Role of SIRPG gene in type 1 diabetes and lichen planus

Ji Wu, MD[a] , Honglei Zhao, MD[a],*

Abstract

Type 1 diabetes (T1D) is a form of diabetes caused by pancreatic β-cell destruction and absolute insulin deficiency. Lichen planus (LP) is an idiopathic inflammatory skin disease of unclear etiology. The role of SIRPG gene dysregulation in T1D and LP remains unclear. Mendelian randomization (MR) using matched samples was employed to study causal relationship between T1D and increased risk of LP. T1D-related single nucleotide polymorphism identification was conducted. Datasets GSE156035 for T1D and GSE52130 for LP were obtained from gene expression omnibus. Differentially expressed genes were identified, analyses included weighted gene co-expression network analysis, functional enrichment, gene set enrichment analysis, and protein–protein interaction network construction and analysis. Heatmaps of gene expression levels were generated. Comparative toxicogenomics database was used to identify diseases most relevant to core genes. Inverse variance weighted, MR-Egger, weighted median methods estimated genetic predisposition between T1D and LP, showing consistent positive correlations using both weighted median and inverse variance weighted methods. Horizontal pleiotropy analysis with MR-Egger intercept indicated no evidence of significant directional pleiotropy ($P = .70645$) for LP. There was no evidence of directional pleiotropy effects between T1D and LP. One hundred eighteen differentially expressed genes were identified. In biological processes, they were mainly enriched in apoptosis, inflammatory response, insulin receptor signaling pathway, glucose metabolism. In cellular components, enrichment was observed in mediator complex and replication fork. In molecular function, they were concentrated in leukotriene receptor activity and helicase activity. Kyoto Encyclopedia of Genes and Genomes analysis revealed enrichment in metabolic pathways, PI3K-Akt signaling pathway, cell cycle, p53 signaling pathway, AGE-RAGE signaling pathway in diabetic complications. Weighted gene co-expression network analysis with a soft threshold power of 4. SIRPG showing high expression in both T1D and LP samples. There is a positive causal relationship between T1D and LP. Comparative toxicogenomics database analysis revealed associations of core genes with metabolic syndrome, lipid metabolism disorders, cardiovascular diseases, immune system diseases, peripheral neuropathic pain, and inflammation. SIRPG is highly expressed in both T1D and LP, providing a new insight into the pathogenesis of T1D and LP.

Abbreviations: CTD = comparative toxicogenomics database, DEGs = differentially expressed genes, GO = gene ontology, GSEA = gene set enrichment analysis, IVW = inverse variance weighted, KEGG = Kyoto Encyclopedia of Genes and Genomes, LP = lichen planus, MR = Mendelian randomization, PPI = protein–protein interaction, STRING = Search Tool for the Retrieval of Interacting Genes, T1D = type 1 diabetes, WGCNA = weighted gene co-expression network analysis.

Keywords: bioinformatics, lichen planus, Mendelian randomization, SIRPG, type 1 diabetes

1. Introduction

Type 1 diabetes (T1D) is an autoimmune disease characterized by insufficient or complete cessation of insulin production.[1] The incidence of T1D varies significantly across different ethnicities and geographic regions. In recent years, the global incidence of T1D has been rising, especially in developed countries.[2] T1D is most commonly diagnosed in children and adolescents, with the typical onset age ranging from 4 to 14 years, hence sometimes referred to as "juvenile diabetes."[3] However, it can occur at any age, including adulthood. The incidence rates of T1D in males and females are similar, though in some regions and age groups, males may have a slightly higher prevalence than females.[4] Early symptoms of T1D are usually pronounced, including abnormal blood glucose levels and symptoms of ketoacidosis. Due to the inability to regulate blood sugar levels automatically, T1D patients face a higher risk of acute complications. Poor long-term glycemic control may also lead to chronic complications such as cardiovascular disease, retinopathy, nephropathy, and neuropathy.[5] Factors such as viral infections, breastfeeding

ᵃ Department of Dermatology, Beilun People's Hospital, Ningbo, Zhejiang, China.

* Correspondence: Honglei Zhao, Department of Dermatology, Beilun People's Hospital, District of Beilun, Ningbo, Zhejiang 315800, China (e-mail: zhaohonglei88@sina.com).

duration, dietary habits, and lifestyle may be associated with the onset of T1D.

Lichen planus (LP) is a common chronic inflammatory skin disease primarily affecting the skin and mucous membranes.[6] It is characterized by flat-topped, purple-red or purplish-brown papules with clear boundaries, often accompanied by itching. LP may appear on different areas of the skin, and in the oral mucosa, it can present as white striae or a reticular pattern, occasionally with ulcers or erythema.[7] The prevalence of LP in the general population ranges from approximately 0.1% to 4%, but the specific incidence varies across different regions and populations.[8] LP primarily affects middle-aged adults, with onset typically occurring between the ages of 30 and 60, and is relatively uncommon in children and the elderly. While it can affect any gender, females may be slightly more affected than males. LP occurs worldwide, and its prevalence may differ across various ethnic and racial groups. Treatment for LP focuses on relieving symptoms and preventing complications, often involving topical or oral medications as needed, with specific treatment plans determined by physicians based on the patient's condition.[9] Although the exact cause of LP remains unclear, it is associated with abnormal immune responses, drug reactions, and certain infections.

Mendelian randomization (MR) is a statistical method that uses genetic data to evaluate causal relationships.[10] Based on Mendel's laws of inheritance, it uses genetic variants related to an exposure factor as instrumental variables to infer whether the exposure has a causal effect on an outcome in observational studies. This method can avoid confounding factors and reverse causality issues, making it increasingly popular in epidemiological research. MR has various applications in medical research, particularly in studying disease etiology, validating therapeutic targets, and identifying new risk factors.[11]

Bioinformatics involves the use of computational methods and tools to process and analyze biological data, helping scientists extract useful information from genome data, protein sequences, and other biological data.[12] By developing algorithms, software tools, and data analysis methods, bioinformatics enables scientists to derive meaningful insights from large-scale biological data. It plays a critical role not only in basic research areas such as gene function prediction, protein structure analysis, and molecular evolution studies but also in applied fields like personalized medicine, disease diagnosis, and breeding strategies. Through bioinformatics, researchers can
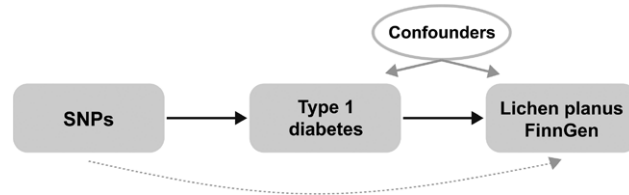

Figure 1. Flow chart of Mendelian randomization analysis.

Table 1

List of genetic instruments for T1D and log odds ratios of LP risk by each instrumental SNPs (GWAS significance with $P < 5 \times 10^8$ and linkage disequilibrium threshold with $R^2 < 0.001$).

No.	SNP	Gene	Chr.	EA	OA	EAF.T1D	EAF.LP	T1D β (SE)	LP β (SE)
1.	rs10944479	BACH2	6	A	G	NA	0.1197	0.195587 (0.0285564)	0.0541 (0.0524)
2.	rs11938508	LOC102724158	4	A	T	NA	0.5079	0.144032 (0.0230277)	0.0561 (0.0341)
3.	rs12722496	IL2RA	10	G	A	NA	0.06469	−0.407803 (0.0400993)	−0.0292 (0.0685)
4.	rs12927355	CLEC16A	16	T	C	NA	0.3035	−0.199659 (0.0235221)	0.028 (0.0366)
5.	rs12982646	MADCAM1	19	A	G	NA	0.3626	0.185982 (0.0281312)	−0.0516 (0.0356)
6.	rs13415465	–	2	G	T	NA	0.3158	−0.144031 (0.0229026)	0.0102 (0.0363)
7.	rs144451833	PRR11	17	T	C	NA	0.03313	−0.486288 (0.0710328)	−0.1085 (0.0953)
8.	rs1487916	PDGFD	11	C	T	NA	0.4118	0.131927 (0.0228193)	−0.0457 (0.0343)
9.	rs151233	APOBR	16	T	C	NA	0.1236	0.222076 (0.0336273)	0.0904 (0.0514)
10.	rs1790946	CD226	18	T	C	NA	0.537	−0.122841 (0.0221911)	−0.0229 (0.0337)
11.	rs2111485	LOC105373724	2	G	A	NA	0.5821	0.155003 (0.0229948)	0.1022 (0.0341)
12.	rs2281808	SIRPG	20	C	T	NA	0.702	0.13141 (0.0237435)	0.072 (0.0369)
13.	rs2289702	CTSH	15	T	C	NA	0.1267	−0.261465 (0.0398964)	−0.0955 (0.0509)
14.	rs231779	CTLA4	2	T	C	NA	0.5033	0.187758 (0.023269)	−0.0679 (0.0337)
15.	rs28746962	–	6	A	C	NA	0.1754	−1.29171 (0.0344348)	−0.1414 (0.0455)
16.	rs3024505	–	1	A	G	NA	0.1572	−0.176552 (0.0322297)	−0.068 (0.0463)
17.	rs4759229	ERBB3	12	G	A	NA	0.6722	−0.235657 (0.0230788)	−0.1063 (0.0359)
18.	rs4820827	HORMAD2	22	C	T	NA	0.6825	−0.167344 (0.0222857)	0.0601 (0.0363)
19.	rs55993634	–	16	G	C	NA	0.08743	0.258682 (0.0379206)	−0.0362 (0.0597)
20.	rs592625	HLA-DOA	6	C	T	NA	0.1109	−0.17977 (0.0314301)	0.0542 (0.0554)
21.	rs62397561	–	6	T	C	NA	0.3133	−0.215666 (0.0271179)	−0.0082 (0.0362)
22.	rs6679677	PHTF1	1	A	C	NA	0.1472	0.666685 (0.0370344)	−0.0016 (0.0475)
23.	rs67140765	MICA	6	T	G	NA	0.05644	−0.587305 (0.0539422)	−0.2231 (0.0758)
24.	rs689	INS	11	T	A	NA	0.7909	0.813401 (0.0328454)	0.0672 (0.0417)
25.	rs7024686	GLIS3	9	C	G	NA	0.4822	−0.131302 (0.0221532)	−0.0892 (0.0337)
26.	rs7237497	–	18	C	T	NA	0.859	−0.253723 (0.0300717)	−0.1375 (0.0485)
27.	rs7310615	SH2B3	12	G	C	NA	0.5866	−0.26601 (0.0230222)	−0.0896 (0.0342)
28.	rs78037977	–	1	G	A	NA	0.1028	−0.227462 (0.0361848)	0.0769 (0.0554)
29.	rs80054410	UBASH3A	21	C	T	NA	0.2823	0.164331 (0.0224537)	0.1142 (0.0375)
30.	rs9266775	LOC124901300	6	G	A	NA	0.1422	0.867307 (0.030985)	−0.0479 (0.0504)
31.	rs927292	ZFP36L1	14	G	C	NA	0.6707	0.149268 (0.0240294)	−0.0972 (0.0357)
32.	rs9468204	–	6	T	C	NA	0.1933	0.31845 (0.0245153)	0.0988 (0.0443)
33.	rs9468618	–	6	T	C	NA	0.05585	−0.214578 (0.0366778)	−0.2484 (0.0741)
34.	rs9600451	–	13	A	G	NA	0.413	0.133753 (0.022926)	−0.0053 (0.0342)
35.	rs9911533	–	17	T	C	NA	0.6067	0.126685 (0.0229105)	−0.0549 (0.0344)

