Induced Expression of VEGFC, ANGPT, and EFNB2 and Their Receptors Characterizes Neovascularization in Proliferative Diabetic Retinopathy

Yaping Li,1,2 Dong Chen,3,4 Luguo Sun,1 Yannan Wu,1 Ying Zou,2 Chen Liang,1 Yongli Bao,1 Jingwen Yi,1 Yu Zhang,3 Jing Hou,3 Zhen Li,3 Fengyun Yu,1 Yanxin Huang,1 Chunlei Yu,1 Lei Liu,1 Zaoxia Liu,2 Yi Zhang,3,4 and Yuxin Li1

1National Engineering Laboratory for Druggable Gene and Protein Screening, Northeast Normal University, Changchun, China
2Department of Ophthalmology, Second Hospital, Jilin University, Changchun, China
3Center for Genome Analysis, ABLife, Inc., Wuhan, People’s Republic of China
4Laboratory for Genome Regulation and Human Health, ABLife, Inc., Wuhan, People’s Republic of China

Correspondence: Yuxin Li, National Engineering Laboratory for Druggable Gene and Protein Screening, Northeast Normal University, Changchun 130117, China; liyx486@nenu.edu.cn.

Yi Zhang, Center for Genome Analysis, ABLife, Inc., 388 Gaxin 2nd Road, East Lake High-Tech Development Zone, Wuhan 430075, People’s Republic of China; yizhang@ablife.cc.

Zaoxia Liu, Department of Ophthalmology, Second Hospital, Jilin University, Changchun 130021, China; liuli15898@sina.com.

Lei Liu, National Engineering Laboratory for Druggable Gene and Protein Screening, Northeast Normal University, 2555 Jingyue Avenue, Changchun 130117, China; liu905@nenu.edu.cn.

Submitted: January 29, 2019
Accepted: August 16, 2019


PURPOSE. To investigate whole transcriptional differences between proliferative diabetic retinopathy (PDR) neovascular membranes (NVMs) and retinas, and the regulatory genes participating in retinal neovascularization in PDR.

METHODS. We used high-throughput sequencing technology to capture the whole-genome gene expression levels of all participants, including 23 patients with PDR or branch retinal vein occlusion (BRVO), 3 normal retinal samples, and 2 retinal samples from type II diabetic (T2D) eyes by donation, followed by analyses of expression patterns using bioinformatics methods, then validation of the data by in situ hybridization and Western blotting.

RESULTS. We showed that transcriptional profiles of the NVMs were distinct from those of the retinas. Angiogenesis growth factors VEGFA, ANGPT1, ANGPT2, and EFNB2, and their receptors FLT4, TIE1, TIE2, and EPHB4, respectively, were overexpressed. Expression of VEGFA was highly upregulated in T2D retina, but low in the NVMs, while angiogenesis transcription factors, including ETS1 and ERG, were coordinately upregulated in NVMs.

CONCLUSIONS. This study described a PDR neovascularization model in which pathological retinal-secreted vascular endothelial growth factor A (VEGFA) enhanced the expression of a set of angiogenesis transcription factors and growth factors, to cooperatively induce the retinal neovascularization. Based on these results, novel potential therapeutic targets and biomarkers for PDR treatment and diagnosis are suggested.

Keywords: proliferative diabetic retinopathy, transcriptome profile, VEGF, ETS transcription factors, neovascularization

Diabetes mellitus is mainly caused by impaired insulin production or action, and does significant long-term damage to many organs, including eyes.1,2 Diabetics have high risks of macrovascular and microvascular complications induced by oxidative stress.3 Diabetic retinopathy (DR) is one of the most common microvascular complications of diabetes.4,5 The advanced stage of DR, proliferative diabetic retinopathy (PDR) with retinal neovascularization, is closely related to vision loss.6–8 Investigations of the molecular mechanisms involved in retinal neovascularization have shown that vascular endothelial growth factor A (VEGFA), also known as VEGF, is a master causative factor for retinal neovascularization.4,9 VEGF was identified as a diffusible mitogenic heparin-binding protein in the 1980s,10,11 belonging to a gene family that includes VEGF, VEGFC, VEGFD, and placental growth factor (PIGF).12 These proteins, as well as other vascular endothelium-specific growth factors, play differential and cooperative roles in promoting and regulating angiogenesis or neovascularization, by interacting with their receptors and then stimulating the following processes.1,2–17

During the pathogenesis of PDR, it has been widely reported that VEGFA-VEGFR2 (VEGF receptor 2, also named KDR or Flk-1) mediates endothelial cell mitogenesis and vascular permeability, via VEGFR2-mediated signaling cascades.9,16 VEGFA could be increased by altered glucose levels, insulin, and other growth factors.16–20 The secreted VEGFA level from retinal pigment...
epithelial cells was increased in the ocular fluid of DR patients and in the retinal ischemic monkey model21–24 and its expression level in DR patients has been correlated with disease progression.25 Consequently, anti-VEGFA therapy has been developed, with the first inhibitor approved in 2004.9,32 The ocular anti-VEGFA therapies have shown clear efficacies in clinical trials25,26. In addition to VEGFA, other proteins of the VEGF family also have important roles in angiogenesis. Competitive and collaborative relationships among VEGFs and their receptors constitute an important regulatory network during angiogenesis.27–31 In human retinal pigment epithelial cells, VEGFA promotes the expression of VEGFC and VEGFR3.31 Therefore, the roles of VEGF family members and angiopoietin and ephrin family members,32–34 as well as their potential cooperation during retinal neovascularization in humans, should be further studied with advanced technologies.35

