# Identification of differentially expressed genes in the endothelial precursor cells of patients with type 2 diabetes mellitus by bioinformatics analysis

ZHIDA SHEN[1*], QI CHEN[1*], HANGYING YING[1], ZETAO MA[1], XUKUN BI[1], XIAOTING LI[1], MEIHUI WANG[2], CHONGYING JIN[1], DONGWU LAI[1], YANBO ZHAO[1] and GUOSHENG FU[1]

[1]Department of Cardiology; [2]Biomedical Research Center, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310016, P.R. China

**Abstract.** Type 2 diabetes mellitus (DM) is a metabolic disease with worldwide prevalence that is associated with a decrease in the number and function of endothelial progenitor cells (EPCs). The aim of the present study was to explore the potential hub genes of EPCs in patients with type 2 DM. Differentially expressed genes (DEGs) were screened from a public microarray dataset (accession no. GSE43950). Pathway and functional enrichment analyses were performed using the Database for Annotation, Visualization and Integrated Discovery. The protein-protein interaction (PPI) network was visualized. The most significantly clustered modules and hub genes were identified using Cytoscape. Furthermore, hub genes were validated by quantitative PCR analysis of EPCs isolated from diabetic and normal subjects. Subsequently, weighted gene co-expression network analysis (WGCNA) was performed to identify the modules incorporating the genes exhibiting the most significant variance. A total of 970 DEGs were obtained and they were mainly accumulated in inflammation-associated pathways. A total of 9 hub genes were extracted from the PPI network and the highest differential expression was determined for the interleukin 8 (IL8) and CXC chemokine ligand 1 (CXCL1) genes. In the WGCNA performed to determine the modules associated with type 2 DM, one module incorporated IL8 and CXCL1.

Finally, pathway enrichment of 10% genes in the pink module ordered by intramodular connectivity (IC) was associated with the IL17 and the chemokine signaling pathways. The present results revealed that the expression of IL8 and CXCL1 may serve important roles in the pathophysiology of EPCs during type 2 DM and inflammatory response may be critical for the reduced number and hypofunction of EPCs isolated from patients with diabetes.

## Introduction

Diabetes mellitus (DM) is a worldwide health care problem that places a heavy burden on patients and society. Diabetes and other parameters of metabolic disorder are considered risk factors for cardiovascular disease. Diabetic patients exhibit a two- to four-fold increased risk of cardiovascular disease, while endothelial cell dysfunction has a major role in the initiation and progression of vascular complications (1). The endothelium is a semipermeable monolayer of spindle-shaped endothelial layers that help maintain vascular homeostasis under physiological conditions (2). In diabetes, endothelial function is compromised, including reduced plasma nitric oxide, increased generation of reactive oxygen species, increased leukocyte infiltration and subsequent inflammation load (3,4). Therefore, ameliorating endothelial dysfunction is a major focus for the prevention and treatment of diabetic vascular complications. Compared with type 1 DM, which is characterized by the autoimmune destruction of β-cells, type 2 DM accounts for 90-95% of all cases of diabetes (5). Type 2 DM is insulin-independent and usually characterized by a partial reduction of insulin secretion and insulin resistance (6).

Endothelial precursor cells (EPCs) are a group of cells with the inherent capacity to differentiate into mature cells (7). Asahara *et al* (8) reported on the isolation of these CD34-positive mononuclear cells from human peripheral blood. EPCs have been indicated to integrate into the capillary-vessel endothelium of rodent hindlimbs induced by ligation of the artery (9). Previous studies also revealed the potency of EPCs in the treatment of endothelial dysfunction induced by diabetes (10,11). However, compared with those in healthy subjects, EPC counts were lower, and the function was

*Correspondence to:* Dr Yanbo Zhao or Dr Guosheng Fu, Department of Cardiology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, 3 East Qingchun Road, Hangzhou, Zhejiang 310016, P.R. China
E-mail: zhaoyb@zju.edu.cn
E-mail: fugs@zju.edu.cn

*Contributed equally

also disturbed in patients with either type of DM (12-14). The development of strategies to improve the counts and activity of EPCs in patients with DM is a major focus in the field of autologous cell therapy. As EPCs from patients with diabetes exhibit different behaviors compared with those from healthy subjects, several different approaches have been investigated to restore their dysregulation and dysfunction by targeting specific sites (15-18).

In the present study, differentially expressed genes (DEGs) in EPCs from patients with type 2 DM vs. healthy subjects were identified. These DEGs were then subjected to gene ontology (GO) and pathway enrichment analyses. A protein-protein interaction (PPI) network was then constructed and visualized, and hub genes were identified by molecular complex detection (MCODE). The top 9 hub genes were subsequently verified by reverse transcription-quantitative (RT-q)PCR in an independent sample set originating from our study center. To further explore the genes that may be associated with the hub genes, a weighted gene co-expression network analysis (WGCNA) was performed to determine a relevant module that incorporates the hub genes, especially interleukin 8 (IL8) and CXC chemokine ligand 1 (CXCL1).

**Materials and methods**

*Obtainment and pre-processing of microarray data*. Gene expression profiles of EPCs from healthy and type 2 diabetic subjects were obtained from the Gene Expression Omnibus database (GEO; www.ncbi.nlm.nih.gov/geo/). The accession number was GSE43950, and this dataset included a total of 14 samples: A total of 9 type 2 diabetes late stage EPC samples and 5 healthy late EPC stage samples. The definition of early and late stage EPCs is discussed in a previous study (19). Late EPC samples referred to the EPCs appeared aged 2-4 weeks and exhibited a cobblestone-like morphology (8). The 9 type 2 diabetes late EPC samples consisted of 5 samples obtained from the type 2 diabetes patients with microvascular complications and 4 samples from the type 2 diabetes patients without clinical microvascular injuries. To identify the DEGs in patients with type 2 diabetes vs. non-diabetic controls, the 9 diabetes samples were analyzed together. The platform used was the Rosetta/Merck Human RSTA Custom Affymetrix 2.0 microarray GPL10379. First, the expression matrix from the GEO database was pre-processed using the robust multi-array analysis method. The probe ID for each gene was then converted to a gene symbol using annotation files obtained from the platform. DEGs were identified by the limma algorithm (http://www.bioconductor.org/packages/2.9/bioc/html/limma.html) in R software (20,21). A P-value of <0.05 and |$\log_2$ fold change|≥1 were used as the cutoff criteria for this analysis.

*Enrichment analysis of DEGs*. GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway functional enrichment analyses were performed using the database for annotation, visualization and integrated discovery (DAVID; https://david.ncifcrf.gov/) (22). P<0.05 was considered to indicate statistical significance and the GO results were ranked by P-value. The significant terms for biological process (BP), cellular component (CC) and molecular function (MF) were visualized in Cytoscape (23). Furthermore, the top 10 GO terms in the categories BP, CC and MF, and the top 10 KEGG pathways were presented in bubble plots generated with ggplot2 (https://cran.rproject.org/web/packages/ggplot2/index.html) in R.

*PPI network construction*. A PPI network was generated to infer the interactions among proteins. The PPI network was constructed using the Search Tool for the Retrieval of Interacting Genes and proteins (STRING) database (https://string-db.org/) (24). To explore the regulatory mechanisms, interactions with the highest confidence of a combined score >0.900 were imported into Cytoscape to construct the PPI network.

*Module selection*. To obtain clusters of genes in the PPI network, MCODE was used to identify the modules in the PPI network. The cutoff criteria were 'degree cutoff=2', 'node score cutoff=0.2', 'k-core=2' and 'maximum depth=100'.

*Analysis of hub genes*. To obtain a balance between the core genes and to avoid missing any key gene, hub genes were extracted using cytoHubba. A total of 12 topological analyses were provided by the cytoHubba plugin (25). In accordance with previous reports, a total of 3 most widely used topological analysis methods, including maximal clique centrality (MCC), maximum neighborhood component (MNC) and density of maximum neighborhood component (DMNC), were used to identify potential hub genes (25,26). The overlapping genes were selected as the hub genes using Venn diagrams.

*Construction of the co-expression network and identification of significant modules*. The WGCNA package in R was used to construct a co-expression network (27). After calculating the correlation of all pairwise genes by Pearson's correlation matrices, the appropriate soft-threshold power β was selected for the construction of modules using the pickSoftThreshold function ranging from 1 to 30. The power of β=16 (scale-free $R^2$=0.818) was selected to construct a scale-free network. The topological overlap matrix (TOM) was therefore constructed to measure the network connectivity of each gene, defined as the sum of its adjacency with all other genes for network generation (28). The interaction analysis among different co-expression modules was performed using the flashClust function, and the average linkage hierarchical clustering was used with a minimum size of 500 for the gene dendrogram (29). The module-trait association was estimated using the correlation between the module eigengene and the phenotype. Gene significance (GS) was defined as the absolute value of the correlation between expression profile and clinical trait. Module membership (MM) was designated as the correlation of the expression profile and each module eigengene (30). Genes with higher intramodular connectivity (IC) were thought to have biological significance compared with other genes in the module (31). Therefore, the top 10% genes ordered by IC were functionally annotated. Finally, KEGG functional enrichment analysis of pathways in the pink module was performed using the clusterProfiler package (32).

