BMC Pharmacology and Toxicology

RESEARCH

Open Access

Development and validation of potential molecular subtypes and signatures of thyroid eye disease based on angiogenesis-related gene analysis

Zixuan Wu[1†], Jun Peng[2†], Xi Long[1], Kang Tan[1], Xiaolei Yao[2,3*] and Qinghua Peng[1,2*]

Abstract

Background Thyroid eye disease (TED) is an autoimmune inflammatory disorder of the orbit, associated with a range of potential clinical sequelae. Tumor cells in TED overexpress pro-angiogenic factors, driving the formation of heterogeneous and immature neovascularization. This dysregulated angiogenesis often leads to a hypoxic microenvironment due to insufficient perfusion. Despite its importance, the role of angiogenesis-related genes (ARGs) in TED pathophysiology remains poorly understood.

Methods To bridge this knowledge gap, our study aimed to identify and validate ARGs implicated in TED using a comprehensive bioinformatics strategy. By intersecting differential gene expression analyses with a curated list of 103 known ARGs, we aimed to pinpoint those with potential roles in TED. Advanced methodologies, including GSEA and GSVA, facilitated an in-depth exploration of the biological functions and pathways associated with these ARGs. Further refinement through Lasso regression and SVM-RFE enabled the identification of key hub genes and the evaluation of their diagnostic potential for TED. Additionally, we investigated the relationship between these hub ARGs and relevant clinical parameters. To corroborate our findings, we analyzed expression data from datasets GSE58331 and GSE105149, focusing on the six ARGs identified as potentially crucial to TED pathology.

Results Our investigation unveiled six ARGs (CRIP2, DUSP1, CTSL, DOCK5, ERAP1, SCG2) as intimately connected to TED. Functional analyses highlighted their involvement in processes such as response to ameboidal-type cell migration, epithelial cell migration, epithelium migration. Importantly, the diagnostic capabilities of these ARGs demonstrated promising efficacy in distinguishing TED from non-affected states.

†Zixuan Wu and Jun Peng contributed equally to this article as the first author.

Qinghua Peng and Xiaolei Yao contributed equally to this article as the corresponding author of this article.

*Correspondence:
Xiaolei Yao
yxlshh@126.com
Qinghua Peng
pqh410007@126.com

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

Conclusions This study identifies six ARGs as novel biomarker candidates for TED, elucidating their potential roles in the disease's pathogenesis.

Keywords Thyroid eye disease (TED), Angiogenesis-related gene (ARGs), Biomarker candidates, CTSL, DOCK5

Introduction

Thyroid eye disease (TED), also known as Graves' ophthalmopathy or Graves' orbitopathy, is an inflammatory condition affecting the tissues surrounding the eyes. TED is commonly associated with Graves' disease, a systemic autoimmune disorder with endocrine manifestations that can have a significant impact on patients' quality of life [1, 2]. The severity of TED lies in its potential to cause visual impairment, functional disability, and disfigurement [3]. Although TED is most commonly observed in patients with Graves' hyperthyroidism, it can also occur in patients who are hypothyroid or euthyroid [4]. The onset of TED typically occurs within 18 months of the manifestation of endocrine symptoms, with approximately 80% of patients experiencing concurrent presentations. While TED and Graves' hyperthyroidism can affect individuals of any age, it predominantly affects women in their third to fifth decade of life. The estimated annual incidence of thyroid eye disease is approximately 16 cases per 100,000 women and three cases per 100,000 men [5, 6]. Several risk factors have been identified, including cigarette smoking, longer duration of Graves' hyperthyroidism, poorly controlled thyroid dysfunction, and prior treatment with radioactive iodine [7]. The diagnosis of TED is typically straightforward, based on a thorough medical history and physical examination. Ophthalmic manifestations occur in up to 50% of patients with Graves' hyperthyroidism [8]. A deeper understanding of the molecular mechanisms underlying TED is crucial for the development of novel treatments that can prevent disease recurrence and enhance patient outcomes.

Angiogenesis, the process of forming new blood vessels from pre-existing vasculature, is a critical physiological event with significant roles in both health and disease. Dysregulation of angiogenesis is implicated in numerous pathological conditions, including cancer, diabetic retinopathy, and chronic inflammatory diseases [9]. In the context of oncology, tumor angiogenesis is a hallmark of cancer, facilitating tumor growth and metastasis. The angiogenic switch, a critical step in tumor progression, involves the upregulation of angiogenic inducers and downregulation of inhibitors [10]. Therapeutic strategies targeting angiogenesis, such as VEGF inhibitors, have shown promise in controlling tumor growth and improving patient outcomes. However, resistance to anti-angiogenic therapies and the heterogeneity of tumor vasculature remain significant challenges. Diabetic retinopathy exemplifies the pathological angiogenesis associated with chronic diseases [11]. High glucose levels lead to the overproduction of VEGF, promoting the growth of fragile and leaky blood vessels in the retina, ultimately resulting in vision loss. Anti-VEGF therapies have become the cornerstone of treatment, offering significant improvements in visual acuity for many patients [12]. Chronic inflammatory diseases, including rheumatoid arthritis, are marked by pathological angiogenesis, which sustains and amplifies inflammation within the highly vascularized synovium. Targeting angiogenic pathways presents a promising approach to alleviating inflammation and mitigating tissue damage [13]. However, significant challenges persist, including the intricate regulation of angiogenesis, tissue-specific heterogeneity of angiogenic signals, and disease adaptability to anti-angiogenic therapies. These complexities necessitate further investigation to refine therapeutic strategies and enhance their efficacy.

The advent of high-throughput transcriptome sequencing and the extensive clinical annotations provided by the TED Initiative have paved new pathways for investigating the transcriptional dynamics and molecular mechanisms underlying TED in biological research [14–16]. Bioinformatics analysis of these rich datasets has revealed compelling insights, shedding light on the intricate pathophysiology and multifaceted mechanisms of TED. Despite significant advancements, there remains a notable gap in the literature regarding the specific exploration of TED using bioinformatics tools. Therefore, the primary objective of this study was to leverage TED-related GEO datasets to explore the significance of various factors within the TED landscape, as depicted in Fig. 1.

Materials and methods

We adopted the methodologies delineated by Zi-Xuan Wu and colleagues in 2023 [17].

Raw data and differentially expressed genes (DEGs)

We utilized two foundational datasets from the GEO series, GSE58331 and GSE105149. GSE58331 was used for training, while GSE105149 was reserved for validation (Table 1). Additionally, the MSigDB provided a comprehensive list of 333 ARGs (Table S1). mRNA profiles were extracted using Perl scripts to match and sort transcriptional data from GSE58331. Following normalization, DEGs among the ARGs were identified using stringent criteria: FDR < 0.05 and $|\log2FC| \geq 1$. Pearson's correlation coefficient was then employed, using the corrplot package in R. R [18] is a versatile and powerful open-source programming language primarily used for statistical

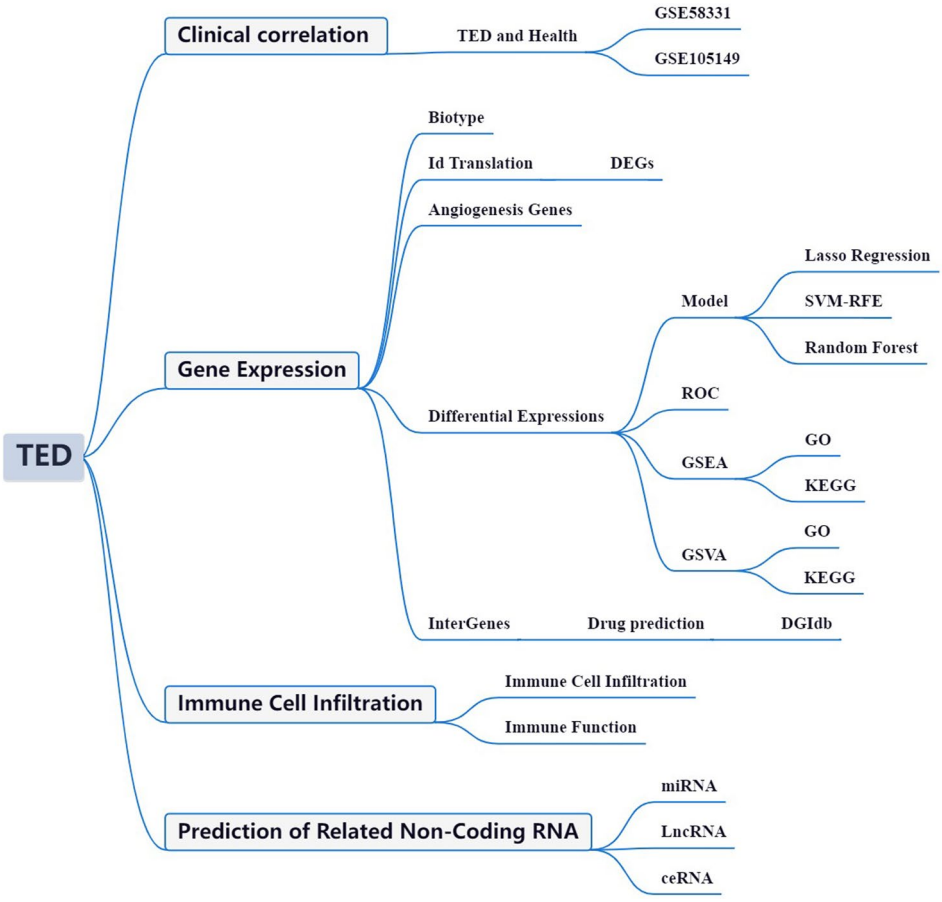

Fig. 1 Framework

Table 1 The clinical characteristics of patients

GSE58331 (Train)		GSE105149 (Test)	
Variables	**Number of samples**	**Variables**	**Number of samples**
Gender		Gender	
Male/Female	15/49	Male/Female	5/6
Diagnosis		Diagnosis	
TED/Normal	35/29	TED/Normal	4/7
Tissue		Tissue	
Anterior Orbit/ Lacrimal gland	49/15	Anterior Orbit/ Lacrimal gland	0/11

computing and data analysis. Widely adopted across various disciplines, including bioinformatics, epidemiology, and computational biology, R provides an extensive range of tools for data manipulation, statistical modeling, and graphical visualization.

