

Identification of diabetes related phenotype and diagnostic biomarkers in coronary artery disease

Gulandanmu Aihemaiti^{1,*}, Abulikemu Abula^{2,*}, Junyi Luo¹, Qian Zhao¹, Ning Song¹, Fen Liu^{1,3}, Niyaziaili Adili¹, Yining Yang^{3,4} and Xiaomei Li^{1,3}

- ¹ Department of Cardiology, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, China
- ² Emergency Trauma Center, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, China
- ³ Xinjiang Key Laboratory of Cardiovascular Disease Research, Clinical Medical Research Institute of First Affiliated Hospital of Xinjiang Medical University, Urumqi, Xinjiang, China
- ⁴ People's Hospital of Xinjiang Uyghur Autonomous Region, Urumqi, Xinjiang, China
- These authors contributed equally to this work.

ABSTRACT

Background. Diabetes mellitus (DM) and coronary artery disease (CAD) are closely interrelated clinical conditions. However, the combination analysis based on DM related CAD diagnostic model remains a gap. The primary objective of this study was to identify diagnostic models and diagnostic markers for CAD based on the association of diabetic phenotypes and attempt to explore them further in a mouse model.

Methods. We used data integration as well as multiple datasets for both coronary artery disease and diabetes to exclude bias as well as to improve reliability. We employed the least absolute shrinkage and selection operator (LASSO) regression algorithms to construct the CAD diagnostic model. Furthermore, we established mouse CAD model (low-density lipoprotein receptor deficient mice with high fat diet) to explore the crosstalk between the screened biomarkers and severe CAD progress.

Results. The intersecting genes from differential analysis and weighted correlation network analysis (WGCNA) results yielded 32 diabetes-related biomarkers. We then identified two diabetes-related phenotypes through the consensus clustering in CAD patients. Microenvironmental analysis revealed that phenotype 1 exhibited higher expression of most cytokines, inflammatory factors, interleukins, and related receptors. Immune cell composition in phenotype 1 showed increased infiltration compared to phenotype 2. The LASSO regression identified 16 diabetes-related genes and we further constructed a diagnostic model based on these genes, which the area under the curve (AUC) reached 0.8. Additionally, single cell immune analysis exhibited the location of these genes. KCNQ1, ATP6V1B1, MTDH, and ITPK1 were predominantly located in macrophages, indicating their potential in regulating macrophage during myocardial injury. Furthermore, We elucidated that KCNQ1 and ITPK1 exhibited high expression level in mouse CAD model in tissue level. exhibited similar expression trends with macrophage biomarkers (CD31 and CD68). The result of qPCR also indicated the elevated level of KCNQ1 and ITPK1, which exhibited crosstalk with CD31 and CD68 in mouse CAD model.

Conclusion. This study delves into the microenvironmental characteristics of diabetes-related phenotypes in CAD, constructing an optimal diagnostic model and validated

Submitted 11 November 2024 Accepted 1 September 2025 Published 14 October 2025

Corresponding authors Yining Yang, yangyn5126@163.com Xiaomei Li, lixm505@163.com

Academic editor Fanglin Guan

Additional Information and Declarations can be found on page 18

DOI 10.7717/peerj.20117

© Copyright 2025 Aihemaiti et al.

Distributed under Creative Commons CC-BY-NC 4.0

OPEN ACCESS

the significance of diagnostic markers in mouse CAD model, which may offer insights that could be beneficial for clinical management in the near future.

Subjects Bioinformatics, Biotechnology, Cardiology, Diabetes and Endocrinology **Keywords** Diabetes mellitus, Coronary artery disease, Mouse CAD model, Biomarkers

INTRODUCTION

Coronary artery disease (CAD) is predominantly characterized by atherosclerosis, a complex disorder characterized by damage to the artery intima (Ben Braiek et al., 2021). It manifests in forms such as stable angina, unstable angina, myocardial infarction (MI), or sudden cardiac death (Iida, Harada & Takebayashi, 2019). Key risk factors like hypertension, hyperlipemia, diabetes mellitus (DM), and smoking are identified as significant contributors to the disease at the population level (Malakar et al., 2019). In 2017, CAD was accountable for roughly 8.93 million deaths, making up about 16.0% of global death, which is expected to remain the leading cause of death worldwide until 2030 (Xu et al., 2023). Coronary angiography is acknowledged as the definitive method for diagnosing the presence and extent of CAD (Li & Kronenberg, 2021). Despite considerable progress in medical treatments, CAD remains a critical public health issue globally (Tseng et al., 2020). Additionally, individuals diagnosed with coronary atherosclerotic heart disease (CHD) are at an increased risk of subsequent cardiovascular events, including MI, stroke, and chronic heart failure (CHF) (Xia et al., 2021), which may reduce the possibility of long-term survival of patients. Hence, Identification of effective diagnostic markers for CAD patients remains an urgent problem.

DM is identified by a variety of heterogeneous disorders resulting from either insulin deficiency or insulin resistance, leading to high blood glucose levels (Bao, 2022). Both type 1 and type 2 diabetes cause organ-specific tissue damage due to multiple metabolic factors (Fan et al., 2022). Interestingly, DM is also a significant risk factor for CHD. Chronic hyperglycemia in DM (Eid et al., 2019) not only harms various organs and tissues but also leads to microvascular complications such as retinopathy (Pignalosa et al., 2021), nephropathy, and neuropathy, as well as macrovascular complications like cardiovascular disorders (Darenskaya, Kolesnikova & Kolesnikov, 2021). The mechanisms behind hyperglycemia-induced vascular damage are complex and not fully understood. However, it is thought that increased intracellular glucose levels enhance the production of reactive oxygen species, subsequently disrupting vital downstream pathways (Cole & Florez, 2020; Vergès, 2020). Additionally, the hyperglycemic state resulting from diabetes may activate protein kinase C signaling and the late glycation pathway, leading to vascular inflammation as well as CAD (Henning, 2018). Enhancement of the hexosamine biosynthetic pathway and activation of the protein O-GlcNAcylation have also been shown to play an important role in diabetes-associated vasculopathy (Wang et al., 2023). Macrophages may be the crosstalk between DM and CAD. Peng & Gao (2022) found that NLPR3⁺ macrophages can exacerbate in vivo inflammation in diabetic patients.

Macrophages also play an important role in cardiac remodeling in patients with CAD (*Yap et al.*, 2023). Interestingly, plaque in coronary arteries is usually larger in diabetic patients than in nondiabetic patients, mainly due to inflammation caused by macrophages and T lymphocytes (*Yahagi et al.*, 2017). These fundamental studies involved cellular and clinical models reveals the deep mechanisms of the DM and CAD. However, it lacked the comprehensiveness of integrated large-sample data and combinational analyses to some extent. Ideal clinical markers may be identified form the perspective of DM related CAD analysis.

In our study, we first explored diagnostic markers as well as diagnostic models for CAD patients from crosstalk with diabetes. Based on multiple fused datasets, we try to validate the identified diagnostic markers in a mouse model to further explore their diagnostic value in patients with CAD.