An omics study on gene expression alterations of fibrovascular membranes (FVMs) between PDR patients and normal retinal controls has revealed 91 genes that are preferentially expressed in FVMs, when compared with normal retinas.54 In the present study, we obtained qualified cDNA libraries from the small neovascular membranes (NVMs) of patients and retinas from normal or type II diabetic (T2D) individuals, and then sequenced the qualified cDNA libraries (RNA-Seq). The transcriptome profiles of the NVMs were distinct from retinal membranes. We also discovered that a number of genes belonging to the master angiogenesis transcription factor (TF) ETSI were systematically upregulated during retinal neovascularization. This expression network suggested distinct VEGF regulatory programs, involving VEGFA acting as a trigger of the program switch.

**Materials and Methods**

**Study Approval, Sample Collection, Treatment, and RNA Extraction**

This study was approved by the Ethics Committee of the Second Hospital of Jilin University, and followed the tenets of the Declaration of Helsinki. Informed consent was obtained from the participants after explanation of the nature and possible consequences of the study. NVMs were surgically obtained from 23 eyes of patients with PDR or branch retinal vein occlusion (BRVO) during vitrectomies, which were performed at the Second Hospital of Jilin University. Three normal and two T2D retinas were from donated eyes. Surgical indications included persistent vitreous hemorrhage, NVM formation, and NVM traction retinal detachment. The surgical specimen for each case was frozen in liquid nitrogen immediately after removal.

After homogenizing the sample with TRIzol reagent, the samples were incubated at room temperature to permit complete dissociation of the nucleoprotein complex. The homogenate was then centrifuged at 4°C, and the supernatant was collected and added to chloroform. The solution was shaken vigorously and centrifuged at 4°C to produce a clear upper aqueous layer (containing RNA), which was precipitated from the aqueous layer with isopropanol. The precipitated RNA was washed with 75% ethanol to remove impurities, and then resuspended with diethyl pyrocarbonate (DEPC)-treated water for use in downstream applications.

**RNA-Seq Library Construction**

The quality, integrity, and quantity of the purified RNAs were determined using a 2100 Bioanalyzer system (Agilent Technologies, Santa Clara, CA, USA). For each sample, 50 ng total RNA was used for the library preparation. Ribosomal RNAs were depleted with a Ribo-Zero rRNA depletion kit for humans (Epicentre, San Diego, CA, USA) before being used for directional RNA-Seq library preparation. Purified RNA was fragmented, and then reverse transcribed with random primers containing adapter sequences and a randomized hexamer. The cDNAs were then synthesized with terminal-tagging oligo-cDNA using the ScriptSeq v2 RNA-Seq Library Preparation Kit (Epicentre). The cDNAs were purified and amplified. PCR products corresponding to 300 to 500 base pair (bp) were purified, quantified, and stored at −80°C until sequenced.

The libraries were prepared following the manufacturer’s instructions and applied to an Illumina NextSeq 500 platform for 151 nucleotide pair-end sequencing (ABLife, Inc., Wuhan, China). The RNA-Seq data have been deposited in the GEO database with accession number GSE102485.

**Data Processing and Differentially Expressed Gene (DEG) Analysis**

Cutadapt55 was used to trim adapters (with a 10% error) and low-quality bases (less than 16) of raw reads. Reads less than 16 nucleotides in both ends were discarded. We used TopHat256 software to align the filtered reads to the human reference genome (GRCh38) with four read mismatches using an end-to-end style. After alignment, total aligned reads for each gene were counted and saved as the initial expression level. Due to the nonuniform normalization of FPKM value,57 we used a tag-wise dispersion mode with the estimated size factors method58 to normalize the gene transcriptional levels. After normalization, we selected the mRNA and long noncoding RNA (lncRNA) genes whose average transcriptional levels were no less than 0.1 as the expressed genes.

We then calculated the Z-score for each gene across all samples. To retrieve genes showing diverse expression levels, we set the mean to Z-score 0.7 as the threshold to filter the genes with mediocre expression in all samples. DEGs were identified using the edgeR59 package from R software (R Project for Statistical Computing, Vienna, Austria). The false discovery rate was < 0.05 and log2 fold change > 1.

**Correlation Network Analysis**

To obtain the expression modules and distinguish genes from a union set by expression features, we used weighted gene coexpression network analysis (WGCNA).40 The normalized expression file of selected genes was the input. The output was the gene modules according to their expression patterns. For each gene module, eigengene was chosen to represent the expression pattern.

