

Quantitative Proteomics and Weighted Correlation Network Analysis of Tear Samples in Adults and Children With Diabetes and Dry Eye

Xinrong Zou^{1,2,*}, Shanshan Wang^{3,*}, Pei Zhang⁴, Lina Lu¹, and Haidong Zou^{1,3}

¹ Department of Preventative Ophthalmology, Shanghai Eye Disease Prevention and Treatment Center, Shanghai, China

² Department of Ophthalmology, Fengcheng Hospital, Fengxian District, Shanghai, China

³ Department of Ophthalmology, Shanghai General Hospital, Shanghai Jiao Tong University, Shanghai, China

⁴ Department of Ophthalmology, Gonghui Hospital, Jingan District, Shanghai, China

Correspondence: Haidong Zou, Department of Preventative Ophthalmology, Shanghai Eye Disease Prevention and Treatment Center, Shanghai, 200040, China. e-mail: zouhaidong@sjtu.edu.cn

Received: May 28, 2020

Accepted: October 20, 2020

Published: December 4, 2020

Keywords: adults; children; diabetes; dry eye; tear proteomics; WGCNA

Citation: Zou X, Wang S, Zhang P, Lu L, Zou H. Quantitative proteomics and weighted correlation network analysis of tear samples in adults and children with diabetes and dry eye. *Trans Vis Sci Tech.* 2020;9(13):8, <https://doi.org/10.1167/tvst.9.13.8>

Purpose: Diabetics are more prone to suffer from dry eye (DE). The ages of diabetes are decreasing, so ocular surface status in younger generations is worthy of attention. We used tandem mass tag (TMT)-labeled proteomics and weighted correlation network analysis (WGCNA) to identify differentially expressed proteins in the tear proteome of adults and children with diabetic DE.

Methods: Study subjects were divided into six groups of 10, including three groups each for adults and children. The adult groups included diabetics with DE (A), diabetics without DE (B), and normal controls (C); the corresponding groups of children were identified as (D), (E), and (F). DE tests were performed on all subjects. We extracted total proteins and labeled them with TMTs for analysis. WGCNA was used to recognize hub genes.

Results: Tear film function was poorer in patients with diabetic DE. In adults, 1922 proteins were identified, and WGCNA analysis revealed three hub genes related to diabetic DE. For children, 2709 proteins were identified, and WGCNA analysis identified one hub gene related to diabetic DE. Kyoto Encyclopedia of Genes and Genomes analysis found similarities among metabolic pathways involved in differential expression of proteins in adult and child tear samples.

Conclusions: The pathogenesis of diabetic DE was highly similar in adults and children. The differentially expressed tear proteins in type 2 diabetes of adults and children was associated with inflammation, immune factors, and lipid metabolism.

Translational Relevance: Our findings found high similarities in the pathogenesis of diabetic DE in adults and children.

Introduction

Diabetic dry eye (DE), which greatly reduces the quality of life of patients with diabetes, is a common condition treated in healthcare outpatient departments.¹ DE may cause ocular discomfort, such as dryness, foreign body sensation, eye pain, vision fluctuations, and even serious consequences including secondary bacterial infections, scarring, and corneal perforations.² Previous studies on diabetic populations in hospitals and communities revealed that the prevalence

of DE in patients with diabetes was significantly higher than in the general population.^{2–7} The mean age of diabetic patients is decreasing, and the incidence of diabetes in children is increasing worldwide.^{8,9} In a recent study, we found that DE disease is more prevalent in children with diabetes than in those without diabetes.¹⁰ The ocular surface status of this population also deserves greater attention. However, the reasons underlying the elevated prevalence of DE in diabetics are yet to be elucidated; microvascular lesions of the lacrimal gland vessels and autonomic neuropathy of the lacrimal gland system are reportedly among

the possible mechanisms involved.^{11,12} Such lesions may result in lacrimal gland injuries at the onset of diabetes. Diabetic keratopathy can cause damage to the corneal epithelium basement membrane and goblet cells, leading to altered expression of tear proteins. Tear proteins play an important role in the composition and function of tear film, so in general, DE occurs because of severely impaired tear film function.¹³

Tear fluid is rich in proteins that are closely related to the functions of tear film^{14,15}; however, because of limitations of experimental methods, only a limited number of tear proteins have been identified so far. In view of advances in proteomics and mass spectrometry technology, increasing numbers of tear proteins can now be identified. Furthermore, relationships between the occurrence of DE and tear proteins,^{16,17} and new information regarding the functions and interactions of tear proteins can be investigated. Nevertheless, it remains to be investigated whether there are differences between the pathogenesis of diabetic DE in adults and children. In this study, a tandem mass tag (TMT)-based global quantitative proteomics analysis of tear samples from adults and children with diabetic DE was carried out to quantify protein expression levels in tear samples. In addition, weighted correlation network analysis (WGCNA, also known as weighted gene coexpression network analysis) was adopted to explore novel theoretical bases involving possible mechanisms underlying the development of diabetic DE in adults and children.

Materials and Methods

Subjects

Our study was in accordance with the tenets of the Declaration of Helsinki and the principles of ARVO statement for protecting human subjects in an experiment. It was a part of the Shanghai Children and Adolescent DM Eye study and the Shanghai Cohort Study of Diabetic Eye Disease study. The name of institution review board was the Ethics Committee of Shanghai General Hospital and the ethics committee of the Children's Hospital affiliated to Shanghai Fudan University. The certification number of this study was 2016ky005 and (2018) 01. In August 2017, three groups of adult subjects (10 in each) were recruited at the Ophthalmology Clinic of Shanghai General Hospital: Group A (diabetes with DE), Group B (diabetes without DE), and Group C (normal controls). There were no differences in gender and age among the groups. From January to February 2018, children diagnosed with diabetes at the Pediatric Hospital of Fudan University were recruited as research subjects,

and children without diabetes and matched for age and sex were also recruited as a control group. The grouping of child subjects (10 in each) was the same as that of adult: DE for children with diabetes (Group D), non-DE for children with diabetes (Group E), and normal control children (Group F). In total, 60 subjects were included in this study.

All adult subjects provided written informed consent. The inclusion criteria for children included in the standard subjects were the following: (1) availability of written informed consent signed by the guardian; (2) age between five and 18 years. The following conditions were excluded from this study. Patients with ocular status that would affect tear production or quality would be removed, including eyelid diseases (ectropion, entropion, trichiasis, ptosis) and eyelid movement disorder(s) caused by facial paralysis, conjunctiva diseases (conjunctivochalasis and pterygium), ocular surgeries within six months or refractive surgeries within two years, history of ocular chemical injuries or use of ocular medications or nutritional tear supplements, and systemic diseases such as Sjogren's syndrome, Parkinson's disease, rheumatoid arthritis, Grave's disease, and systemic lupus erythematosus.⁷

Tear Sample Collection

All adult subjects underwent DE tests, including tear film break-up time (BUT), Schirmer 1 test (S1t), and corneal fluorescein staining (FL). The DE tests were also performed in children, including BUT and S1t. The Schirmer strip from the S1t test was placed in a 5 mL centrifuge tube (Eppendorf, Hauppauge, NY, USA) and stored at -80°C before further processing.

Diagnostic Standard

Diagnostic criteria for type 2 diabetes in adults were adopted from the World Health Organization guidelines.¹⁸ Children with diabetes (including type 1 and type 2 diabetes) were diagnosed based on the WHO diagnostic criteria, which include the following: (1) Typical symptoms (polydipsia, polyuria and unexplained weight loss), fasting blood glucose >7 mmol/L or postprandial blood glucose >11.1 mmol/L. (2) No typical symptoms with fasting blood glucose >7 mmol/L or postprandial blood glucose >11.1 mmol/L. Those who attained the above values were diagnosed as having diabetes mellitus. (3) No typical symptoms were observed with fasting blood glucose >7 mmol/L or postprandial blood glucose >11.1 mmol/L, and glucose tolerance test two-hour blood glucose >11.1 mmol/L.¹⁰

Because there is no international consensus regarding diagnostic standards for DE, our study used the

modified DE Workshop classification and Chinese Academy of Medical Sciences Consensus of clinical experts on DE (2013). DE was considered to be present in subjects with BUT < 5 seconds or SIt < 5 mm, and absent in those with BUT > 10 seconds and SIt > 10 mm.^{10,19}

Protein Extraction

All Schirmer strips collected from subjects were sliced and placed in ice-cold lysis buffer (8 M urea; Sigma, St. Louis, MO, USA; 1% Calbiochem protease inhibitor cocktail III, Calbiochem, German). Next, samples were sonicated three times with a high-intensity ultrasonic processor (Ningbo Scientz Biotechnology Co., Ltd. Ningbo, Zhejiang, China). Centrifugal removal of residual debris was performed by centrifugation at 12,000g at 4°C for 10 minutes. Supernatants were collected and protein concentrations determined according to the manufacturer's instructions using a commercially available bicinchoninic acid assay kit (Beyotime, Beijing, China).