*Isolation and cultivation of circulating EPCs*. The EPCs used in the present study were obtained from patients diagnosed

Table I. Primers used for quantitative PCR.

| | Primer sequence (5'→3') | |
| --- | --- | --- |
| Gene | Forward | Reverse |
| IL8 | TTTTGCCAAGGAGTGCTAAAGA | AACCCTCTGCACCCAGTTTTC |
| FPR1 | AAGGCCATGGGAGGACATTG | CAGGGCCCAATGATCACCTT |
| CXCL1 | CTGGCGGATCCAAGCAAATG | GCCCCTTTGTTCTAAGCCAG |
| GNAI3 | ATCGACCGCAACTTACGGG | AGTCAATCTTTAGCCGTCCCA |
| FPR2 | CTGAATGGATCAGAAGTGGTGG | CCCAAATCACTAGTCCATTGCC |
| GNAI2 | TACCGGGCGGTTGTCTACA | GGGTCGGCAAAGTCGATCTG |
| ANXA1 | GCGGTGAGCCCCTATCCTA | TGATGGTTGCTTCATCCACAC |
| GNB1 | GTGAGCTTGACCAGTTACGG | TGTGATCTGAGAGAGAGTTGCAT |
| LPAR5 | CGCCATCTTCCAGATGAAC | TAGCGGTCCACGTTGATG |
| GAPDH | CCTGCACCACCAACTGCTTA | GGCCATCCACAGTCTTCTGAG |

with type 2 DM or from non-diabetic controls at Sir Run Run Shaw Hospital (Hangzhou, China). Type 2 DM was diagnosed according to the American Diabetes Association criteria (33). Inclusion/exclusion criteria for the patients enrolled in the present study were as follows: Age >50 years, willingness to provide written informed consent, no acute myocardial infarction and no acute stroke. The patients with type 2 DM were diagnosed >5 years previously. Furthermore, major diabetic complications (diabetic retinopathy, diabetic nephropathy, diabetic neuropathy, lower limb arteriopathy) were recorded according to criteria described previously (34). A total of 8 patients were enrolled in the present study between April 2017 and April 2018. EPCs were isolated, cultured and characterized according to methods previously reported (35). The study procedure was approved by the Ethics Review Board of Sir Run Run Shaw Hospital, School of Medicine. In brief, peripheral blood mononuclear cells were isolated by density gradient centrifugation. Cells were cultured in Endothelial Basal Medium-2 (Lonza Group, Ltd.) supplemented with corresponding nutrient factors including 10% FBS (Gibco; Thermo Fisher Scientific, Inc.). Non-adherent cells were removed after 4 days of primary culture. The culture was maintained for another 7 days following addition of fresh medium. The purity of the EPCs was verified using tetramethylindocarbocyanine-labelled acetylated low-density lipoprotein (DiLDL) uptake and lectin binding under a laser scanning confocal microscope (data not shown) (Zeiss AG). The third generation of EPCs was selected for the subsequent analysis. Detailed information on these patients is provided in Supplemental Table SI.

*Validation of microarrays by RT-qPCR.* To verify the major conclusions drawn from the microarray data, the expression levels of genes encoding for identified hub genes in EPCs from patients were determined. The RNA of EPCs was extracted using TRIzol reagent (CWBIO) and converted into complementary (c)DNA using a PrimeScript RT reagent Kit (Takara Biotechnology, Co. Ltd) according to the manufacturer's protocol. Subsequently, qPCR was performed using UltraSYBR Mixture (Low ROX; CWBIO) on the Lightcycler

480 II system (Roche Life Science). For each well, the reaction mixture contained 2 $\mu$l cDNA, 0.4 $\mu$l forward primer (0.2 $\mu$M), 0.4 $\mu$l reverse primer (0.2 $\mu$M), 7.2 $\mu$l RNA-free water and 10 $\mu$l 2X UltraSYBR Mix. The reactions were incubated at 95°C for 10 min, followed by 40 cycles of 95°C for 15 sec and 60°C for 1 min. The expression data were normalized to the reference GAPDH. FCs of relative mRNA expression were calculated using the change-in-threshold method ($2^{-\Delta\Delta Cq}$) (37). The primer sets used for qPCR are listed in Table I.

*Statistical analysis.* All experiments were performed at least three times. All results are expressed as the mean ± standard deviation. Statistical analyses were performed using Student's t-test with GraphPad Prism 7 software (GraphPad Software, Inc.). P<0.05 was considered to indicate statistical significance.

**Results**

*Identification of DEGs and clustering analysis.* A total of 970 DEGs were finally screened from the expression profiles and are presented in a volcano plot in Fig. 1A.

Clustering analysis was then performed using the pheatmap package in R and 970 DEGs were divided into two major clusters (Fig. 1B). They included 822 upregulated and 148 downregulated genes (Fig. 1C). The top 10 upregulated and top 10 downregulated DEGs ranked by P-value are presented in Table II.

*Enrichment analysis of DEGs.* To explore the GO terms and pathways in which the DEGs were mainly involved, the 970 DEGs were uploaded onto the DAVID website. The most enriched GO terms in the category BP included 'inflammatory response', 'cellular response to lipopolysaccharide', 'innate immune response', 'immune response' and 'lipopolysaccharide-mediated signaling pathway' (Fig. 2A). In the category CC, the GO terms with the highest enrichment were mainly associated with 'extracellular exosomes', 'plasma membranes', 'lysosomes and cell surface' (Fig. 2B). As for the category MF, the GO terms with the highest accumulation of DEGs were
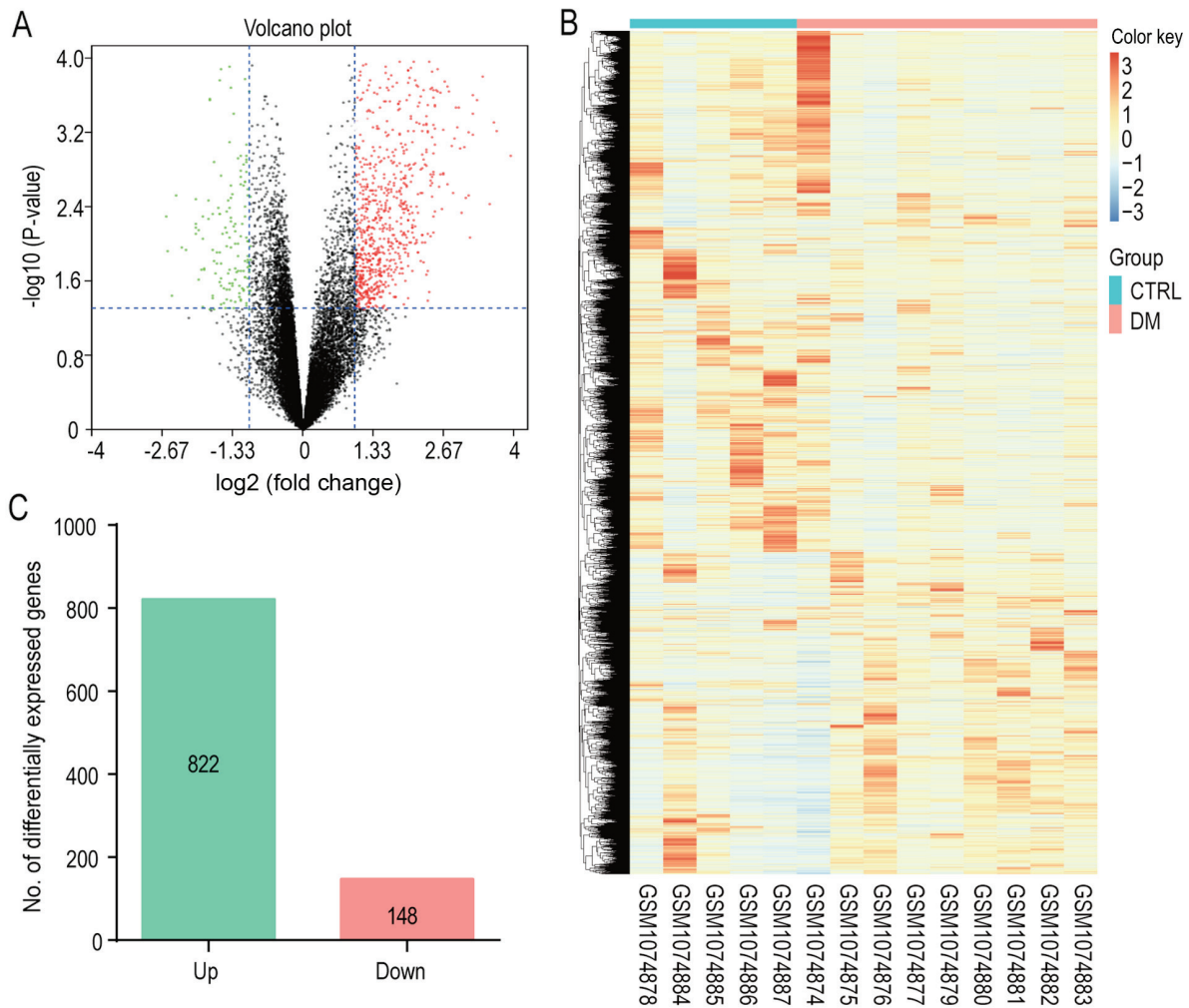
Figure 1. The volcano plot and heatmap of differentially expressed genes in type 2 DM. (A) Volcano plot of DEGs. The x-axis represents the log2 FC and the y-axis represents the log10 (P-value). The green dots represent downregulated genes and red dots represent upregulated genes. (B) Heat map of DEGs following clustering analysis. The vertical axis represents the sample, and the horizontal axis represents DEGs. Red indicates upregulated and blue represents downregulated genes. (C) DEGs were selected according to P<0.05 and |log2 FC|≥1. DEGs, differentially expressed genes; DM, diabetes mellitus; CTRL, control; FC, fold change.

