Diabetes mellitus induces a novel inflammatory network involving cancer progression: Insights from bioinformatic analysis and in vitro validation

Yejun Tan  
Central South University

Jin Kang  
the Second Xiangya Hospital of Central South University

Hongli Li  
Central South University

Aifang Zhong  
Central South University

Yaqiong Liu  
National University of Ireland Galway

Zheyu Zhang  
Central South University

Roujie Huang  
Peking Union Medical College, Chinese Academy of Medical Science

Xin Cheng  
Central South University

Weijun Peng (pengweijun87@csu.edu.cn)  
Central South University

Research Article

Keywords: diabetes and cancer-related genes (DCRG), diabetes-based cancer-associated inflammation network (DCIN), diabetes mellitus, inflammation, cancer

Posted Date: November 11th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2244212/v1

License: Creative Commons Attribution 4.0 International License. Read Full License
Abstract

Background

The diabetes patients have a higher incidence of malignant tumors than people without diabetes. However, the molecular mechanisms of the relationship between diabetes and malignant tumors remain largely unknown.

Methods

By exploiting available public databases, diabetes and cancer-related genes (DCRGs) were screened, and a diabetes-based cancer-associated inflammation network (DCIN) was constructed. Then, the role of DCRGs in different tumors were analyzed from various perspectives. Additionally, drug sensitivity and single-cell sequencing data were analyzed using colon cancer (COAD) as an example. Finally, the expression of DCRGs and arachidonic acid metabolism pathway was verified in vitro.

Results

Seven identified DCRGs, including PPARG, MMP9, CTNNB1, TNF, TGFB1, PTGS2, and HIF1A, were integrated to construct a DCIN. The bioinformatics analysis showed that the expression of the seven DCRGs in different tumors was significantly different, which had varied effects on diverse perspectives. Single-cell sequencing analyzed in COAD showed that the activity of the DCRGs was highest in M1 macrophage and the lowest in Plasma B. In vitro experiments showed that the DCRGs verified by western bolt and PEG2 verified by ELISA were all highly expressed in COAD epithelial cells stimulated by high glucose.

Conclusion

This study, for the first time, constructed a DCIN, which provides novel insights into the underlying mechanism of how diabetes increases the occurrence and development of tumors. Although further research is required, our results offer clues for new potential therapeutic strategies to prevent and treat malignant tumors.

Introduction

Cancer is one of the major diseases threatening people's health worldwide [1], and 8–18% of cancer patients also have diabetes as a coexisting illness [2]. Diabetes Mellitus (DM) is associated with higher cancer risk, peaking approximately eight years after diagnosis [3]. Observational studies consistently indicate that people with DM have an increased risk for liver, pancreatic, endometrial, colorectal, breast, and bladder cancers [4, 5]. Between 1988–1994 and 2010–2015, cancer mortality rates dropped in the
United States but remained about 30% higher in adults with diabetes compared to those without diabetes [6]. This suggests that effective control of DM could decrease the healthcare burden and improve the quality of life [7]. Additionally, it has been reported that managing hyperglycemia and insulin resistance in patients with DM who are also suffering from cancer may improve their overall quality of life. Given the substantial medical burden of these diseases worldwide, understanding the association between cancer and DM might be essential to public health [8, 9].

Preliminary investigations to understand the mechanisms underlying the relationship between diabetes and cancer demonstrated the involvement of apoptotic and proliferation pathways [9, 10]. They showed the involvement of factors such as hyperglycemia, hyperinsulinemia, insulin-like growth factor 1 (IGF-1), oxidative stress, and sex hormones. The association between DM and cancer may be causal (resulting from hyperinsulinemia or hyperglycemia) or confounding (resulting from common risk factors such as obesity) [11]. Despite the prevalence of cancer and diabetes, the extent to which diabetes affects cancer remains unclear. Investigating the correlation between diabetes and cancer will contribute to a better understanding of cancer's pathogenesis, as well as provide a better reference for individualized treatment of diabetic tumors.

Emerging evidence suggests that chronic inflammation is a crucial link between DM and cancer [12]. Diabetes is associated with insulin secretory defects related to inflammation [13]. Chronic inflammation, characterized by high levels of oxidative stress and reactive oxygen species (ROS), activation of pro-inflammatory pathways, and abnormal adipokine production, may develop a microenvironment that promotes tumor cell growth, facilitates metastasis, increases angiogenesis, and impairs natural killer cells and macrophages. [14] A cross-talk between cancer and diabetes is influenced by oxidative stress. Hyperglycemia could increase the production of superoxides [15]. Additionally, insulin could stimulate the production of reactive oxygen species (ROS) [16]. Strong evidence suggests that oxidative stress influences the expression of several genes and signal transduction pathways that play an essential role in tumorigenesis [17]. As a result of cytokine-dependent activation of nuclear factor (NF)-kB pathways, ROS interferes with cell proliferation and apoptosis [18].

Studies show that NF-kB is hyperactivated in colorectal cancer, breast cancer, blood neoplasms, and pancreatic cancer cell lines. Inflammation facilitates the development and progression of tumors and their resistance to treatment. In contrast, acute inflammatory responses often stimulate dendritic cells (DCs) to mature and present antigens, resulting in anti-tumor immune responses [19][20].