MATERIALS AND METHODS

Data sources and preparation

The datasets concerning diabetes (2025) and CAD (Chen et al., 2023) predominantly derive from the GEO database. The diabetes datasets include GSE20966 (20 samples with diabetes), GSE25724 (six of 13 total samples have diabetes), and GSE38642 (nine of 63 total samples have diabetes), and the CAD datasets include GSE12288 (222 samples in total, with 110 samples in CAD), GSE20680 (195 samples with CAD), GSE20681 (198 samples in CAD), and GSE42148 (24 samples, with 13 samples in CAD). The single cell dataset was GSE32678. Additionally, to increase the sample size and confidence level, various datasets of patients with diabetes mellitus and coronary heart disease were merged, with batch effects excluded by "combat" packages for optimal analysis. For missing values, we fill them with the mean value. For data cleansing, we correct formatting errors, remove duplicate values, and remove unneeded data to ensure data cleanliness. For data with large variance, we usually use the "normalize" function to normalize the data, making the data comparable. In addition, for outliers, we usually remove them or fill them with mean values. Specific dataset inclusions and workflow are detailed in the supplementary material.

WGCNA analysis and differential analysis to identify diabetes related markers

Weighted correlation network analysis (WGCNA) is a systems biology approach for characterizing gene association patterns between samples (*Langfelder & Horvath*, 2008). This analysis focuses on identifying co-expressed genes based on gene-set introgression and gene-set-phenotype associations. We utilized WGCNA to pinpoint modules co-expressed with diabetes and the core genes in the network.

Identification of diabetes phenotypes in CAD patients

The intersecting genes from WGCNA and differential expression analysis were employed to identify immune-related diabetes phenotypes. We identified these phenotypes using the unsupervised clustering algorithm (consensus clustering). The main basis for the classification is the CDF value and delta region in the consensus clustering algorithm, when

the CDF curve is the most downward and the delta region has the largest change, it means that this algorithm has the best classification effect.

Identification diabetes related coronary artery disease diagnostic system

For more efficient results, we applied the LASSO regression algorithms to reduce dimensionality and identified 16 diabetes-related genes. These genes were used to construct the diagnostic monitoring model for CAD and to measure the area under the curve (AUC) to evaluate the diagnostic capacity. The model was constructed by "glmnet" package (seed(123), alpha = 1, nfold=100).

Single-cell landscape analysis

The "Seurat" package was applied for single-cell sequencing analysis. The criteria for cell screening were defined as nCount_RNA >100, nFeature_RNA <5,000, percent.mt <25, and nFeature_RNA >100. Only the top 3,000 variable genes were then chosen for additional examination. Using principal component analysis (PCA), significant dimensions with p < 0.05 were found. Dimensionality reduction was made possible via the t-distributed stochastic neighbor embedding (tSNE) approach, which was combined with an 11-cluster classification analysis conducted on all cells. The "singleR" package was used to annotate cell clusters based on trends in the composition of marker genes.

Vertebrate animal study methods

This study was approved by Animal Experimental Medicine Ethics Committee of the First Affiliated Hospital of Xinjiang Medical University, Approval number: 20210301-215. Twenty-eight-week-old male C57/B6J wild-type and LDLR—/— mice were obtained from the Xinjiang Medical University experimental animal center's specialized pathogen-free (SPF) barrier system. A high-fat diet consisting of 78.85% chow diet, 21% lard, and 0.15% cholesterol was provided to the model group whereas the control group was fed chow diet. An injection of pentobarbital was used to induce anesthesia. We use carbon dioxide for the euthanasia of mice.

Mouse euthanasia operation procedures

We placed the mice to be executed in empty boxes and used carbon dioxide asphyxiation to execute them. Criteria established for euthanizing animals prior to the planned end of the experiment was the State Standard of the People's Republic of China, GB/T 39760-2021.

Gene silence of KCNQ1 and ITPK1

We knocked down of KCNQ1 and ITPK1 through siRNA. The sequence was as follow: KCNQ1(SiRNA): CCGCCUGAACCGAGUAGAATT; ITPK1(SiRNA): AUCUUGUCCAGCUCCGUCATT.

CAD mouse model construction

Twenty-eight-week-old male C57/B6J wild-type and LDLR—/— mice were grown in the Xinjiang Medical University experimental animal center's specialized pathogen-free (SPF) barrier system. Following randomization, a high-fat diet consisting of 78.85% chow diet,

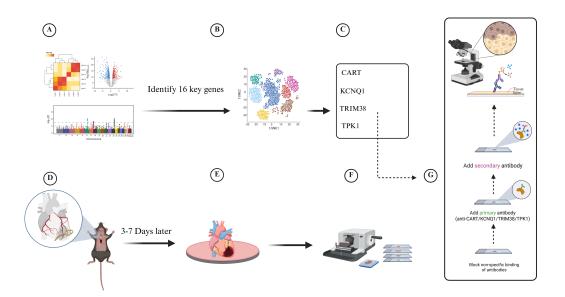


Figure 1 Experimental flowchart for transcriptomic and single-cell data to develop a diabetes-related phenotype and its microenvironment influence in coronary artery disease. (A) Analyze the gene information of patients in GEO database, (B) Analyze single-cell data from patients with coronary myocardial injury, (C) Identify four key genes in diabetes-related CAD, (D) Establish myocardial infarction model, (E) Euthanasia of mice and removal of heart tissue, (F) Collect and section heart tissue, (G) Immunohistochemistry assay.

21% lard, and 0. 15% cholesterol was provided to the CAD model group (LDLR—/— mice) continuously for 12 weeks, whereas the control group was fed chow diet. To create a model of a myocardial infarction, the anterior descending branch of the mice was clamped for seven days at the conclusion of the twelfth week. Changes in electrocardiogram before and after modeling were observed, and blood lipids and cardiac enzymes were measured 2 weeks after surgery to determine whether coronary artery disease occurred in mice. Sections of mouse heart tissue were obtained for immunohistochemistry (IHC) staining. The IHC staining of tissue slices from the cardiac tissues of the two groups (five mice in each group) was then examined (Fig. 1).

Flow cytometry analysis

Apoptosis rates were determined using the Annexin V/PE Apoptosis Detection Kit (BD Biosciences, Franklin Lakes, NJ, USA). Cells were collected with $1\times$ binding buffer. Subsequently, five μL of Annexin V-PE and five μL of 7-AAD were added, and the incubation was carried out for 5 min in the absence of light. The analysis of apoptosis was finally conducted using a FACSCalibur ffow cytometer.

Immunohistochemistry staining

Detailed procedures for immunohistochemistry can be found in the classic literature on *Tavitian et al.* (2025). Dimethylbenzene was used to dewax pathological sections, and gradient alcohols were used to hydrate them. Sections were then subjected to boiling sodium citrate (pH 6.0) for antigen retrieval and 3% H2O2 to deactivate endogenous peroxidase.

Sections were then treated in turn with primary antibody, HRP-linked secondary antibody, and 5% goat serum. A DAB substrate was used to further visualize the indicated antibody staining. Sections were then examined under a bright-field microscope after being restained with hematoxylin and sealed with neutral glue. The staining intensity (negative, 0; weak, 1; moderate, 2; strong, 3) and the percentage of positive signal (negative, 0; 1–30%, 1; 31–60%, 2; above 60%, 3) were used to assess the protein expression. The IHC score was obtained by multiplying the two scores mentioned above. The main antibodies we used were KCNQ1 (A2174, 1:400; ABclonal, Waltham, MA, USA), ITPK1 (10568-AP-1, 1:500: Proteintech, Rosemont, IL, USA), CD68 (28058-1-AP, 1:300: Proteintech, Rosemont, IL, USA).

