



## OPEN ACCESS

## EDITED BY

Giulia Bernardini,  
University of Siena, Italy

## REVIEWED BY

Davide Dalfovo,  
University of Trento, Italy  
Cheng-wei Lu,  
First Affiliated Hospital of Jilin  
University, China

## \*CORRESPONDENCE

Li-Wei Zhang,  
✉ drzhangliwei@163.com  
Hong-Lei Zhang,  
✉ hlzhang2014@163.com  
Hai Liu,  
✉ herryhai@ynu.edu.cn

<sup>†</sup>These authors have contributed equally to this work and share first authorship

RECEIVED 08 November 2024

ACCEPTED 25 November 2024

PUBLISHED 11 December 2024

## CITATION

Zhang L-W, Yang J, Jiang H-W, Yang X-Q, Chen Y-N, Ying W-D, Deng Y-L, Zhang M-h, Liu H and Zhang H-L (2024) Identification of biomarkers and immune microenvironment associated with pterygium through bioinformatics and machine learning. *Front. Mol. Biosci.* 11:1524517. doi: 10.3389/fmolb.2024.1524517

## COPYRIGHT

© 2024 Zhang, Yang, Jiang, Yang, Chen, Ying, Deng, Zhang, Liu and Zhang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Identification of biomarkers and immune microenvironment associated with pterygium through bioinformatics and machine learning

Li-Wei Zhang<sup>1\*†</sup>, Ji Yang<sup>1†</sup>, Hua-Wei Jiang<sup>1,2†</sup>, Xiu-Qiang Yang<sup>3†</sup>, Ya-Nan Chen<sup>1</sup>, Wei-Dang Ying<sup>1</sup>, Ying-Liang Deng<sup>1</sup>, Min-hui Zhang<sup>1</sup>, Hai Liu<sup>1\*</sup> and Hong-Lei Zhang<sup>4\*</sup>

<sup>1</sup>Department of Ophthalmology, The Affiliated Hospital of Yunnan University, Second People's Hospital of Yunnan Province, The Eye Hospital of Yunnan Province, The Eye Disease Clinical Medical Research Center of Yunnan Province, The Eye Disease Clinical Medical Center of Yunnan Province, Kunming, China, <sup>2</sup>Department of Ophthalmology, Wuhan Children's Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, <sup>3</sup>Department of Ophthalmology, Diqing Tibetan Autonomous Prefecture People's Hospital, Diqing, Yunnan, China, <sup>4</sup>Center for Scientific Research, Yunnan University of Chinese Medicine, Kunming, China

**Background:** Pterygium is a complex ocular surface disease characterized by the abnormal proliferation and growth of conjunctival and fibrovascular tissues at the corneal-scleral margin. Understanding the underlying molecular mechanisms of pterygium is crucial for developing effective diagnostic and therapeutic strategies.

**Methods:** To elucidate the molecular mechanisms of pterygium, we conducted a differential gene expression analysis between pterygium and normal conjunctival tissues using high-throughput RNA sequencing. We identified differentially expressed genes (DEGs) with statistical significance (adjust  $p < 0.05$ ,  $|\log_{2}FC| > 1$ ). Enrichment analyses were performed to assess the biological processes and signaling pathways associated with these DEGs. Additionally, we utilized weighted correlation network analysis (WGCNA) to select module genes and applied Random Forest (RF) and Support Vector Machine (SVM) algorithms to identify pivotal feature genes influencing pterygium progression. The diagnostic potential of these genes was validated using external datasets (GSE2513 and GSE51995). Immune cell infiltration analysis was conducted using CIBERSORT to compare immune cell populations between pterygium and normal conjunctival tissues. Quantitative PCR (qPCR) was used to confirm the expression levels of the identified feature genes. Furthermore, we identified key miRNAs and candidate drugs targeting these feature genes.

**Results:** A total of 718 DEGs were identified in pterygium tissues compared to normal conjunctival tissues, with 254 genes showing upregulated expression and 464 genes exhibiting downregulated expression. Enrichment analyses revealed that these DEGs were significantly associated with inflammatory processes and key signaling pathways, notably leukocyte migration and IL-17 signaling. Using WGCNA, RF, and SVM, we identified KRT10 and NGEF as pivotal feature genes influencing pterygium

progression. The diagnostic potential of these genes was validated using external datasets. Immune cell infiltration analysis demonstrated significant differences in immune cell populations between pterygium and normal conjunctival tissues, with an increased presence of M1 macrophages and resting dendritic cells in pterygium samples. qPCR analysis confirmed the elevated expression of KRT10 and NGEF in pterygium tissues.

**Conclusion:** Our findings emphasize the importance of gene expression profiling in unraveling the pathogenesis of pterygium. The identification of pivotal feature gene KRT10 and NGEF provide valuable insights into the molecular mechanisms underlying pterygium progression.

#### KEYWORDS

pterygium, RNA sequencing, bioinformatics, WCGNA, machine learning, immunoinfiltration

## Introduction

Pterygium, a multiple ocular surface disease, arises from abnormal proliferation and growth of conjunctival and fibrovascular tissues at the corneal scleral margin, subsequently invading the cornea (Singh et al., 1988; Guimares et al., 2022). This condition significantly impacts visual health, altering appearance, destabilizing the tear film, and inducing discomfort. Furthermore, pterygium severely affects corneal astigmatism, leading to decreased vision, and in extreme cases, can cause eye movement disorders and blindness (Chu et al., 2020). The prevalence of pterygium varies globally, from 0.074% in Saudi Arabia to 53% in Taiwan, China. Combined with global data (Alqahtani, 2013; Rezvan et al., 2018), the overall prevalence of pterygium is 12% (Rezvan et al., 2018). Numerous studies, both domestic and international, have established a strong correlation between ultraviolet radiation and the onset as well as progression of pterygium (Moran and Hollows, 1984; Modenese et al., 2023). There have also been studies on the relationship between surface air pollution and the publication of pterygium (Lu et al., 2023; Taylor, 2013). Epidemiological research further supports this finding, indicating a higher incidence among rural populations and individuals who engage in frequent outdoor activities, potentially due to increased exposure to ultraviolet radiation. The prevalence and postsurgical recurrence rate of pterygium are notably high, posing a significant threat to human visual health, as untreated cases may result in blindness (Prabhasawat et al., 1997; Chang et al., 2023). Surgical resection is the main treatment of pterygium. However, there are some serious complications in this treatment, such as postoperative recurrence and secondary infection (Hacıoğlu and Erdöl, 2017). Ocular demodicosis, environment pollution and ultraviolet radiation are risk factors for pterygium recurrence (Huang et al., 2013).

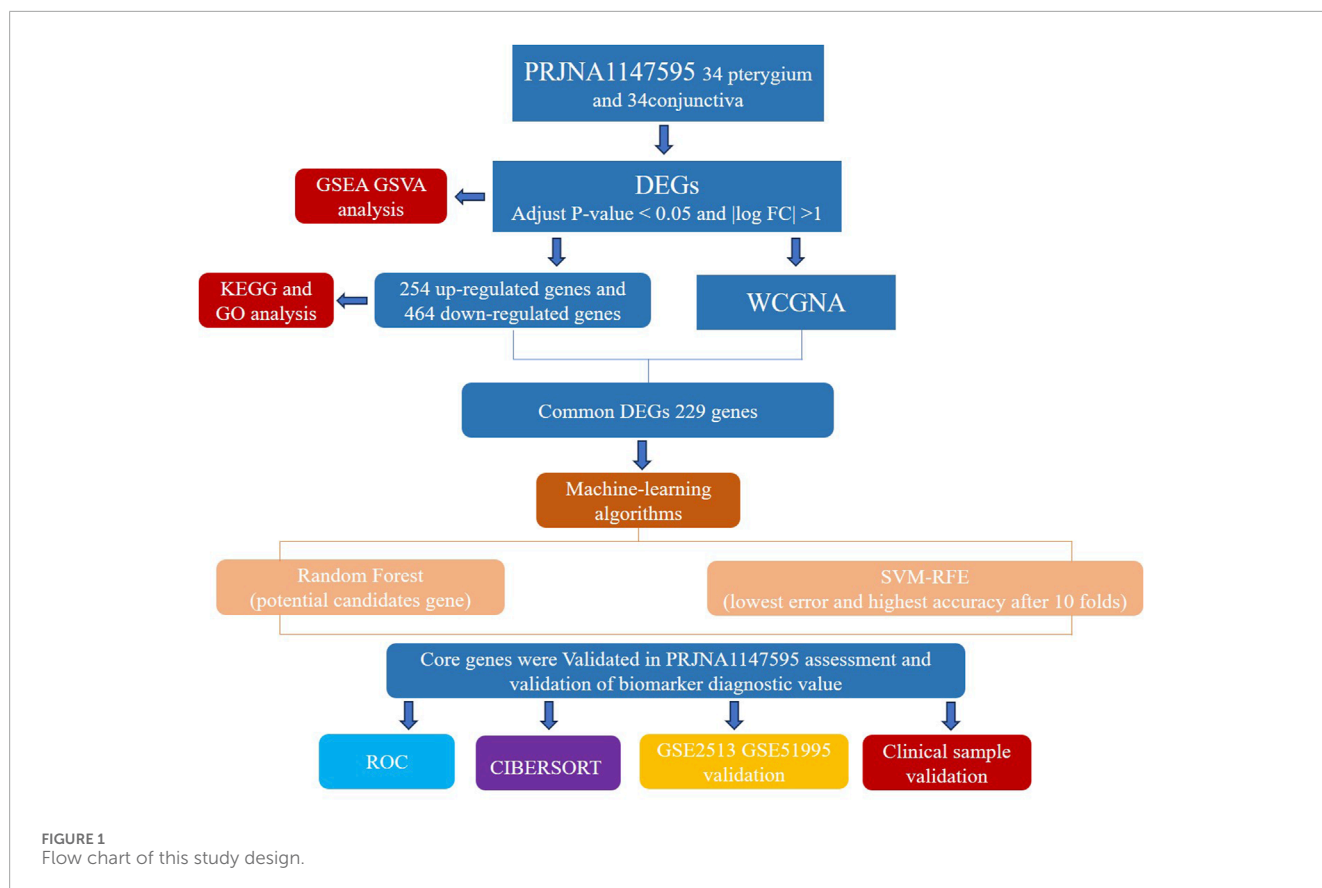
Current research on pterygium has largely focused on its clinical manifestations and surgical management. However, the molecular mechanisms underlying pterygium development remain poorly understood, creating a gap in knowledge that could inform future therapeutic strategies. In addition, there are also immune inflammation, cell proliferation and apoptosis disorders and lipid metabolism disorders and other related mechanisms involved in the development of pterygium (Shahraki et al., 2021; Chalkia et al., 2019; Kalogeropoulos et al., 2020; Rubeshkumar et al., 2020). Recent

