

RESEARCH

Open Access



# Characterizing the metabolic divide: distinctive metabolites differentiating CAD-T2DM from CAD patients

Yingjian Liu<sup>1,2,3†</sup>, Ju-e Liu<sup>2†</sup>, Huafeng He<sup>1,2,3†</sup>, Min Qin<sup>1,2,3</sup>, Heping Lei<sup>3</sup>, Jinxiu Meng<sup>3</sup>, Chen Liu<sup>4</sup>, Xiaoping Chen<sup>5</sup>, Wenwei Luo<sup>2\*</sup> and Shilong Zhong<sup>1,2,3\*</sup>

## Abstract

**Objective** To delineate the metabolomic differences in plasma samples between patients with coronary artery disease (CAD) and those with concomitant CAD and type 2 diabetes mellitus (T2DM), and to pinpoint distinctive metabolites indicative of T2DM risk.

**Method** Plasma samples from CAD and CAD-T2DM patients across three centers underwent comprehensive metabolomic and lipidomic analyses. Multivariate logistic regression was employed to discern the relationship between the identified metabolites and T2DM risk. Characteristic metabolites' metabolic impacts were further probed through hepatocyte cellular experiments. Subsequent transcriptomic analyses elucidated the potential target sites explaining the metabolic actions of these metabolites.

**Results** Metabolomic analysis revealed 192 and 95 significantly altered profiles in the discovery (FDR < 0.05) and validation (P < 0.05) cohorts, respectively, that were associated with T2DM risk in univariate logistic regression. Further multivariate regression analyses identified 22 characteristic metabolites consistently associated with T2DM risk in both cohorts. Notably, pipercolinic acid and L-pipercolic acid, lysine derivatives, exhibited negative association with CAD-T2DM and influenced cellular glucose metabolism in hepatocytes. Transcriptomic insights shed light on potential metabolic action sites of these metabolites.

**Conclusions** This research underscores the metabolic disparities between CAD and CAD-T2DM patients, spotlighting the protective attributes of pipercolinic acid and L-pipercolic acid. The comprehensive metabolomic and transcriptomic findings provide novel insights into the mechanism research, prophylaxis and treatment of comorbidity of CAD and T2DM.

**Keywords** Metabolomics, Lipidomics, Coronary artery disease, Type 2 diabetes mellitus, Glucose metabolism

<sup>†</sup>Yingjian Liu, Ju-e Liu and Huafeng He have contributed equally to this work.

\*Correspondence:

Wenwei Luo

luowenwei@gdph.org.cn

Shilong Zhong

shz2020@qq.com

Full list of author information is available at the end of the article



## Introduction

Cardiovascular disease coupled with diabetes poses significant health challenges globally [1]. Diabetes is a risk factor for coronary artery disease (CAD), independent of traditional risk factors such as hyperlipidemia, hypertension and smoking [2]. Type 2 diabetes mellitus (T2DM) is prevalent in more than half of patients with CAD. People with both CAD and T2DM have a higher risk of cardiovascular events than those with CAD alone [3–5]. T2DM and CAD are often associated with each other, and both of them are metabolic syndrome [6, 7]. T2DM leads to an approximately two-fold increased risk of CAD, which in return serves as a major contributor to death and disability in T2DM patients [8]. CAD caused by T2DM is a complex metabolic disease process involving multiple physiological networks and different pathways. Identifying the key factors that contribute to disease progression is therefore critical in the context of effective disease management. Because of the limitations of traditional risk markers and risk prediction models for predicting cardiovascular disease in diabetic patients [9], it is particularly important to identify new specific markers for CAD-T2DM.

Metabolism is a fundamental feature of life, allowing organisms to adapt to changes in the internal and external environment to maintain vital functions. Metabolites are terminals or by-products of cellular regulatory or biochemical processes, and thus have the potential to provide more detail on the biological pathways that are disturbed [10]. Detecting these key metabolite signatures and understanding the function of these metabolic pathways are fundamental to identifying valuable biomarkers for better disease management. With the advancement of metabolomic technology, more than 21,000 annotated metabolites have been discovered in the human body, including 1,581,000 unannotated metabolite entries [11], highlighting the immense potential of metabolomics in disease research. Metabolomics, which includes lipidomics, can provide a comprehensive view of metabolites in biological systems and has become a powerful tool for unraveling complex biochemical pathways in various diseases [12–14]. Numerous studies have harnessed metabolomic techniques to uncover differential metabolites and potential biomarkers in CAD (e.g., creatine, lysophosphatidylcholine (16:0) and T2DM (e.g., tryptophan, branched chain amino acids, phospholipids), which hold clinical relevance for disease onset, stratification, and prognosis [15–22]. However, there is a glaring research void exists concerning the metabolic contrasts between CAD patients and those concurrently diagnosed with both CAD and T2DM [23]. This gap underscores the imperative for a dedicated exploration into the distinct metabolic shifts in this particular subset of patients.

Metabolites play a critical role in maintaining metabolic homeostasis by influencing energy metabolism (e.g., L-tryptophan, phenylalanine, triglycerides [24–26]). Although fatty acid metabolism dominates cardiac energy metabolism, glucose metabolism is also an important source of additional energy for the damaged heart [27]. However, in patients with T2DM cardiomyopathy, abnormal myocardial fatty acids activate PPAR- $\alpha$ , which upregulates fatty acid uptake, storage and  $\beta$ -oxidation while inhibiting glucose utilization [28]. Therefore, understanding the distinct small-molecule profiles of metabolites in the combined disease pattern can not only provide a basis for elucidation the potential pathogenesis of CAD by T2DM, but may also reveal potential therapeutic targets for improving the prognosis of patients with CAD through key hubs in energy metabolism.

To explore the specific metabolic biomarkers of CAD-T2DM, based on a 3-center cohort of 1,465 CAD patients, our study distinguished the plasma metabolomic characteristics of CAD patients between with and without T2DM through a widely targeted metabolomics and lipidome atlas, and constructed a high-performance model for predicting the all-cause mortality using the identified biomarkers. In addition, through transcriptome sequencing and cell validation experiments, the action pathways of key metabolites were explored, which provided a new basis for clinical diagnosis, treatment and risk prediction of CAD/T2DM comorbid disease.

## Methods

### Inclusion criteria

This study used a single-center discovery cohort from the Department of Cardiology, Guangdong Provincial People's Hospital. From 2010 to 2014, a total of 952 patients diagnosed with CAD were enrolled. The multicenter validation cohort incorporated patients from three distinct hospitals: Guangdong Provincial People's Hospital (n=350), Xiangya Hospital of Central South University (n=178), and the First Affiliated Hospital of Sun Yat-sen University (n=33), enrolling a total of 561 CAD patients between September 2017 and October 2018. A significant portion of the patients in this study had also been part of our prior metabolomics and lipidomics research [29].

**Diagnostic Criteria for CAD:** Patients admitted for symptoms such as chest tightness or chest pain undergo a comprehensive evaluation on admission. After exclusion of contraindications, coronary angiography is performed to assess the extent and severity of coronary artery lesions. If the results show direct narrowing of one or more coronary artery lumina by less than 50% and the patient reports typical angina symptoms on admission, a diagnosis of coronary artery disease can be confirmed.

Diagnostic criteria by the American Diabetes Association [30] include the following: A fasting plasma glucose level of 126 mg/dL (7.0 mmol/L) or higher, a 2 h plasma glucose level of 200 mg/dL (11.1 mmol/L) or higher during a 75 g oral glucose tolerance test, a random plasma glucose of 200 mg/dL (11.1 mmol/L) or higher in a patient with classic symptoms of hyperglycemia or hyperglycemic crisis, or A hemoglobin A1c level of 6.5% (48 mmol/mol) or higher.

Inclusion criteria were as follows: (1) Chinese Han individuals aged between 18 and 80 years without consanguinity; (2) Patients diagnosed with CAD via coronary angiography and underwent percutaneous coronary intervention (PCI); (3) Patients diagnosed with T2DM based on the established T2DM diagnostic criteria or those with a T2DM history, as well as those without diabetes. Exclusion criteria (Additional file 1: Figure S1): (1) Patients with renal dysfunction, i.e., those with a history of renal transplantation, dialysis, or serum creatinine concentration >3 times the upper limit of normal (345  $\mu$ mol/L); (2) Patients with hepatic impairment, diagnosed cirrhosis, or serum transaminase concentration >3 times the upper limit of normal (120 U/L); (3) Pregnant or lactating women; (4) Patients with severe infections, or other terminal illnesses; (5) Patients with autoimmune disease diseases that affect glucose and lipid metabolism, such as advanced cancer, history of thyroid problems and taking antithyroid drugs or thyroid hormone medication; (6) Patients with poor compliance who are unable to complete the study; (7) Patients with type 1 diabetes, gestational diabetes, or certain types of diabetes; (8) Patients currently taking medications (e.g., hormones) that affect glucose and lipid metabolism.

#### Baseline data collection

Comprehensive baseline clinical data were collected for all participants, including general details such as age, gender, height, weight, and medical history, including the presence of CAD, diabetes, hypertension, hyperlipidemia. Clinical biochemical indicators such as blood lipids, fasting blood glucose (GLUC), hemoglobin A1c (HbA1c), liver function indicators, creatinine (CREA) and creatine kinase (CK) were also recorded.

#### Plasma sample collection

Patients were instructed to fast for at least 8 h to reduce the effect of nutrient intake on analyte levels. On the day of enrollment, 4 mL of venous blood was collected and stored at 4 °C in ethylenediaminetetraacetic acid (EDTA) anticoagulant tubes. Within 2 h, samples were centrifuged at 4 °C for 10 min at 3000 g to separate plasma and blood cells. The separated plasma and blood cells were

then aliquoted, labeled, and stored at – 80 °C for later use.

#### Widely targeted metabolomics analysis

Widely targeted metabolomic profiling was conducted in the plasma sample of the discovery and multicenter validation cohorts in March 2017 and May 2019, respectively.

In the discovery cohort, for sample extraction, plasma was firstly thawed at 4 °C and then vortexed for 10 s. Then, 50  $\mu$ L of plasma was transferred to 150  $\mu$ L of pre-cooled methanol to precipitate proteins, vortexed for 3 min at room temperature, and then centrifuged at 12,000 rpm for 10 min at 4 °C. Thereafter, the supernatant was centrifuged again at 12,000 rpm for 3 min at 4 °C. Finally, an aliquot of the resulting supernatant was used for metabolomic analysis. Metabolomic profiling was conducted on a Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometric (LC–ESI–MS/MS) system (UPLC, Shim-pack UFLC SHIMADZU CBM30A; MS, Applied Biosystems 4500 QTRAP). In total, 202 annotated metabolites and 667 lipid species were identified and quantified.

