliwiec M, Iwaszkiewicz-Grześ D, et al. Factors affecting long-term efficacy of T regulatory cell-based therapy in type 1 diabetes. *J Transl Med*. 2016;14(1):1–11. doi:10.1186/s12967-016-1090-7
11. Babon JAB, DeNicola ME, Blodgett DM, et al. Analysis of self-antigen specificity of islet-infiltrating T cells from human donors with type 1 diabetes. *Nature Med*. 2016;22(12):1482–1487. doi:10.1038/nm.4203
12. Mandrup-Poulsen T, Pickersgill L, Donath MY. Blockade of interleukin 1 in type 1 diabetes mellitus. *Nat Rev Endocrinol*. 2010;6(3):158–166. doi:10.1038/nrendo.2009.271
13. Lin J, Lu Y, Wang B, et al. Analysis of immune cell components and immune-related gene expression profiles in peripheral blood of patients with type 1 diabetes mellitus. *J Transl Med*. 2021;19(1):1–16. doi:10.1186/s12967-021-02991-3
14. Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 2012;41(D1):D991–D995. doi:10.1093/nar/gks1193
15. Chen B, et al. Profiling tumor infiltrating immune cells with CIBERSORT. *Cancer Sys Bio*. 2018;1:243–259.
16. Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739–1740. doi:10.1093/bioinformatics/btr260
17. Liberzon A, Birger C, Thorvaldsdóttir H, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417–425. doi:10.1016/j.cels.2015.12.004
18. Yu G, Wang L-G, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–287. doi:10.1089/omi.2011.0118
19. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *bioinformatics*. 2010;26(1):139–140. doi:10.1093/bioinformatics/btp616
20. Guo K, McGregor B, Hur J. VennDetail: a package for visualization and extract details. 2021;1:1.
21. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49(D1):D605–D612. doi:10.1093/nar/gkaa1074
22. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–2504. doi:10.1101/gr.1239303
23. Chin C-H, Chen S-H, Wu -H-H, et al. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014;8(4):1–7. doi:10.1186/1752-0509-8-S4-S11
24. Blanco JL, Porto-Pazos AB, Pazos A, et al. Prediction of high anti-angiogenic activity peptides in silico using a generalized linear model and feature selection. *Sci Rep*. 2018;8(1):15688. doi:10.1038/s41598-018-33911-z
25. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf*. 2011;12(1):1–8. doi:10.1186/1471-2105-12-77
26. Shen Y, et al. MicroRNA–Disease Network Analysis Repurposes Methotrexate for the Treatment of Abdominal Aortic Aneurysm in Mice. *Geno Pro Bioin*. 2022;1:1.
27. Szablewski L. Role of immune system in type 1 diabetes mellitus pathogenesis. *Int Immunopharmacol*. 2014;22(1):182–191. doi:10.1016/j.intimp.2014.06.033
28. Shao S, et al. Th17 cells in type 1 diabetes. *Cellular Immun*. 2012;280(1):16–21. doi:10.1016/j.cellimm.2012.11.001
29. Sitrin J, Ring A, Garcia KC, et al. Regulatory T cells control NK cells in an insulitic lesion by depriving them of IL-2. *J Exp Med*. 2013;210(6):1153–1165. doi:10.1084/jem.20122248
30. Gianchecchi E, Delfino DV, Fierabracci A. Natural killer cells: potential biomarkers and therapeutic target in autoimmune diseases? *Front Immunol*. 2021;12:616853. doi:10.3389/fimmu.2021.616853