'protein binding, receptor activity', 'carbohydrate binding' and 'lipopolysaccharide receptor activity' (Fig. 2C).

To identify pathways significantly enriched by the DEGs, KEGG pathway analysis was performed. The enrichment analysis of the KEGG pathways demonstrated that DEGs were mainly enriched in the 'NF-κB signaling pathway', 'tumor necrosis factor (TNF) signaling pathway' and 'inflammatory bowel disease' (Fig. 2D).

*PPI network construction and module analysis*. A total of 394 nodes and 1,486 interactional pairs were included in the PPI network (Fig. 3). Among them, proteins with degrees of ≥20 were visualized in detail. A total of 5 proteins had degrees of interaction of ≥30 [IL8, 43; ubiquitin conjugating enzyme E2 D1 (UNE2D1), 35; G protein subunit beta 1 (GNB1), 33; E1A binding protein p300 (EP300), 32; Cbl proto-oncogene (CBL), 30]. In addition, a subnetwork clustering analysis was performed using the MCODE plugin. The top 3 subnet modules were identified (Fig. 4A-C). Module a (score=20.634) included 42 nodes and 423 interactional pairs (Fig. 4A). Module b (score=12.000) included 12 nodes and 66 interactional pairs

(Fig. 4B). Module c (score=7.641) included 40 nodes and 149 interactional pairs (Fig. 4C).

*Hub gene selection*. The hub genes were determined by overlapping the top 20 genes obtained using three topological analysis methods, MCC, MNC and DMNC (Fig. 5A). A total of 9 genes were selected, including 8 upregulated genes and 1 downregulated gene (Fig. 5B; Table III).

*Validation of hub genes by RT-qPCR*. RT-qPCR was used to validate hub genes obtained in the above analysis. EPCs were isolated from 4 non-diabetic controls and 4 diabetic patients. The detailed information of these patients were presented in Table S1. The patients that met the criteria were included from April 2017 to April 2018. The mRNA expression levels of IL8, formyl peptide receptor 1 (FPR1), CXCL1, GNAI3, FPR2, GNAI2, annexin A1 (ANXA1), GNB1 and lysophosphatidic acid receptor 5 (LPAR5) were determined. Overall, the results were similar to those of the microarray analysis, while the RT-qPCR results indicated greater fold changes than those calculated from the gene array data (Fig. 6). Of

Table II. Top 10 up- and downregulated differentially expressed genes.

A, Upregulated genes

| Gene | Average expression value | Log$_2$ FC | P-value |
|---|---|---|---|
| NLRP3 | 7.94243 | 3.05234 | 2.44x10$^{-10}$ |
| NINJ1 | 8.64930 | 2.94253 | 2.39x10$^{-9}$ |
| G0S2 | 9.91064 | 6.10854 | 1.88x10$^{-8}$ |
| CKAP4 | 6.42605 | 2.65448 | 8.77x10$^{-8}$ |
| MAFB | 9.35177 | 3.80994 | 1.78x10$^{-7}$ |
| C20orf24 | 9.77482 | 1.48972 | 2.72x10$^{-7}$ |
| C5AR1 | 7.62619 | 4.01330 | 3.48x10$^{-7}$ |
| PTX3 | 6.21338 | 3.11387 | 3.64x10$^{-7}$ |
| PFKFB3 | 8.33607 | 2.69583 | 4.12x10$^{-7}$ |
| ID2 | 10.66494 | 1.77079 | 5.55x10$^{-7}$ |

B, Downregulated genes

| Gene | Average expression value | Log$_2$ FC | P-value |
|---|---|---|---|
| TRUB2 | 4.16713 | -1.14507 | 8.59x10$^{-7}$ |
| PKIA | 5.22172 | -1.33779 | 3.09x10$^{-5}$ |
| GPATCH4 | 7.44535 | -1.25382 | 3.87x10$^{-5}$ |
| EIF5B | 8.92228 | -1.07249 | 4.75x10$^{-5}$ |
| CNNM3 | 6.16580 | -1.40637 | 1.23x10$^{-4}$ |
| CD72 | 3.27212 | -1.55625 | 1.30x10$^{-4}$ |
| ZDHHC14 | 5.93256 | -1.11224 | 1.68x10$^{-4}$ |
| ENO2 | 4.21881 | -1.58752 | 1.72x10$^{-4}$ |
| PPM1K | 3.60385 | -1.38364 | 2.05x10$^{-4}$ |
| DAZAP1 | 9.80455 | -1.03973 | 2.29x10$^{-4}$ |

FC, fold change.

all 9 genes, IL8 and CXCL1 exhibited the most significant changes.

*WGCNA and key module identification.* To further explore the roles of chemokines in EPCs under diabetic conditions, WGCNA was used to identify the relevant modules incorporating IL8 and CXCL1. Prior to construction of the co-expression network, the soft-threshold power of 16 was selected to define the adjacency matrix based on the criterion of approximate scale-free topology (Fig. 7A and B). After determining the soft-threshold, the TOM was built. A total of 9 distinct gene co-expression modules for EPCs were constructed and are presented in different colors (Fig. 7C). IL8 and CXCL1 are included in the pink module (Fig. 8A). The values between module membership (MM) and gene significance (GS) in the pink module were presented in Fig. 8B. Genes with high IC were indicated to be tightly associated with other genes. The top 10% genes in the pink module ordered by IC were mainly enriched in the pathway of IL17 and the chemokine signaling pathway using KEGG pathway analysis (Fig. 8C).

**Discussion**

EPCs are a group of multipotent precursor cells that have the capacity to differentiate into blood vessels and blood cells (38). However, in previous studies, the number and function of EPCs were indicated to be decreased in patients with type 2 DM (12,39). Thus, identification of DEGs may be beneficial for the elucidation of the pathophysiology of EPCs in patients with diabetes. By analyzing the expression profiles in the dataset GSE43950 downloaded from the GEO database, 970 DEGs were identified, including 822 upregulated and 148 downregulated genes. Subsequent analyses included GO enrichment analysis of DEGs, KEGG pathway analysis, module selection by MCODE and identification of hub genes by cytoHubba. Thereafter, hub genes selected from the PPI network were validated by RT-qPCR and associated modules containing hub genes with the highest FC were investigated. To the best of our knowledge, the present study was the first aiming to determine potential genes associated with EPCs in patients with type 2 DM. The present study provides novel insight into the molecular mechanisms of dysfunctional EPCs in patients with type 2 DM.

Through the use of DAVID, GO enrichment analysis of the 970 DEGs was performed and the most enriched GO terms in the category BP were associated with 'inflammatory response, response to lipopolysaccharide' and 'innate immune response' in patients with type 2 DM. In the category CC, enriched GO terms were mainly associated with 'extracellular exosomes', 'plasma membranes' and 'lysosomes'. In the category MF, GO terms enriched for DEGs in DM groups included 'protein binding' and 'receptor activity'. KEGG pathway analysis indicated that the DEGs of the type 2 DM group were mainly enriched in the 'TNF signaling pathway' and 'NF-κB signaling pathway'. The GO terms 'inflammatory response', 'response to lipopolysaccharide' and 'innate immune response' fit well with concept of an inflammatory phenotype in the EPCs of type 2 DM as described previously (15). Indeed, it has been indicated that anti-diabetic drugs that possess the capacity to increase EPC numbers and inhibit premature apoptosis may protect EPCs from injury by DM partly through alleviation of inflammation (40). In the category CC, the GO term 'extracellular exosome' was indicated to be of high relevance. Exosomes have been reported to have an important role in the performance of EPCs under certain physiological and pathological conditions (41-43). It is reasonable to speculate that the decreased number and inhibited function of EPCs may partially be attributed to the altered exosome content of EPCs in patients with type 2 DM. These exosomes may perform multiple biological functions, including angiogenesis and wound healing (43,44). In the category MF, DEGs were mainly enriched in 'protein binding', 'receptor activity' and 'lipopolysaccharide receptor activity'. As for the KEGG pathways, the 'NF-κB signaling pathway' and 'TNF signaling pathway' were enriched in EPCs of patients with type 2 DM. Above all, these results indicated that inflammation may have a major role in the initiation and progression of dysfunction of EPCs in patients with DM, in accordance with previous studies (15,43).