Building a model, immune cell infiltration and functional enrichment analysis

To explore the biological significance and pathway involvement of DEGs, we performed GO and KEGG pathway analyses. Using R, we examined the impact of differentially expressed ARGs on BP, MF, and CC. Lasso regression, implemented through the glmnet package with cross-validation, was applied to refine our model. For further validation, we employed the support vector machine-recursive feature elimination (SVM-RFE) algorithm using the e1071 package, constructing a machine learning model with precision. The analysis began with feature selection using the SVM-RFE method, implemented through the svmRFE function. In this approach, the number of selected features (denoted as k) was set to 10, with the threshold for halving feature importance established at 50%. To evaluate the robustness and generalizability of the selected features, 10-fold cross-validation was employed. The dataset, consisting of geneNum genes, was randomly partitioned into 10 distinct subsets, ensuring each gene was assigned to a fold. This partitioning was accomplished by creating a random assignment of fold labels (folds) and subsequently applying a function to group the genes into 10 equal folds. For each fold, feature selection was repeated using the SVM-RFE method, and the results were computed through the svmRFE. wrap function. This function executed the feature selection process for each fold, maintaining the parameters set

earlier (10 features and the halving threshold of 50%). The overall outcomes from each fold were then consolidated to assess the consistency and effectiveness of the selected features across the entire dataset.

Cross-validation was pivotal in assessing model performance, minimizing error rates, and enhancing accuracy to ensure robustness. Key gene analysis and disease classification were conducted using the ggplot2 package, enabling us to pinpoint critical genes. The analysis begins by converting the input data rt into a matrix format using as.matrix(rt). The row names of rt are then processed to extract the second part of each row name (after the underscore) using the gsub function, with the result stored in the variable y. The relationship between the matrix x and the processed response variable y is then modeled using logistic regression via the glmnet function, specifying the "binomial" family and an elastic net regularization (alpha = 1). To refine the model, cross-validation is performed using cv.glmnet, again specifying the "binomial" family and an elastic net penalty (alpha = 1). The cross-validation uses a 10-fold procedure (nfolds = 10) and evaluates the model's performance based on the deviance metric (type.measure = 'deviance'). This cross-validation step allows for optimal model selection and assessment of its generalizability. Furthermore, immune cell composition was examined through the CIBERSORT algorithm, providing comprehensive insights into the immune landscape associated with the disease.

Gene set enrichment and variation analyses and drug-gene interaction insights

To unveil functional dynamics and pathway alterations across diverse samples, we employed GSEA and GSVA. Using R, we evaluated the impact of differentially expressed ARGs on BP, MF, and CC, and pathways, yielding a detailed understanding of their roles in disease mechanisms. Recognizing validated biomarkers as key components of therapeutic strategies, accurate drug prediction becomes essential. To achieve this, we utilized the DGIdb to predict potential drug interactions for the identified hub genes.

Construction of an mRNA-miRNA-lncRNA network

Non-coding RNA transcripts play pivotal roles in shaping the genetic regulatory network. miRNAs modulate gene expression by promoting mRNA degradation and inhibiting translation, while lncRNAs, typically over 200 nucleotides in length, influence diverse cellular processes through chromatin modification, transcriptional regulation, and interference mechanisms. Recent studies underscore the intricate crosstalk between miRNAs and lncRNAs, leading to competitive binding interactions with other regulatory molecules. This phenomenon, termed the ceRNA network, reveals how lncRNAs regulate gene expression by sequestering miRNAs, thereby modulating their activity. To investigate these interactions, we curated target gene information from miRTarBase and PrognoScan, databases that provide validated miRNA-lncRNA-target relationships.

Results

Identification and enrichment analysis of DEGs

Among the 103 ARGs analyzed, a subset showed significant differences in expression levels between treatment and control groups. Specifically, genes such as SHH, BCAM, CARHSP1, NOTCH1, CRIP2, ZCCHC14, DLL4, UPP1, DAPK3, GATA2, PDCD1, TBXA2R clustered in the treatment group, while genes like CANX, ARRDC3, BACE2, AGGF1, CNIH1, KRAS, C1QBP, DUSP6, ERAP1, PDCD10, CAPZA2, DEGS1 clustered in the control group (Fig. 2a). These DEGs were further analyzed for correlation, and a correlation matrix was visualized (Fig. 2b) (Table S2). GO enrichment analysis identified 935 core targets. MFs primarily included enzyme inhibitor activity (GO:0004857), cadherin binding (GO:0045296), DNA-binding transcription activator activity, RNA polymerase II-specific (GO:0001228). CCs mainly involved the external side of plasma membrane (GO:0009897), cell-cell junction (GO:0005911), apical part of cell (GO:0045177). BPs mainly response to ameboidal-type cell migration (GO:0001667), epithelial cell migration (GO:0010631), epithelium migration (GO:0090132). KEGG analysis revealed that overexpressed genes were primarily involved in Pathways in cancer (hsa05200),Tight junction (hsa04530),Hepatitis C (hsa05160) (Fig. 2c-d and Table S3a–b).

Model construction

We employed a methodological involving LASSO regression analysis and Cox proportional hazards regression analysis(Fig. 3a-b). Subsequently, we constructed a machine learning model using SVM-RFE to validate the predictive accuracy and reliability of the model. This model demonstrated an impressive accuracy of 0.74, with a minimal error rate of 0.26 (Fig. 3c-d). A cross-reference of six ARGs identified through both LASSO and SVM methodologies (Fig. 3e). Comparative analysis with the six hub genes revealed significantly high AUC values, indicating robust predictive power: CRIP2 (AUC = 0.759), DUSP1 (AUC = 0.716), CTSL (AUC = 0.703), DOCK5 (AUC = 0.798), ERAP1 (AUC = 0.778), SCG2 (AUC = 0.665) (Fig. 3f). Notably, an AUC of 0.949 (95% Confidence Interval: 0.889 – 0.992) was achieved in the GSE58331 dataset, underscoring the model's exceptional precision and stability (Fig. 3g and Table 2 and S4).

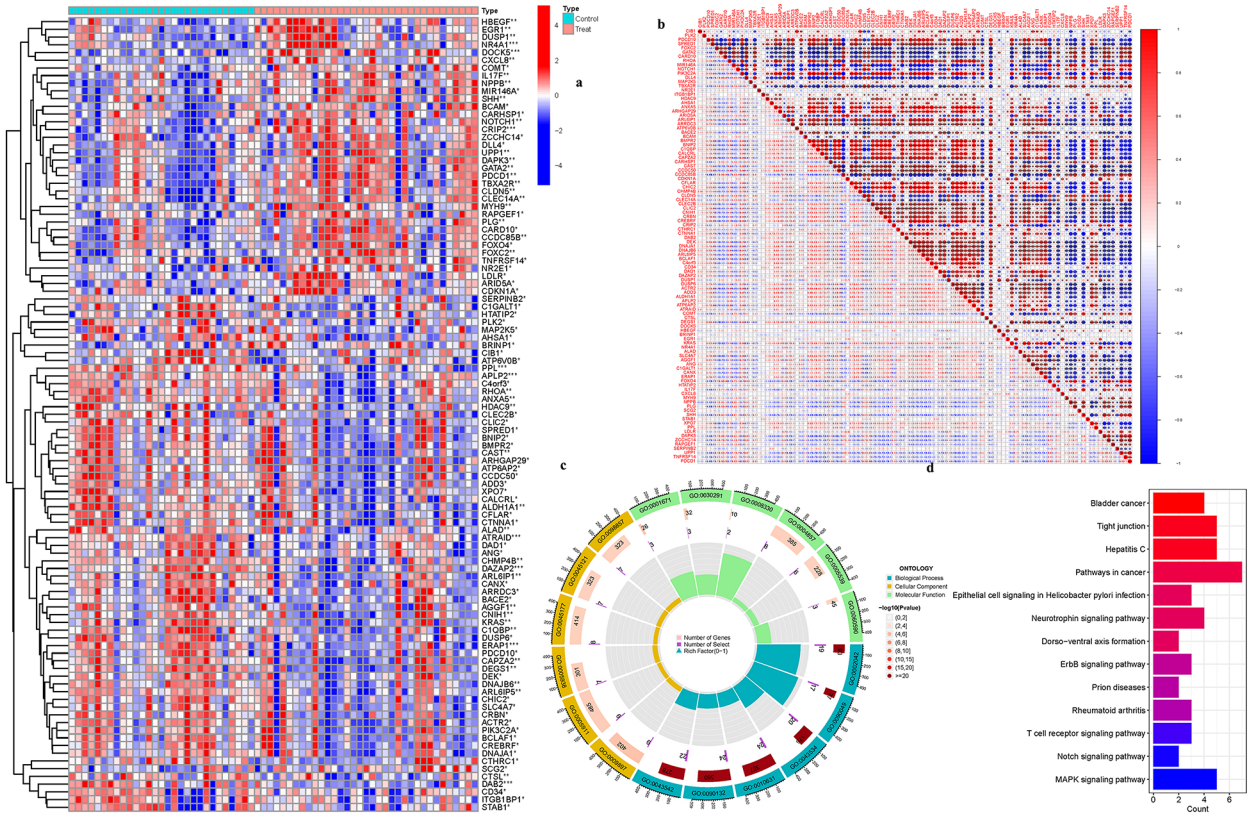

Fig. 2 Identification and enrichment analysis. (**a**) Difference analysis. (**b**) Correlation analysis. (**c**) GO. (**d**) KEGG