studies have highlighted the importance of inflammatory processes in pterygium pathology, revealing that the disease is associated with a heightened inflammatory response and altered immune cell infiltration within the conjunctival tissue (Labbé et al., 2010). This suggests that targeting inflammatory pathways may offer new avenues for intervention.

Currently, there is a lack of consensus regarding the risk factors and clear molecular mechanisms underlying pterygium. Our study aims to contribute to a deeper understanding of pterygium by exploring these risk factors and potential molecular mechanisms. High-throughput RNA sequencing technologies have emerged as powerful tools for investigating gene expression profiles associated with various diseases, including pterygium. Studies employing RNA sequencing have demonstrated significant differences in gene expression between pterygium and normal conjunctival tissues, providing insights into the molecular pathways involved in pterygium development (Yoon et al., 2023). Furthermore, machine learning techniques such as random forests (RF) and support vector machines (SVM) have been applied to identify key gene signatures associated with pterygium, enhancing the predictive power of genomic data (Uddin et al., 2019). Figure 1 depicts our research protocol.

Therefore, an in-depth exploration of the molecular mechanisms underlying pterygium pathogenesis holds significant importance for disease prevention, treatment, and the identification of potential drug targets. Given the complexity and uncertainty surrounding the risk factors and molecular mechanisms of pterygium, we aim to screen for differentially expressed genes in pterygium and bulbar conjunctiva tissues through transcriptomics research. Additionally, we seek to identify module gene sets associated with the pterygium phenotype and analyze their biological functions. On a cellular level, we will assess differences in immune cell infiltration between pterygium and bulbar conjunctiva groups. This comprehensive approach not only advances our understanding of the molecular mechanisms of pterygium but also offers new insights into its transcription profile. Ultimately, this research strives to identify novel treatment methods and therapeutic targets, aiming to optimize disease prevention, treatment, and reduce the postoperative recurrence rate, thereby carrying substantial clinical significance.





## Materials and methods

### Patients and specimens

The research protocol was approved by the Ethics Review Committee of the Affiliated Hospital of Yunnan University, and all study participants gave written informed consent (approval number: 2022198). The study was carried out in accordance with the Declaration of Helsinki. All surgical procedures were performed under local anesthesia by the same surgeon. Thirty-four patients underwent elective pterygium surgery, including 5 males and 29 females. The average age was 56.8 years and the average duration of illness was 4.19 years (Supplementary Figure S1). During the surgical procedure, the affected pterygium tissue was removed alongside the surrounding excess loose conjunctiva. The integrity of the total RNA extracted from both pterygium and bulbar conjunctiva samples was evaluated using an Agilent 2,100 bioanalyzer. Subsequently, rRNA was eliminated from the total RNA to isolate the sample mRNA. This mRNA was then subjected to random fragmentation using divalent cations in NEB Fragmentation Buffer, followed by chain-specific fragmentation for mRNA construction. Initial library quantification was carried out with a Qubit2.0 Fluorometer, and the library was subsequently diluted to a concentration of 1.5 ng/ul. The size of the library inserts was determined using an Agilent 2,100 bioanalyzer, and QRT-PCR was utilized to accurately quantify the library's effective concentration, which needed to exceed 2 nM. After assessing the quality of the genomic DNA, it was fragmented

through mechanical interruption (ultrasound). The fragmented DNA then underwent purification, end-repair, 3' end adenylation, ligation to a sequencing adapter, and size-selection using agarose gel electrophoresis. The resulting polymerase chain reaction (PCR) product was amplified to generate the sequencing library. Sequencing was conducted on the Illumina NovaSeq 6,000 platform with a read length of 150 bp. Through quality control, trimming, deduplication, and alignment of the original fastq data by a high-throughput sequencing service provider, we obtained the gene expression matrix of transcriptome sequencing. The data presented in this publication has been archived in NCBI's Sequence Read Archive (SRA) database (accession number: PRJNA1147595, URL: <https://dataview.ncbi.nlm.nih.gov/object/PRJNA1147595?reviewer=3k3nnr66jke53qo77la1sbt91b>). Supplementary datasets can be found in NCBI's Gene Expression Omnibus (GEO) (accession numbers: GSE2513, URL: <https://www.ncbi.nlm.nih.gov/geo/>). GSE2513 includes four conjunctival samples and eight matched pairs of pterygium and control conjunctival samples (Wong et al., 2006). GSE51995 includes four conjunctival samples and four pairs of pterygium and control conjunctival samples (Hou et al., 2014).

### Detection of differentially expressed genes (DEGs) in conjunctiva versus pterygium samples

To detect DEGs, we employed the "DESeq2" R package. The DESeq2 package was used for difference analysis of the original

Counts matrix, and the analysis was carried out according to the standard process. The Variance Stabilizing Transformations (VST) method provided by package DESeq2 was used to Normalize the original Counts matrix. A Wilcoxon rank sum test was conducted to assess gene expression differences between the conjunctiva and pterygium samples. Genes were considered significantly differentially expressed if they met the criteria of adjusted  $p < 0.05$  and  $|\log_2FC| > 1$ . For data visualization, we utilized the “pheatmap”, “ggpubr” and “ggplot2” R packages.

## Functional annotation and enrichment analysis of DEGs

To analyze the DEGs, we employed the R tool known as “clusterProfiler.” (Yu et al., 2012) Using the “clusterProfiler” R package, we carried out both Gene Ontology (GO) (Ashburner et al., 2000) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses (Ogata et al., 1999). Gene set enrichment analysis (GSEA) was conducted on all genes (previously ranked based on their  $\log_2FC$  between analyzed groups) using the cluster profiler package. Enrichment was considered significant if the nominal false discovery rate (FDR) was  $< 0.25$  and the P-value was  $< 0.05$ , referencing the ‘c2.cp.all.v2022.1.Hs.symbols.gmt’ gene set. Additionally, we conducted Gene Set Variation Analysis (GSVA) by utilizing the “GSEABase” and “GSVA” R packages (Hänzelmann et al., 2013) and referencing the ‘h.all.v2023.2.Hs.symbols.gmt’ gene set. These packages offer utilities for evaluating the enrichment or variability of gene sets within gene expression data, thereby enabling a thorough examination of pathway or gene set activities across different samples. P-values were calculated with the Benjamini–Hochberg method, and the terms with P-values  $< 0.05$  were considered to be significant.

## Weighted gene co-expression network analysis

For the construction of expression networks within the PRJNA1147595 dataset, we utilized the weighted gene co-expression network analysis (WGCNA) approach. This method was steered by adhering to the scale-free topology criterion, ensuring that the resultant network displayed a scale-free structure characterized by a power-law distribution of node connections. To ascertain the appropriate soft threshold power and establish adjacencies, we employed the “pickSoftThreshold” function from the WGCNA package. Once determined, the adjacency matrix was transformed into a topological overlap matrix (TOM). Hierarchical clustering analysis was then conducted based on the dissimilarity derived from the TOM.

To delineate co-expressed gene modules, we adopted the dynamic tree cutting technique, setting a minimum module size of 50. This approach allowed us to categorize genes into distinct modules according to their expression patterns. Subsequently, we evaluated the relationship between these gene modules and the trait of interest, pterygium, by considering both gene significance (GS) and module membership (MM) values. GS reflects the correlation between gene expression and pterygium, whereas MM measures the

extent of co-expression within a given module. By integrating GS and MM values, we pinpointed the key modules linked to pterygium, suggesting their potential functional importance in the context of the disease.

## Identification of hub genes based on machine learning methods

The gene sets of the significant difference modules analyzed by WGCNA were intersected with the expression difference genes obtained by DESeq2 analysis. To identify the key hub genes, two distinct machine learning algorithms were employed: support vector machine (SVM) and random forest (RF). The SVM algorithm was executed using the “e1071” package, offering a comprehensive toolkit for SVM model training and classification. The RF algorithm was applied using the “randomForest” package, where the error rate was computed across 1 to 500 trees. The optimal tree count was determined by selecting the value that minimized the error rate while ensuring stability. Subsequently, the RF classifier was leveraged to compute feature importance scores. This selection utilized the Gini coefficient method to pinpoint the most pertinent genes. The intersection of the top five feature genes from both the SVM and RF models was used to pinpoint crucial hub genes involved in the disease process.

Utilizing the R programming language, receiver operating characteristic (ROC) curves were generated for the hub genes. To evaluate the distinctive capabilities between pterygium tissues and conjunctiva tissues, the area under the curve (AUC) was calculated for the respective ROC curves of these hub genes.