In the multicenter validation cohort, plasma was thawed on ice, and 150  $\mu$ L of ice-cold methanol was added to 50  $\mu$ L of plasma. The mixture was vortexed for 3 min and then centrifuged at 12,000 rpm for 10 min at 4 °C. The supernatant was collected and then centrifuged at 12,000 rpm for 5 min at 4 °C. Finally, the resulting supernatant was used for Ultraperformance Liquid Chromatography Tandem Mass Spectrometry (UPLC–MS/MS) analysis. The sample extracts were analyzed using an LC–ESI–MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A; MS, Applied Biosystems 6500+QTRAP). In total, more than 600 metabolites and 600 lipid species have been annotated. 161 metabolites and 309 lipid species are identical to the discovery cohort.

#### Qualitative and quantitative analysis

Qualitative analysis of the precursor ion and fragments spectra detected was carried out based on self-built Metware database (MWDB) with retention time and ion pairs, as well as the public database of metabolites information. We used MS/MS spectra to search against public databases to improve confidence in metabolite identification. Some of these substances are qualitatively analyzed with removing isotopic signals, repetitive signals containing K<sup>+</sup> ions, Na<sup>+</sup> ions, and NH<sub>4</sub><sup>+</sup> ions, and repeated signals of fragmented ions that themselves are of larger molecular weight. Metabolite structure resolution is referenced in existing mass spectrometry public databases such as MassBank (<http://www.massbank.jp/>), HMDB (<http://www.hmdb.ca/>) and METLIN (

[metlin.scripps.edu/index.php](https://metlin.scripps.edu/index.php)). Quantitation of metabolites was accomplished using multiple reaction monitoring (MRM) of triple quadrupole mass spectrometry. Under the MRM mode, the quadrupole rod first screened precursor ions (parent ions) of the target substance to exclude ions corresponding to other molecular weight substances and to preliminarily eliminate the interference. The precursor ions were induced to ionize in the collision cell to form many fragment ions, fragment ions. The fragment ions were then filtered through the triple four-stage bar to select the necessary characteristic fragment ion needed, which eliminated the interference of non-target ions, thereby making the quantification of better accuracy and repeatability. Raw data from the analysis were processed using Analyst 1.6 software (AB Sciex). After obtaining the metabolite spectrum analysis data of different samples, integration of peak areas was performed for the mass spectrum peaks, and the integral correction was performed for the mass spectrum peaks of the same metabolite in different samples. The Quality Control (QC) sample is a mixture prepared from all samples. After every 10 samples analyzed, a QC sample is inserted to monitor and correct the stability and repeatability of the instrument (Additional file 1: Figure S2). The ionization detection modes and ion pair information for metabolites and lipids are detailed in Additional file 2: Tables S1, S2.

#### Quality control

Patients were screened based on the above inclusion and exclusion criteria to determine the final samples included in the analysis. In the baseline characteristics of the study subjects, continuous variables with non-Gaussian distribution were presented as median (interquartile range) and compared using the Mann–Whitney U test (for non-normal distribution). Categorical variables were presented as counts (percentages) and compared using the chi-squared test. Statistical significance was set at  $p < 0.05$ .

Before analysis, QC-RLSC (quality control-based robust LOESS signal correction) was used to correct the levels of metabolites detected in different batches to minimize errors caused by batch-to-batch variation [31]. After correction, the data were normalized using the Pareto scaling method [32], which includes mean normalization and square root variance transformation (Additional file 1: Figure S3). After the above corrections and standardizations, the data were ready for further analysis.

#### Data analysis for metabolomics

Logistic regression and correlation analyses were performed using built-in functions in R (version 4.2.2).

The basic function “glm” in R is used to perform univariate and multivariate logistic regression. For the differentiation of metabolomics between CAD patients without diabetes and those with diabetes, univariate logistic regression was employed. To control for potential confounders, traditional clinical risk factors such as age, alanine aminotransferase (ALT), aspartate aminotransferase (AST), body mass index (BMI), estimated glomerular filtration rate (eGFR), hypertension, hyperlipidemia, and gender were incorporated into the multivariate logistic regression. Metabolites with a Benjamini–Hochberg False Discovery Rate (FDR) adjusted  $p$  value less than 0.05 were subjected to Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis using the MetaboAnalyst online tool (<https://www.metaboanalyst.ca>). In the multivariate logistic regression results, metabolites with  $FDR < 0.05$  in the discovery cohort and  $P < 0.05$  in the validation cohort were defined as characteristic metabolites (CMs). After excluding a single metabolite with a contrasting OR direction, only 22 CMs were remained. The correlation of these metabolites with HbA1c and fasting glucose concentrations was determined using the built-in R function (“cor”). For the development of a diagnostic model for classifying endpoint events of all-cause mortality in coronary heart disease, Least Absolute Shrinkage and Selection Operator (LASSO) [33] selection analysis was applied to clinical variables and characteristic metabolite variables. The optimal value of the adjustment parameter  $\lambda$  was determined by fivefold cross-validation (200 iterations). The selected biomarkers were then used as covariates to construct a binary logistic regression model. The predictive ability and effectiveness of the model were assessed using the area under the curve (AUC) of the receiver operating characteristic (ROC). The DeLong test was then used to compare the AUC values of different models. Visualization of major metabolites was facilitated by the R packages “forestplot” and “pheatmap”.

#### Reagents and cells

HepaRG cells were obtained from our research group. Pilocolinic acid (CID: 849), L-pipecolic acid (CID: 439227), and 5-oxoproline (CID: 7405) were purchased from Shanghai Macklin Biochemical Technology Co., Ltd. Dulbecco's modified eagle medium (DMEM) culture medium and fetal bovine serum (FBS) were provided by Gibco. The 1% antibiotics (100× streptomycin-penicillin) were purchased from Procell, and Trizol RNA extraction reagent was purchased from Axygen Biosciences (AG). In addition, the glucose and lactate assay kits were provided by Nanjing Jiancheng Tech Co., Ltd. and the ATP assay kit and bovine serum albumin (BSA) reagent kit were provided by Shanghai Beyotime Biotechnology.

### Preparation of metabolite stock solution

For metabolite intervention, metabolites were dissolved in complete DMEM culture medium (containing 10% FBS and 1% antibiotics) to formulate a 75 mM stock solution. This solution was stored at  $-20^{\circ}\text{C}$  and used within one week. During the experiment, the stock solution was diluted to the desired concentrations (50  $\mu\text{M}$ , 250  $\mu\text{M}$ , 500  $\mu\text{M}$ ).

### Assay for glucose concentration

5E4 HepaRG cells were seeded into each well of a 12-well plate. After 24 h of attachment, the supernatant was discarded and the cells were rinsed three times with phosphate-buffered saline (PBS). Culture medium containing different drug concentrations (50  $\mu\text{M}$ , 250  $\mu\text{M}$ , 500  $\mu\text{M}$ ) was then added to each well. No drug treatment was given in the negative control group. A blank well without cells was also prepared to which standard complete DMEM culture medium was added. After 60 h of drug treatment, 500  $\mu\text{M}$  supernatant was collected and centrifuged at 1000 rpm for 10 min. Glucose concentration was then determined according to the guidelines of the Jiancheng Glucose Assay Kit.  $\text{Glucose Consumption} = \text{Glucose Concentration in Complete Medium} - \text{Glucose Concentration in Supernatant after 60 h}$ . Glucose consumption was expressed as the ratio of glucose consumption in each intervention group to the negative control group.

### Measurement of lactate levels

Under metabolite intervention conditions, a concentration of 50  $\mu\text{M}$  was chosen. After 60 h of intervention, 200  $\mu\text{M}$  trypsin was added to each well for 2 min of digestion. Once the cells were completely detached, 200  $\mu\text{M}$  complete culture medium was added to stop the digestion and the samples were collected in 1.5 mL centrifuge tubes. The samples were then centrifuged at 1000 rpm for 5 min, the supernatant was gently discarded, and 200  $\mu\text{L}$  PBS was added to each tube for resuspension. Cells were lysed using an ultrasonic cell disruptor. The supernatant was then collected, and lactate levels were determined using the Jiancheng lactate assay kit according to its instructions.

### ATP assay

10E4 HepaRG cells were seeded in each well of a 6-well plate. The metabolite was adjusted to a concentration of 50  $\mu\text{M}$ . After treatment, 200  $\mu\text{L}$  ATP lysis buffer was added to each well and the cells were lysed on ice for 5 min. The lysate was then centrifuged at 12,000 g for 5 min. A portion of the supernatant was collected and mixed with the ATP detection working solution in a

96-well plate. The concentrations of the calibration curve were determined. After a minimum reaction time of 3 s, chemiluminescence was detected. The results were expressed as the ratio of ATP content in each intervention group to the negative control group.

### Bovine serum albumin (BSA) assay

Similarly, 10E4 HepaRG cells were seeded in each well of a 6-well plate. Each metabolite was added at a concentration of 50  $\mu\text{M}$ . After treatment, 200  $\mu\text{L}$  lysis buffer was added to each well and the cells were lysed on ice for 5 min. The lysate was then centrifuged at 13,000 rpm for 18 min. A portion of the supernatant was removed and mixed with the BSA detection working solution in a 96-well plate. The concentrations of the calibration curve were determined. After a minimum reaction time of 3 s, the absorbance at 540 nm was measured.

### Statistical analysis

Each sample within the same batch of data was measured at least twice and the mean was calculated. For biological replicates, the experiment was independently repeated at least three times. One-way analysis of variance was used to test for statistical differences in GraphPad Prism 9, with a P value of less than 0.05 considered significant.

### RNA sequence assay and transcriptomics analysis

10E4 HepaRG cells were seeded in each well of a 6-well plate. After 24 h of attachment, cells were treated with regular complete culture medium, 50  $\mu\text{M}$  pipercolic acid, and 50  $\mu\text{M}$  L-pipercolic acid complete culture medium for 60 h. Total RNA was extracted using the Trizol reagent kit according to the manufacturer's instructions. RNA quality was evaluated and verified. After total RNA extraction, eukaryotic mRNA was enriched. The enriched mRNA was then fragmented and reverse transcribed to cDNA. The purified double-stranded cDNA fragments underwent several processes before sequencing by Gene Denovo Biotechnology Co. on the Illumina Novaseq6000 platform. Post-filtering of clean reads, the data was aligned first with Ribosome RNA (rRNA) and then with the reference genome. To ensure the precision of subsequent analyses, corrections were made for gene length or transcript length, followed by a correction for sequencing depth. After obtaining Transcripts Per Million (TPM) values for genes, further analyses were performed. Principal component analysis (PCA) was performed using the R package "gmodels". Differential expression analysis of RNAs was performed using "DESeq2". Transcripts with a p-value below 0.05 and an absolute fold change of  $\geq 1.3$  were classified as differentially expressed genes/transcripts. The temporal expression profile of these genes was then examined, followed by KEGG enrichment

analyses for each profile. All of the above statistical analyses were performed using the OmicShare tools.

