

RESEARCH

Open Access



Identification of biomarkers associated with metabolic cardiovascular disease using mRNA-SNP-miRNA regulatory network analysis

Zhiyuan Fan¹, Wenjuan Peng¹, Zhiwen Wang², Ling Zhang¹ and Kuo Liu^{1*}

Abstract

Background: CVD is the leading cause of death in T2DM patients. However, few biomarkers have been identified to detect and diagnose CVD in the early stage of T2DM. The aim of our study was to identify the important mRNAs, micro (mi)RNAs and SNPs (single nucleotide polymorphisms) that are associated with metabolic cardiovascular disease.

Materials and methods: Expression profiles and GWAS data were obtained from Gene Expression Omnibus (GEO) database. MiRNA-sequencing was conducted by Illumina HiSeq 2000 platform in T2DM patients and T2DM with CVD patients. EQTL analysis and gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted. mRNA-miRNA co-expression network and mRNA-SNP-miRNA interaction network were established and visualized by Cytoscape 3.7.2.

Results: In our study, we identified 56 genes and 16 miRNAs that were significantly differentially expressed. KEGG analyses results indicated that B cell receptor signaling pathway and hematopoietic cell lineage were included in the biological functions of differentially expressed genes. mRNA-miRNA co-expression network and mRNA-SNP-miRNA interaction network illustrated that *let-7i-5p*, *RASGRP3*, *KRT1* and *CEP41* may be potential biomarkers for the early detection and diagnosis of CVD in T2DM patients.

Conclusion: Our results suggested that downregulated *let-7i-5p*, and upregulated *RASGRP3*, *KRT1* and *CEP41* may play crucial roles in molecular mechanisms underlying the initiation and development of CVD in T2DM patients.

Keywords: CVD, T2DM, mRNA, SNP, miRNA, Interaction network

Background

It is estimated that in 2019 463 million adults aged 20–79 years would develop diabetes with type 2 diabetes (T2DM) accounting for >90% of cases [1]. The incidence of T2DM has greatly increased in recent years and has become a great threat to human health worldwide. T2DM is one of major risk factors of cardiovascular

diseases (CVD), but its mechanism of action is not fully understood [2]. Even newly diagnosed diabetics were reported to have at least one vascular complication, and CVD is the leading cause of death in T2DM patients [3, 4]. Furthermore, people with T2DM are 2–6 times more likely to die of CVD than those without diabetes [5]. However, few biomarkers have been identified to diagnose CVD at the early stage of T2DM.

MicroRNAs (miRNAs) bind to complementary sequences of their target mRNA by base pairing to induce mRNA degradation and/or inhibit translation, thus affecting gene expression after transcription [6].

*Correspondence: liukuo@cmmu.edu.cn

¹ Department of Epidemiology and Health Statistics, School of Public Health, Capital Medical University, and Beijing Municipal Key Laboratory of Clinical Epidemiology, Beijing, China

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



MiRNA has been found to have high specificity for disease status and may be used as a potential biomarker to predict disease progression. In recent years, multiple miRNAs have been demonstrated to be involved in angiogenesis and endothelial cells dysfunction [7–9]. Expression of miRNAs can be affected by many factors. The miRNA seed region can specifically bind to miRNA recognition element (MRE) in the 3'UTR (3' untranslated region) of their target mRNA [10]. Therefore, any anomalies of miRNA-MRE interaction may have an impact on gene expression and lead to diseases [11–13]. Single nucleotide polymorphism (SNP), the most common genetic variants, can interfere with the base pairing between miRNA and its target mRNA, affecting normal expression of genes and eventually contributing to disease pathogenesis [14]. Genome-Wide Association Studies (GWAS) have found that genetic variants play an important role in pathogenesis of CVD [15], and association between some genetic variants and CVD has critical biological significance for T2DM patients [16, 17]. Genome-wide expression quantitative trait locus (eQTL) analysis is an effective method to study the effect of SNPs on gene expression [18, 19]. However, eQTL analysis mainly focuses on mRNA expression and rarely involves miRNA, which leads to incomplete interaction patterns [20].

To more comprehensively reveal the complicated association between mRNA, miRNA and SNP, and to find potential biomarkers with high specificity and sensitivity to diagnose CVD at the early stage of T2DM, we performed an integrative analysis. eQTL analysis and gene ontology (GO), KEGG pathway enrichment analyses were also conducted to better understand the connection between mRNA, miRNA and SNP, and their potential effect on CVD in T2DM patients.

Materials and methods

Sample processing and miRNA profiling

Six diabetes patients and five diabetes with ischemic heart disease were recruited from communities in Beijing after informed consent was obtained from each participant. T2DM was diagnosed according to American

Diabetes Association Criteria [24357215]. Ischemic heart disease was defined by clinical history, including acute myocardium infarction, angina pectoris and/or ischemic electrocardiographic alterations. All the participants did not have diabetic retinopathy, nor diabetic microvascular complications. This study was approved by the Ethics Committee of Capital Medical University (No. 2016SY24).

Total RNA extraction was performed using TRIzol (Invitrogen, USA) according to manufacturer's instructions. After removing DNA contamination by DNase I treatment, total RNA was assessed by NanoDrop spectrophotometer (NanoDrop, USA). A total of 2 µg of RNA/sample was used for the miRNA library. TruSeq Small RNA Sample Preparation Kit (Illumina, Inc., San Diego, CA, USA) was used to generate miRNA sequencing libraries. The library concentration was assessed by Qubit Spectrophotometer and the miRNA sequencing library quality was obtained by using the Agilent 2100 Bioanalyzer system with a High Sensitivity DNA Kit (Agilent Technologies). Sequencing was performed on an Illumina HiSeq 2000 platform and 50 bp of single-end reads were generated. We used fastqc v0.10.1 (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) to check the quality of raw reads. The reads were mapped to hg19 reference genome to identify mature miRNA, and the expression profile was generated by using miRDeep2 software (<https://www.mdc-berlin.de/8551903/en/>).

Microarray datasets and preprocessing

We searched GEO data repository (<https://www.ncbi.nlm.nih.gov/geo/>) for eligible studies until January 20, 2020. The following terms and different combination of them were used: "type 2 diabetes", "cardiovascular disease", "atherosclerosis", "coronary artery disease", "heart disease", "angina pectoris", "myocardial infarction", "coronary angiography", "chronic myocardial ischemia syndrome", and "acute coronary syndrome". Based on the search strategy, three datasets were downloaded from GEO, among which GSE90074 and GSE90073 (see Additional file 1: Table S1 for more details) were from same study. Details of each dataset are provided in Table 1.