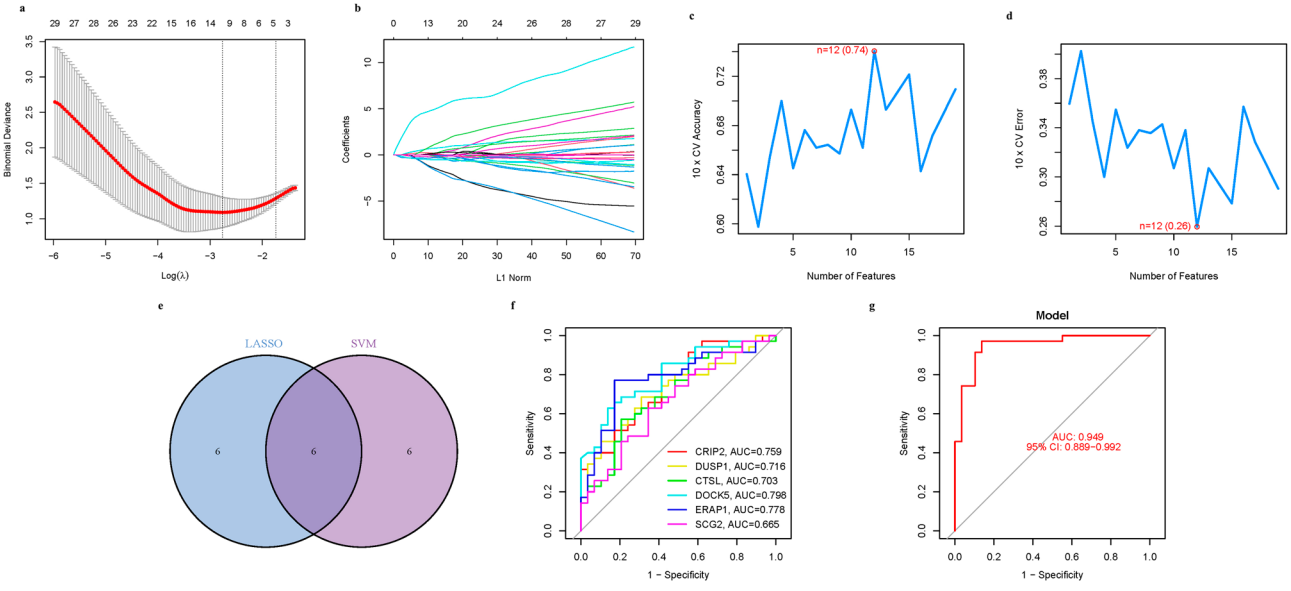

Fig. 3 The ARGs signature. (**a**) Regression of LASSO. (**b**) Cross-validation. (**c–d**) Accuracy and error. (**e**) Venn. (**f**) AUC of 6 hub genes. (**g**) AUC of train group

GSEA of analysis

Through literature review and sensitivity analysis within the model, we identified CTSL and DOCK5 as potentially the most relevant genes to TED. In terms of GO analysis, CTSL mainly involves BP response to peptide,BP response to peptide hormone,BP skin development. DOCK5 mainly involves CC cell substrate junction,CC nuclear speck,MF smad binding (Fig. 4a). In KEGG analysis, CTSL is primarily associated with citrate cycle tca cycle, complement and coagulation cascades, insulin

Table 2 The characteristics of model

Label	LASSO	SVM-RFE
Sensitivity	0.909091	1
Specificity	0.857143	0.928571
Pos Pred Value	0.833333	0.916667
Neg Pred Value	0.923077	1
Precision	0.833333	0.916667
Recall	0.909091	1
F1	0.869565	0.956522
Prevalence	0.44	0.44
Detection Rate	0.4	0.44
Detection Prevalence	0.48	0.48
Balanced Accuracy	0.883117	0.964286

signaling pathway, while DOCK5 is linked to pathways in cancer, spliceosome, tgf beta signaling pathway (Fig. 4b) (Table S5).

Immunological environment in TED

The immunological environment plays a crucial role in the initiation and progression of TED. We created a violin plot to illustrate the levels of immune cells. T cells follicular helper and Neutrophils activated were highly expressed in the treatment group (Fig. 5a). Furthermore, we conducted a correlation study between these genes and immune cells (Fig. 5b). Additionally, we analyzed the immune infiltration of CTSL and DOCK5 separately (Fig. 5c-d).

GSVA of analysis

In terms of GO analysis, CTSL mainly involves BP urea transmembrane transport, MF urea transmembrane transporter activity, MF copper ion transmembrane transporter activity, MF cell adhesive protein binding involved in bundle of his cell purkinje myocyte communication, BP cardiac muscle cell cardiac muscle cell adhesion, MF structural constituent of tooth enamel. DOCK5 mainly involves CC immunoglobulin complex circulating, CC microvesicle, BP negative regulation of fibroblast growth factor production, BP flavone metabolic process, CC immunoglobulin complex, BP flavonoid glucuronidation, BP xenobiotic glucuronidation (Fig. 6a). In terms of KEGG analysis, CTSL mainly involves maturity onset diabetes of the young, adherens junction, thyroid cancer,