31. Kucuksezer UC, et al. The role of natural killer cells in autoimmune diseases. *Front Immunol*. 2021;12:622306.
32. Lien E, Zipris D. The role of Toll-like receptor pathways in the mechanism of type 1 diabetes. *Curr Molec Med*. 2009;9(1):52–68. doi:10.2174/156652409787314453
33. Meyers AJ, Shah RR, Gottlieb PA, et al. Altered Toll-like receptor signaling pathways in human type 1 diabetes. *J Mol Med*. 2010;88(12):1221–1231. doi:10.1007/s00109-010-0666-6
34. Bartolome A, Zhu C, Sussel L, et al. Notch signaling dynamically regulates adult β cell proliferation and maturity. *J Clin Invest*. 2019;129(1):268–280. doi:10.1172/JCI98098
35. Eom YS, Gwon A-R, Kwak KM, et al. Notch1 has an important role in β-cell mass determination and development of diabetes. *Diabetes Metab J*. 2021;45(1):86–96. doi:10.4093/dmj.2019.0160
36. Wang Y, Wang S, Zhang W, et al. Notch1 participates in the activation of autophagy in the hippocampus of type I diabetic mice. *Neurochem Int*. 2021;150:105156. doi:10.1016/j.neuint.2021.105156
37. Cetkovic-Cvrlje M, Dragt AL, Vassilev A, et al. Targeting JAK3 with JANEX-1 for prevention of autoimmune type 1 diabetes in NOD mice. *Clin Immunol*. 2003;106(3):213–225. doi:10.1016/S1521-6616(02)00049-9
38. Szypowska A, Stelmaszczyk-Emmel A, Demkow U, et al. High expression of OX40 (CD134) and 4-1BB (CD137) molecules on CD4+ CD25high cells in children with type 1 diabetes. *Adv Med Sci*. 2014;59(1):39–43. doi:10.1016/j.advms.2013.07.003
39. Bresson D, Fousteri G, Manenkova Y, et al. Antigen-specific prevention of type 1 diabetes in NOD mice is ameliorated by OX40 agonist treatment. *J Autoimmun*. 2011;37(4):342–351. doi:10.1016/j.jaut.2011.10.001
40. Pixley FJ, Stanley ER. CSF-1 regulation of the wandering macrophage: complexity in action. *Trends Cell Biol*. 2004;14(11):628–638. doi:10.1016/j.tcb.2004.09.016
41. Dikilitaş A, et al. Gingival crevicular fluid CSF-1 and IL-34 levels in patients with stage III grade C periodontitis and uncontrolled type 2 diabetes mellitus. *J Period Imp Sci*. 2021;1:52.
42. Bhagadurshah RR, Eagappan S, Kasthuri Santharam R, et al. The Impact of Body Mass Index, Residual Beta Cell Function and Estimated Glucose Disposal Rate on the Development of Double Diabetes and Microvascular Complications in Patients With Type 1 Diabetes Mellitus. *Cureus*. 2023;15(11):e48979. doi:10.7759/cureus.48979

43. NOOR N, et al. 961-P: the Effect of Obesity on HbA1c among Adults with Type 1 Diabetes: a U.S. *Based Multi Study Dia*. 2022;71 (Supplement_1):1.
44. Purnell JQ, et al. Effect of excessive weight gain with intensive therapy of type 1 diabetes on lipid levels and blood pressure: results from the DCCT. *DIAB Con Comp Tril Jama*. 1998;280(2):140–146.
45. McAlpine CS, Kiss MG, Rattik S, et al. Sleep modulates haematopoiesis and protects against atherosclerosis. *Nature*. 2019;566(7744):383–387. doi:10.1038/s41586-019-0948-2
46. Sehgal A, Irvine KM, Hume DA. Functions of macrophage colony-stimulating factor (CSF1) in development, homeostasis, and tissue repair. *Semin Immunol*. 2021;54:101509. doi:10.1016/j.smim.2021.101509
47. Harman-Boehm I, Bluher M, Redel H, et al. Macrophage Infiltration into Omental Versus Subcutaneous Fat across Different Populations: effect of Regional Adiposity and the Comorbidities of Obesity. *J Clin Endocrinol Metab*. 2007;92(6):2240–2247. doi:10.1210/jc.2006-1811