Table 1 Details of microarray datasets from GEO database

GSE	Type	Sample size		Chip
		T2DM	T2DM + CVD	
GSE90074	mRNA	17	38	Agilent-014850 whole human genome microarray 4 × 44 K G4112F
GSE66175	mRNA	48	57	Affymetrix human genome U133A 2.0 array
GSE90073	SNP	13	25	Affymetrix genome-wide human SNP 6.0 array

Quality evaluation and data preprocessing

The quality of the original CEL files from the dataset GSE66175 (see Additional file 1: Table S1 for more details) was evaluated in this study. We plotted the relative log expression boxplot and normalized unscaled standard errors boxplot to evaluate the consistency of chip quality using affyPLM package (<https://git.bioconductor.org/packages/affyPLM/>). The degradation of RNA would have a great impact on the quality of chips, thus RNA degradation plot was plotted to show the trend. Then RMA algorithm was used for background correction and normalization of the data. We could obtain the matrix file after quality control and preprocessing in dataset GSE90074, thus we did not do further quality control. Then we converted the probe names from two gene expression profiles into gene names using R 3.6.2 software. And K-Nearest Neighbor in impute package (<https://git.bioconductor.org/packages/impute/>) was used to fill in the missing value of the gene. Next, we merged two intact gene expression profiles by dplyr package (<https://CRAN.R-project.org/package=dplyr>). To eliminate difference between batches, we used the combat function in SVA package (<https://bioconductor.org/packages/sva/>) of R 3.6.2 software.

The quality of raw data for miRNA were quality checked by visualization of base quality distributions. To ensure the quality of data analysis, the raw data were filtered and clean data were obtained by using FASTX-Toolkit software (http://hannonlab.cshl.edu/fastx_toolkit/). The following sequences were removed: linker sequence; sequences without 3' adapter sequences and insert fragments; sequence with Q20 percentage below 60%; sequences outside the length range of 18–36 bp. After the quality control, BLAST alignment (<http://blast.ncbi.nlm.nih.gov/>) was performed between the miRNA data and the mature miRNA sequences of the corresponding species in the miRBase database (<http://www.mirbase.org/>), and comparison with the RFAM database (<http://rfam.xfam.org/>) and the reference genome were performed to conduct a preliminary evaluation of the sequencing results.

Differential expression analysis

We performed differential expression analysis to find mRNA and miRNA whose $|\text{Log two-fold change}| > 1.5$ and p value < 0.05 by comparing T2DM patients with T2DM patients complicated with CVD in the R computing environment using limma (<https://bioconductor.org/packages/limma/>) and DESeq2 packages (<https://bioconductor.org/packages/DESeq2/>). Next, we plotted volcano plot and heatmap to visualize differentially expressed genes and miRNA using pheatmap package (<https://CRAN.R-project.org/package=pheatmap>).

Construction of mRNA-miRNA co-expression network

We performed Pearson correlation analysis to test correlation of differentially expressed mRNA and miRNA. Then, mRNA-miRNA pairs (p value < 0.05 , absolute value of correlation coefficient > 0.5) were used to construct mRNA-miRNA co-expression network by using software Cytoscape 3.7.2.

GO and KEGG pathway enrichment analyses

GO and KEGG pathway enrichment analyses were performed to explore the biological process, cellular component, molecular function of differentially expressed genes and other pathways they might get involved in.

mRNA expression quantitative trait locus analysis

We performed the mRNA expression quantitative trait locus analysis to explore the association between SNP and mRNA expression level by using MatrixEQTL package (<https://CRAN.R-project.org/package=MatrixEQTL>). Gender was used as a covariate to adjust its effect on mRNA expression. There are two types of eQTL, cis-eQTL which means the difference of gene expression may be caused by the gene itself and trans-eQTL which means the difference of gene expression may be caused by other genes. Then, cis-eQTL and trans-eQTL (p value < 0.05) were used for the next analysis.

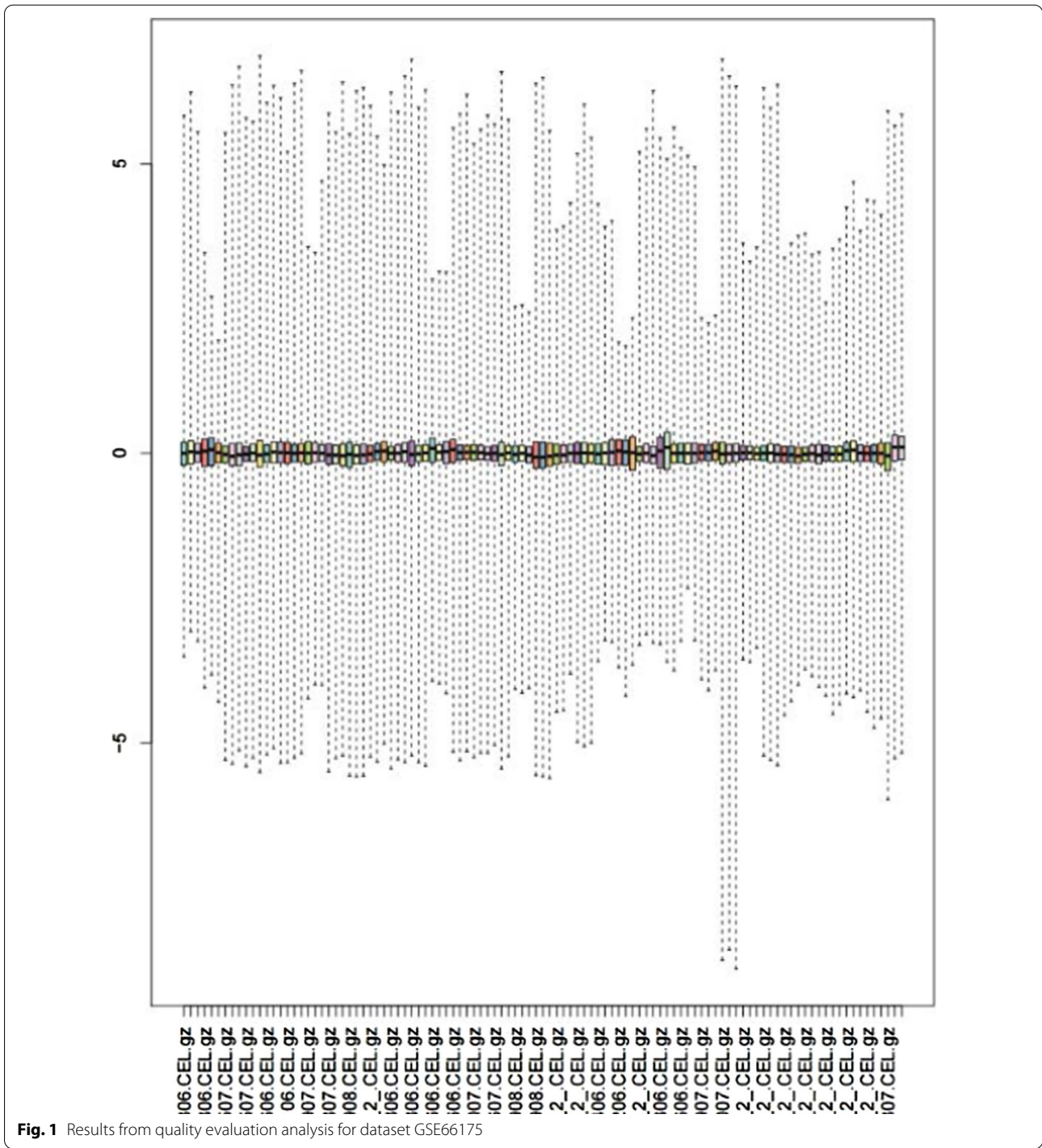
Construction of mRNA-SNP-miRNA interaction network

