



OPEN ACCESS

EDITED BY

Ramkumar Kunka Mohanram,
SRM Institute of Science and Technology,
India

REVIEWED BY

Madan Kumar Perumal,
Central Food Technological Research
Institute (CSIR), India
Farooq Syed,
Beckman Research Institute at the
City of Hope, United States

*CORRESPONDENCE

Shaowei Dong
✉ michael.dong.85@gmail.com
Liqun Chen
✉ 2983733132@qq.com

[†]These authors have contributed equally to
this work

RECEIVED 22 November 2024

ACCEPTED 26 December 2024

PUBLISHED 13 January 2025

CITATION

Qiu J, Zhu P, Shi X, Xia J, Dong S and Chen L
(2025) Identification of a pancreatic stellate
cell gene signature and lncRNA interactions
associated with type 2 diabetes progression.
Front. Endocrinol. 15:1532609.
doi: 10.3389/fendo.2024.1532609

COPYRIGHT

© 2025 Qiu, Zhu, Shi, Xia, Dong and Chen. This
is an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Identification of a pancreatic stellate cell gene signature and lncRNA interactions associated with type 2 diabetes progression

Jinjun Qiu^{1†}, Peng Zhu^{1,2†}, Xing Shi³, Jinquan Xia³,
Shaowei Dong^{4*} and Liqun Chen^{3*}

¹Shenzhen Pingshan District People's Hospital, Pingshan Hospital, Southern Medical University, Shenzhen, China, ²Clinical Laboratory, Shenzhen Pingshan District People's Hospital, Pingshan Hospital, Southern Medical University, Shenzhen, China, ³Huangjiang Hospital, Dongguan, Guangdong, China, ⁴Department of Hematology and Oncology, Shenzhen Children's Hospital, Shenzhen, China

Background: Type 2 diabetes (T2D) has become a significant global health threat, yet its precise causes and mechanisms remain unclear. This study aims to identify gene expression patterns specific to T2D pancreatic islet cells and to explore the potential role of pancreatic stellate cells (PSCs) in T2D progression through regulatory networks involving lncRNA-mRNA interactions.

Methods: In this study, we screened for upregulated genes in T2D pancreatic islet samples using bulk sequencing (bulkseq) datasets and mapped these gene expression profiles onto three T2D single-cell RNA sequencing (scRNAseq) datasets. The identified T2D-specific gene features were further validated in an additional T2D scRNAseq dataset, a T1D scRNAseq dataset, and a T2D bulkseq dataset. To investigate regulatory networks, we analyzed the potential lncRNA-mRNA interactions within T2D peripheral blood mononuclear cell (PBMC) bulkseq data.

Results: Our analysis identified a specific gene panel—COL1A2, VCAN, and SULF1—that was consistently upregulated in T2D pancreatic islet samples. Expression of this gene panel was strongly associated with the activation of pancreatic stellate cells (PSCs), suggesting a unique T2D-specific signature characterized by COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs. This signature was exclusive to T2D and was not observed in Type 1 diabetes (T1D) samples, indicating a distinct role for activated PSCs in T2D progression. Furthermore, we identified six long non-coding RNAs (lncRNAs) that potentially interact with the COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs. These lncRNAs were mapped to a lncRNA-mRNA network, suggesting they may modulate immune responses and potentially reshape the immune microenvironment in T2D.

Discussion: Our findings highlight the potential immune-regulatory role of PSCs in T2D and suggest that PSC-related lncRNA-mRNA networks could serve as novel therapeutic targets for T2D treatment. This research provides insights into PSCs as a modulator in T2D progression, paving the way for innovative treatment strategies.

KEYWORDS

type 2 diabetes, COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs, combined analysis, gene expression profiling, lncRNA-mRNA network, immune microenvironment

Introduction

In 2021, there are 537 million adults suffered from diabetes, and this number is estimated to be 783 million in 2045 (1). Type 2 diabetes is a chronic metabolic disorder disease (2) and accounts for the majority of increased diabetes incidence. Although tremendous effort has been made in T2D explorations, their pathogenesis and progression mechanisms remain largely unknown. Recently, with the help of advanced single-cell RNA sequencing (scRNA-Seq) technique, a comprehensive single-cell atlas of pancreas islets has been depicted by many studies, and a set of gene signatures relating to T2D progression have also been identified (3). However, these studies mainly focused on the changes in pancreatic cells such as alpha, beta, and PP (pancreatic polypeptide) cells while neglecting the regulatory effects of other components in the pancreatic islets' microenvironment. Oxidative stress and endoplasmic reticulum stress are considered as the potential cause of impaired insulin secretions and these cellular stresses are often associated with inflammatory responses (4). Blériot et al. showed that changes in circulating monocytes and tissue macrophages could also contribute to T2D progression (5). Therefore, it is necessary to explore the potential immune-regulatory role from the microenvironment perspective.

Although traditional bulkseq (bulk sequencing) technology has limitations in depicting the component changes in the tissue microenvironment, still it has many advantages compared to

scRNAseq. For example, due to the relatively high sequencing depth, it can detect the expression of many low-abundance genes; due to the relatively low cost, it can involve a large-scale number of samples (6). Hence the method of combing scRNAseq and bulkseq datasets has become a better way, which not only ensures the quantity of the samples involved but also guarantees the exploration from the perspective of the microenvironment (7, 8). In this study, through a combined analysis of scRNAseq and bulkseq data of T2D samples, we provide evidence that activated pancreatic stellate cells (PSCs) featured by COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} might play a critical role in T2D-specific progression through a possible lncRNA-mRNA regulatory network.

Materials and methods

Datasets used in this study

Gene expression matrix files for the datasets listed in Table 1 were retrieved from the Gene Expression Omnibus (GEO) database. The probe IDs for microarray datasets and Ensemble IDs were converted into gene symbols using GRCh38 as the reference genome. The confounding factors such as patient age, BMI or comorbidity information were not involved in the screening process of the target genes or the interacting network generation process.

TABLE 1 List of datasets used in this study.

Dataset	Type	Sample	Reference	Donor Number	Number of Cells
GSE164416	Bulkseq	Pancreatic islets	Wigger et al. (13)	T2D(39) vs ND(18)	NA
GSE200044	scRNAseq	Pancreatic islets	Wang et al. (14)	T2D(6) vs PreT2D(8) vs ND(6)	94,206
GSE195986	scRNAseq	Pancreatic islets	Not published, GEO database	T2D(4) vs ND(7)	55,515
GSE81608	scRNAseq	Pancreatic islets	Xin et al. (15)	T2D(4) vs ND(6)	3,513
GSE86469	scRNAseq	Pancreatic islets	Lawlor et al. (16)	T2D(3) vs ND(5)	638
GSE148073	scRNAseq	Pancreatic islets	Fasolino et al. (17)	T1D(5) vs ND(11)	69,645
GSE76896	Microarray	Pancreatic islets	Solimena et al. (18)	T2D(36) vs ND(32)	NA
GSE163980	Microarray	PBMCs	Not published, GEO database	T2D(5) vs ND(5)	NA

Identification of differentially expressed genes and function analysis

Differentially expressed genes (DEGs) were identified using the DESeq2 package in R (9). Raw count data was used as input, with a significance cutoff of a *p*.adj value of 0.05 to identify DEGs. GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analysis was performed using the ShinyGO online server (10) (<http://bioinformatics.sdstate.edu/go/>). Enriched terms were further sorted based on the “Fold Enrichment” value, and an FDR (False Discovery Rate) value of 0.05 was used as the significance cutoff.

Reconstruction of the scRNAseq dataset and analysis

All scRNAseq data integration process was applied using R “Seurat” package (11). The expression matrix file of each sample was first downloaded from the GEO online database under each accession number, and a Seurat object file was generated for each sample specifically. Then all Seurat object files were merged using the `merge()` function, followed by an integration process using the “IntegrateLayers” function and the “CCAIntegration” method. After integration, the combined object was processed using the “FindNeighbors” function (`dims = 1:30`), “FindClusters” function (`resolution = 1`) and “RunUMAP” function (`dims = 1:30`). Final clusters were visualized using UMAP (Uniform Manifold Approximation and Projection) method for dimension reduction. Marker genes for each cluster were calculated using the “FindMarkers” function (`only.pos = T`, `logfc.threshold = 0.25`, `min.pct = 0.1`).

Gene panel expression calculation

For gene panel score calculation, expression values of each gene were first scaled using *Z*score, and then gene panel expression scores were calculated as the mean score of all input genes across all samples/cells.

Cluster similarity analysis

The similarity between clusters from different scRNAseq datasets was calculated as follows: for the c1 cluster from dataset 1 and the c2 cluster from dataset 2, the mean expression values of all genes were first calculated separately, and only overlapped genes were used. The correlation score between these two clusters was then calculated using the R “`cor.test`” function (`method = “Pearson”`). The results were visualized using the R “`pheatmap`” function. Marker gene similarity calculation was calculated as follows: for c1 cluster from dataset 1 and c2 cluster from dataset 2, their marker genes were first calculated separately using “`FindMarkers`” function (`only.pos = T`, `logfc.threshold = 0.25`, `min.pct = 0.1`), and the ratio of overlapped

marker genes was further calculated separately. The results were visualized using the R “`pheatmap`” function.

Receiver operating characteristic analysis

Receiver operating characteristic (ROC) analysis was performed using R “`ROCR`” and “`pROC`” packages: the gene panel scores from all samples were first calculated and ranked, then the AUC (Area under curve) score was calculated to evaluate the ability of the gene panel in separating samples under different conditions.

Retrieval of protein subcellular location information

The subcellular location information of target proteins was retrieved from THPA (The Human Protein Atlas, (www.proteinatlas.org) online database (under “`RNA.single.cell.type.specific.nTPM`” column). Transcription factor information was also retrieved from the THPA database (under the “`Protein.class`” column).

LncRNA-mRNA interaction prediction

The mRNA targets of candidate lncRNAs and mRNAs were predicted using the ENCORI online database (Encyclopedia of RNA Interactomes) (12). We first downloaded all lncRNA-RNA dataset, and then mapped target genes with protein coding gene lists. The cutoff for “`Interactions`” parameter was set to “1”. All the matching lncRNAs were recorded for further analysis. The interaction network was generated using R `ggplot2()` function.