Chr = indicates chromosome, EA = effect allele, EAF = effect allele frequency, LP = lichen planus, OA = other allele, SNPs = single nucleotide polymorphisms, T1D = type 1 diabetes.

better understand the complexity of biological systems, providing strong support for advances in life sciences.[13,14]

The relationship between SIRPG and both T1D and LP remains unclear. This study aims to explore core genes between T1D and LP using bioinformatics techniques, performing enrichment and pathway analysis. Public datasets will be used to validate the significant role of SIRPG in T1D and LP.

2. Methods

2.1. Type 1 diabetes and lichen planus Mendelian randomization analysis

2.1.1. Overall study design. The summary data for this study were derived from published research that received approval from its institutional review board. Therefore, no further sanction was required. A two-sample MR study was conducted to investigate the causal relationship between T1D and LP risk increase.

2.1.2. Identification of SNPs associated with T1D. Thirty-five single nucleotide polymorphisms (SNPs) associated with T1D were obtained from the latest GWAS database, identified from 7467 T1D cases and 10,218 control participants. To identify independently significant SNPs within the genomic region, the following criteria were applied: (1) SNPs showing significant association with T1D at the genome-wide level ($P < 5 \times 10^{-8}$), (2) exclusion of SNPs in linkage disequilibrium with other SNPs ($r^2 < 0.001$, window size = 10,000 kb), and (3) SNPs with F statistics (F-statistic = $(\beta/SE)^2$) < 10 were excluded to avoid bias in instrumental variables in MR.

2.1.3. Study results: LP. Data on LP were sourced from the FinnGen project, available at https://gwas.mrcieu.ac.uk/datasets/finn-b-L12_LICHENPLANUS/. This included 1865 LP cases and 212,242 control participants. The study obtained approval from its institutional review board, and all participants in the original study provided informed consent.

2.1.4. Statistical analysis. The "TwoSampleMR" package in R software (version 0.5.8) was used for two-sample MR analysis (R version 4.2.2, http://www.r-project.org). Bonferroni-corrected analyses with two-tailed $P < .05$ were used in all statistical tests. The inverse variance weighted (IVW) method was employed to estimate the causal relationship between T1D and LP, with sensitivity analyses using the weighted median method and MR-Egger method. The MR-Egger intercept estimated pleiotropic effects and whether the intercept significantly differed from zero statistically. These methods were utilized to investigate the causal relationship between T1D and LP.

2.2. Differential gene bioinformatics analysis

2.2.1. T1D and LP datasets. In this study, the T1D dataset GSE156035 and the LP dataset GSE52130 profiles were downloaded from the gene expression omnibus database (http://www.ncbi.nlm.nih.gov/geo/) generated from GPL20844 and GPL10558. GSE156035 included 20 T1D and 20 normal samples, while GSE52130 included 11 LP and 12 normal samples, used for identifying differentially expressed genes (DEGs) between T1D and LP.

2.2.2. Selection of DEGs. Log2 transformation was applied to the GSE156035 and GSE52130 datasets. The lmFit function was used for multivariate linear regression, and empirical Bayesian moderation was applied to compute moderated t-statistics, moderated F-statistics, and the log odds of differential expression approaching a common value, resulting in volcano plots. The
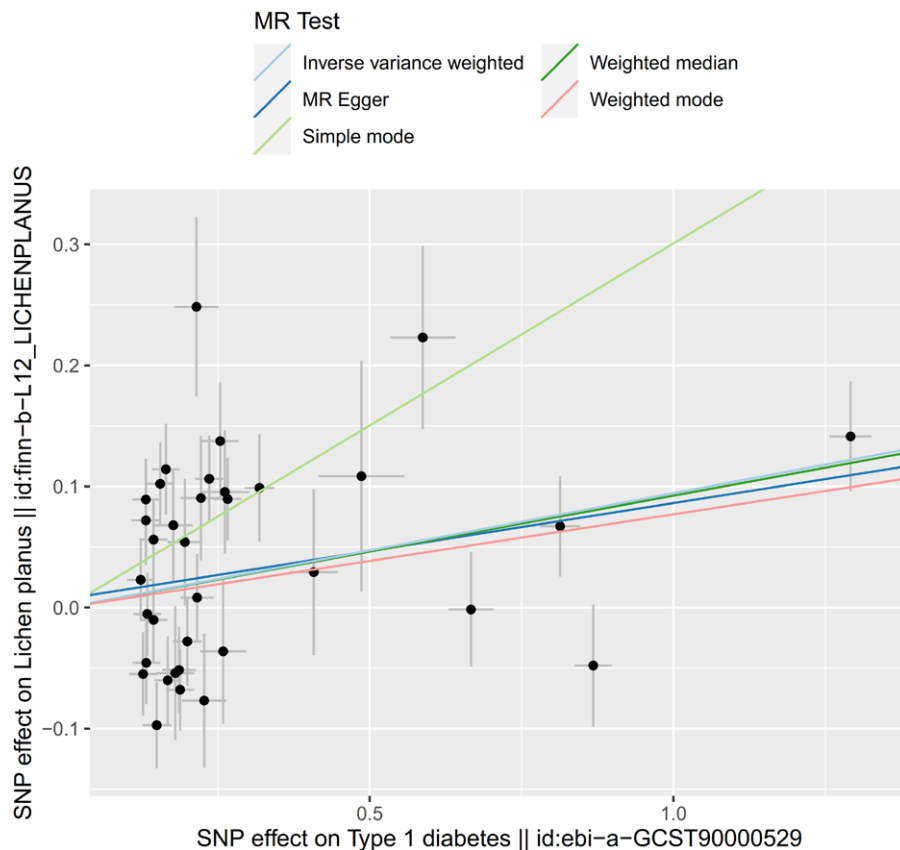

Figure 2. Genetic instrument variables for T1D. T1D = type 1 diabetes.

intersection of DEGs from GSE156035 and GSE52130 datasets was obtained.

2.2.3. Weighted gene co-expression network analysis.
Using the gene expression matrix of GSE52130, median absolute deviation was computed for each gene, removing the least median absolute deviation 50% of genes. The goodSamplesGenes method from the WGCNA package in R was used to eliminate outlier genes and samples, constructing a scale-free co-expression network using the WGCNA package. Gene modules were categorized based on topological overlap measure-based average

linkage hierarchical clustering, with a minimum module size of 30 genes. A sensitivity of 3 was set. Module characteristics were analyzed further by calculating module eigengenes, selecting a cutoff for the module dendrogram, and merging modules with a distance <0.25, where the gray module was considered unassignable to any specific module.