To explore the interaction between diabetes and tumor, we screened seven diabetes and cancer-related genes (DCRGs) using public databases and constructed a diabetes-based cancer-associated inflammation network (DCIN). We analyzed the role of the seven DCRGs in cancer from different perspectives. Drug sensitivity analysis and single-cell sequencing data were analyzed using colon adenocarcinoma (COAD). The expression of DCRGs and arachidonic acid metabolism pathway was further verified in vitro. Therefore, for the first time, our study constructed a diabetes-based cancer-associated inflammation network, offering novel insights into the underlying mechanism of diabetes
increasing the occurrence and development of tumors. Although further research is still required, our results offer clues for new potential therapeutic strategies to prevent and treat malignant tumors.

Materials And Methods

Screening for genes associated with DM and various cancers

Firstly, DisGeNET (https://www.disgenet.org/) was used to explore genes related to DM, pancreatic cancer, liver cancer, and colon cancer [21]. These genes were filtered according to a correlation score greater than 0.1. The intersection of genes related to these four diseases was taken from DisGeNET. Similarly, GeneCards (https://www.genecards.org/) was also used to search for genes related to DM, pancreatic cancer, liver cancer, and colon cancer. These genes were filtered according to a correlation score greater than 1. The intersection of genes related to these four diseases was taken from GeneCards. The “ggplot2” software package was used to visualize the intersection in Venn diagrams.

Data Downloading And Preprocessing

Gene expression profiles, phenotypic information, and survival data of 33 TCGA tumor samples and adjacent tissues (11057 samples in total) were downloaded from the UCSC Xena database (http://xena.ucsc.edu/). The gene expression profiles were set in Fragments Per Kilobase Per Million (FPKM) and HTseq-counts format. Demographic, tumor information and follow-up data were also extracted from the database for all patients. Subsequently, the expression profiles of PPARG, MMP9, CTNNB1, TNF, TGFB1, PTGS2, and HIF1A were extracted from 33 TCGA tumor samples and adjacent tissues for further analysis.

Differential And Co-expression Analysis

For all the TCGA tumor types analyzed, the expression of the seven DCGRs between tumor and normal tissues was assessed using the Wilcox test and visualized using “ggplot2”. The differential expression of the seven DCGRs between tumor and normal tissues across cancers was presented as a log2 fold change in the heatmap. Co-expression between these seven DGRGs at the transcriptional level was analyzed by the “corrplot” software package. In addition, the STRING database (https://string-db.org/) was used to construct a protein-protein interaction network among these genes.

Clinical Correlation Analysis

Kaplan-Meier plots of DCRGs were generated using the R package to analyze the differences in overall survival outcomes between patients with high and low expression of the seven DCRGs. The phenotype and survival data of 33 TCGA cancer types obtained from the UCSC Xena database were analyzed. These
DCRGs were divided into high and low-expression groups for survival analysis according to the median expression level. The software packages “Survival” and “SurvMiner” were used to plot survival curves. In addition, the hazard ratios of DCRGs in each TCGA tumor type were obtained by Cox proportional hazards regression. Cox regression analysis was performed using “survival” and “Forestplot” software packages to determine the pan-cancer relationship between the seven DCRGs and survival. Furthermore, the expression of these seven DCRGs was evaluated in COAD patients with different stages. \( P \)-values < 0.05 were considered significant.

Gene Set Variation Analysis (Gsva) Of DcrGs

In order to investigate the potential pathways of these DCRGs, we performed GSVA enrichment analysis on the pathways of the seven DCRGs across cancers. GSVA gene sets were obtained from the MSigDB database (c2. Cp. Kegg. V7.1. Symbols. gmt). Firstly, the “GSVA” software package was used to generate GSVA scores for pan-cancer expression profiles of all seven genes. Then, the “limma” software package was used to analyze the differences between pan-cancer tumor tissues and paracancerous tissues. The pathways with \(| t\text{-value of GSVA score} | > 2\) were considered significantly enriched. Finally, the “ggplot2” software package was used to visualize the differences, and R software was used to count the significantly enriched pathways of each pathway across cancers.

Relationship Between DcrGs Expression And Tumor Immune Cell Infiltration

To better understand the relationship between DCRGs and immune cells, the relationship between the gene expression levels of 7 DCRGs and the infiltration levels of 26 immune-related cells were estimated. CIBERSORT (https://cibersort.stanford.edu/) was used to estimate the extent of immune cell infiltration across cancer samples. Finally, “ggplot2”, “ggpubr”, and “ggExtra” software packages were used to assess the correlation between the levels of the 7 DCRGs and each immune cell infiltration in cancer (\( P < 0.05 \) was considered significant).

Immune Subtype Analysis

Immuno-tumor microenvironment (TME) has therapeutic and prognostic significance in anti-tumor therapy. Studies have identified six immune subtypes of tumor types based on five representative immune signatures, which include wound healing (C1), IFN-\( \gamma \) dominant (C2), inflammatory (C3), lymphocyte depleted (C4), immunologically quiet (C5), and TGF-\( \beta \) dominant (C6) [21]. Differential expression analysis was performed using the Kruskal test to understand the mRNA expression levels of DCRGs in the six different immune subtypes of tumor types. Furthermore, the mRNA expression levels of the seven DCRGs were analyzed in COAD.
Stemness Indices And Tme Across Cancers

We evaluated the ESTIMATE (Estimation of STormal and Immune cells in MAalignant Tumor tissues using Expression data) in pan-cancer [22]. The ESTIMATE score is calculated based on gene expression characteristics and can reflect tumor purity with good prediction accuracy. Therefore, Spearman correlation analysis was performed between the expression levels of the seven DCRGs genes and matrix scores by “Estimate” and “limma” packages.