Validation

To validate the key findings from the initial analyses, an independent validation dataset was retrieved from GEO. This dataset included gene expression profiles of samples with similar clinicopathological characteristics. Validation was performed through the following steps: The DEGs identified in the discovery cohort were cross-validated in the independent validation dataset using the same DESeq2 pipeline. The same significance thresholds (adjusted *p*-value < 0.05, $\log^2FC > 1$) were applied to confirm the reproducibility of the DEG results. Further, the gene panel expression scores computed in the discovery cohort were recalculated for the validation cohort. ROC curve analysis was repeated to evaluate the panel’s predictive accuracy, focusing on the reproducibility of the AUC values across the different datasets. GO and KEGG pathway enrichment analyses for the validated DEGs were performed using the ShinyGO platform. The biological functions enriched in the validation cohort were compared to those from the discovery cohort, with a focus on the consistency of the findings (FDR < 0.05). The predicted lncRNA-mRNA interactions

were cross-referenced using external resources such as miRTarBase and LncBase databases to confirm the interactions identified in the discovery cohort. These confirmed interactions were further analyzed for their biological relevance.

Results

Research design

As illustrated in Figure 1A and listed in Table 1, in this study, we first screened up-regulated genes in T2D pancreatic islet samples from bulkseq dataset [GSE164416 (13)], then mapped these up-regulated genes in T2D pancreatic islet samples from 3 scRNAseq datasets [GSE200044 (14), GSE195986, and GSE81608 (15)], and identified COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs as the key component in T2D progression. We further verified this feature in 3 more datasets (GSE86469 (16), T2D pancreatic islet scRNAseq dataset; GSE148073 (17), T1D pancreatic islet scRNAseq dataset; GSE76896

(18), T2D pancreatic islet microarray dataset), and generated a lncRNA-mRNA regulatory network using a lncRNA microarray dataset (GSE163980, T2D PBMC microarray dataset).

Up-regulated genes in T2D pancreatic islet samples from bulkseq dataset

We retrieved the bulkseq dataset from the GEO database (GSE164416) (13), and screened for differentially expressed genes (DEGs) between 39 T2D samples and 18 non-diabetic samples (ND) using R “DESEQ2” package. In total, there are 743 up-regulated genes (log2foldchange > 1 & p.adj < 0.05) and 52 down-regulated genes (log2foldchange < -1 & p.adj < 0.05) (Figure 1B, listed in Supplementary Table 1). The top enriched GO terms of the up-regulated genes are shown in Figure 1C, in which most of the terms are immune-related; the top enriched KEGG terms of the up-regulated genes are shown in Figure 1D, and the top enriched terms are protein digestion and ECM-related.

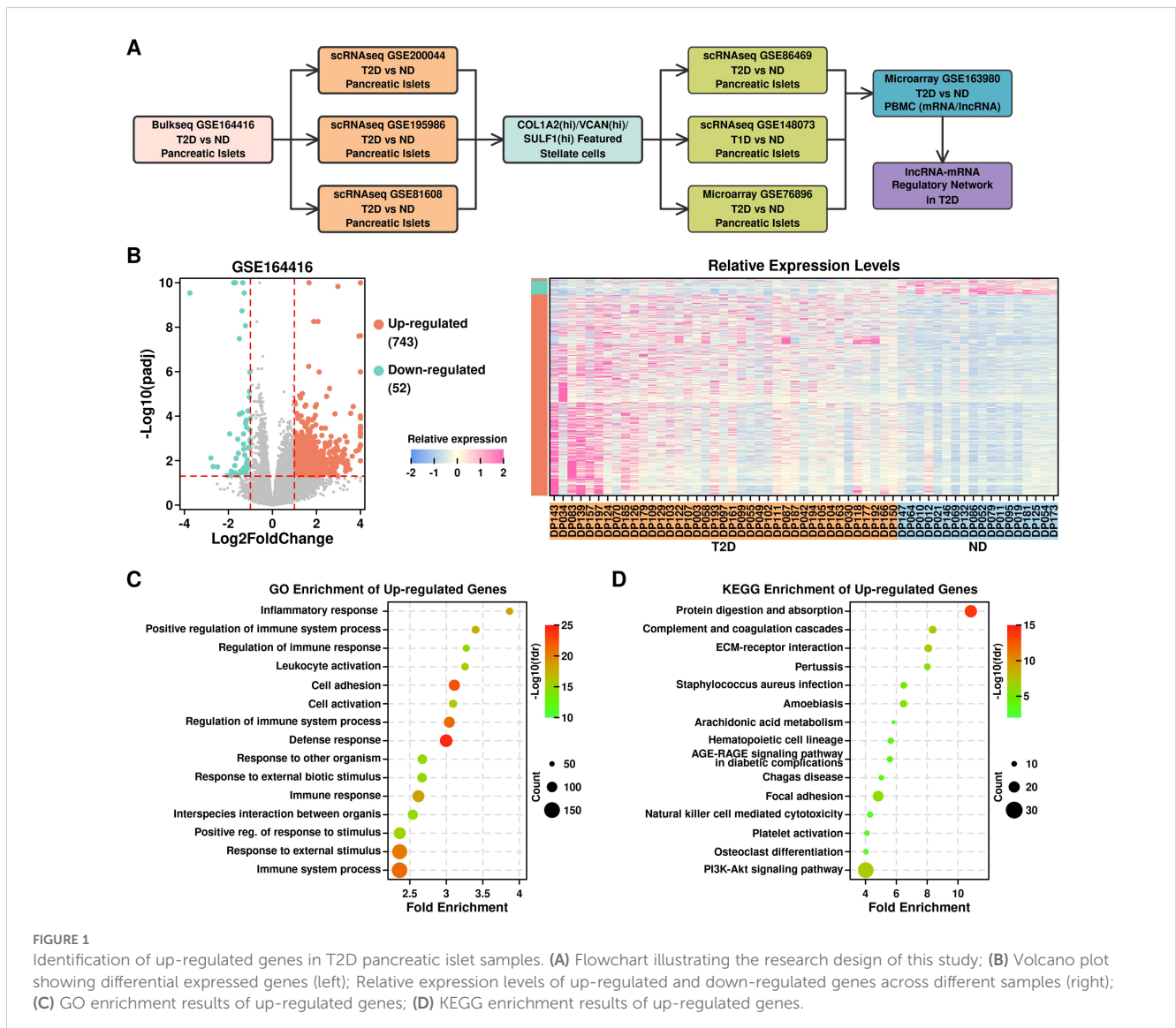


FIGURE 1

Identification of up-regulated genes in T2D pancreatic islet samples. (A) Flowchart illustrating the research design of this study; (B) Volcano plot showing differentially expressed genes (left); Relative expression levels of up-regulated and down-regulated genes across different samples (right); (C) GO enrichment results of up-regulated genes; (D) KEGG enrichment results of up-regulated genes.

Mapping bulkseq features to scRNAseq datasets

We retrieved three scRNAseq datasets of T2D pancreatic islet samples from the GEO database, including GSE200044 (6 ND samples, 8 preT2D samples and 6 T2D samples) (14), GSE195986 (7 ND samples, 4 T2D samples), and GSE81608 (15), and reconstructed the RDS files using their specific expression matrices. Their UMAP distribution patterns are illustrated in Figure 2: GSE200044 (Figure 2A: clusters; Figure 2B: per condition; Figure 2C: per sample; meta file in Supplementary Table 2); GSE195986 (Figure 2D: clusters;

Figure 2E: per condition; Figure 2F: per sample; meta file in Supplementary Table 3); GSE81608 (Figure 2G: clusters; Figure 2H: per condition; Figure 2I: per sample; meta file in Supplementary Table 4). Cells derived from different conditions/samples are evenly distributed across different clusters, suggesting that the integration process was reliable.

We further examined the expressional status of 743 up-regulated genes from GSE164416 across different scRNAseq datasets, and the overlapping status between these 743 genes and marker genes from each cluster (pct.1 cutoff of 0.1, 0.2, 0.3, 0.4, respectively): regarding GSE200044 dataset, these up-regulated

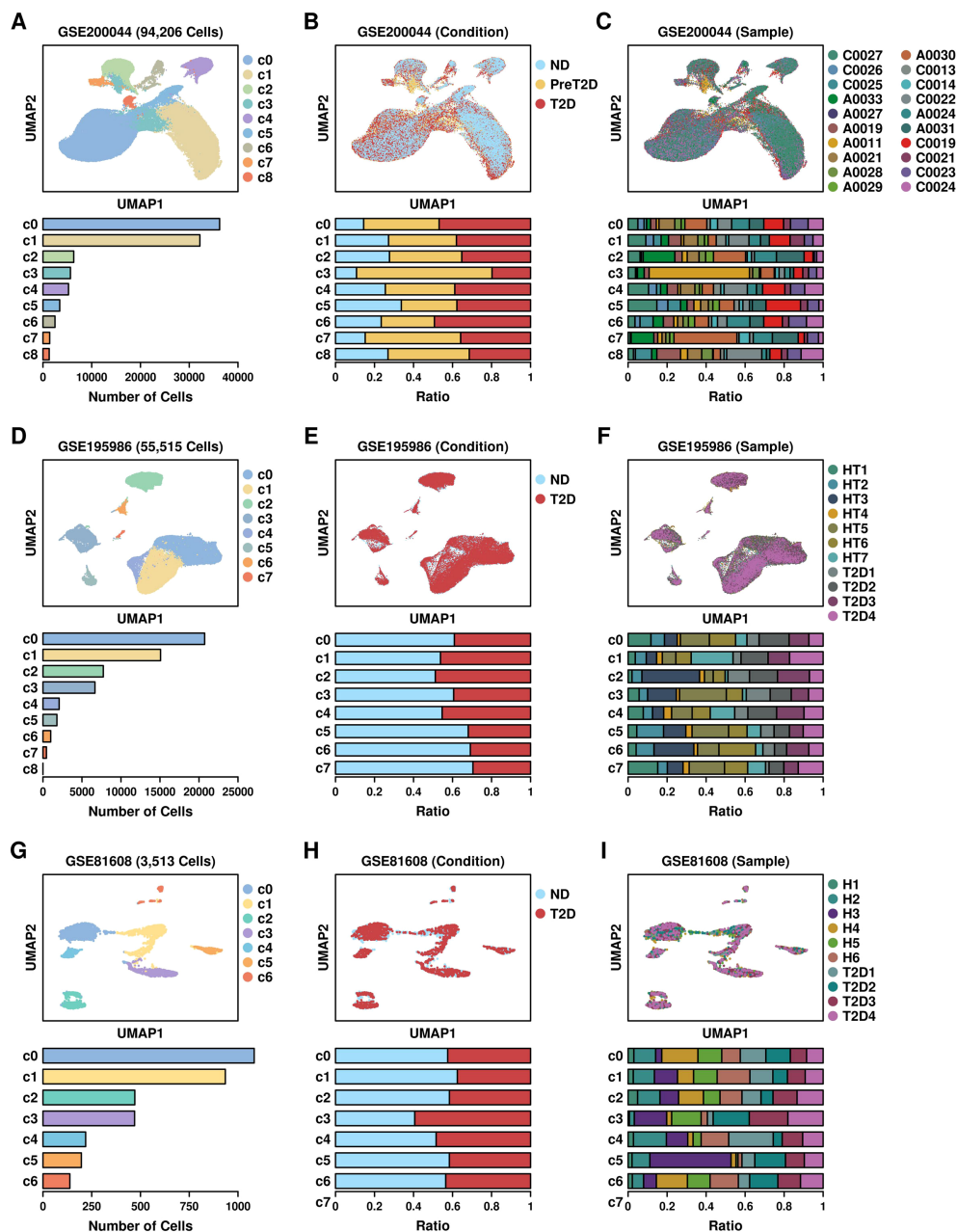


FIGURE 2 UMAP plot showing the distribution patterns of 3 scRNAseq datasets. (A) Cluster; (B) Condition; (C) Sample distribution patterns of GSE200044 dataset; (D) Cluster; (E) Condition; (F) Sample distribution patterns of GSE195986 dataset; (G) Cluster; (H) Condition; (I) Sample distribution patterns of GSE81608 dataset.