2.2.4. Functional enrichment analysis.
Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were conducted to assess gene functions and biological pathways using clusterProfiler (version 3.14.3) in R software
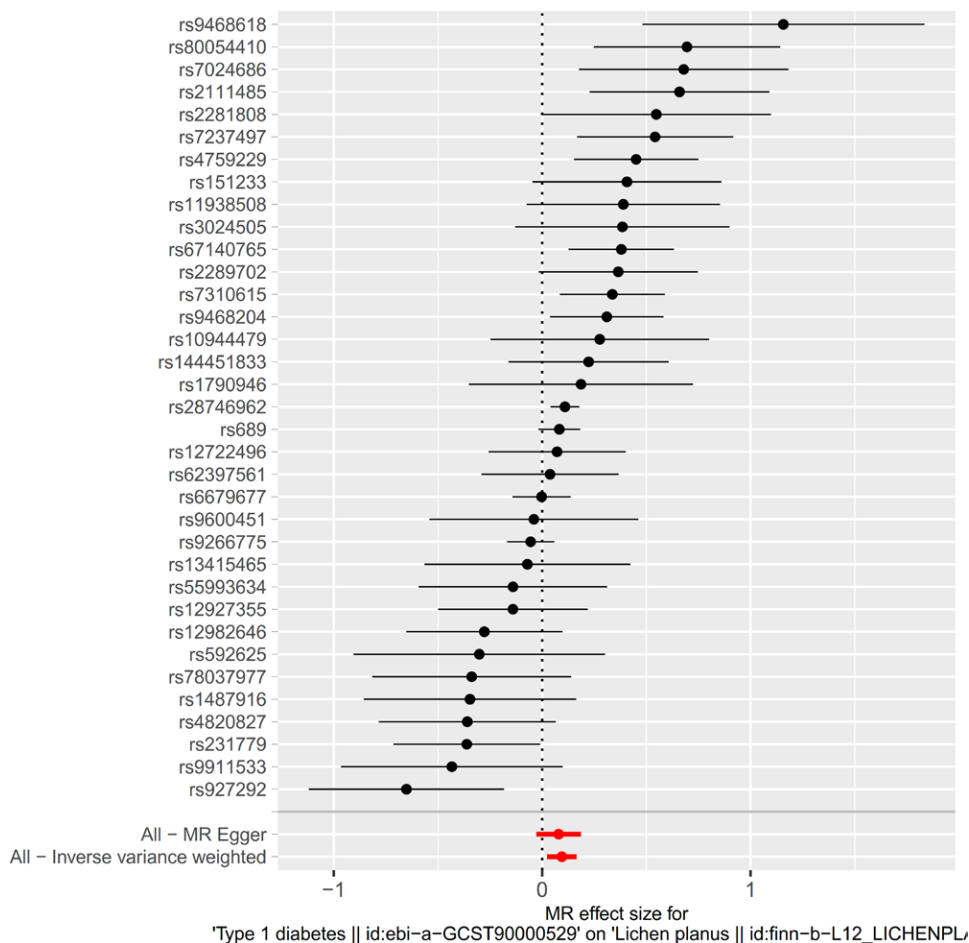

Figure 3. LP risk associations for each genetic variant. LP = lichen planus.

Figure 4. Mendelian randomization analysis of T1D and LP. IVW, MR-Egger, and weighted median regression were used to estimate the causal relationship between genetically predicted T1D and LP. IVW = inverse variance weighted, LP = lichen planus, MR = Mendelian randomization, T1D = type 1 diabetes.

4

(version 4.2.2). Inputting the differential gene list from the Venn diagram into the KEGG API (https://www.kegg.jp/kegg/rest/keggapi.html), the latest KEGG Pathway gene annotations were obtained, mapping genes to the background set. The org.Hs.eg. db package (version 3.1.0) provided GO annotations of genes, serving as the background set. A minimum gene set of 5 and a maximum gene set of 5000 with a *P* value < .05 and FDR < 0.25 were considered statistically significant measures.

2.2.5. Gene set enrichment analysis.

Gene set enrichment analysis (GSEA) software (version 3.0) was obtained from the GSEA website (DOI: 10.1073/pnas.0506580102, http://software.broadinstitute.org/gsea/index.jsp). Groups of disease and normal samples were divided into 2 groups, and subsets from the Molecular Signatures Database (DOI: 10.1093/bioinformatics/btr260, http://www.gsea-msigdb.org/gsea/downloads.jsp) were downloaded (c2.cp.kegg.v7.4.symbols.gmt) to evaluate related pathways and molecular mechanisms. Based on gene expression profiles and phenotype grouping, a minimum gene set of 5, a maximum gene set of 5000, 1000 permutations, with a *P* < .05 and FDR < 0.25 were considered statistically significant. GO and KEGG analyses were performed on the entire genome.

2.2.6. Construction and analysis of protein–protein interaction networks.

The Search Tool for the Retrieval of Interacting Genes (STRING) database (http://string-db.org/) was used to collect, score, and integrate all publicly available protein–protein interaction (PPI) information sources, supplemented by predicted interactions. The differential gene list was input into the STRING database to construct a predicted core gene PPI network (confidence > 0.4). Cytoscape software provided biological network analysis and 2D visualization. The PPI network from STRING was visualized and core genes were predicted using Cytoscape software. The PPI network was imported into Cytoscape, and MCODE was used to find the best modules. Five algorithms (MCC, DMNC, EPC, Radiality, EcCentricity) were used to calculate the best-related genes and intersections were taken, visualized, and core gene lists exported.

2.2.7. Gene expression heatmap.

Using the R package heatmap, expression heatmaps were generated for core genes identified in the PPI network, showing expression differences between T1D and LP samples compared to normal samples in GSE156035 and GSE52130.

2.2.8. Comparative toxicogenomics database analysis.

The comparative toxicogenomics database (CTD) integrates a large amount of data on interactions between chemicals, genes, phenotypes, and diseases, providing great convenience for research on environmental exposure factors related to diseases and potential mechanisms of drugs. Core genes were input into the CTD website, identifying the most relevant diseases related to core genes, and radar plots of gene expression differences were plotted using Excel.

3. Results

3.1. MR analysis of T1D and LP

3.1.1. Genetic instrument variables for T1D.

Using T1D as an exposure factor, after removing 2 SNPs that may exhibit heterogeneity (rs3129880, *P* = .0148, and rs719654, *P* < .0074), we extracted 35 SNP results. Table 1 lists genetic instruments
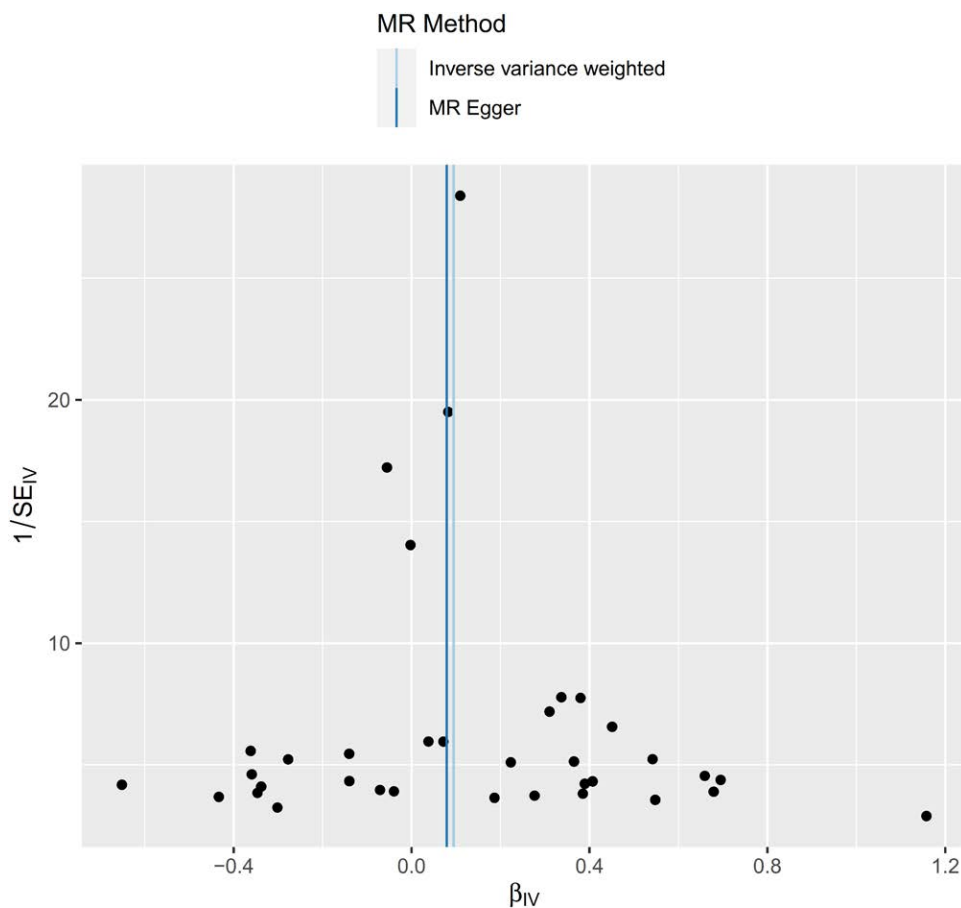

Figure 5. Horizontal pleiotropy analysis.

related to T1D and LP. Flow chart of MR analysis (Fig. 1). Each black dot represents an SNP, where the *X*-axis shows the impact of the SNP on exposure, and the *Y*-axis represents the SNP's effect on the outcome. The slope of each line reflects the potential causal relationship (Fig. 2), and the results support a positive correlation between T1D and LP. LP risk associations for each genetic variant (Fig. 3).

3.1.2. MR analysis of T1D and LP.
We used IVW, MR-Egger, and weighted median regression to estimate the causal relationship between genetically predicted T1D and LP (Fig. 4). Both the weighted median and IVW methods supported a positive correlation between T1D and LP. The IVW results were the most decisive, showing evidence of a causal relationship between T1D and LP (Weighted median odds ratio [OR] = 1.09682 [95% CI, 1.03391–1.16355], $P = .00217$. IVW odds ratio [OR] = 1.09908 [95% CI, 1.02388–1.17981]; $P = .00899$).

3.1.3. Horizontal pleiotropy analysis.
A funnel plot illustrated the individual Wald ratio of each SNP against its precision, where asymmetry suggested directional horizontal pleiotropy. However, with a small sample of SNPs, assessing the symmetry was challenging (Fig. 5). The MR-Egger intercept in our study indicated no significant evidence of directional pleiotropy ($P = .70645$). Thus, no directional pleiotropic effects were observed between T1D and LP.