In order to further analyze the association between DCRGs and pan-cancer stemness features, the stemness index of tumor samples was calculated using a one-class logistic regression (OCLR) algorithm. Subsequently, the Spearman correlation analysis was performed based on gene expression and stemness score [23]. Here, two types of dryness indices were obtained: DNA methylation-based dryness indices (DNAss) and mRNA expression-based dryness indices (RNAss).

Mutations In The Seven Dcrgs Across Cancers

Mutations in the seven DCRGs across cancers

Mutations in the seven T2DM genes associated with multiple cancers in all tumor tissues were analyzed by the cBioPortal platform. cBioPortal (http://www.cbioportal.org/) is a tumor database with heredity and variation data, which can provide researchers with multidimensional visual data. Mutations in the seven genes related to multiple cancers in 2,922 samples of 38 cancers from the Pan-Cancer Analysis of whole Genomes (ICGC/TCGA, Nature 2020) were analyzed. The types and frequencies of mutations in the seven DCRG genes in all tumors were analyzed in “OncoPrint” and “CancerTypesSummary”. “OncoPrint” displayed the mutation, copy number, and expression of the target gene in all samples as a heat map. In addition, “CancerTypesSummary” showed the mutation rate of target genes in generalized cancers as a bar graph. Finally, the overall survival between variant and wild-type patients was compared.

Correlation Of The Expression Levels Of The Seven Dcrgs With Tmb And MsI

Tumor mutation burden (TMB) is a quantifiable biomarker of immune response that reflects the number of mutations in tumor cells [24]. Microsatellite instability (MSI) is caused by MMR deficiency and is associated with the prognosis of patients [25]. TMB and MSI are intrinsically related to immune checkpoint inhibitor sensitivity. We investigated whether there was a correlation between the expression levels of the seven DCRGs with TMB and MSI. A Perl script was used to calculate the TMB score, corrected by the total length of exons. MSI scores were determined for all samples based on somatic mutation data downloaded from TCGA. Spearman correlation coefficients were used to analyze the relationship of DCRGs expression with TMB and MSI. The result was displayed as a heat map generated by the “ggplot2” software package.
Correlation Analysis Between The Seven Dcrgs And Immune Checkpoint-related Genes In Coad

Due to the high frequency of mutations of the seven genes in COAD, it was selected for further analysis. R software was used to analyze the correlation between the seven DCRGs and 46 immune checkpoint-related genes in COAD. The result was also visualized as a heatmap by the “ggplot2” software package.

Pan-cancer Drug Sensitivity Analysis

CellMiner is a web-based tool containing genomic and pharmacological information for investigators to utilize transcript and drug response data from the NCI-60 cell line set compiled by the National Cancer Institute. RNA-seq spectrum data of the seven DCRGs and their pharmacological activity were collected from the CellMiner database (https://discover.nci.nih.gov/cellminer/). The “Impute” software package was used to preprocess the raw data. The correlation between the transcriptional expression of the seven DCRGs and compound sensitivity was investigated using Pearson correlation analysis. When the $P$-value was less than 0.05 and the correlation coefficient was greater than 0.3, the relevant DCRGs were considered sensitive to the corresponding chemotherapeutic drugs.

Analysis Of Single-cell Transcriptome Sequencing Data

Single-cell sequencing data of colon cancer were analyzed to further validate the function of the seven DCRGs in COAD. The single-cell sequencing dataset (GSE161277) was downloaded from Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) database [26]. Quality control and single-cell data analysis were performed using the R software package “Seurat”. The “Harmony” R software package was also used to integrate multiple samples. Principal component analysis (PCA) was used to reduce dimensionality, and then the UMAP function was used for visualization. The FindClusters function was performed to identify clusters of cells, and stromal and immune cells were annotated based on specific markers from previous studies [26]. Finally, the AddModuleScore function was used to measure the activity of the seven DCRGs in each cell cluster and tissue group.

Cell Culture

The SW480 cells were grown in Dulbecco’s modified Eagle’s medium (DMEM, GENVIEW, GD3103) supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (PS). The NCM460 cells were grown in 1640 medium (GENVIEW, GR3101) with 10% FBS and 1% PS (NCM Biotech, C100C5). Cells were grown at 37°C in an atmosphere of 5% CO2 and 95% relative humidity, and the medium was changed every 2 days. When cells reached approximately 90% confluency, they were detached with 0.1% trypsin-ethylendiaminetetraacetic acid (NCM Biotech, C100C1) and seeded in a 6-well plate. Subsequently, the cells were cultured with high glucose (HG, 50 mmol/l) or normal glucose (NG, 5.5 mmol/l) for 2 days at 37°C.
Western Blot Analysis

Western blot analysis
The total proteins of SW480 and NCM460 cells were prepared using RIPA buffer containing protease inhibitors and phosphatase inhibitors. A BCA protein assay kit was used to measure protein concentrations. 20 µg proteins were loaded per lane, separated by electrophoresis, and then transferred to polyvinylidene fluoride (PVDF) membranes (C3117, Millipore). The membrane was blocked and then incubated for 1 h with β-actin (1:100000; ABclonal, AC026), PPAR gamma (1:5000; Proteintech, 16643-1-AP), MMP9 (1:1000; Proteintech, 10375-2-AP), TNF alpha (1:3000; Proteintech, 60291-1-IG), TGF beta1 (1:2000; Proteintech, 21898-1-AP), COX2 (1: 500; Proteintech, 27308-1-AP), HIF1A (1:2500; Proteintech, 20960-1-AP) and Beta-catenin (1:10000; Proteintech, 51067-2-AP). Immunoblot analysis was performed with horseradish peroxidase (HRP)-conjugated anti-mouse antibodies or anti-rabbit antibodies (1:5000; ZSGB-BIO, ZB-5301, and ZB-5305) and developed with the ECL kit (Beyotime Biotechnology, P0018FM). The level of β-actin was used as a loading control, and the ratios of the gray value of the target protein bands to the gray value of the corresponding internal control bands were defined as the expression level of the target protein.