Statistical analysis

R version 4.1.0 was the primary tool for all bioinformatics analyses and related statistical calculations. For biological experiment statistical analysis, GraphPad Prism version 8.1 was used. Regarding the multiple testing methods for the statistical analysis, we have used FDR control methods (the BH method). Regarding cross-validation of diagnostic models. We divided the fused dataset into training and testing datasets according to 1:1, and performed cross-validation and correction in the two datasets,

RESULTS

WGCNA and differential analysis to identify key diabetes-related gene markers

To reduce data analysis inaccuracies due to small sample sizes, data from three GEO platform datasets were merged, carefully excluding batch effects to create a combined dataset of diabetic patients and control samples. Initial analysis showed significant data dimension disparities between datasets (Fig. 2A). However, post-fusion, the data exhibited nearly uniform distribution patterns (Figs. 2B, 2C). Subsequent analysis of variance identified differential genes between diabetic patients and healthy controls (p < 0.05and abs (log (FC)) >1), revealing 2,250 differential genes (Fig. 2D). Further, WGCNA analysis was conducted to isolate modular genes strongly linked to diabetes (Fig. 2E). In the WGCNA procedure, the soft threshold β was set at 9. For network visualization, 400 of these genes were randomly selected (Fig. S1). The WGCNA network was constructed using the partitioning method, merging and analyzing modules with high similarity, ultimately identifying 22 different gene modules (Fig. 2E). Correlation analysis between gene modules and clinical data of diabetic patients indicated that the yellow module had the strongest correlation with diabetes (r = 0.53, p = 5e - 6) (Fig. 2F). To pinpoint hub genes within the yellow module, 368 genes with GS >0.2 and MM >0.6 were classified as central diabetes-related genes. The intersecting genes from differential analysis and WGCNA results yielded 32 diabetes-related biomarkers (Fig. 2G).

Diabetes related phenotype and characteristics in coronary artery disease

We employed the same methodology from the analysis of diabetes to boost sample size and lower error in the bioinformatics analysis of CAD patients. We reduced the batch

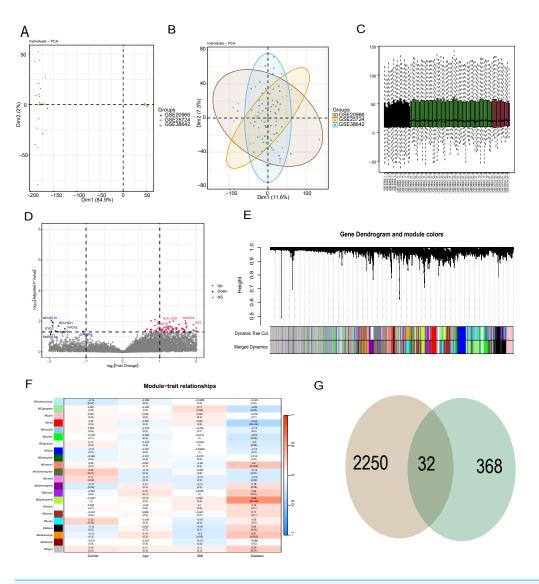


Figure 2 WGCNA and differential analysis to identify key diabetes-related gene markers. (A) The data difference among datasets. (B, C) The data distribution after fusion. (D) The volcano plot between diabetic patients and healthy controls. (E) WGCNA analysis to identify different gene modules. (F) Correlation analysis between gene modules and clinical data of diabetic patients. (G) The intersecting genes from differential analysis and WGCNA results.

effect by combining four datasets from the GEO platform. When the batch effect was excluded, the initial examination of the four datasets showed significant differences (Fig. 3A) that eventually became negligible (Fig. 3B). We also looked into the capacity of diabetes-related markers to classify patients with CAD, given the possible influence of diabetes on the development of CAD. We classified the gene expressions of 32 diabetes biomarkers in CAD patients using the consensus clustering method, which allowed us to identify diabetes-related phenotypes in CAD patients. Binary typing proved to be the most efficient method when comparing the delta areas and cumulative distribution

function (PDF) curves of various categorization outcomes (Fig. S2). Consequently, CAD patients were divided into phenotype 1 and phenotype 2. Furthermore, we discovered that phenotype 1 had marginally higher expression of genes linked to diabetes than phenotype 2 (Fig. 3C). Additional examination of the microenvironmental scores revealed that phenotype 1 outperformed phenotype 2 in terms of stromal, immune, and estimation scores (Figs. 3D–3F).

Identification of microenvironment characteristics of diabetes related phenotype in CAD

We delved into the microenvironmental differences between diabetes-related phenotypes in CAD. We compared the phenotypes with regard to cytokines and inflammatory factors, which are important determinants of the immunological and inflammatory milieu. Typically, phenotype 1 had higher amounts of cytokines and inflammatory factors such CXCL4, CCL5, CXCR8, CXCR9, and other regulatory molecules, while phenotype 2 had lower levels of these factors (Fig. 4A). Furthermore, phenotype 1 also exhibited significant expression levels of numerous interleukins and associated receptors, such as IL2, IL3, PDGFD, etc., which are known to regulate inflammation (Fig. 4A). We also examined each phenotype's immune cell composition. The estimation of immune cell infiltration was made using several algorithms. Phenotype 2 showed less immune cell infiltration than phenotype 1, which showed greater infiltration of immune cells such as CD8+ T cells, NK cells, etc. (Fig. 4B). Additionally, we examined the strength of endogenous and exogenous immunomodulation between the phenotypes. While endogenous immunity strength was similar between phenotypes, exogenous immunity was significantly stronger in phenotype 1 compared to phenotype 2 (Fig. 4C). Comparative analysis of classical immune features showed that promotive immune features were significantly more robust in phenotype 1 than in phenotype 2 (Fig. 4D).

Diabetes related phenotypes differ in myocardial lesion signaling and metabolic characteristics

Recognizing the significant distinctions between the two diabetes-related phenotypes, we looked at the differential pathways in the 2,180 differential genes that were found. Figure 5A shows a volcano diagram of these genes. We conducted Gene Set Variation Analysis (GSVA) for 10 changed pathways common to CAD and compared these between the two phenotypes, indicating considerable route variations (Fig. 5B). The subgroups' different metabolic and myocardial injury-related pathways were revealed using enrichment analysis. Gene Ontology (GO) analysis identified biological process differences primarily in "fatty acid metabolic process," "response to glucocorticoid," "heart morphogenesis," and other metabolism-related processes. Additionally, immune-related processes like "leukocyte cell–cell adhesion" and "regulation of innate immune response" were significantly enriched (Fig. 5C). KEGG analysis further enriched metabolic signals such as "Lipid and atherosclerosis" and "Carbon metabolism," along with myocardial lesion signals like "Diabetic cardiomyopathy" and "Cardiac muscle contraction" (Fig. 5D). We investigated metabolic differences using the MSigDB database, taking into account the variation in metabolic signals among different phenotypes. The most important route variations