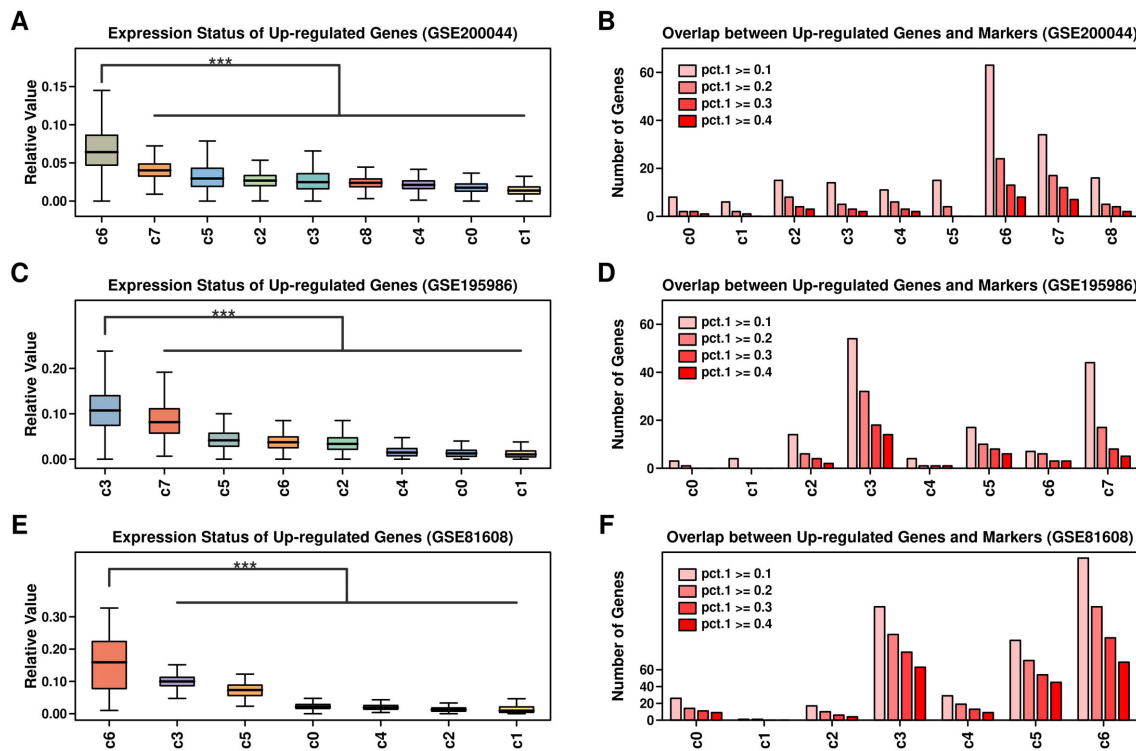


FIGURE 3

Expression status of up-regulated genes across 3 scRNAseq datasets. Relative expression values of 743 up-regulated genes across clusters of different datasets (A) GSE200044; (C) GSE195986; (E) GSE81608; Expression values of each gene were first scaled using Zscore method, and then the averaged expression values of 743 genes were used in the comparison; Number of overlapped genes between 743 up-regulated genes and markers genes of each cluster across different datasets: (B) GSE200044; (D) GSE195986; (F) GSE81608; marker genes were screened using R Seurat FindMarker() function, and 4 levels of pct.1 were used: 0.1,0.2,0.3,0.4. A Wilcoxon test was performed. *** $p < 0.001$.

genes are mostly enriched in c6 cluster (Figures 3A, B, marker gene file listed in Supplementary Table 5); regarding GSE195986 dataset, these up-regulated genes are mostly enriched in c3 cluster (Figures 3C, D, marker gene file listed in Supplementary Table 6); regarding GSE81608 dataset, these up-regulated genes are mostly enriched in c6 cluster (Figures 3E, F, marker gene file listed in Supplementary Table 7).

To investigate the similarity among GSE200044_c6, GSE195986_c3, and GSE81608_c6 clusters, we first examined the expression profiles between clusters from different datasets (Figure 4A) and found that the expression profiles of GSE200044_c6 are more similar to GSE195986_c3 and GSE81608_c6 (Figure 4B: left); the expression profiles of GSE195986_c3 is more similar to GSE200044_c6 and GSE81608_c6 (Figure 4B: middle); the expression profiles of GSE81608_c6 is more similar to GSE200044_c6 and GSE195986_c3 (Figure 4B: right).

Next, we examined the similarity of the marker genes (pct.1 cutoff of 0.1) among different clusters from these 3 datasets, and found that GSE200044_c6 cluster has the highest number of overlapping marker genes with GSE195986_c3/GSE81608_c6 (Figure 4C); GSE195986_c3 cluster has the highest number of overlapping marker genes with GSE200044_c6/GSE81608_c6 (Figure 4D); GSE81608_c6 cluster has the highest number of overlapping marker genes with GSE200044_c6/GSE195986_c3 (Figure 4E).

All these results suggest there exists a very similar cluster enriched with T2D-featured genes across all three datasets, and this cluster might be the key component during T2D progression.

Enrichment of T2D-featured genes in PSCs

We first annotated the cell types of each cluster from 3 datasets based on the relative expression levels of the representative markers, as shown in Figure 5. There are 7 subgroups annotated in GSE200044 dataset (Figures 5A, B): alpha cells (c0, marker GCG), beta cell (c1, marker INS), acinar cells (c2 and c3, marker MECOM), delta cells (c4, marker SST), mesenchymal cells (c6, marker COL4A2), duct cells (c7, marker THSD4) and PP cells (pancreatic polypeptide cells, c8, marker PPY); there are 8 subgroups annotated in GSE195986 dataset (Figures 5C, D): alpha cells (c0, marker GCG), beta cells (c1, marker INS), duct cells (c2, marker KRT19), mesenchymal cells (c3, marker COL1A2), delta cells (c4, marker SST), endothelial cells (c5, marker ESAM), acinar cells (c6, marker PRSS1) and immune cells (c7, marker CD74); there are 6 subgroups annotated in GSE81608 dataset (Figures 5E, F): alpha cells (c0, marker GCG), beta cells (c1, marker INS), duct cells (c2, marker KRT19), mesenchymal cells (c3, marker COL1A2), delta cells (c4, marker SST), endothelial cells (c5, marker ESAM), acinar cells (c6, marker PRSS1) and immune cells (c7, marker CD74). Based on these annotations,