Enzyme-linked Immunosorbent Assay (Elisa)

After treatment, the concentration of Leukotriene B4 (LT-B4), Leukotriene C4 (LT-C4), and Prostaglandin E2 (PG-E2) was tested using an enzyme-linked immunosorbent assay (ELISA) kit. These ELISA kits included Human Leukotriene B4, LT-B4 ELISA Kit (CSB-E08033h), Human Prostaglandin E2, PG-E2 ELISA Kit (CSB-E07965h), and Leukotriene C4 Assay Kit (H556-1). In brief, cells were seeded into 6-well plates, followed by 2 days’ incubation at 37°C. Then, the cultivating supernatant was collected, and the consistency of these factors was determined by ELISA following the manufacturer's protocols.

Results

Differential expression analysis and co-expression analysis of DCRGs

First, we used DisGeNET to explore gene sets related to T2DM, pancreatic cancer, liver cancer, and colon cancer (r > 0.1) and obtained 9 intersection genes. Then we used GeneCards to search for T2DM, pancreatic cancer, liver cancer, and colon cancer-related genes (r > 0.1) and obtained 59 intersection genes. Finally, by taking the intersection of the nine genes from DisGeNET and 59 genes from gene cards, we obtained seven DCRGs: PPARG, MMP9, CTNNB1, TNF, TGFB1, PTGS2, HIF1A (Fig. 1A-C). The expression of DCRGs varied in different tumors. MMP9 showed high expression in all tumors except THYM, TNF showed high expression in most tumors, and PTGS2 showed low expression in most tumors. Meanwhile, the DCRGs exhibited no statistical differences in MESO and UVEM (Fig. 2A-G, sFig1). It is worth noting that the correlation analysis showed a correlation between HIF1A and PTGS2 (r = 0.36),
TGFB1 and HIF1A ($r = 0.3$), MMP9 and TGFB1 ($r = 0.38$), TNF and MMP9 ($r = 0.3$, Fig. 2I). Meanwhile, the seven DCRGs were used to constitute a protein-protein interaction network (Fig. 2J). As each DCRG was associated with inflammation, we defined the network as DCIN.

**Clinical Correlation Analysis**

To investigate the association between expression levels and the prognosis of the seven DCRGs, we performed an overall survival analysis of 33 cancers. Patients were divided into high-expressing and low-expressing groups according to the median gene expression levels. High and low expressions of the seven DCRGs were associated with overall survival in multiple cancers. **CTNNB1** was a protective factor for KIRC, READ, and UVM and a risk factor for ACC, HNSC, and LGG. **HIF1A** was a protective factor for SKCM and a risk factor for LIHC, MESO, and PCPG. **MMP9** was a protective factor for DLBC and a risk factor for ACC, BICA, KIRC, and LIHC. **PPARG** was a protective factor for BLCA, BRCA, KIRC, KEAD, and UVM and a risk factor for HINSC, LGG, LIHC, PAAD, and PCPG. **PTGS2** was a protective factor for COAD and PCPG and a risk factor for TGCT and UVM. **TGFB1** was a protective factor for SKCM, and a risk factor for LAML, LGG, MESO, and STAD. **TNF** was a protective factor for SARC and SRCM and a risk factor for THYM ($P < 0.05$, sFig2). Cox proportional hazard regression was used to detect the prognostic effects of the DCRGs in 33 TCGA tumors (Fig. 3). The expression level of the DCRGs varied in different tumors, and the effect on prognosis was inconsistent. Therefore, high expression of the same DCRG may be a protective factor in one type of cancer and a risk factor in another.

**GsVA Pathway Enrichment Analysis Of DCRGs Across Cancers**

In order to investigate the involvement of these seven DCRGs in cancer-related pathways and the pathways related to DM and carcinogenesis, we performed GSVA enrichment analysis on the DCRGs across cancers. We considered pathways with $|t| \text{ value of GSVA score}| > 2$ as significantly enriched pathways. GSVA enrichment analysis revealed several metabolically related pathways being enriched across cancers. Among them, the PPAR signaling pathway, MTOR signaling pathway, VEGF signaling pathway, arachidonic acid signaling pathway, which are related to metabolism, appeared multiple times in various cancers. Additionally, type I and type II diabetes mellitus also appeared in various cancer pathways (Supplementary Fig. 2), indicating that these seven DCRGs are closely related to diabetes and metabolism.

**Immune Cell Infiltration Analyses Of DCRGs Across Cancers**

We next examined the relationship between the seven DCRGs and the infiltration levels of 26 immune-related cells. Our data showed that in most cancers, the level of immune cell infiltration was significantly correlated with the expression of the DCRGs ($|R| \geq 0.5$, $P < 0.01$). Among them, **MMP9** was
associated with immune cells in 12 tumors, \textit{PTGS2} with immune cells in seven tumors, and \textit{TNF} with immune cells in five tumors. (sFig 3).