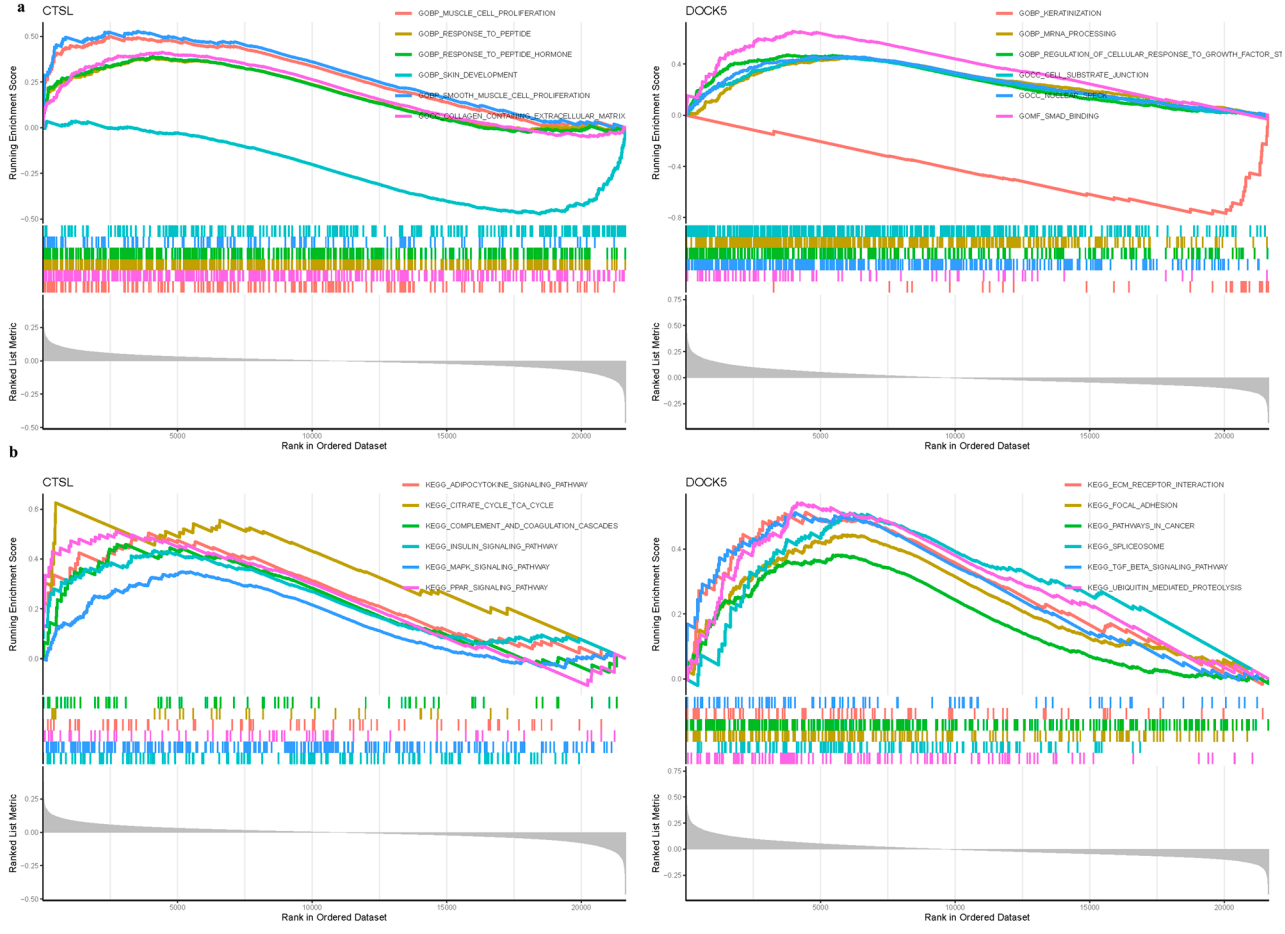

Fig. 4 GSEA of analysis in CTSL and DOCK5. (**a**) GO. (**b**) KEGG

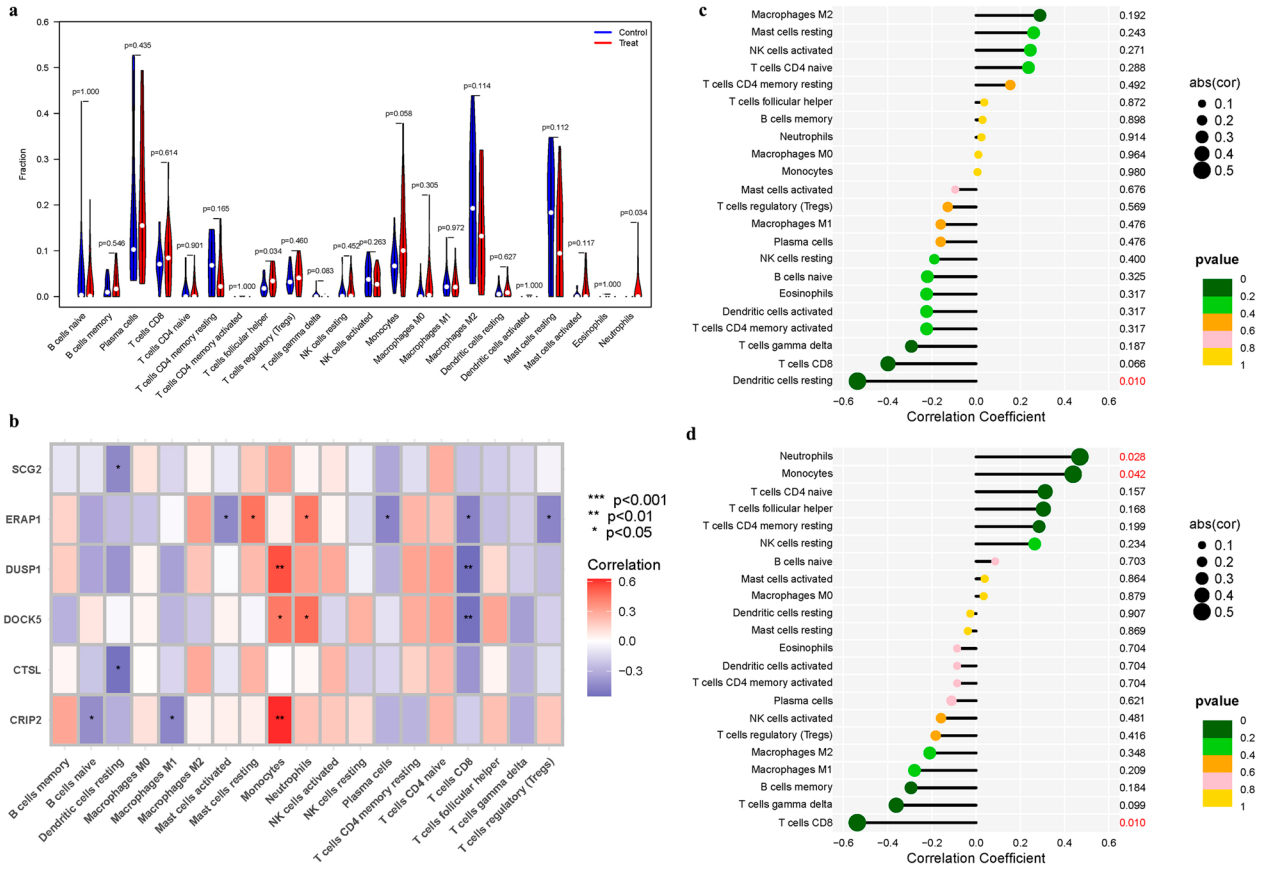

Fig. 5 Immunological Environment. (**a**) Immune cell. (**b**) Correlation between ARGs and immune cells. (**c**) CTSL. (**d**) DOCK5

glycosphingolipid biosynthesis lacto and neolacto series, rna degradation, riboflavin metabolism, asthma, intestinal immune network for iga production. DOCK5 mainly involves glyoxylate and dicarboxylate metabolism, intestinal immune network for iga production, steroid biosynthesis, maturity onset diabetes of the young, folate biosynthesis, olfactory transduction, glycosphingolipid biosynthesis lacto and neolacto series, primary immunodeficiency (Fig. 6b).

Drug-gene interactions and gene regulatory networks

The six hub genel predicted 10 drugs. These include albuterol, vasopressin, hydroxyurea, gallinamide a, bortezomib, tosedostat, chembl531269, chembl587247, umbelliferone, esculetin, scopoletin (Table S6). In addition, the drug-gene interactions were visualized by Cytoscape (Fig. 7). To elucidate the underlying mechanisms of ARGs, we established an extensive gene regulatory network (Fig. 8). Within this network, CTSLRA, UCP3, CTSLRG, UCP1 demonstrated robust interactions with ARGs, known for its roles in inflammation and immunity. Consequently, we hypothesize that ARGs potentially contributes to the pathogenesis of TED by modulating KLF6, PPP1R15A, BTG2, MYLIP, ANGPTL3, and associated downstream genes involved in immune and inflammatory pathways. Furthermore, albuterol, vasopressin, hydroxyurea emerges as a promising therapeutic candidate. These findings offer expansive avenues for future research endeavors in this domain.

Identification of common RNAs and construction of miRNA-LncRNA shared genes network

We conducted searches in three databases to identify 191 miRNAs and 191 lncRNAs linked with TED (Table S7a–b). Table S7 demonstrates the matching of these genes against the corresponding miRNA database. The databases used for this search were miRanda [19], miRDB [20], and TargetScan [21]. Each match with the relevant miRNA in these databases was assigned a score of 1. Notably, a match in all three databases received a score of 3. The corresponding lncRNA data was obtained by matching the miRNAs with the spongeScan database. The network of miRNAs-lncRNAs-genes was constructed by taking the intersection of these datasets with the shared genes obtained from Lasso regression and SVM-RFE. Finally, the miRNA-genes network included 157 lncRNAs, 169 miRNAs, and several common genes,