Trypsin Digestion

For protein digestion, the solution was reduced for 30 minutes with 5 mM dithiothreitol at 56°C, and then alkylated with 11 mM iodoacetamide at room temperature (25°C) in darkness for 15 minutes. The urea concentration of protein sample was then diluted to less than 2 M by adding 100 mM triethylammonium bicarbonate (TEAB). Finally, trypsin was added and the ratio of trypsin/protein mass in the first digestion (overnight) was 1:50, while that in the second digestion (4 h) was 1:100.

TMT Labeling

After trypsin digestion, peptides were desalted using a solid-phase extraction column (strata × C18, Phenomenex, Torrance, CA, USA) and then, vacuum dried. Peptides were resolvated in 0.5 M TEAB and treated in accordance with the manufacturer's operating instructions for the TMT kit.

HPLC Fractionation

Trypsin peptides were separated by high pH reverse-phase high performance liquid chromatography (HPLC) with C18 columns (5 μm particles, 4.6 mm ID, 250 mm length, 300Extend; Agilent, Santa Clara, CA, USA). In short, the peptide was first divided into 60 fractions in 60 minutes using a gradient of 8% to 32% acetonitrile (pH 9.0). The peptides were then

combined into eight fractions for adult samples and 18 fractions for child samples and dried in a vacuum centrifuge.

Liquid Chromatography-Mass Spectrometry Analysis

Tryptic peptides were dissolved in liquid chromatography mobile phase solution (0.1% (V/V) formic acid aqueous solution) (solvent A) and directly loaded onto a homemade reverse-phase analytical column (15 cm length, 75 μm internal diameter). The gradients included: solvent B from 6% to 23% (0.1% formic acid in 98% acetonitrile) in 26 minutes, 23% to 35% in eight minutes, and 80% over three minutes, and then kept at 80% for the final three minutes at a constant flow rate of 400 nL/min on an easy NLC 1000 UPLC system (ThermoFisher Scientific, Waltham, MA, USA). The peptides were subjected to nanospray ion source followed by tandem mass spectrometry (MS/MS; Q Exactive Plus, ThermoFisher Scientific) for adult tear samples and Orbitrap Fusion Lumos for children's tear samples, coupled with an ultra-HPLC instrument.

The applied electrospray voltage was 2.0 kV. The m/z range of the complete scan was 350 to 1800 and the resolution for detecting intact peptides in the Orbitrap was 70,000. Peptides for MS/MS were then selected using the normalized collision energy setting 28 and fragments were detected at a resolution of 17,500 using the Orbitrap. A data-dependent process was performed 20 MS/MS scans after one MS scan, with a dynamic exclusion time of 15.0s. The automatic gain control was set to 5E4; the fixed first quality was set to 100 m/z.

We used Maxquant (v1.5.2.8) to retrieve secondary mass spectrometry data. Search parameter settings were as follows: the database was human SwissProt (20,130 sequences), the anti-library was added to calculate the false-positive rate (false discovery rate [FDR]) caused by random matching, and a general pollution library was added to the database to eliminate contaminating proteins in the identification results. Cysteine alkylation was set as fixed modification, while oxidation of methionine and N-terminal acetylation of proteins were set as variable modification. The quantitative method was set to TMT-10plex and the FDR of protein identification and peptide-spectrum match identification was set to 1%.

Bioinformatics Methods

WGCNA can be used to study the expression patterns of multiple genes with similar expression patterns. It can construct large networks in an unsuper-

vised manner and describe correlation patterns; it is one of the most powerful methods for network construction. Genes with similar expression trends in different samples could be defined as a “gene module.” The scale-free topological network model used by WGCNA is more suitable for protein-protein interaction networks. Therefore we sought to explore the relationships between biological characteristics and modules in a more detailed, comprehensive, and systematic fashion. Using this analytical method, the most relevant gene modules in clinical practice are revealed, which can be used to study the basic mechanisms involved in disease development. To further investigate the mechanisms underpinning the influence of module genes on related clinical characteristics, genes of interest in the module were uploaded to the UniProt-GOA Database (<http://www.ebi.ac.uk/GOA/>) for Gene Ontology (GO) functional annotation and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. An FDR < 0.01 was used as the cutoff criterion.

WGCNA includes gene coexpression, network construction, module identification, module-phenotype correlation recognition, and key driver gene identification. In the current study, we applied this method to construct correlation patterns among our protein samples. The nodes of networks refer to protein expression profiles, and edges between proteins were determined by pairwise correlations of their expression. Because LC-MS/MS data inevitably have missing values, it is necessary to filter data before subsequent analyses to guarantee reliability of the de novo network construction. For data preprocessing, we used the “good samples genes” function in the WGCNA R package to iteratively remove samples and proteins with > 50% missing entries (default setting).

The GO annotation proteome was from the UniProt Go database, supplemented by InterProScan soft. The proteins were classified according to biological processes, cell components and molecular functions. Wolf PSORT software was used to predict subcellular localization. The KEGG database was used to annotate protein pathways.

Identification of Hub Genes

Hub genes may play vital roles in the development of diabetes in patients with DE, so it is necessary to identify hub genes belonging to the modules. Visualization of module structures and network connections can be realized by several methods. The function of exporting the network to Cytoscape allows the networks to be exported in a format fit for analysis by Cytoscape. The flowchart to describe the data management process in

the bioinformatic methods is included in Supplementary Material.

Statistical Methods

Fisher’s exact test, one-way analysis of variance, and independent sample *t*-tests were performed with SPSS version 24.0 (IBM Corporation, Chicago, IL, USA); *P* < 0.05 was considered statistically significant. Data were expressed as means ± standard deviation. For each category in the GO and KEGG analyses, Student’s two-tailed *t*-test with *P* values < 0.05 was considered statistically significant.

Results

Demographic and Baseline Characteristics of the Subjects

The age and gender of adult or child subjects were matched (Table 1 for adults, Table 2 for children). Compared with patients with diabetes alone, adults with diabetes and DE had lower BUT scores, SIt values, and greater duration of diabetes (Table 1). Child subjects with diabetic DE had lower BUT scores (Table 2). Adults had type 2 diabetes, whereas children had type 1 diabetes.

Tear Protein Identification

In our study, quantitative proteomics analysis of TMT-labeled in adults and children was performed. Adult tear samples presented a total of 1922 proteins, of which 1814 contained quantitative information, and 650 had quantitative values for 30 samples. A total of 2709 proteins were identified in children’s tear samples, of which 2357 were quantifiable, and 1089 had quantitative values in 30 samples. We performed subsequent WGCNA analysis on these 650 and 1089 proteins. The quality control test results of mass spectrum data are included in Supplementary Figures S1 and S2.

WGCNA Results

In order to determine the major characteristics of the proteome of tear samples, the WGCNA software package in R was applied to the coexpression network using expression values and clinical data involving 650 differentially expressed genes from 30 adult tear samples, and 1089 differentially expressed genes from 30 child tear samples.