We used Pearson’s correlation coefficient (PCC) analysis to investigate the coexpressed genes with VEGFA. We set absolute (PCC) ≥ 0.5 and P values ≤ 0.05 as the thresholds to filter the coexpressed genes. We then obtained the human TFs from AnimalTFDB 2.0,41 and selected the TFs that were coexpressed with VEGFA from the coexpressed genes.

**In Situ Hybridization (ISH) Experiments**

ISH was performed using a digoxigenin mRNA detection kit with some modifications (Boster Biological Technology, Wuhan, China). Single-stranded sense and antisense digoxigenin-labeled corresponding mRNA probes were generated by in vitro cDNA transcription with T3 or T7 RNA polymerase. The tissues were fixed in 4% paraformaldehyde (containing 1/1000 DEPC) for 1 hour, then dehydrated in ethanol gradients of 70%, 80%, 90%, 95%, and 100%, and incubated with xylene and embedded in paraffin. Next, 4-μm-thick paraffin sections were baked face-up for 2 hours at 50°C and deparaffinized. The tissue sections were then treated with proteinase K (5 mg/ml; Sigma-Aldrich Corp.,...
St. Louis, MO, USA) in 10 mM Tris-HCl (pH 8.0), 1 mM EDTA at 37°C for 15 minutes before treatment with 3% H2O2 for 8 minutes, then rinsed with phosphate-buffered saline and postfixed in 1% paraformaldehyde containing 1/1000 DEPC at room temperature for 10 minutes. Sections were then prehybridized for 4 hours at 40°C with 50 μL prehybridization buffer. After the prehybridization, hybridization was performed in buffer containing prehybridization fluid and 1.0 mg L-1 either antisense or sense digoxigenin-labeled mRNA probes at 40°C overnight. Next, the sections were washed at 37°C for 30 minutes in a buffer containing 50% formamide and 2× SSC, followed by digestion with ribonuclease A (10 mg/mL; Sigma-Aldrich Corp.) at 37°C for 30 minutes. After being washed in 2× SSC at 37°C for 20 minutes, and twice in 0.2× SSC at 37°C for 30 minutes, blocking was performed in blocking buffer containing 1% bovine serum albumin for each section for 30 minutes at 37°C. After blocking buffer was removed, the sections were incubated with biotinylated mouse anti-digoxigenin antibody in a humidified chamber at 37°C for 60 minutes, then postrinsed and incubated with Strept Avidin Biotin-peroxidase Complex (SABC) solution for 30 minutes, and finally reacted with a solution containing biotinylated peroxidase for 30 minutes at room temperature. After the sections were rinsed, a coloring reaction was performed with a DAB Color Development Kit (Solarbio Science & Technology, Beijing, China). When the color reaction was visible, it was stopped by washing in buffer solution, and the sections were then counterstained with hematoxylin. Finally, the section was washed, dehydrated, and mounted.

Western Blotting

Total protein was extracted by cell lysis. The extracts were electroblotted onto a polyvinylidene difluoride (PVDF) membrane following separation by 12% SDS-PAGE. The PVDF membrane was incubated with a blocking solution of 5% skim milk for 1.5 hours at room temperature, followed by incubation with a specific antibody (VEGF-C, 22601-1-AP; Proteintech, Wuhan, China) overnight at 4°C and then with horseradish peroxidase–conjugated secondary antibody (SA00001-2; Proteintech) for 1 hour at room temperature. Finally, the blots were visualized using an ECL reagent (Amersham, Buckinghamshire, UK).

Statistical Analysis

To describe the biologically functional significance of selected gene sets, we used the hypergeometric test to calculate the enrichment of Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. Other statistical analysis was performed using R software (R Project for Statistical Computing) unless otherwise stated. The significance of differences was evaluated with either Student’s t-test, when only two groups were compared, or the hypergeometric test for a Venn diagram. No statistical methods were used to predetermine sample size. *P < 0.05; **P < 0.01; ***P < 0.001. Hierarchical clustering was performed by Cluster 3.0 (http://bonsai.hgc.jp/~mdehoon/software/cluster/software. htm, in the public domain) or the heat map function in R.

RESULTS

Transcriptome Profiles Distinguish NVMs From Retinal Membranes

In order to establish a link between the gene expression profiles in NVMs and PDR, we collected pathological tissues from retinal NVMs of patients that comprehensively included the major types of retinal neovascular disorders. Qualified transcriptome data were obtained from 23 NVM samples from patients, including 17 from T2D patients, 3 from T1D patients, and 3 from BRVO patients (Supplementary Table S1). T2D NVM samples observed at seven different regions were collected (Fig. 1A; Supplementary Table S1). This sampling roughly reflected the incidence of PDR in our hospital. Qualified transcriptome data were also obtained from three retinal membranes from normal individuals and two from T2D patients. A total of 1.43 billion reads were sequenced with an average of 47.51 million reads per sample. After adaptor and low-quality read filtering, approximately 40 million reads per sample were kept and aligned to the human genome (Supplementary Table S2). Similarly to a previous study,32 the sequencing results contained a large fraction of reads (19%-50%) within intronic regions (Supplementary Fig. S1A), and approximately 80% of the reads were not spliced (Supplementary Fig. S1B), indicating that a large number of nascent transcripts were captured. We obtained 25,920 expressed mRNA and lncRNA genes with normalized expressions ≥ 0.1 at least in one sample.