## Analysis of immune cells infiltration

According to the CIBERSORT deconvolution algorithm (Newman et al., 2015), we analyzed the infiltration of 22 kinds of immune cells in pterygium and conjunctival tissues, and compared whether there was a significant difference between the two groups. To assess the degrees of immune cell infiltration in both conjunctiva and pterygium samples, we utilized the CIBERSORT technique. This method enabled us to measure gene expression levels quantitatively and ascertain the infiltration levels among patients. The immune cell populations encompassed both immune-stimulating and immune-inhibiting cells. To examine the relationships between these immune cells, we conducted a correlation analysis using the Spearman coefficient and depicted the results via correlation heatmaps. At the same time, we also analyzed the correlation between hub gene and immune cells. We considered findings statistically significant if they met the threshold of  $p < 0.05$ .

## Validation of the mRNA expression of hub genes by quantitative real-time PCR

**Patient Sample Preparation.** There were 4 primary pterygium specimens obtained in Affiliated Hospital of Yunnan University during pterygium surgery. The control conjunctival tissues were

derived from healthy conjunctival tissues on the temporal side of the surgical eye of the same patient, with a size of 1.5 mm × 1.5 mm.

**Tissue RNA Extraction and RT-qPCR.** All specimens were collected during the operation and immediately placed in RNA protective agent, stored at -20°C for usage. RNA was extracted from 80 mg specimen tissue using TRIZOL reagent. The relative expression of hub genes was measured using a reverse kit (HiScript<sup>®</sup> III RT SuperMix for qPCR (+gDNA wiper)) and ChamQ<sup>™</sup> Universal SYBR<sup>®</sup> qPCR Master Mix (Vazyme company, Nanjing, China) on CFX96 fluorescence qPCR instrument. And the PCR reaction was carried as follows: predenaturation at 95°C for 30 s, denaturation at 95°C for 3–10 s, and annealing extension at 60°C for 10–30 s, for 40 cycles. Using GAPDH as internal references, the relative expression of genes was calculated by the 2- $\Delta$ Ct method.

## Identification of key miRNAs and potential drug candidates

The identification of key miRNAs targeting the two characteristic genes, as well as the screening of potential drug candidates, was conducted through a rigorous and comprehensive process. We utilized the Targetscan database (version 7.2, accessed via the Enrichr platform at <http://amp.pharm.mssm.edu/Enrichr/>) to forecast miRNAs that could potentially target the selected genes (Agarwal et al., 2015). This involved inputting the gene symbols into the Targetscan interface and retrieving a list of predicted miRNAs based on the presence of conserved target sites within the 3'UTRs of the genes. The predicted miRNAs were then filtered based on their conservation across species, expression levels in relevant tissues, and previous reports of their involvement in related biological processes. Subsequently, to identify potential drug candidates that could modulate the expression or activity of these genes, we accessed the DSigDB database through the Enrichr platform. This database integrates drug signatures from various sources, allowing us to search for drugs that have been reported to affect the expression or activity of the selected genes.

## Statistical analysis

The statistical analyses were performed using R software (version 4.2.1). All *p* values were two-sided, and significance was indicated by *p* < 0.05.

## Results

### Identification of the DEGs

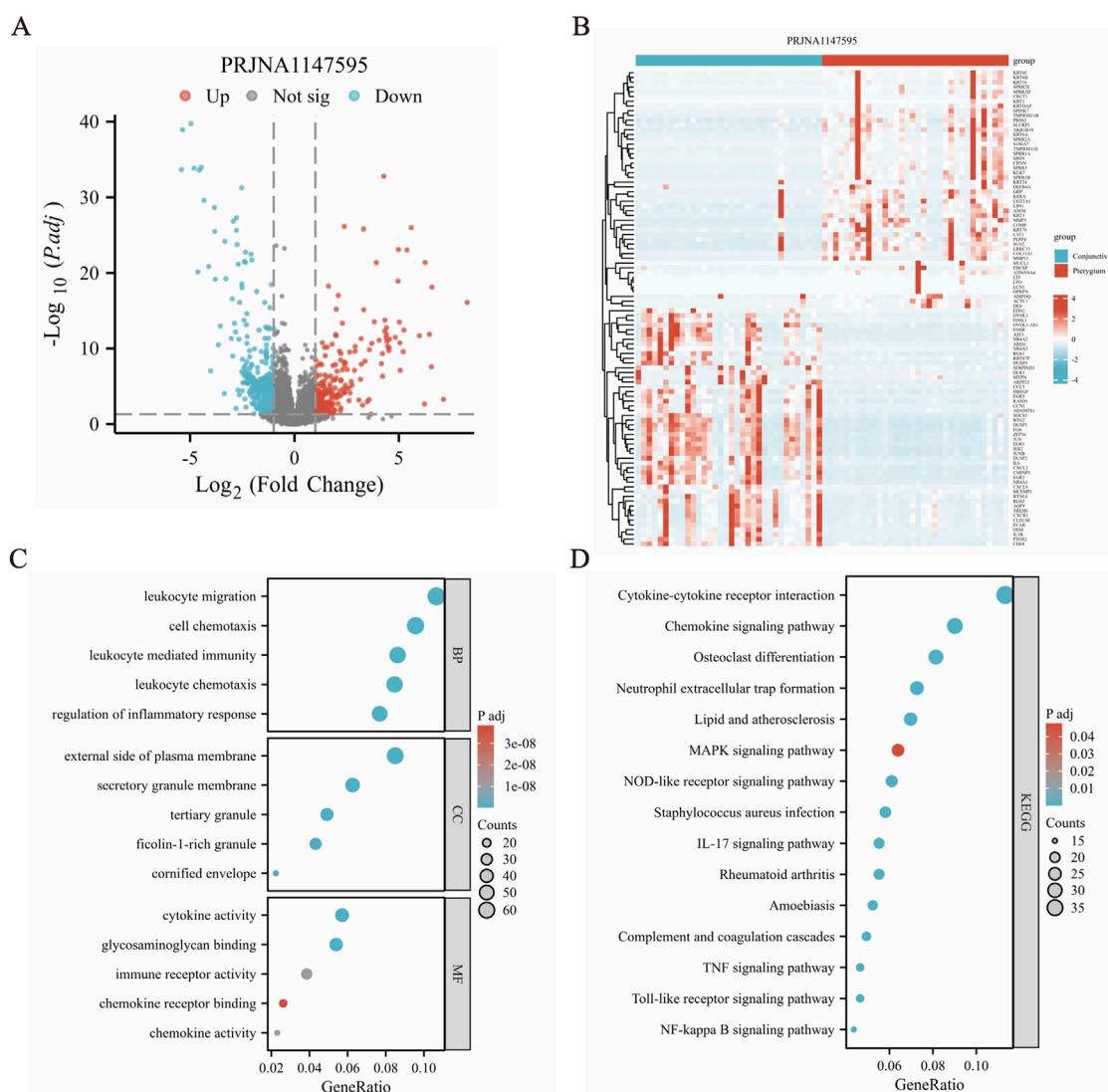
Based on the RNA-Seq data, a total of 718 DEGs were obtained, and these genes reached the threshold adjust *p*-value of <0.05 (Supplementary Table S1). Among these DEGs, 254 transcripts were upregulated in the disease group against the health one, while 464 transcripts were found to decrease in pterygium tissues (Figure 2A). In addition, the relative expression levels are shown in a heatmap plot and the top 50 upregulated and downregulated genes were shown using a volcano plot (Figure 2B).

### Gene function analysis for the DEGs

GO and KEGG pathway enrichment analyses were conducted to assess the function of these 718 DEGs. The clusterProfiler package was used for GO and KEGG pathway enrichment analysis to find potential biological pathways. The top 5 BP (biological process), CC (cellular component) and MF (molecular function) terms with the lowest *p*-values in each category were selected and visualized using bubble plots (Figure 2C). The DEGs were primarily enriched in BP terms such as “leukocyte migration (GO:0050900)”, “leukocyte chemotaxis (GO:0030595)”, “cell chemotaxis (GO:0060326)”, “leukocyte mediated immunity (GO:0002443)” and “regulation of inflammatory response (GO:0050727)”. Genes were also enriched in CC terms like “external side of plasma membrane (GO:0009897)”, “tertiary granule (GO:0070820)”, “secretory granule membrane (GO:0030667)”, “ficolin-1-rich granule (GO:0101002)” and “cornified envelope (GO:0001533)”. Additionally, the DEGs were enriched in MF terms including “cytokine activity (GO:0005125)”, “glycosaminoglycan binding (GO:0005539)”, “immune receptor activity (GO:0140375)”, “chemokine activity (GO:0008009)” and “chemokine receptor binding (GO:0042379)”. In the KEGG pathway enrichment results, many pathways related to inflammation and signal transduction were enriched, such as IL-17 signaling pathway (hsa04657), chemokine signaling pathway (hsa04062), TNF signaling pathway (hsa04668), MAPK signaling pathway (hsa04010) and cytokine-cytokine receptor interaction (hsa04060) (Figure 2D) and so on (Supplementary Table S2).