First, cluster analysis was performed on the 650 and 1089 proteins of the tear samples to check whether

Table 1. Baseline Characteristics of 30 Adult Tear Samples

Characteristics	A Diabetes with DE (n = 10)	B Diabetes Without DE (n = 10)	C Normal Subjects (n = 10)	Statistic	<i>P</i>
Gender (n, %)				—	0.861
Female	4(40.00)	3(30.00)	3(30.00)		
Male	6(60.00)	7(70.00)	7(70.00)		
Age ($\bar{x} \pm$ SD, years)	58.80 \pm 4.26	57.70 \pm 7.17	58.00 \pm 4.32	0.110	0.896
BUT ($\bar{x} \pm$ SD, s)	3.58 \pm 0.81	10.97 \pm 2.00	11.63 \pm 1.78	76.720	<0.001
S1T ($\bar{x} \pm$ SD, mm)	3.30 \pm 1.57	14.80 \pm 6.63	18.90 \pm 7.82	18.231	<0.001
FL	0.6 \pm 0.577	0	0		
Diabetic duration ($\bar{x} \pm$ SD, years)	12.40 \pm 4.50	5.70 \pm 3.02	—	6.31	0.001

Gender was analyzed with Fisher's exact test. Three groups were compared using the Kruskal-Wallis test, and pairwise comparison between groups was based on the Student-Newman-Keuls method. The BUT and S1t in Group A were significantly lower than that in other groups.

BUT, tear break-up time; S1t, Schirmer 1 test; FL, fluorescein staining.

Table 2. Baseline Characteristics of 30 Child Tear Samples

Characteristics	D Diabetes with DE (n = 10)	E Diabetes Without DE (n = 10)	F Normal Subjects (n = 10)	Statistic	<i>P</i>
Gender (n, %)				—	1.000
Female	4 (40.00)	4 (40.00)	4 (40.00)		
Male	6 (60.00)	6 (60.00)	6 (60.00)		
Age ($\bar{x} \pm$ SD, years)	11.70 \pm 2.79	12.00 \pm 3.30	11.20 \pm 1.32	0.240	0.788
BUT [Median (Q1,Q3), s]	4 (3.00,4.00)	9 (7.75,12.00)	10 (8.00,12.25)	14.706	<0.001
S1T ($\bar{x} \pm$ SD, mm)	9.20 \pm 3.46	13.50 \pm 5.79	13.90 \pm 6.17	2.435	0.107
FL	0.4 \pm 0.843	0	0		
Diabetic duration [Median (Q1,Q3), years]	4.50 (1.00,6.25)	4.50 (2.00,6.25)	—	6.31	0.971

Gender was calculated with Fisher's exact test. Age and S1T used one-way analysis of variance. BUT comparison used Kruskal-Wallis test, and the pairwise comparison between groups used the Student-Newman-Keuls method. The BUT in Group D was significantly lower than that in other groups.

S1t, Schirmer 1 test; FL, fluorescein staining.

there were any outliers. According to the clustering results, A_1 to A_8 and C_1 in the adult samples were outlier samples, which would be eliminated in subsequent analysis. (Fig. 1A) For the children's samples, C_2 to C_5 was an outlier sample, which would also be eliminated in subsequent analysis. (Fig. 1B)

Our analysis showed that strong modularity was found in analysis of tear samples proteome. (Fig. 2A and Fig. 2B)

Identification of Hub Genes

The WGCNA package provides a suitable way to export networks to CytoScape for network visualization. The turquoise module coexpression networks

for adult and child tear samples were exported and then visualized using CytoScape (open source, www.cytoscape.org) software, in which line thickness represents protein weight. (Fig. 3A and Fig. 3B) Under the threshold of module connectivity (core gene Module Membership) > 0.8 and clinical trait relationship (core gene Trait Significance) > 0.2, 3 genes (*LYZ*, *ZAG*, *DNAJC3*) with high connectivity (module in turquoise) were selected as hub genes in the adult group (Fig. 3A), and 1 gene (*PGKI*) with high connectivity (module in turquoise) was selected as a hub gene in the group of children (Fig. 3B). The molecular functions of the hub genes were listed in Table 3.

The correlation analysis between hub genes and clinical characteristics were listed in Table 4. The data

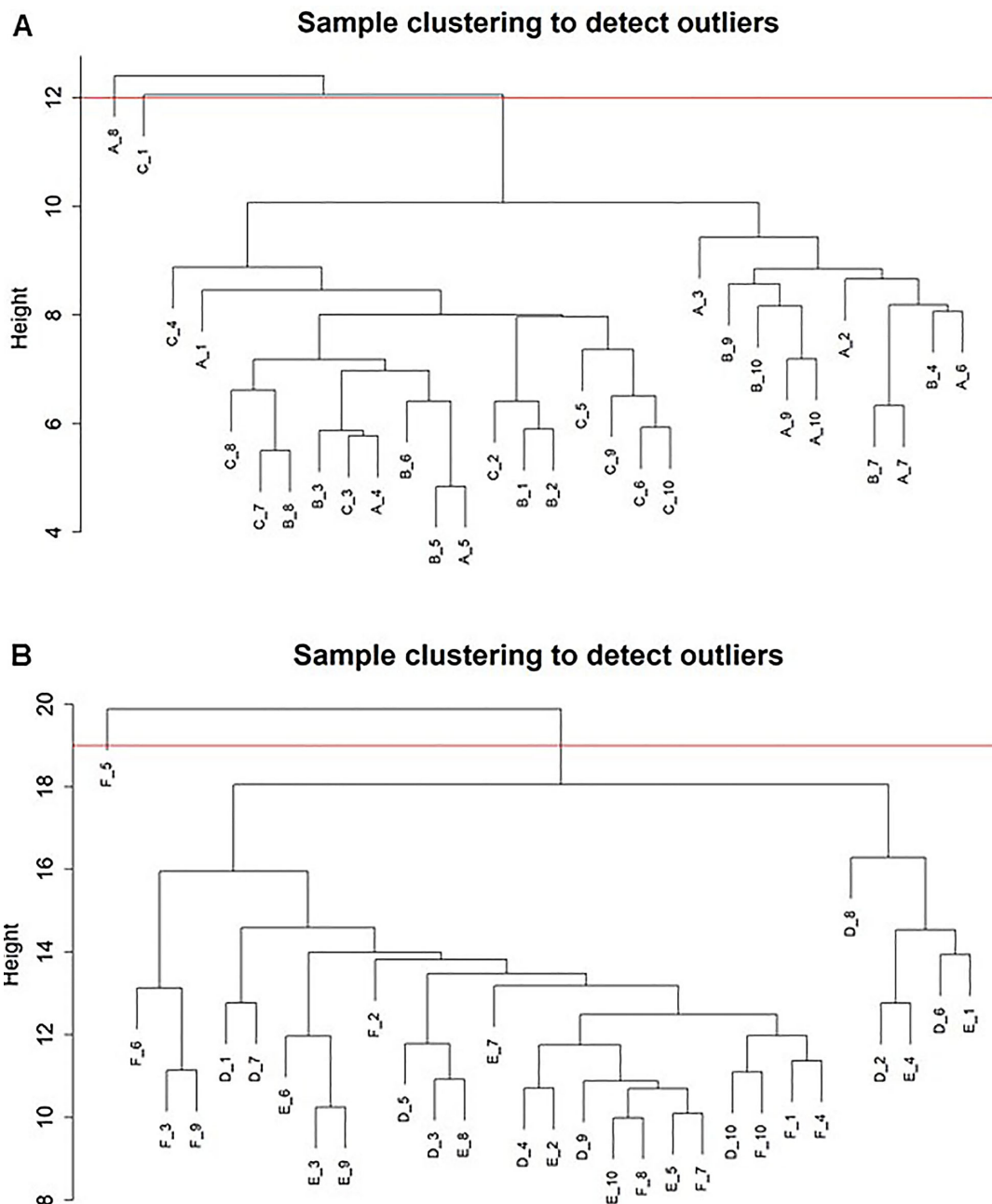


Figure 1. (A) Outliers in adult samples. A_1 to A_8 and C_1 were outlier samples, and they were eliminated for subsequent analysis. (B) Outliers in child samples. F_2 to F_5 was an outlier sample, and it was eliminated for subsequent analysis.

obtained from LC-mass (either from adults or children) is included in Supplementary Excels S1 and S2. The detail of identifying hub genes is included in Supplementary Excels S3 and S4.