**Results**

**Study workflow overview**

The study workflow is shown in Fig. 1, created with BioRender.com.

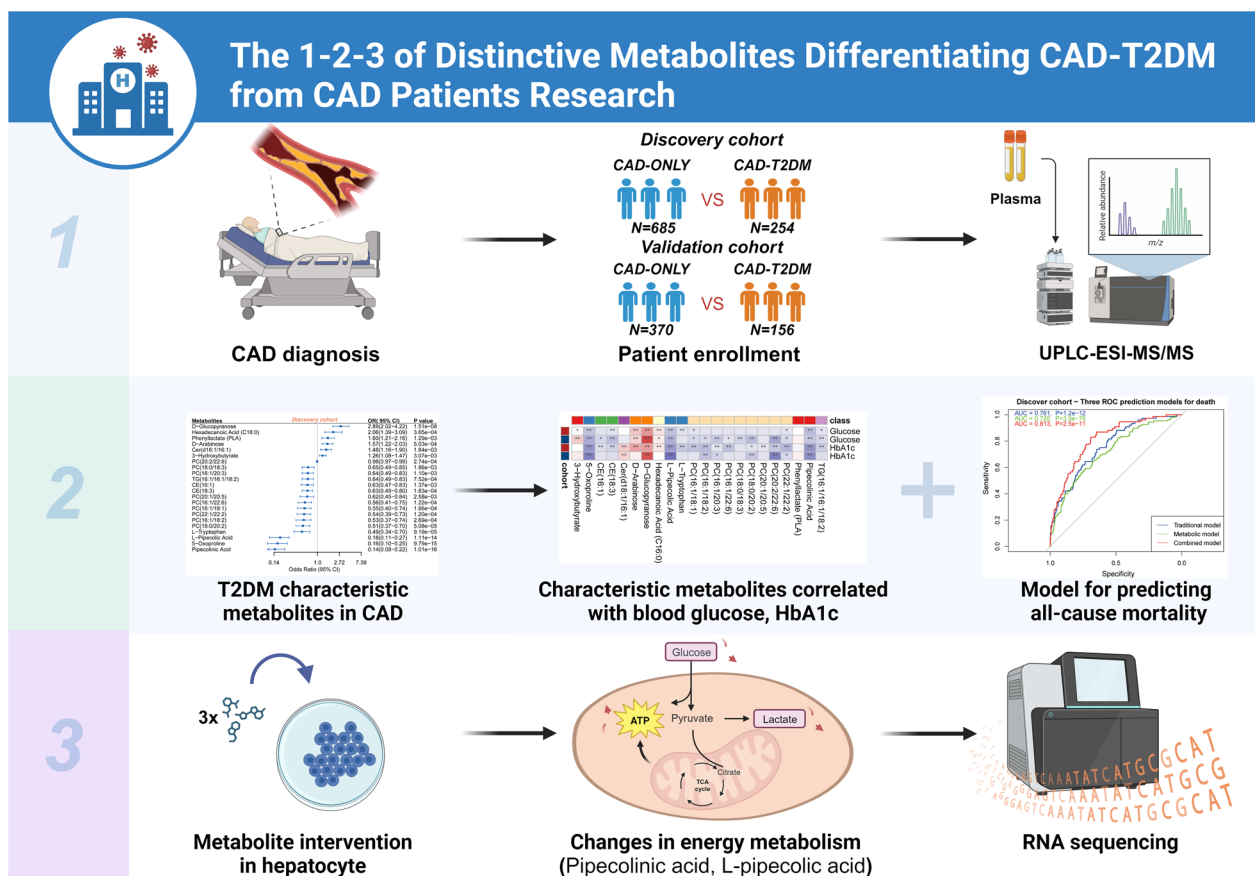
**Participant profile overview**

In this study, we enrolled a total of 1465 patients diagnosed with CAD from three different medical centers in China. The discovery cohort, which originated from the Guangdong Provincial People’s Hospital, included 685 patients with CAD and an additional subset of 254 patients who manifested a concomitant presence of CAD and T2DM (Table 1). Compared to the CAD-ONLY group, the CAD-T2DM group had a slightly older mean age and a female predominance. Hypertension was more

common in the CAD-T2DM group, while the incidence of hyperlipidemia was similar in both groups. Biochemical markers such as triglyceride (TG), fasting blood glucose (GLUC) and hemoglobin A1c (HbA1c) were significantly higher in the CAD-T2DM group ( $P < 0.05$ ). The baseline information for the validation cohort of 526 patients is presented in Additional file 2 Table S3.

**Metabolomic landscape in CAD-T2DM patients**

To delineate the intricate relationship between specific metabolites and CAD-T2DM, initial univariate logistic regression analyses were conducted on an array of 202 annotated metabolites and 667 lipidomic entities. Subsequent to FDR adjustment, a total of 192 metabolites remained significantly correlated with T2DM ( $FDR < 0.05$ ) (Additional file 2: Table S4). The validation cohort corroborated the significance of 95 metabolites ( $P < 0.05$ ) (Additional file 2: Table S5). Further



**Fig. 1** Study design There were 3 steps to the design of this study. Phase 1: Patients with CAD, a total of 1465 individuals from three centers, were stratified into discovery and validation groups based on different enrollment times. widely metabolomics was used to detect plasma metabolites and lipids. Phase 2: Using multivariate logistic regression, 22 characteristic metabolites (CMs) were identified that were associated with CAD-T2DM (vs. CAD-ONLY). Relationships between these CMs and fasting blood glucose, HbA1c and all-cause mortality were established. Phase 3: At the cellular level, three CMs significantly negatively correlated with T2DM were validated. Only pipecolic acid and L-pipecolic acid induced changes in cellular energy metabolism. Finally, transcriptional data provided insight into potential targets and pathways affected by these two metabolites

**Table 1** Clinical baseline data of patients in the discovery cohort

	CAD-ONLY n = 685	CAD-T2DM n = 254	P value
Demographic data			
Age (years)	62.60 (56.00,70.20)	64.40 (57.55,71.85)	0.030
Sex (male), %	83.50	72.05	< 0.001
BMI (kg/m <sup>2</sup> )	24.00 (22.00,26.00)	24.50 (22.00,27.00)	0.011
Comorbidities			
Hypertension, %	57.08	65.75	0.016
Hyperlipidemia, %	10.51	12.60	0.365
Biochemical measurements			
ALT, U/L	24.00 (18,34)	24.00 (18,34)	0.633
AST, U/L	24.00 (20,30)	23.00 (19,30)	0.065
eGFR, ml/min/1.73 m <sup>2</sup>	88.81 (74.24,104.60)	87.53 (69.28,103.51)	0.475
CK, U/L	88.00 (64.00,129.00)	85.00 (57.50,110.00)	0.028
CKMB, U/L	6.25 (4.50,8.73)	7.00 (5.00,9.20)	0.038
CREA, μmol/L	83 (71,96)	82 (69,97)	0.611
CHOL, mmol/L	4.07 (3.49,4.85)	4.18 (3.56,4.91)	0.243
LDLC, mmol/L	2.44 (1.88,3.16)	2.46 (1.91,3.03)	0.777
HDLc, mmol/L	0.92 (0.79,1.10)	0.955 (0.79,1.09)	0.799
APOA, g/L	1.00 (0.86,1.17)	1.01 (0.86,1.20)	0.579
TG, mmol/L	1.30 (0.97,1.78)	1.48 (1.11,2.11)	< 0.001
Lp(a), mg/L	179.40 (86.74,428.64)	178.09 (79.86,361.14)	0.333
proBNP, pg/mL	197.10 (63.19,672.28)	254.70 (85.75,878.00)	0.082
GLUC, mmol/L	5.45 (4.90,6.22)	7.90 (6.33,10.84)	< 0.001
HbA1c, %	6.00 (5.70,6.30)	7.40 (6.60,8.55)	< 0.001

Data are shown as median (IQR) or n (%). Chi-square ( $\chi^2$ ) tests were used for categorical variables. T-tests or Mann–Whitney U tests were used for normally and non-normally distributed variables, respectively. *BMI* body mass index, *ALT* alanine aminotransferase, *AST* aspartate aminotransferase, *eGFR* estimated glomerular filtration rate, *CK* creatine kinase, *CKMB* MB isoenzyme of creatine kinase, *CREA* Creatinine, *CHOL* cholesterol, *LDLC* low-density lipoprotein cholesterol, *HDLc* high-density lipoprotein cholesterol, *APOA* apolipoprotein, *TG* triacylglycerol, *Lp(a)*: lipoprotein(a), *proBNP* N-terminal pro-B-type natriuretic peptide, *GLUC* fasting blood glucose, *HbA1c* hemoglobin A1c

multivariate logistic regression analyses, adjusted for variables such as age, ALT, AST, BMI, eGFR, gender, hypertension and hyperlipidemia, revealed that 65 metabolites, including pipercolinic acid (OR: 0.14 [0.08–0.22]), retained their statistical significance (FDR < 0.05) (Additional file 2: Table S6). These metabolites were predominantly implicated in the biosynthetic pathways of phenylalanine, tyrosine, and tryptophan, as well as glycerophospholipid metabolism (Additional file 1: Figure S4).