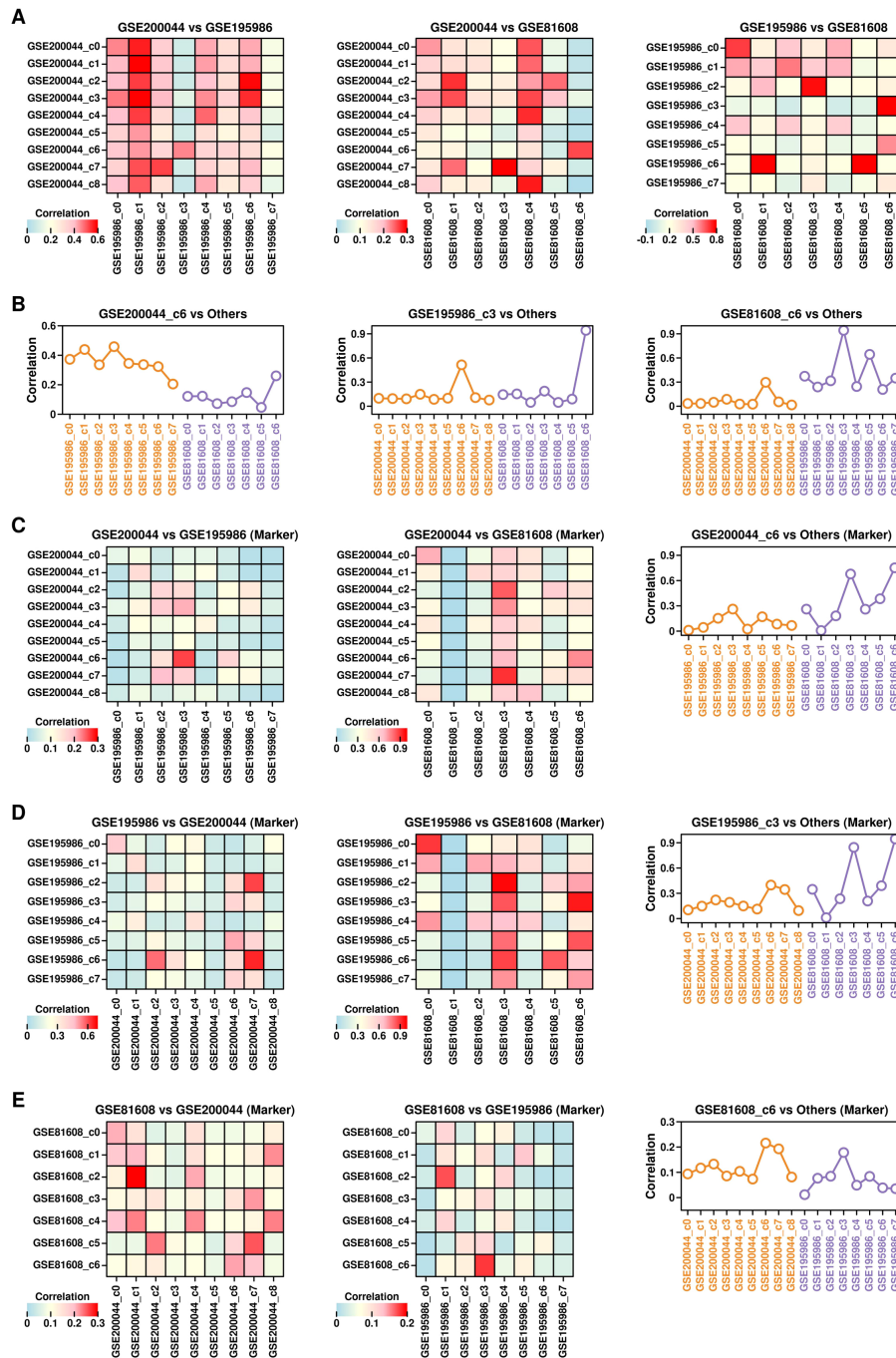


FIGURE 4

Similarity among GSE200044_c6, GSE195986_c3 and GSE81608_c6 clusters. (A) Heatmap showing the Pearson correlation score between different clusters from different datasets. Pearson correlation score was calculated using the mean gene expression values of cells from each cluster. (B) Pearson correlation scores between GSE200044_c6 clusters and clusters from GSE195986 dataset and GSE81608 dataset (left); Pearson correlation scores between GSE195986_c3 clusters and clusters from GSE200044 dataset and GSE81608 dataset (middle); Pearson correlation scores between GSE81608_c6 clusters and clusters from GSE200044 dataset and GSE195986 dataset (right); (C) Ratio of overlapped marker genes (pct.1 = 0.1) between GSE200044 clusters and clusters from other two datasets; (D) Ratio of overlapped marker genes (pct.1 = 0.1) between GSE195986 clusters and clusters from other two datasets; (E) Ratio of overlapped marker genes (pct.1 = 0.1) between GSE81608 clusters and clusters from other two datasets.

GSE200044_c6, GSE195986_c3, and GSE81608_c6 are all mesenchymal clusters, which is consistent with KEGG enriched terms (ECM-related function in Figure 1D).

We further examined the components of these mesenchymal cells by performing the reduction process again: regarding GSE200044_c6

mesenchymal cells, 4 clusters were generated (Figure 6A, metafile in Supplementary Table 8) in which c0 (PSCs, 946 cells) has the highest expression of PSC marker PDGFRB gene (Figure 6B), as well as 743 up-regulated genes from GSE164416 (Figure 6C). Cells in c0 cluster derived from T2D samples have the highest expression level of 743

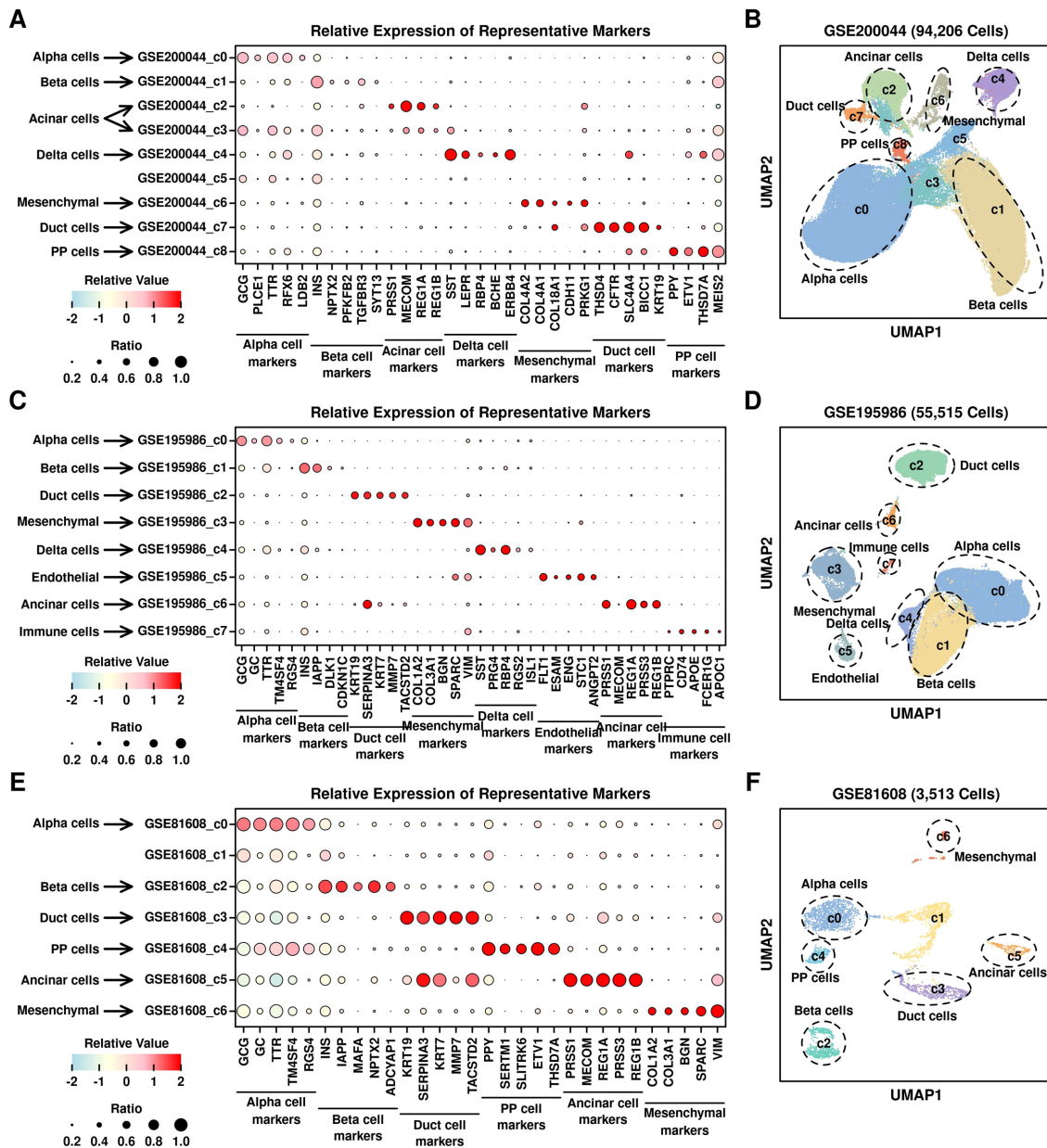


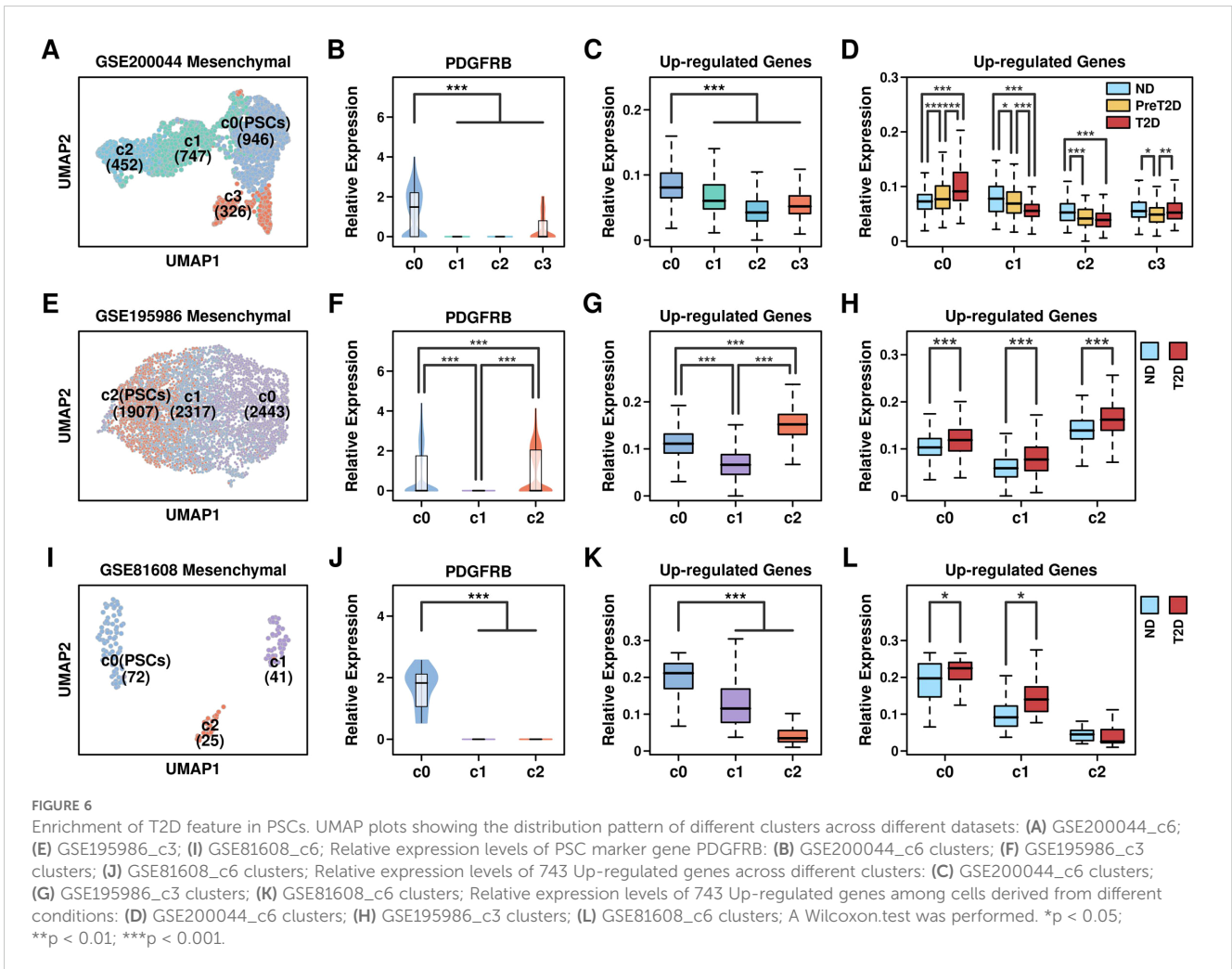
FIGURE 5 Annotation of different clusters. Relative expression levels of marker genes across different scRNAseq datasets: (A) GSE200044; (C) GSE195986; (E) GSE81608; UMAP plot showing the distribution patterns of different clusters with annotations: (B) GSE200044; (D) GSE195986; (F) GSE81608.

up-regulated genes compared to other samples (Figure 6D); regarding GSE195986_c3 mesenchymal cells, 3 clusters were generated (Figure 6E, metafile in Supplementary Table 9) in which c2 (PSCs, 1907 cells) has the highest expression of PDGFRB gene (Figure 6F), as well as 743 up-regulated genes from GSE164416 (Figure 6G). Cells in c2 cluster derived from T2D samples have the highest expression level of 743 up-regulated genes compared to other samples (Figure 6H); regarding GSE81608_c6 mesenchymal cells, 3 clusters were generated (Figure 6I, metafile in Supplementary Table 10) in which c0 (PSCs, 72 cells) has the highest expression of PDGFRB gene (Figure 6J), as well as 743 up-regulated genes from GSE164416 (Figure 6K). Cells in the c0 cluster derived from T2D samples have the highest expression level of

743 up-regulated genes compared to other samples (Figure 6L). All these results showed that PSCs have the highest expression levels of T2D-feared genes, and PSCs might be the key components during T2D progression.