To obtain a global understanding of transcriptional dynamics from NVMs and retinas, we performed principal component analysis (PCA) based on the normalized expression level of all samples (Fig. 1B). NVMs and retinas could be clearly separated by the second principal component (PC2), indicating the presence of a large number of DEGs between NVMs and retinas. NVMs originating from different regions or pathological states could not be clearly separated because of the large individual variation reflected by the PC1 direction. We found that six NVMs (P_I_W , two P_II_VAs, P_II_W, P_BRVO, and P_II_S) were separated from other samples (Fig. 1B).

We also calculated PCCs for all samples. From the hierarchical clustering results, samples were clustered into three major groups (Supplementary Fig. S1C, blue frames). Group I (top left) contained 11 NVM samples and one retina sample; group II (middle) contained two normal and two T2D retinal samples; and group III (bottom right) contained 12 NVM samples. Group II could be treated as the retina group. Six proliferative samples (P_BRVO3, P_II_S2, P_II_VA1, P_II_W4, P_II_VA2, and P_I_W3) in group I showed low coefficients (R < 0.75) with most of the samples in group III (Supplementary Fig. S1C), consistent with the PCA results (Fig. 1B). The PCA and PCC analyses both showed that the transcriptional profile of NVMs was different from the retinal membranes, consistent with the fact that the NVM was essentially different from the retina.43

We then compared the DEGs between each type of NVM and normal retinas. DEGs in NVMs from the surrounding area of the eye (P_II_S) were more distinct from other groups, and those in the vascular arch and optic disc (P_II_VA_OD) displayed the largest DEG number (3322, Fig. 1C), among which 911 (27.4%) were unique to this group and the rest of DEGs were shared by at least one other group. A total of 1387 DEGs (32.35%) were shared by at least four groups, and 1601 DEGs (37.35%) were identified only from one specific group. A total of 153 DEGs were common in all NVMs versus retina comparisons (Fig. 1C). Only 19 of the 87 DEGs from Ishikawa et al.33 were not included in our DEG set (Fig. 1C), implying the high sensitivity of the RNA-seq method used in this study.

Expression of the VEGF Family and Two Other Vascular Endothelium-Specific Growth Factors in NVMs and Retinas

It has been reported that several proteins are major contributors to angiogenesis and vascular development, including five
proteins from the VEGF family, four proteins from the angiopoietin family, and at least one protein from the large ephrin family.\textsuperscript{15,16,44,45} By presenting the transcriptional levels of all these growth factors and their receptors, we found a significant decline of \textit{VEGFA} and its major receptor \textit{KDR} in the NVM group (Fig. 2A). This expression pattern suggested that the VEGFA pathway was repressed in NVM samples. Expression of \textit{VEGFB} did not show significant alterations between NVMs and diabetic retinas. In contrast, \textit{VEGFC} and its receptor gene, \textit{VEGFR-3 (FLT4)}, exhibited an increased expression pattern (Fig. 2A, left). Both angiopoietin genes and their receptors showed a consistently higher expression level in the NVMs than the diabetic retinas (Fig. 2A, middle). For the ephrin family, expression of \textit{EFNB2}, \textit{EFNA1}, and the receptor \textit{EPHB4} was increased in the NVMs (Fig. 2A, right).

To further validate the above observations, we included normal retina in the comparison group. We analyzed the relative expression patterns of some known neovascularization-related genes by sample-based normalized expression levels (Supplementary Fig. S2A). Notably, we found that \textit{VEGFA}, \textit{FLT1}, \textit{KDR}, and tumor neovascularization-related TFs \textit{HIF1A} and \textit{PROX1}\textsuperscript{46,47} were highly expressed in normal retinas. In the diabetic state, expression of \textit{VEGFA} increased dramatically in two T2D retinal samples, while expression of \textit{PROX1} was exclusively shut down. Remarkably, \textit{VEGFA} expression in all NVMs from T2D patients was drastically...