To illustrate the interaction of the DEGs, a PPI network was constructed using the STRING database. The top 3 subnet modules based on the PPI network were then selected by
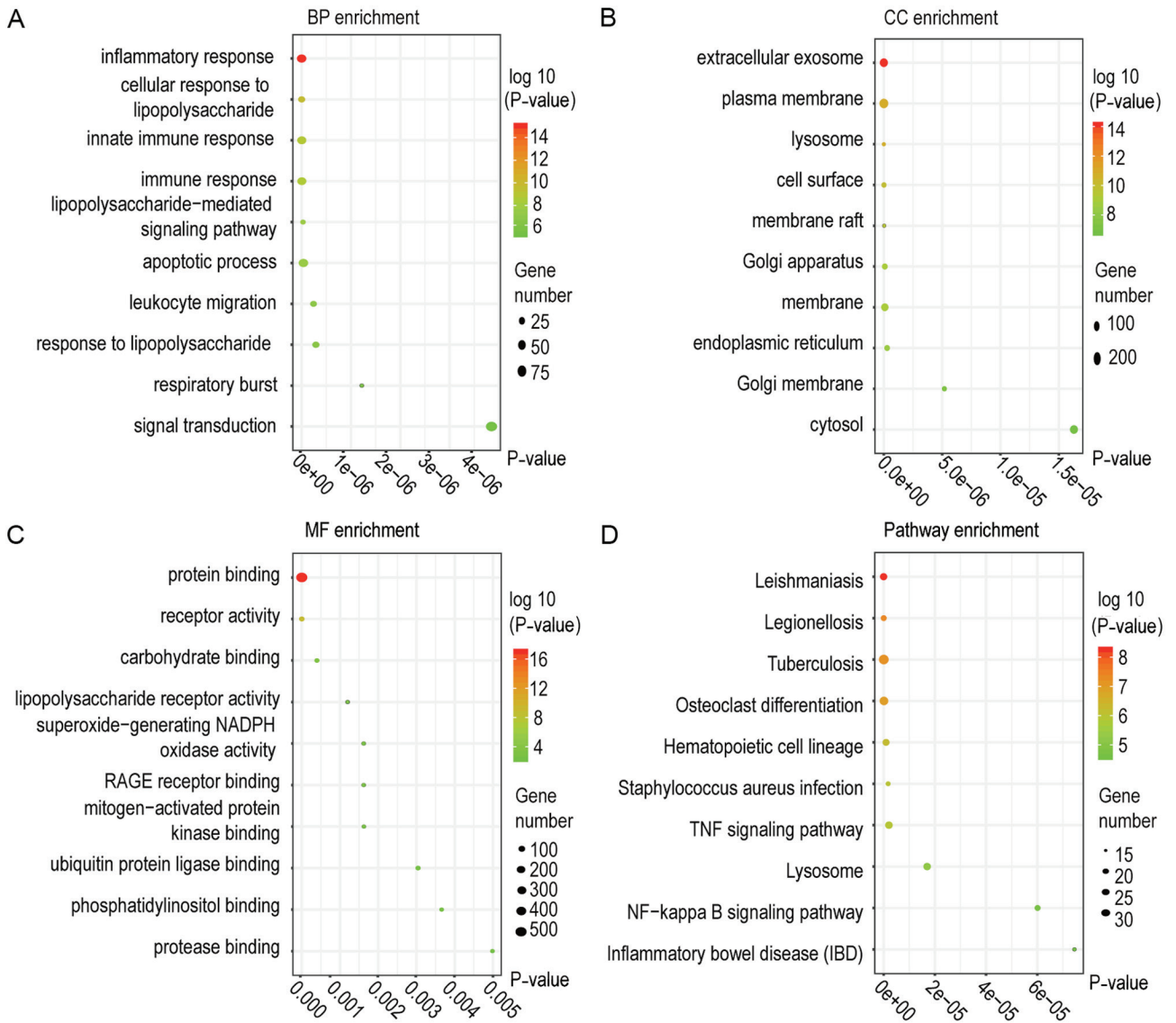
Figure 2. GO and KEGG enrichment analyses of DEGs. (A-C) Scatterplots of the top 10 enriched (A) BP, (B) CC and (C) MF terms for the DEGs. (D) The top 10 KEGG pathways for the DEGs. The size and color of the dots represent the gene number and the range of P-values, respectively. DEGs, differentially expressed genes; GO, Gene Ontology; BP, Biological Process; CC, Cellular Component; MF, Molecular Function; KEGG, Kyoto Encyclopedia of Genes and Genomes.

MCODE. GO enrichment analysis of these modules (including subnet modules a, b, and c) revealed that these subnet modules were mainly enriched in the terms inflammation, endocytosis and Golgi vesicle-mediated transport processes. The results of the enrichment analysis in the subnet modules further demonstrated that inflammation may be crucial for the dysfunction of EPCs in patients with type 2 DM.

By overlapping the top 20 genes obtained from the MCC, MNC and DMNC methods, 9 genes were selected (IL8, FPR1, CXCL1, GNAI3, FPR2, GNAI2, ANXA1, GNB1 and LPAR5). Among these genes, IL8 exhibited the most significant difference between the diabetic and control samples. IL8, as a member of the chemokines, has been reported to activate the motile apparatus of neutrophils, inducing the surface adhesion of inflammatory cells (45). Monocytes and macrophages are usually considered the principal cellular source of IL8. However, a wide variety of nucleated cells, including EPCs, are potential sources of IL8 (46). Previous studies have indicated

that in patients with type 2 DM, IL8 is elevated in the plasma compared with that in healthy subjects (47). IL8 is considered a canonical angiogenic factor and has a crucial role in the protection of EPCs for attenuating injury and prompting recovery of damaged tissues (48-50). Consistent with a previous study, the current study indicated that IL8 was significantly higher in the microarrays of EPCs isolated from patients with diabetes (51). Furthermore, the results of detecting mRNA expression in EPCs under type 2 DM by *RT-qPCR* were similar. All these results indicated that elevated IL8 may be involved in the dysfunction of EPCs during the long-term injury associated with type 2 DM.

Besides IL8, CXCL1 was also significantly elevated with a similar degree to that of IL8. CXCL1, best known for its chemotactic activity toward neutrophils and monocytes/macrophages, is also a potent angiogenic factor (52). Herlea-Pana *et al* (53) reported that CXCL1 and its cognate receptor, CXC motif chemokine-receptor-2 (CXCR2), were
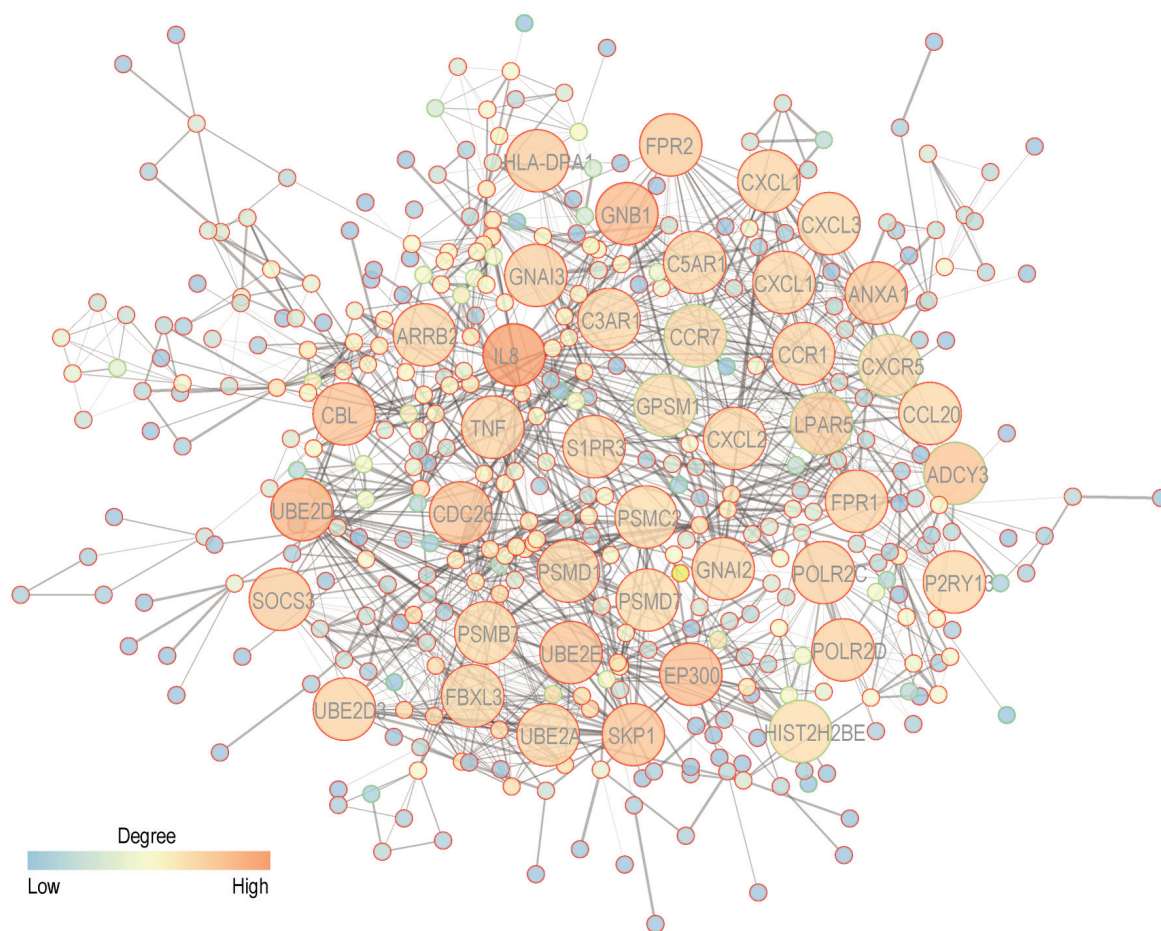
Figure 3. Protein-protein interaction network of differentially expressed genes. Proteins with degrees ≥20 were visualized in detail. Circles with green and red margins represent the downregulated and upregulated genes, respectively. Lines indicate protein-protein interactions.