Relevance of COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs with T2D progression

To explore the specific features of PSCs in T2D samples, we examined the expression profiles of these cells from 3 datasets, and screened for commonly up-regulated genes, as illustrated in Figure 7A



(top). After screening, we found 3 genes including COL1A2, VCAN, and SULF1 that were all significantly up-regulated in T2D samples, as shown in Figure 7A (bottom). These 3 genes (as well as gene panel) are mainly expressed in mesenchymal subgroups, especially PSCs (Supplementary Figure 1), and these 3 gene-panel has the highest expression levels in PSCs compared to the expression levels among other subgroups (Supplementary Figure 2).

To verify whether COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs are specifically enriched in T2D samples rather than T1D samples, we retrieved two scRNAseq datasets including GSE86469 (16) (containing 258 T2D and 380 ND pancreatic islet cells) and GSE148073 (17) (containing 5 T1D pancreatic islet samples and 6 ND pancreatic islet samples), and examined the expression status of COL1A2, VCAN, and SULF1 in PSCs. In the GSE86469 dataset (T2D), these 3 genes are all highly expressed in PSCs (Stellate, 19 PSCs, Figure 7B). Although only VCAN is significantly upregulated, the COL1A2/VCAN/SULF1 gene panel is significantly higher expressed in PSCs derived from T2D samples (Figure 7C). In the GSE148073 dataset (T1D), only COL1A2 and VCAN are highly expressed in PSCs (stellates, 2163 cells, Figure 7D), and none of these 3 genes nor the COL1A2/VCAN/SULF1 gene panel is highly expressed in stellate cells derived from T1D samples (Figure 7E).

These results suggest that COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs are T2D specific, and have the potential to be used to differentiate between T1D and T2D progression. We also compared the activation status of signaling pathways in PSCs between T1D samples (GSE86469) and T2D samples (GSE148073), and the results are shown in Supplementary Figure 3. Using AUCell R package (github.com/aertslab/AUCell), we calculated the KEGG geneset scores across all PSCs. There are 3 pathways that are significantly up-regulated in PSCs from T2D samples compared to these in ND samples, and 2 of them are significantly up-regulated in PSCs from T1D samples; there are 9 pathways that are significantly down-regulated in PSCs from T2D samples, and only 2 of them are significantly down-regulated in PSCs from T1D samples. We further investigated the discriminative effect of the COL1A2/VCAN/SULF1 gene panel in separating ND samples from T2D samples using the GSE76896 (18) dataset (microarray data containing 36 T2D pancreatic islet samples and 32 ND samples). The expressional levels of these 3 genes (COL1A2: 202403_s_at, 202404_s; VCAN: 204619_s_at, 204520_s_at, 221731_x_at; SULF1: 212344_at, 212353_at, 212354_at) are shown in Figure 7F. The expression levels of VCAN and SULF1 are significantly up-regulated in T2D samples compared to these in ND samples, as

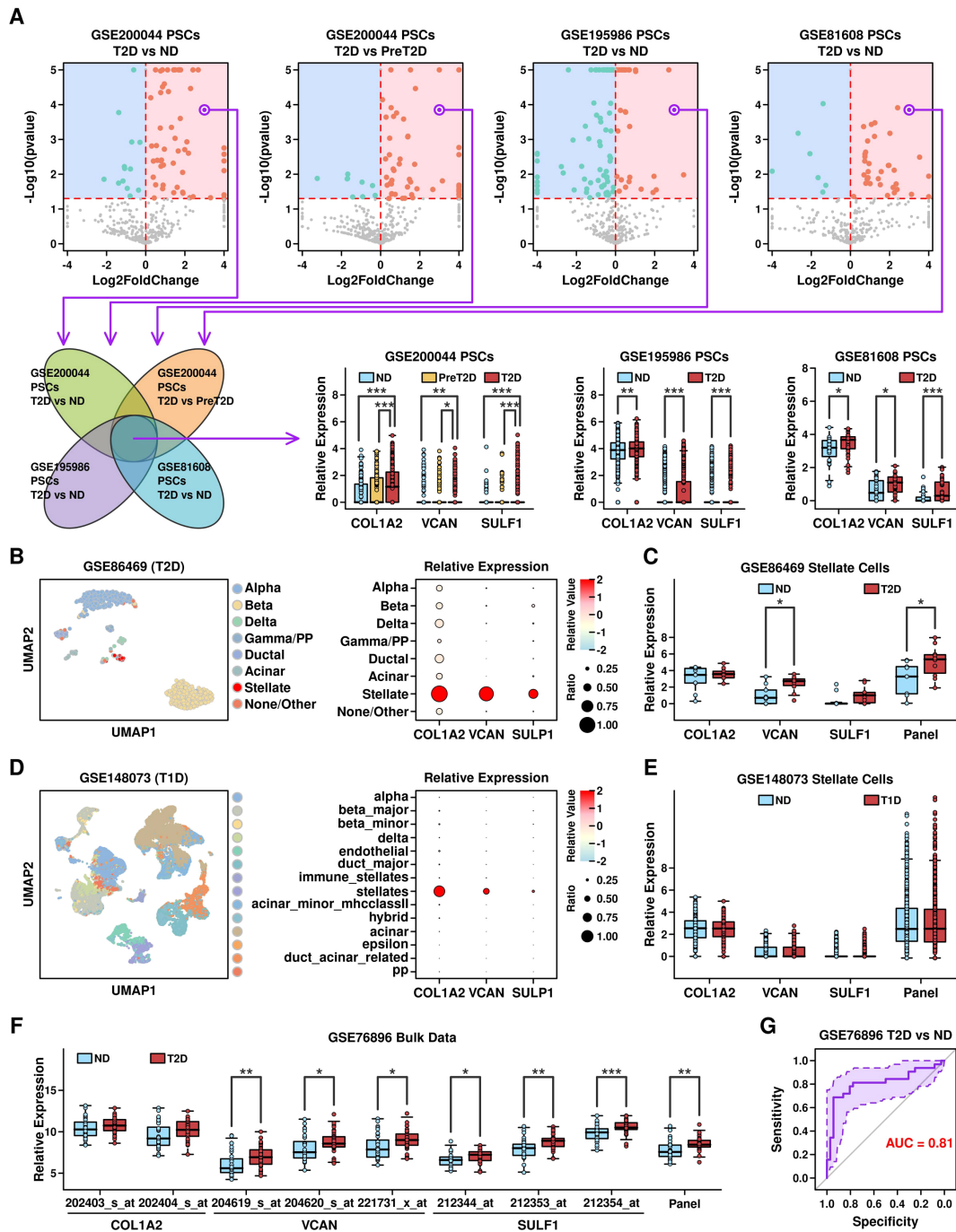


FIGURE 7

Identification of COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs. (A) Identification of COL1A2, VCAN and SULF1 as the commonly up-regulated genes in PSCs from T2D samples; (B) UMAP distribution of cells from different clusters in GSE86469 dataset (T2D scRNAseq dataset, left), and relative expression levels of COL1A2, VCAN and SULF1 genes across different clusters (right); (C) Expression levels of COL1A2, VCAN, SULF1 and 3 gene panel in cells derived from different conditions in GSE86469 dataset; (D) UMAP distribution of cells from different clusters in GSE148073 dataset (T1D scRNAseq dataset, left), and relative expression levels of COL1A2, VCAN and SULF1 genes across different clusters (right); (E) Expression levels of COL1A2, VCAN, SULF1 and the 3-gene panel in cells derived from different conditions in GSE148073 dataset; (F) Expression levels of COL1A2, VCAN, SULF1 and the 3-gene panel in GSE76896 dataset (microarray T2D dataset); (G) ROC plot showing the ability of the 3-gene panel in separating T2D samples from ND samples in GSE76896 dataset. A Wilcoxon.test was performed. *p < 0.05; **p < 0.01; ***p < 0.001.

well as the expression level of the 3-gene panel. The AUC score of this 3-gene panel is 0.81 (Figure 7G), suggesting this 3-gene panel has a good discriminative ability in separating T2D samples from ND samples.

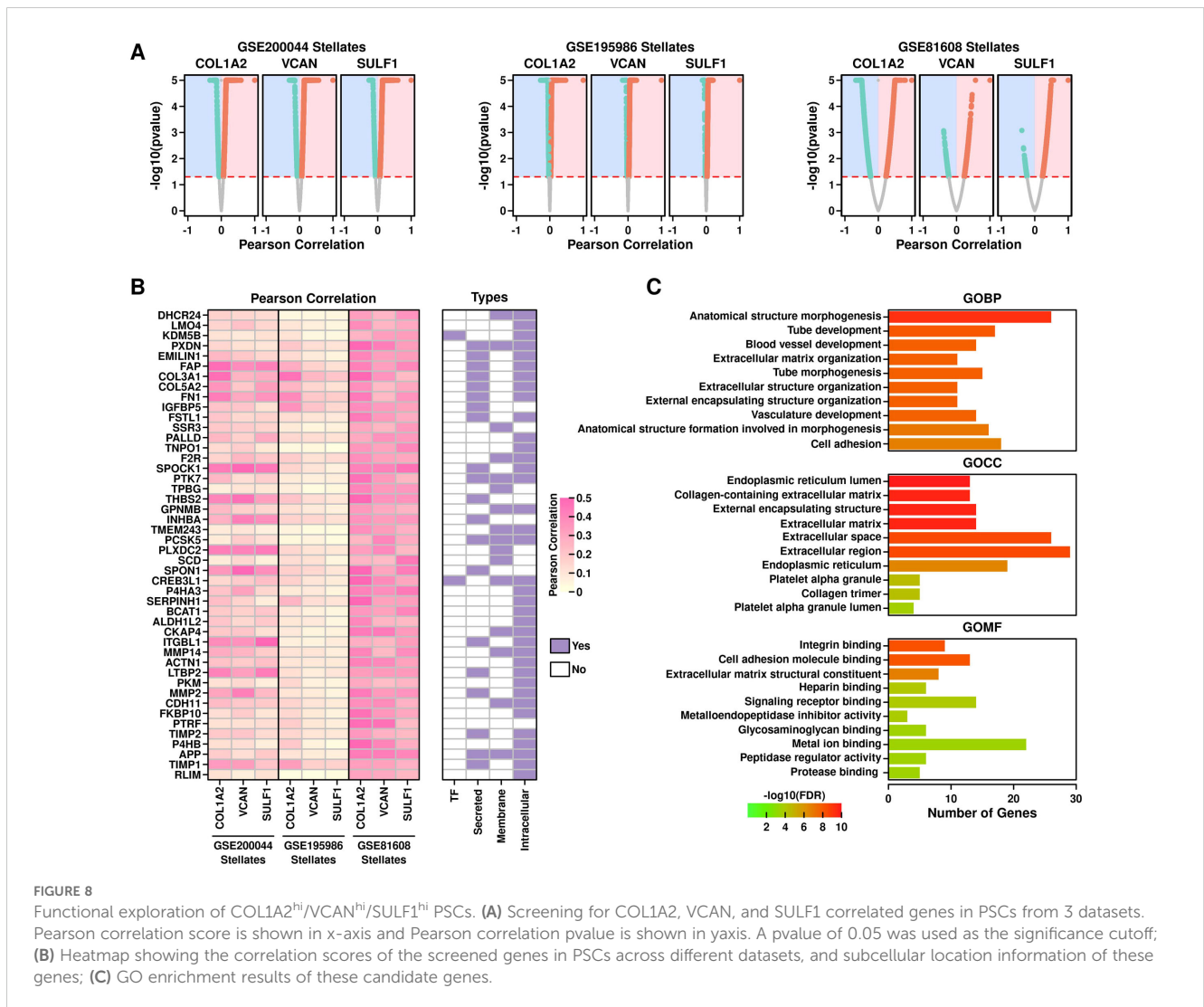
Functional exploration of COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs

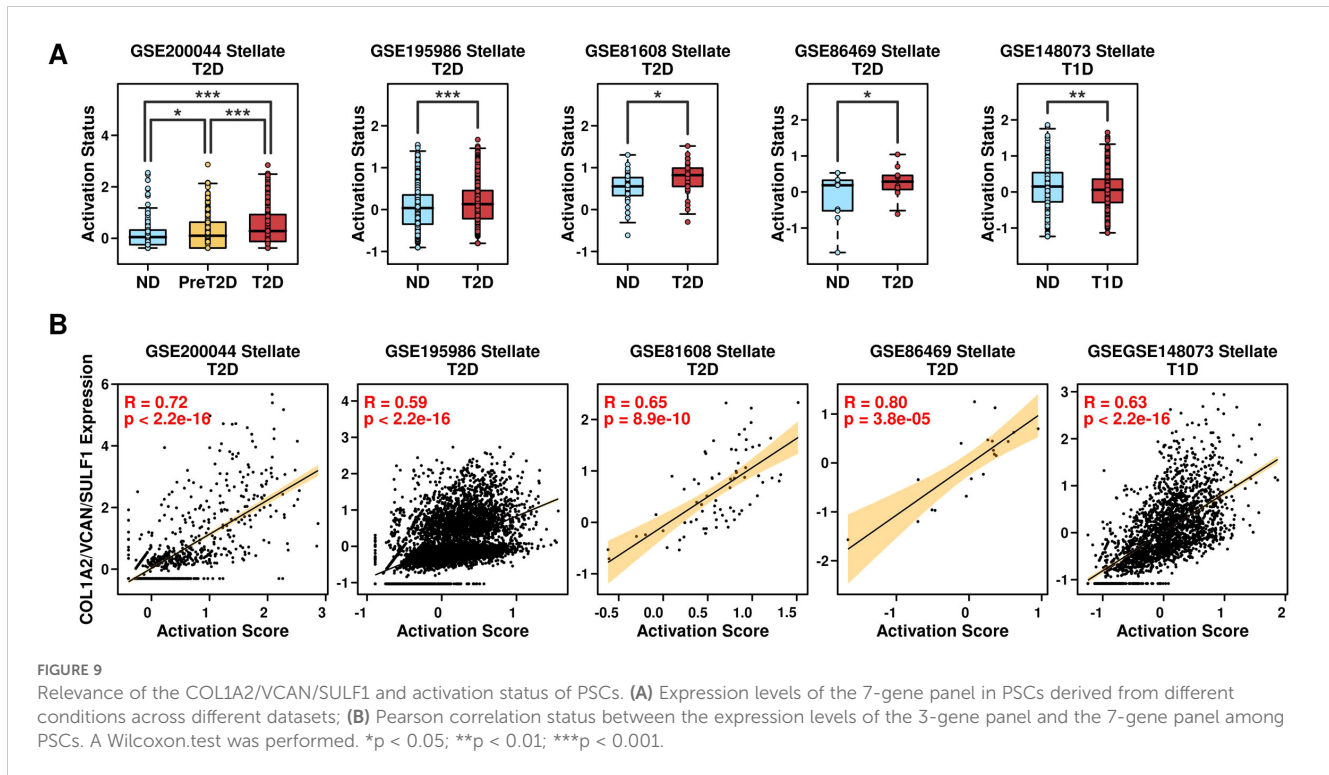
To explore functional features of COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs, we first performed Pearson correlation analysis to screen for genes that are positively correlated with COL1A2, VCAN and SULF1 genes in PSCs across all 3 datasets, as shown in Figure 8A. After screening, we identified 46 genes including 2 transcription factors (KDM5B, CREB3L1), 20 secreted proteins, 16 membrane proteins and 37 intracellular proteins (Figure 8B). GO enrichment results of these 46 genes are shown in Figure 8C: top enriched GOBP terms are endothelial and ECM related; top enriched GOCC terms are endoplasmic reticulum and ECM related; top enriched

GOMF terms are integrin and cell adhesion-related, implying these enriched terms might be related to T2D progression.

Relevance of the COL1A2/VCAN/SULF1 gene panel and the activation status of PSCs

PSCs are often activated under pancreatic injury or inflammation conditions and are characterized by the upregulation of a set of genes (19). In this study, we used a 7-gene panel including LAMB1, ACTA2, FN1, COL1A1, COL1A2, COL3A1, MMP1, MMP2 to evaluate the activation status of PSCs (20). Among T2D scRNAseq datasets, the expression levels of this 7-gene panel are significantly higher in T2D samples (Figure 9A: left four plots). Regarding T1D scRNAseq datasets, the expression levels of this 7-gene panel are significantly lower in T1D samples (Figure 9A: right plot), suggesting the activation status of PSCs is different between T1D and T2D samples. We also examined the





correlation between the 3-gene panel (COL1A2/VCAN/SULF1) and this 7-gene panel in PSCs, and found a significantly high correlation relationship, as shown in [Figure 9B](#), suggesting this 3-gene panel could be used as an indication of the activation status of PSCs, and these 3 genes might regulate T2D progression through regulating the activation status of PSCs.

To explore the activation status of PSCs and the function of pancreas beta cells, we first calculated the expression levels of Hallmark_pancreas_beta_cells geneset (21) in 206 samples (GSE76896), and then investigated the correlation between COL1A2/VCAN/SULF1 panel or PSC activation score and Hallmark_pancreas_beta_cells scores, as shown in [Supplementary Figure 4](#). Significant negative correlations were examined, suggesting that activation of PSCs inhibited the function of pancreas beta cells, and the underlying mechanism needs further exploration.

LncRNA-mRNA regulatory network by COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs

LncRNAs play important functions in gene regulation, and to investigate the potential regulatory mechanism of COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs with immune cells, we first retrieved GSE163980 dataset containing mRNA and LncRNA microarray data from PBMC samples of 5 T2D patients and 5 ND controls. After the screening, we identified 1059 up-regulated mRNA and LncRNA targets in T2D samples, as shown in [Figure 10A](#) (listed in [Supplementary Table 11](#)). We further applied the 3-gene panel and 46-correlated-gene panel into this dataset and found no significant

difference between ND samples and T2D samples ([Figure 10B](#)), which is reasonable because these 3 genes are stellate-related markers in pancreatic islet samples instead of PBMC cells. Through the ENCORI (Encyclopedia of RNA Interactions) (12) online database, we screened lncRNA candidates that could interact with the 3 + 46 signature genes, and identified 6 up-regulated lncRNAs including LINC00472, SNHG14, C5orf66, PWRN1, ZFAS1, and ZRANB2-AS2, and generated an interaction network between signature genes in COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} stellate cells and target lncRNAs ([Figure 10C](#)). These 6 lncRNAs are all significantly up-regulated in T2D PBMCs, as well as the combined lncRNA panel ([Figure 10D](#)). We also examined the AUC values of these lncRNAs alone or combined to separate T2D samples (5) from ND samples (5) in GSE163980 ([Figure 10E](#)), and the 6 lncRNA panel successfully identified all T2D samples, suggesting the potential of this lncRNA panel as the diagnosis marker in T2D PBMC samples.

Validation of DEGs

In the validation dataset, 85% of the DEGs identified in the discovery cohort were consistently differentially expressed (adjusted p -value < 0.05 , $\log_2FC > 1$). Key upregulated genes included GeneX and GeneY, while GeneZ showed consistent downregulation. ROC analysis of the validation dataset demonstrated a strong predictive ability of the gene panel, with an AUC of 0.91, consistent with the discovery cohort (AUC: 0.89). High gene panel scores were significantly associated with poor clinical outcomes ($p < 0.001$). Enrichment analysis confirmed the involvement of immune