**Immune Subtype Analysis**

The mRNA expression of the seven DCRGs in six immune subtypes of 33 TCGA tumor types was analyzed by the Kruskal test. The seven DCRGs were differentially expressed in the C1-C6 subtypes across cancers ($P < 0.001$). The expression of \textit{CTNNB1} was the highest in the C1-C6 immune subtypes (Fig. 4A). Similarly, the seven DCRGs were differentially expressed in the C1-C6 immune subtypes of COAD ($P < 0.05$). Moreover, the expression of \textit{CTNNB1} in the C1-C6 immune subtype was highest across cancers compared with the other 6 DCRGs. (Fig. 4B).

**Stemness Indices And Microtumor Environment**

The matrix fraction of TCGA cancer samples was calculated by applying the ESTIMATE algorithm. Spearman correlation analysis was used to describe the correlation between the expression levels of the DCRGs and the pan-cancer matrix score. We found that the expression of \textit{MMP9}, \textit{TGF\textbeta}1, and \textit{TNF} were positively correlated with immune and stromal scores ($P < 0.05$, $R > 0.5$, sFig4A, B). In order to analyze the correlation between the DCRGs and pan-cancer stemness characteristics, we calculated the stemness index of the TCGA tumor samples using a Class I logistic regression (OCLR) algorithm. Subsequently, spearman correlation analysis was performed based on gene expression and stemness score. Two types of dryness indices were calculated: DNA methylation-based dryness index (DNAss) and mRNA expression-based dryness index (RNAss). The correlation between the two stemness indexes and the expression levels of the seven DCRGs varied in the TCGA tumors. DNAss exhibited a strong correlation between OV and TGFB1, THYM and PTGS2, and TCGT and CTNNB1. RNAss demonstrated that \textit{TGF\textbeta}1 was negatively correlated with many cancers, and \textit{CTNNB1} was negatively correlated with THYM (sFig4C, D).

**Mutations In The Seven DCRGs Across Cancers**

We analyzed mutations in the seven DCRGs in 2,922 samples of 38 cancers from the Pan-Cancer Analysis of whole Genomes (ICGC/TCGA, Nature 2020) using the cBioPortal platform. Among the seven DCRGs, \textit{PTGS2} had the highest mutation frequency across cancers (11%), followed by \textit{MMP9} (10%, Fig. 5A). Of the 38 types of cancer, the mutation frequency of colorectal cancer was the highest (65%), followed by lung cancer (52%, Fig. 5B). Overall survival analysis showed that the mutated group had a worse prognosis than the non-mutated group (Fig. 5C).

**Correlations Of The DCRGs With The Immunomodulators, TMB And MSI**
We then investigated whether the expression levels of the seven DCRGs correlated with TMB and MSI, both of which are intrinsically related to immune checkpoint inhibitor sensitivity. The results showed that, except for KICH, READ, CHOL, TGCT, and LAML, the TMB of the remaining 28 tumors was correlated with the expression of at least one of the seven DCRGs (sFig5A). Meanwhile, except for CHOL, OV, UCS, and LAML, the MSI of the other 29 cancers was correlated with the expression of at least one of the seven DCRGs (sFig5B). Interestingly, all seven DCRGs were positively correlated with TMB and MSI of TNYM. In addition, TNF was negatively correlated with the target tumor in all TMB and MSI except THYM.

**Seven Dcrgs Predict The Response To Cancer Immunotherapy**

Due to the high frequency of mutations of the seven DCRGs in COAD, it was selected for further analysis. We analyzed the association of the seven DCRGs with multiple cancers and 46 immune checkpoint-related genes in COAD. All 46 immune checkpoint-related genes were associated with at least one of the seven DCRGs. Among them, $CD274$ (PD-L1) was correlated with 6 genes except for $MMP9$ (sFig5C).

**Drug Sensitivity Analysis Across Cancers**

To analyze the potential effect of the DCRGs on drug response, we performed a Pearson correlation analysis between the transcriptional expression of the seven DCRGs in the NCI-60 cell line and the drug activity of 263 anti-tumor drugs retrieved from the CellMiner database. Scatterplots (sFig6) sorted by value showed significant correlations between drug sensitivity and gene expression. Notably, Nelarabine and $TNF$ showed a high correlation ($|R|=0.747$). Together, our results showed that multiple drugs were sensitive to the seven DCRGs ($P<0.001$, $|R|>0.4$).

**Single Cells Analysis Of The Seven Dcrgs In Coad**

Next, we analyzed the activity of the DCRGs in COAD single cells based on scRNA-seq, and the marker genes identified eight cell types. Then, we used the AddModuleScore function to measure the activity of seven DRGs in each cell cluster. Among them, the activity of the DCRGs was highest in M1 macrophage and the lowest in Plasma B(Figure 6A-C). The activity of the DCRGs was highest in carcinoma and lowest in the blood of colon cancer patients, which was lower than that of normal colorectal tissue Fig. 6D,E.

**Verification Of The Seven Dcrgs And Arachidonic Acid Pathway In Coad**

In order to verify the expression of the DCRGs and arachidonic acid metabolic pathways in tumors, we stimulated colon cancer epithelial cells and normal intestinal epithelial cells with high glucose and physiological levels of glucose, respectively. All seven DCRGs were significantly elevated in the colon cancer epithelial cells under a high glucose environment (Fig. 7A-H). However, in normal intestinal
epithelial cells, there was no significant difference in the expression of the DCRGs under a normal or high glucose environment (Fig. 7I-P). These results suggest that high glucose stimulation can lead to high expression of the seven DCRGs in colon cancer epithelium, while it does not affect their expression in normal colon epithelial cells.