We downloaded SNP data related to differentially expressed miRNA from open sources MirSNP (<http://bioinfo.bjmu.edu.cn/mirsnp/search/>) and PolymiRTS 3.0 (<https://compbio.uthsc.edu/miRSNP/>), including SNPs in the binding regions of miRNA and its target mRNA. Next, the SNPs were matched with the cis- and trans-eQTLs of mRNA to find the overlapping SNPs, which could be used to connect mRNA and miRNA. TargetScan (http://www.targetscan.org/mamm_31/) and miRTarBase databases (<http://mirtarbase.cuhk.edu.cn/php/index.php>) were used to predict the target mRNA of miRNA. Then, we used the overlapping SNPs and the prediction target relationships between mRNAs and miRNAs to construct mRNA-SNP-miRNA interaction networks. Visualization of the mRNA-SNP-miRNA interaction network was performed by Software Cytoscape 3.7.2.

Results

Analysis of chip quality

To ensure the quality of chips included in the study, we evaluated the original CEL files from the dataset GSE66175 to eliminate the chips with poor quality. The relative log expression boxplot showed that the central values of all samples were close to the Y-axis "0" and basically consistent, indicating that all samples were of high



quality (Fig. 1). To further evaluate the chip quality, we plotted normalized unscaled standard errors boxplot, in which the standard errors of the samples were all close to the Y-axis "1" with small deviation, indicating that the chip quality was reliable and consistent with the results of the relative log expression boxplot (Additional file 1: Figure S1). The RNA degradation plot showed that the

lines of each chip are roughly parallel, indicating that the chip quality is high, which can be used for subsequent data analysis (Additional file 1: Figure S2). We downloaded matrix file of dataset GSE90074 after quality control and preprocessing and integrated it with the dataset GSE66175, then a total of 160 samples (including 65

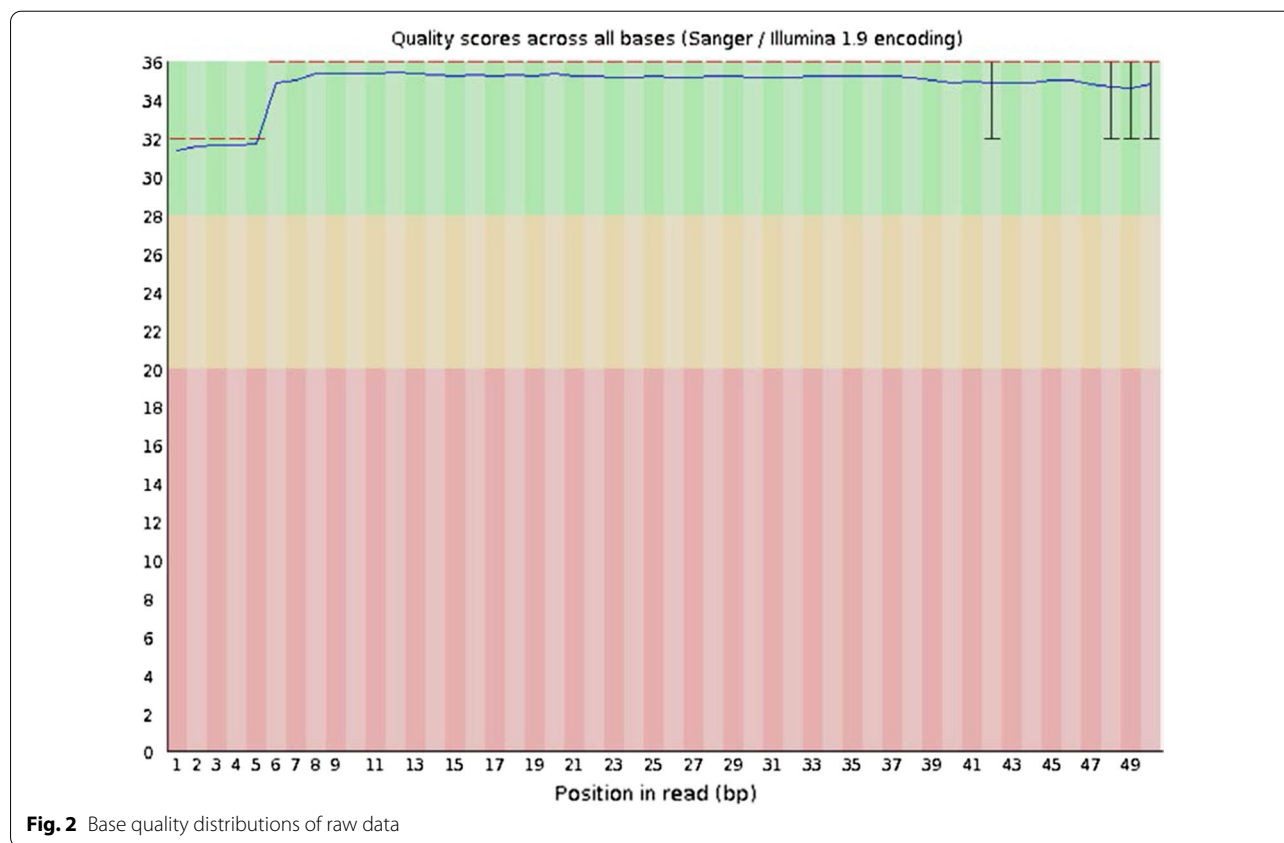


Table 2 Quality control results

Sample ID	N after filtration	N of high quality	N with 18–36 bp	N alignments to mature miRNAs in miRBase
DM-01	12,250,616	12,238,294	12,233,798	1530
DM-02	12,281,076	12,268,427	12,264,970	1499
DM-03	10,744,486	10,731,569	10,728,194	1468
DM-04	10,689,640	10,678,807	10,676,732	1440
DM-05	12,995,117	12,981,009	12,975,479	1588
DM-06	9,843,907	9,832,067	9,828,794	1498
DM + CVD-01	13,441,194	13,427,396	13,420,180	1584
DM + CVD-02	15,087,835	15,070,108	15,060,695	1572
DM + CVD-03	12,995,412	12,982,855	12,980,420	1518
DM + CVD-04	9,701,266	9,692,259	9,676,519	1521
DM + CVD-05	11,029,427	11,017,673	11,014,271	1427

N: number of sequence

T2DM patients and 95 T2DM patients with CVD) were included in this study.

To ensure the quality of data analysis, we visualized the base quality distributions to check the data quality (Fig. 2). In addition, we filtered the raw data and performed data quality control. The quality control results were listed in Table 2.

Differential expression analysis and visualization

Differentially expressed genes were screened by differential expression analysis using limma and DESeq2 packages. The results showed that a total of 56 genes were significantly differentially expressed, among which 47 were up-regulated and 9 were down-regulated (Additional file 1: Table S2). Also, miRNAs were selected by

differential expression analysis using DESeq2 packages. There were 16 miRNAs were significantly differentially expressed, among which 8 were up-regulated and 8 were down-regulated (Additional file 1: Table S3). Then, we plotted heatmap to visualize the results by using pheatmap package (Fig. 3a, b).