3.1.4. Effect of individual genetic instruments on LP.
An exclusion analysis was performed to assess the impact of each SNP on the overall causal estimate. When systematically removing individual SNPs and repeating the MR analysis, no substantial differences in the estimated causal effect were observed (Fig. 6). Therefore, no single genetic instrument was driving the observed effects.

3.2. DEGs of T1D and LP

3.2.1. DEGs analysis.
In this study, DEGs were identified in the T1D dataset GSE156035 (Fig. 7A) and the LP dataset GSE52130 (Fig. 7B), based on preset cutoff values. A Venn diagram was used to intersect DEGs from both datasets, resulting in 118 DEGs (Fig. 7C).

3.2.2. Functional enrichment analysis.

3.2.2.1. GO and KEGG analysis.
GO and KEGG analyses were performed on the DEGs. In the biological processes analysis, genes were mainly enriched in the apoptotic process, inflammatory response, insulin receptor signaling pathway, and glucose metabolism (Fig. 8A). In the cellular component analysis, they were mainly enriched in mediator complex and replisome (Fig. 8B). In the molecular function analysis, the genes were enriched in interleukin-2 receptor activity and helicase activity (Fig. 8C). The KEGG analysis showed enrichment in metabolic pathways, PI3K-Akt signaling pathway, cell cycle, P53 signaling pathway, and AGE-RAGE signaling pathway in diabetic complications (Fig. 8D).

3.2.2.2. Gene set enrichment analysis.
GSEA was performed on the entire genome to identify possible enriched pathways. The results from the GSEA aligned with the GO and KEGG
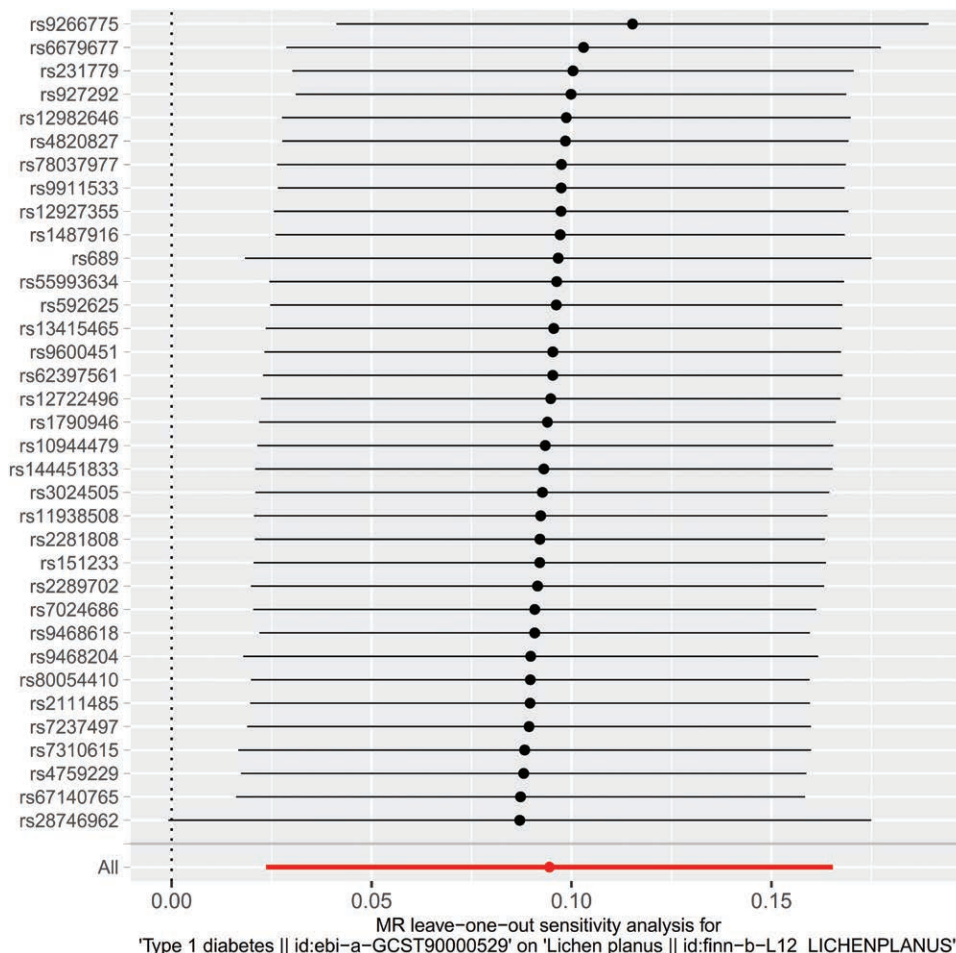

Figure 6. Effect of individual genetic instruments on LP. When systematically removing individual SNPs and repeating the MR analysis, no substantial differences in the estimated causal effect were observed. LP = lichen planus, MR = Mendelian randomization, SNPs = single nucleotide polymorphisms.

results, showing enrichment in inflammatory response, insulin receptor signaling pathway, mediator complex, helicase activity, cell cycle, and P53 signaling pathway (Fig. 9).

3.2.2.3.   Metascape   enrichment   analysis. Metascape revealed GO enrichment in cytokine signaling in the immune system, hormone transport, and positive regulation of protein phosphorylation (Fig. 10A). Enrichment networks were also generated, showing connections and confidence levels between enriched terms (Fig. 10B and C).

3.2.3. Weighted gene co-expression network analysis. The selection of soft threshold power is an important step in

weighted gene co-expression network analysis (WGCNA). Network topology analysis was performed to determine the soft threshold power. The soft-threshold power in the WGCNA was set to 4 (Fig. 11A). Hierarchical clustering trees of all genes were constructed, 34 important modules were generated (Fig. 11B), and interactions between important modules were analyzed (Fig. 11C). Heat maps of correlation between modules and phenotypes (Fig. 11D) and scatter plots of correlation between GS and MM of related hub genes (Fig. 11E) were generated. We calculated the expression correlation between module feature vectors and genes to obtain MM, and according to the cutoff criterion ($|MM| > 0.8$), 6 genes with high connectivity
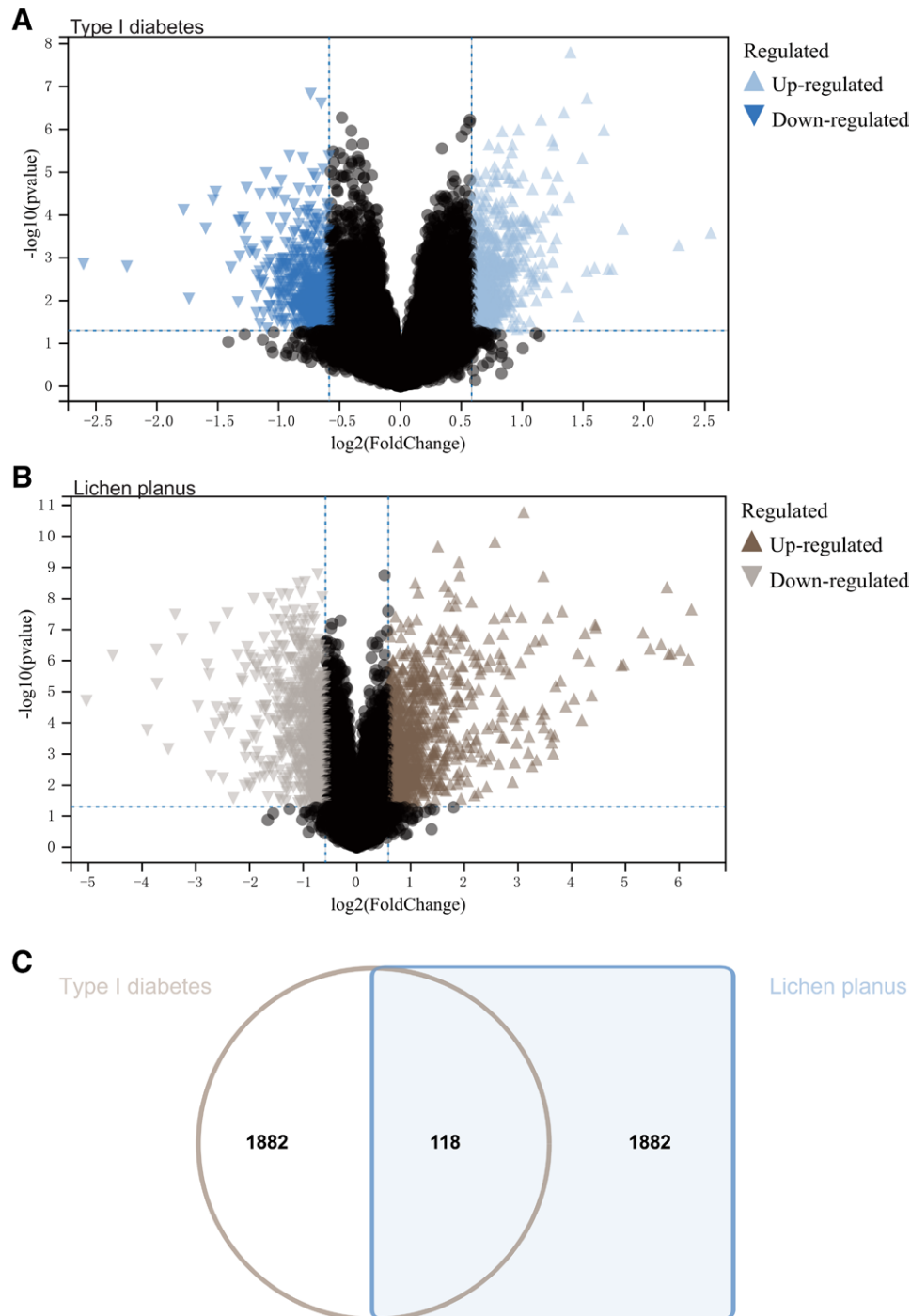

Figure 7. DEGs of T1D and LP. (A) T1D dataset GSE156035. (B) LP dataset GSE52130. (C) A Venn diagram was used to intersect DEGs from both datasets, resulting in 118 DEGs. DEGs = differentially expressed genes, LP = lichen planus, T1D = type 1 diabetes.

in clinically significant modules were identified as hub genes. The DEGs screened by WGCNA and DEGs were drawn a Venn diagram and intersected (Fig. 11F), and these DEGs were used for subsequent analysis.