increased in Reversa mice, an animal model apt to develop plaques when fed an atherosclerosis-inducing western diet. Of note, elevated CXCL1-CXCR2 was indicated to have a protective effect to prevent the progression of plaque formation, while CXCR2 was able to recruit EPCs to the plaques to repair the injured endothelium. Likewise, the chemokine system was indicated to be associated with homing and engraftment of EPCs (54). Above all, these results indicated that the chemokine system has an important role in the regulation of EPCs under diabetic conditions.

To further explore the genes that may be associated with IL8 and CXCL1 in the microarray data, WGCNA was performed to determine the relevant module that incorporates IL8 and CXCL1. A total of 9 modules were recognized by dynamic tree cut, of which the pink module contained IL8 and CXCL1. Most genes in this module were enriched in the TNF signaling pathway, the chemokine signaling pathway and the NOD-like receptor signaling pathway. These pathways fit well with the concept of an inflammatory and immune system disorder in type 2 DM (55,56).

The microarray data analyzed in the present study were deposited in the GEO as the dataset GSE43950. By analyzing these microarray data, the present study aimed to identify candidate genes that may be helpful for understanding the roles of EPC dysfunction under type 2 DM conditions and the associated mechanisms. Furthermore, to minimize the bias in the original expression profiles and to better understand the

functional terms and pathways of the DEGs in the EPCs influenced by type 2 DM, the array data were analyzed in three steps. First, DEGs between healthy subjects and diabetes patients were determined from the microarrays using the classical bioinformatics analysis tool, limma. Subsequent enrichment analysis was performed by DAVID. Hub genes and associated modules based on PPI network were selected using cytohubba in Cytoscape. Thereafter, the predicted hub genes were validated by RT-qPCR of EPCs isolated from patients at Sir Run Run Shaw Hospital (Hangzhou, China). IL8 and CXCL1 were found to exhibit the greatest variation in the EPCs of patients. To identify the genes that were co-expressed with IL8 and CXCL1, WGCNA, which is an algorithm recently wide used in microarray data analyses to identify the modules with similar expression patterns, was applied in the current study (57). After selecting the appropriate soft-threshold power β, the pink module was indicated to incorporate IL8 and CXCL1.

The present results indicated that inflammation may be an important mechanism underlying the difference in function and number of EPCs isolated from healthy individuals and diabetic patients. EPCs have long been regarded as a potential therapy to accelerate re-endothelialization of impaired vessels. Elevated IL8 expression is also thought to be involved in the repair progress (18). Compared to the successful preclinical results, the results of clinical trials on the efficacy parameters are conflicting (58,59). Certain studies indicated no beneficial
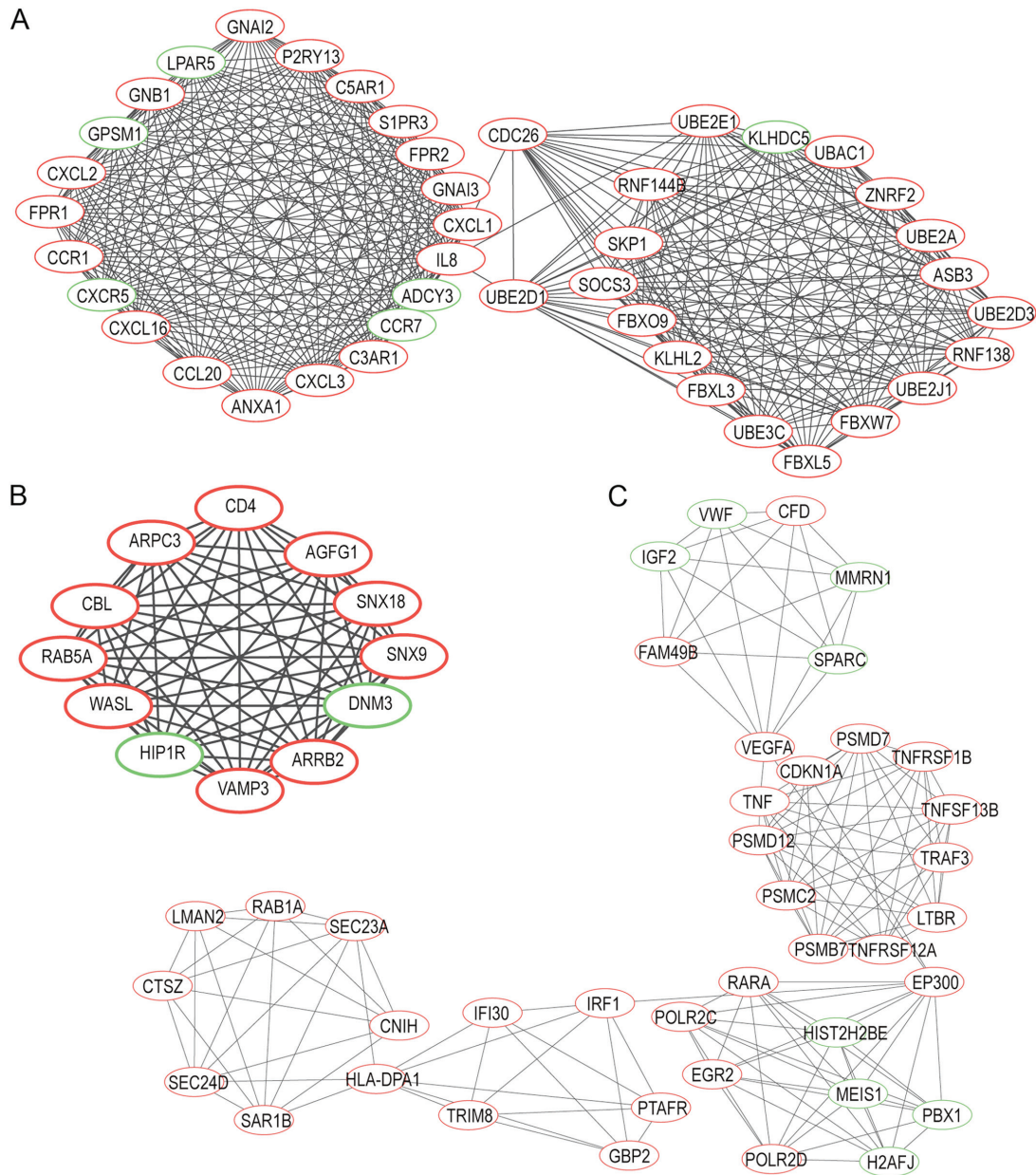
Figure 4. Results of the subnet analysis of the PPI network. Three modules, (A) Module a, (B) module b and (C) module c were extracted from the PPI network in Cytoscape using the MCODE plugin. Circles with green and red margins represent the downregulated and upregulated genes, respectively. Lines indicate PPIs. PPI, protein-protein interaction.
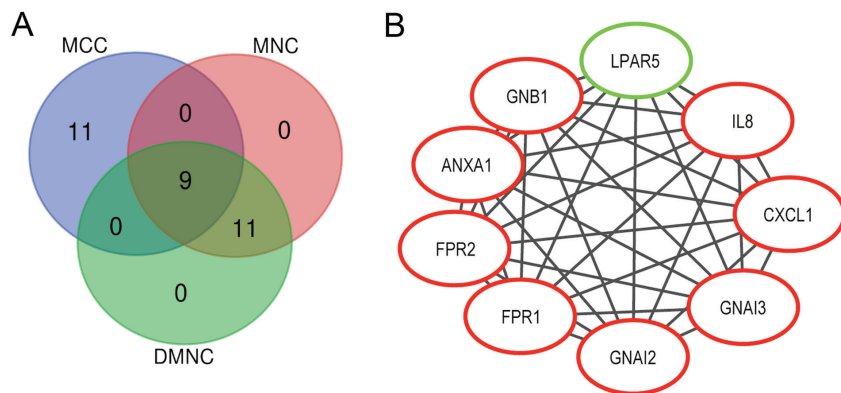


Figure 5. Obtainment of 9 hub genes, construction of PPI network and enrichment analysis of hub genes. (A) Overlapping DEGs among the three topological cytoHubba methods including, maximal clique centrality, maximum neighborhood component and density of maximum neighborhood component. (B) PPI interaction network of DEGs determined by STRING and visualized in Cytoscape. Lines indicate protein-protein interactions. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein-protein interaction; DEG, differentially expressed gene.

