Single-cell RNA Sequencing Identifies Natural Kill Cell-Related Transcription Factors Associated With Age-Related Macular Degeneration

Yili Luo[1], Jianpeng Liu[2], Wangqiang Feng[1], Da Lin[1], Mengji Chen[1] and Haihua Zheng[1] (iD)

[1]Department of Ophthalmology, The Second Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang, China. [2]Department of Pathology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, Zhejiang, China.

ABSTRACT

BACKGROUND: Age-related Macular Degeneration (AMD) poses a growing global health concern as the leading cause of central vision loss in elderly people.

OBJECTION: This study focuses on unraveling the intricate involvement of Natural Killer (NK) cells in AMD, shedding light on their immune responses and cytokine regulatory roles.

METHODS: Transcriptomic data from the Gene Expression Omnibus database were utilized, employing single-cell RNA-seq analysis. High-dimensional weighted gene co-expression network analysis (hdWGCNA) and single-cell regulatory network inference and clustering (SCENIC) analysis were applied to reveal the regulatory mechanisms of NK cells in early-stage AMD patients. Machine learning models, such as random forests and decision trees, were employed to screen hub genes and key transcription factors (TFs) associated with AMD.

RESULTS: Distinct cell clusters were identified in the present study, especially the T/NK cluster, with a notable increase in NK cell abundance observed in AMD. Cell-cell communication analyses revealed altered interactions, particularly in NK cells, indicating their potential role in AMD pathogenesis. HdWGCNA highlighted the turquoise module, enriched in inflammation-related pathways, as significantly associated with AMD in NK cells. The SCENIC analysis identified key TFs in NK cell regulatory networks. The integration of hub genes and TFs identified *CREM, FOXP1, IRF1, NFKB2*, and *USF2* as potential predictors for AMD through machine learning.

CONCLUSION: This comprehensive approach enhances our understanding of NK cell dynamics, signaling alterations, and potential predictive models for AMD. The identified TFs provide new avenues for molecular interventions and highlight the intricate relationship between NK cells and AMD pathogenesis. Overall, this study contributes valuable insights for advancing our understanding and management of AMD.

KEYWORDS: Age-related macular degeneration, NK cell, transcription factors, machine learning, single-cell RNA sequencing

Introduction

As the global population ages, projections indicate a rise in the number of age-related macular degeneration (AMD) patients, reaching 196 million by 2020 and escalating to 288 million by 2040.[1] AMD, characterized by the degeneration of the macula, results in damage to photoreceptors and is the leading cause of central vision loss in elderly people.[2,3] The pathogenesis of AMD involves various factors, including genetics, environment, and immunology.[4-6] Environmental and lifestyle variables, such as smoking, hypertension, and body mass index, exert a significant influence on AMD risk.[7] Similarly, genetic factors play a crucial role, with individuals having a family history of AMD facing an elevated risk compared to those without such familial predisposition.[8]

There is an increasing consensus suggesting that immunology also plays an important role. Despite the immune privilege status of the eye, abnormal immune and inflammatory responses occur, intensifying during the development of AMD.[9,10] Complement C3a, C3d, Ba, Bb, C5a, and CFH levels are reportedly higher in AMD patients than healthy controls.[11] Besides, the abnormal recruitment of macrophages/microglia to tissue lesions reportedly contributes to the development of AMD.[12] It is now understood that Natural Killer (NK) cells, integral to the immune system, impact cytokine production and cytotoxicity. They exert antiviral and antitumoral effects by killing cells with low expression of the MHC antigen presentation system.[13] Goverdhan et al[14] found that the NK receptor AA haplotype, coupled with the human leukocyte antigen (HLA)-Cw*0701 allele, was associated with AMD. In an additional study of gene expression profiles related to AMD, the hub genes, including *C1S, ADM*, and *1ER5L*, exhibited a positive correlation with the activation of NK cells during the progression of AMD.[15]