GO and KEGG Analysis

GO enrichment analysis in adults showed that most of the differentially expressed proteins originated from extracellular exosomes, vesicles, and organelles, which

perform molecular functions involving catalytic activity, cell adhesion molecule binding, cadherin binding, and hydrolase activity, and participate in biological processes such as retinal homeostasis, myeloid leukocyte activation, cell secretion, and activation of immune responses (Fig. 4). GO enrichment analysis in children showed that most of the differentially expressed proteins originated from extracellular exosomes, vesicles, and organelles, which perform molecular functions involving catalytic activity, same

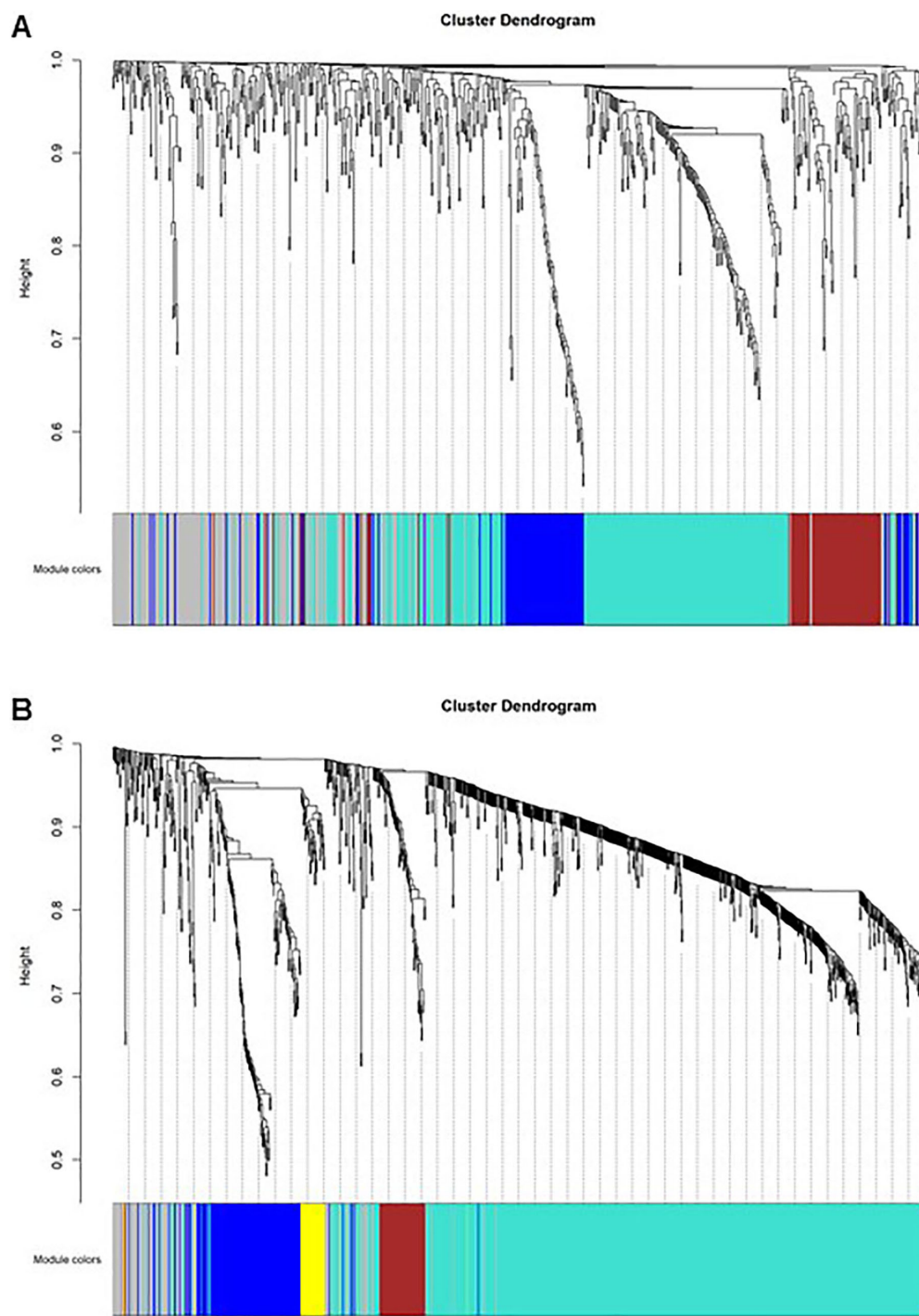


Figure 2. (A) Cluster dendrogram and module overview for adult samples. A total of three nonoverlapping modules with highly correlated proteins were detected in adult tear samples, covering 52 to 340 proteins. According to conventions involving WGCNA, modules are tagged here with different colors. One hundred seventy-four proteins were not assigned to any module and are marked in *gray*. (B) Cluster dendrogram and module overview for child samples. In total, five nonoverlapping modules with highly correlated proteins were detected in child tear samples, encompassing 34 to 725 proteins. Modules were also tagged with different colors. One hundred nine proteins were not assigned to any modules, and they are labeled in *gray*.

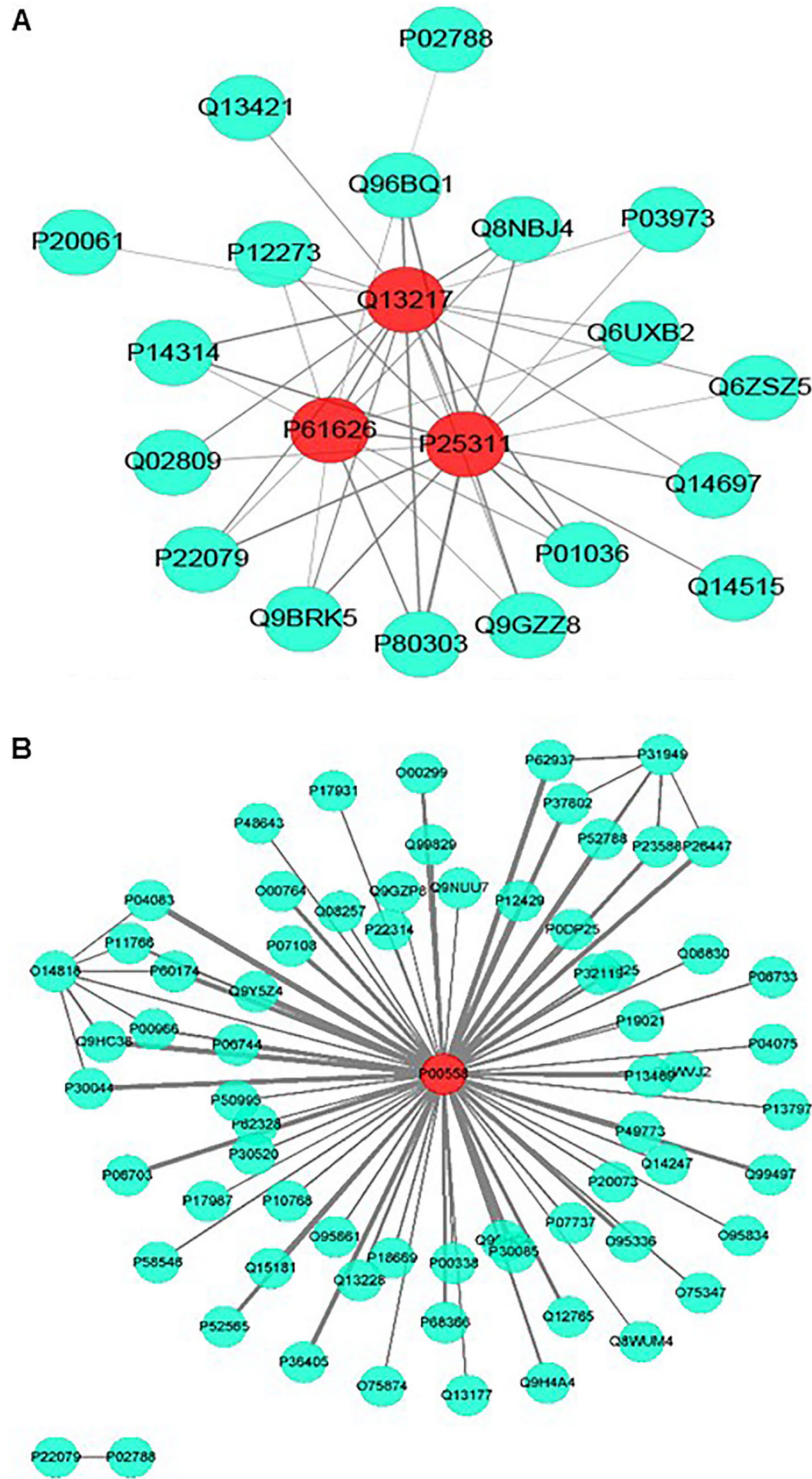


Figure 3. (A) Coexpression network visualization for adult tear samples. P61626 (LYZ), P25311 (ZAG), and Q13217 (DNAJC3) with high connectivity (module in *turquoise*) were identified as hub genes in the adult group. (B) Coexpression network visualization for child tear samples. P00558 (PGK1) with high connectivity (module in *turquoise*) was selected as a hub gene in child group.