3.2.4. PPI network construction and analysis. A PPI network of the DEGs was built using the STRING database and visualized in Cytoscape (Fig. 12A). Five algorithms (MCC, DMNC, EPC, Radiality, and EcCentricity) were used to identify the hub genes (Fig. 12B–F), and the Venn diagram was made to obtain the merge as the core genes (Fig. 12G). Finally, 3 core genes (SIRPG, MED12, and MAT2A) were obtained.

3.2.5. Core gene expression heatmap. The expression levels of core genes in GSE184050 and GSE160997 were visualized and heat maps were made respectively, and the expression results of core gene SIRPG were found to be more significant. SIRPG was highly expressed in T1D samples and low expression in normal samples (Fig. 13A), and SIRPG was highly expressed in LP samples. The expression was low in normal samples (Fig. 13B). Taken together with the results of the MR analysis described above, a positive causal relationship between T1D and LP is suggested. The core gene SIRPG was also found in the SNP result list and was highly expressed in both T1D and LP.

3.2.6. CTD analysis. The hub genes were input into the CTD website, revealing associations with metabolic syndrome, lipid metabolism disorders, cardiovascular diseases, immune system diseases, painful peripheral neuropathy, and inflammation (Fig. 13C).

4. Discussion

T1D is a severe chronic disease, primarily harmful due to acute and chronic complications caused by poor long-term blood glucose control.[15] Persistent hyperglycemia can damage the nervous system, leading to peripheral neuropathy and autonomic neuropathy.[16] LP is a relatively common non-infectious chronic inflammatory disease affecting the skin and mucous membranes.[17] Its pathological characteristics include hyperkeratosis, basal cell liquefaction degeneration, sawtooth-shaped rete ridges, band-like lymphocytic infiltration in the superficial dermis, and immunopathological changes in the basement membrane zone. LP tends to recur, with most patients experiencing alternating periods of remission and relapse.[18] Exploring the molecular mechanisms underlying T1D and LP is crucial for the development of targeted therapies. The primary result of this study is that SIRPG is highly expressed in both T1D and LP, with higher expression levels associated with worse prognoses.
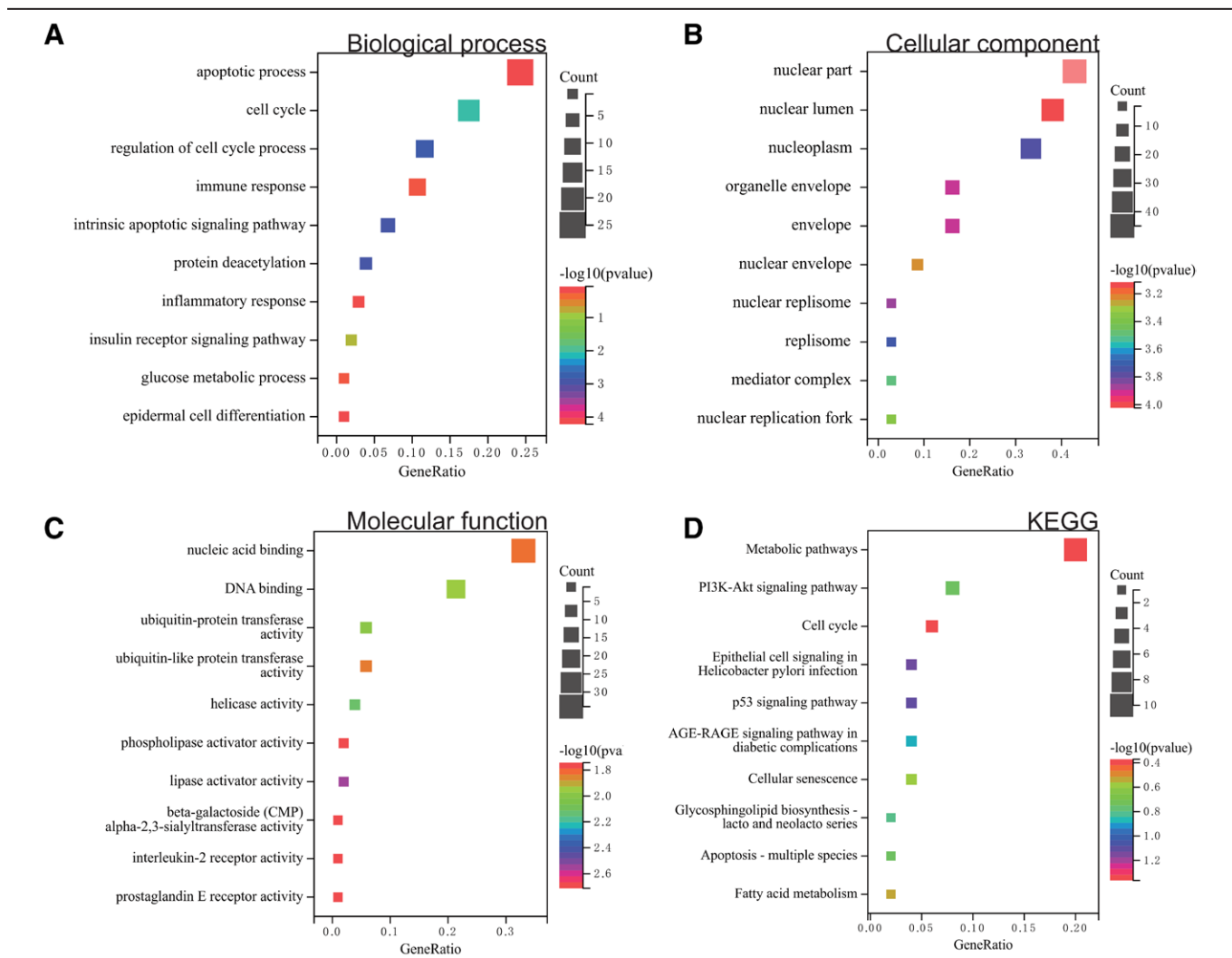

Figure 8. Functional enrichment analysis. GO and KEGG analysis. (A) BP analysis. (B) CC analysis. (C) MF analysis. (D) KEGG analysis. BP = biological processes, CC = cellular components, GO = gene ontology, KEGG = Kyoto Encyclopedia of Genes and Genomes, MF = molecular function.

The core pathological feature of T1D is the absolute lack of insulin secretion, caused by the immune system mistakenly attacking and destroying pancreatic islet β-cells.[19] Since insulin is a key hormone for regulating blood glucose levels, its destruction results in an inability to produce sufficient insulin, leading to hyperglycemia, and thus patients require insulin injections to control blood glucose levels. The immune system mistakenly recognizes β-cells as "foreign substances" and produces
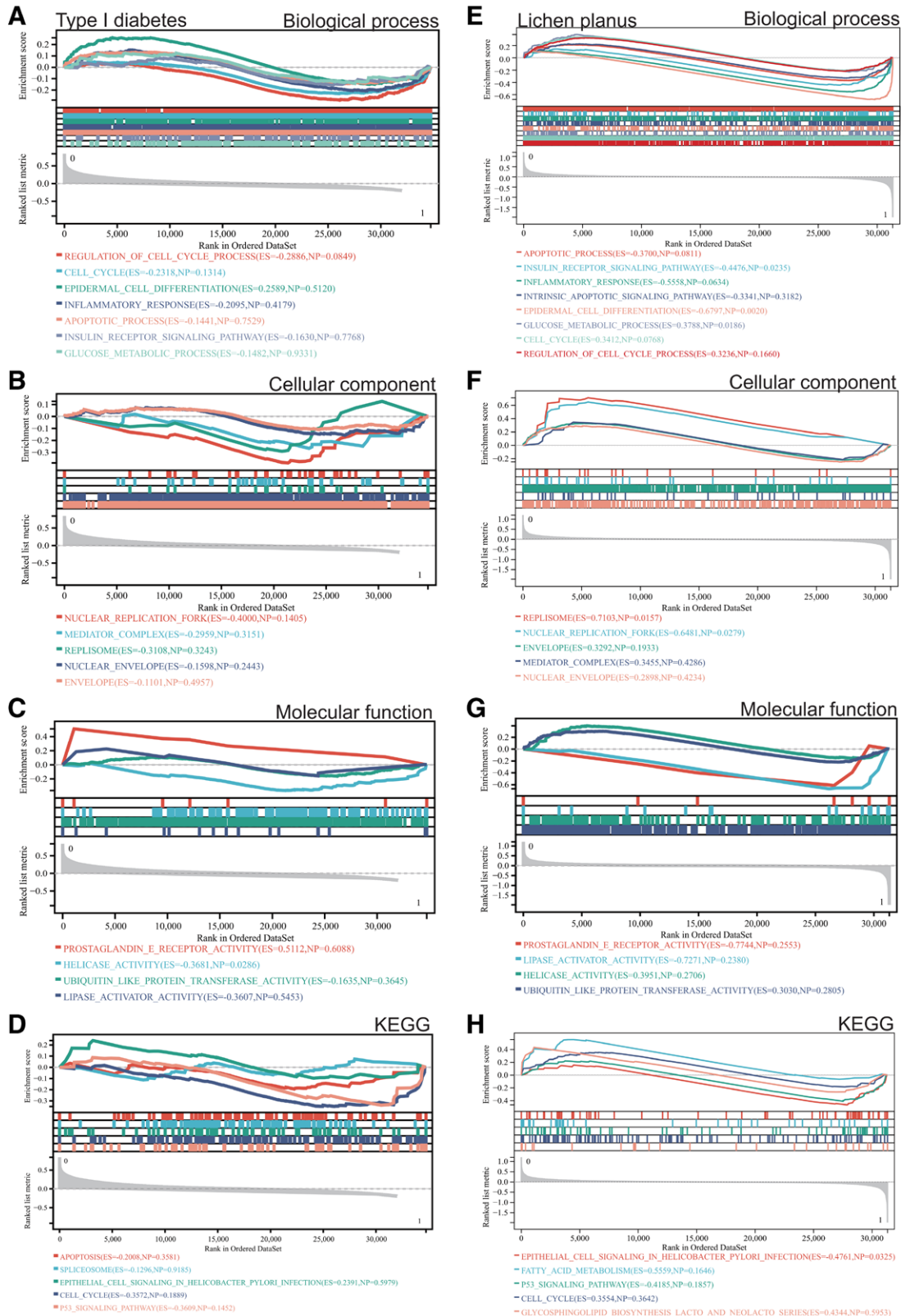