The independent validation cohort was utilized to verify the metabolites identified in our study (Additional file 2: Table S7). Then, a confluence of 23 candidate characteristic metabolites was discerned (Additional file 2: Table S8), intersecting between the 65 metabolites that retained statistical significance following multivariate adjustment and those substantiated by the validation cohort (P < 0.05). Upon the exclusion of 3-indolebutyric acid, which demonstrated an inverse Odds Ratio (OR) (OR in discovery cohort: 0.56 [0.41–0.75]; OR in

validation cohort: 2.74 [1.4–5.49]), a total of 22 characteristic metabolites (hereafter called CMs.) were retained as salient biomarkers intricately associated with T2DM risk (Additional file 2: Table S9). Among these, 16 metabolites were found to be negative correlated with T2DM susceptibility, while 6 metabolites were positively correlated. The distribution of the normalized concentrations of the CMs in the two groups is as follows (Additional file 1: Figures S5, S6). Notably, the monosaccharides D-glucopyranose (OR: 2.89 [2.02–4.22]) and D-arabinose (OR: 1.57 [1.22–2.03]) were implicated in augmenting the risk of T2DM in the discovery cohort. Hexadecanoic acid was identified as a significant risk factor for T2DM (OR: 2.06 [1.39–3.09]), second only to D-glucopyranose. Phenyllactate, a metabolic byproduct of phenylalanine catabolism, was conspicuously elevated in patients diagnosed with phenylketonuria. Ceramide (Cer) (d18:1/16:1) emerged as the sole lipid moiety positively correlated with T2DM risk (OR: 1.48 [1.16–1.90]). In our results, pipercolinic acid and L-pipercolinic acid (OR: 0.18 [0.11–0.27]) were

identified as the most important potential protective factors against T2DM. Moreover, our investigation unveiled an array of Phosphatidylcholines (PCs) and Cholesteryl Esters (CEs), suggesting their putative protective role in the pathogenesis of T2DM (Fig. 2).

### Correlation of characteristic metabolites with GLUC and HbA1c

Fasting blood glucose and glycated hemoglobin are established clinical biomarkers of T2DM that reflect fluctuations in blood glucose concentrations over time. To determine whether the CMs of T2DM are directly associated with changes in body glucose levels, we first examined the correlation of 22 metabolites with GLUC and HbA1c (Fig. 3, Additional file 2: Tables S10, S11).

For GLUC, both the discovery and validation cohorts showed that D-glucopyranose and D-arabinose were significantly positively correlated with GLUC, whereas pipercolinic acid, L-pipercolic acid, 5-oxoproline, TG(16:1/16:1/18:2), CE(18:3), and PC(18:0/20:2) were all significantly negatively correlated with GLUC. For HbA1c in both cohorts, D-glucopyranose and Cer(d18:1/16:1) were significantly positively correlated with HbA1c, whereas pipercolinic acid, 5-oxoproline, L-pipercolic acid, PC(16:1/18:2), PC(22:1/22:2), PC(18:0/20:2), PC(16:1/20:3), and CE(18:3) were all significantly negatively correlated with HbA1c. Although these metabolites were correlated with T2DM in logistic regression, they were not correlated with GLUC or HbA1c, suggesting their potential role as stage markers not directly related to blood glucose levels.

### Pipercolinic acid and L-pipercolic acid promote ATP production in hepatocytes

To further elucidate whether the CMs influence glucose consumption and utilization, we validated three metabolites that were significantly negatively correlated with diabetes. After treating HepaRG cell with different concentrations (50  $\mu$ M, 250  $\mu$ M, 500  $\mu$ M) of pipercolinic acid, 5-oxoproline, and L-pipercolic acid for 60 h, the 50  $\mu$ M concentration of pipercolinic acid and L-pipercolic acid (In the following, PA and L-PA refer specifically to the groups treated with 50  $\mu$ M pipercolinic acid and 50  $\mu$ M L-pipercolic acid, respectively.) significantly increased cellular glucose consumption (Fig. 4A). Both significantly increased ATP levels (Fig. 4B). BSA protein levels showed no difference, ruling out the influence of cell proliferation (Additional file 1: Figure S7). Changes in intracellular lactate levels indicated that pipercolinic acid and L-pipercolic acid convert excess glucose to ATP through oxidative phosphorylation (Fig. 4C).

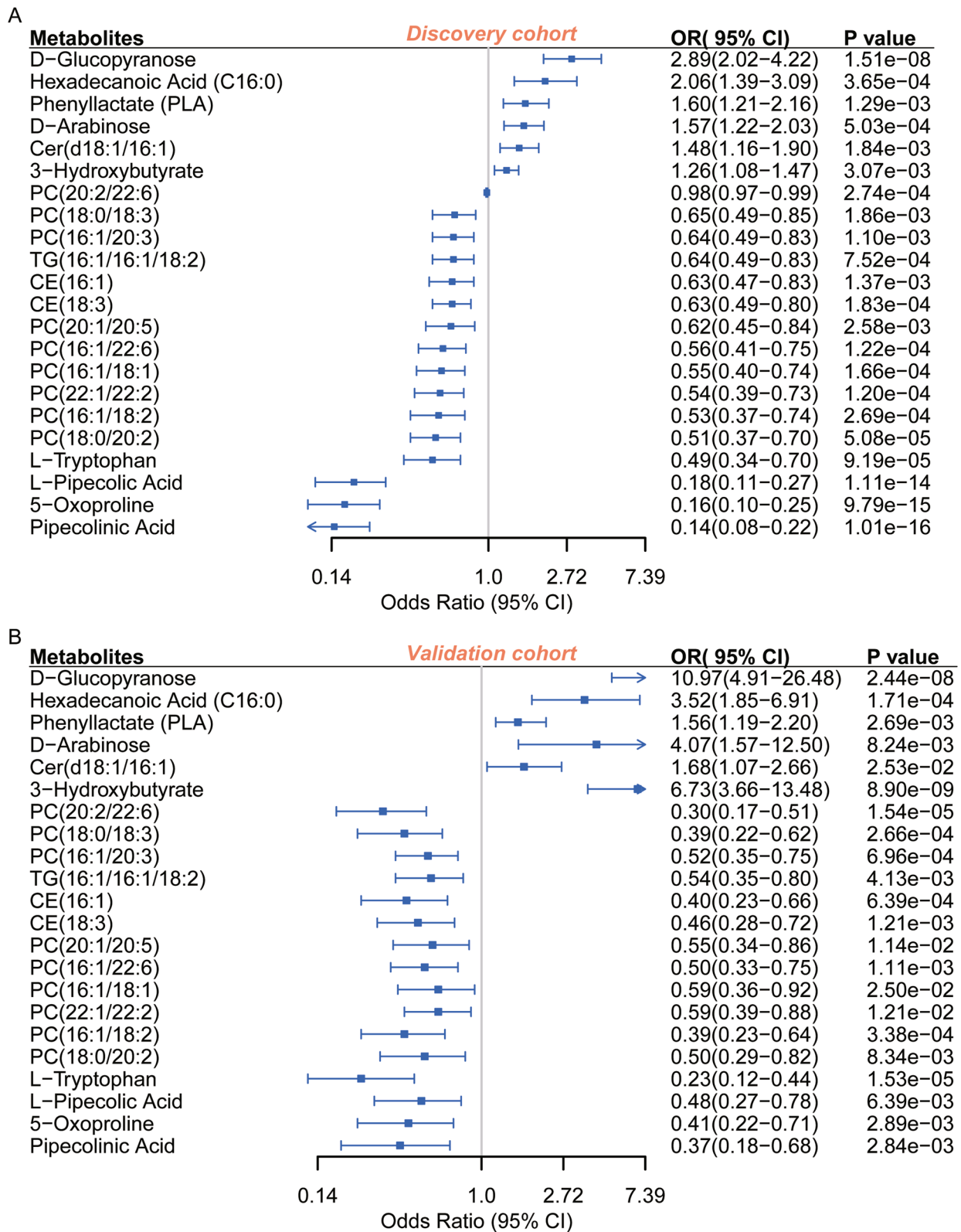
### RNA Sequencing identifies genes driven by different isomers of pipercolic acid

To uncover potential targets influenced by pipercolinic acid and L-pipercolic acid, we used high-throughput RNA Sequencing (RNA-Seq) to identify differences at the transcriptional level. In Principal Component 1 (PC1), the two drug treatment groups and the negative control (NC) group exhibited a clear separation, indicating significant differences in gene expression. The PA and L-PA showed some degree of separation in Principal Component 2 (PC2) (Fig. 4D). The results of differential gene expression analysis with  $P < 0.05$  after PA and L-PA treatment are listed in Additional file 2: Tables S12, S13, respectively.

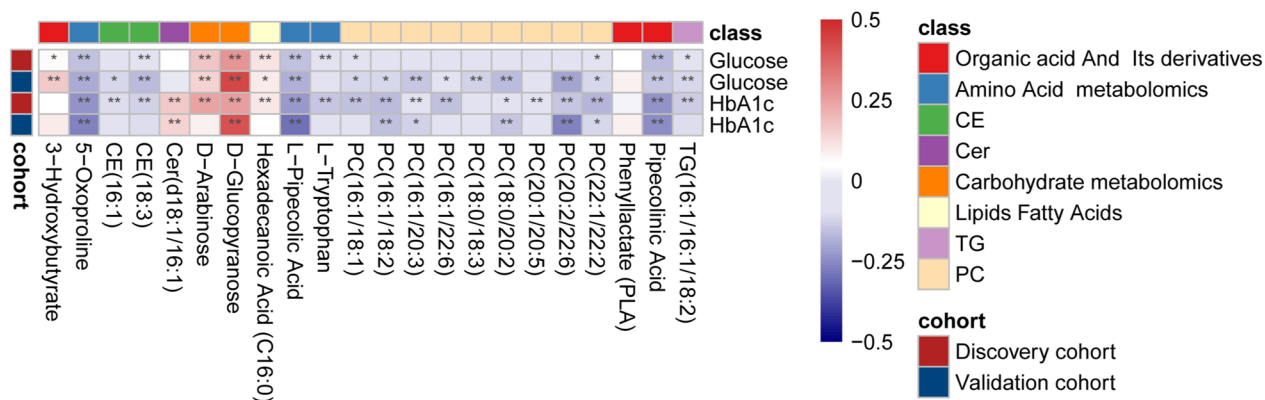
Using a threshold of  $P < 0.05$  and a fold change greater than 1.3, we identified 296 and 182 differentially expressed genes (DEGs) in the two comparison groups (NC vs. PA and NC vs. L-PA), respectively (Fig. 4D). 66 DEGs were common to both groups (Fig. 4E), and both sets of DEGs showed similar expression patterns (Additional file 1: Figure S8). Volcano plots for the DEGs of the two comparison groups are shown in Fig. 4F, G, with genes with FDR less than 0.05 labeled. The common differential gene *HID1* (Fig. 4F) was significantly downregulated in both groups. Previous RNA-Seq transcriptomic studies found that *HID1* mRNA levels were significantly correlated with obesity and glucose metabolic parameters in human subcutaneous and adipose tissues after Bonferroni correction for multiple testing [34]. After PA treatment, the transcript levels of the glycolysis inhibitor *TIGAR* [35] and the glucose neogenesis-related gene *CA5B* were downregulated, while the transcript levels of *LCN6*, *FAM156A* and *CITED4* were upregulated. In the comparison group (NC vs. L-PA), *ARID5B* was an upregulated differential gene, while transcript levels of *AKR1B1*, *SLC2A6*, and *GALNT4* all decreased.