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Transcriptome profiles of PDR samples were separated from nonproliferative samples. (A) A fundus photograph shows the right eye of a patient with 20/200 visual acuity, no signs of hypertension. \textit{Black arrow} shows the proliferative membranes (pale yellow) around the vascular arcade. The \textit{white arrow} shows the blood vessel arcade. Other parts of the retina were labeled on the photograph that was related to Supplementary Table S1. (B) Principal component analysis of the sequenced samples revealed the separation between proliferative membrane and other samples. (C) An UpSet plot showing the overlapped differentially expressed genes between proliferative membrane and normal retina. Only groups with more than 10 genes are shown. DEGs from Ishikawa et al.\textsuperscript{17} were also included and are shown as a \textit{brown color}.}
\end{figure}
decreased compared with the T2D retinas (Supplementary Fig. S2A). In contrast, we found that VEGFC was at a low level in normal and T2D retinas, but at a high level in most of the NVM samples (Supplementary Fig. S2A). The differentially regulated expression of VEGFA and VEGFC was also plotted by their gene-based normalized expression levels (Supplementary Fig. S2B). To validate the induced expression of VEGFC, we performed RNA ISH in NVM samples (Fig. 2B). VEGF expression (dark brown) emerged at the margin or the inner membrane (arrows, Fig. 2B) of the tissues. The VEGFC signal was quite strong, similar to that of VEGFA (Fig. 2B, upper). We also validated the increased protein levels of VEGFC in T2D NVMs (Fig. 2C). These results together showed that expression of VEGFC was highly induced in the NVMs, while VEGFA was the most expressed in diabetic retinas.

**Genes Coexpressed With VEGFA and Their Network in the NVMs of PDR Patients**

We then characterized the VEGFA-related regulatory network by using a coexpression network approach. A correlation matrix between VEGFA and the other 25,919 genes was
generated by computing their PCCs. With criteria of $P$ value $\leq 0.01$ and absolute PCCs $\geq 0.5$, a total of 1615 genes were coexpressed with VEGFA, including 1421 positive and 194 negative pairs (Supplementary Table S3). Among these coexpressed genes, we found HMGA1, a TF that controls insulin receptor (INSR) gene transcription, was positively coexpressed with VEGFA (Fig. 3A), coincident with the previous finding that HMGA1 positively regulated the expression level of VEGFA. ANGPT2 was negatively coexpressed with VEGFA ($R = -0.729$; Fig. 3B), consistent with the upregulation of ANGPT2 in NVMs. In a similar manner as the VEGFA expression patterns, genes positively coexpressed with VEGFA were higher in diabetic and normal retinas than in NVMs (Supplementary Fig. S3A). In contrast, 194 genes that negatively coexpressed with VEGFA showed a reciprocal expression pattern (Supplementary Fig. S3A).

Based on the concept that coexpressed genes are functionally related, we further explored the VEGFA-regulated networks by functional enrichment analysis. The genes positively coexpressed with VEGFA were mostly enriched in the glycolysis and glucose metabolic processes (Fig. 3C; Supplementary Fig. S3B). Consistent with the role of hypoxia in increasing VEGFA expression, the HIF-1 signaling pathway (KEGG) was also enriched (Supplementary Fig. S3B). The biological process terms, including blood coagulation and platelet activation, as well as regulation of the apoptotic process, and vascular smooth muscle contraction were enriched by genes negatively coexpressed with VEGFA (Fig. 3D; Supplementary Fig. S3C). We also plotted the expression profiles of genes in VEGF signaling pathways including 19 genes. Expression of these genes was classified into two subgroups, with one group (GI) showing a constitutively higher expression in retina but lower expression in NVMs (Fig. 3E). Genes of another group (GII) were expressed at lower levels in normal retina and higher in T2D retina, but expressed at a much higher level in NVMs (Fig. 3E), suggesting that modulation of the VEGF signaling expression network could be a key event for initiating neovascularization.

**A Specific Group of ETS Transcription Factors Promote the Progression of PDR**

To further explore VEGFA-mediated gene expression networks in diabetic retinas and NVMs, we analyzed the TFs involved in VEGFA regulation. We found that $ETS1$ was negatively coexpressed with VEGFA. $ETS1$ has been recently shown to amplify the VEGF signal through acetylation to promote angiogenesis. Two ETSs ($ETS1$ and $ETS2$) and two ERGs ($ERG$ and $FLI1$) domain-containing genes showed a reverse expression pattern with VEGFA, that is, expressing at high levels in NVMs (Fig. 4A). Three of these four genes ($ETS1$, $ERG$, and $FLI1$) are known to promote angiogenesis by activating $VEGFR$ and $TIE1/TIE2$. We also found that VEGFA was negatively coexpressed with $TIE1$ and $TIE2$ (also known as $TEK$; $R = -0.45$). By adding VEGFC to this network, we observed the positive coexpression between VEGFC and ETS family members (Fig. 4A).

It is known that the ETS family contains both transcriptional activators and repressors for angiogenesis. To define the global function of ETS family genes in PDR, we searched for all ETS family genes in humans and studied their expression patterns. A clustering heat map of 28 expressed ETS family

---

**Figure 3.** Functions of genes coexpressed with VEGFA and their expression patterns. (A, B) The dot plot represents the coexpression relationship between HMGA1 (A), ANGPT2 (B), and VEGFA (C). The bar plot of enriched biological process terms by genes positively coexpressed with VEGFA. (D) The bar plot of enriched biological process terms by genes negatively coexpressed with VEGFA. (E) A heat map presentation of the genes belonging to the VEGF signaling pathway. Two group gene sets of the VEGF signaling pathway were labeled by blue and red colors.
genes exhibited two distinct expression clusters (Fig. 4B). Cluster I contained 16 genes that showed sporadic expression patterns. Among cluster I genes, *ELF2* showed a similar overexpression pattern as *VEGFA* in the T2D retinas. In contrast, cluster II contained 12 genes, which showed a low expression level in retina samples but globally overexpressed levels in NVM samples (Fig. 4B), implying their potential role in promoting PDR-related neovascularization.54,55