Figure 9. Functional enrichment analysis. GSEA. (A) Biological process of Type I diabetes. (B) Cellular component of Type I diabetes. (C) Molecular function of Type I diabetes. (D) KEGG of Type I diabetes. (E) Biological process of Lichen planus. (F) Cellular component of Lichen planus. (G) Molecular function of Lichen planus. (H) KEGG of Lichen planus. GSEA = gene set enrichment analysis.

autoantibodies against these cells. These autoantibodies, along with T-cell-mediated immune responses, collectively lead to the destruction of β-cells.[20] β-cells are responsible for producing and secreting insulin, which is crucial for regulating blood glucose.[21] When these cells are destroyed, insulin levels in the body drop significantly or are completely absent. The absolute lack

of insulin prevents cells from effectively utilizing glucose, leading to elevated blood glucose levels, causing a hyperglycemic state. In the pancreatic tissue of T1D patients, a typical pathological change is the infiltration of lymphocytes into the islets, a phenomenon known as insulitis. Lymphocytes gather around the islets, and as the disease progresses, the number of β-cells
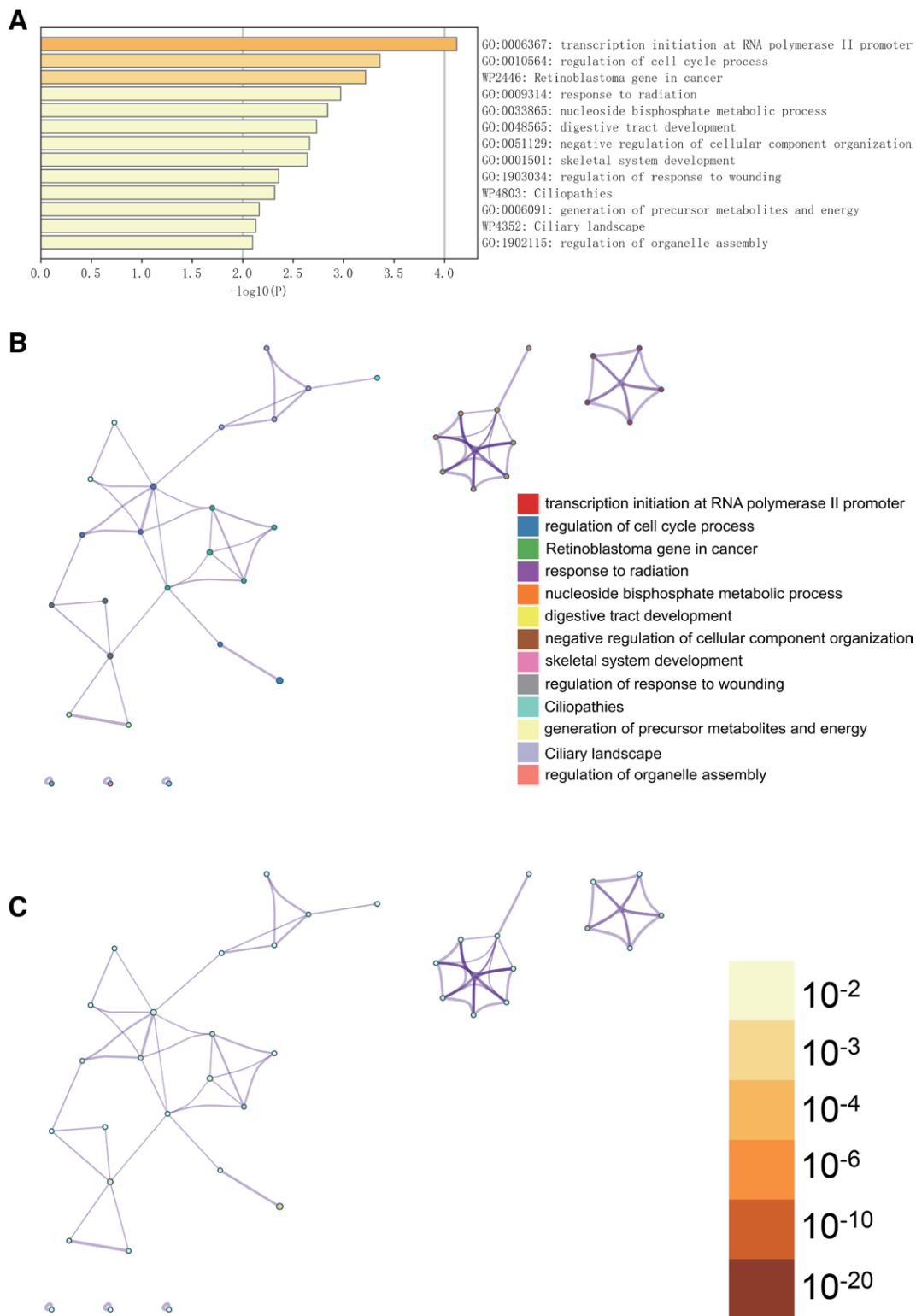

Figure 10. Metascape enrichment analysis. (A) Metascape revealed GO enrichment in cytokine signaling in the immune system, hormone transport, and positive regulation of protein phosphorylation. (B and C) Enrichment networks were also generated, showing connections and confidence levels between enriched terms. GO = gene ontology.

gradually decreases and may eventually disappear entirely. Due to the lack of insulin, the body cannot use glucose as the primary energy source and instead turns to fat breakdown for energy.[22,23] This metabolic shift produces a large number of ketone bodies, leading to diabetic ketoacidosis, which can occur during the initial onset of T1D or when insulin treatment is inadequate.

LP is an autoimmune disease related to immune system dysregulation, particularly T-cell-mediated immune responses, in which CD8+ T-cells attack the basal layer cells of the skin and
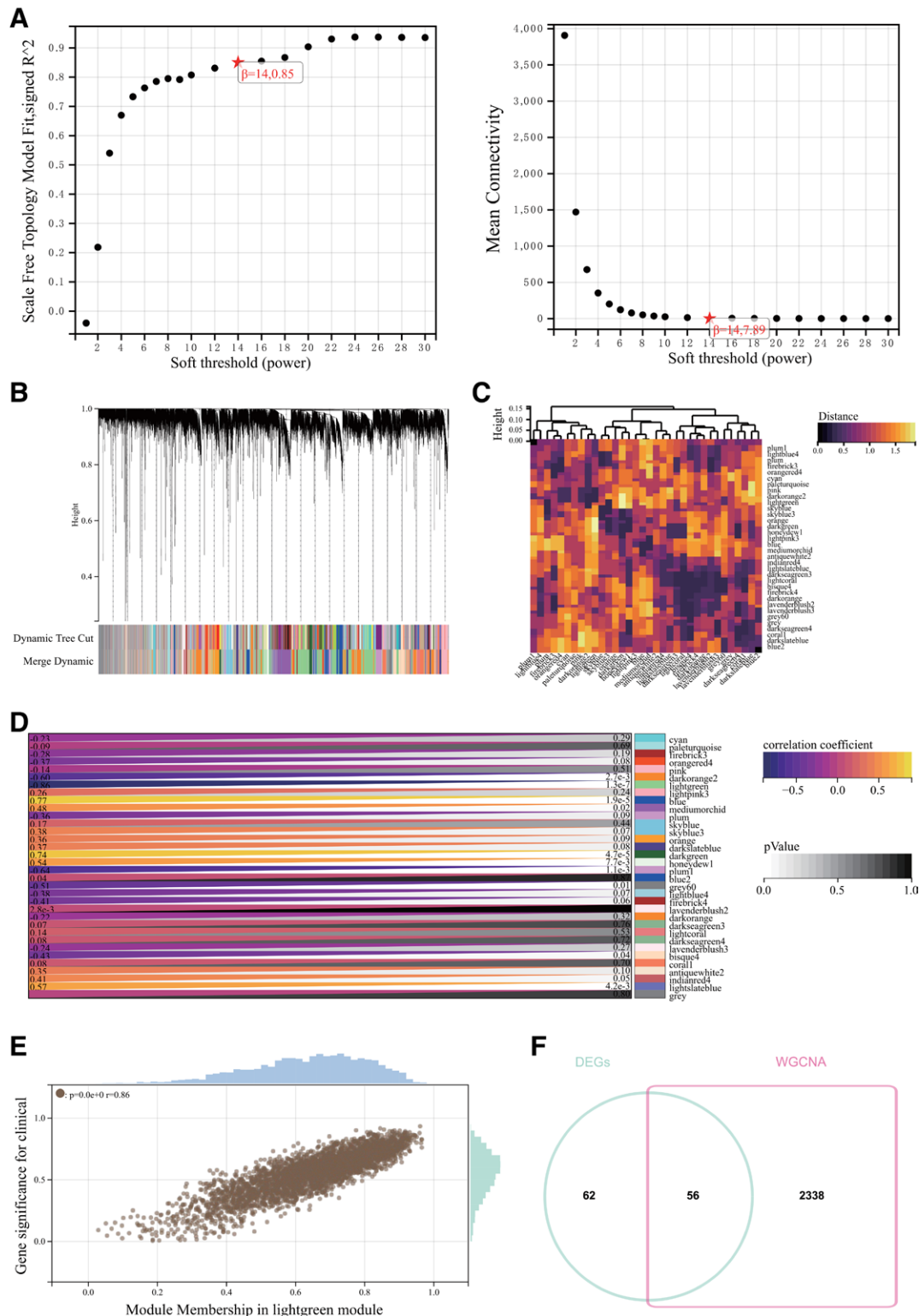