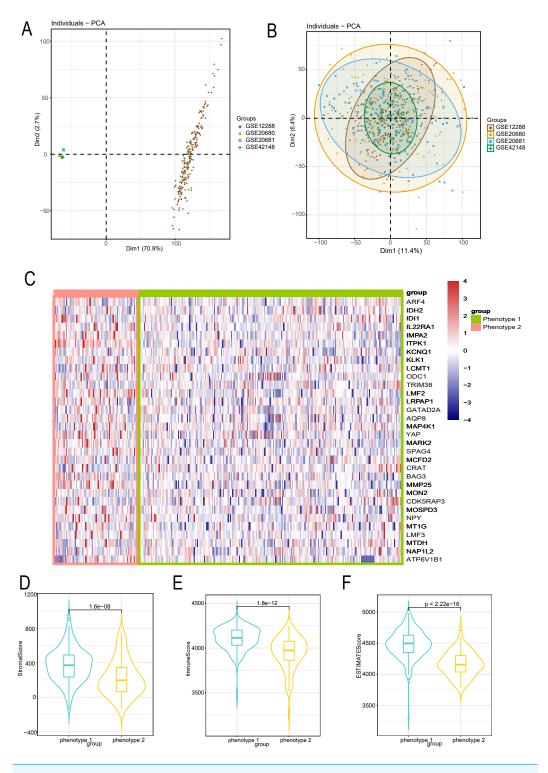


Figure 3 Diabetes related phenotype and characteristics in coronary artery disease (CAD). (A) The data difference among datasets. (B) The data distribution after fusion. (C) The CAD phenotype based on consensus clustering. (D–F) The microenvironmental score analysis between different phenotypes.

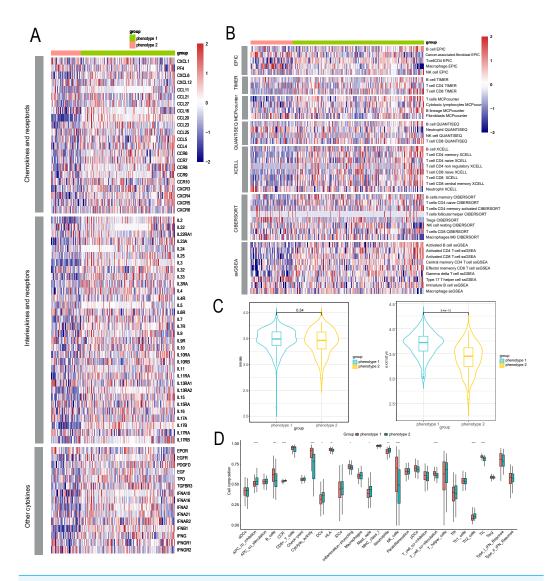


Figure 4 Identification of microenvironment characteristics of diabetes related phenotype in CAD. (A) The inflammation feature between different phenotypes. (B) The immune microenvironment feature between different phenotypes. (C) The endogenous and exogenous immunity feature between different phenotypes. (D) The classical immune features between different phenotypes.

are represented by a heatmap (Fig. 5E), which mainly shows the metabolic transport and absorption of proteins, fatty acids, glucose, and heart disease-related pathways. The significant distinctions between these symptoms, as revealed by the GSEA analysis, include alterations in metabolism, chemokines, and diabetic cardiomyopathy (Fig. 5F).

CAD diagnostic model constructed with diabetes-related genetic features shows excellent performance

To develop a more precise CAD diagnostic system, we further selected key genes for constructing a diabetes-related CAD typing using Lasso regression. A total of 32 core genes were included in the screening model (seed= 1, alpha=1, family="gaussian"), and final model

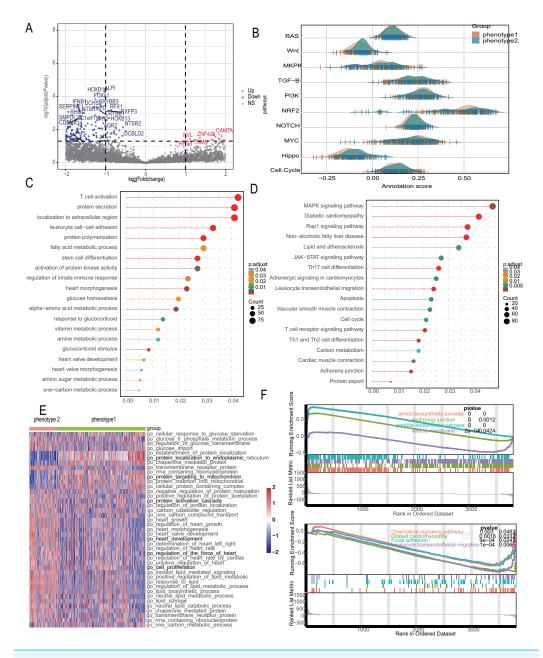


Figure 5 Diabetes related phenotypes differ in myocardial lesion signaling and metabolic characteristics. (A) Volcano diagram to show the differential gene between diabetes-related phenotypes. (B) Pathway variances in different phenotypes. (C) GO analysis based on the differential genes. (D) KEGG analysis based on the differential genes. (E) A heatmap to visualize the metabolic difference. (F) GSEA analysis between the phenotypes.

scores were derived using the predict algorithm. The final model incorporated 16 diabetes-related genes: "TRIM38," "SAPG4," "ODC1," "NPY," "MTDH," "MT1G," "LMF2," "KCNQ1," "ITPK1," "GATAD2A," "CRAT," "CDK5RAP3," "BAG3," "ATP6V1B1," "ARF4," "AQP8." (Figs. 6A, 6B). The modeling algorithm is as follows: model score= -0.0012369*exp (TRIM38) + 0.0017723*exp(SAPG4)+ 0.0014256*exp(ODC1)+

 $\begin{array}{lll} 0.0014256^* exp(NPY) + & 0.0014256^* exp(MTDH) + & 0.0016259^* exp(MT1G) \\ -0.0050854^* exp(LMF2) + & 0.0015382^* exp(KCNQ1) + & 0.0011388^* exp(ITPK1) + \\ 0.0087858 exp(GATAD2A) + & 0.0020088^* exp(CRAT) & -0.0065762^* exp(CDK5RAP3) \\ -0.00023941^* exp(BAG3) + & 0.00817318 exp(ATP6V1B1) + & 0.0065567^* exp(ARF4) + \\ 0.0075323^* exp(AQP8). \end{array}$

The predictive performance for each gene and the total LASSO diagnostic model was calculated. The areas under the receiver operating characteristic (ROC) curves were in Figs. 6C, 6D, 6E, respectively. Notably, when all 16 genes were included in a single diagnostic model, the AUC in the training dataset reached 0.823, and 0.758 in testing dataset (Fig. 6F, Fig. 6G).