![Image](image_url)
We then obtained a coexpressed gene network of ETS1/2, EGR, and FLI1 to further explore their potential functions in NVMs. The coexpressed genes with these four TFs significantly overlapped. For example, 732 genes were coexpressed with at least three of these four TFs (26.65%, \( P = 2.30e-155 \); hypergeometric test; Fig. 4C). Consistently, functional analysis of ETS-coexpressed genes showed significant enrichment of angiogenesis and related biological processes (Fig. 4D; Supplementary Figs. S4A–C). To assess whether these coexpressed genes resulted from transcriptional activation by TF binding, we downloaded and analyzed the DNA binding profile data of ETS1 from human umbilical vein endothelial cells.\(^{52}\) We found 305 coexpressed genes with ETS1, which were also bound by ETS1 (\( P = 1.24 e-5 \); hypergeometric test; Fig. 4E), suggesting that a significant fraction of the coexpressed genes with these TFs involved direct transcriptional regulation.

We then characterized the coexpression relationship between these four TFs and vascular growth factors in Figure 2A. Most notably, besides VEGF, ANGPT2, EFNB2, and EFNA1, their receptors TIE1/TEK, VEGFR1/3, and EPHB4 were positively coexpressed with all these four TFs (Supplementary Table S4). This implied that the ETS and ERG domains containing TFs could participate in PDR progression by promoting angiogenesis. Based on the coexpression results, we propose that ETS TF family genes, especially ETS1/2, ERG, and FLI1, participated in the formation and progression of PDR and angiogenesis via positively regulating the expression of TIE, VEGFR, and ephrin family genes (Fig. 4F).

### Gene Expression in T2D Retina Is Reprogrammed

We identified gene sets associated with specific sample groups using the WGCNA method. We first calculated the mean Z-score for each gene (see Materials and Methods and Supplementary Fig. S5A) to select genes with a mean Z-score no less than 0.7 (12,293 genes), and then clustered these genes using the WGCNA method\(^{10}\) to obtain a consensus coexpression network. We identified 12 modules with characteristic expression patterns (Supplementary Fig. S5B). These modules could be further classified into two major groups based on clustering of eigenegene values (Supplementary Fig. S5C): retinal group (four modules in the right branch), and NVM group (six modules in the left branch). The four retinal modules had higher expression levels in Rnorm and RII samples than others (Figs. 5A–D). Genes in the blue (2341 genes) module were specifically expressed in normal and T2D retinas, and classified as the constitutive retina genes (Fig. 5A). Green (662 genes) and green-yellow modules (99 genes) had higher and specific expression levels in normal retina, which we named T2D-repressed modules (Figs. 5B, 5D). The purple module (214 genes) exhibited consistently high levels in T2D retina samples, which we named T2D-induced genes (Fig. 5C). In summary, the T2D-repressed and -induced genes comprised 29.4% of the four retina groups, indicating that the diabetic state profoundly altered the transcriptional profile of the retina.

We further explored the functions of genes in the four retina groups using GO and KEGG enrichment analysis. Among the top 10 enriched GO biological process (BP) terms in the blue module, seven of them were directly retina related, including phototransduction, retinal bipolar neuron differentiation, retina development in camera-type eye, and rhodopsin-mediated signaling pathway (Fig. 5A). This enrichment pattern was highly consistent with the constitutive expression of these genes in the retina, regardless of the pathological state. For enrichment of KEGG pathways, NOD-like receptor, phosphatidylinositol, and neurotrophin, as well as Fc epsilon RI signaling pathways were among the top 10 enriched pathways (Supplementary Fig. S5D).

T2D-repressed genes were mainly located in the green module. GO enrichment analysis revealed that these genes were highly enriched in synaptic transmission and neuronal functions, including vision perception, sensory perception of sound, and adult walking behavior, as well as response to stimulus (Fig. 5B). KEGG analysis revealed that a large number of synaptic functions were selectively repressed under the T2D state, including serotonergic synapse, cholinergic synapse, GABAergic synapse, and glutamatergic synapse (Supplementary Fig. S5E). In contrast to the T2D-repressed genes, we noticed that the T2D-induced genes in the purple module were specifically enriched in glucose metabolic processes and small molecule metabolic processes, and responses to stress and protein targeting to mitochondrion (Fig. 5C; Supplementary Fig. S5F). T2D-induced genes included VEGFA, ALDOA, ENO1, EGLN1, and LDHA were enriched in the HIF-1 signaling pathway. These results suggested that the transcriptional profiles in T2D retina had adapted to a stressful living condition containing high levels of glucose.