To further comprehend the biological context of these differentially expressed genes, KEGG pathway enrichment analysis was executed. Treatment with PA and L-PA revealed distinct as well as overlapping pathway enrichment profiles, contributing to a deeper understanding of the molecular mechanisms underlying Pipercolic Acid's impact on gene expression and metabolic regulation. In the PA comparison group (NC vs. PA), 242 pathways were identified (Additional file 2: Table S14). Beyond the lipid and carbohydrate metabolism pathways (Fig. 4H), significant enrichment was observed in the IL-17 and NF-kappa B signaling pathways. In addition, the L-PA comparison group (NC vs. L-PA) exhibited enrichment in 206 pathways (Additional file 2: Table S15). These pathways showed notable associations with glucose and lipid metabolism-related pathways, including other types of O-glycan biosynthesis, mucin-type O-glycan biosynthesis, folate biosynthesis, citrate cycle (TCA cycle),





**Fig. 2** Forest plot of characteristic plasma metabolites associated with T2DM risk in CAD patients. (Discovery cohort FDR < 0.05, Validation cohort P < 0.05). Multivariate logistic regression in the discovery **A** and validation **B** cohorts, adjusted for age, ALT, AST, BMI, eGFR, hypertension, hyperlipidemia, and gender



**Fig. 3** Spearman correlation heatmap of characteristic metabolites with GLUC and HbA1c levels in CAD patients. The intensity of the color in the heatmap represents the strength of the correlation, with deeper colors indicating stronger associations. An asterisk (\*) indicates statistical significance at  $P < 0.05$

and regulation of lipolysis in adipocytes (Fig. 4I). Notably, no pathways related to glycation modifications were observed in the PA comparison group (NC vs. PA).

### Characteristic metabolites enhance predictive accuracy for clinical endpoints in cardiovascular diseases

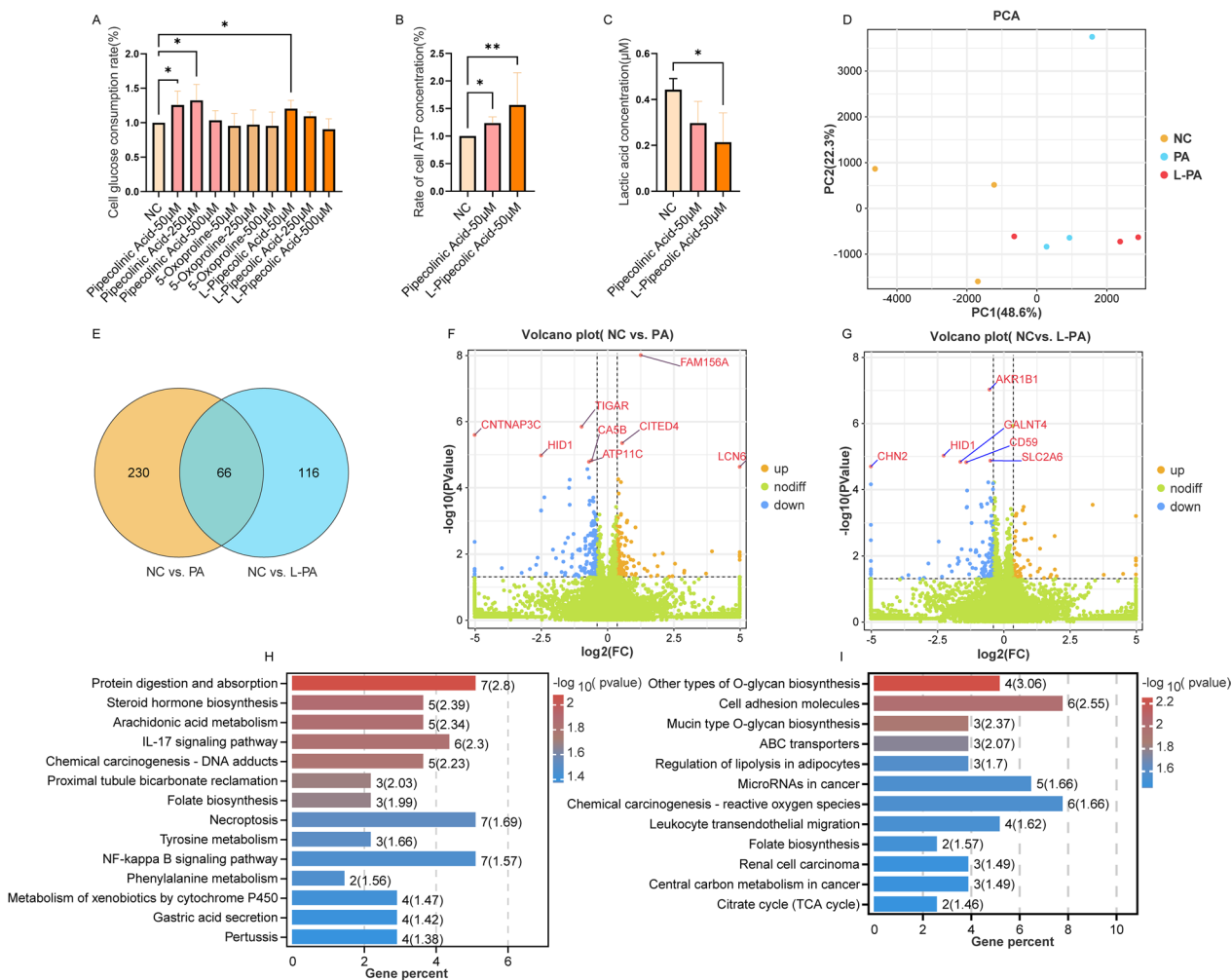
We focused on the association between the 22 identified CMs and CAD all-cause mortality. For this analysis, clinical endpoints were used as predictive indicators. Using LASSO, we screened traditional clinical variables and metabolomic variables separately (fivefold cross-validation, 200 replicates, frequency of occurrence > 100 times, Additional file 2: Table S16) to identify the most important features. We then compared the predictive power of the clinical model, the metabolic model, and the combined model. In the discovery cohort, the combined model (AUC=0.813, P value (vs. Clinical model): 0.001) had a higher predictive power than both the clinical model (AUC=0.761) and the metabolic model (AUC=0.720) (Fig. 5A). This result was also confirmed in the validation cohort (Fig. 5B). Differences between the models were determined using the Delong test (Additional file 2: Table S17).

### Discussion

In this study, we have investigated the metabolomic profile of plasma sample from CAD as well as patients with CAD and T2DM. Statistical analysis of the metabolomic data revealed 192 and 95 significantly altered metabolic signatures of CAD-T2DM patients in discovery (FDR < 0.05) and validation cohort ( $P < 0.05$ ), respectively. Further multivariate logistic regression analyses, adjusted for variables revealed that 65 metabolites retained their statistical significance in discovery cohort. After intersection of 65 metabolites in the validation cohort and

the discovery cohort, 23 candidate CMs were identified. Upon the exclusion of 3-indolebutyric acid, which demonstrated an inverse OR, a total of 22 CMs were retained as salient biomarkers intricately associated with T2DM risk. Among these, 16 metabolites exhibited negative correlation with T2DM susceptibility, while 6 metabolites were positively correlated. The positive metabolites including D-glucopyranose, hexadecanoic acid, PLA, D-arabinose, Cer(d18:1/16:1) and 3-hydroxybutyrate were implicated in augmenting the risk of T2DM in the two cohorts.

We observed the exclusive accumulation of D-glucopyranose and D-arabinose in CAD-T2DM patients. D-glucopyranose represents the natural form of glucose, while D-arabinose, a rare pentose, serves as a carbon source for bacteria [36]. Previous studies suggest a marked growth inhibitory effect of D-arabinose on the nematode *Caenorhabditis elegans*, possibly due to disruption of D-ribose and D-fructose metabolism [37]. Regarding PLA, it is significantly elevated in patients with phenylketonuria, and PLA levels show a positive correlation with T2DM in our study. However, in a cohort study of 134 individuals (healthy and T2DM), PLA levels decreased significantly in the T2DM group. This suggests that compared to individuals with T2DM alone (vs. healthy individuals) [38], our CAD-T2DM may have different PLA metabolite characteristics than CAD-ONLY. In addition, there is evidence that treatment activates PPAR- $\gamma$ 2, which promotes adipocyte differentiation and glucose uptake [39]. Further studies are needed to gain a deeper understanding of these differences. 3-Hydroxybutyrate, originating from the increased  $\beta$ -oxidation of free fatty acids [40], is found in higher concentrations in the plasma of T2DM patients [41] and has been shown to ameliorate insulin resistance in T2DM mice through the

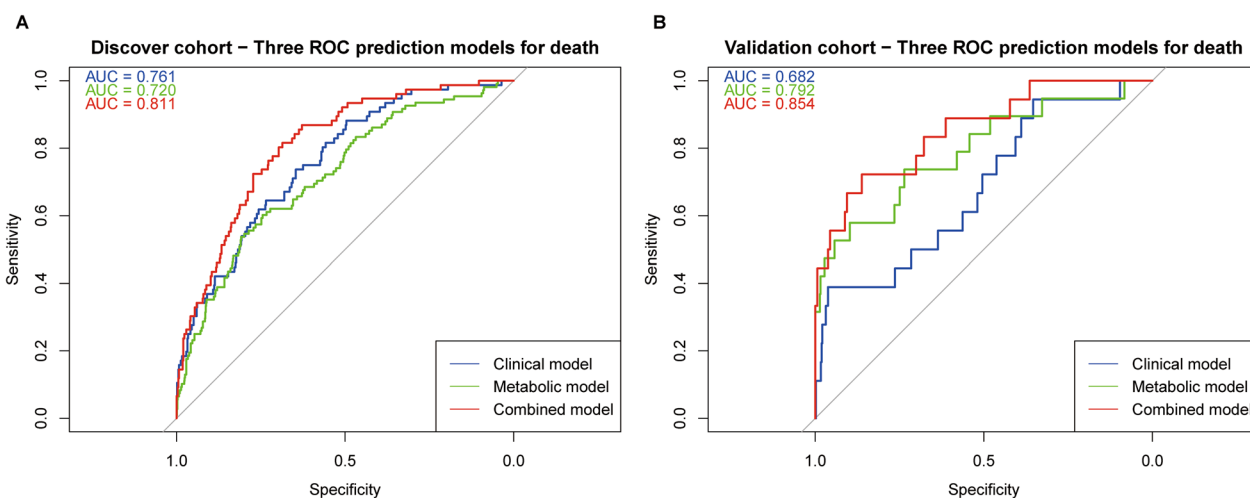


**Fig. 4** Effects of PA and L-PA on glucose metabolism and transcriptional changes in hepatocytes. **A** Glucose consumption in HepaRG cells influenced by metabolites at different concentrations over a period of 60 h. **B** Intracellular ATP levels and **C** intracellular lactate levels under the same conditions. Multiple group comparisons were performed using one-way analysis of variance (ANOVA). Significance is indicated by \* for  $P < 0.05$  and \*\* for  $P < 0.01$ . **D** Principal component analysis (PCA) visually depicts the differences between groups after PA and L-PA intervention. **E** Venn diagram shows the overlap of differentially expressed genes (DEGs) between the comparison groups (NC vs. PA) and (NC vs. L-PA). **F, G** Volcano plots show the DEGs for the comparison groups (NC vs. PA) and (NC vs. L-PA), respectively, with DEGs with  $FDR < 0.05$  labeled. **H, I** KEGG pathway enrichment analysis of DEGs from the comparison groups (NC vs. PA) and (NC vs. L-PA). The x-axis indicates the percentage of DEGs relative to the total genes in the pathway (expressed in %), with numbers in parentheses representing  $-\log_{10}(p\text{-value})$