For the arachidonic acid pathway, the expression of \textit{PEG2} in colon cancer epithelial cells was significantly increased under high glucose, while the difference in \textit{LTC4} was not statistically significant compared to the controls. \textit{PEG2} was not significantly different in normal intestinal epithelial cells, while \textit{LTC4} expression was significantly higher under high glucose stimulation. Moreover, \textit{LTB4} did not exhibit significant differences in either colon cancer cells or normal intestinal epithelial cells (Fig. 7Q-S).

**Discussion**

Using bioinformatics, we screened seven DCRGs, including \textit{PPARG}, \textit{MMP9}, \textit{CTNNB1}, \textit{TNF}, \textit{TGFB1}, \textit{PTGS2}, and \textit{HIF1A}, and established a DCIN to comprehensively understand the diabetes-inflammation-cancer interaction (Fig. 8). In addition, we found that the level of immune infiltration, immune subtypes, tumor microtumor environment, mutation, the correlation with TMB and MSI, and the sensitivity of immune checkpoint-related drugs were strongly different in this DCIN. Furthermore, the expression of DCRGs and arachidonic acid metabolism pathway was verified \textit{in vitro}. Together, we defined a DCIN and illustrated the potential mechanism through which diabetes may influence cancer.

Among DCIN, Tumor necrosis factor (TNF) is a pro-inflammatory cytokine produced and secreted by cytotoxic lymphocytes upon tumor target recognition. MMP9 is a member of matrix metalloproteinases (MMPs), which have been widely studied in various cancers [27, 28]. Both the tumor cells and surrounding stromal cells can synthesize MMPs [29]. MMP9 is one of the key proteolytic enzymes in the breakdown and reconstruction of the extracellular matrix in colorectal cancer (CRC) invasion and metastasis [30]. TNF-α upregulates MMP9 expression via c-Src, MAPKs, and NF-κB pathways [31]. It has been shown that MMP9 regulates the vascular endothelial growth factor (VEGF) signaling axis by cleaving membrane-bound VEGF, leading to increased bioavailability of its receptor VEGFR2 [32]. Thus, MMP9 may contribute to local angiogenesis by increasing levels of VEGF in tumors [33]. Our results showed that TNF was highly expressed in almost all tumors, confirming its carcinogenicity. MMP9 was also highly expressed in all tumors except THYM, suggesting that MMP9 inhibitors may be used in the treatment of various tumors. β-Catenin (\textit{CTNNB1}) is a multitasking protein involved in transcription and cell adhesion [34]. \textit{CTNNB1} affects tumors in two ways. On the one hand, \textit{CTNNB1} plays a vital role in the development and progression of tumors through Survivin, which inhibits apoptosis, promotes cell cycle progression, and enhances angiogenesis [35]. On the other hand, \textit{CTNNB1} knockdown inhibits the Wnt/β-catenin signaling pathway and downregulates the expression of downstream genes, including axin 2, lymphoid enhancer binding factor 1 (\textit{LEF1}), and cyclin D1. Therefore, low expression of \textit{CTNNB1} can inhibit tumor proliferation [36]. In our study, \textit{CTNNB1} was highly expressed in all the tumors except CESC, OV, UCEC, and UCS, which also confirmed its carcinogenicity.
HIF-1α is part of the Hypoxia-inducible factor-1 (HIF-1) family [37], which is a key factor regulating cell adaptation to hypoxia [38]. HIF1 can induce the expression of several pro-angiogenic factors, including vascular endothelial growth factor (VEGF) and VEGF receptors FLT-1 and FLK-1. Among all of these pro-angiogenic factors, VEGF-A, a potent endothelial mitogen, is a notable protein since it is highly expressed in many human tumors [39, 40]. However, during the later stages of tumor development, TGFβ1 functions as a tumor promoter by inducing the epithelial-mesenchymal transition (EMT) in cancer cells, resulting in increased invasion and metastasis [41]. Meanwhile, PTGS2, also known as COX-2, contributes to angiogenesis [42]. COX-2 overexpression in colon cancer cells leads to the production of prostaglandins and the induction of pro-angiogenic factors such as vascular endothelial growth factor (VEGF) and basic fibroblast growth factor (bFGF), stimulating endothelial cell migration and tube formation [43, 44]. In our study, the high and low expression levels of TGFβ1 in different tumors showed great differences, which also reflected the dual role of TGFβ1. Peroxisome proliferator-activated receptors (PPARs) include three isotypes: PPARα, PPARγ, and PPARδ. The PPARγ gene encodes PPARγ [45]. The anti-tumor functions of PPARα and PPARγ are currently inconclusive and controversial [46]. Downregulation of PPARγ could inhibit the proliferation of T24 cells, which might be caused by cell cycle arrest in the G0/G1 phase [47]. PPARγ has been demonstrated to have anti-neoplastic effects by arresting the cell cycle, causing terminal differentiation, and inhibiting inflammation [48]. Here, over-expression of PPARγ was a protective factor for five tumors and a risk factor for another five. It also confirmed that the role of PPARγ varies in the different cancer types. Together, we found that high glucose stimulation can lead to high expression of the seven DCRGs in colon cancer epithelium, while it does not affect their expression in normal colon epithelial cells. These results suggest that diabetes may promote the occurrence and development of tumors through DCIN.