To compare the differential expression of genes in the retina modules and NVM modules, the top 25 most enriched BP terms and KEGG pathways in the four retina-specific modules were selected and plotted (Fig. 5E; Supplementary Fig. S5G), revealing that T2D-induced retina functions were generally shared more by NVM groups than the T2D repressed.

### The Neovascularization-Specific Gene Expression Programs

We then analyzed NVM-related modules, including eight modules exhibiting high levels in some or most NVM samples, which were called NVM-related modules (7837 genes). The turquoise (4306 genes) and brown (1073 genes) modules exhibited high expression in less than half of the NVMs (Supplementary Figs. S6A–B). The other six modules (2459 genes), including red, black, yellow, pink, magenta and tan, were prevalently high in NVM samples, showing similar or complementary expression patterns (Figs. 6A–F). For example, the red module showed elevated levels of gene expression in most of P_I_W and P_II_W samples and some other NVM samples (Fig. 6A). These genes were named neovascularization-related genes, which could serve as possible PDR markers for further study. We provided detailed module information in the supplementary information section (Supplementary Table S5).

We then plotted the top 25 BP and KEGG pathways in the six neovascularization-specific modules (Fig. 6G; Supplementary Fig. S6C). Most of these enriched terms were commonly from multiple neovascularization-specific modules. Impressively, most NVM terms were shared among neovascularization-specific modules, but were excluded from the retina groups. The pink module included VEGFC, EPHB4, EFNB1, and other angiogenesis-related genes such as PDGF\(_A\), TGFBR1, FZD6, ITGAV, ITGAS, TGFBR1, FGFR1, and SHC1. Some angiogenesis growth factors and receptors described in Figure 2 were also included in the magenta module, including ANGPT2, EFN1A, TIE2, FLT4, FLT1, and EPHB4 (Supplementary Table S5).

Furthermore, the pink (281 genes) and magenta (188 genes) modules were highly enriched in angiogenesis and extracellular matrix organization pathways (Fig. 6G). The expression profile of genes in angiogenesis and the related functional pathways had a global higher expression level in NVM samples (Supplementary Fig. S6D). In addition, several angiogenesis-related pathways such as blood coagulation, platelet activation,\(^{50}\) and leukocyte migration\(^{50}\) were also highly enriched in NVM modules (Fig. 6G; Supplementary Fig. S6C).
FIGURE 5. WGCNA analysis of the varied genes revealed distinct expression pattern modules. (A–D) A bar plot showing the eigengene value (left) and the top 10 enriched biological processes (right) of retinal related modules. Samples were sorted by sample ID order. (E) A heat map presentation showing the top enriched biological process terms for genes in the retina-related modules. Enrichment scores (−log_{10} P value) were normalized by terms for the presentation.
Although PDR is one of the leading causes of preventable blindness, a better understanding of its pathogenetic mechanisms is needed to improve clinical treatments.8 In the present study, we have presented a new approach to study retinal angiogenesis and neovascularization associated with PDR, through robustly comparing the deep-sequenced transcriptomes of NVMs from T2D, T1D, and BRVO patients with those from T2D and normal retinas. To the best of our knowledge, we are the first to characterize the global alterations of transcriptome profiles between NVMs and the retina, which provides an unprecedented opportunity to understand the mechanisms of retinal angiogenesis and neovascularization, as well as to develop novel biomarkers and therapies against PDR.

**Gene Expression in T2D Retinas Is Extensively Reprogrammed to Adapt to High-Glucose Stress**

It has been reported that loss of neuroretinal adaptation was detected before observable vascular pathologies,6 suggesting preceding transcriptome changes before DR diagnosis. Using the WGCNA method, we identified three modules of retina-specific genes: constitutive retina genes (2341, 70.6%), T2D-
repressed retina genes (761, 22.9%), and T2D-induced retina genes (214, 6.5%). Although we obtained only five retinal samples in this study, we could clearly separate the retinal samples from the NVMs to obtain three distinct retina-specific modules. Functional clustering analysis showed that genes in these modules were enriched in either expected or interesting functional pathways, which supported the validity of the results. We observed that the glucose metabolic process and small molecule metabolic process pathways were enriched only in T2D-induced modules. The different pathways between T2D-induced and T2D-repressed modules suggested that the transcriptional profile of retinal cells had been reprogrammed to adapt to a stressful condition including high levels of glucose, and some retinal functions were compromised in T2D patients. These findings are worthy of further investigation in animal models and in the early clinical diagnosis of DR.61