HCAR2/ $Ca^{2+}$ /cAMP/PKA/Raf1/ERK1/2/PPAR $\gamma$  pathway [42].

Pipecolic acid, a metabolite of lysine found in human physiological fluids such as urine, plasma and CSE, is an important regulator of immunity in plants and humans alike [43]. Normal adults excrete pipecolic acid primarily as the D-enantiomer even though it is present in the blood stream mainly as the L-enantiomer. Despite some studies reporting a significant increase in D-pipecolic acid, L-pipecolic acid concentrations in the Human Diabetic Corneal Stroma compared to healthy

individuals [44], other research [45] revealed that 81 metabolites were significantly changed before T2DM onset, and pipecolic acid was reported to be negatively significantly associated with future T2DM risk. Previous studies have reported a reduction in pipecolic acid levels in individuals with T2DM [46]. Similarly, our study shows that pipecolic acid and L-pipecolic acid have a protective effect in CAD-T2DM. Moreover, our results innovatively revealed at the cellular and molecular level that pipecolic acid and L-pipecolic acid both convert excess glucose to ATP for energy through oxidative phosphorylation in hepatocytes.



**Fig. 5** Models for predicting mortality in CAD patients based on characteristic metabolites. **A** For the discovery cohort and **B** For the validation cohort. Clinical model: The clinical variables selected by LASSO regression are age, AST, CREA, proBNP, and gender. Metabolic model: The metabolites selected by LASSO regression are 3-hydroxybutyrate, CE(16:1), CE(18:3), Cer(d18:1/16:1), D-arabinose, D-glucopyranose, hexadecanoic acid(C16:0), L-pipecolic acid, PC(16:1/18:1), PC(16:1/20:3), PC(18:0/18:3), PC(18:0/20:2), PC(20:1/20:5), PC(22:1/22:2), phenyllactate, pipecolic acid, and TG(16:1/16:1/18:2). Combined Model: This model integrates both clinical and metabolite variables for prediction

Elevated levels of tryptophan are often accompanied by an increased risk of T2DM [47]. However, in patients with CAD, elevated Tryptophan levels are associated with a reduced mortality rate from coronary artery diseases [48]. Metabolites involved in glutathione metabolism, including 5-oxoproline, have been previously identified as markers for diabetes [49]. The first longitudinal analysis conducted in SOALS suggests that 5-oxoproline is associated with a 50% higher risk of T2DM [50]. The alteration in amino acid metabolism may impact the disease mechanism in CAD patients with T2DM [49]. This change implies that some biomarkers effective in healthy populations may not be suitable for patients with comorbid conditions.

Furthermore, to the best of our knowledge, this is the first metabolomic comparison including lipidomics between CAD-ONLY and CAD-T2DM. The presence of a large number of lipids among the CMs indicates a concomitant alteration in lipid metabolism. A portion of phosphatidylcholine (PC), cholesteryl ester (CE), and triglycerides (TG) are negatively correlated with T2DM, whereas hexadecanoic acid(C16:0) and Cer(d18:1/16:1) are positively correlated with T2DM. The elevated levels of hexadecanoic acid in individuals with diabetes may be explained by an increase in harmful complex lipid synthesis, impairment of cellular organelle function, and promotion of receptor-mediated inflammation [51–54]. In simple terms, intracellular levels of palmitic acid rise above the mitochondrial oxidation limit and are converted into harmful complex fatty acid-derived lipids such as diacylglycerol and Cer. PC, as one of the most

abundant phospholipids in mitochondria, plays a critical role in maintaining mitochondrial function [55]. Excessive lipid supply and reduced mitochondrial oxidative capacity can impair  $\beta$ -oxidation, leading to the accumulation of Cer [56].

In our investigation, comprehensive high-throughput RNA-Seq analysis identified 296 and 182 differentially expressed genes in the two comparison groups (NC vs. PA and NC vs. L-PA), respectively, with 66 DEGs being common to both groups and showing similar trends. For example, TIGAR, an intrinsic inhibitor of glycolysis [35, 57], and ARK1B1, a known activator of the polyol pathway often associated with long-term diabetic complications [58], as well as the glucose neogenesis-related gene CA5B [59, 60], were all significantly decreased after PA and L-PA intervention. Conversely, transcript levels of LCN6, which enhances skeletal muscle mitochondrial function [61, 62], increased significantly after intervention. KEGG enrichment analysis revealed a primary clustering of alterations in glucose and lipid metabolism and inflammatory pathways for both comparison groups (NC vs. PA and NC vs. L-PA). Folate biosynthesis emerged as a commonly enriched pathway, where chronic folate deficiency leads to disruptions in glucose and lipid metabolism and subsequent cognitive dysfunction in mice [63]. In addition, DEGs in L-PA affected the O-glycan biosynthesis and TCA cycle pathways, which may explain the increased ATP generation observed in the L-PA group compared to the PA group. In contrast, DEGs from the NC vs. PA group showed significant enrichment in the IL-17 and NF-kappa B pathways [64, 65]. In conclusion,

both pipercolinic acid and L-pipercolic acid are capable of modifying hepatic glucose utilization, with the transcriptomic data shedding light on the potential target sites and pathway differences for the metabolic actions of these two metabolites, including their impact on ATP production.

Given the enrichment of DEGs in NC vs. PA and NC vs. L-PA in diverse pathways, including those related to immune and signaling activation, and acknowledging the well-documented influence of diabetes on cardiovascular disease progression. There have been many studies on the relationship between diabetes and cardiovascular disease [66–68]. Globally, cardiovascular disease affects approximately 32.2% of all persons with T2DM, and the major cause of death for T2DM patients is cardiovascular disease [67, 68]. Diabetes mellitus is a risk factor for cardiovascular disease and has been associated with 3- to fourfold higher mortality [66]. We conducted LASSO analysis to value the prediction of metabolites on cardiovascular death. the combined model with metabolites is more comprehensive than clinical model and the metabolic model.

Our study has several limitations. First, because all the patients in this study came from a cohort of CAD patients without healthy individuals, the characteristics examined in this article are based only on CAD patients, and the difference in the metabolic spectrum between CAD patients and healthy individuals could not be determined. Secondly, the differential timeline for enrollment (2010–2014 in discovery vs. 2017–2018 for validation) and the low participant numbers in the Sun Yat-sen University and the Xiangya Hospital of Central South University may introduce confounding for center-specific features. Third, we only recorded the presence of T2DM at the time of enrollment and did not follow up the incidence of new-onset diabetes, thereby losing information on the time variable. Fourth, the lack of information on the use of hypoglycemic agents prevents correction for potential bias introduced by such medications. Fifth, more than 80% of the patients in this cohort are male. Although we corrected for this covariate during the analysis, our results may still have some gender bias. Sixth, the metabolomic data in this study were not collected at one time, and with the iterative updating of technology, the range of detectable metabolites is different. Therefore, we only included repeatedly appearing metabolites, which undoubtedly lost some metabolite information. In addition, multiple testing can lead to batch effects, which may still influence the results despite correction. Seventh, A triple quadrupole low resolution mass spectrometer with sMRM mode was used to establish the database information of the three samples of pipercolic acid (pipercolinic acid, D-pipercolic acid, and L-pipercolic acid). We

only verified the effect of pipercolinic acid and L-pipercolic acid on glucose metabolism, and the role of D-pipercolic acid was neglected in this study, so the role of D-pipercolic acid in metabolism needs to be further investigated. Finally, the direct addition of metabolites to cells without validation in a disease model may not fully elucidate the role of these metabolites in disease contexts. Future studies using a more comprehensive approach, including the use of disease models and a diverse patient cohort, are essential to validate and extend our findings.

## Conclusion

This study is based on a cohort of 1465 CAD patients from three centers. Plasma metabolic characteristics were differentiated between CAD patients without T2DM and those with concomitant T2DM through comprehensive, widely targeted metabolomics and lipidomics profiling analyses. We identified potential metabolic biomarkers and constructed a powerful model using characteristic metabolites to predict all-cause mortality endpoints. Moreover, we explored the potential functional pathways of characteristic metabolites through transcriptomic sequencing and cellular experiments. Intriguingly, pipercolic acid and L-pipercolic acid, the key metabolites impact on ATP production, may act as the intervention points to improve energy metabolism of CAD-T2DM patients. In conclusion, our study provides novel insights into the further mechanism research, prophylaxis and treatment of comorbidity of CAD and T2DM.