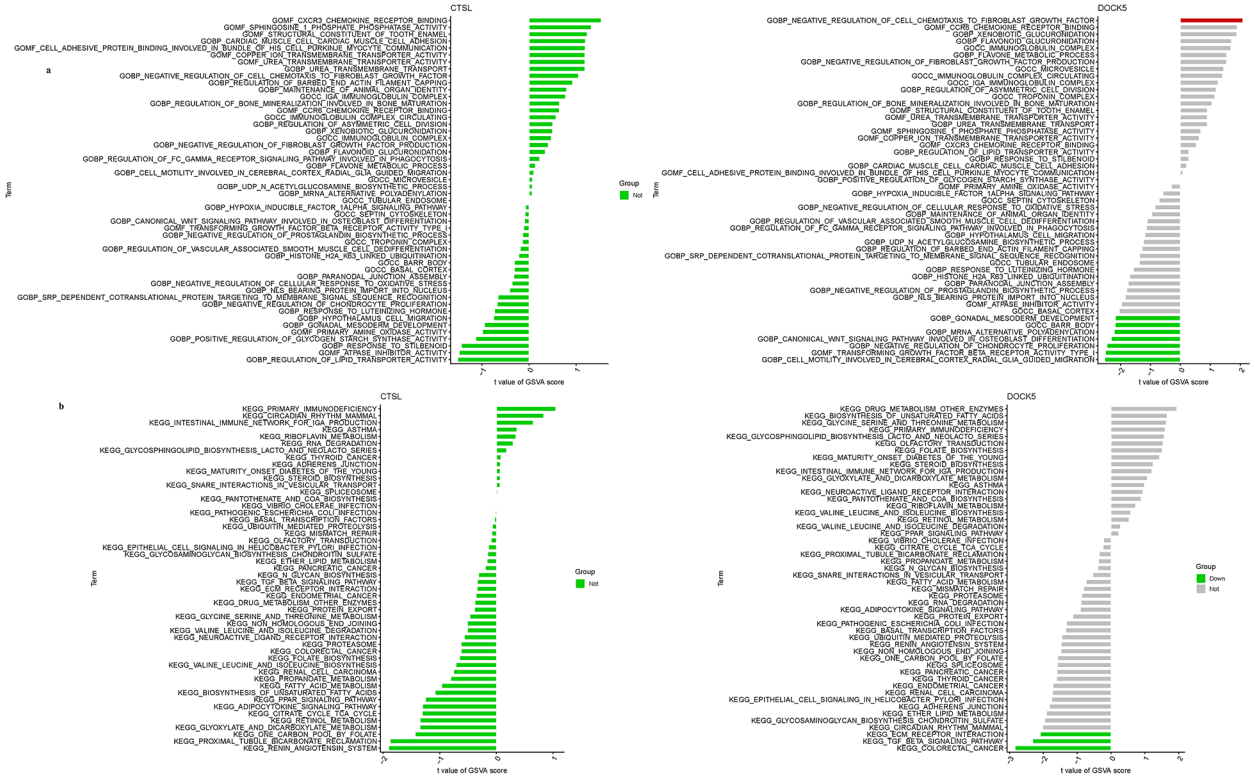

Fig. 6 GSVA of analysis in CTSL and DOCK5. (**a**) GO. (**b**) KEGG

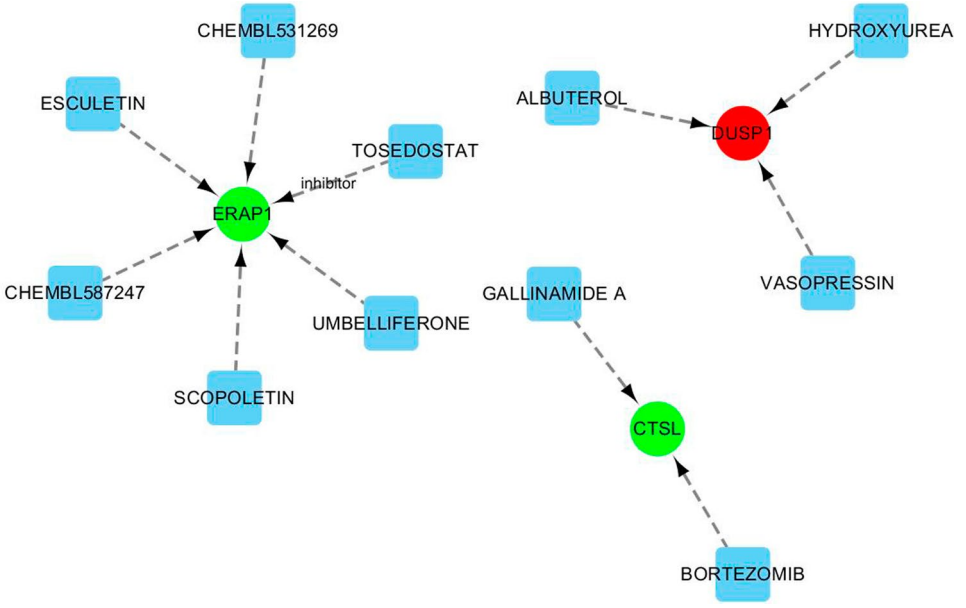

Fig. 7 Drug-gene interactions. *Note*: Red circles are up-regulated genes, green hexagons are down-regulated genes, and blue squares are associated drugs

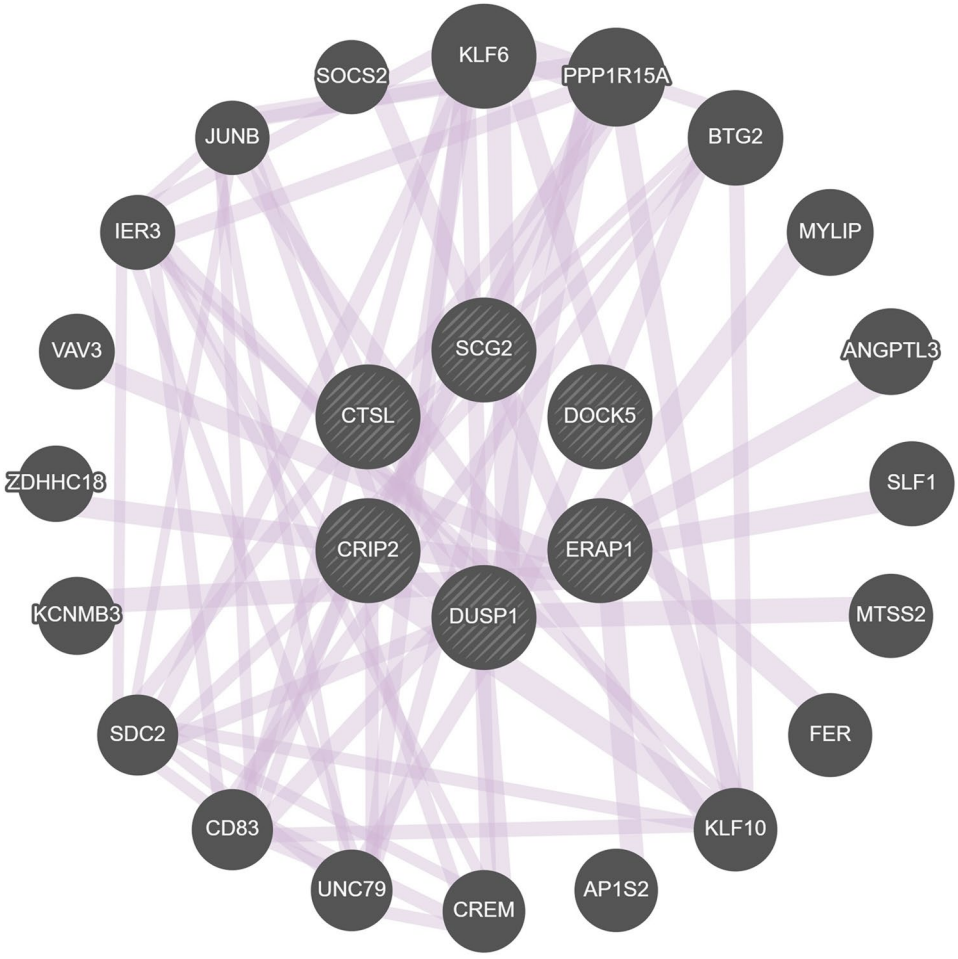

Fig. 8 Gene regulatory networks of ARGs

including five hub genes (DOCK5, SCG2, DUSP1, CRIP2, ERAP11) (Fig. 9).

Validation of hub genes and model

We used GSE105149 for validation of the hub genes. However, among the six ARGs, only CTSL and DOCK5 showed significant differences in the GSE105149 analysis (Fig. 10). Upon recalibration of the data, we observed differences in sample sizes between the two datasets, as well as differences in patient sources, which may have contributed to bias in the results. The boxplots illustrated the residual expression patterns of these genes in TED (Fig. 11a). There were differences in the proportions of the four different modes (Fig. 11b). The diagnostic capacity of the ARGs in distinguishing TED from control samples revealed a satisfactory diagnostic value, with an AUC of RF: 1.000; SVM: 1.000; XGB: 1.000; GLM: 1.000 (Fig. 11c). Additionally, an AUC of 1.000 (95% CI 1.000–1.000) was achieved in GSE105149 (Fig. 11d).

Discussions

TED is a rare inflammatory disorder of the orbit, often leading to visual impairment, facial deformity, and, in severe cases, permanent vision loss [22]. Ophthalmic manifestations include diplopia, dry eyes, and optic nerve damage, alongside soft tissue alterations such as proptosis, eyelid retraction, orbital fat expansion, and swelling [23]. TED is commonly classified into two distinct phases, as proposed by Francis Rundle. Risk factors such as smoking and exposure to radioactive iodine (RAI) are known to increase TED incidence, with their combined effect being additive [24]. Genetic studies have implicated polymorphisms in genes encoding human leukocyte antigen, cytotoxic T-lymphocyte antigen 4, interleukin 23 receptor, protein tyrosine thyroglobulin, and thyroid-stimulating hormone receptor in the disease's pathogenesis [25, 26]. Current therapeutic strategies focus on managing inflammation during the active phase [27]. Angiogenesis, the formation of new blood vessels from pre-existing vasculature, is a crucial process for growth, development, and wound healing, tightly regulated by a balance of pro-angiogenic factors such as VEGF and

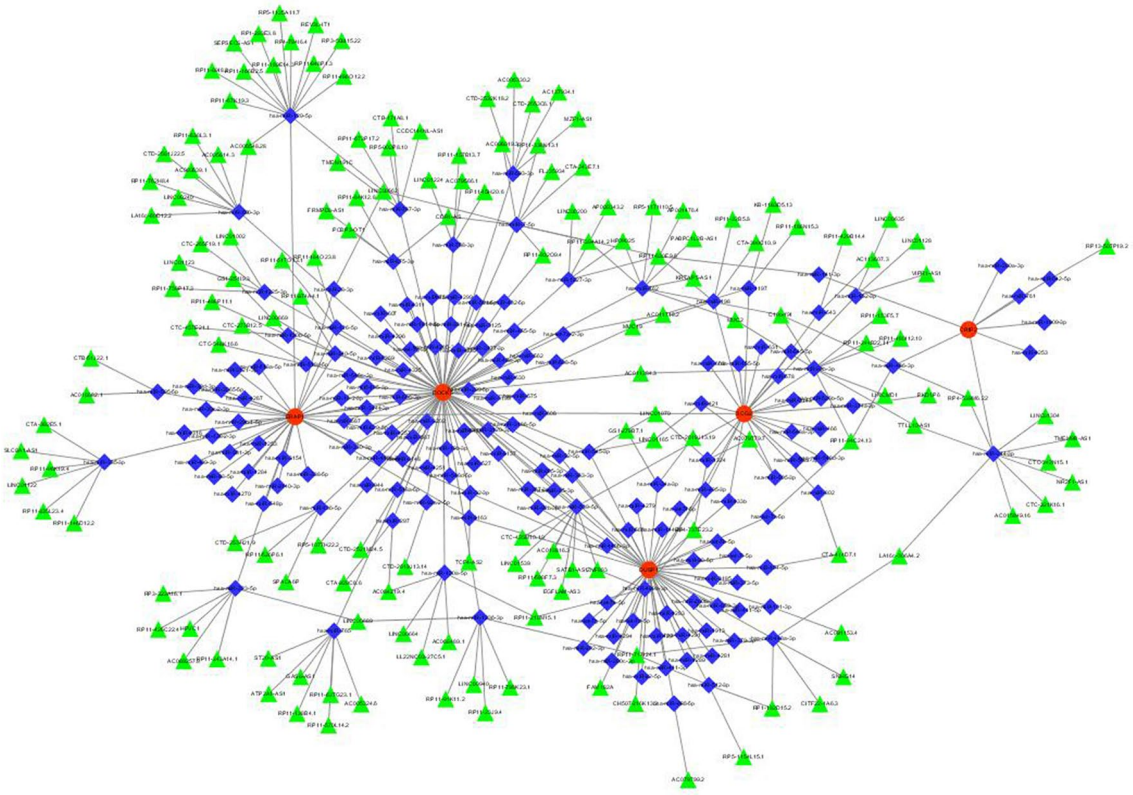

Fig. 9 miRNAs-LncRNAs shared genes network. *Note* Red circles are mrnas, blue quadrangles are miRNAs, and green triangles are lncRNAs