The GSVA analysis showed that DCRGs were closely related to the arachidonic acid metabolism pathway. Arachidonic acid, specifically its metabolites, has attracted a great deal of attention in cancer biology, especially in the context of inflammation [49]. The importance of arachidonic acid in biology lies in the fact that it can be metabolized by three distinct enzyme systems, cyclooxygenases (COXs, also referred to as PGG/H synthases), lipoxygenases (LOXs), and cytochrome P450 (CYP) enzymes (ω-hydroxylases and epoxygenases) to generate an impressive spectrum of biologically active fatty acid mediators [50]. The signaling of cyclooxygenase 2-prostaglandin E2-prostaglandin E2 receptors (COX-2-PGE2-EPs) is the central inflammatory pathway involved in carcinogenesis [51]. COX is the primary enzyme in the synthesis of eicosanoids and exists in two isoforms: COX-1, which is ubiquitously expressed [52], and COX-2, which is expressed predominantly in inflammatory cells and upregulated in chronic and acute inflammations [53]. Prostaglandins derived from COX-2 contribute to cancer progression and metastasis [54]. The COX-2 expression is stimulated by different growth factors, cytokines, and prostaglandins, which are associated with inflammatory responses and have been shown as prognostic factors for malignancy [55]. Furthermore, upregulation of COX-2 and PGE2 has been identified in many human cancers and precancerous lesions, and COX-inhibitory drugs have shown protective effects in colorectal cancer [56, 57]. In addition to colorectal cancer, nonsteroidal anti-inflammatory drug (NSAIDs) have also been associated with a reduced risk of breast, esophageal, stomach, bladder, ovary, and lung cancers [58,
We showed that the expression of PEG2 in colon cancer epithelial cells was significantly increased under high glucose compared to the control. In contrast, PEG2 had no significant difference in ordinary intestinal epithelial cells. These results suggest that diabetes may promote the occurrence and development of cancers through the pathway of arachidonic acid metabolism.

In the immune infiltration analysis of the seven DCRGs and 26 immune cells, we found that MMP9 was associated with 12 tumors, PTGS2 was associated with 7 tumors, and TNF was associated with 5 tumors. There was a positive correlation between MMP9 and all the immune cells. This suggests that DCRGs may influence tumor development and prognosis through the immune cells. Our genetic analysis showed a high frequency of copy number variations of DCRGs, and the prognosis of the mutational group was lower than the non-alterative group, indicating that mutations could affect the occurrence and development of tumors. According to research, breast cancer has been linked to mutations in MMP9[60]. Additionally, it may promote the invasion and metastasis of bladder cancer[61]. PPARG mutations are closely associated with cancers of the digestive tract (colon, stomach, oesophagus, and pancreas), melanoma, breast cancer, prostate cancer, and bladder cancer[62]. The role of these seven DCRGs mutations in different tumors needs further investigation. Moreover, TMB and MSI analysis of the DCRGs suggested that TMB and MSI are correlated with the expression of the seven DCRGs in most tumors. Interestingly, both TMB and MSI were positively correlated with the expression of all DCRG in TNYM. Moreover, except for THYM, TMB and MSI were negatively correlated with TNF expression in the rest of all the DCRGs. This suggests that the treatment of TNYM at the level of immune checkpoints may differ from others. In addition, drug sensitivity analysis suggested that the seven DCRGs potentially affected tumor drug response. Among them, Nelarabine was highly correlated with TNF, which has guiding clinical significance for the selection of anti-tumor therapy.

To further investigate the role of the seven DCRGs in colon cancer cells, single-cell sequencing data of colon cancer were analyzed. As a result, the activity of the seven DCRGs was found to be the highest in M1 macrophage and the lowest in plasma B. Previous study has reported that macrophages can promote pro-tumor inflammation by secreting pro-inflammatory cytokines. On the one hand, it can induce immune responses but also support tumor growth and survival of malignant cells [63]. Macrophages also play diverse roles in cancer development, ranging from anti-tumor activity at early stages of progression to tumor promotion in the most commonly established cancers [64]. For diabetes, it has previously been suggested that hyperglycemia enhances cancer immune evasion by increasing O-GlcNAcylation to induce alternative macrophage polarization [65]. Thus, macrophages play a crucial role in the development of diabetes and cancer. MMP9 is secreted by macrophages and acts on PAR1 of PDAC cells to induce epithelial-to-mesenchymal transition. This macrophage-induced mesenchymal transition supports the tumor-promoting effect of macrophage influx, explaining the dual contribution of these immune cells to tumor growth [66]. Notably, studies have shown that TNF-α regulates diabetic macrophage function through the histone acetyltransferase MOF [67]. Therefore, based on the single-cell and our experimental validation results, we concluded that the seven DCRGs were associated with immune inflammation, especially macrophages.
This study is the first to propose the concept of DCIN, but it still has some limitations. Firstly, all the samples involved in this study were from China, so we are not quite sure about the applicability of the prediction model in other parts of the world. Secondly, the results of this study have not been verified by other independent databases. However, we have followed up the results with our molecular biology experiments to make the results more convincing. Finally, the analysis in this study focused on the correlation between the DCRGs; however, biostatistical correlations cannot elucidate direct interactions and direct regulatory mechanisms. Therefore, our future experiments aim to verify the interaction of various molecules in the DCIN and the potential mechanism of molecular action on tumors through more molecular experiments.

**Conclusion**

A DCIN was constructed for the first time in this study (Fig. 8), allowing novel insights into the mechanism by which diabetes might increase the occurrence and development of tumors. Although further research is necessary, the results of our study also provide clues for new potential therapeutic strategies aimed at preventing and treating malignant tumors.

**Declarations**

**Data Availability:**

The datasets generated and/or analyzed during the current study are available in Supplementary Materials and TCGA program (https://portal.gdc.cancer.gov).