## Abbreviations

CAD	Coronary artery disease
T2DM	Type 2 diabetes mellitus
PCI	Percutaneous coronary intervention
GLUC	Fasting blood glucose
HbA1c	Hemoglobin A1c
CREA	Creatinine
CK	Creatine kinase
EDTA	Ethylenediaminetetraacetic acid
LC-ESI-MS/MS	Liquid chromatography electrospray ionization tandem mass spectrometric
UPLC-MS/MS	Ultraperformance liquid chromatography tandem mass spectrometry
MWDB	Metware database
MRM	Multiple reaction monitoring
QC	Quality control
QC-RLSC	Quality control-based robust LOESS signal correction
ALT	Alanine aminotransferase
AST	Aspartate aminotransferase
BMI	Body mass index
eGFR	Estimated glomerular filtration rate
FDR	False discovery rate
KEGG	Kyoto Encyclopedia of Genes and Genomes
CMs	Characteristic metabolites
LASSO	Least absolute shrinkage and selection operator
AUC	Area under the curve
ROC	Receiver operating characteristic
DMEM	Dulbecco's modified eagle medium
PBS	Phosphate-buffered saline
FBS	Fetal bovine serum

AG	Axygen biosciences
BSA	Bovine serum albumin
PBS	Phosphate-buffered saline
rRNA	Ribosome RNA
TPM	Transcripts per million
PCA	Principal component analysis
TG	Triglycerides
OR	Odds ratio
Cer	Ceramide
PC	Phosphatidylcholine
CE	Cholesteryl ester
PA	Groups treated with 50 $\mu$ M pipercolinic acid for 60 h
L-PA	Groups treated with 50 $\mu$ M L-pipercolic acid for 60 h
PC1	Principal component 1
RNA-Seq	RNA sequencing
NC	Negative control
PC2	Principal component 2
DEGs	Differentially expressed genes

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12933-023-02102-0>.

**Additional file 1: Figure S1.** Flowchart for patient enrolment. **Figure S2.** Mass spectra of quality control (QC) samples. **Figure S3.** Density plots of the data. **Figure S4.** Pathway enrichment analysis of metabolites associated with T2DM (FDR<0.05). **Figure S5.** Distribution of characterized metabolite levels in the discovery cohort. **Figure S6.** Distribution of characterized metabolite levels in the validation cohort. **Figure S7.** BSA content after 60 h metabolite intervention. **Figure S8.** Heatmap of differentially expressed genes (DEGs).

**Additional file 2: Table S1.** The ionization modes and ion pairs of the metabolites. **Table S2.** The ionization modes and ion pairs of the lipid species. **Table S3.** Clinical baseline data of patients in the validation cohort. **Table S4.** One factor logistic regression of diabetes in the discovery cohort. **Table S5.** One factor logistic regression of diabetes in the validation cohort. **Table S6.** Multivariate logistic regression of diabetes in the discovery cohort. **Table S7.** Multivariate logistic regression of diabetes in the validation cohort. **Table S8.** Multivariate logistic results of candidate characteristic metabolites in two cohort. **Table S9.** Multivariate logistic results of characteristic metabolites in two cohorts. **Table S10.** Relationship between plasma glucose and metabolites associated with diabetes. **Table S11.** Relationship between plasma HbA1c and metabolites associated with diabetes. **Table S12.** Difference in transcription levels after PA treatment (P<0.05). **Table S13.** Difference in transcription levels after L-PA treatment (P<0.05). **Table S14.** KEGG pathway analysis of DEGs in NC vs. PA. **Table S15.** KEGG pathway analysis of DEGs in NC vs. L-PA. **Table S16.** Feature inclusion frequency using LASSO based feature selection for death prediction. **Table S17.** Model performance measures (95% CIs) for death prediction.

**Additional file 3:** Transcriptome Data.

## Acknowledgements

We thank the Guangdong Provincial People's Hospital, the First Affiliated Hospital of Sun Yat-sen University and the Xiangya Hospital of Central South University for their support and cooperation.

## Author contributions

SLZ and YJL conceived the idea. YJL and MQ analyzed and interpreted the data. YJL and H.F.H. performed all cell biology experiments. YJL and JEL wrote the manuscript. HPL, CL, JXM and XPC collected and organized the data. WWL and SLZ reviewed the final manuscript. All authors participated in the experiments and studies. All authors have discussed the results and have contributed to the final draft of the manuscript.

## Funding

This study was supported by the National Natural Science Foundation of China (No. 82274016, 82003746, 81872934), the Science and Technology

Program of Guangzhou (No. 2023B03J1251, 2023A04J0542), Guangdong Basic and Applied Basic Research Foundation (No. 2023A1515011133), the Key-Area Research and Development Program of Guangdong Province, China (No. 2019B020229003), and National key research and development program (No. 2017YFC0909301).

## Availability of data and materials

Patient-related data, including profiles, will not be made publicly available for privacy reasons. However, RNA sequencing expression data are available in the Additional file 3 for use by researchers who comply with ethical and legal requirements. For data access or inquiries, please contact the study data administrator or the relevant research institution.

## Declarations

### Ethics approval and consent to participate

The study adhered to the guidelines of the Helsinki Declaration and received ethics approval from the Medical Ethical Review Committee of Guangdong Provincial People's Hospital (GDREC2010137 and GDREC2017071H). All participants provided written informed consent.

### Consent for publication

Not applicable.

### Competing interests

Authors have nothings to be disclosed. The authors declare that they have no competing interests.

### Author details

<sup>1</sup>School of Medicine, South China University of Technology, Guangzhou 510006, Guangdong, China. <sup>2</sup>Department of Pharmacy, Guangdong Provincial People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University, 106 Zhongshan 2nd Road, Guangzhou 510080, China. <sup>3</sup>Guangdong Provincial Key Laboratory of Coronary Heart Disease Prevention, Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou 510080, Guangdong, China. <sup>4</sup>Department of Cardiology, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China. <sup>5</sup>Department of Clinical Pharmacology, Xiangya Hospital, Central South University, Changsha, China.

Received: 9 November 2023 Accepted: 25 December 2023

Published online: 06 January 2024

## References

- Wu T, Qiao S, Shi C, Wang S, Ji G. Metabolomics window into diabetic complications. *J Diabetes Invest*. 2018;9:244–55.
- Lima EG, Hueb W, Garcia RMR, Pereira AC, Soares PR, Favarato D, et al. Impact of diabetes on 10-year outcomes of patients with multivessel coronary artery disease in the medicine, angioplasty, or surgery study II (MASS II) trial. *Am Heart J*. 2013;166:250–7.
- Shen M, Xie Q, Zhang R, Yu C, Xiao P. Metabolite-assisted models improve risk prediction of coronary heart disease in patients with diabetes. *Front Pharmacol*. 2023. <https://doi.org/10.3389/fphar.2023.1175021>.
- Arnold SV, Bhatt DL, Barsness GW, Beatty AL, Deedwania PC, Inzucchi SE, et al. Clinical management of stable coronary artery disease in patients with type 2 diabetes mellitus: a scientific statement from the American Heart Association. *Circulation*. 2020;141:e779–806.
- Goodarzi MO, Rotter JI. Genetics insights in the relationship between type 2 diabetes and coronary heart disease. *Circ Res*. 2020;126:1526–48.
- Metabolic Syndrome as a Precursor of Cardiovascular Disease and Type 2 Diabetes Mellitus | *Circulation*. <https://www.ahajournals.org/doi/https://doi.org/10.1161/circulationaha.105.539528>
- PGC-1 $\alpha$ , glucose metabolism and type 2 diabetes mellitus—Abstract—Europe PMC. <https://europepmc.org/article/med/27094040>
- Xiao H, Ma Y, Zhou Z, Li X, Ding K, Wu Y, et al. Disease patterns of coronary heart disease and type 2 diabetes harbored distinct and shared genetic architecture. *Cardiovasc Diabetol*. 2022;21:276.