Single-cell analysis explored CAD patient landscapes as well as expression patterns of modeled key genes

Further analysis was conducted on the expression landscape of diabetes-related genes using single-cell data from patients with coronary myocardial injury. After calculating and filtering the mitochondrial gene and ribosome ratios, as well as applying downscaling and tSNE clustering analyses (Fig. S3), we classified the patient cells into 11 subpopulations. These subpopulations were annotated using the "SingleR" package (Fig. 7A), identifying six cell types: macrophages, endothelial cells, monocytes, NK cells, muscle cells, and T cells (Fig. 7B). Macrophages and endothelial cells showed closer associations with other cells, indicating their potential roles in pathogenesis (Fig. 7C, Fig. S4). Dot plots were used to illustrate the top five genes differing among individual cells (Fig. 7D). The distribution of key markers in individual cells of the diabetes diagnostic model revealed that LMF2 and CDK5RAP3 were predominantly found in macrophages, while ARF4 was widely distributed among various immune and endothelial cells, as well as in muscle cells. KCNQ1, ATP6V1B1, MTDH, and ITPK1 were mainly distributed in macrophages (Fig. 7E, Fig. S5). Additionally, the heatmap exhibited the differential genes of various immune cells (Fig. S6).

KCNQ1 and ITPK1 were found to influence immune microenvironment of coronary artery disease

Building on the diagnostic model analysis, we further investigated the clinical relevance of these diagnostic markers. A mouse coronary artery disease model and a normal mouse control were established, with myocardial tissues removed, fixed, embedded, and subjected to IHC staining (Fig. 8A). We assessed the differential expression of the 16 gene diagnostic markers and some common immunological indicators. The results indicated upregulated expression of KCNQ1 and ITPK1 in mice with coronary artery disease, along with increased expression of CD31 and CD68. This suggests that KCNQ1 and ITPK1 are significant in the pathogenesis of CAD, potentially influencing disease progression by affecting the immune microenvironment in patients with CAD (Fig. 8B). We further compared mRNA levels in different subgroups and found the results to be consistent with those of histochemical staining (Fig. 8C). We further treated cells with CHD inducing conditions such as high glucose and fat, and then knocked down ITPK1 or KCNQ1. We found that high-glucose and high-fat conditions promoted the apoptosis rate of HUVEC, which was reversed by

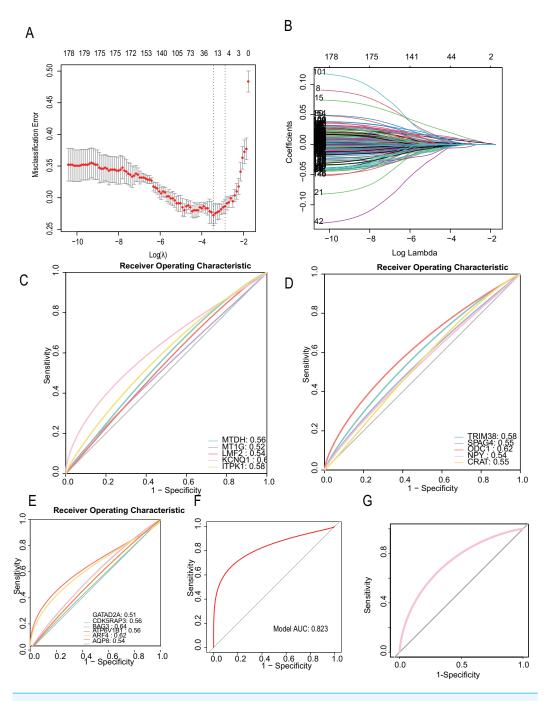


Figure 6 CAD diagnostic model constructed with diabetes-related genetic features shows excellent performance. (A–B) LASSO regression to construct the diagnostic model (the model constructing gene are "TRIM38," "SAPG4," "ODC1," "NPY," "MTDH," "MT1G," "LMF2," "KCNQ1," "ITPK1," "GATAD2A," "CRAT," "CDK5RAP3," "BAG3," "ATP6V1B1," "ARF4," "AQP8." Their Lasso coefficients are −0.0012369, 0.0017723, 0.0014256, 0.0016259, −0.0050854, 0.0015382, 0.0011388, 0.0087858, 0.0020088, −0.0065762, −0.00023941, 0.013747, 0.0081731, 0.0065567, 0.0075323). (C–E) The areas under the receiver operating characteristic (ROC) curves (continued on next page...)

Full-size DOI: 10.7717/peerj.20117/fig-6

Figure 6 (...continued)

on the model constructing genes (MTDH (AUC = 0.56), MT1G (AUC = 0.56), LMF2 (AUC = 0.54), KCNQ1 (AUC = 0.62), TTPK1 (AUC = 0.58), TRIM38 (AUC = 0.58), SPAG4 (AUC = 0.55), ODC1 (AUC = 0.62), NPY (AUC = 0.54), CRAT (AUC = 0.55), GATAD2A (AUC = 0.51), CDK5RAP3 (AUC = 0.56), BAG3 (AUC = 0.64), ATP6V1B1 (AUC = 0.56), ARF4 (AUC = 0.62), AQP8 (AUC = 0.54)). (F, G) The ROC curve on the diagnostic model (Training dataset). (G) The ROC curve on the diagnostic model (Testing dataset).

knockdown of KCNQ1 or ITPK1, suggesting the important role of ITPK1 or KCNQ1 in the development of CAD (Figs. 8D–8F).

DISCUSSION

In our current study, 32 diabetes-related biomarkers were identified based on differential expression analysis and WGCNA results. We further identified a diabetes related CAD phenotype based on the consensus clustering and exhibited its variation on cytokines, inflammatory factors, and immune cells infiltration. Utilizing a prediction algorithm, we identified 16 core diabetes related genes and construct a diabetes related CAD diagnostic model. Additionally, we found that KCNQ1 and ITPK1 may influence the immune microenvironment of CAD. CAD is a hazardous cardiovascular disease with complex pathology, the underage of which has made macro-control of the disease increasingly difficult (Vergès, 2020). Interestingly, patients with coronary artery disease often also have concomitant diabetes, which simultaneously exacerbates the progression of coronary artery disease. Previous literature indicated that CAD patients with type 2 diabetes mellitus (T2DM) often have increased epicardial adipose tissue (EAT) volume, correlating with higher cardiovascular mortality risk (Liu et al., 2023). Currently, combined diabetes and coronary heart disease analyses based on multiple datasets remain a vacuum; therefore, diagnostic models and biomarkers identified by combined analyses may lead to new directions and clinical practices for CAD treatment.