MRNA-miRNA co-expression network

Pearson correlation analysis was performed to explore correlation of differentially expressed mRNA and miRNA, and the co-expression network was visualized, in which 19 mRNA (including 15 upregulated mRNA and 4 downregulated mRNA) and 12 miRNAs (including 5 upregulated miRNAs and 7 downregulated miRNAs) were included (Fig. 4). The let-7i-5p and miR-320c were both correlated to 6 genes.

GO and KEGG pathway enrichment analysis

To explore the potential biological functions of differentially expressed genes, GO and KEGG pathway enrichment analyses were performed. The GO results showed that the differentially expressed gene were most likely associated with plasma membrane, extracellular space and receptor activity (Fig. 5). KEGG results revealed that the biological functions of differentially expressed genes included B cell receptor signaling pathway and hematopoietic cell lineage (Table 3).

MRNA-SNP-miRNA interaction network

We performed eQTL analysis to explore the association between SNP and mRNA expression level, then we found 60 mRNA and cis-eQTL pairs, and 44,454 mRNA and trans-eQTL pairs. We searched MirSNP and PolymiRTS 3.0 to obtain the SNPs related to differentially expressed miRNAs, eventually we screened 16,018 miRNA and SNP pairs including 16 differentially expressed miRNAs and 15,741 SNPs. Then SNPs related to differentially expressed miRNAs were matched with cis- and trans-eQTL, 16 mRNA-SNP-miRNA trios were obtained by overlapping SNPs. In addition, the effect and binding energy of 8 miRNA and SNP pairs in mRNA-SNP-miRNA trios were obtained (Additional file 1: Table S6). The results showed that rs17093783, rs2235364, rs1285935 and rs11562803 could enhance the binding region, rs2502607, rs325009 and rs11574860 could result in a decrease in the stability of the binding region, and rs17047863 could break the binding region. We also predicted the target mRNA of differentially expressed miRNA, and found that target mRNA of hsa-miR-581 included *CEP41* which was consistent with the mRNA-SNP-miRNA trios. We visualized the relationship between 16 mRNA-SNP-miRNA trios and prediction

of differentially expressed miRNA by using software Cytoscape 3.7.2 (Fig. 6).

Discussion

CVD is the most common cause of death among patients with T2DM, imposing a heavy burden on the economy of individuals and societies. Some studies have investigated the underlying mechanism of CVD and T2DM, and found that miRNA and SNP play important role in the occurrence and development. However, there is still a lack of biomarkers with good sensitivity and specificity for early detection and diagnosis of CVD in T2DM patients. Constructing the mRNA-SNP-miRNA interaction network can help to reveal the underlying correlation between genetic variants and diseases. But there is a lack of miRNA sequencing data of CVD in T2DM patients, leading to few evidence and certain limitations. We combined mRNA, miRNA and SNP data by searching GEO database and conducting miRNA sequencing. Our present study identified 16 mRNA-SNP-miRNA interaction trios, which revealed that SNPs in the binding region between miRNA and mRNA can interfere gene expression. Previous studies have proved that miRNAs play an essential role in the occurrence and development of various diseases [21], and our findings may increase understanding of the underlying mechanism.

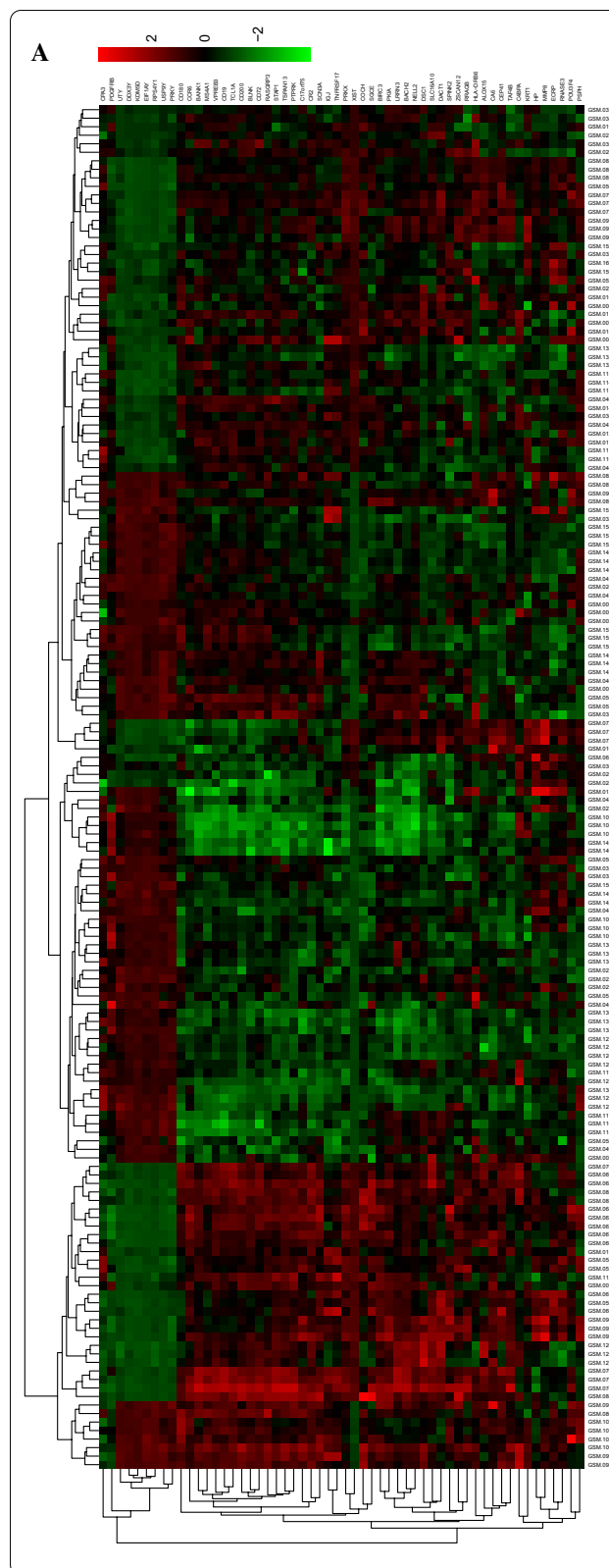
In this study, we identified 56 differentially expressed genes and 16 differentially expressed miRNAs between T2DM patients and T2DM patients with CVD. According to mRNA-miRNA co-expression network, let-7i-5p and miR-320c were both related to 6 genes. Previous study showed that let-7i-5p was involved in the regulation of cardiac cell cycle, and inhibition of let-7i-5p may be a potential strategy for cardiac repair after ischemic injury [22]. In addition, downregulated let-7i was observed in dilated cardiomyopathy and low expression of let-7i was associated with poor clinical outcomes of patients with dilated cardiomyopathy [23]. Downregulated let-7i-5p was observed in cardiomyocytes during hypoxia injury and let-7i-5p pathway was used to suppress hypoxia-induced apoptosis and mitochondrial energy metabolism dysfunction in vitro [24]. Taken together, let-7i-5p was demonstrated to be associated with the pathogenesis of CVD, and may play an important role in the pathophysiological process of CVD. Interestingly, the six genes associated with downregulated miR-320 were up-regulated which indicated that miR-320 may bind to complementary sequences of mRNA of these genes to induce mRNA degradation or inhibit translation, thus affecting gene expression.