Table 3. Data of 4 Hub Genes in the Turquoise Module

Protein Accession	Protein Description	Gene Name	Molecular Function	Quantitative Value Ratio in the 3 Groups	P Value
Hub genes in adult group					
P61626	Lysozyme C	<i>LYZ</i>	Bacteriolytic function; lysozyme activity; identical protein binding; catalytic activity	B/A 1.72 C/B 1.28	<0.05
P25311	Zinc-alpha-2-glycoprotein	<i>AZGP1</i>	Stimulates lipid degradation in adipocytes	B/A 1.61 C/B 1.31	<0.05
Q13217	DNA J homolog subfamily C member 3	<i>DNAJC3</i>	Chaperone binding, misfolded protein binding, protein kinase binding, protein kinase inhibitor activity	B/A 1.55 C/B 1.35	<0.05
Hub gene in child group					
P00558	Phosphoglycerate kinase 1	<i>PGK1</i>	Encoding 3-phosphoglycerate kinase, a polymerase alpha cofactor protein	E/D 1.03 F/E 0.93	>0.05

Table 4. Correlation Analysis Between Hub Genes and Clinical Characteristics

Group	Hub Gene	Age	BUT	S1T	FL	DM Duration
Adult						
	<i>LYZ</i>	-0.114	0.699***	0.709***	-0.260	-0.607**
	<i>AZGP1</i>	-0.066	0.679***	0.696***	-0.265	-0.590**
	<i>DNAJC3</i>	-0.090	0.684***	0.692***	-0.274	-0.586**
Child						
	<i>PGK1</i>	0.183	-0.062	-0.266	-0.069	0.309

The correlation analysis between adult hub genes and age, BUT, FL adopts Pearson correlation coefficient, and the correlation analysis with S1T and DM duration adopts the Spearman correlation coefficient.

The correlation analysis between child hub gene and age, S1T adopts Pearson correlation coefficient, the correlation analysis with BUT, FL, DM duration adopts spearman correlation coefficient. **, *** indicate that the hypothesis test is statistically significant at the 1% and 0.1% confidence levels, respectively.

S1t, Schirmer 1 test; FL, fluorescein staining.

protein binding, protein complex binding, and carbohydrate derivative binding. With respect to biological processes, activation of myeloid leukocytes, leukocytes in immune responses, and regulation of cellular amine metabolism were associated with the differentially expressed genes (Fig. 5). We speculate that inflammation and immunity are likely involved in the development of diabetic DE.

KEGG analysis showed that the differentially expressed proteins participated in metabolic pathways. In both adults and children tear samples, differentially expressed proteins were mainly involved in

the same metabolic pathways ,including glycolysis, pentose phosphate pathway, and proteasomes. (Figs. 6, 7) Our results suggested that there were similarities between the metabolic pathways involved in differentially expressed proteins in adult and child tear samples.

Discussion

In the current study, we carried out a TMT-based global quantitative proteomics analysis in adults and children with diabetic DE. In addition, we used

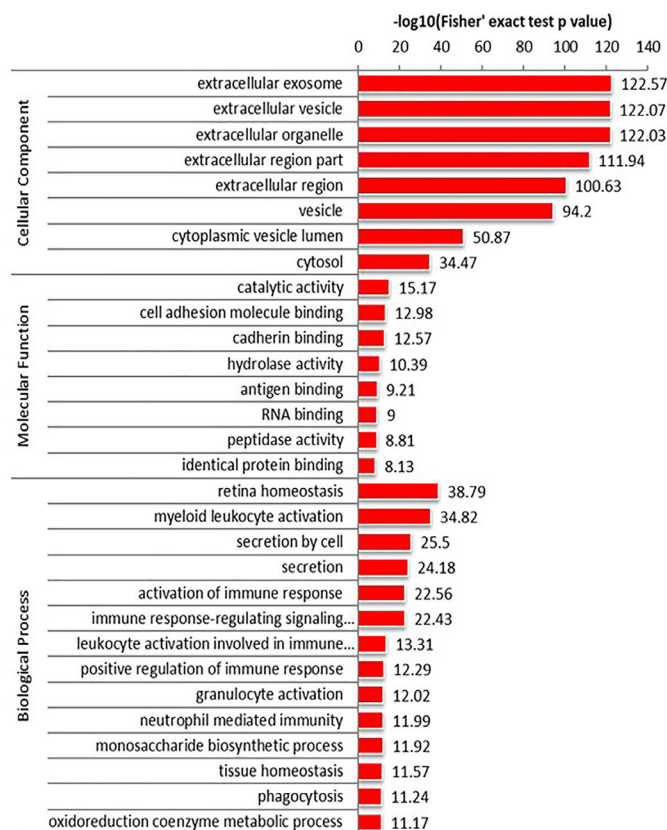


Figure 4. Analysis of GO enrichment in adult tear samples. Differentially expressed proteins were derived from extracellular exosomes, vesicles, and organelles, which perform molecular functions including catalytic activity, cell adhesion molecule binding, cadherin binding, and hydrolase activity, and participate in biological process such as retinal homeostasis, myeloid leukocyte activation, cell secretion, and activation of immune responses.

WGCNA to investigate the proteomes of tear samples from adults and children for the first time. A potential novel theory could be proposed for the mechanisms underlying diabetic DE.

Biological significance could be represented using WGCNA clustering criteria, which is entirely different from clustering methods based on the geometric distance between data. The stability of each module was calculated after identifying the genetic modules. Then, correlations between the stable modules and clinical characteristics (such as age and gender) were analyzed. Using this method, we were able to identify the most clinically related gene modules, which could be further used to explore the main causes of disease development and provide additional information for understanding the pathogenesis of DE disease.²⁰

GO enrichment analysis indicated that the differentially expressed proteins in tear samples from adults and children originated from the same sources and the molecular functions and biological processes involved

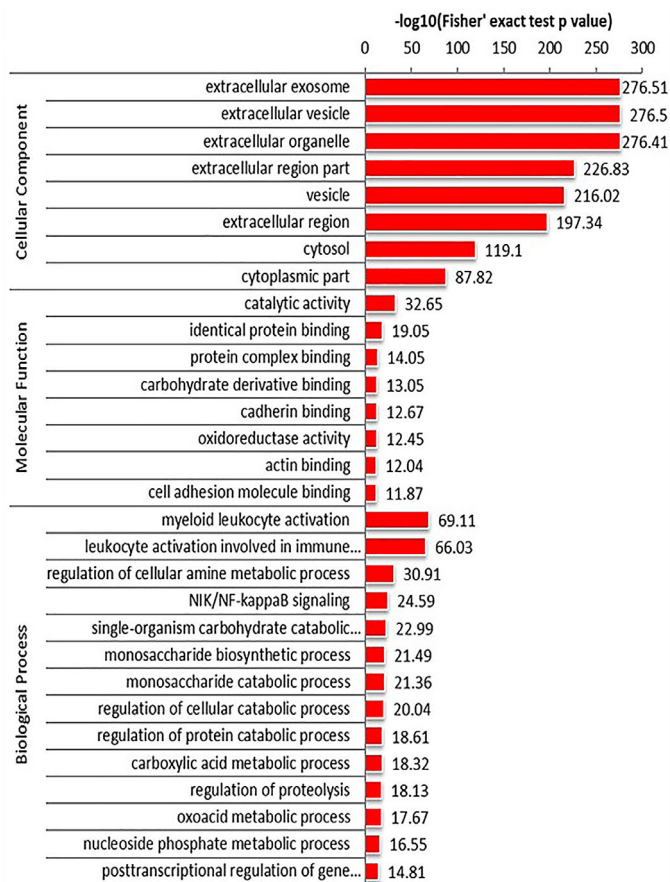


Figure 5. Analysis of GO enrichment in child tear samples. Differentially expressed proteins were derived from extracellular exosomes, vesicles and organelles, which perform molecular functions including catalytic activity, the same protein binding, protein complex binding, and carbohydrate derivative binding. With regard to biological processes, activation of myeloid leukocytes, leukocytes in immune responses, and regulation of cellular amine metabolism were associated with these differentially expressed genes.