### GSEA and GSVA analysis

Our reference gene set was ‘c2\_cp.all.v2022.1.Hs.symbols.gmt’. The datasets underwent GSEA enrichment analysis to identify significant enrichment based on the criteria of FDR <0.25 and *p* < 0.05. In gene set enrichment analysis (GSEA), the significant enrichment in upregulated pathways including REACTOME\_MET\_ACTIVATES\_PTK2\_SIGNALING, REACTOME\_COLLAGEN\_FORMATION, REACTOME\_MET\_PROMOTES\_CELL\_MOTILITY, KEGG\_OXIDATIVE\_PHOSPHORYLATION and so on (Figures 3A–D). The significant enrichment in downregulated pathways including WP\_MAPK\_SIGNALING\_PATHWAY (Figure 3E) and so on (Supplementary Table S3). GSVA enrichment analysis was conducted on the PRJNA1147595 dataset, revealing distinct pathways (Supplementary Table S4). By utilizing the gene set variation analysis (GSVA) package and referencing the “h.all.v2023.2.Hs.symbols.gmt” gene set (Hänzelmann et al., 2013), the gene expression matrix data were subjected to GSVA. The GSVA variance analysis was performed using the limma package of R software. Differential pathways were filtered based on an *p*-value <0.05 and |log<sub>2</sub>FC| > 0.25 (Supplementary Table S4). The differential pathways in the PRJNA1147595 dataset encompassed HALLMARK\_TNFA\_SIGNALING\_VIA\_NFKB, HALLMARK\_INFLAMMATORY\_RESPONSE, HALLMARK\_OXIDATIVE\_PHOSPHORYLATION, HALLMARK\_IL6\_JAK\_STAT3\_SIGNALING and HALLMARK\_ALLOGRAFT\_REJECTION (Figure 3F).



**FIGURE 2** Differential gene expression analysis revealed significant variations between conjunctiva and pterygium. **(A)** Volcano diagram shows that there are 254 upregulated genes and 464 downregulated genes in the pterygium group compared with the conjunctival group, a total of 718 differentially expressed genes. **(B)** Cluster heat map shows the expression of the top 50 upregulated and downregulated genes. **(C)** GO functional enrichment analysis results. **(D)** Results of KEGG pathway enrichment analysis.

## WGCNA analysis

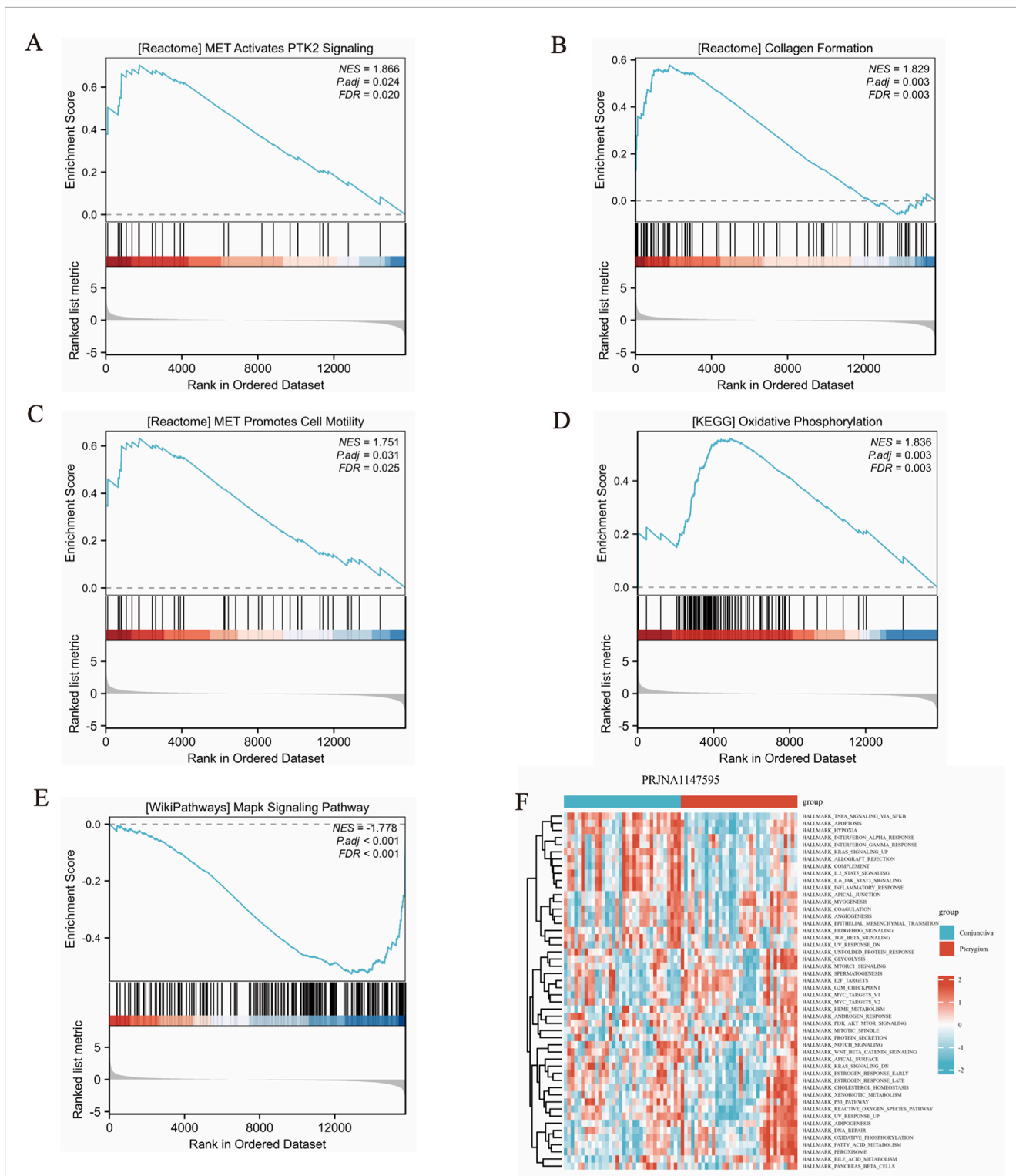
Initially, expression data from 34 pterygium samples and 34 conjunctiva samples were utilized to establish co-expression modules through the WGCNA algorithm. We focused on the top 25% of genes exhibiting the greatest variability for subsequent investigation. Utilizing the “flashClust” package, we conducted cluster analysis on these samples, with the findings illustrated in Figure 4A. At a power value of 11, an independence level of 0.9 was attained, accompanied by an elevated mean connectivity (Figures 4B, C). By employing the dynamic cutting technique, we discerned seven unique gene co-expression modules within pterygium, and a heatmap of the TOM was also generated (Figures 4D, E). The genes within these seven modules were concurrently employed to explore the relationship between module

eigengenes and clinical characteristics (Figure 4F). Notably, 3,956 genes were divided into 7 modules. Among them, salmon module (cor = -0.66, P = 1e-9) and turquoise module (cor = 0.34, p = 0.005) were significantly correlated with pterygium phenotype (Figure 4F).

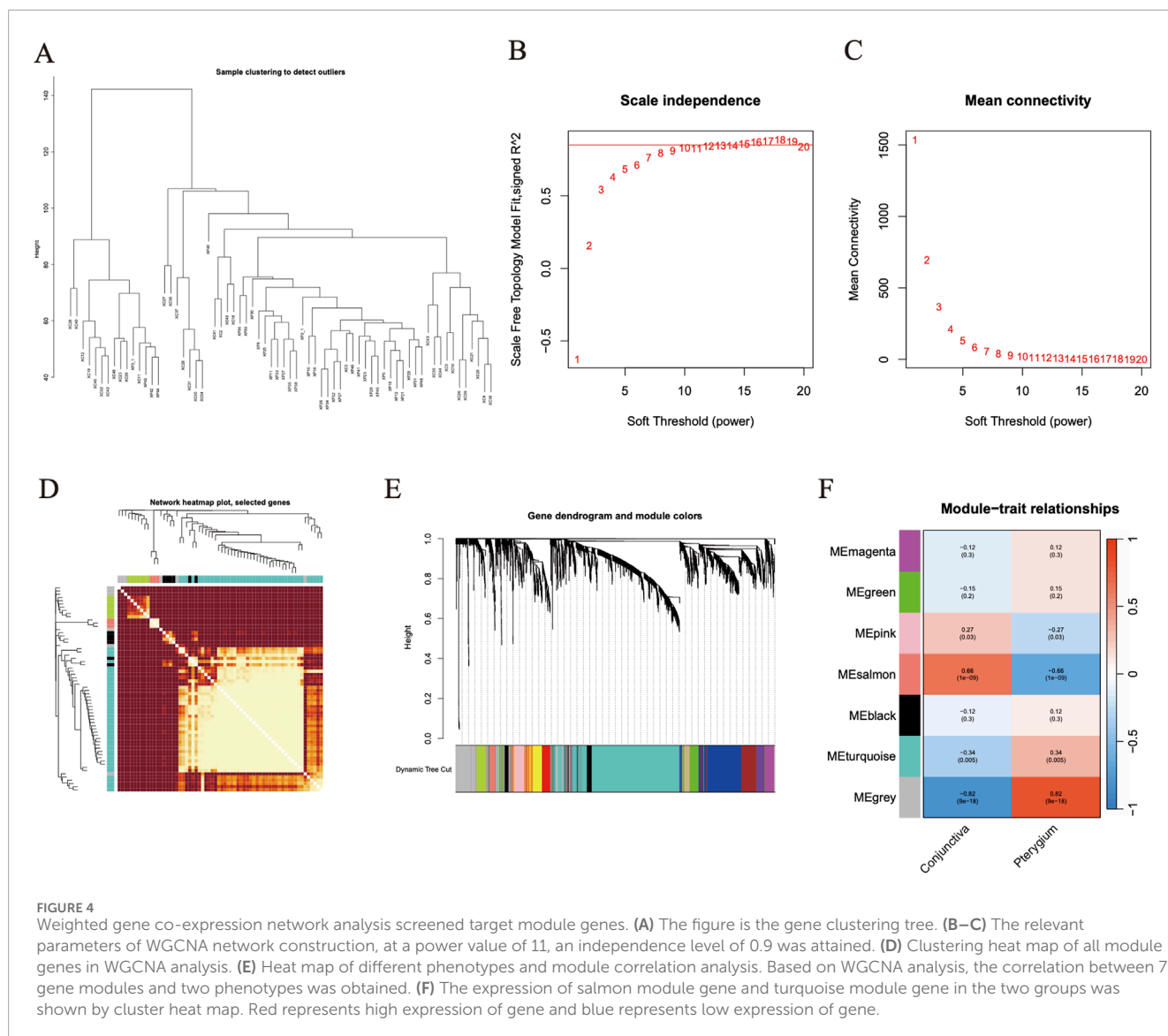
## Identification of hub genes in pterygium using machine learning algorithms

A Venn diagram was constructed to visually represent the relationships among the differentially expressed genes (DEGs) and the salmon and turquoise modules identified through Weighted Gene Co-expression Network Analysis (Figure 5A). Subsequently, two distinct machine learning algorithms—SVM and RF—were leveraged to meticulously screen and pinpoint key signature genes





**FIGURE 3** Gene set enrichment analysis (GSEA) and gene set variation analysis (GSVA) results. **(A–D)** The significant enrichment in upregulated pathways. The members of the line marker pathway gene set appear in the gene ranking list, and the bottom is the rank value distribution of all genes. **(E)** The significant enrichment in downregulated pathway. The results obtained from the GO database and the KEGG database for the pre-defined gene set. **(F)** The differential pathways in the dataset encompassed, showing in cluster heat map.