In our present study, KEGG pathway enrichment analysis indicated significant enrichment in pathways including B cell receptor signaling pathway and hematopoietic

Fig. 3 Results from differential expression analysis. **a** The heatmap for differentially expressed genes. (The corresponding number of file names of Fig. 3A was listed in Additional file 1: Table S4) **b** The heatmap for differentially expressed miRNAs (The corresponding number of file names of Fig. 3B was listed in Additional file 1: Table S5). The red represented upregulated genes or miRNAs, the green represented downregulated genes or miRNAs, and the black represented the genes or miRNAs with no significant difference in expression level

cell lineage. Previous studies suggested that compared with healthy people, *CD19* was related to B cell receptor and significantly downregulated in patients with acute myocardial infarction [25]. However, this study showed that *CD19* associated with B cell receptor signaling pathway was upregulated in T2DM patients with CVD, which could attribute to the mechanism of CVD in T2DM patients. Moreover, mutations in the *TET2* that can promote clonal hematopoiesis were associated with an increased risk of atherosclerosis [26]. Interestingly, we found that *TET2* was target gene of 7 differentially expressed miRNAs (miR-133a-3p, miR-6817-3p, miR-873-5p, miR-581, miR-210-3p, miR-202-5p, miR-2355-3p, let-7i-5p, miR-196a-5p, miR-760) in our study which highlighted that the hematopoietic cell lineage and differentially expressed miRNAs may be involved in the progression of CVD in T2DM patients. In addition, KEGG pathway enrichment analysis showed PI3K-Akt signaling pathway was associated with *PDGFRB*, *CD19* and *TCL1A*. Some studies confirmed that PI3K-Akt signaling pathway plays an important role in the pathophysiology of vascular diseases [27]. Activated PI3K-Akt signaling pathway has been proven to improve insulin sensitivity, regulate glucose and lipid metabolism, and protect vascular endothelial cells [28]. In addition, activating IRS/PI3K/Akt pathway activity may play an anti-atherosclerotic role [29]. Interestingly, KEGG pathway enrichment analysis indicated that linoleic acid metabolism was related to *ALOX15*, and a growing number of studies reported that linoleic acid was associated with the prevention of T2DM and CVD [30, 31]. Therefore, we can reasonably speculate that PI3K-Akt signaling pathway and linoleic acid metabolism may play a role in the pathogenesis of T2DM complicated with CVD.

MRNA-SNP-miRNA interaction network included 16 SNPs, 9 mRNA and 11 miRNAs, but only miR-581, rs325009 and *CEP41* were related to each other. Previous studies suggested that *CEP41* was a new regulator for angiogenesis which promoted angiogenesis through the HIF1A-Aurka-VEGF pathway and was involved in endothelial cell migration which was involved in the



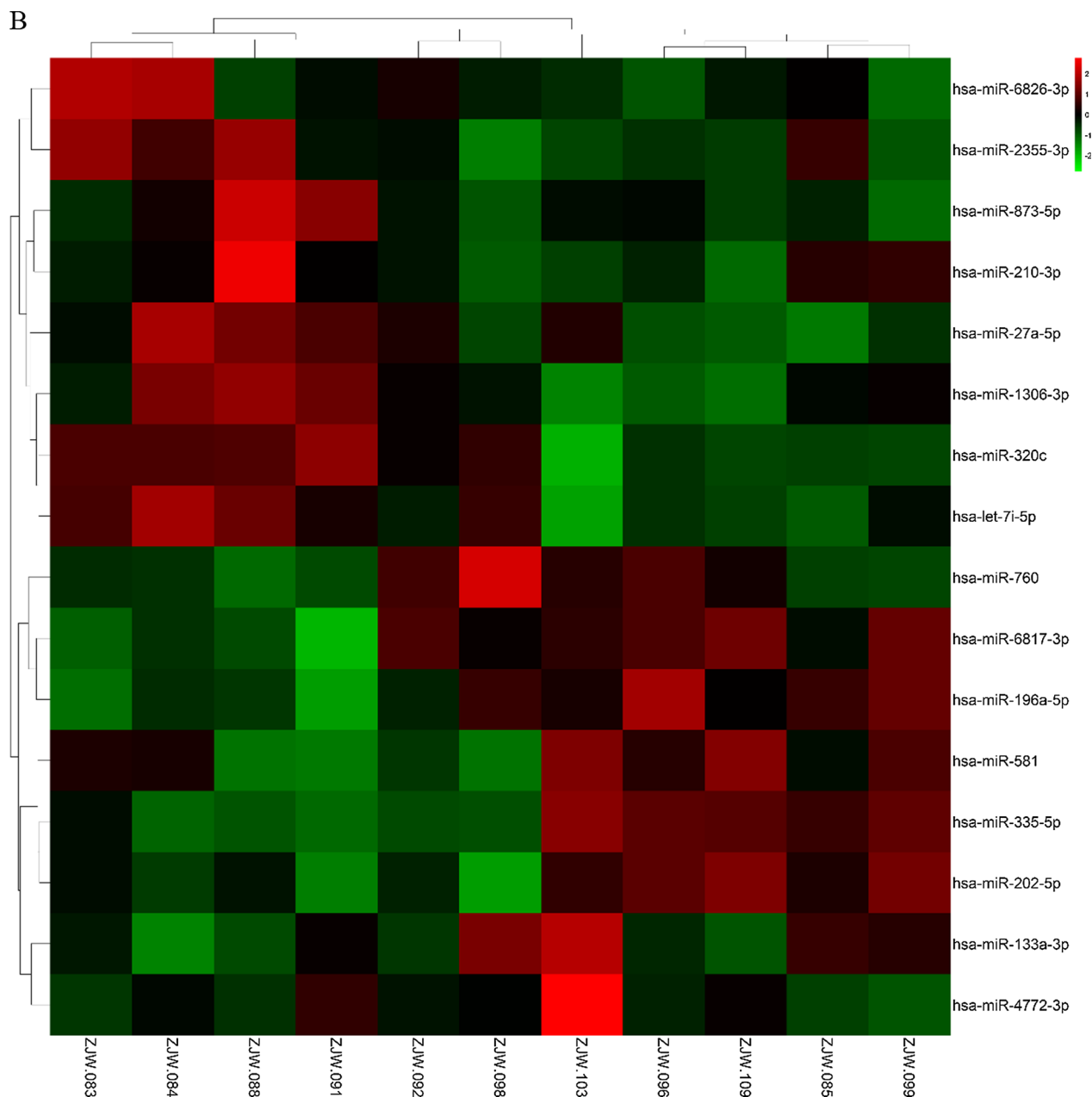


Fig. 3(b) (continued)

pathogenesis of CVD [32]. Our study showed that *RASGRP3* was associated with 4 SNPs in the mRNA-SNP-miRNA interaction network, and each intermediated one pairs of miRNA-mRNA correlations, which indicated that *RASGRP3* may play a key role in the underlying mechanism. Paramjeet and colleagues observed that *RASGRP3* could affect the role of endothelial cells in angiogenesis in diabetic mice by mediating endothelial cell signal transduction [33]. We further explored functions of the rest of mRNA in the mRNA-SNP-miRNA interaction network, and found that *KRT1* might be

associated with CVD. Gao et al. found that inhibiting *KRT1* can activate Notch signaling pathway, thereby inhibiting the inflammatory response and endoplasmic reticulum stress of vascular endothelial cells in coronary atherosclerosis [34]. Furthermore, it has been reported that inhibition of *KRT1* expression can improve myocardial ischemia-reperfusion injury by activating Notch signaling pathway [35].