Table III. Information on the 9 hub genes.

| Gene | Log$_2$ FC | P-value | Degree | Expression alteration |
|---|---|---|---|---|
| IL8 | 3.89158 | $9.02 \times 10^{-6}$ | 43 | Upregulated |
| FPR1 | 3.04300 | $4.82 \times 10^{-5}$ | 22 | Upregulated |
| CXCL1 | 2.49797 | $3.30 \times 10^{-4}$ | 22 | Upregulated |
| GNAI3 | 2.09317 | $1.38 \times 10^{-3}$ | 24 | Upregulated |
| FPR2 | 2.83408 | $2.84 \times 10^{-3}$ | 26 | Upregulated |
| GNAI2 | 1.01805 | $3.26 \times 10^{-3}$ | 22 | Upregulated |
| ANXA1 | 2.18436 | $4.56 \times 10^{-3}$ | 26 | Upregulated |
| GNB1 | 1.12879 | $1.48 \times 10^{-2}$ | 33 | Upregulated |
| LPAR5 | -1.20699 | $2.81 \times 10^{-2}$ | 25 | Downregulated |

FC, fold change; IL8, interleukin 8; FPR1, formyl peptide receptor 1; CXCL1, C-X-C motif chemokine ligand 1; GNAI3, G protein subunit alpha I3; FPR2, formyl peptide receptor 2; GNAI2, G protein subunit alpha 2; ANXA1, annexin A1; GNB1, G protein subunit beta 1; LPAR5, lysophosphatidic acid receptor 5.
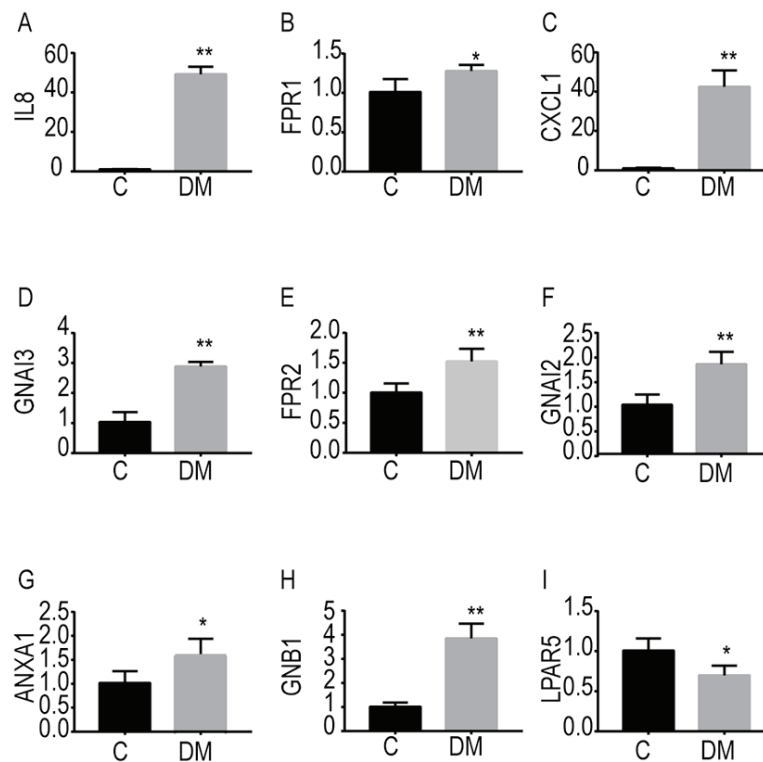


Figure 6. Validation of microarray data by quantitative PCR. The expression of the top 9 hub genes identified from the microarray in detected in endothelial progenitor cells from patients with DM and healthy controls is provided in bar graphs. (A) IL8, (B) FPR1, (C) CXCL1, (D) GNAI3, (E) FPR2, (F) GNAI2, (G) ANXA1, (H) GNB1, (I) LPAR5. Values are expressed as the mean ± standard deviation (n=4 in each experiment). *P<0.05, **P<0.01 vs. control group. DM, diabetes mellitus; C, control; IL8, interleukin 8; FPR1, formyl peptide receptor 1; CXCL1, C-X-C motif chemokine ligand 1; GNAI3, G protein subunit alpha I3; FPR2, formyl peptide receptor 2; GNAI2, G protein subunit alpha 2; ANXA1, annexin A1; GNB1, G protein subunit beta 1; LPAR5, lysophosphatidic acid receptor 5.

effect of autologous EPC transplantation in diabetic patients with critical limb ischemia (58). The controversy may be attributed to the heterogeneity of survival and different sources of EPCs in these trials (7,58). The present results revealed that inflammatory responses and associated genes may be partially involved in the abnormal function and number of EPCs in patients with diabetes.

Taken together, the present study identified DEGs in the EPCs of patients with type 2 DM vs. healthy subjects, which may be involved in the changes in the number and function of EPCs in patients with type 2 DM. Hub genes (IL8, FPR1, CXCL1, GNAI3, FPR2, GNAI2, ANXA1, GNB1 and LPAR5) extracted from the PPI network may be helpful as potential diagnostic and prognostic biomarkers of disordered EPCs. In addition,
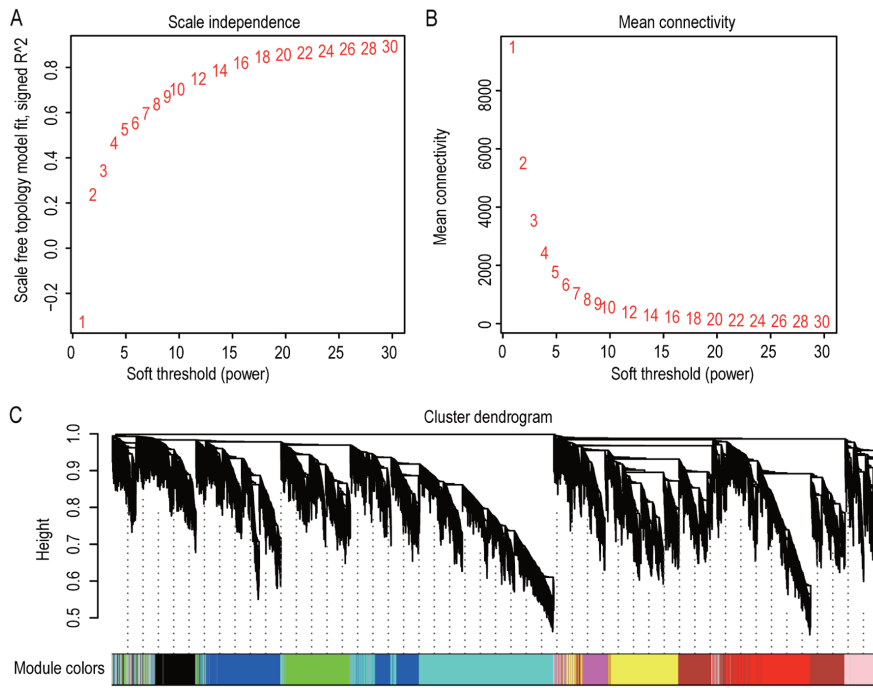
Figure 7. Appropriate soft-threshold power selection and construction of the cluster dendrogram. (A) Analysis of the scale-free fit index for various soft-threshold powers. (B) Analysis of the mean connectivity for various soft-threshold powers. (C) The co-expression modules constructed by weighted gene co-expression network analysis and visualized in the cluster dendrogram. Each colored row represents a color-coded module that contains a group of highly connected genes.
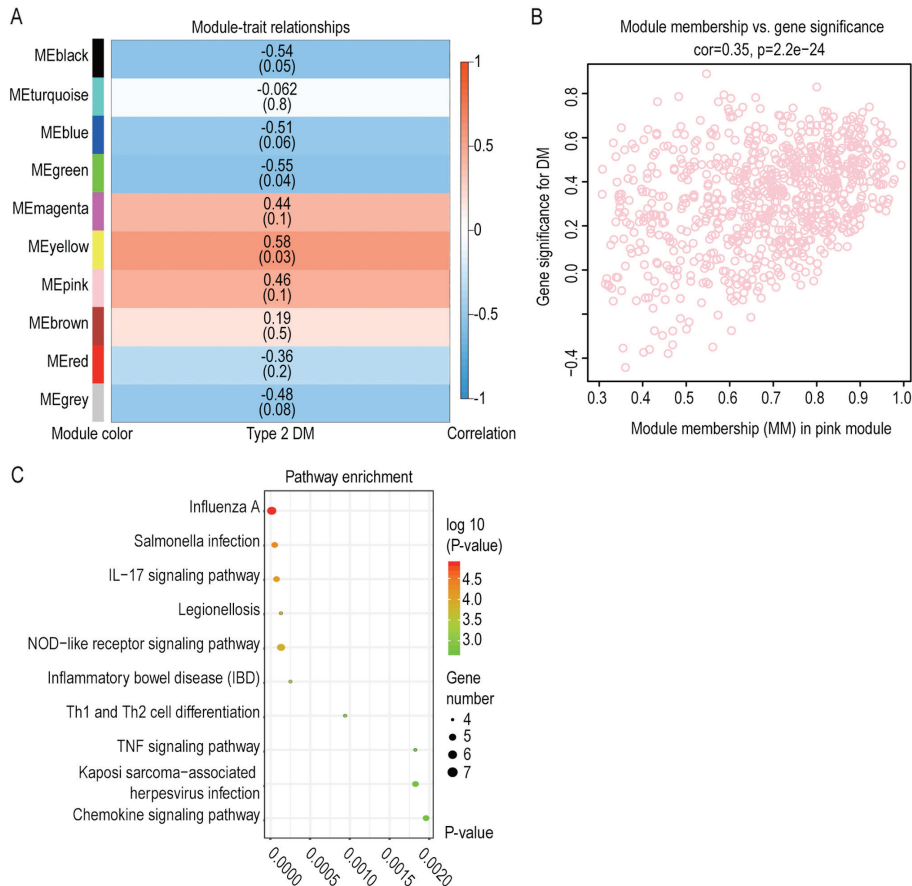