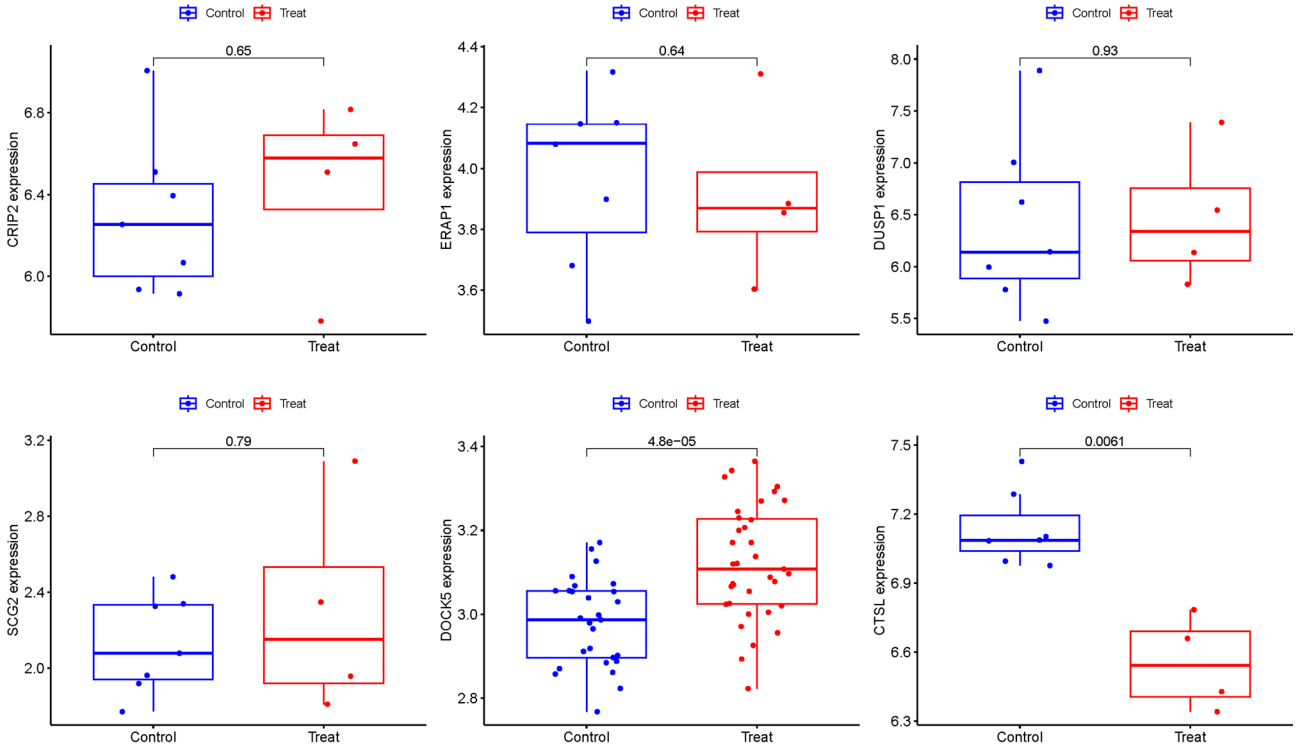

Fig. 10 Validation of hub genes

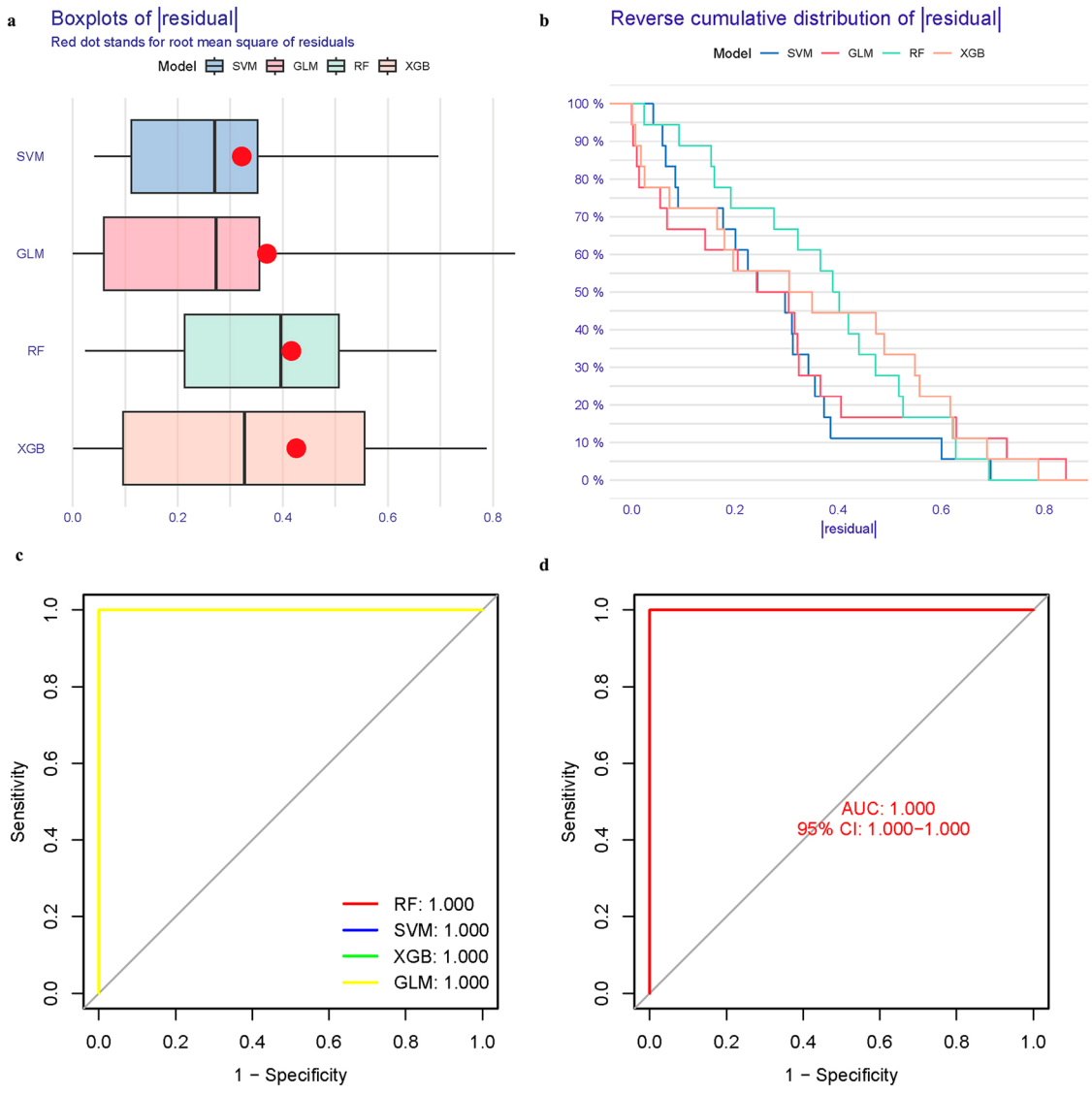

Fig. 11 Model verification. (**a**) Residual expression patterns. (**b**) Model expression patterns. (**c**) AUC of the model. (**d**) AUC of the test group

anti-angiogenic molecules. Dysregulated angiogenesis is a hallmark of numerous ocular diseases, leading to significant morbidity and vision loss [28]. Diabetic retinopathy, for instance, results from chronic hyperglycemia-induced VEGF overexpression, causing fragile, leaky retinal blood vessels, hemorrhages, and retinal detachment [29]. Similarly, age-related macular degeneration (AMD), particularly its neovascular form, is characterized by excessive VEGF production, stimulating the growth of choroidal neovascular membranes that leak fluid and blood, resulting in macular edema, scarring, and rapid vision loss [30]. In our study of TED, we identified 103 DEGs strongly associated with angiogenesis through an integrative analysis combining DEGs with advanced computational techniques. This underscores the pivotal role of ARGs in TED pathophysiology.

Through meticulous crossover analysis, we pinpointed six central ARGs (CRIP2, DUSP1, CTSL, DOCK5, ERAP1, SCG2) with significant diagnostic potential for TED. CRIP2, DUSP1, ERAP1, and SCG2 are implicated in a range of pathological conditions through their dysregulated expression. CRIP2, a zinc-finger protein, has been associated with neurodegenerative diseases, where its altered expression affects cellular processes such as apoptosis and synaptic plasticity [31]. DUSP1, a dual-specificity phosphatase, plays a critical role in modulating inflammatory responses and is frequently upregulated in conditions involving oxidative stress and immune dysregulation [32]. ERAP1, an aminopeptidase involved in antigen presentation, is linked to autoimmune disorders, with polymorphisms in its gene conferring susceptibility to diseases such as ankylosing spondylitis [33]. SCG2,

a neuroendocrine secretory protein, is dysregulated in various cancers and neurological disorders, suggesting its involvement in tumorigenesis and neuronal dysfunction [34]. Collectively, these genes underscore the complex molecular mechanisms underlying disease pathogenesis. Our study provides key insights into the interaction of specific genes with transcription factors in the context of angiogenesis. A comprehensive review of the literature highlights the pivotal roles of CTSL and DOCK5 in linking TED with ARGs. Further analysis of their biological functions reveals their involvement in immune-related processes, including oxidative stress response, aging, and peptide-mediated cellular reactions. These findings suggest that ARGs exert broad regulatory effects across multiple biological pathways, particularly in immune modulation. Our research emphasizes the complex interplay between angiogenesis, ARGs, and TED, offering a foundation for the development of novel therapeutic strategies to mitigate angiogenesis-related disease progression and enhance patient outcomes.