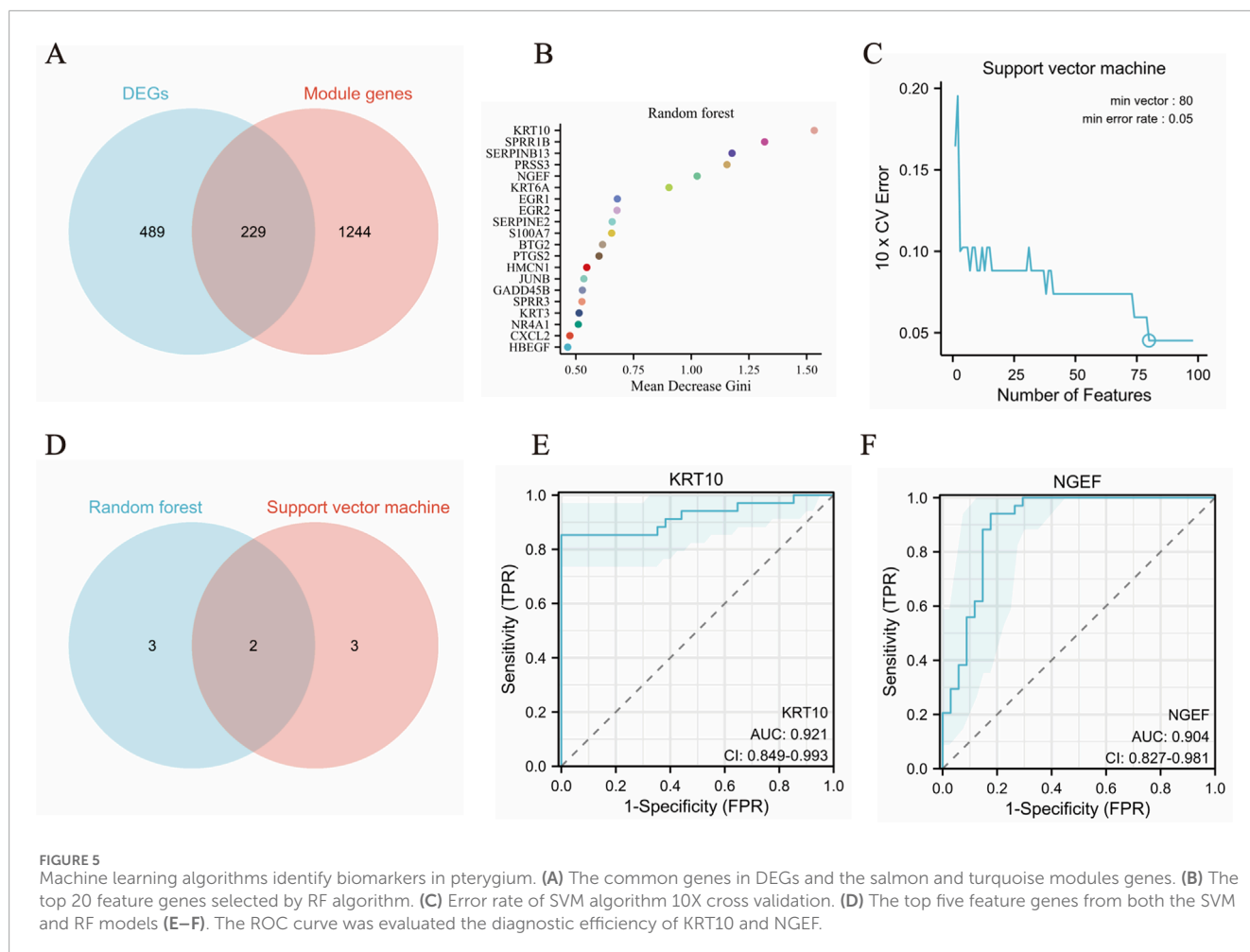
within the pterygium dataset. For the RF classifier, an optimal number of 500 trees was selected, a decision grounded in an evaluation of error rates and classifier stability across varying tree counts. By analyzing the Mean Decrease Gini results, we identified top five pivotal hub genes: KRT10, SPRR1B, SERPINB13, PRSS3 and NGEF (illustrated in Figure 5B). The SVM algorithm was also employed to rigorously select hub genes associated with pterygium, revealing five significant genes: KRT10, KRT6B, BTG2, ASPG and NGEF (Figure 5C).

To identify crucial hub genes involved in pterygium, we examined the intersection of the top five feature genes from both the SVM and RF models (Figure 5D). The model's performance was evaluated using the area under the ROC curve (AUC) (Figures 5E, F). Impressively, the dataset yielded an AUC of 0.921 (KRT10) and 0.904 (NGEF), indicating a high level of accuracy in classifying gene expression data. Notably, KRT10 and NGEF were consistently identified as hub genes associated with pterygium across all two machine learning methods and WGCNA.

We downloaded the original data of GSE2513 and GSE51995 from GEO website, and conducted probe annotation and data standardization on the original expression data to obtain the gene expression matrix. The datasets GSE2513 and GSE51995 were merged utilizing the R packages dplyr and purrr, and the shared genes were identified based on the intersection of genes derived from sequencing analysis. Subsequently, the cleaned data were standardized using the stats, preprocessCore, and limma R packages. The resultant standardized data were then subjected to further analysis (Smyth and Gentleman, 2013). KRT10 achieved an AUC of 0.906 in GSE2513 and GSE51995 combined data and NGEF achieved an AUC of 0.625 (Supplementary Table S2).

## Immuno-infiltration landscape between conjunctiva and pterygium

The CIBERSORT deconvolution algorithm was employed to quantify the infiltration of 22 distinct immune cell types within