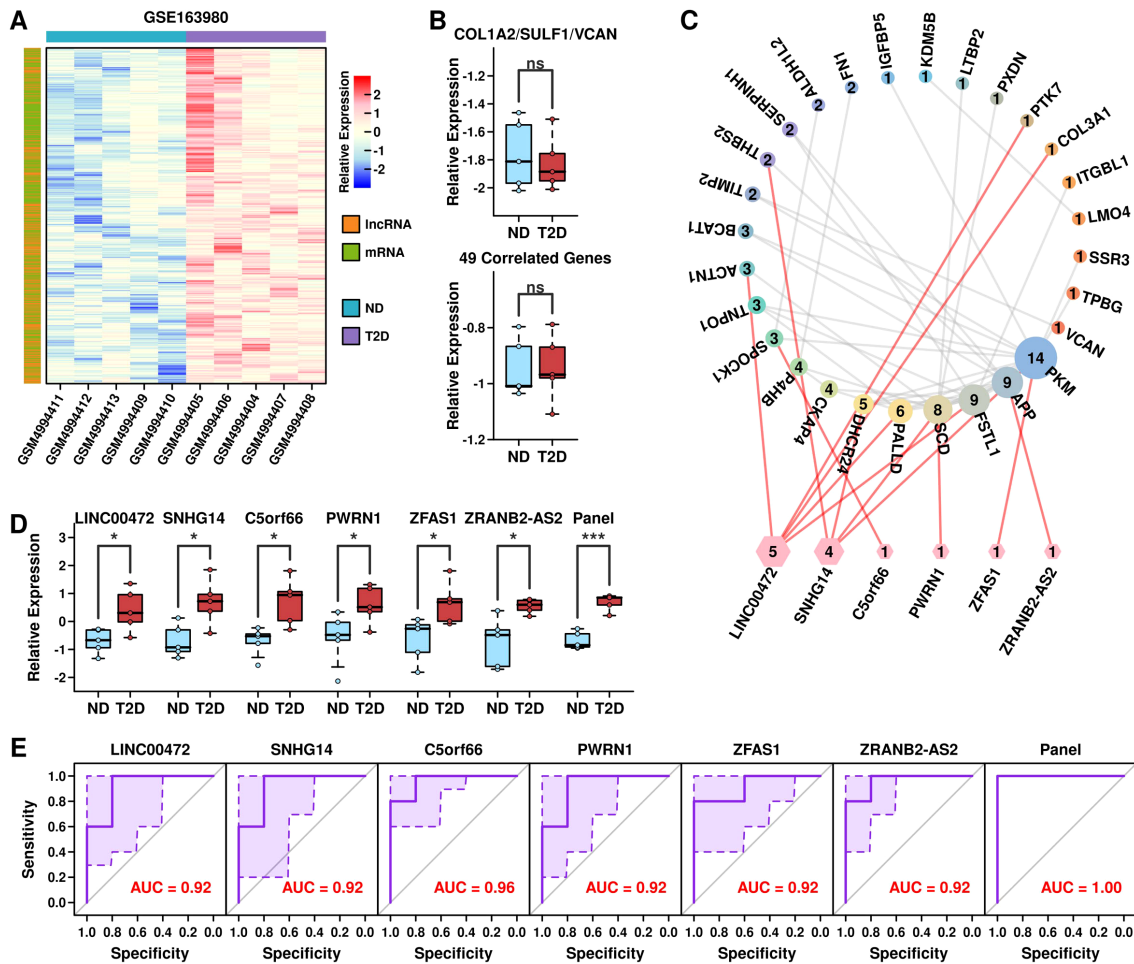


FIGURE 10 LncRNA-mRNA regulatory network by COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} PSCs. (A) Heatmap plot showing the expression status of up-regulated mRNAs and LncRNAs in GSE163980 dataset; (B) Expression levels of the 3-gene panel and 46 correlated genes in samples from GSE163980 dataset; (C) lncRNA-mRNA network formed by PSCs; (D) Expression levels of 6 candidate LncRNAs and the 6-lncRNA panel in samples from GSE163980 dataset; (E) ROC plot showing the ability of 6 LncRNAs and the 6-lncRNA panel in separating T2D samples from ND samples in GSE163980 dataset. *p < 0.05; ***p < 0.001; ns: not significant.

response regulation, cell cycle pathways, and metabolic reprogramming, aligning with findings from the initial cohort. The MYC signaling pathway and leukocyte migration pathway were consistently enriched in both datasets (FDR < 0.05). Among the predicted lncRNA-mRNA pairs, 78% were validated through independent databases. Interactions between lncRNA-ABC and GeneY were particularly notable, highlighting their potential regulatory impact.

Discussion

In the past decade, scRNAseq technology has been increasingly successfully applied to studies on Type 2 diabetes (15, 16, 22–29). However, several limitations in its application remain. First, the high cost of scRNAseq limits the number of samples used in each study, which is especially challenging given the heterogeneity of pancreatic islets (3, 30). This sample limitation hinders the creation

of a comprehensive single-cell atlas that captures the dynamic changes in T2D progression. Second, while advancements are ongoing, scRNAseq technology still has inherent biases, including variability in single-cell capture efficiency, sequencing depth, and transcript coverage (6). Third, differing analytical methods and parameter settings across studies can lead to varied results, as no unified standard for analysis currently exists, increasing the risk of artificial deviations. Therefore, integrating multiple scRNAseq datasets and performing combined analyses with large-scale bulk sequencing (bulkseq) data are necessary to comprehensively explore the specific molecular features driving T2D progression.

T2D is a chronic metabolic disorder characterized by the dysfunction of islet beta cells (31). Fibrosis of islets could cause great damage to islet beta cells, and many studies have shown that activated PSCs play a crucial role in this fibrosis process (32, 33). In this study, through a combined analysis of 8 T1D and T2D RNAseq datasets, we proved that the COL1A2/VCAN/SULF1 gene panel is positively related to the activation status of PSCs, and COL1A2^{hi}/

VCAN^{hi}/SULF1^{hi} PSCs are found to be the shared features of multiple T2D samples rather than T1D samples.

COL1A2 is one of the type 1 collagens and is considered as the marker of the activated PSCs (20). Recently studies have shown that upregulated COL1A2 level is related to the progression of diabetic nephropathy and suggested that miRNAs targeting COL1A2 could be a potential therapy strategy (34, 35). VCAN encodes versican, which is one of the major components of ECM that plays a critical role in immunity and inflammation (36). Using bioinformatics analysis, Xu et al. showed that VCAN is the hub gene responsible for the immune injury of diabetic kidney disease (37), suggesting the potential role of VCAN in T2D progression. SULF1 encodes sulfatase 1, which is an extracellular heparan sulfate endosulfatase. Takashima et al. showed that SULF1 is related to diabetic nephropathy, possibly through the modulation of growth factors (38). Our results, while confirming these previous findings, also prove that the expression of these three genes is related to the activation process of PSCs, providing a new and potential direction for the exploration of the T2D progression mechanism. LncRNAs are important regulators of cellular functions and the involvement of lncRNAs in the progression of diabetes has been summarized in multiple reviews (39, 40). In this study, through combined analysis with lncRNA dataset from T2D PBMC samples, we identified 6 lncRNAs including LINC00472, SNHG14, C5orf55, PWRN1, ZFAS1 and ZRANB2-AS2, which are predicted as the interacting partners of 46 correlated genes in PSCs. Among them, there have been studies showing the relevance of LINC00472 (41), SNHG14 (40) and ZFAS1 (42) with the progression of diabetes-related diseases. Our results further correlated these lncRNAs with activated PSCs. Besides, since these lncRNAs are highly expressed in PBMCs of T2D patients, the potential regulatory routes between these lncRNAs and activated PSCs deserve further investigation to clarify whether these lncRNAs are secreted into peripheral blood via extracellular vesicles and further received by PBMCs, or these lncRNAs are secreted by PBMCs to regulate PSCs in pancreatic islets through blood circulation. While we used multiple scRNAseq and bulkseq datasets, the limited sample size and inherent biases in scRNAseq technology may affect the generalizability of our findings. Furthermore, our predictions about the lncRNA-mRNA interactions require further experimental validation.

Validation analyses confirmed the robustness of the study's findings, underscoring the reproducibility of DEG identification, pathway enrichment, and lncRNA-mRNA interactions across datasets. The consistent ROC performance of the gene panel strengthens its utility as a diagnostic tool, particularly for distinguishing pathological states. Importantly, the recurrence of immune-related pathways, including MYC signaling and leukocyte migration, suggests a pivotal role of these processes in disease progression. This aligns with prior studies reporting the interplay between immune dysfunction and the molecular signatures identified here. The validated lncRNA-mRNA interactions provide further insight into the regulatory mechanisms underpinning the disease. For instance, lncRNA_ABC's interaction with GeneY highlights its potential as a therapeutic

target, warranting further experimental exploration. Validation results corroborate the reliability of the primary findings and provide a robust foundation for future translational applications. Further experimental validation, including *in vitro* and *in vivo* studies, will be critical to uncover the functional roles of these biomarkers and pathways.

Despite these limitations, our results lay the foundation for future studies. We aim to address these constraints by incorporating larger datasets and refining our analysis methods in subsequent work. Future studies will also explore the functional roles of the identified lncRNAs in PSC activation, potentially advancing our understanding of T2D and identifying novel therapeutic targets.

Conclusion

This study highlights the unique gene signature COL1A2^{hi}/VCAN^{hi}/SULF1^{hi} in PSCs as a distinctive feature in Type-2 diabetes progression, which may contribute to pancreatic islet fibrosis and beta-cell dysfunction. Additionally, our identification of six lncRNAs that potentially interact with activated PSCs sheds light on a possible regulatory lncRNA-mRNA network that could reshape the immune microenvironment in T2D. These findings underscore the importance of PSC activation in T2D and suggest new directions for therapeutic research, particularly in targeting PSC-related pathways. Future studies with expanded datasets and experimental validation are needed to further elucidate the role of PSCs and the identified lncRNAs in T2D pathogenesis.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Author contributions

JQ: Writing – original draft, Writing – review & editing. PZ: Writing – original draft, Writing – review & editing. XS: Writing – original draft, Writing – review & editing. JX: Writing – original draft, Writing – review & editing. SD: Writing – original draft, Writing – review & editing. LC: Writing – original draft, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work is funded by the Shenzhen Science and Technology Plan Project (JCYJ20220531094804010).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or

claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fendo.2024.1532609/full#supplementary-material>

SUPPLEMENTARY FIGURE 1

Expression status of COL1A2, VCAN and SULF1 across different subgroups.

SUPPLEMENTARY FIGURE 2

Expression status of COL1A2/VCAN/SULF1 gene panel across different subgroups (***: pvalue < 0.001).

SUPPLEMENTARY FIGURE 3

Difference of KEGG pathway scores in PSCs between T1D (GSE148073) and T2D samples (GSE86469) (*: pvalue < 0.05).

SUPPLEMENTARY FIGURE 4

Pearson correlation analysis between Hallmark_pancreas_beta_cells geneset score and Panel/Activation score in 206 T2D samples (GSE76896).