Several studies have underscored the importance of various genes in crosstalk between diabetes and CAD (*Deda et al.*, 2022). A review highlighted CHI3L1, CD36, LEPR, RETN, IL-18, RBP-4, and RARRES2 as potential genetic markers for insulin resistance (IR) and atherosclerosis in T2DM patients, which ultimately promote the progression of CAD (Liu et al., 2023). Additionally, genes like LDLR, CETP, APOA5, Apo E, Apo B, and Apo A-I are linked to CAD onset in T2DM patients, suggesting the potential for personalized diet-gene intervention therapies (Raj et al., 2015). In our current research, differential analysis and WGCNA analysis were used to extract diabetes-related pathogenic genes, and the immune and inflammatory microenvironment differences of diabetic comorbidity phenotypes were verified in the CAD cohort, which further verified the potential regulation of diabetesrelated genes on CAD. Additionally, we found similarities in endogenous immunity between the two subtypes as well as differences between exogenous immunity. The endogenous immune system consists of mast cells, eosinophils and basophils, and natural killer cells, while the exogenous immunity consists mainly of T cells and B cells (Sun et al., 2020). This suggests that T cells and B cells play a potential role in the differentiation of CAD subtypes. Interestingly, this is the first time to identify a coronary heart disease classification based on

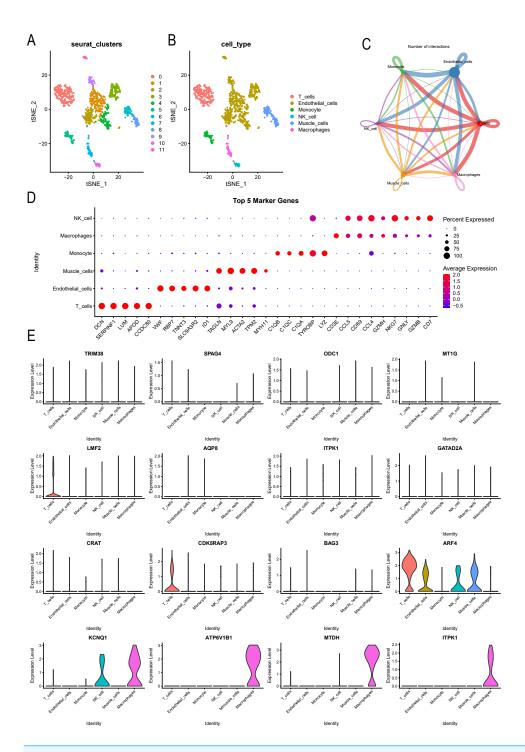


Figure 7 Single-cell analysis explored CAD patient landscapes as well as expression patterns of modeled key genes. (A–B) The subgroup annotation based on single cell analysis. (C) The association between different immune cells. (D) Dot plots to illustrate the top five genes differing among individual cells. (E) The distribution of key markers in individual cells.

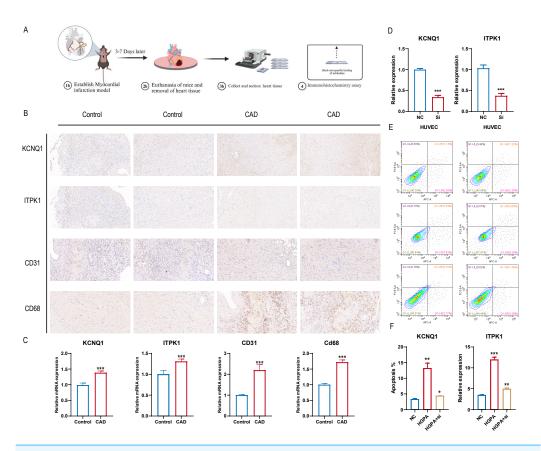


Figure 8 KCNQ1 and ITPK1 were found to influence immune microenvironment of coronary artery disease. (A) Workflow of mouse coronary artery disease model. (B) The co expression of KCNQ1 and ITPK1 with macrophage biomarkers in immunohistochemical staining. (C) The mRNA level of ITPK1 and KCNQ1 in various models. (D) Knockdown of KCNQ1 and ITPK1 in HUEVC cells. (E–F) Apoptosis rate of various groups.

diabetes-related characteristics, filling the gap in the clinical practice. Our classification is based on the consensus algorithm and is based on gene expression levels. By evaluating the expression levels of different genes, it accurately locates the internal microenvironment of different patients, which is conducive to precisely identifying the immune infiltration status of the patients. Similarly, our subtypes also have some limitations, including confounding factors such as sample heterogeneity when identifying diabetes-related genes and coronary heart disease-related genes, which may interfere with our results to a certain extent. In future research, it might be possible to assess the expression levels of the corresponding genes through expression probe, thereby further promoting and applying our classification.

We also found that KCNQ1 and ITPK1 may influence the immune microenvironment of CAD through the crosstalk with macrophages. Previous studies have identified the involvement of macrophages in the pathogenesis of diabetes and related CAD (*Henning*, 2018; *Peng & Gao*, 2022; *Yap et al.*, 2023). Dysregulation of macrophage polarization between M1 and M2 phenotypes alters the immune-inflammatory response in CAD, leading to plaque progression. Therefore, metabolic dysregulation of macrophages may

be one of the key features of CAD progression (Wang et al., 2024). CD31 and CD68 are key regulators of macrophages (Andreata et al., 2018; Garaicoa-Pazmino et al., 2019), and our study identified a co-expression profile between KCNQ1 and ITPK1 and these two regulators, thus speculating that they may regulate the pathogenesis and progression of CAD by regulating macrophage metabolism. In previous study, Zhu, Deng & Zhou (2018) identified KCNQ1 as a potential immunomodulatory gene in gout, and identified the pharmacological value of KCNQ1 in type 2 diabetes. Abnormalities in KCNQ1 have also been found to be significantly associated with arrhythmia (Kekenes-Huskey et al., 2022; Khatami, Mohajeri-Tehrani & Tavangar, 2019). Interestingly, KCNQ1 also plays a significant role in heart-related diseases. KCNQ1 exerts a crucial function in regulating potassium ions within the heart. Herbert et al. (2002) reported that mutations in KCNQ1 are closely associated with the occurrence of long QT syndrome, and they may cause cardiac dysfunction through a dominant negative mechanism. Similarly, KCNQ1 has been found to be closely associated with the occurrence of gestational diabetes. This might be caused by the imbalance in the release of insulin related to KCNQ1 (Herbert et al., 2002; Majcher et al., 2022). In our research, we also discovered that KCNQ1 plays a certain role in the apoptosis of vascular smooth muscle cells, which opens up a new research direction. To conclude, KCNQ1 has a certain impact on the occurrence of diabetes and heart-related diseases. This might be a new and significant biomarker for coronary heart disease. Research on ITPK1 remains scarce, with only one study labeling ITPK1 as being associated with migraine and thyroid dysfunction (Tasnim & Nyholt, 2023). In future studies, we may be able to obtain different macrophage phenotypic characteristics by reprogramming KCNQ1 as well as ITPK1 expression in adult macrophages, which may provide new ideas for the treatment of CAD.

In order to uncover potential molecular mechanisms and targets for diabetes-related CAD modifications, we have attached various concerns regarding the enrichment and analysis of pathways and molecular variances between different CAD phenotypes. Our data imply that phenotype -associated genes lead to changes in different pathways, including the Hippo system and NRF2 pathway. Interestingly, of the 32 biomarkers, both YAP and BAG3 were found to be associated with the Hippo and NRF2 pathways, which could be a potential cause of their changes (Moya & Halder, 2019; Wackerhage et al., 2019). Additionally, NRF2 was also found to be correlated with MT1G, the potential biomarker in CAD (Tang & Kang, 2024). we discovered that changed genes are primarily associated with pathway enrichment in pathways related to immunological, metabolic, and cardiac damage. Additional differentiation analysis of metabolic pathway enrichment revealed that protein, carbohydrate, and lipid metabolism are the primary metabolic reprogramming changes. Patients with CAD tend to experience changes in metabolic patterns, including alterations in glucose and fatty acid oxidative phosphorylation (Jiang et al., 2021). Previous studies also found that MAP4K1 may be involved in mitochondrial injury to alter energy metabolism. The upregulated MAP4K1 may be involved in the metabolic reprogramming process of CAD. These findings imply that metabolic reprogramming frequently occurs in tandem with changes in the diabetes-related phenotype. In general, synergistic investigation of the immune microenvironment and metabolism may lead to new discoveries regarding CAD treatment targets.