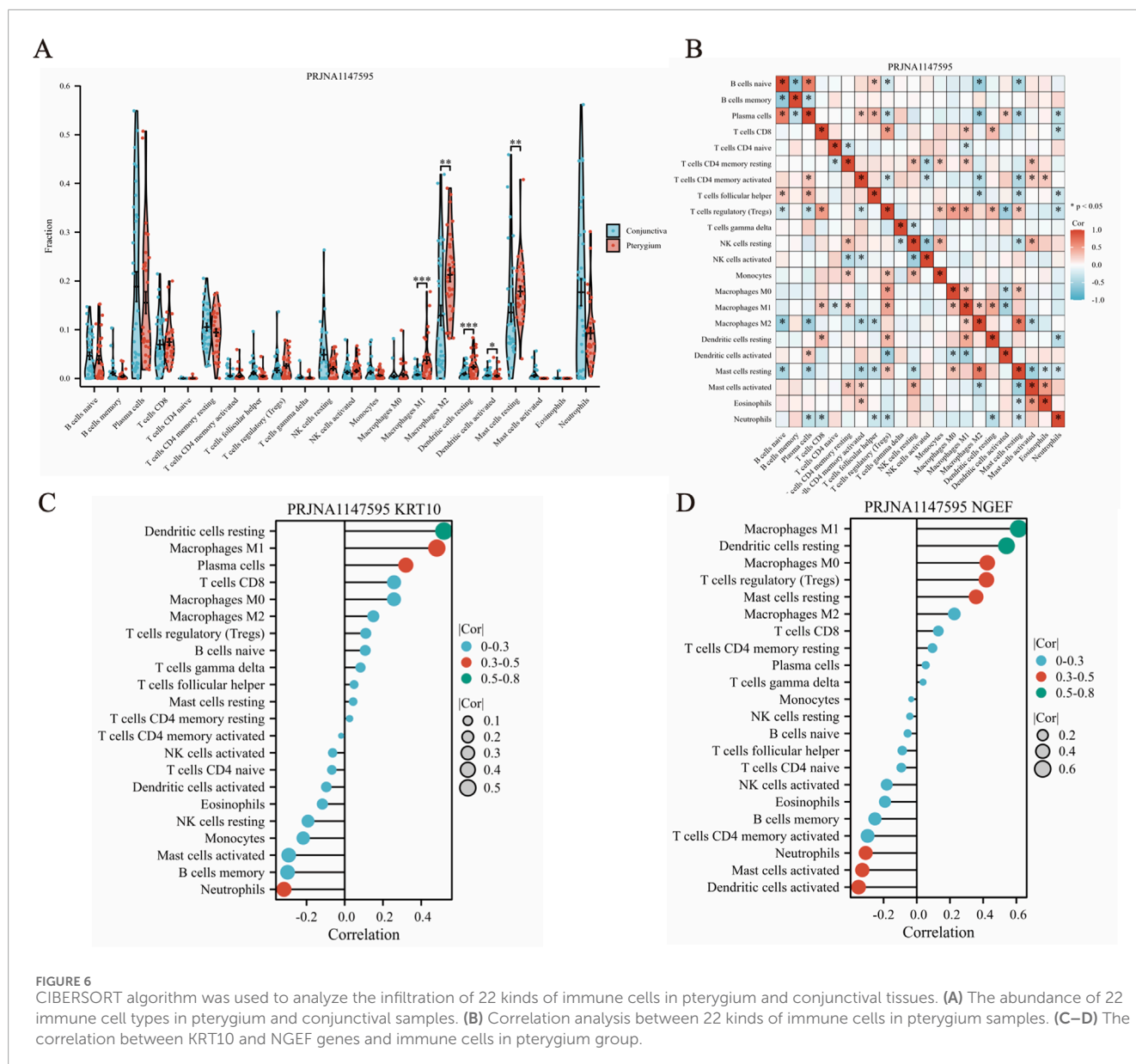


pterygium and bulbar conjunctiva tissues (Figure 6A). Our analysis revealed significant differences in the levels of five immune cells between these two tissue types. Specifically, M1 macrophages, resting dendritic cells, and activated dendritic cells were notably elevated in the pterygium group, whereas resting mast cells and M2 macrophages were significantly increased in the bulbar conjunctiva group (Figure 6A). Furthermore, a general correlation was observed among the 22 immune cell types within pterygium samples (Figure 6B). Correlation analysis between these immune cells and KRT10 (a hub gene) in pterygium tissues indicated that dendritic resting cells ( $\text{cor} = 0.58$ ,  $p = 1.63 \times 10^{-7}$ ) and M1 macrophages ( $\text{cor} = 0.43$ ,  $p = 3.21 \times 10^{-5}$ ) were positively correlated with KRT10 expression, while Neutrophils also exhibited a significant negative correlation with KRT10 ( $\text{cor} = -0.38$ ,  $p = 0.0016$ ) (Figure 6C). Correlation analysis between these immune cells and NGEF in pterygium tissues indicated that Macrophages M1 ( $\text{cor} = 0.61$ ,  $P = 2.57 \times 10^{-8}$ ), Dendritic cells resting ( $\text{cor} = 0.54$ ,  $P = 1.95 \times 10^{-6}$ ), Macrophages M0 ( $\text{cor} = 0.42$ ,  $P = 0.0003$ ), T cells regulatory ( $\text{cor} = 0.42$ ,  $P = 0.0004$ ) and Mast cells resting ( $\text{cor} = 0.36$ ,  $P = 0.0029$ ) were positively correlated with NGEF expression, while Dendritic cells activated ( $\text{cor} = -0.35$ ,  $p = 0.0034$ ) and Mast cells activated ( $\text{cor} = -0.33$ ,  $p = 0.0063$ ) exhibited a significant negative correlation with NGEF (Figure 6D).

To validate our findings, we employed the combined dataset of GSE2513 and GSE51995 from the GEO database as a verification cohort. Following CIBERSORT analysis on this integrated dataset, we observed that there were no statistically significant differences in the proportions of M1 macrophages, resting dendritic cells, activated dendritic cells, resting mast cells, and M2 macrophages between the pterygium and conjunctiva groups. Additionally, a comprehensive correlation analysis revealed notable associations among the 22 immune cell types within the pterygium samples (Supplementary Table S3).

## Validation of the mRNA expression of hub genes by quantitative real-time PCR

In our analysis of RNA-Seq data derived from 68 samples, we observed a statistically significant elevation in the expression levels of KRT10 and NGEF genes in the pterygium group compared to the bulbar conjunctiva group, with a P-value  $< 0.05$  (as depicted in Figures 7A, B). To validate these findings, we conducted a qPCR experiment, which confirmed a significantly higher expression of KRT10 and NGEF in the pterygium group relative to the bulbar conjunctiva group, with a P-value of 0.04



**FIGURE 6** CIBERSORT algorithm was used to analyze the infiltration of 22 kinds of immune cells in pterygium and conjunctival tissues. **(A)** The abundance of 22 immune cell types in pterygium and conjunctival samples. **(B)** Correlation analysis between 22 kinds of immune cells in pterygium samples. **(C–D)** The correlation between KRT10 and NGEF genes and immune cells in pterygium group.

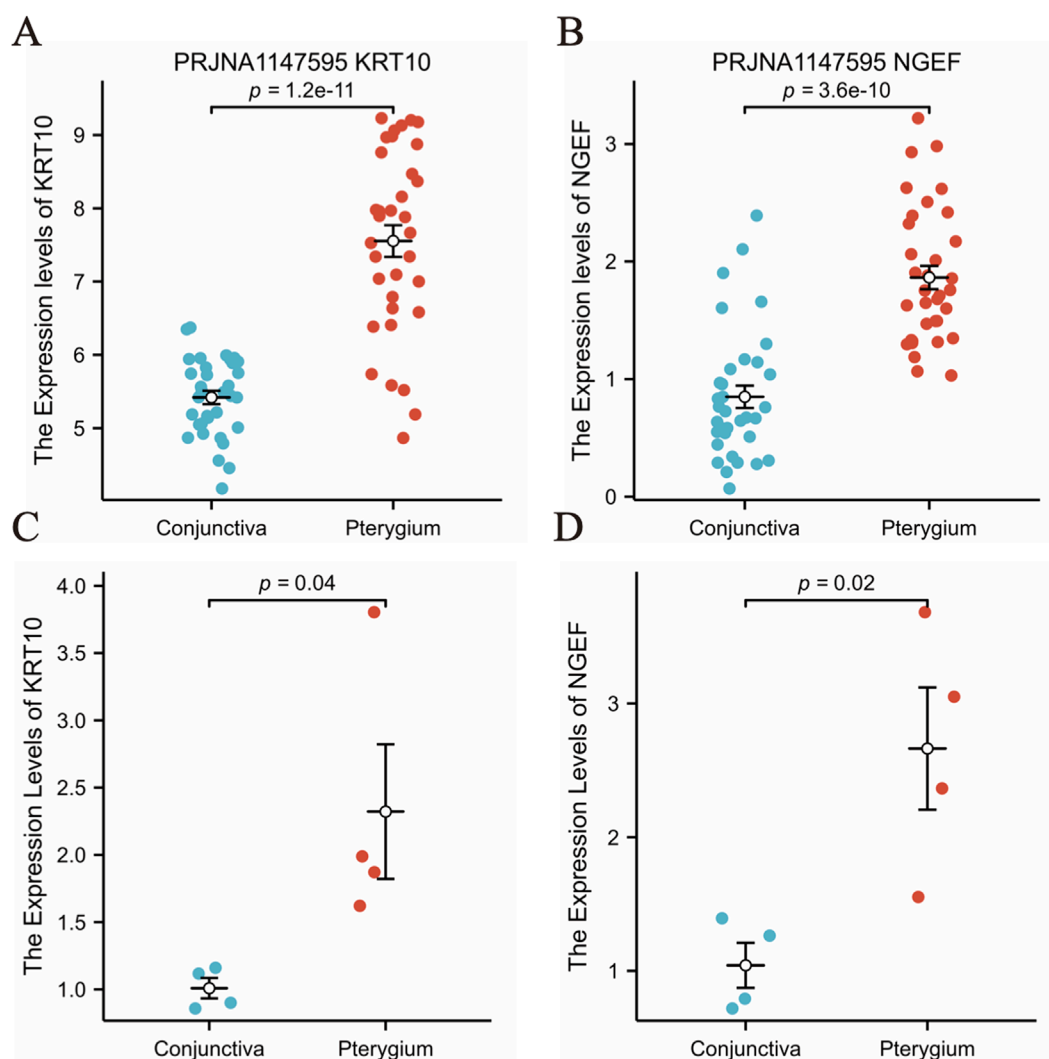
and 0.02 as illustrated in **Figures 7C, D** and primer sequences are shown in **Supplementary Table S5**.

### Identification of key miRNAs and potential drug candidates

To pinpoint crucial miRNAs and potential drug candidates that target the two identified feature genes, we conducted a comprehensive analysis using data sourced from the Targetscan and DSigDB databases. This rigorous screening process, employing a statistical significance threshold of  $p < 0.05$ , led to the identification of three pivotal miRNAs (depicted in **Figure 8A**). Our findings revealed intricate interaction networks: KRT10 was found to interact with one distinct miRNA (hsa-miR-4750), while NGEF was associated with hsa-miR-450a and hsa-miR-4737.

Furthermore, we undertook an extensive screening to identify the drug molecules, guided by P-value  $< 0.05$  as per the DSigDB database (illustrated in **Figures 8B, C**). We conducted a rigorous screening process and identified KRT10 and NGEF as crucial genes implicated in the pathogenesis of pterygium. Subsequently, we presented only the top four drugs that significantly influenced KRT10 expression, along with benzo [a] pyrene CTD 00005488, which demonstrated the ability to jointly regulate both KRT10 and NGEF genes. In the case of the NGEF gene, our analysis revealed only four drugs with statistical significance ( $p < 0.05$ ), including benzo [a] pyrene CTD 00005488, and we have accordingly reported these four drugs. Among these, benzo [a]pyrene (CTD 00005488) emerged as a prominent candidate, demonstrating interactions with all two feature genes and achieving a notable combined score of 94002.9. The remaining drug candidates exhibited interactions with KRT10 and NGEF, offering valuable insights for the advancement of pterygium treatment research and development.