Transcription factors (TFs), as key regulators of gene expression, play a vital role in NK cell maturation, development, and

differentiation. Ets family members play a regulatory role in the development of NK cells. Depletion of ETS proto-oncogene 1 (Ets-1) in mice results in a significantly diminished NK cell compartment.[16] TFs of the notch family operate as the common lymphoid progenitor (CLP) stage, contributing to lymphocyte development. Their role involves, at least in part, inhibiting the commitment to a myeloid lineage fate. Notably, the Notch ligands Jagged1/2 and Delta-like protein-1 demonstrate a preference for fostering the in vitro development of NK cells.[17,18]

The advancement of single-cell sequencing technology presents an opportunity to explore NK cell subpopulations and examine gene expression patterns related to transcription factors. Employing high-dimensional weighted gene coexpression network analysis (hdWGCNA) and single-cell regulatory network inference and clustering analysis (SCENIC), we identified hub transcription factor genes in NK cells within the choroid of early-stage AMD patients. By revealing the expression profile, our objective is to comprehensively grasp the distinctive regulatory mechanisms of these transcription factors in the pathogenesis of AMD. This approach provides a new perspective for a more comprehensive understanding of AMD's underlying mechanisms.

Materials and Methods

Data acquisition and processing

Transcriptomic data were collected from the Gene Expression Omnibus (GEO) database. Single-cell analyses were conducted using dataset GSE203499.The GSE203499 dataset includes samples from 20 healthy controls, 24 individuals with early AMD, and 6 donors with geographic atrophy. The majority of samples were collected from four distinct anatomical regions of the eye: the macular choroid, macular retina, peripheral choroid, and peripheral retina. We only included samples from the macular choroid for the healthy control and early AMD groups. Machine learning models were trained with datasets GSE135092 and GSE29801, and accuracy assessments were conducted with dataset GSE99248 (sFigure 1). The GSE135092 dataset comprises samples from 106 donors with no history of ocular diseases and 23 donors previously diagnosed with AMD, collected from the macular or peripheral regions of donor eyes. The GSE29801 dataset includes 177 samples from the macular or extramacular regions of RPE-choroids and 118 samples from the macular or extramacular regions of retinas, encompassing cases with no reported ocular disease, possible preclinical AMD, or AMD. The GSE99248 dataset contains retinal and RPE-choroid-scleral tissue punches from 8 normal and 8 AMD donor eye samples. For all dataset, we only enrolled samples from the macular choroid for the healthy control and AMD groups

Single-cell RNA-seq analysis

The data from dataset GSE203499 were further analyzed using Seurat in R.[19] Potential doublets and low-quality cells with both small and large library sizes within individual datasets were systematically excluded. The application of Sctransform facilitated the normalization and variance stabilization of molecular count data. The identification of differentially expressed marker genes for each cluster was achieved using the FindAllMarkers function. To visualize the distribution of cells, Uniform Manifold Approximation and Projection (UMAP) analysis was conducted.

Cell–cell communication analysis

To infer and analyze cell-cell communication, we employed CellChat, a comprehensive repository encompassing ligands, receptors, cofactors, and their interactions.[20] Facilitated by the versatile and user-friendly toolkit CellChat, we explored novel intercellular communications and constructed cell-cell communication atlases. For the analysis of cell interactions, expression levels were computed relative to the total count mapped to the identical set of coding genes across all transcriptomes. The resulting expression values were then averaged within each single-cell cluster.

Differential expression analysis and gene set enrichment analysis (GSEA)

In comparing NK cells between the AMD and Normal groups, we conducted differential expression analysis. Differentially expressed genes (DEGs) were identified through pseudobulk RNA analysis, where the RNA count values for each gene were aggregated across all cells in each sample.[21] DESeq2 was then employed for the differential expression analysis.[22] Additionally, GSEA was utilized to investigate the functional pathways enriched by the differentially expressed genes.

Pseudo–time analysis

Monocle2 was utilized to sequence the 2 CD8+ T cell clusters along a "pseudotime" trajectory.[23] Following successful quality control, genes were incorporated into Monocle's Reversed Graph Embedding algorithm to define the trajectory. Subsequently, Monocle executed dimensionality reduction on the data, resulting in the ordered arrangement of cells based on pseudotime.

hdWGCNA analysis

hdWGCNA is a tool designed for conducting weighted gene co-expression network analysis (WGCNA) on high-dimensional data, including single-cell RNA-seq.[24] It serves to construct cell type-specific co-expression networks, identify resilient modules of interconnected genes and offer the biological context for these modules. In this study, hdWGCNA was employed to identify the key module associated with NK/T clusters. Subsequently, hub genes within this key module were identified through screening.