Although previous studies have shown that miRNA-related SNPs play a role in the pathogenesis of T2DM complicated with CVD, there is a lack of relevant

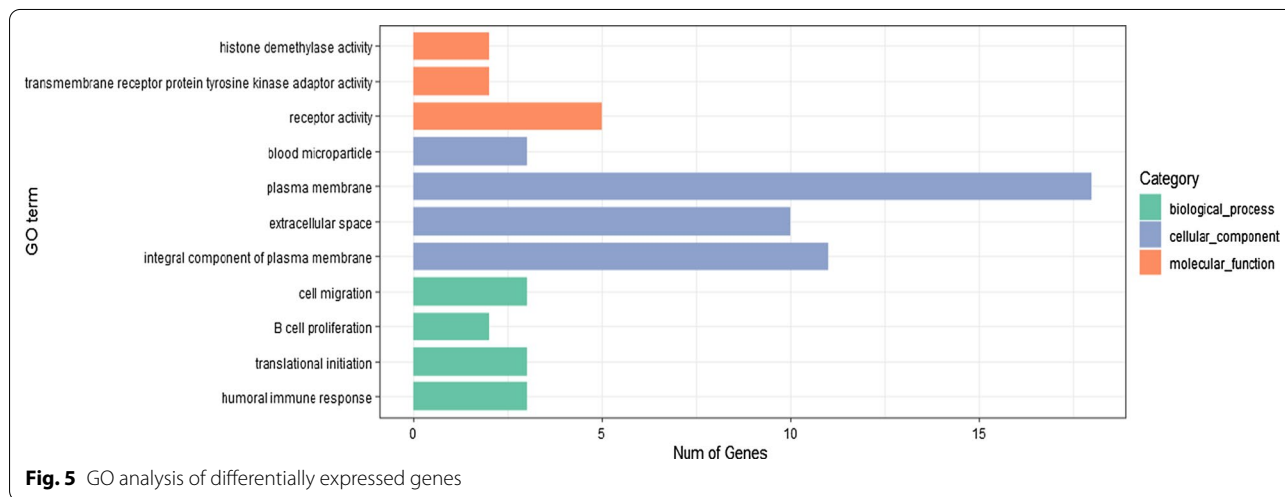
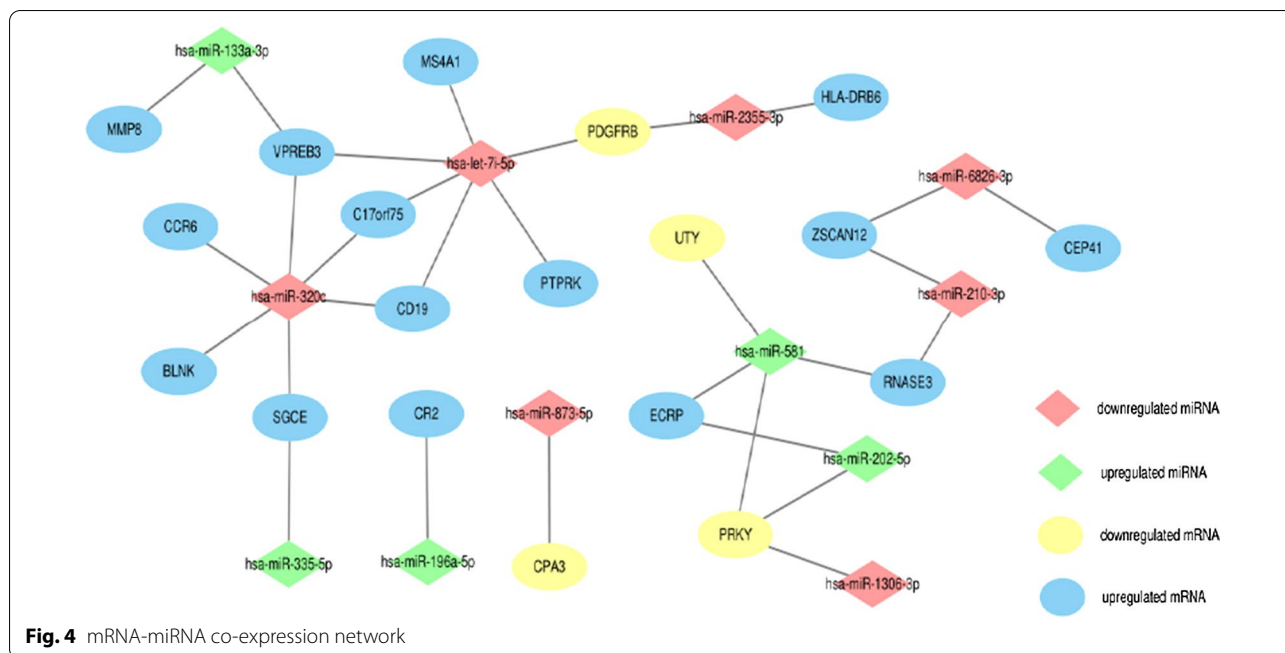
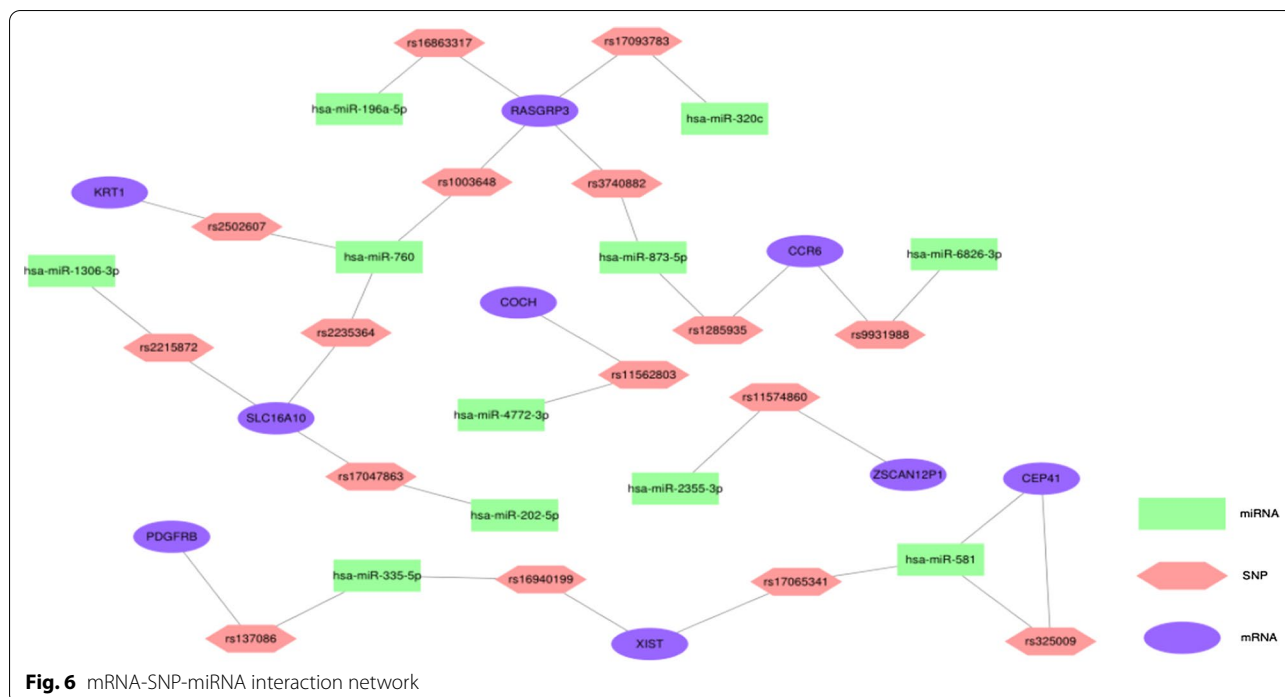