CTSL and DOCK5 play pivotal roles in the development of ocular diseases. Emerging evidence suggests that mutations or aberrant expression of these genes are implicated in various ophthalmic disorders. CTSL encodes a lysosomal protease crucial for protein degradation and extracellular matrix remodeling. Dysregulated expression of CTSL is closely associated with corneal diseases, retinal degenerative conditions, and glaucoma [35]. Studies indicate that upregulation of CTSL can lead to corneal stromal disruption, thereby precipitating corneal pathologies. Moreover, elevated CTSL expression in retinal tissues is linked to damage in retinal pigment epithelial cells and apoptosis of retinal ganglion cells, which are hallmark features of retinal degenerative diseases [36]. In glaucoma, research further suggests that CTSL may contribute to the pathogenesis by modulating intraocular pressure and optic nerve damage [37]. DOCK5 encodes a Rac-GEF protein involved in cytoskeletal reorganization and cell migration. Mutations in DOCK5 are associated with retinal dystrophies and macular degeneration. Functional impairments of DOCK5 in retinal pigment epithelial cells can lead to structural and functional degradation of the retina [38]. In patients with retinal dystrophy, DOCK5 mutations may cause photoreceptor cell degeneration, resulting in progressive vision loss. In macular degeneration, DOCK5 might influence the disease process by altering the dynamics of the extracellular matrix and retinal vasculature formation [39]. Our findings reveal a complex interplay between CTSL and DOCK5 in regulating retinal function and intraocular pressure, providing novel insights into the pathophysiology of ocular diseases. Using dataset GSE105149, our study explores the involvement of these genes in TED, suggesting that angiogenesis-related traits may serve as potential prognostic biomarkers.

A variety of clinical trials have delved into the research of numerous immune-modulators during the last 6–8 years, spurred by an improving understanding of the molecular processes underlying the pathogenesis of TED [40]. It is commonly assumed that TED is caused by thyroid autoantibodies, which cause a cross-reactive response on orbital fibroblasts to the thyrotropin receptor (TSHR) and the insulin-IGF-1 R. The ensuing cascade is triggered by B cell-driven processes and includes autoantigen recognition on orbital fibroblasts, subsequent T cell activation, recruitment of T cells, macrophages, and mast cells to the orbit, and an increasing release of proinflammatory cytokines [41, 42]. Therefore, orbital fibroblasts become activated, initiating a self-sustaining cycle of increased cytokine production, myofibroblast and adipocyte proliferation, and hyaluronic acid excretion. Strategies focusing on targeted therapy, fully rooted in this sophisticated pathophysiological paradigm, would obviously include T cell activation and depletion inhibition [43]. Agents targeting CD3, such as ciclosporin, otelixizumab, and teplizumab, have clearly demonstrated potential in limiting T cell activity. Furthermore, rituximab treatment provides as a supplementary therapeutic pathway for B cell depletion [44]. Notably, the inhibition of cytokine activity, exemplified by tocilizumab, and the employment of anti-TNF alpha monoclonal antibodies, have been pursued [45]. Furthermore, the comprehensive dampening of immune cell cycling through the application of antimetabolites like azathioprine and mycophenolate has exhibited encouraging outcomes within the realm of TED treatment trials [46]. These research further explored the expression patterns of RGs within the immunological microenvironment. T follicular helper cells and activated neutrophils were prominently expressed in the treatment group. TED presents significant therapeutic challenges, necessitating the development of novel treatment strategies. Future research must elucidate how ARGs influence inflammatory and immune responses in TED, thereby advancing our understanding of its pathogenesis and identifying potential therapeutic targets. Investigating the interactions between ARGs, immune cells, and cytokines will provide insights into the regulatory networks driving TED.

The study of biomarkers and their relationship to TED has gotten little attention in the existing literature. Several research are now being conducted to investigate the association between metabolism and eye problems using bioinformatics analysis [47–49]. For example, Liu et al. used WGCNA to identify TED hub genes. Hu et al. created a bioinformatics model for thyroid eye illness and discovered 11 hub genes. Huang et al. used advanced comprehensive bioinformatics analysis and in vivo validation to

identified six genes as relevant for diabetic retinopathy (CD44, CDC42, TIMP1, BMP7, RHOC, FLT1). However, no research has been conducted on the link between ARGs and TED. Our study presents an innovative methodological approach by integrating a comprehensive set of ARGs from GEO datasets to explore cell metabolism in the context of TED. Multiple machine learning methods were used in this study, and the results obtained by multiple machine learning methods were intersected. Then the key genes were subjected to functional enrichment analysis, GSEA analysis and GSVE analysis to explore the key biological functions and pathways. In addition, we also performed upstream and downstream gene prediction and related miRNA and LncRNA prediction based on the obtained key genes. It provides a reference for future research. This strategy distinguishes our research from prior studies by offering a multifaceted analysis that combines computational data with a focus on therapeutic strategies. By leveraging publicly available genomic resources, we have not only identified key ARGs involved in TED but also advanced our theoretical understanding of their potential roles in disease pathogenesis, thus paving the way for novel therapeutic insights. Despite these advancements, our investigation acknowledges the need for a deeper understanding of the molecular mechanisms through which ARGs interact with TED. The complexities of these interactions suggest that further in vivo and in vitro studies are essential to elucidate the precise biological processes at play. Our work also highlights the need for future exploration into the relationship between prognostic genes and familial aggregation in TED. By further investigating this connection, we anticipate uncovering important insights into the predictive and prognostic value of ARGs, which could refine the clinical management of TED and guide the development of targeted therapies. These directions underscore the potential of our methodological approach to spark future research that will enrich our knowledge of TED and improve therapeutic outcomes.

Conclusions

The pathogenesis and progression of TED involve complex, multifactorial interactions among various targets, pathways, signaling entities, and regulatory frameworks. CTSL and DOCK5 are key players, capable of exerting significant influence on aging program.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40360-025-00880-9.

Supplementary Material 1

Author contributions

Zixuan Wu and Jun Peng drafted and revised the manuscript. Xi Long and Kang Tan were in charge of data collection. Zixuan Wu and Jun Peng were in charge of design of frame. Xiaolei Yao and Qinghua Peng conceived and designed this article, in charge of syntax modification and revised of the manuscript. All the authors have read and agreed to the final version manuscript.

Funding

Financial support was provided by the National Natural Science Foundation of China (30772824,81574031); Key Laboratory of TCM Prevention and Treatment of Ent Diseases of Hunan Province (2017TP1018); Changsha Science and Technology Plan Project (K1501014-31, KC1704005); Central government financial support for the construction of local universities (2018-2019); State Administration of Traditional Chinese Medicine Key Discipline of Ophthalmology construction project; Key discipline construction project of TCM Five Senses Science in Hunan Province; Hunan Graduate Research Innovation Project (CX20240072); "Yifang" Graduate Innovation Project, Hunan University of Chinese Medicine (2022YF01). Hunan Science and Technology Talent Promotion Project (2023TJ-N25)

Data availability

The datasets generated during and/or analyzed during the current study are available in the appendix.

Declarations

Ethics approval and consent to participation

This manuscript is not a clinical trial, hence the ethics approval and consent to participation are not applicable.

Consent for publication

All authors have read and approved this manuscript to be considered for publication.

Competing interests

The authors declare no competing interests.

Author details

[1]Hunan University of Chinese Medicine, Changsha, Hunan Province 410208, China
[2]The First Hospital of Hunan University of Chinese Medicine, Changsha, Hunan Province 410007, China
[3]Ophthalmology Department, The First Hospital of Hunan University of Chinese Medicine, Changsha 410011, China

Received: 3 August 2024 / Accepted: 24 February 2025
Published online: 10 March 2025

References

1. Belliveau MJ, Jordan DR. Thyroid eye disease. CMAJ. 2013;185(9):797.
2. Ugradar S, Kang J, Kossler AL, Zimmerman E, Braun J, Harrison AR, Bose S, Cockerham K, Douglas RS. Teprotumumab for the treatment of chronic thyroid eye disease. Eye (Lond). 2022;36(8):1553–59.
3. Girnita L, Smith TJ, Janssen J. It takes two to Tango: IGF-I and TSH receptors in thyroid eye disease. J Clin Endocrinol Metab. 2022;107(Suppl_1):S1–S12.
4. Weiler DL. Thyroid eye disease: a review. Clin Exp Optom. 2017;100(1):20–25.
5. Douglas RS, Kahaly GJ, Patel A, Sile S, Thompson E, Perdok R, Fleming JC, Fowler BT, Marcocci C, Marino M, et al. Teprotumumab for the treatment of active thyroid eye disease. N Engl J Med 2020;382:341–52.
6. Nie T, Lamb YN. Teprotumumab: a review in thyroid eye disease. DRUGS. 2022;82(17):1663–70.
7. Jain AP, Jaru-Ampornpan P, Douglas RS. Thyroid eye disease: redefining its management-A review. Clin Exp Ophthalmol. 2021;49(2):203–11.
8. Men CJ, Kossler AL, Wester ST. Updates on the understanding and management of thyroid eye disease. Ther Adv Ophthalmol. 2021;13:970386640.
9. La Mendola D, Trincavelli ML, Martini C. Angiogenesis in Disease. Int J Mol Sci. 2022;23(18).