This study has limitations that must be acknowledged. The sample size was restricted, and some critical indices for CAD and DM were not consistent, such as probe cleaning, correction, annotation details, and sample clinical information. Although we have provided some additional information regarding functional assays for KCNQ1 and ITPK1, the experimental validation remains quite limited in scope. our conclusions need validation through more mechanism experiments.

Our study has some potential future applications in clinical flow. Our identified genetic markers are significantly upregulated in patients with coronary artery disease and have a high co-diagnostic value, they may be used for early CAD diagnosis in diabetic patients. We can integrate the diagnostic value of these genes into the clinical flow by constructing relevant gene probes, fitting them in patient blood samples, and detecting the relative expression of these biomarkers.

CONCLUSION

To conclude, we constructed a diabetes related CAD phenotype and an optimal diagnostic model. We found that different CAD phenotypes have large differences in inflammatory and immune microenvironment, and found that KCNQ1 and ITK1 are potential pathogenic genes and immune regulators of CAD.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was supported by the National Natural Science Foundation of China (No. 8216020109), the "Tianshan Talents" High-level Talent Training Program in Medicine and Healthcare (2023TSYCLJ0035), and the Key Project of the Natural Science Foundation of the Xinjiang Uygur Autonomous Region (2023D01D12). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

The National Natural Science Foundation of China: No. 8216020109.

The "Tianshan Talents" High-level Talent Training Program in Medicine and Healthcare: 2023TSYCLJ0035.

The Key Project of the Natural Science Foundation of the Xinjiang Uygur Autonomous Region: 2023D01D12.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

 Gulandanmu Aihemaiti performed the experiments, prepared figures and/or tables, and approved the final draft.

- Abulikemu Abula performed the experiments, prepared figures and/or tables, and approved the final draft.
- Junyi Luo analyzed the data, prepared figures and/or tables, and approved the final draft.
- Qian Zhao analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Ning Song analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Fen Liu analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Niyaziaili Adili analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Yining Yang conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Xiaomei Li conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

Animal Ethics

The following information was supplied relating to ethical approvals (*i.e.*, approving body and any reference numbers):

Animal Experimental Medicine Ethics Committee of the First Affiliated Hospital of Xinjiang Medical University, Approval number: 20210301-215 Approval date: 2021-03-01.

Data Availability

The following information was supplied regarding data availability: Data is available in the Supplemental Files.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.20117#supplemental-information.

REFERENCES

Andreata F, Syvannarath V, Clement M, Delbosc S, Guedj K, Fornasa G, Khallou-Laschet J, Morvan M, Even G, Procopio E, Gaston AT, Le Borgne M, Deschamps L, Nicoletti A, Caligiuri G. 2018. Macrophage CD31 signaling in dissecting aortic aneurysm. *Journal of the American College of Cardiology* 72:45–57 DOI 10.1016/j.jacc.2018.04.047.

Anonymous. 2025. 4. Comprehensive medical evaluation and assessment of comorbidities: standards of care in diabetes-2025. *Diabetes Care* **48**:S59–S85 DOI 10.2337/dc25-S004.

Bao J. 2022. Weighted gene coexpression network analysis identifies crucial genes involved in coronary atherosclerotic heart disease. *Disease Markers* **2022**:6971238 DOI 10.1155/2022/6971238.

- Ben Braiek A, Chahed H, Dumont F, Abdelhak F, Hichem D, Gamra H, Baudin B. 2021. Identification of biomarker panels as predictors of severity in coronary artery disease. *Journal of Cellular and Molecular Medicine* 25:1518–1530 DOI 10.1111/jcmm.16244.
- Chen W, Ni M, Huang H, Cong H, Fu X, Gao W, Yang Y, Yu M, Song X, Liu M, Yuan Z, Zhang B, Wang Z, Wang Y, Chen Y, Zhang C, Zhang Y. 2023. Chinese expert consensus on the diagnosis and treatment of coronary microvascular diseases (2023 Edition). *MedComm* 4:e438 DOI 10.1002/mco2.438.
- **Cole JB, Florez JC. 2020.** Genetics of diabetes mellitus and diabetes complications. *Nature Reviews Nephrology* **16**:377–390 DOI 10.1038/s41581-020-0278-5.
- **Darenskaya MA, Kolesnikova LI, Kolesnikov SI. 2021.** Oxidative stress: pathogenetic role in diabetes mellitus and its complications and therapeutic approaches to correction. *Bulletin of Experimental Biology and Medicine* **171**:179–189 DOI 10.1007/s10517-021-05191-7.
- Deda O, Panteris E, Meikopoulos T, Begou O, Mouskeftara T, Karagiannidis E, Papazoglou AS, Sianos G, Theodoridis G, Gika H. 2022. Correlation of serum acylcarnitines with clinical presentation and severity of coronary artery disease. *Biomolecules* 12 DOI 10.3390/biom12030354.
- Eid S, Sas KM, Abcouwer SF, Feldman EL, Gardner TW, Pennathur S, Fort PE. 2019.

 New insights into the mechanisms of diabetic complications: role of lipids and lipid metabolism. *Diabetologia* 62:1539–1549 DOI 10.1007/s00125-019-4959-1.
- Fan W, Pang H, Xie Z, Huang G, Zhou Z. 2022. Circular RNAs in diabetes mellitus and its complications. *Frontiers in Endocrinology* 13:885650 DOI 10.3389/fendo.2022.885650.
- Garaicoa-Pazmino C, Fretwurst T, Squarize CH, Berglundh T, Giannobile WV, Larsson L, Castilho RM. 2019. Characterization of macrophage polarization in periodontal disease. *Journal of Clinical Periodontology* 46:830–839 DOI 10.1111/jcpe.13156.
- **Henning RJ. 2018.** Type-2 diabetes mellitus and cardiovascular disease. *Future Cardiology* **14**:491–509 DOI 10.2217/fca-2018-0045.
- Herbert E, Trusz-Gluza M, Moric E, Smiłowska-Dzielicka E, Mazurek U, Wilczok T. 2002. KCNQ1 gene mutations and the respective genotype-phenotype correlations in the long QT syndrome. *Medical Science Monitor* **8**:Ra240–Ra248.
- **Iida M, Harada S, Takebayashi T. 2019.** Application of metabolomics to epidemiological studies of atherosclerosis and cardiovascular disease. *Journal of Atherosclerosis and Thrombosis* **26**:747–757 DOI 10.5551/jat.RV17036.
- **Jiang M, Xie X, Cao F, Wang Y. 2021.** Mitochondrial metabolism in myocardial remodeling and mechanical unloading: implications for ischemic heart disease. *Frontiers in Cardiovascular Medicine* **8**:789267 DOI 10.3389/fcvm.2021.789267.
- Kekenes-Huskey PM, Burgess DE, Sun B, Bartos DC, Rozmus ER, Anderson CL, January CT, Eckhardt LL, Delisle BP. 2022. Mutation-Specific Differences in Kv7.1 (KCNQ1) and Kv11.1 (KCNH2) channel dysfunction and long QT syndrome phenotypes. *International Journal of Molecular Sciences* 23 DOI 10.3390/ijms23137389.