were highly analogous. Thus, we speculate that inflammation and immune responses are mainly related to the occurrence of diabetic DE. Zhou et al.¹⁶ demonstrated that the severity of DE can be classified on the basis of the expression levels of a group of inflammation-related proteins, which includes α 1-acid glycoprotein, S100 A8, and S100 A9. These proteins are upregulated in DE groups. The higher the levels of these proteins, the more severe was the degree of DE. Li et al.¹⁷ reported that patients with diabetes and DE exhibit increased expression of immune and inflammation-related proteins, including neutrophil elastase 2 and clusterin, when compared to healthy subjects. Other studies have reported alterations in the protein expression profiles of inflammatory cytokines, such as increased levels of interleukin-8 (IL-8), IL-1 α , IL-1 β , IL-12, tumor necrosis factor- α (TNF- α), and interferon- γ (INF- γ), in the tears of DE patients; this

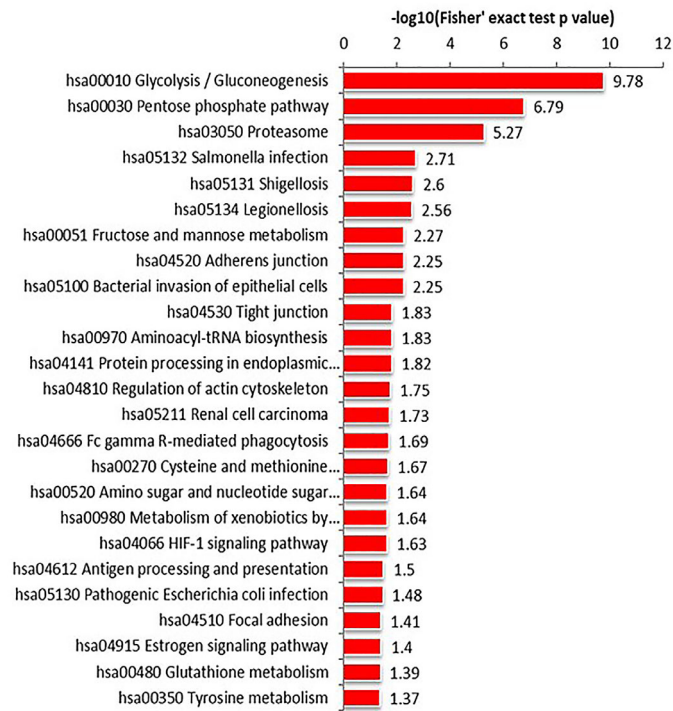


Figure 6. KEGG pathway enrichment of adult tear samples. Differentially expressed proteins were mainly involved in glycolysis, pentose phosphate pathway, and in proteasomes.

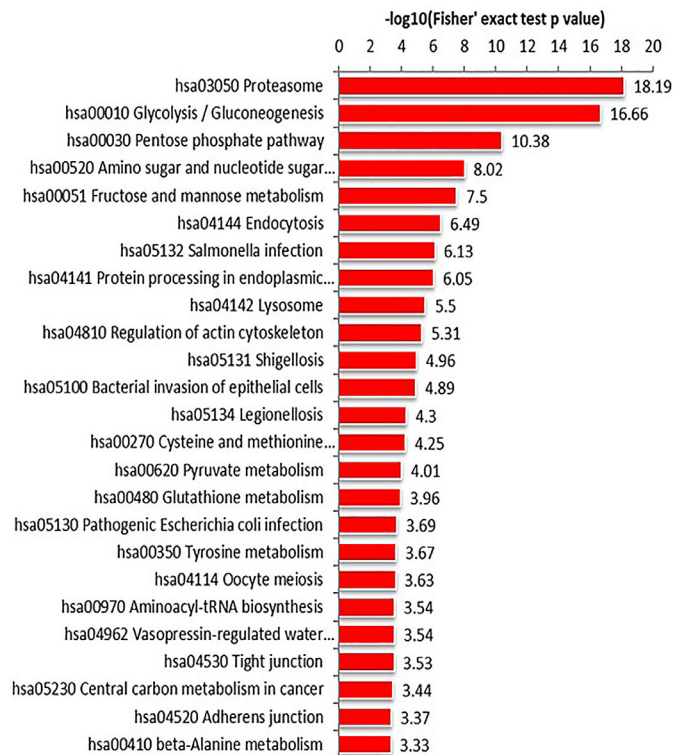


Figure 7. KEGG pathway enrichment of child tear samples. Differentially expressed proteins were involved in proteasomes, glycolysis, and the pentose phosphate pathway.

indicates the involvement of inflammatory processes in DE development.^{21,22} Therefore inflammation and immune factors are highly correlated with DE occurrence.

Retina homeostasis in the biological process was specifically identified in adult GO enrichment analysis rather than child. The reason may be the differences in retinal status between adults and children. The diabetic duration of the adult diabetic DE group was significantly longer than that of child diabetic DE group, and some subjects of adult diabetic DE group suffered from mild to moderate diabetic retinopathy. Csoz et al.²³ reported that some proteins were differentially expressed in DR patients, including the inflammation-related protein lactotransferrin (LTF), along with the other five identified in their study (lipocalin 1 [LCN1], lacritin [LACRT], lysozyme C [LYZ], lipophilin A [SCGB1D1], and immunoglobulin λ chain [IGLC1]), they were considered to be candidate biomarkers of DR diagnosis. In Hagan's study, B-2-microglobulin (B2M) showed a decreased expression in DR patients compared to healthy controls and diabetic patients without DR.²⁴ In our study, LTF, LCN1, LYZ, and B2M were all related to the biological process of retina homeostasis. The relationship between tear proteins and diabetic retinopathy needs further research.

The results of KEGG analysis showed similarities between the metabolic pathways involved in differential expression of proteins in tear samples from adults and children. The differentially expressed proteins in tear samples from children were mainly enriched in the proteasome, glycolysis, and pentose phosphate pathways. Differentially expressed proteins were involved in the same metabolic pathways in the tear samples from adults and children; these pathways included glycolysis, pentose phosphate pathway, and proteasome. Therefore we speculate that the pathogenesis of DE in adults and children is highly correlated with the three above-mentioned metabolic pathways.

Our investigations finally focused on hub genes; a total of four hub genes were identified: *LYZ*, *ZAG*, and *DNAJC3* in adult tear samples, and *PGK1* in child tear samples. In comparisons between the samples from adults and children, GO and KEGG results were highly similar, although the hub genes involved were different. Such differences may be associated with the diabetes type. The adult subjects were all type 2 diabetics, whereas all the child subjects had type 1 diabetes. It is well known that type 1 diabetes is an autoimmune disorder caused by immune-mediated destruction of pancreatic beta cells. In contrast, type 2 diabetes is a polygenic disease caused by multiple genetic and environmental factors. Hence, differences in the pathogenesis may be the result of different hub genes.