**Conflicts of Interest**

The authors declare that there is no conflict of interests.

**Authors’ Contributions**

WP designed this study and revised the manuscript; YT, and JK conducted the data collection, bioinformatic and statistical analysis, figures visualization and manuscript writing. HL conducted WB and ELISA experiments. YL, AZ, RH, ZZ and XC revised the manuscript. All of coauthors have approved the final version of manuscript.

**Acknowledgments**

This work was financially supported by the National Natural Science Foundation of China (No. 81873169), Hunan Youth Science and Technology Innovation Talent Project (No.2022RC1220), Hunan Provincial Natural Science Foundation of China (No.2020JJ4803), and Hunan Flagship Department of Integrated Traditional Chinese and Western Medicine. We also thank Bullet Edits Limited for the linguistic editing and proofreading of the manuscript.
References


44. S. Vimalraj, A concise review of VEGF, PDGF, FGF, Notch, angiopoietin, and HGF signalling in tumor angiogenesis with a focus on alternative approaches and future directions. Int. J. Biol. Macromol. 221, 1428-1438 (2022)


**Abbreviations**

adrenocortical carcinoma (ACC), Bladder Urothelial Carcinoma (BLCA), breast invasive carcinoma (BRCA), cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), Cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC), esophageal carcinoma (ESCA), glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSC), Kidney Chromophobe (KICH), kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), Acute Myeloid Leukemia (LAML), Brain Lower Grade Glioma (LGG), liver hepatocellular carcinoma (LIHC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), Mesothelioma (MESO), ovarian serous cystadenocarcinoma (OV), pancreatic adenocarcinoma (PAAD), Pheochromocytoma and Paraganglioma (PCPG), prostate adenocarcinoma (PRAD), rectum adenocarcinoma (READ), Sarcoma (SARC), Skin Cutaneous Melanoma (SKCM), stomach
adenocarcinoma (STAD), Testicular Germ Cell Tumors (TGCT), thyroid carcinoma (THCA), Thymoma (THYM), Uterine Corpus Endometrial Carcinoma (UCEC), Uterine Carcinosarcoma (UCS), and Uveal Melanoma (UVM)

Figures

Figure 1

**Screening for differential DCRGs**  (A) Venn diagram showing nine intersection genes of five diseases: T2DM, pancreatic cancer, liver cancer, and colon cancer based on DisGeNET (r > 0.1).(B) The 59 intersection genes of five diseases: T2DM, pancreatic cancer, liver cancer, and colon cancer, based on GeneCards (r > 0.1). (C) The seven DCRGs obtained by the intersection of DisGeNET and GeneCards.
**Figure 2**

**Differential expression analysis** (A-G) The box plots showing differential expression of the seven DCRGs in normal and tumor tissues (*** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$). (H) The heatmap showing the transcriptional level of the seven DCRGs between normal and tumor tissues in various cancers. The gradient colors represent the log Fold Change (logFC) value. (I) The heatmap showing the pairwise correlation of the seven DCRGs. (J) The protein-protein interaction network of the seven DCRGs.
Figure 3

Survival analysis of the seven DCRGs across cancers Cox proportional hazard analyses illustrating the hazard ratios (HRs) of the seven DCRGs in 33 TCGA tumors. DCRGs with HR > 1 were regarded as risk factors, while DCRGs with HR < 1 were regarded as protective factors.
Figure 4

Expression pattern of the seven DCRGs in six different immune subtypes (A) Transcriptional expression of the seven DCRGs in C1-C6 immune subtypes across all TCGA cancers. (B) Box plots showing the expression level of the DCRGs in the immune subtypes in COAD (*** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$).
Figure 5

Alterations in the DCRGs in cancer based on the cBioPortal platform (A) Mutation types and frequencies of the seven DCRGs in all samples. (B) Mutation types and frequencies of the seven DCRGs in 27 cancer types. (C) Kaplan-Meier plots of the seven DCRGs across cancers showing the differential survival outcomes of the altered and unaltered group ($P<0.05$).
Figure 6

Analysis of single-cell RNA sequencing of COAD (A) UMAP plot of the cell clusters of GSE161277. (B) Top 10 marker genes among the eight cell clusters. (C) The expression of the DCRGs among the eight cell clusters. (D) The expression of the seven DCRGs among tissues of adenoma, blood, carcinoma, normal, and paracancer. (E) UMAP plot of the seven DCRGs expression patterns among the eight cell clusters.
Figure 7

Verification of the seven DCRGs and arachidonic acid pathway in COAD  (A-H) Western blots of the seven DCRGs in SW480 treated with NG or HG. (I-P) Western blots of the seven DCRGs in NCM460 treated with NG or HG. (Q-S) ELISA of \( \text{PG}E_2 \), \( \text{LTC}4 \), and \( \text{LTB}4 \) in SW480 and NCM460 treated with NG or HG. * \( P < 0.05 \) compared to controls. HG: high glucose, NG: normal glucose.
Figure 8

The mechanistic diagram of diabetes-based cancer-associated inflammation network (DCIN) The legend appears in the upper right corner of the image. Cell membrane is represented by the double-framed line in the upper part of the picture, and nucleus is represented by the double-framed line in the middle circle. The orange gene represents the DCRGS that was screened out in this study, while the gray gene represents an important biological factor.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementFig1.tif
- SupplementFig2.tif
- SupplementFig3.tif
- SupplementFig4.tif
- SupplementFig5.tif
• SupplementFig6.tif