Figure 8. Network analysis of gene expression in type 2 DM and investigation of the module containing IL8 and CXCL1. (A) Module-trait associations. Each row corresponds to an ME, containing the corresponding correlation and P-value. The column represents the clinical trait of type 2 DM. The table was colored according to the correlation between ME and clinical traits. (B) The MM vs. GS plot for the pink module indicated that MM and GS are highly correlated. (C) Scatterplot of enriched Kyoto Encyclopedia of Genes and Genomes pathways for the pink module. The size and color of the dots represent the gene number and the range of P-values, respectively. ME, module eigengene; GS, gene significance; DM, diabetes mellitus; MM, module membership; IL8, interleukin 8; CXCL1, CXC motif chemokine ligand 1.

inflammatory responses are critical for the reduced number and hypofunction of EPCs isolated from diabetic patients.

The present study has several limitations that should be pointed out. First, it was not possible to conclude whether changes in gene expression were the cause or consequence of the complex physiological environment in patients with type 2 DM. Furthermore, the sample size of EPCs was relatively small (n=4 for each group). Third, due to the lack of a sorting machine, the EPCs in the present study were isolated and cultured in culture dishes. EPCs isolated from the patients in Sir Run Run Shaw Hospital were cultured *in vitro* while CD34+ cells were directly sorted in GSE43950. Therefore, the fold changes of identified hub genes may not fit well with the predicted outcomes of microarrays in GSE43950. Finally, as the clinical traits in the dataset GSE43950 were limited, the exact clinical characteristics of these patients were not known. In the current, the probes were annotated using the platform annotation file, which was last updated in July 09, 2014 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL10379). Therefore, some proteins discovered in the past few years may not be identified in the present study.

## Availability of data and materials

The datasets used and/or analyzed in the present study are available from the corresponding author on reasonable request.

## Authors' contributions

GF and YZ were involved in the conception and design of the study. ZS and QC acquired the data and prepared the manuscript. HY, ZM, XB, XL, MW, CJ and DWL analyzed and interpreted the data and performed the statistical analysis. All authors read and approved the final manuscript.

## Ethics approval and consent to participate

The study was approved by the Ethics Committee of Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University (Hangzhou, China). All of the patients provided written informed consent to participate in the study.

## Patient consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Shi Y and Vanhoutte PM: Macro- and microvascular endothelial dysfunction in diabetes. J Diabetes 9: 434-449, 2017.
2. Yang C, Kelaini S, Caines R and Margariti A: RBPs play important roles in vascular endothelial dysfunction under diabetic conditions. Front Physiol 9: 1310, 2018.
3. Zhang HN, Xu QQ, Thakur A, Alfred MO, Chakraborty M, Ghosh A and Yu XB: Endothelial dysfunction in diabetes and hypertension: Role of microRNAs and long non-coding RNAs. Life Sci 213: 258-268, 2018.
4. Giacco F and Brownlee M: Oxidative stress and diabetic complications. Circ Res 107: 1058-1070, 2010.
5. Zimmet P, Alberti KG and Shaw J: Global and societal implications of the diabetes epidemic. Nature 414: 782-787, 2001.
6. Vijan S: In the clinic. Type 2 diabetes. Ann Intern Med 152: ITC31-15; quiz ITC316, 2010.
7. Pyšná A, Bém R, Němcová A, Fejfarová V, Jirkovská A, Hazdrová J, Jude EB and Dubský M: Endothelial progenitor cells biology in diabetes mellitus and peripheral arterial disease and their therapeutic potential. Stem Cell Rev 15: 157-165, 2019.
8. Asahara T, Murohara T, Sullivan A, Silver M, van der Zee R, Li T, Witzenbichler B, Schatteman G and Isner JM: Isolation of putative progenitor endothelial cells for angiogenesis. Science 275: 964-967, 1997.
9. Asahara T and Kawamoto A: Endothelial progenitor cells for postnatal vasculogenesis. Am J Physiol Cell Physiol 287: C572-579, 2004.
10. Dubsky M, Jirkovska A, Bem R, Nemcova A, Fejfarova V and Jude EB: Cell therapy of critical limb ischemia in diabetic patients-State of art. Diabetes Res Clin Pract 126: 263-271, 2017.
11. Amann B, Luedemann C, Ratei R and Schmidt-Lucke JA: Autologous bone marrow cell transplantation increases leg perfusion and reduces amputations in patients with advanced critical limb ischemia due to peripheral artery disease. Cell Transplant 18: 371-380, 2009.
12. Liao YF, Chen LL, Zeng TS, Li YM, Fan Yu, Hu LJ and Ling Yue: Number of circulating endothelial progenitor cells as a marker of vascular endothelial function for type 2 diabetes. Vasc Med 15: 279-285, 2010.
13. Sibal L, Aldibbiat A, Agarwal SC, Mitchell G, Oates C, Razvi S, Weaver JU, Shaw JA and Home PD: Circulating endothelial progenitor cells, endothelial function, carotid intima-media thickness and circulating markers of endothelial dysfunction in people with type 1 diabetes without macrovascular disease or microalbuminuria. Diabetologia 52: 1464-1473, 2009.
14. Tepper OM, Galiano RD, Capla JM, Kalka C, Gagne PJ, Jacobowitz GR, Levine JP and Gurtner GC: Human endothelial progenitor cells from type II diabetics exhibit impaired proliferation, adhesion, and incorporation into vascular structures. Circulation 106: 2781-2786, 2002.
15. Yiu KH and Tse HF: Specific role of impaired glucose metabolism and diabetes mellitus in endothelial progenitor cell characteristics and function. Arterioscler Thromb Vasc Biol 34: 1136-1143, 2014.
16. Chen LL, Liao YF, Zeng TS, Yu F, Li HQ and Feng Y: Effects of metformin plus gliclazide compared with metformin alone on circulating endothelial progenitor cell in type 2 diabetic patients. Endocrine 38: 266-275, 2010.
17. Chang TT, Lin LY and Chen JW: Inhibition of macrophage inflammatory protein-1β improves endothelial progenitor cell function and ischemia-induced angiogenesis in diabetes. Angiogenesis 22: 53-65, 2019.
18. Ambasta RK, Kohli H and Kumar P: Multiple therapeutic effect of endothelial progenitor cell regulated by drugs in diabetes and diabetes related disorder. J Transl Med 15: 185, 2017.
19. Cheng CC, Chang SJ, Chueh YN, Huang TS, Huang PH, Cheng SM, Tsai TN, Chen JW and Wang HW: Distinct angiogenesis roles and surface markers of early and late endothelial progenitor cells revealed by functional group analyses. BMC Genomics 14: 182, 2013.

20. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, *et al*: Bioconductor: Open software development for computational biology and bioinformatics. Genome Biol 5: R80, 2004.