- **Khatami F, Mohajeri-Tehrani MR, Tavangar SM. 2019.** The importance of precision medicine in Type 2 Diabetes Mellitus (T2DM): from pharmacogenetic and pharmacoepigenetic aspects. *Endocrine, Metabolic & Immune Disorders Drug Targets* **19**:719–731 DOI 10.2174/1871530319666190228102212.
- **Langfelder P, Horvath S. 2008.** WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**:559 DOI 10.1186/1471-2105-9-559.
- **Li DL, Kronenberg MW. 2021.** Myocardial perfusion and viability imaging in coronary artery disease: clinical value in diagnosis, prognosis, and therapeutic guidance. *American Journal of Medicine* **134**:968–975 DOI 10.1016/j.amjmed.2021.03.011.
- Liu Z, Neuber S, Klose K, Jiang M, Kelle S, Zhou N, Wang S, Stamm C, Luo F. 2023. Relationship between epicardial adipose tissue attenuation and coronary artery disease in type 2 diabetes mellitus patients. *Journal of Cardiovascular Medicine* 24:244–252 DOI 10.2459/jcm.0000000000001454.
- Majcher S, Ustianowski P, Malinowski D, Czerewaty M, Tarnowski M, Safranow K, Dziedziejko V, Pawlik A. 2022. KCNJ11 and KCNQ1 gene polymorphisms and placental expression in women with gestational Diabetes Mellitus. *Gene* 13 DOI 10.3390/genes13081315.
- Malakar AK, Choudhury D, Halder B, Paul P, Uddin A, Chakraborty S. 2019. A review on coronary artery disease, its risk factors, and therapeutics. *Journal of Cellular Physiology* 234:16812–16823 DOI 10.1002/jcp.28350.
- **Moya IM, Halder G. 2019.** Hippo-YAP/TAZ signalling in organ regeneration and regenerative medicine. *Nature Reviews Molecular Cell Biology* **20**:211–226 DOI 10.1038/s41580-018-0086-y.
- Peng Y, Gao Y. 2022. NLRP3+ macrophages aggravate inflammatory cystitis in diabetes. *Frontiers in Immunology* **13**:1057746 DOI 10.3389/fimmu.2022.1057746.
- Pignalosa FC, Desiderio A, Mirra P, Nigro C, Perruolo G, Ulianich L, Formisano P, Beguinot F, Miele C, Napoli R, Fiory F. 2021. Diabetes and cognitive impairment: a role for glucotoxicity and dopaminergic dysfunction. *International Journal of Molecular Sciences* 22 DOI 10.3390/ijms222212366.
- **Raj R, Bhatti JS, Badada SK, Ramteke PW. 2015.** Genetic basis of dyslipidemia in disease precipitation of coronary artery disease (CAD) associated type 2 diabetes mellitus (T2DM). *Diabetes/Metabolism Research and Reviews* **31**:663–671 DOI 10.1002/dmrr.2630.
- Sun L, Wang X, Saredy J, Yuan Z, Yang X, Wang H. 2020. Innate-adaptive immunity interplay and redox regulation in immune response. *Redox Biology* 37:101759 DOI 10.1016/j.redox.2020.101759.
- **Tang D, Kang R. 2024.** NFE2L2 and ferroptosis resistance in cancer therapy. *Cancer Drug Resistance* 7:41–41 DOI 10.20517/cdr.2024.123.
- **Tasnim S, Nyholt DR. 2023.** Migraine and thyroid dysfunction: co-occurrence, shared genes and biological mechanisms. *European Journal of Neurology* **30**:1815–1827 DOI 10.1111/ene.15753.
- **Tavitian A, Lax E, Song W, Szyf M, Schipper HM. 2025.** Hippocampal reelin and GAD67 gene expression and methylation in the GFAP.HMOX1 mouse model

- of schizophrenia. *Biochimica et Biophysica Acta: Molecular Cell Research* 119899 DOI 10.1016/j.bbamcr.2025.119899.
- Tseng PY, Chen YT, Wang CH, Chiu KM, Peng YS, Hsu SP, Chen KL, Yang CY, Lee OK. 2020. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. *Critical Care* 24:478 DOI 10.1186/s13054-020-03179-9.
- **Vergès B. 2020.** Cardiovascular disease in type 1 diabetes: a review of epidemiological data and underlying mechanisms. *Diabete et Metabolisme* **46**:442–449 DOI 10.1016/j.diabet.2020.09.001.
- Wackerhage H, Schoenfeld BJ, Hamilton DL, Lehti M, Hulmi JJ. 2019. Stimuli and sensors that initiate skeletal muscle hypertrophy following resistance exercise. *Journal of Applied Physiology* 126:30–43 DOI 10.1152/japplphysiol.00685.2018.
- Wang Y, Eshwaran R, Beck SC, Hammes HP, Wieland T, Feng Y. 2023. Contribution of the hexosamine biosynthetic pathway in the hyperglycemia-dependent and independent breakdown of the retinal neurovascular unit. *Molecular Metabolism* 73:101736 DOI 10.1016/j.molmet.2023.101736.
- Wang J, Wu Q, Wang X, Liu H, Chen M, Xu L, Zhang Z, Li K, Li W, Zhong J. 2024. Targeting macrophage phenotypes and metabolism as novel therapeutic approaches in atherosclerosis and related cardiovascular diseases. *Current Atherosclerosis Reports* 26:573–588 DOI 10.1007/s11883-024-01229-z.
- **Xia M, Wu Q, Chen P, Qian C. 2021.** Regulatory T cell-related gene biomarkers in the deterioration of atherosclerosis. *Frontiers in Cardiovascular Medicine* **8**:661709 DOI 10.3389/fcvm.2021.661709.
- Xu J, Zhao Q, Li J, Yuan Y, Cao X, Zhang X, Fang J, Yan W, Wang B, Li Y, Chu Y. 2023. Validation of a predictive model for coronary artery disease in patients with diabetes. *Journal of Cardiovascular Medicine* 24:36–43 DOI 10.2459/jcm.0000000000001387.
- Yahagi K, Kolodgie FD, Lutter C, Mori H, Romero ME, Finn AV, Virmani R. 2017. Pathology of human coronary and carotid artery atherosclerosis and vascular calcification in diabetes mellitus. *Arteriosclerosis, Thrombosis, and Vascular Biology* 37:191–204 DOI 10.1161/atvbaha.116.306256.
- Yap J, Irei J, Lozano-Gerona J, Vanapruks S, Bishop T, Boisvert WA. 2023. Macrophages in cardiac remodelling after myocardial infarction. *Nature Reviews Cardiology* 20:373–385 DOI 10.1038/s41569-022-00823-5.
- **Zhu W, Deng Y, Zhou X. 2018.** Multiple membrane transporters and some immune regulatory genes are major genetic factors to gout. *The Open Rheumatology Journal* **12:94**–113 DOI 10.2174/1874312901812010094.