Table 3 Results of KEGG pathway enrichment analysis for differentially expressed genes

Pathway	Gene name	Regulation	p value
B cell receptor signaling pathway	<i>CD72</i>	Upregulation	< 0.01
	<i>BLNK</i>	Upregulation	
	<i>CD19</i>	Upregulation	
	<i>CR2</i>	Upregulation	
	<i>RasGRP3</i>	Upregulation	
hematopoietic cell lineage	<i>CD19</i>	Upregulation	0.03
	<i>MS4A1</i>	Upregulation	
	<i>CR2</i>	Upregulation	

researches that combine mRNA, miRNA and SNP for analysis. We searched GEO database and found that there are few studies involving miRNA sequencing for T2DM patients and T2DM patients with CVD. Hence, we recruited six diabetes patients and five diabetes patients with ischemic heart disease from communities in Beijing to obtain their miRNA expression profile. Previous studies suggested that miR-196a-5p and miR-202-5p may play a role in the pathogenesis of diabetes and its cardiovascular complications. Omer et al. revealed that compared with normal group, miR-196a-5p was downregulated in CVD group without statistically significance, and they



indicated that miR-196a-5p may be a potential biomarker for the diagnosis of coronary artery disease and acute coronary syndrome [36]. However, our present study suggested that miR-196a-5p was differentially upregulated in CVD group, which was inconsistent with previous study. We attributed the difference to the limited sample size, and studies with large samples are needed to verify the results. In addition, up-regulated miR-202-5p was found to have a protective effect on the heart of mice with myocardial ischemia–reperfusion injury [37]. But our present study identified that miR-202-5p was differentially upregulated in CVD group, and we considered that it was due to the different species in two studies which weakened the comparability.

We successfully constructed the mRNA-SNP-miRNA interaction network to visualize the relationships between mRNAs, SNP and miRNAs. However, there was no statistical significance in Pearson correlation analysis of mRNA and miRNA connected by overlapping SNPs in the mRNA-SNP-miRNA interaction network. Considering this study, the reasons may be associated with limitations in this study. First, the sample size of miRNA is so small that its results are not representative enough, and future studies including larger samples are required; second, mRNA and miRNA expression data were obtained from different samples, resulting in bias although we searched eligible databases by defining the CVD clearly and setting covariates to eliminate the bias, and miRNA, mRNA and SNP data from same experiment are needed

to provide more convincing research results. Moreover, some miRNAs cannot inhibit mRNA expression but can inhibit protein translation. Hence, the co-expression analysis of mRNA and miRNA was not statistically significant, which did not indicate that there was no regulatory relationship between them. Further study on the significance of miRNA and mRNA at protein level is required.

Conclusion

We demonstrated mRNA-SNP-miRNA interaction network and identified important pathways contributing to metabolic cardiovascular disease. These results suggested that downregulated let-7i-5p, and upregulated *RASGRP3*, *KRT1* and *CEP41* may play crucial roles in molecular mechanisms underlying the initiation and development of CVD in T2DM patients. However, further studies regarding the role of these genes and miRNAs in progression of CVD in T2DM patients are required.

Abbreviations

CVD: Cardiovascular diseases; T2DM: Type 2 diabetes mellitus; miRNA: MicroRNA; SNP: Single nucleotide polymorphisms; GEO: Gene expression omnibus; GO: Gene ontology; KEGG: Kyoto encyclopedia of genes and genomes; MRE: miRNA recognition element; GWAS: Genome-wide association studies; eQTL: Genome-wide expression quantitative trait locus.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12872-021-02166-4>.

Additional file 1: Figure S1. The normalized unscaled standard errors boxplot from quality evaluation analysis for dataset GSE66175. **Figure S2.** The RNA degradation plot from quality evaluation analysis for dataset GSE66175. **Figure S3.** KEGG pathway enrichment analysis for differentially expressed genes. **Table S1.** Sample IDs. **Table S2.** Top 10 mRNA in absolute value of log₂(fold change). **Table S3.** Top 10 miRNA in absolute value of log₂(fold change). **Table S4.** The corresponding number of file names in Fig 3A. **Table S5.** The corresponding number of file names in Fig. 3B. **Table S6.** The relationship between miRNA and SNP in mRNA-SNP-miRNA trios.

Acknowledgements

We would like to express our sincere appreciation to the support and cooperation of the doctors and the participants from Zhangjiawan community health service center of Tongzhou District.

Authors' contributions

ZF, WP and KL were responsible for design of the study, data analysis. ZF, ZW, WP and KL took part in the discussion of the paper. ZF wrote the manuscript which was reviewed and revised by LZ and KL. All authors critically revised and approved the final version of the manuscript. All authors have read and approved the manuscript.

Funding

This study was funded by research grant from the National Natural Science Foundation of China (81602908).

Availability of data and materials

Microarray datasets (GSE90074, GSE66175, and GSE90073) for this study are openly available in Gene Expression Omnibus database at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE90074>, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66175>, and <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE90073>, respectively (last accessed on 23 Jan 2020). The datasets of miRNA used during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

All participants provided written consent before entering the study. The study was approved by the Ethics Committee of Capital Medical University, with the ethical Approval Number: 2016SY24. The study was carried out in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Competing interests

All authors declare that they have no competing interests.