21. Reiner A, Yekutieli D and Benjamini Y: Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics 19: 368-375, 2003.

22. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC and Lempicki RA: DAVID: Database for annotation, visualization, and integrated discovery. Genome Biol 4: P3, 2003.

23. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T: Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 13: 2498-2504, 2003.

24. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, *et al*: The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res 45: D362-D368, 2017.

25. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT and Lin CY: CytoHubba: Identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 4 (Suppl 8): S11, 2014.

26. Liu J, Yu Z, Sun M, Liu Q, Wei M and Gao H: Identification of cancer/testis antigen 2 gene as a potential hepatocellular carcinoma therapeutic target by hub gene screening with topological analysis. Oncol Lett 18: 4778-4788, 2019.

27. Langfelder P and Horvath S: WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics 9: 559, 2008.

28. Yip AM and Horvath S: Gene network interconnectedness and the generalized topological overlap measure. BMC Bioinformatics 8: 22, 2007.

29. Ravasz E, Somera AL, Mongru DA, Oltvai ZN and Barabasi AL: Hierarchical organization of modularity in metabolic networks. Science 297: 1551-1555, 2002.

30. Shi K, Bing ZT, Cao GQ, Guo L, Cao YN, Jiang HO and Zhang MX: Identify the signature genes for diagnose of uveal melanoma by weight gene co-expression network analysis. Int J Ophthalmol 8: 269-274, 2015.

31. Wang T, Zheng X, Li R, Liu X, Wu J, Zhong X, Zhang W, Liu Y, He X, Liu W, *et al*: Integrated bioinformatic analysis reveals YWHAB as a novel diagnostic biomarker for idiopathic pulmonary arterial hypertension. J Cell Physiol 234: 6449-6462 2019.

32. Huang da W, Sherman BT and Lempicki RA: Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37: 1-13, 2009.

33. American Diabetes A: Diagnosis and classification of diabetes mellitus. Diabetes Care 37 (Suppl 1): S81-S90, 2014.

34. Del Coco L, Vergara D, De Matteis S, Mensà E, Sabbatinelli J, Prattichizzo F, Bonfigli AR, Storci G, Bravaccini S, Pirini F, *et al*: NMR-based metabolomic approach tracks potential serum biomarkers of disease progression in patients with type 2 diabetes mellitus. J Clin Med 8: E720, 2019.

35. Zhao Y, Qiu F, Xu S, Yu L and Fu G: Thymosin β4 activates integrin-linked kinase and decreases endothelial progenitor cells apoptosis under serum deprivation. J Cell Physiol 226: 2798-2806, 2011.

36. Mason MJ, Fan G, Plath K, Zhou Q and Horvath S: Signed weighted gene co-expression network analysis of transcriptional regulation in murine embryonic stem cells. BMC Genomics 10: 327, 2009.

37. Shen Z, Chen Q, Jin T, Wang M, Ying H, Lu J, Wang M, Zhang W, Qiu F, Jin C, *et al*: Theaflavin 3,3'-digallate reverses the downregulation of connexin 43 and autophagy induced by high glucose via AMPK activation in cardiomyocytes. J Cell Physiol 234: 17999-18016, 2019.

38. Raval Z and Losordo DW: Cell therapy of peripheral arterial disease: From experimental findings to clinical trials. Circ Res 112: 1288-1302, 2013.

39. Egan CG, Lavery R, Caporali F, Fondelli C, Laghi-Pasini F, Dotta F and Sorrentino V: Generalised reduction of putative endothelial progenitors and CXCR4-positive peripheral blood cells in type 2 diabetes. Diabetologia 51: 1296-1305, 2008.

40. Lois N, McCarter RV, O'Neill C, Medina RJ and Stitt AW: Endothelial progenitor cells in diabetic retinopathy. Front Endocrinol (Lausanne) 5: 44, 2014.

41. Ma C, Wang J, Liu H, Chen Y, Ma X, Chen S, Chen Y, Bihl JI and Yang YI: Moderate exercise enhances endothelial progenitor cell exosomes release and function. Med Sci Sports Exerc 50: 2024-2032, 2018.

42. Sun J, Zhang Z, Ma T, Yang Z, Zhang J, Liu X, Lu D, Shen Z, Yang J and Meng Q: Endothelial progenitor cell-derived exosomes, loaded with miR-126, promoted deep vein thrombosis resolution and recanalization. Stem Cell Res Ther 9: 223, 2018.

43. Zhang J, Chen C, Hu B, Niu X, Liu X, Zhang G, Zhang C, Li Q and Wang Y: Exosomes derived from human endothelial progenitor cells accelerate cutaneous wound healing by promoting angiogenesis through Erk1/2 signaling. Int J Biol Sci 12: 1472-1487, 2016.

44. Ke X, Yang D, Liang J, Wang X, Wu S, Wang X and Hu C: Human endothelial progenitor cell-derived exosomes increase proliferation and angiogenesis in cardiac fibroblasts by promoting the mesenchymal-endothelial transition and reducing high mobility group Box 1 protein B1 expression. DNA Cell Biol 36: 1018-1028, 2017.

45. Sikora J, Smycz-Kubanska M, Mielczarek-Palacz A and Kondera-Anasz Z: Abnormal peritoneal regulation of chemokine activation-the role of IL-8 in pathogenesis of endometriosis. Am J Reprod Immunol 77: 2017.

46. Remick DG: Interleukin-8. Crit Care Med 33 (12 Suppl): S466-S467, 2005.

47. Cai W, Qiu C, Zhang H, Chen X, Zhang X, Meng Q and Wei J: Detection of circulating natural antibodies to inflammatory cytokines in type-2 diabetes and clinical significance. J Inflamm (Lond) 14: 24, 2017.

48. Barcelos LS, Duplaa C, Kränkel N, Graiani G, Invernici G, Katare R, Siragusa M, Meloni M, Campesi I, Monica M, *et al*: Human CD133+ progenitor cells promote the healing of diabetic ischemic ulcers by paracrine stimulation of angiogenesis and activation of wnt signaling. Circ Res 104: 1095-1102, 2009.

49. Chen C, Lin X, Wang J, Tang G, Mu Z, Chen X, Xu J, Wang Y, Zhang Z and Yang GY: Effect of HMGB1 on the paracrine action of EPC promotes post-ischemic neovascularization in mice. Stem Cells 32: 2679-2689, 2014.

50. Liang X, Ding Y, Zhang Y, Tse HF and Lian Q: Paracrine mechanisms of mesenchymal stem cell-based therapy: Current status and perspectives. Cell Transplant 23: 1045-1059, 2014.

51. Montane J, Cadavez L and Novials A: Stress and the inflammatory process: A major cause of pancreatic cell death in type 2 diabetes. Diabetes Metab Syndr Obes 7: 25-34, 2014.

52. Belo AV, Leles F, Barcelos LS, Ferreira MA, Bakhle YS, Teixeira MM and Andrade SP: Murine chemokine CXCL2/KC is a surrogate marker for angiogenic activity in the inflammatory granulation tissue. Microcirculation 12: 597-606, 2005.

53. Herlea-Pana O, Yao L, Heuser-Baker J, Wang Q, Wang Q, Georgescu C, Zou MH and Barlic-Dicen J: Chemokine receptors CXCR2 and CX3CR1 differentially regulate functional responses of bone-marrow endothelial progenitors during atherosclerotic plaque regression. Cardiovasc Res 106: 324-337, 2015.

54. Chavakis E, Urbich C and Dimmeler S: Homing and engraftment of progenitor cells: A prerequisite for cell therapy. J Mol Cell Cardiol 45: 514-522, 2008.

55. Eyileten C, Kaplon-Cieslicka A, Mirowska-Guzel D, Malek L and Postula M: Antidiabetic effect of brain-derived neurotrophic factor and its association with inflammation in type 2 diabetes mellitus. J Diabetes Res 2017: 2823671, 2017.

56. Domingueti CP, Dusse LM, Carvalho MD, de Sousa LP, Gomes KB and Fernandes AP: Diabetes mellitus: The linkage between oxidative stress, inflammation, hypercoagulability and vascular complications. J Diabetes Complications 30: 738-745, 2016.

57. Wan Q, Tang J, Han Y and Wang D: Co-expression modules construction by WGCNA and identify potential prognostic markers of uveal melanoma. Exp Eye Res 166: 13-20, 2018.

58. Wils J, Favre J and Bellien J: Modulating putative endothelial progenitor cells for the treatment of endothelial dysfunction and cardiovascular complications in diabetes. Pharmacol Ther 170: 98-115, 2017.

59. Teraa M, Sprengers RW, Schutgens RE, Slaper-Cortenbach IC, van der Graaf Y, Algra A, van der Tweel I, Doevendans PA, Mali WP, Moll FL and Verhaar MC: Effect of repetitive intra-arterial infusion of bone marrow mononuclear cells in patients with no-option limb ischemia: The randomized, double-blind, placebo-controlled rejuvenating endothelial progenitor cells via transcutaneous intra-arterial supplementation (JUVENTAS) trial. Circulation 131: 851-860, 2015.