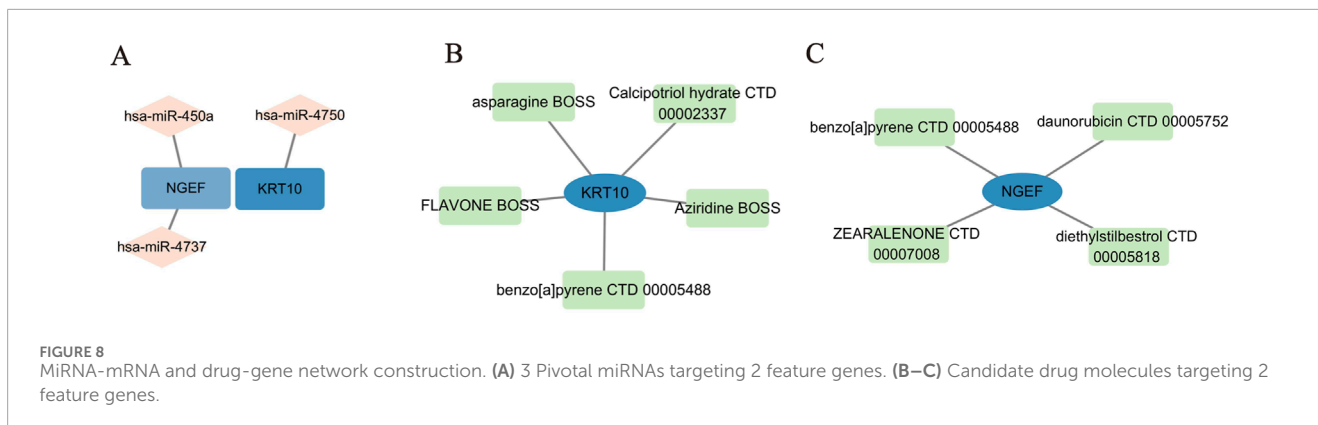
**FIGURE 7**  
 Verification of KRT10 and NGEF genes expression. (A–B) The expression of KRT10 and NGEF in the pterygium group and the conjunctival group was compared in our RNA-seq data. (C–D) The expression of KRT10 and NGEF in 4 pairs of pterygium and conjunctival tissues was analyzed.

## Discussion

Pterygium, a prevalent ophthalmologic condition, is distinguished by abnormal proliferation of subconjunctival fibrous tissue accompanied by angiogenesis, ultimately invading the central cornea. This condition derives its name from its resemblance to insect wings (Lan et al., 2018). The clinical management of pterygium primarily involves surgical excision; however, the high recurrence rates following surgery remain a significant challenge, necessitating further exploration of its underlying pathophysiological mechanisms (Rim et al., 2017).

Our study aims to delve into the molecular underpinnings of pterygium by investigating the differential gene expression between pterygium-affected and normal conjunctival tissues. Utilizing high-throughput RNA sequencing coupled with bioinformatics analyses, we identified 718 differentially expressed genes (DEGs) that may elucidate the biological processes involved in pterygium pathogenesis. Of particular note is the stark contrast between the

upregulated (254) and downregulated (464) genes, suggesting that the pathogenesis of pterygium involves extensive dysregulation of gene expression. The biological functions of these DEGs warrant further exploration, especially concerning their roles in cellular proliferation, inflammation, and tissue remodeling. For instance, genes associated with inflammatory responses, such as those implicated in the IL-17 signaling pathway, could elucidate the chronic inflammatory nature of pterygium development. Additionally, the distinct gene expression profiles identified through heatmaps and volcano plots could serve as a foundation for developing targeted therapies. Importantly, the use of Weighted Correlation Network Analysis (WGCNA) in conjunction with machine learning algorithms like Random Forest (RF) and Support Vector Machine (SVM) to pinpoint key feature genes such as KRT10 and NGEF exemplifies a robust methodological framework that enhances the reliability of the findings, thereby paving the way for subsequent biological validation and clinical application (Lindsay and Sullivan, 2001; Ono et al., 2021).



The pathway enrichment analysis, encompassing Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG), underscores the significant involvement of inflammation-related pathways in pterygium pathology. The enrichment in leukocyte migration and the IL-17 signaling pathway suggests that immune dysregulation plays a pivotal role in the disease's etiology. This aligns with previous findings that highlight the involvement of inflammatory mediators in ocular surface diseases, indicating that targeting these pathways could represent a novel therapeutic strategy to mitigate pterygium progression. Future interventions could focus on modulating these inflammatory pathways to improve patient outcomes, particularly for those with recurrent disease. Furthermore, the integration of Gene Set Variation Analysis (GSVA) and Gene Set Enrichment Analysis (GSEA) results could facilitate a deeper understanding of the temporal dynamics of these pathways during disease progression (Zhang et al., 2024; Gupta et al., 2024). Pterygium is a chronic inflammatory proliferative lesion (Zhang et al., 2011). The CIBERSORT analysis revealing significant differences in immune cell infiltration, particularly the elevated levels of M1 macrophages and resting dendritic cells in pterygium tissues, emphasizes the immune landscape's complexity (Zhao et al., 2023). The presence of pro-inflammatory M1 macrophages suggests a sustained inflammatory response, which could exacerbate tissue damage and drive disease progression. The interaction between various immune cell types and their roles in modulating the tumor microenvironment merits further investigation (Labbé et al., 2010; Celeva et al., 2011).

The validation of KRT10 and NGEF expression through quantitative PCR further solidifies the credibility of the RNA-Seq findings. This experimental confirmation not only reinforces the role of these DEGs in pterygium but also suggests their potential utility as biomarkers for disease diagnosis and prognosis (Adiguzel et al., 2007; March et al., 2019). The identification of potential drug molecules interacting with KRT10 and NGEF, such as benzo [a]pyrene, opens new avenues for therapeutic development (Zhang et al., 2022). These findings underscore the importance of drug repurposing strategies that leverage existing pharmacological agents to target specific pathways implicated in pterygium. The ability to identify compounds that modulate the expression or activity of key genes could lead to more effective treatment modalities. Future research should focus on the pharmacodynamics

and pharmacokinetics of these identified compounds in the context of pterygium, as well as their safety profiles in clinical populations.

This study acknowledges several limitations that warrant consideration. Firstly, the sample size utilized in our analysis was relatively modest, which may restrict the generalizability of our findings. A larger cohort would enhance the robustness of our conclusions and facilitate a more comprehensive understanding of the genetic landscape associated with pterygium. Secondly, the absence of long-term clinical follow-up limits our ability to assess the prognostic implications of the identified biomarkers. Additionally, batch effects inherent in high-throughput sequencing could influence the consistency of gene expression profiles, potentially confounding the interpretation of results. Lastly, while we have established a correlation between differential gene expression and disease pathology, the lack of experimental validation for all findings presents a challenge in confirming causative relationships. Future studies should aim to address these limitations through larger, longitudinal investigations and rigorous experimental validations to bolster the reliability of our insights.

## Conclusion

In conclusion, this study offers an exhaustive examination of the molecular mechanisms underlying pterygium, emphasizing distinct gene expression profiles and their diagnostic and therapeutic implications. The discovery of pivotal genes, namely, KRT10 and NGEF, coupled with insights into immune cell infiltration patterns and pertinent signaling pathways, establishes a foundation for future research endeavors aimed at deciphering the pathophysiology of pterygium. Moreover, the identification of potential pharmacological agents that interact with these key genes presents promising opportunities for the development of novel therapies. As we progress, it is crucial for subsequent studies to validate these findings in larger and more heterogeneous populations. This will ultimately facilitate the development of targeted interventions and personalized treatment strategies for patients suffering from pterygium.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repository and accession number(s) can be found in the article/[Supplementary Material](#).

## Ethics statement

The studies involving humans were approved by Ethics Review Committee of the Affiliated Hospital of Yunnan University. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants's; legal guardians/next of kin.

## Author contributions

L-WZ: Conceptualization, Supervision, Writing – review and editing, Funding acquisition. JY: Formal Analysis, Writing – original draft, Investigation, Software. H-WJ: Investigation, Software, Validation, Writing – original draft, Writing – review and editing. X-QY: Resources, Writing – review and editing. Y-NC: Data curation, Resources, Writing – review and editing. W-DY: Data curation, Resources, Writing – review and editing. Y-LD: Data curation, Resources, Writing – review and editing. M-hZ: Data curation, Resources, Writing – review and editing. HL: Funding acquisition, Methodology, Supervision, Writing – review and editing. H-LZ: Conceptualization, Project administration, Writing – review and editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was carried out with the support of the National Natural Science Foundation of China (Grant No. 82460201 and No. 81860171). Yunnan Provincial Health Committee Training