Author details

¹Department of Epidemiology and Health Statistics, School of Public Health, Capital Medical University, and Beijing Municipal Key Laboratory of Clinical Epidemiology, Beijing, China. ²Department of Urology, Beijing Friendship Hospital, Capital Medical University, Beijing, China.

Received: 23 February 2021 Accepted: 14 July 2021

Published online: 23 July 2021

References

- International Diabetes Federation. IDF Diabetes Atlas. 9th ed. Brussels, Belgium: International Diabetes Federation; 2019.
- Stirban A, et al. Vascular effects of advanced glycation endproducts: clinical effects and molecular mechanisms. *Mol Metab*. 2013;3:94–108.
- Ali A, et al. Prevalence of microvascular complications in newly diagnosed patients with type 2 diabetes. *Pak J Med Sci*. 2013;29(4):899–902.
- WHO. Non-communicable diseases country profile 2014. <http://www.who.int/nmh/publications/ncd-profiles-2014/en/>.
- Gaede P, et al. Multifactorial intervention and cardiovascular disease in patients with type 2 diabetes. *N Engl J Med*. 2003;348:383–93.
- Lewis BP, et al. Prediction of mammalian microRNA targets. *Cell*. 2003;115(7):787–98.
- Caporali A, et al. Deregulation of microRNA-503 contributes to diabetes mellitus-induced impairment of endothelial function and reparative angiogenesis after limb ischemia. *Circulation*. 2011;123(3):282–91.
- Caporali A, et al. p75(NTR)-dependent activation of NF-κB regulates microRNA-503 transcription and pericyte-endothelial crosstalk in diabetes after limb ischaemia. *Nat Commun*. 2015;6:8024.
- La Sala L, et al. Oscillating glucose induces microRNA-185 and impairs an efficient antioxidant response in human endothelial cells. *Cardiovasc Diabetol*. 2016;15:71.
- Saunders MA, et al. Human polymorphism at microRNAs and microRNA target sites. *Proc Natl Acad Sci USA*. 2007;104(9):3300–5.
- Nam JW, et al. Human microRNA prediction through a probabilistic co-learning model of sequence and structure. *Nucleic Acids Res*. 2005;33(11):3570–81.
- Davis BN, et al. Smad proteins bind a conserved RNA sequence to promote microRNA maturation by Drosha. *Mol Cell*. 2010;39(3):373–84.
- Haas U, et al. MicroRNA-mediated regulation of gene expression is affected by disease-associated SNPs within the 3'-UTR via altered RNA structure. *RNA Biol*. 2012;9(6):924–37.
- Sethupathy P, et al. MicroRNA target site polymorphisms and human disease. *Trends Genet*. 2008;24(10):489–97.
- CARDIoGRAMplusC4D Consortium, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat Genet*. 2013;45(1):25–33.
- Clarke P, et al. The impact of diabetes-related complications on health-care costs: results from the United Kingdom Prospective Diabetes Study (UKPDS Study No. 65). *Diabet Med*. 2003;20(6):442–450.
- Qi L, et al. Association between a genetic variant related to glutamic acid metabolism and coronary heart disease in individuals with type 2 diabetes. *JAMA*. 2013;310(8):821–8.
- Pickrell JK, et al. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature*. 2010;464(7289):768–72.
- Heinig M, et al. A trans-acting locus regulates an anti-viral expression network and Type 1 diabetes risk. *Nature*. 2010;467(7314):460–4.
- Lin X, et al. Genome-wide integrative analysis identified SNP-miRNA-mRNA interaction networks in peripheral blood mononuclear cells. *Epigenomics*. 2017;9(10):1287–98.
- Kamiński MJ, et al. In-silico identification of cardiovascular disease-related SNPs affecting predicted microRNA target sites. *Pol Arch Med Wewn*. 2013;123(7–8):355–63.
- Hu Y, et al. Suppression of miRNA let-7i-5p promotes cardiomyocyte proliferation and repairs heart function post injury by targeting CCND2 and E2F2. *Clin Sci (Lond)*. 2019;133(3):425–41.
- Satoh M, et al. A cellular microRNA, let-7i, is a novel biomarker for clinical outcome in patients with dilated cardiomyopathy. *J Card Fail*. 2011;17:923–9.
- Zhang H, et al. Silencing TTTY15 mitigates hypoxia-induced mitochondrial energy metabolism dysfunction and cardiomyocytes apoptosis via TTTY15/let-7i-5p and TLR3/NF-κB pathways. *Cell Signal*. 2020;76:109779.
- Choi J, et al. Gelidium elegans extract ameliorates type 2 diabetes via regulation of MAPK and PI3K/Akt signaling. *Nutrients*. 2018;10(1):51.
- Feng C, et al. SHBG expression is correlated with PI3K/AKT pathway activity in a cellular model of human insulin resistance. *Gynecol Endocrinol*. 2018;34(7):567–73.
- King GL, et al. Selective insulin resistance and the development of cardiovascular diseases in diabetes: The 2015 Edwin Bierman Award Lecture. *Diabetes*. 2016;65(6):1462–71.
- Wu JHY, et al. Omega-6 fatty acid biomarkers and incident type 2 diabetes: pooled analysis of individual-level data for 39740 adults

- from 20 prospective cohort studies. *Lancet Diabetes Endocrinol*. 2017;5(12):965–74.
29. Marklund M, et al. Biomarkers of Dietary Omega-6 Fatty Acids and Incident Cardiovascular Disease and Mortality. *Circulation*. 2019;139(21):2422–36.
 30. Yan W, et al. Characteristics of B cell-associated gene expression in patients with coronary artery disease. *Mol Med Rep*. 2016;13(5):4113–21.
 31. Fuster JJ, et al. Clonal hematopoiesis associated with TET2 deficiency accelerates atherosclerosis development in mice. *Science*. 2017;355(6327):842–7.
 32. Ki SM, et al. Mediated ciliary tubulin glutamylation drives angiogenesis through AURKA-dependent deciliation. *EMBO Rep*. 2020; 21(2):48290.
 33. Randhawa PK, et al. The Ras activator RasGRP3 mediates diabetes-induced embryonic defects and affects endothelial cell migration. *Circ Res*. 2011;108(10):1199–208.
 34. Gao ZF, et al. microRNA-107 protects against inflammation and endoplasmic reticulum stress of vascular endothelial cells via KRT1-dependent Notch signaling pathway in a mouse model of coronary atherosclerosis. *J Cell Physiol*. 2019;234(7):12029–41.
 35. Fang HC, et al. KRT1 gene silencing ameliorates myocardial ischemia-reperfusion injury via the activation of the Notch signaling pathway in mouse models. *J Cell Physiol*. 2019;234(4):3634–46.
 36. Omer Faruk Karacorlu, et al. Circulating miR-196a-5p miR-373-3p and miR-375: novel candidate biomarkers for diagnosis of acute coronary syndrome. *Meta Gene*. 2018; 17.
 37. Li Y, et al. miR-202-5p protects rat against myocardial ischemia reperfusion injury by downregulating the expression of Trpv2 to attenuate the Ca²⁺ overload in cardiomyocytes. *J Cell Biochem*. 2019;120(8):13680–93.

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