Figure 11. WGCNA. (A) The soft-threshold power in the WGCNA was set to 4. (B) Hierarchical clustering trees of all genes were constructed, 34 important modules were generated. (C) Interactions between important modules. (D) Heat maps of correlation between modules and phenotypes. (E) Scatter plots of correlation between GS and MM of related hub genes. (F) Venn diagram and intersected. WGCNA = weighted gene co-expression network analysis.

mucosa, leading to damage and inflammation.[24] While LP is not typically fatal, its chronic relapsing nature, potential risk of malignancy, and possible functional impairment can cause significant distress for patients.

SIRPG belongs to the SIRP family and is a gene associated with immune regulation. The SIRP family[25,26] includes members such as SIRPA, SIRPB, and SIRPG, which typically influence

immune cell migration, phagocytosis, and activation by interacting with CD47. The protein encoded by the SIRPG gene interacts with immune cells, particularly in regulating the activation and inhibition of T-cells and dendritic cells. SIRPG is usually expressed on the surface of immune system cells and is involved in mediating interactions between immune cells.[27] SIRPG plays a role in immune checkpoint regulation, similar to SIRPA and
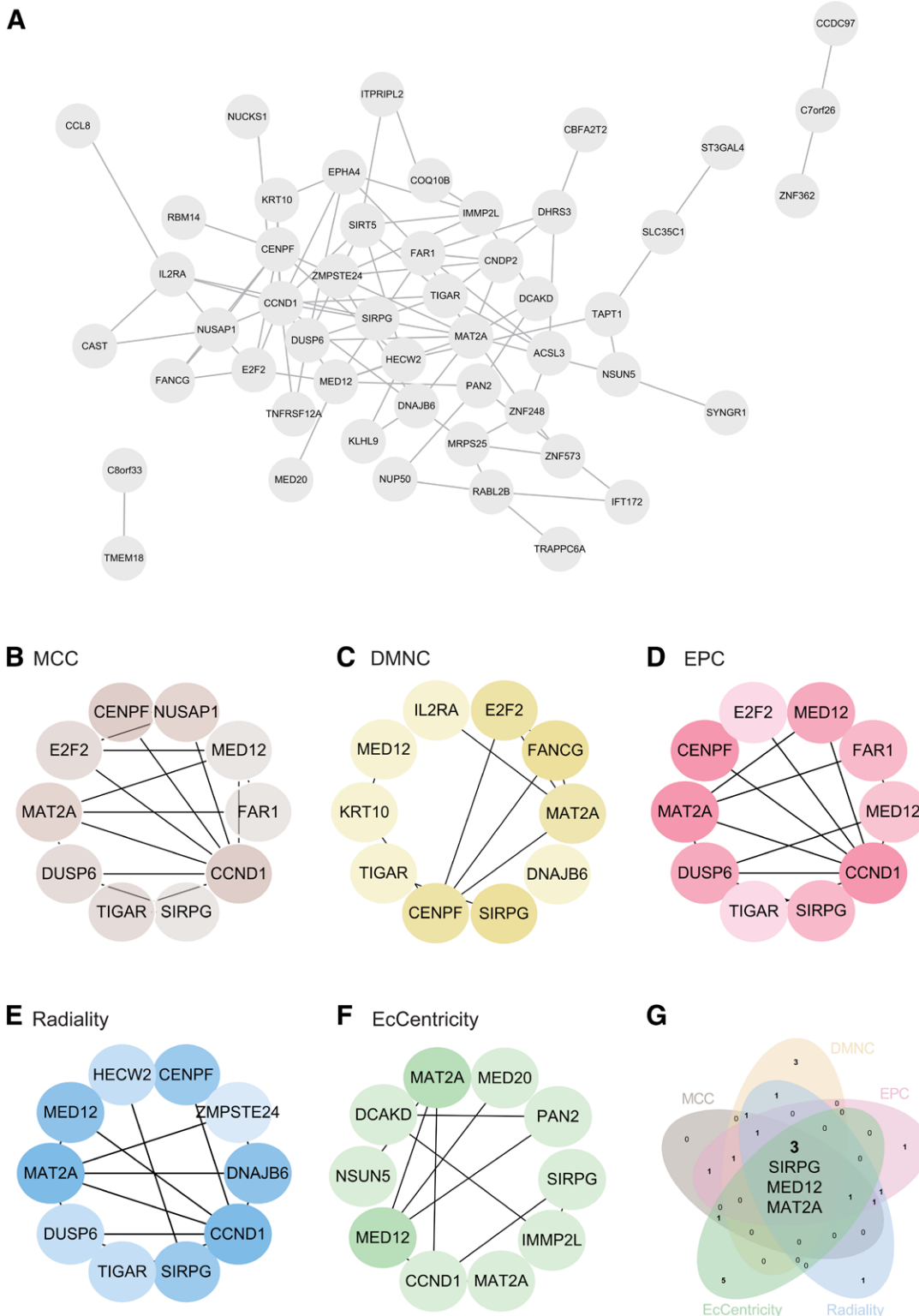

Figure 12. PPI network construction and analysis. (A) PPI network. (B–F) Five algorithms (MCC, DMNC, EPC, Radiality, and EcCentricity) were used to identify the hub genes. (G) Venn diagram. PPI = protein–protein interaction.

other immune inhibitory molecules, and may contribute to autoimmune diseases.

SIRPG may play a role in regulating immune system self-tolerance and autoimmune responses. Through interaction with CD47, SIRPG regulates T-cell activation and function, which is crucial for maintaining immune system self-tolerance. The imbalance of self-tolerance is often a key factor in the development of autoimmune diseases. In T1D, the attack on pancreatic