## References

- Adiguzel, U., Karabacak, T., Sari, A., Oz, O., and Cinel, L. (2007). Cyclooxygenase-2 expression in primary and recurrent pterygium. *Eur. J. Ophthalmol.* 17 (6), 879–884. doi:10.1177/112067210701700602
- Agarwal, V., Bell, G. W., Nam, J. W., and Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *eLife* 4, e05005. doi:10.7554/eLife.05005
- Alqahtani, J. M. (2013). The prevalence of pterygium in Alkhobar: a hospital-based study. *J. Fam. and community Med.* 20 (3), 159–161. doi:10.4103/2230-8229.121980
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556
- Celeva, M. V., Stankovic, B. G., and Zdravkowska, J. M. (2011). Comparative study of pterygium surgery. *Prilozi* 32 (2), 273–287.
- Chalkia, A. K., Bontzos, G., Spandidos, D. A., and Detorakis, E. T. (2019). Human papillomavirus infection and ocular surface disease (Review). *Int. J. Oncol.* 54 (5), 1503–1510. doi:10.3892/ijo.2019.4755
- Chang, J., Cao, Q., Yong, J., Ling, X., Zhang, X., Kang, Z., et al. (2023). The effect of different pterygium surgery techniques on the ocular surface parameters in different durations: a systematic review and meta-analysis. *Graefes Archive Clin. Exp. Ophthalmol.* 262 (5), 1383–1396. doi:10.1007/s00417-023-06191-1
- Chu, W. K., Choi, H. L., Bhat, A. K., and Jhanji, V. (2020). Pterygium: new insights. *Eye (Lond)*. 34 (6), 1047–1050. doi:10.1038/s41433-020-0786-3
- Guimares, J. A. D., Hounpke, B. W., Duarte, B., Boso, A. L. M., Vitorino, M. G. M., Baptista, L. C., et al. (2022). Transcriptomics and network analysis highlight potential pathways in the pathogenesis of pterygium. *Sci. Rep.* 12 (1), 286. doi:10.1038/s41598-021-04248-x
- Gupta, M., Arya, S., Agrawal, P., Gupta, H., and Sikka, R. (2024). Unravelling the molecular tapestry of pterygium: insights into genes for diagnostic and therapeutic innovations. *Eye* 38 (15), 2880–2887. doi:10.1038/s41433-024-03186-y
- Hacıoğlu, D., and Erdöl, H. (2017). Developments and current approaches in the treatment of pterygium. *Int. Ophthalmol.* 37 (4), 1073–1081. doi:10.1007/s10792-016-0358-5
- Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14 (7), 7. doi:10.1186/1471-2105-14-7

program for leading medical talents (L2019029). China Scholarship Council (202208535051). Yunnan Applied Basic Research Project (No. 202101AZ070001-002).

## Acknowledgments

We are grateful to the contributors to the public databases used in this study.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2024.1524517/full#supplementary-material>

- Hou, A., Lan, W., Law, K. P., Khoo, S. C., Tin, M. Q., Lim, Y. P., et al. (2014). Evaluation of global differential gene and protein expression in primary Pterygium: S100A8 and S100A9 as possible drivers of a signaling network. *PLoS one* 9 (5), e97402. doi:10.1371/journal.pone.0097402
- Huang, Y., He, H., Sheha, H., and Tseng, S. C. (2013). Ocular demodicosis as a risk factor of pterygium recurrence. *Ophthalmology* 120 (7), 1341–1347. doi:10.1016/j.ophtha.2013.01.001
- Kalogeropoulos, D., Papoudou-Bai, A., Lane, M., Goussia, A., Charchanti, A., Moschos, M. M., et al. (2020). Antigen-presenting cells in ocular surface diseases. *Int. Ophthalmol.* 40 (6), 1603–1618. doi:10.1007/s10792-020-01329-0
- Labbé, A., Gheek, L., Iordanidou, V., Mehanna, C., Brignole-Baudouin, F., and Baudouin, C. (2010). An *in vivo* confocal microscopy and impression cytology evaluation of pterygium activity. *Cornea*. 29 (4), 392–399. doi:10.1097/ICO.0b013e3181bd44ce
- Lan, W., Hou, A., Lakshminarayanan, R., Lim, Y. P., and Tong, L. (2018). Linc-9432 is a novel pterygium lincRNA which regulates differentiation of fibroblasts. *FEBS Lett.* 592 (7), 1173–1184. doi:10.1002/1873-3468.13027
- Lindsay, R. G., and Sullivan, L. (2001). Pterygium-induced corneal astigmatism. *Clin. and Exp. optometry* 84 (4), 200–203. doi:10.1111/j.1444-0938.2001.tb05026.x
- Lu, C. W., Fu, J., Liu, X. F., Cui, Z. H., Chen, W. W., Guo, L., et al. (2023). Impacts of air pollution and meteorological conditions on dry eye disease among residents in a northeastern Chinese metropolis: a six-year crossover study in a cold region. *Light, Sci. and Appl.* 12 (1), 186. doi:10.1038/s41377-023-01207-1
- March, O. P., Lettner, T., Klausegger, A., Ablinger, M., Kocher, T., Hainzl, S., et al. (2019). Gene editing-mediated disruption of epidermolytic ichthyosis-associated KRT10 alleles restores filament stability in keratinocytes. *J. investigative dermatology* 139 (8), 1699–1710. doi:10.1016/j.jid.2019.03.1146
- Modenese, A., Chou, B. R., Adam, B., Loney, T., Thomas, T., Gobba, F., et al. (2023). Occupational exposure to solar radiation and the eye: a call to implement health surveillance of outdoor workers. *Med. Lav.* 114 (4), e2023032. doi:10.23749/mdl.v114i4.14657
- Moran, D. J., and Hollows, F. C. (1984). Pterygium and ultraviolet radiation: a positive correlation. *Br. J. Ophthalmol.* 68 (5), 343–346. doi:10.1136/bjo.68.5.343
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337
- Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., and Kanehisa, M. (1999). KEGG: Kyoto Encyclopedia of genes and Genomes. *Nucleic acids Res.* 27 (1), 29–34. doi:10.1093/nar/27.1.29
- Ono, T., Mori, Y., Nejima, R., Iwasaki, T., Miyai, T., Aihara, M., et al. (2021). Comparison of corneal irregularity after recurrent and primary pterygium surgery using fourier harmonic analysis. *Transl. Vis. Sci. and Technol.* 10 (11), 13. doi:10.1167/tvst.10.11.13
- Prabhasawat, P., Barton, K., Burkett, G., and Tseng, S. C. G. (1997). Comparison of conjunctival autografts, amniotic membrane grafts, and primary closure for pterygium excision. *Ophthalmology* 104 (6), 974–985. doi:10.1016/s0161-6420(97)30197-3
- Rezvani, F., Khabazkhoob, M., Hooshmand, E., Yekta, A., Saatchi, M., and Hashemi, H. (2018). Prevalence and risk factors of pterygium: a systematic review and meta-analysis. *Surv. Ophthalmol.* 63 (5), 719–735. doi:10.1016/j.survophthal.2018.03.001
- Rim, T. H., Kang, M. J., Choi, M., Seo, K. Y., and Kim, S. S. (2017). The incidence and prevalence of pterygium in South Korea: a 10-year population-based Korean cohort study. *PLoS one* 12 (3), e0171954. doi:10.1371/journal.pone.0171954
- Rubeshkumar, P. C., Ponnaiah, M., Anandhi, D., and John, D. (2020). Association between exposure to artificial sources of ultraviolet radiation and ocular diseases: a systematic review protocol. *JBI Evid. Synth.* 18 (8), 1766–1773. doi:10.11124/jbisrir-d-19-00206
- Shahraki, T., Arabi, A., and Feizi, S. (2021). Pterygium: an update on pathophysiology, clinical features, and management. *Ther. Adv. Ophthalmol.* 13, 25158414211020152. doi:10.1177/25158414211020152
- Singh, G., Wilson, M. R., and Foster, C. S. (1988). Mitomycin eye drops as treatment for pterygium. *Ophthalmology* 95 (6), 813–821. doi:10.1016/s0161-6420(88)33104-0
- Smyth, G. K. (2013). “Limma: linear models for microarray data,” in *Bioinformatics and computational biology solutions using R and Bioconductor*. Editor R. Gentleman (Springer).
- Taylor, H. R. (2013). Risk factors for pterygium. *Ophthalmology* 120 (2), 441. doi:10.1016/j.ophtha.2012.09.048
- Uddin, S., Khan, A., Hossain, M. E., and Moni, M. A. (2019). Comparing different supervised machine learning algorithms for disease prediction. *BMC Med. Inf. Decis. Mak.* 19 (1), 281. doi:10.1186/s12911-019-1004-8
- Wong, Y. W., Chew, J., Yang, H., Tan, D. T., and Beuerman, R. (2006). Expression of insulin-like growth factor binding protein-3 in pterygium tissue. *Br. J. Ophthalmol.* 90 (6), 769–772. doi:10.1136/bjo.2005.087486
- Yoon, C. H., Seol, B. R., and Choi, H. J. (2023). Effect of pterygium on corneal astigmatism, irregularity and higher-order aberrations: a comparative study with normal fellow eyes. *Sci. Rep.* 13 (1), 7328. doi:10.1038/s41598-023-34466-4
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OmicS a J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118
- Zhang, L. W., Chen, B. H., Xi, X. H., Han, Q. Q., and Tang, L. S. (2011). Survivin and p53 expression in primary and recurrent pterygium in Chinese patients. *Int. J. Ophthalmol.* 4 (4), 388–392. doi:10.3980/j.issn.2222-3959.2011.04.13
- Zhang, S., Liu, T., Zhang, M., Chen, Q., Su, M., Bai, T., et al. (2022). Study on molecular mechanism of benzo (a) pyrene on CMA by HSP90 $\alpha$  and HIF-1 $\alpha$ . *Toxicol. vitro* 83, 105372. doi:10.1016/j.tiv.2022.105372
- Zhang, X., Han, P., Qiu, J., Huang, F., Luo, Q., Cheng, J., et al. (2024). Single-cell RNA sequencing reveals the complex cellular niche of pterygium. *Ocul. Surf.* 32, 91–103. doi:10.1016/j.jtos.2024.01.013
- Zhao, Z., Mak, T. K., Shi, Y., Huang, H., Huo, M., and Zhang, C. (2023). The DNA damage repair-related lncRNAs signature predicts the prognosis and immunotherapy response in gastric cancer. *Front. Immunol.* 14, 1117255. doi:10.3389/fimmu.2023.1117255