**Genes Overexpressed in NVMs as Potential Biomarkers**

Approximately half of patients fail to respond to current anti-VEGF treatment,8 necessitating the study of other factors contributing to PDR to serve as more comprehensive prognostic and diagnostic biomarkers. The study of additional genes contributing to PDR progression is hindered by technical difficulties in studying the very small amounts of patient tissues. A microarray analysis of gene expression in six NVMs from patients with PDR and three normal human retinas has only revealed 91 genes preferentially expressed in active neovascular patients, when compared with normal retinas, excluding VEGFs and angiopoietins.54 In this study, we reported the detection of 1387 overlapped DEGs among four diabetic proliferative membrane groups and the normal retina, including reported active neovascular genes. WGCNA analysis identified six modules with prevalent higher expression in NVM samples, when compared with normal retinas, containing 2459 genes. These genes were highly enriched in angiogenesis, blood coagulation, platelet activation, leukocyte migration, and extracellular matrix organization, consistent with the current understanding of the mechanism of angiogenesis,17,52 and suggested that a large number of angiogenesis-related genes were highly expressed in NVMs. Genes in NVM modules therefore represented potential biomarkers and candidates for developing PDR diagnostic and therapeutic methodologies.62

**Expression of VEGFC, ANGPT, EFNB3, and Their Receptors Is Induced During Retina Neovascularization**

The three major families, including VEGFs and the angiopoietin and ephrin protein families, play substantially differential and cooperative roles in promoting and regulating angiogenesis and neovascularization.16,17 In this study, analysis of the expression patterns of all genes in these three families showed that expression of VEGFC, ANGPT1 and ANGPT2, EFNB3, and their receptor genes were coordinately increased in NVMs compared with those of retinas, and these observations were confirmed by ISH and Western blot experiments. The essential role of VEGFC in angiogenesis has been reported in embryonically developing mice and during cancer.53–56 The high expression of VEGFC in T2D-induced NVMs suggested its potential role in maintaining the progress of NVM development. These findings represent the first report of a network of three major families of vascular endothelium-specific growth factors that may cooperate during retinal neovascularization, which is consistent with recent reports.27,51

**VEGFA Expression Is Highly Regulated During Retina Pathogenesis and Neovascularization**

The essential function of VEGFA in promoting neovascularization in ischemic retinopathy has been proposed based on multiple lines of evidence.9,25–33,67 Our results showed that VEGFA was expressed at high levels in normal retinas, diabetic retinas, and NVMs, while levels of VEGFA in diabetic retinas were significantly higher than in normal retinas and NVMs. The finding that VEGFA expression is specifically increased in T2D retina and reduced in NVMs was unexpected. However, it is consistent with the finding that the onset of DR may be present at the time of diagnoses for T2D patients, and suggests a model in which VEGFA plays an important role for the pathological transformation of retina from T2D to PDR.61 It also provides mechanistic insight into clinical observations that earlier anti-VEGF therapy produces better results.68

Based on the results presented in this study, we propose a dynamic gene regulatory model that initiates and stabilizes retinal neovascularization and neovascularization. In this model, expression of VEGFA is high in diabetic retina and is sensitive to pathological states. Upon pathological induction, VEGFA secretion into ocular fluid is increased. The increased level of extracellular VEGFA stimulates the expression of ETS TFs in endothelial cells.52 Based on these coexpression relationships, ETS TFs may in turn regulate the expression of growth factors, VEGFC, angiopoietin, and ephrin, as well as their receptors,54 (Figs. 2, 6E). These growth factors then coordinately drive the gene expression in endothelial cells to a pattern toward the NVMs rather than retina cells. It is very possible that the relatively high basal levels of VEGFA, and its main receptor, KDR, maintain the proper response of retinal endothelial cells to extracellular VEGFA, and this newly established expression pattern collectively supports the initiation and stabilization of retinal neovascularization.

In summary, the transcriptome sequencing and bioinformatics analysis presented in this study have not only generated the most informative transcriptome data, a comprehensive list of candidate genes for developing PDR treatment, and biomarkers, but have also provided mechanistic insights into the mechanism of retina angiogenesis and neovascularization. The dynamic expression of VEGFA, the induced expression of endothelium-specific growth factors, and a group of ETS family of TFs form the framework of an elaborate network that coordinates angiogenesis in PDR. Based on these observations, further studies with higher-resolution methods in animal and cell models are needed to further clarify, validate, and extend these regulatory mechanisms that occur during retinal neovascularization in PDR.69

**Acknowledgments**

The authors thank Jian Zhou for the discussion about data processing and Hong Wu, ABLife, Inc., for language polishing.

Supported by the National Natural Science Foundation of China (No. 81272242 and No. 81502284), National Key New Drug Creation and Manufacturing Program of Ministry of Science and Technology (No. 2013ZX09103003004), the Fundamental Research Funds for the Central Universities and the Grant of Jilin Province Science & Technology Committee, China (No. 20170414028GH, No. 20170520031JH, No. 201706230TC, No. 20150204038YY, No. 20150101188JC, No. 20150204038YY and No. 20150309033YY). This work was also supported by Appreciate the Beauty of Life, Inc. (No. 2014-11032, YZ).

Disclosure: Y. Li, None; D. Chen, None; L. Sun, None; Y. Wu, None; Y. Zou, None; C. Liang, None; Y. Bao, None; J. Yi, None; Y. Zhang, None; J. Hou, None; Z. Li, None; F. Yu, None; Y.
Huang, None; C. Yu, None; L. Liu, None; Z. Liu, None; Y. Zhang, None; Y. Li, None

References