10. Ren Y, Luo H. Metformin: the next angiogenesis panacea? SAGE Open Med. 2021;9:393114839.
11. Costa L, Reis RL, Silva-Correia J, Oliveira JM. Microfluidics for Angiogenesis Research. Adv Exp Med Biol. 2020;1230:97–119.
12. Kretschmer M, Rudiger D, Zahler S. Mechanical aspects of Angiogenesis. Cancers (Basel). 2021;13(19).
13. Trabalzini L, Tundis R, Valacchi G, Evelson PA, Finetti F. Editorial: angiogenesis and nutraceuticals. Front Pharmacol. 2022;13:943158.
14. Li P, Yi N, Ding CS, et al. Research on classification diagnosis model of psoriasis based on deep residual network. Digit Chin Med 2021;4:92–101.
15. Peng CD, Wang L, Jiang DM, et al. Establishing and validating a spotted tongue recognition and extraction model based on multiscale convolutional neural network. Digit Chin Med 2022;5:49–58.
16. Zhou XY, Yang SW, Ou JT, et al. Screening influencing factors of blood stasis constitution in traditional Chinese medicine. Digit Chin Med 2022;5:169–77.
17. Wu Z, Liu P, Huang B, Deng S, Song Z, Huang X, Yang J, Cheng S. A novel Alzheimer's disease prognostic signature: identification and analysis of glutamine metabolism genes in immunogenicity and immunotherapy efficacy. Sci Rep. 2023;13(1):6895.
18. Su S, Tian L, Dong X, Hickey PF, Freytag S, Ritchie ME. CellBench: r/Bioconductor software for comparing single-cell RNA-seq analysis methods. Bioinformatics. 2020;36(7):2288–90.
19. De Carvalho TR, Giaretta AA, Teixeira BF, Martins LB. New bioacoustic and distributional data on Bokermannohyla sapiranga Brandao et al., 2012 (Anura: hylidae): revisiting its diagnosis in comparison with B. pseudopseudis (Miranda-Ribeiro, 1937). Zootaxa. 2013;3746:383–92.
20. Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res. 2020;48(D1):D127–D131.
21. Mon-Lopez D, Tejero-Gonzalez CM. Validity and reliability of the TargetScan ISSF Pistol & Rifle application for measuring shooting performance. Scand J Med Sci Sports. 2019;29(11):1707–12.
22. Hoang TD, Stocker DJ, Chou EL, Burch HB. 2022 Update on clinical management of graves disease and thyroid eye disease. Endocrinol Metab Clin North Am. 2022;51(2):287–304.
23. Khong JJ, McNab A. Medical treatment in thyroid eye disease in 2020. Br J Ophthalmol. 2021;105(3):299–305.
24. Lee MH, Chin YH, Ng CH, Nistala K, Ow Z, Sundar G, Yang SP, Khoo CM. Risk factors of thyroid eye disease. Endocr Pract. 2021;27(3):245–53.
25. Rashad R, Pinto R, Li E, Sohrab M, Distefano AG. Thyroid eye disease. Life (Basel). 2022;12(12).
26. Szelog J, Swanson H, Sniegowski MC, Lyon DB. Thyroid eye disease. Mo Med. 2022;119(4):343–50.
27. Yu CY, Ford RL, Wester ST, Shriver EM. Update on thyroid eye disease: regional variations in prevalence, diagnosis, and management. Indian J Ophthalmol. 2022;70(7):2335–45.
28. Vega R, Carretero M, Bonilla LL. Anomalous Angiogenesis in Retina. Biomedicines. 2021;9(2).
29. Liang JH, Akhanov V, Ho A, Tawfik M, D'Souza SP, Cameron MA, Lang RA, Samuel MA. Dopamine signaling from ganglion cells directs layer-specific angiogenesis in the retina. Curr Biol. 2023;33(18):3821–34.
30. Kitahata S, Inoue M, Yanagi Y, Tanaka S, Inoue T, Yamakawa T, Uchio E, Kadonosono K. Angiogenesis and Anastomosis on Graft Retina after Autologous Retinal Transplantation. Ophthalmol Retina. 2022;6(10):972–74.
31. Chen L, Li N, Zhang M, Sun M, Bian J, Yang B, Li Z, Wang J, Li F, Shi X, et al. APEX2-based Proximity Labeling of Atox1 Identifies CRIP2 as a nuclear copper-binding protein that regulates autophagy activation. Angew Chem Int Ed Engl 2021;60:25346–55.
32. Shi L, Zha H, Pan Z, Wang J, Xia Y, Li H, Huang H, Yue R, Song Z, Zhu J. DUSP1 protects against ischemic acute kidney injury through stabilizing mtDNA via interaction with JNK. Cell Death Dis. 2023;14(11):724.
33. Reeves E, Islam Y, James E. ERAP1: a potential therapeutic target for a myriad of diseases. Expert Opin Ther Targets. 2020;24(6):535–44.
34. Weng S, Liu Z, Ren X, Xu H, Ge X, Ren Y, Zhang Y, Dang Q, Liu L, Guo C, et al. SCG2: a prognostic marker that pinpoints chemotherapy and immunotherapy in colorectal cancer. Front Immunol. 2022;13:873871.
35. Xu J, Xu W, Choi J, Brhane Y, Christiani DC, Kothari J, McKay J, Field JK, Davies M, Liu G, et al. Large-scale whole exome sequencing studies identify two genes,CTSL and APOE, associated with lung cancer. PLOS Genet 2023;19:e1010902.
36. Huang L, Lou N, Xie T, Tang L, Han X, Shi Y. Identification of an antigen-presenting cells/T/NK cells-related gene signature to predict prognosis and CTSL to predict immunotherapeutic response for lung adenocarcinoma: an integrated analysis of bulk and single-cell RNA sequencing. Cancer Immunol Immunother. 2023;72(10):3259–77.
37. Zhang L, Zhao Y, Yang J, Zhu Y, Li T, Liu X, Zhang P, Cheng J, Sun S, Wei C, et al. CTSL, a prognostic marker of breast cancer, that promotes proliferation, migration, and invasion in cells in triple-negative breast cancer. Front Oncol. 2023;13:1158087.
38. Qu H, Liu X, Zhu J, Xiong X, Li L, He Q, Wang Y, Yang G, Zhang L, Yang Q, et al. Dock5 deficiency promotes proteinuric kidney diseases via modulating podocyte lipid metabolism. Adv Sci (Weinh) 2024;11:e2306365.
39. Liu C, Li G, Zheng S, She L, Lu S, Wang Y, Huang D, Zhang X, Sun L, Liu Y, et al. PHF5A regulates the expression of the DOCK5 variant to promote HNSCC progression through p38 MAPK activation. Biol Direct 2023;18:39.
40. Rosenbaum JT, Choi D, Wong A, Wilson DJ, Grossniklaus HE, Harrington CA, Dailey RA, Ng JD, Steele EA, Czyz CN, et al. The role of the immune response in the pathogenesis of thyroid eye disease: a reassessment. PLoS ONE 2015;10:e137654.
41. Nallu R, Madhavan P, Chirch L, Luthra P. Thyroid eye disease due to immune reconstitution inflammatory syndrome as a consequence of antiretroviral therapy in the setting of AIDS. Case Rep Endocrinol. 2020;2020:1728423.
42. Lehmann GM, Feldon SE, Smith TJ, Phipps RP. Immune mechanisms in thyroid eye disease. Thyroid. 2008;18(9):959–65.
43. Sagiv O, Kandl TJ, Thakar SD, Thuro BA, Busaidy NL, Cabanillas M, Jimenez C, Dadu R, Graham PH, Debnam JM, et al. Extraocular muscle enlargement and thyroid eye disease-like orbital inflammation associated with immune checkpoint inhibitor therapy in cancer patients. Ophthalmic Plast Reconstr Surg 2019;35:50–52.
44. Chau C, Shih KC, Chow L, Lee V. Considerations for use of immune checkpoint inhibitors in cancer therapy for patients with co-existing thyroid eye disease. Ophthalmol Ther. 2021;10(1):5–12.
45. Yoon JS, Kikkawa DO. Thyroid eye disease: from pathogenesis to targeted therapies. Taiwan J Ophthalmol. 2022;12(1):3–11.
46. Sears CM, Azad AD, Amarikwa L, Pham BH, Men CJ, Kaplan DN, Liu J, Hoffman AR, Swanson A, Alyono J, et al. Hearing dysfunction after treatment with Teprotumumab for thyroid eye disease. Am J Ophthalmol. 2022;240:1–13.
47. Liu H, Chen L, Lei X, Ren H, Li G, Deng Z. Identification of hub genes associated with nonspecific orbital inflammation by weighted gene coexpression network analysis. Dis Markers. 2022;2022:7588084.
48. Hu J, Zhou S, Guo W. Construction of the coexpression network involved in the pathogenesis of thyroid eye disease via bioinformatics analysis. Hum Genomics. 2022;16(1):38.
49. Huang J, Zhou Q. Gene biomarkers related to Th17 cells in Macular Edema of diabetic retinopathy: cutting-edge comprehensive bioinformatics analysis and in vivo validation. Front Immunol. 2022;13:858972.

Publisher's Note