β-cells by T-cells is one of the primary pathological mechanisms. Polymorphisms in the SIRPG gene may be related to susceptibility to T1D in certain populations. Variants in SIRPG that regulate T-cell and dendritic cell function may influence individual immune responses and the risk of developing T1D. The immune regulatory protein encoded by the SIRPG gene is involved in modulating immune cell activation. As T1D is an autoimmune disease, where the immune system mistakenly
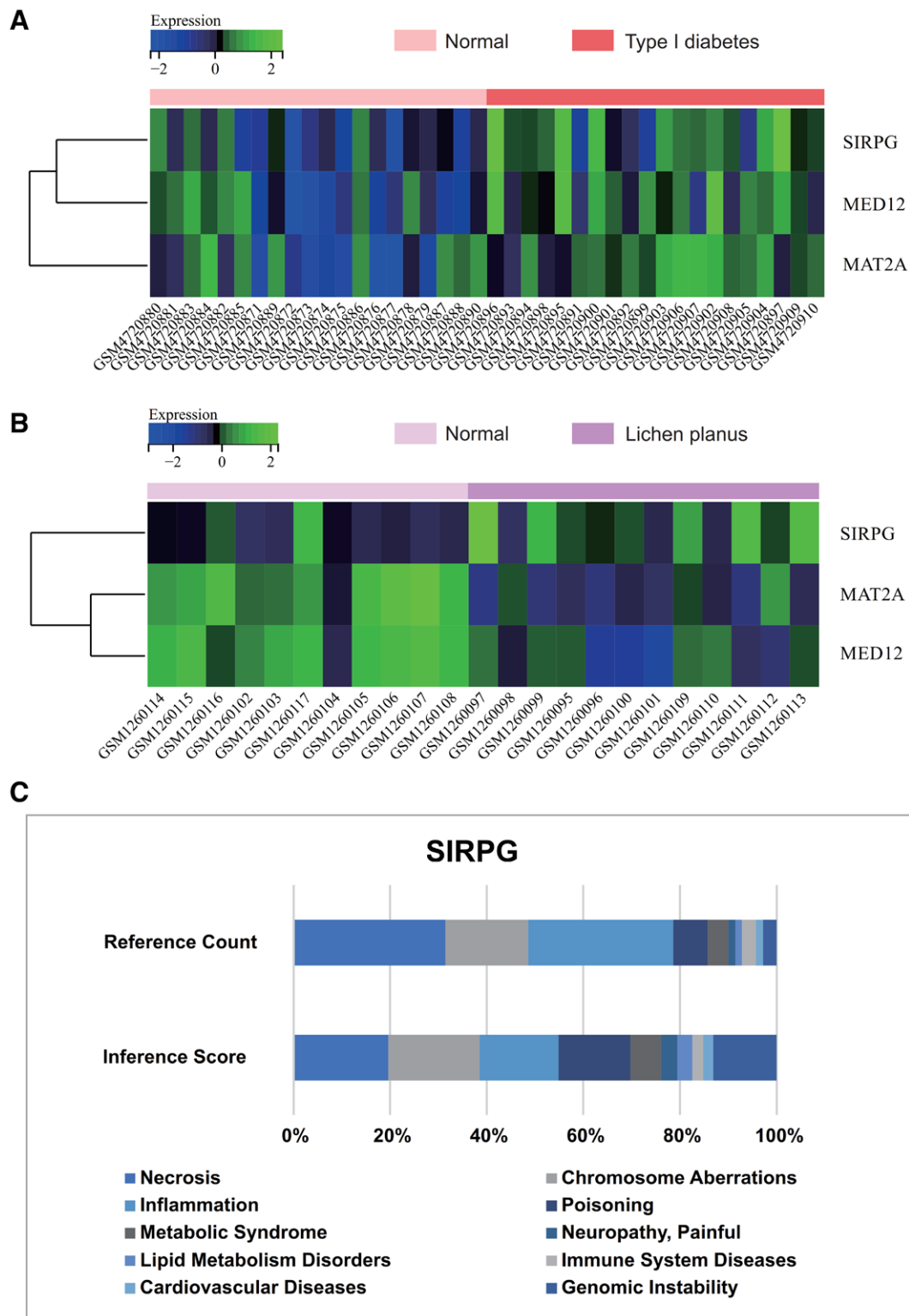

Figure 13. Core gene expression heatmap of (A) type 1 diabetes and (B) lichen planus. (C) CTD analysis. CTD = comparative toxicogenomics database.

attacks pancreatic β-cells, SIRPG, through its expression on the surface of immune cells, regulates the activation and inhibition of these cells and may influence the occurrence of autoimmune responses. Studies suggest that abnormalities in the immune system are associated with the development of T1D.[28] As a regulator of immune responses, SIRPG may play a role in the immunopathology of T1D. Splicing regulation of SIRPG could modulate the risk of T1D.[29] SIRPG also regulates immune cell activity, particularly in the regulation of T-cells and natural killer cells, which may play a role in certain immune-related skin diseases, though the direct link remains unclear. SIRPG may influence the immune system and autoimmune responses,[30] and its role in diseases like LP deserves further investigation. While there is no direct evidence that the SIRPG gene plays a clear role in the pathogenesis of LP, its key position in regulating T-cell immune function suggests that it may be related to the progression of LP. SIRPG's function may be linked to the chronic inflammation seen in LP. The persistent inflammatory response characteristic of LP may be influenced by SIRPG's regulation of interactions between immune cells, potentially contributing to either overactivation or suppression of immune responses in the disease. Through its interaction with CD47, SIRPG regulates T-cell activation and function, which may indirectly affect T-cell activity in LP.

Although this study presents rigorous bioinformatic analysis, certain limitations remain. The study lacks gene overexpression or knockout animal experiments to further validate SIRPG's function. Future research should delve deeper into this aspect.

5. Conclusion

The high expression of SIRPG is significantly associated with both T1D and LP. The upregulation of SIRPG may exacerbate disease progression by regulating immune responses, promoting inflammation, and causing tissue damage. This finding provides new insights into the pathogenesis of T1D and LP and suggests that SIRPG could serve as a potential therapeutic target. Intervening in its expression or function may help improve patient outcomes.

Author contributions

Conceptualization: Ji Wu, Honglei Zhao.
Data curation: Ji Wu, Honglei Zhao.
Formal analysis: Ji Wu.
Methodology: Ji Wu, Honglei Zhao.
Writing – original draft: Ji Wu, Honglei Zhao.
Writing – review & editing: Ji Wu, Honglei Zhao.

References

[1] Desai S, Deshmukh A. Mapping of type 1 diabetes mellitus. Curr Diabetes Rev. 2020;16:438–41.

[2] Primavera M, Giannini C, Chiarelli F. Prediction and prevention of type 1 diabetes. Front Endocrinol (Lausanne). 2020;11:248.

[3] Lu X, Zhao C. Exercise and type 1 diabetes. Adv Exp Med Biol. 2020;1228:107–21.

[4] Haak T, Gölz S, Fritsche A, et al. Therapy of type 1 diabetes. Exp Clin Endocrinol Diabetes. 2019;127:S27–38.

[5] Redondo MJ, Morgan NG. Heterogeneity and endotypes in type 1 diabetes mellitus. Nat Rev Endocrinol. 2023;19:542–54.

[6] McDermott E, Solomon N, Silva AM, Khoshpouri P. Lichen planus. Radiographics. 2023;43:e220200.

[7] Louisy A, Humbert E, Samimi M. Oral lichen planus: an update on diagnosis and management. Am J Clin Dermatol. 2024;25:35–53.

[8] van Hees C, van der Meij EH. Lichen planus. Ned Tijdschr Tandheelkd. 2023;130:221–6.

[9] Solimani F, Forchhammer S, Schloegl A, Ghoreschi K, Meier K. Lichen planus—a clinical guide. J Dtsch Dermatol Ges. 2021;19:864–83.

[10] Bowden J, Holmes MV. Meta-analysis and Mendelian randomization: a review. Res Synth Methods. 2019;10:486–96.

[11] Skrivankova VW, Richmond RC, Woolf B, et al. Strengthening the reporting of observational studies in epidemiology using Mendelian randomisation (STROBE-MR): explanation and elaboration. BMJ. 2021;375:n2233.

[12] Wang Y, Zhao Y, Bollas A, Wang Y, Au KF. Nanopore sequencing technology, bioinformatics and applications. Nat Biotechnol. 2021;39:1348–65.

[13] Gauthier J, Vincent AT, Charette SJ, Derome N. A brief history of bioinformatics. Brief Bioinform. 2019;20:1981–96.

[14] Azad RK, Shulaev V. Metabolomics technology and bioinformatics for precision medicine. Brief Bioinform. 2019;20:1957–71.

[15] Ilonen J, Lempainen J, Veijola R. The heterogeneous pathogenesis of type 1 diabetes mellitus. Nat Rev Endocrinol. 2019;15:635–50.

[16] Toschi E. Type 1 diabetes and aging. Endocrinol Metab Clin North Am. 2023;52:389–403.

[17] Böer-Auer A, Lütgerath C. Lichen planus: fundamentals, clinical variants, histological features, and differential diagnosis. Hautarzt. 2020;71:1007–21.

[18] Daye M, Işık B, Kılınç F. Lichen planus due to hirudotherapy. Turkiye Parazitol Derg. 2021;45:149–52.

[19] Ramos EL, Dayan CM, Chatenoud L, et al. Teplizumab and β-cell function in newly diagnosed type 1 diabetes. N Engl J Med. 2023;389:2151–61.

[20] Roep BO, Thomaidou S, van Tienhoven R, Zaldumbide A. Type 1 diabetes mellitus as a disease of the β-cell (do not blame the immune system?). Nat Rev Endocrinol. 2021;17:150–61.

[21] Blanchi B, Taurand M, Colace C, et al. EndoC-βH5 cells are storable and ready-to-use human pancreatic beta cells with physiological insulin secretion. Mol Metab. 2023;76:101772.

[22] Norton L, Shannon C, Gastaldelli A, DeFronzo RA. Insulin: the master regulator of glucose metabolism. Metabolism. 2022;129:155142.

[23] Sabu C, Henna TK, Raphey VR, Nivitha KP, Pramod K. Advanced biosensors for glucose and insulin. Biosens Bioelectron. 2019;141:111201.

[24] El-Howati A, Thornhill MH, Colley HE, Murdoch C. Immune mechanisms in oral lichen planus. Oral Dis. 2023;29:1400–15.

[25] Shen Q, Zhao L, Pan L, et al. Soluble SIRP-alpha promotes murine acute lung injury through suppressing macrophage phagocytosis. Front Immunol. 2022;13:865579.

[26] Zhang H, Li F, Yang Y, Chen J, Hu X. SIRP/CD47 signaling in neurological disorders. Brain Res. 2015;1623:74–80.

[27] Luo L, Jiang M, Wu H, et al. SIRPG expression positively associates with an inflamed tumor microenvironment and response to PD-1 blockade. Cancer Immunol Immunother. 2024;73:147.

[28] Eizirik DL, Szymczak F, Mallone R. Why does the immune system destroy pancreatic β-cells but not α-cells in type 1 diabetes. Nat Rev Endocrinol. 2023;19:425–34.

[29] Smith MJ, Pastor L, Newman J, Concannon P. Genetic control of splicing at SIRPG modulates risk of type 1 diabetes. Diabetes. 2022;71:350–8.

[30] Mao G, Li J, Wang N, et al. SIRPG promotes lung squamous cell carcinoma pathogenesis via M1 macrophages: a multi-omics study integrating data and Mendelian randomization. Front Oncol. 2024;14:1392417.