9. Wong ND, Sattar N. Cardiovascular risk in diabetes mellitus: epidemiology, assessment and prevention. *Nat Rev Cardiol*. 2023;20:685–95.
10. Holmes E, Wilson ID, Nicholson JK. Metabolic phenotyping in health and disease. *Cell*. 2008;134:714–7.
11. Wishart DS, Guo A, Oler E, Wang F, Anjum A, Peters H, et al. HMDB 5.0: the human metabolome database for 2022. *Nucleic Acids Res*. 2021;50:D622–31.
12. Liu X, Locasale JW. Metabolomics: a primer. *Trends Biochem Sci*. 2017;42:274–84.
13. Metabolomics: beyond biomarkers and towards mechanisms *Nature reviews molecular cell Biology*. <https://www.nature.com/articles/nrm.2016.25>
14. Meikle PJ, Christopher MJ. Lipidomics is providing new insight into the metabolic syndrome and its sequelae. *Curr Opin Lipidol*. 2011;22:210.
15. Fan Y, Li Y, Chen Y, Zhao Y-J, Liu L-W, Li J, et al. Comprehensive metabolomic characterization of coronary artery diseases. *J Am Coll Cardiol*. 2016;68:1281–93.
16. Roberts LD, Koulman A, Griffin JL. Towards metabolic biomarkers of insulin resistance and type 2 diabetes: progress from the metabolome. *Lancet Diabetes Endocrinol*. 2014;2:65–75.
17. Urpi-Sarda M, Almanza-Aguilera E, Tulipani S, Tinahones FJ, Salas-Salvadó J, Andres-Lacueva C. Metabolomics for biomarkers of type 2 diabetes mellitus: advances and nutritional intervention trends. *Curr Cardiovasc Risk Rep*. 2015;9:12.
18. Würtz P, Soininen P, Kangas AJ, Rönnemaa T, Lehtimäki T, Kähönen M, et al. Branched-chain and aromatic amino acids are predictors of insulin resistance in young adults. *Diabetes Care*. 2013;36:648–55.
19. Branched-chain and aromatic amino acids are predictors of insulin resistance in young adults – PubMed. <https://pubmed.ncbi.nlm.nih.gov/23129134/>
20. Yun H, Sun L, Wu Q, Zong G, Qi Q, Li H, et al. Associations among circulating sphingolipids,  $\beta$ -cell function, and risk of developing type 2 diabetes: a population-based cohort study in China. *PLoS Med*. 2020;17: e1003451.
21. Yousri NA, Suhre K, Yassin E, Al-Shakaki A, Robay A, Elshafei M, et al. Metabolic and metabo-clinical signatures of type 2 diabetes, obesity, retinopathy, and dyslipidemia. *Diabetes*. 2022;71:184–205.
22. Chen Z-Z, Gerszten RE. Metabolomics and proteomics in type 2 diabetes. *Circ Res*. 2020;126:1613–27.
23. Liu W, Guo P, Dai T, Shi X, Shen G, Feng J. Metabolic interactions and differences between coronary heart disease and diabetes mellitus: a pilot study on biomarker determination and pathogenesis. *J Proteome Res*. 2021;20:2364–73.
24. Inubushi T, Kamemura N, Oda M, Sakurai J, Nakaya Y, Harada N, et al. L-tryptophan suppresses rise in blood glucose and preserves insulin secretion in type-2 diabetes mellitus rats. *J Nutr Sci Vitaminol*. 2012;58:415–22.
25. Keskin E, Yoldas IH. Fructose consumption correlates with triglyceride-glucose index and glycemic status in healthy adults. *Clin Nutr ESPEN*. 2022;52:184–9.
26. Zhou Q, Sun W-W, Chen J-C, Zhang H-L, Liu J, Lin Y, et al. Phenylalanine impairs insulin signaling and inhibits glucose uptake through modification of IR $\beta$ . *Nat Commun*. 2022;13:4291.
27. Cardiac Energy Metabolism in Heart Failure Circulation Research. <https://doi.org/10.1161/CIRCRESAHA.121.318241>
28. Amaral N, Okonko DO. Metabolic abnormalities of the heart in type II diabetes. *Diab Vasc Dis Res*. 2015;12:239–48.
29. Chen H, Wang Z, Qin M, Zhang B, Lin L, Ma Q, et al. Comprehensive metabolomics identified the prominent role of glycerophospholipid metabolism in coronary artery disease progression. *Front Mol Biosci*. 2021. <https://doi.org/10.3389/fmolb.2021.632950>.
30. American Diabetes Association. 2 classification and diagnosis of diabetes: standards of medical care in diabetes—2021. *Diabetes Care*. 2020;44:15–33.
31. Luan H, Ji F, Chen Y, Cai Z. statTarget: a streamlined tool for signal drift correction and interpretations of quantitative mass spectrometry-based omics data. *Anal Chim Acta*. 2018;1036:66–72.
32. van den Berg RA, Hoefsloot HJ, Westerhuis JA, Smilde AK, van der Werf MJ. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*. 2006;7:142.
33. Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*. 1996. <http://www.socolar.com/Article/Index?aid=100066876931&jid=10000059916>
34. Weingarten A, Turchetti L, Krohn K, Klötting I, Kern M, Kovacs P, et al. Novel genes on rat chromosome 10 are linked to body fat mass, preadipocyte number and adipocyte size. *Int J Obes*. 2016;40:1832–40.
35. Sirtuin 3 Alleviates Diabetic Cardiomyopathy by Regulating TIGAR and Cardiomyocyte Metabolism *Journal of the American Heart Association*. <https://doi.org/10.1161/JAHA.120.018913>
36. Watanabe S, Fukumori F, Nishiwaki H, Sakurai Y, Tajima K, Watanabe Y. Novel non-phosphorylative pathway of pentose metabolism from bacteria. *Sci Rep*. 2019;9:155.
37. Sakoguchi H, Yoshihara A, Shintani T, Okuma K, Izumori K, Sato M. Growth inhibitory effect of D-arabinose against the nematode *Caenorhabditis elegans*: discovery of a novel bioactive monosaccharide. *Bioorg Med Chem Lett*. 2016;26:726–9.
38. Zhao Q, Zhang A, Zong W, An N, Zhang H, Luan Y, et al. Exploring potential biomarkers and determining the metabolic mechanism of type 2 diabetes mellitus using liquid chromatography coupled to high-resolution mass spectrometry. *RSC Adv*. 2017;7:44186–98.
39. Ilavienil S, Kim DH, Valan Arasu M, Srigopalram S, Sivanesan R, Choi KC. Phenyllactic acid from *Lactobacillus plantarum* promotes adipogenic activity in 3T3-L1 adipocyte via up-regulation of PPAR- $\gamma$ . *Molecules*. 2015;20:15359–73.
40. Laffel L. Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes. *Diabetes Metab Res Rev*. 1999;15:412–26.
41. Vigili de Kreutzenberg S, Avogaro A. The role of point-of-care 3-hydroxybutyrate testing in patients with type 2 diabetes undergoing coronary angiography. *J Endocrinol Invest*. 2017;40:627–34.
42. Zhang Y, Li Z, Liu X, Chen X, Zhang S, Chen Y, et al. 3-Hydroxybutyrate ameliorates insulin resistance by inhibiting PPAR $\gamma$  Ser273 phosphorylation in type 2 diabetic mice. *Signal Transduct Target Ther*. 2023;8:190.
43. He M. Pipecolic acid in microbes: biosynthetic routes and enzymes. *J Ind Microbiol Biotechnol*. 2006;33:401–7.
44. Priyadarsini S, McKay TB, Sarker-Nag A, Allegood J, Chalfant C, Ma J-X, et al. Complete metabolome and lipidome analysis reveals novel biomarkers in the human diabetic corneal stroma. *Exp Eye Res*. 2016;153:90–100.
45. Ouyang Y, Qiu G, Zhao X, Su B, Feng D, Lv W, et al. Metabolome-Genome-wide association study (mGWAS) reveals novel metabolites associated with future type 2 diabetes risk and susceptibility loci in a case-control study in a Chinese prospective cohort. *Global Chall*. 2021;5:2000088.
46. Gu X, Al Dubayee M, Alshahrani A, Masood A, Benabdelkamel H, Zahra M, et al. Distinctive metabolomics patterns associated with insulin resistance and type 2 diabetes mellitus. *Front Mol Biosci*. 2020;7: 609806.
47. Chen T, Zheng X, Ma X, Bao Y, Ni Y, Hu C, et al. Tryptophan predicts the risk for future type 2 diabetes. *PLoS ONE*. 2016;11: e0162192.
48. Li Q, You Y, Zeng Y, Wang X, Pan Z, Pang J, et al. Associations between plasma tryptophan and indole-3-propionic acid levels and mortality in patients with coronary artery disease. *Am J Clin Nutr*. 2022;116:1070–7.
49. Zhang L, Zhang Y, Ma Z, Zhu Y, Chen Z. Altered amino acid metabolism between coronary heart disease patients with and without type 2 diabetes by quantitative  $^1\text{H}$  NMR based metabolomics. *J Pharm Biomed Anal*. 2021;206: 114381.
50. Rivas-Tumanyan S, Pacheco LS, Haslam DE, Liang L, Tucker KL, Joshipura KJ, et al. Novel plasma metabolomic markers associated with diabetes progression in older puerto ricans. *Metabolites*. 2022;12:513.
51. Shetty SS, Kumari S. Fatty acids and their role in type-2 diabetes (Review). *Exp Ther Med*. 2021;22:1–6.
52. Wu H-T, Chen W, Cheng K-C, Ku P-M, Yeh C-H, Cheng J-T. Oleic acid activates peroxisome proliferator-activated receptor  $\delta$  to compensate insulin resistance in steatotic cells. *J Nutr Biochem*. 2012;23:1264–70.
53. Palomer X, Pizarro-Delgado J, Barroso E, Vázquez-Carrera M. Palmitic and oleic acid: The Yin and Yang of fatty acids in type 2 diabetes mellitus. *Trends Endocrinol Metab*. 2018;29:178–90.
54. Artemis P. Simopoulos. 2023. The Impact of the Bellagio Report on Healthy Agriculture Healthy Nutrition, Healthy People: Scientific and Policy Aspects and the International Network of Centers for Genetics Nutrition and Fitness for Health. *Lifestyle Genomics*. <https://doi.org/10.1159/000375495>

55. van der Veen JN, Kennelly JP, Wan S, Vance JE, Vance DE, Jacobs RL. The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochimica et Biophysica Acta Biomembranes*. 2017;1859:1558–72.
56. Morze J, Wittenbecher C, Schwingshackl L, Danielewicz A, Rynkiewicz A, Hu FB, et al. Metabolomics and type 2 diabetes risk: an updated systematic review and meta-analysis of prospective cohort studies. *Diabetes Care*. 2022;45:1013–24.
57. Zhou W, Yao Y, Li J, Wu D, Zhao M, Yan Z, et al. TIGAR attenuates high glucose-induced neuronal apoptosis via an autophagy pathway. *Front Mol Neurosci*. 2019;12:193.
58. Penning TM. The aldo-keto reductases (AKRs): overview. *Chem Biol Interact*. 2015;234:236–46.
59. Kucharczyk P, Albano G, Deisl C, Ho TM, Bargagli M, Anderegg M, et al. Thiazides attenuate insulin secretion through inhibition of mitochondrial carbonic anhydrase 5b in  $\beta$ -islet cells in mice. *J Am Soc Nephrol*. 2023;34:1179–90.
60. Shah GN, Rubbelke TS, Hendin J, Nguyen H, Waheed A, Shoemaker JD, et al. Targeted mutagenesis of mitochondrial carbonic anhydrases VA and VB implicates both enzymes in ammonia detoxification and glucose metabolism. *Proc Natl Acad Sci USA*. 2013;110:7423–8.
61. Charkoftaki G, Wang Y, McAndrews M, Bruford EA, Thompson DC, Vasiliou V, et al. Update on the human and mouse lipocalin (LCN) gene family, including evidence the mouse Mup cluster is result of an “evolutionary bloom.” *Hum Genomics*. 2019;13:11.
62. Tian Y, Yang J, Lan M, Zou T. Construction and analysis of a joint diagnosis model of random forest and artificial neural network for heart failure. *Aging (Albany NY)*. 2020;12:26221–35.
63. Zhao M, Yuan MM, Yuan L, Huang LL, Liao JH, Yu XL, et al. Chronic folate deficiency induces glucose and lipid metabolism disorders and subsequent cognitive dysfunction in mice. *PLoS ONE*. 2018;13: e0202910.
64. Abdel-Moneim A, Bakery HH, Allam G. The potential pathogenic role of IL-17/Th17 cells in both type 1 and type 2 diabetes mellitus. *Biomed Pharmacother*. 2018;101:287–92.
65. Tantiwong P, Shanmugasundaram K, Monroy A, Ghosh S, Li M, DeFronzo RA, et al. NF- $\kappa$ B activity in muscle from obese and type 2 diabetic subjects under basal and exercise-stimulated conditions. *Am J Physiol Endocrinol Metab*. 2010;299:E794–801.
66. Raghavan S, Vassy JL, Ho Y, Song RJ, Gagnon DR, Cho K, et al. Diabetes mellitus-related all-cause and cardiovascular mortality in a national cohort of adults. *J Am Heart Assoc*. 2019;8: e011295.
67. Ma C-X, Ma X-N, Guan C-H, Li Y-D, Mauricio D, Fu S-B. Cardiovascular disease in type 2 diabetes mellitus: progress toward personalized management. *Cardiovasc Diabetol*. 2022;21:74.
68. Einarson TR, Acs A, Ludwig C, Panton UH. Prevalence of cardiovascular disease in type 2 diabetes: a systematic literature review of scientific evidence from across the world in 2007–2017. *Cardiovasc Diabetol*. 2018;17:83.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