LYZ is an antibacterial tear protein. LYZ is the most abundant tear protein with antimicrobial activity that kills Gram-positive bacteria through catalytic hydrolysis of cell wall peptidoglycans and may also have antibacterial effects on Gram-negative bacteria, probably via electrostatic interactions between its cation domain and the negatively charged microbial membrane.²⁵ According to previous reports, LYZ is significantly down-regulated in the tears of DE patients.^{14,26,27} Furthermore, Versura and colleagues²⁸ reported that LYZ exhibits a significant association with the prediction of Sjögren syndrome and not-Sjögren syndrome DE. Hanstock et al.²⁹ concluded that LYZ has potential as a biomarker of mucosal immune status and believed that because the eye can represent the entrance point for transmission of infections, the assessment of immune status at the ocular surface may have clinical relevance. LYZ concentrations are lower in certain cases, such as upper respiratory tract infection and contact lens usage.²⁷ In our study, the level of LYZ was quantitatively the lowest in the adult diabetic DE group. The decrease in LYZ indicated that the ocular surface defense system was perturbed and the ocular surface homeostatic microenvironment was disrupted.³⁰ A proposal regarding this concept was presented by Zhang et al.³⁰ in 2017; they pointed out that changes to or absence of one or more components of the ocular surface may lead to the destruction of otherwise stable ocular surfaces and eventually, DE disease ensues. Their results also showed that the level of LYZ was lowest in the diabetic DE group, and our results corroborated this to a certain extent. Ocular surface homeostasis in the diabetic DE group was poor.

Zinc alpha-2-glycoprotein (*ZAG*), which is involved in lipid metabolism, was among the hub genes identified in this study³¹; it is also reported to be significantly associated with obesity.³² *ZAG* levels in serum are negatively correlated to body weight and body fat percentage. Balaž et al.³² discovered that the expression level of *ZAG* in adipose tissue is positively associated with insulin sensitivity of the whole body and adipose tissue but negatively correlated with the morphological size of adipocytes. They also found that *ZAG* plays an important role in regulating insulin sensitivity of the body and adipose tissue.³³ Versura et al.²⁷ revealed that *ZAG-2* is strongly correlated with BUT and reported a statistically significant decrease in the levels of LYZ (3.06 ± 1.07 vs. 2.15 ± 0.78) and *ZAG* (0.43 ± 0.24 vs. 0.25 ± 0.2) (mg/mL, mean \pm SD) in patients with DE compared to those in normal controls; these proteins are related to the severity of DE.²⁷ Liu et al.³⁴ demonstrated that the serum level of *ZAG* in Chinese individuals is inversely related to

body mass index and body weight. Obesity is one of the main risk factors for the development of type 2 diabetes, and decreased insulin sensitivity results in the occurrence of type 2 diabetes. Therefore identification of the hub gene *ZAG* suggested that there may be a correlation between the occurrence of DE and lipid metabolism, and *ZAG* may be a potential biomarker for diabetic DE diagnosis.

DNA J homologous subfamily C member 3 (*DNAJC3*) is a co-chaperone of BiP, which is a member of the HSP70 chaperone protein family found in the endoplasmic reticulum.³⁵ *DNAJC3* plays a role in insulin secretion deficiency and type 2 diabetes.³⁶ Lu et al.³⁶ found that *DNAJC3* is related to insulin secretory defects and type 2 diabetes by disrupting the endoplasmic reticulum leads to the functional disturbance of insulin-secreting beta cells. Synofzik et al.³⁷ discovered that the loss of *DNAJC3* causes diabetes mellitus and widespread neurodegeneration. Loss-of-function *DNAJC3* mutations result in a monogenic recessive diabetes mellitus in humans.³⁷ Bublitz et al.³⁸ also reported a case of *DNAJC3* mutations in a diabetic patient with multisystem neuropathy in addition to hypothyroidism. Based on their conclusions, it is reasonable to believe that *DNAJC3* mutations are likely to be responsible for diabetes and related disorders. In a study by Laybutt et al.,³⁹ increased production of *DNAJC3* was observed in human pancreatic sections in type 2 diabetics and endoplasmic reticulum stress arises in type 2 diabetics because *DNAJC3* is associated with potential beta cell failure. Relationships involving *DNAJC3* and diabetic DE should be investigated in future studies.

Phosphoglycerate kinase 1 (*PGK-1*) is an X-linked gene encoding 3-phosphoglycerate kinase, which is an essential enzyme for glycolysis in all cells.⁴⁰ *PGK-1* also acts as a polymerase alpha cofactor protein (primer recognition protein).⁴¹ Research by Fu et al.⁴² found that *PGK1* is a potential survival marker and promoter of malignant invasion in breast cancer due to its role in regulating HIF-1 α -mediated epithelial-mesenchymal transition. They believed that high expression of *PGK1* is associated with poor prognosis in breast cancer. *PGK1* and HIF-1 α form a positive feed-forward loop to stimulate breast cancer progression and metastasis, and thus *PGK1* is expected to become a promising biomarker, as well as a means of targeted therapy for breast cancer.⁴² Zhou et al.⁴³ proposed that *PGK1* mediates DNA repair and methylation through the HSP90/ERK pathway and ultimately enhances resistance to cisplatin, which is expected to become a target for chemotherapy in endometrial cancer. However, no reports on the role of *PGK1* in DE were retrieved from literature, and

so, any correlation between them requires further research.

Our study also has certain shortcomings. First, the numbers of cases included was relatively small. Second, the type of diabetes in adults was different from that in children. To add, the number of proteins detected in tear samples from children was much higher than that in samples from adults. In this study, the adult group and the child group were analyzed separately different proteomics platforms, including differentially expressed protein analysis (DEP), DEP functional analysis. Different instrument platforms may affect the number of identified proteins, but this would not affect the functional bias of the analysis of differential proteins. Finally, cooperation was poor when tear secretion test was performed in children, resulting in no statistical difference for results involving SIt.

Conclusions

In summary, the different protein expression profiles in the tears of adults and children with diabetic DE were related to biological processes involving immune and inflammatory responses. Moreover, the similarity of KEGG pathway analysis results for differentially expressed proteins in tear samples from both adults and children was very high. Therefore, the pathogenesis of the two populations studied here has some common grounds. The hub genes that we identified may be potential new biomarkers for the diagnosis of diabetic DE; further research is needed in the future. In view of the background of our study and the advantages of WGCNA, similar results can also be analyzed in other proteomics research.

Acknowledgments

The authors thank all the staff and participants in this study for their valuable skill and support.

Supported by Chinese National Nature Science Foundation (Project number 82071012), Shanghai engineering research center of precise diagnosis and treatment of eye diseases, Shanghai, China (Project No. 19DZ2250100). The Project of Shanghai Shen Kang Hospital Development Centre (Grant No. SHDC2018110, SHDC2020CR30538). The Science and Technology Commission of Shanghai Municipality (Project No. 20DZ1100200). Shanghai Municipal Education Commission—Gaofeng Clinical Medicine Grant Support (Project No. 20172022). Shanghai

General Hospital, Clinical Research (Project No. CTCCR-2018Z01).

Disclosure: **X. Zou**, None; **S. Wang**, None; **P. Zhang**, None; **L. Lu**, None; **H. Zou**, None

* ZX and WS contributed equally to this work.