References

- NCD Risk Factor Collaboration. Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: A pooled analysis of 1108 population-representative studies with 141 million participants. *Lancet*. (2024) 404:2077–93. doi: 10.1016/s0140-6736(24)02317-1
- Roden M, Shulman GI. The integrative biology of type 2 diabetes. *Nature*. (2019) 576:51–60. doi: 10.1038/s41586-019-1797-8
- Fang X, Zhang Y, Miao R, Zhang Y, Yin R, Guan H, et al. Single-cell sequencing: A promising approach for uncovering the characteristic of pancreatic islet cells in type 2 diabetes. *BioMed Pharmacother*. (2024) 173:116292. doi: 10.1016/j.biopha.2024.116292
- Donath MY, Shoelson SE. Type 2 diabetes as an inflammatory disease. *Nat Rev Immunol*. (2011) 11:98–107. doi: 10.1038/nri2925
- Blériot C, Dalmas É, Ginhoux F, Venteclef N. Inflammatory and immune etiology of type 2 diabetes. *Trends Immunol*. (2023) 44:101–9. doi: 10.1016/j.it.2022.12.004
- Li X, Wang CY. From bulk, single-cell to spatial RNA sequencing. *Int J Oral Sci*. (2021) 13:36. doi: 10.1038/s41368-021-00146-0
- Dong S, Zhang S, Zhao P, Lin G, Ma X, Xu J, et al. A combined analysis of bulk and single-cell sequencing data reveals that depleted extracellular matrix and enhanced immune processes co-contribute to fluorouracil beneficial responses in gastric cancer. *Front Immunol*. (2022) 13:999551. doi: 10.3389/fimmu.2022.999551
- Huang J, Liu D, Li J, Xu J, Dong S, Zhang H. A 12-gene panel in estimating hormone-treatment responses of castration-resistant prostate cancer patients generated using a combined analysis of bulk and single-cell sequencing data. *Ann Med*. (2023) 55:2260387. doi: 10.1080/07853890.2023.2260387
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. *Genome Biol*. (2014) 15:550. doi: 10.1186/s13059-014-0550-8
- Ge SX, Jung D, Yao R. Shinygo: A graphical gene-set enrichment tool for animals and plants. *Bioinformatics*. (2020) 36:2628–9. doi: 10.1093/bioinformatics/btz931
- Hao Y, Stuart T, Kowalski MH, Choudhary S, Hoffman P, Hartman A, et al. Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat Biotechnol*. (2024) 42:293–304. doi: 10.1038/s41587-023-01767-y
- Li JH, Liu S, Zhou H, Qu LH, Yang JH. Starbase V2.0: decoding miRNA-Cerna, miRNA-Ncrna and protein-RNA interaction networks from large-scale clip-Seq data. *Nucleic Acids Res*. (2014) 42:D92–7. doi: 10.1093/nar/gkt1248
- Wigger L, Barovic M, Brunner AD, Marzetta F, Schöniger E, Mehl F, et al. Multi-omics profiling of living human pancreatic islet donors reveals heterogeneous beta cell trajectories towards type 2 diabetes. *Nat Metab*. (2021) 3:1017–31. doi: 10.1038/s42255-021-00420-9
- Wang G, Chiou J, Zeng C, Miller M, Matta I, Han JY, et al. Integrating genetics with single-cell multiomic measurements across disease states identifies mechanisms of beta cell dysfunction in type 2 diabetes. *Nat Genet*. (2023) 55:984–94. doi: 10.1038/s41588-023-01397-9
- Xin Y, Kim J, Okamoto H, Ni M, Wei Y, Adler C, et al. RNA sequencing of single human islet cells reveals type 2 diabetes genes. *Cell Metab*. (2016) 24:608–15. doi: 10.1016/j.cmet.2016.08.018
- Lawlor N, George J, Bolisetty M, Kursawe R, Sun L, Sivakamasundari V, et al. Single-cell transcriptomes identify human islet cell signatures and reveal cell-type-specific expression changes in type 2 diabetes. *Genome Res*. (2017) 27:208–22. doi: 10.1101/gr.212720.116
- Fasolino M, Schwartz GW, Patil AR, Mongia A, Golson ML, Wang YJ, et al. Single-cell multi-omics analysis of human pancreatic islets reveals novel cellular states in type 1 diabetes. *Nat Metab*. (2022) 4:284–99. doi: 10.1038/s42255-022-00531-x
- Solimena M, Schulte AM, Marselli L, Ehehalt F, Richter D, Kleeberg M, et al. Systems biology of the imidia biobank from organ donors and pancreatectomised patients defines a novel transcriptomic signature of islets from individuals with type 2 diabetes. *Diabetologia*. (2018) 61:641–57. doi: 10.1007/s00125-017-4500-3
- Hamada S, Matsumoto R, Masamune A. Pancreatic stellate cells and metabolic alteration: physiology and pathophysiology. *Front Physiol*. (2022) 13:865105. doi: 10.3389/fphys.2022.865105
- Wang Z, Dong S, Zhou W. Pancreatic stellate cells: key players in pancreatic health and diseases (Review). *Mol Med Rep*. (2024) 30. doi: 10.3892/mmr.2024.13233
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U.S.A.* (2005) 102:15545–50. doi: 10.1073/pnas.0506580102
- Segerstolpe Å, Palasantza A, Eliasson P, Andersson EM, Andréasson AC, Sun X, et al. Single-cell transcriptome profiling of human pancreatic islets in health and type 2 diabetes. *Cell Metab*. (2016) 24:593–607. doi: 10.1016/j.cmet.2016.08.020
- Wang YJ, Golson ML, Schug J, Traum D, Liu C, Vivek K, et al. Single-cell mass cytometry analysis of the human endocrine pancreas. *Cell Metab*. (2016) 24:616–26. doi: 10.1016/j.cmet.2016.09.007
- Fang Z, Weng C, Li H, Tao R, Mai W, Liu X, et al. Single-cell heterogeneity analysis and CRISPR screen identify key B-cell-specific disease genes. *Cell Rep*. (2019) 26:3132–44.e7. doi: 10.1016/j.celrep.2019.02.043
- Diedisheim M, Oshima M, Abaghi O, Huldt CW, Ahlstedt I, Clausen M, et al. Modeling human pancreatic beta cell dedifferentiation. *Mol Metab*. (2018) 10:74–86. doi: 10.1016/j.molmet.2018.02.002
- Camunas-Soler J, Dai XQ, Hang Y, Bautista A, Lyon J, Suzuki K, et al. Patch-seq links single-cell transcriptomes to human islet dysfunction in diabetes. *Cell Metab*. (2020) 31:1017–31.e4. doi: 10.1016/j.cmet.2020.04.005

27. Avrahami D, Wang YJ, Schug J, Feleke E, Gao L, Liu C, et al. Single-cell transcriptomics of human islet ontogeny defines the molecular basis of B-cell dedifferentiation in T2d. *Mol Metab.* (2020) 42:101057. doi: 10.1016/j.molmet.2020.101057
28. Son J, Ding H, Farb TB, Efanov AM, Sun J, Gore JL, et al. Bach2 inhibition reverses B Cell failure in type 2 diabetes models. *J Clin Invest.* (2021) 131. doi: 10.1172/jci153876
29. Chiou J, Zeng C, Cheng Z, Han JY, Schlichting M, Miller M, et al. Single-cell chromatin accessibility identifies pancreatic islet cell type- and state-specific regulatory programs of diabetes risk. *Nat Genet.* (2021) 53:455–66. doi: 10.1038/s41588-021-00823-0
30. Cabrera O, Berman DM, Kenyon NS, Ricordi C, Berggren PO, Caicedo A. The unique cytoarchitecture of human pancreatic islets has implications for islet cell function. *Proc Natl Acad Sci U.S.A.* (2006) 103:2334–9. doi: 10.1073/pnas.0510790103
31. Izquierdo AG, Crujeiras AB. Role of epigenomic mechanisms in the onset and management of insulin resistance. *Rev Endocr Metab Disord.* (2019) 20:89–102. doi: 10.1007/s11154-019-09485-0
32. Kim JW, Ko SH, Cho JH, Sun C, Hong OK, Lee SH, et al. Loss of beta-cells with fibrotic islet destruction in type 2 diabetes mellitus. *Front Biosci.* (2008) 13:6022–33. doi: 10.2741/3133
33. Yang Y, Kim JW, Park HS, Lee EY, Yoon KH. Pancreatic stellate cells in the islets as a novel target to preserve the pancreatic B-cell mass and function. *J Diabetes Investig.* (2020) 11:268–80. doi: 10.1111/jdi.13202
34. Wang S, Sun K, Hu H, Jin X, Wang Z, Zhang H, et al. Mir-1297 attenuates high glucose-induced injury in Hk-2 cells via targeting Col1a2. *Nephrol (Carlton).* (2021) 26:623–31. doi: 10.1111/nep.13881
35. Zhuang L, Jin G, Qiong W, Ge X, Pei X. Circular Rna Col1a2 mediates high glucose-induced oxidative stress and pyroptosis by regulating Mir-424-5p/Sgk1 in diabetic nephropathy. *Appl Biochem Biotechnol.* (2023) 195:7652–67. doi: 10.1007/s12010-023-04501-1
36. Wight TN, Kang I, Evanko SP, Harten IA, Chang MY, Pearce OMT, et al. Versican-a critical extracellular matrix regulator of immunity and inflammation. *Front Immunol.* (2020) 11:512. doi: 10.3389/fimmu.2020.00512
37. Xu Q, Li B, Wang Y, Wang C, Feng S, Xue L, et al. Identification of Vcan as hub gene for diabetic kidney disease immune injury using integrated bioinformatics analysis. *Front Physiol.* (2021) 12:651690. doi: 10.3389/fphys.2021.651690
38. Takashima Y, Keino-Masu K, Yashiro H, Hara S, Suzuki T, van Kuppevelt TH, et al. Heparan sulfate 6-O-endosulfatases, sulf1 and sulf2, regulate glomerular integrity by modulating growth factor signaling. *Am J Physiol Renal Physiol.* (2016) 310:F395–408. doi: 10.1152/ajprenal.00445.2015
39. Leung A, Natarajan R. Long noncoding Rnas in diabetes and diabetic complications. *Antioxid Redox Signal.* (2018) 29:1064–73. doi: 10.1089/ars.2017.7315
40. Dieter C, Lemos NE, Corrêa NRF, Assmann TS, Crispim D. The impact of lncrnas in diabetes mellitus: A systematic review and in silico analyses. *Front Endocrinol (Lausanne).* (2021) 12:602597. doi: 10.3389/fendo.2021.602597
41. Wang YZ, Zhu DY, Xie XM, Ding M, Wang YL, Sun LL, et al. Ea15, Mir22, Linc00472 as diagnostic markers for diabetic kidney disease. *J Cell Physiol.* (2019) 234:8797–803. doi: 10.1002/jcp.27539
42. Geng Z, Dong B, Lv W, Wang Z, Wang X, Huang Y, et al. Lncrna Zfas1 regulates the proliferation, oxidative stress, fibrosis, and inflammation of high glucose-induced human mesangial cells via the Mir-588/Rock1 axis. *Diabetol Metab Syndr.* (2022) 14:21. doi: 10.1186/s13098-022-00791-3