References

1. Yazdani-Ibn-Taz MK, Han MM, Jonuscheit S, Collier A, Nally JE, Hagan S. Patient-reported severity of DE and quality of life in diabetes. *Clin Ophthalmol.* 2019;13:217–224.
2. Manaviat MR, Rashidi M, Afkhami-Ardekani M, Shoja MR. Prevalence of DE syndrome and diabetic retinopathy in type 2 diabetic patients. *BMC Ophthalmol.* 2008;8:10.
3. Jin J, Chen LH, Liu XL, Jin GS, Lou SX, Fang FN. Tear film function in non-insulin dependent diabetics [in Chinese]. *Zhong Hua Yan Ke Za Zhi.* 2003;39:10–13.
4. Dogru M, Katakami C, Inoue M. Tear function and ocular surface changes in noninsulin-dependent diabetes mellitus. *Ophthalmology.* 2001;108:586–592.
5. Seifart U, Stempel L. The DE and diabetes mellitus. *Ophthalmologie.* 1994;91:235–239.
6. Hom M, De Land P. Self-reported DEs and diabetic history. *Optometry.* 2006;77:554–558.
7. Zou X, Lu L, Xu Y, et al. Prevalence and clinical characteristics of DE disease in community-based type 2 diabetic patients: the Beixinjing eye study. *BMC Ophthalmol.* 2018;18:117.
8. Patterson C, Guariguata L, Dahlquist G, Soltész G, Ogle G, Silink M. Diabetes in the young—a global view and worldwide estimates of numbers of children with type 1 diabetes. *Diabetes Res Clin Pract.* 2014;103(2):161–175.
9. Ogurtsova K, Da Rocha Fernandes JD, Huang Y, et al. Diabetes Atlas: IDF. Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract.* 2017;128:40–50.
10. Wang Shanshan, Jia Yan, Li Tao, et al. DE disease is more prevalent in children with diabetes than in those without diabetes. *Curr Eye Res.* 2019;44(12):1299–1305.
11. Dogru M, Katakami C, Inoue M. Tear function and ocular surface changes in noninsulin-dependent diabetes mellitus. *Ophthalmology.* 2001;108:586–592.
12. Song XJ, Li DQ, Farley W, et al. Neurturin-deficient mice develop dry eye and eratocon-

- junctivitis sicca. *Invest Ophthalmol Vis Sci.* 2003;44:4223–4229.
13. Gipson IK, Argueso P. Role of mucins in the function of the corneal and conjunctival epithelia. *Int Rev Cytol.* 2003;231:1–49.
 14. Zhou L, Liu DN. Proteomics of human tears: towards clinical applications [J]. *Chin J Exp Ophthalmol.* 2015;33(5):385–388.
 15. Zhou L, Liu DN. Paying attention to proteomics of human tear: clinical significance and application in ocular surface diseases [J]. *Chin J Exp Ophthalmol.* 2016;34(2):97–102.
 16. Zhou L, Beuerman RW, Chan CM, et al. Identification of tear fluid biomarkers in DE syndrome using iTRAQ quantitative proteomics. *J Proteome Res.* 2009;8:4889–4905.
 17. Li B, Sheng M, Xie L, et al. Tear proteomic analysis of patients with type 2 diabetes and DE syndrome by two-dimensional nanoliquid chromatography coupled with tandem mass spectrometry. *Invest Ophthalmol Vis Sci.* 2014;55:177–186.
 18. Du J, Shi H, Lu Y, et al. Tagging single nucleotide polymorphisms in the PPAR-c and RXR-a gene and type 2 diabetes risk: a case-control study of a Chinese Han population. *J Biomed Res.* 2011;25:33–41.
 19. Department of Ophthalmology, Chinese Academy of Medical Sciences. Consensus of clinical experts on DE (2013) [J]. *Clin J Ophthalmol.* 2013;49(1):73–75.
 20. Chen L, Yuan L, Qian K, et al. Identification of biomarkers associated with pathological stage and prognosis of clear cell renal cell carcinoma by co-expression network analysis. *Front Physiol.* 2018;9:399.
 21. Boehm N, Riechardt A.I, Wiegand M, Pfeiffer N, Grus F.H. Proinflammatory cytokine profiling of tears from DE patients by means of antibody microarrays. *Invest Ophthalmol Vis Sci.* 2011;52(10):7725–7730.
 22. Massingale ML, Li X, Vallabhajosyula M, Chen D, Wei Y, Asbell PA. Analysis of inflammatory cytokines in the tears of DE patients. *Cornea.* 2009;28(9):1023–1027.
 23. Csoz E, Boross P, Csutak A, et al. Quantitative analysis of proteins in the tear fluid of patients with diabetic retinopathy. *J Proteom.* 2012;75:2196–2204.
 24. Hagan S, Martin E, Enriquez-de-Salamanca A. Tear fluid biomarkers in ocular and systemic disease: Potential use for predictive, preventive and personalised medicine. *EPMA J.* 2016;7:15.
 25. Nash JA, Ballard TN, Weaver TE, Akinbi HT. The peptidoglycan-degrading property of lysozyme is not required for bactericidal activity in vivo. *J Immunol.* 2006;177:519–526.
 26. Srinivasan S, Thangavelu M, Zhang L, Green KB, Nichols KK. TRAQ quantitative proteomics in the analysis of tears in DE patients. *Invest Ophthalmol Vis Sci.* 2012;53:5052–5059.
 27. Versura P, Bavelloni A, Grillini M, Fresina M, Campos EC. Campos1 diagnostic performance of a tear protein panel in early DE. *Molecular Vision.* 2013;19:1247–1257.
 28. Versura P, Giannaccare G, Vukatana G, Mulè R, Malavolta N, Campos EC. Predictive role of tear protein expression in the early diagnosis of Sjögren's syndrome. *Ann Clin Biochem.* 2018;55(5):561–570.
 29. Hanstock HG, Edwards JP, Walsh NP. Tear lactoferrin and lysozyme as clinically relevant biomarkers of mucosal immune competence. *Front Immunol.* 2019;10:1178.
 30. Zhang X, M VJ, Qu Y, et al. DE management: targeting the ocular surface microenvironment. *Int J Mol Sci.* 2017;18:1398.
 31. Gong FY, Zhang SJ, Deng JY, et al. Zinc- α 2-glycoprotein is involved in regulation of body weight through inhibition of lipogenic enzymes in adipose tissue. *Int J Obes.* 2009;33:1023–1030.
 32. Balaž M, Ukropcova B, Kurdiova T, et al. Improved adipose tissue metabolism after 5-year growth hormone replacement therapy in growth hormone deficient adults: the role of zinc- α 2-glycoprotein. *Adipocyte.* 2014;4:113–122.
 33. Balaz M, Vician M, Janakova Z, et al. Subcutaneous adipose tissue zinc- α 2-glycoprotein is associated with adipose tissue and whole-body insulin sensitivity. *Obesity (Silver Spring).* 2014;22(8):1821–1829.
 34. Liu M, Zhu H, Dai Y, et al. Zinc- α 2-glycoprotein is associated with obesity in Chinese people and HFD-induced obese mice. *Front Physiol.* 2018;9:62.
 35. Gething MJ. Role and regulation of the ER chaperone BiP. *Semin Cell Dev Biol.* 1999;10:465–472.
 36. Lu H, Yang Y, Allister EM, Wijesekara N, Wheeler MB. The identification of potential factors associated with the development of type 2 diabetes: a quantitative proteomics approach. *Mol Cell Proteomics.* 2008;7(8):1434–1451.
 37. Synofzik M, Haack TB, Kopajtich R, et al. Absence of BiP co-chaperone DNAJC3 causes diabetes mellitus and multisystemic neurodegeneration. *Am J Hum Genet.* 2014;95(6):689–697.
 38. Bublitz SK, Alhaddad B, Synofzik M, et al. Expanding the phenotype of DNAJC3 mutations:

- a case with hypothyroidism additionally to diabetes mellitus and multisystemic neurodegeneration. *Clin Genet.* 2017;92(5):561–562.
39. Laybutt DR, Preston AM, Akerfeldt MC, et al. Endoplasmic reticulum stress contributes to beta cell apoptosis in type 2 diabetes. *Diabetologia.* 2007;50(4):752–763.
 40. Bollong MJ, Lee G, Coukos JS, et al. A metabolite-derived protein modification integrates glycolysis with KEAP1-NRF2 signaling. *Nature.* 2018;562(7728):600–604.
 41. Jindal HK, Vishwanatha JK. Functional identity of a primer recognition protein as phosphoglycerate kinase. *J Biol Chem.* 1990;265(12):6540–6543.
 42. Fu D, He C, Wei J, et al. PGK1 is a potential survival biomarker and invasion promoter by regulating the HIF-1 α -mediated epithelial-mesenchymal transition process in breast cancer. *Cell Physiol Biochem.* 2018;51(5):2434–2444.
 43. Zhou JW, Tang JJ, Sun W, Wang H. PGK1 facilitates cisplatin chemoresistance by triggering HSP90/ERK pathway mediated DNA repair and methylation in endometrial endometrioid adenocarcinoma. *Mol Med.* 2019;25(1):11.