SCENIC analysis

SCENIC represents a gene regulatory networks (GRNs) algorithm tailored for the analysis of single-cell data.[25] SCENIC introduces a gene co-expression network inferred through TF motifs, subsequently delineating high-reliability GRNs primarily governed by TFs. In this study, the SCENIC algorithm was employed to unravel the specific network of transcription factors associated with NK cells.

Machine learning

To filter the variables with high accuracy and stable performance, we used the method of nested resampling in the mlr3 package.[26] The best model was selected among 4 machine learning models, including logistic regression (LR), random forest (ranger), decision trees (rpart), and adaptive best subset selection (abess). For comparing the performance of different learners, we used the benchmarking function of mlr3 parkage in the training set (GSE135092 and GSE29801). Finally, we used the data from the testing set (GSE99248) to test the performance of each learner and screen the hub gene.

Statistical analyses

Statistical analyses were performed using R software (Version 4.3). Analyses were conducted using two-tailed unpaired *t*-test. Statistical significance was defined as ns = $P > .05$; *$P < .05$; **$P < .01$; ***$P < .001$; and ****$P < .0001$.

Results

Single-cell RNA-seq profiling of different cell clusters in AMD patients and healthy controls

To depict the landscape and dynamics of choroid in the macula, single-cell transcriptomes were assessed with subcluster analysis and visualized by the UMAP approach (Figure 1A). A total of 55 543 cells from AMD and normal eye tissue were classified into 29 clusters (Figure 1B). Based on the gene signature of distinct subclusters, these cells were mainly composed of 5 immune cell clusters, including T lymphocyte and NK cells (T/NK: *PTPRC, CD2*), fibroblasts (Fibro: *DCN, CFH*), pericyte (Peri: *ACTA2, RGS*), melanocyte (Melano: *MITF, MLANA*), schwann cell (Schwa: *CDH19, PLP1*), macrophyte (Macro: *AIF1, CD68*), B lymphocyte (B lym: *CD79A, MS4A1*), erythrocytes (Eryth: *HBB, HBA2*), endothelium cells (Endo: *VWF, PECAM1*), mast cells (Mast: *KIT, CPA3*), retinal pigment epithelium cells (RPE: *RPE65, RLBP1*), Plasma cells (Plasma: *MZB1, SDC1*), Rod cells (ROD: *NRL, PDE6A*) (Figure 1B and C). The T/NK subcluster comprised 16 144 cells, the highest among all cell clusters (Figure 1D). To elucidate the origin of all clusters, a total of 30 479 cells derived from AMD samples and 25 064 cells derived from normal samples were independently subjected to clustering (Figure 1E). We further quantified the absolute numbers and

relative contribution of all clusters in different disease stages (Figure 1F and G). Given that the NK/T subcluster constitutes the highest proportion among all subtypes, it is significant to note that NK cells and T cells play pivotal roles in immune responses and inflammation. Consequently, the NK/T clusters were subjected to further exploration and decoding at the single-cell level.

Characterization of the identities of sub-cluster in NK/T cell

In the overall cluster analysis, considering the close spatial distribution of T cells and NK cells in the UMAP plot, we initially grouped T cells and NK cells together as a single cluster for analysis. However, for more in-depth exploration, the NK/T cluster was re-clustered, resulting in the identification of 5 distinct cell subclusters (Figure 2A). The absolute numbers and relative contributions of all clusters at different disease stages revealed significant differences between AMD and normal control, particularly for the T/NK_3 and T/NK_4 subclusters (Figure 2B and C). Based on marker genes, we further characterized these 5 clusters (Figure 2D). *CD3D* and *CD3E* were commonly expressed in T cells, whereas the T/NK_3 subcluster was the only one that did not express *CD3D* and *CD3E*. Simultaneously, it exhibited high expression of NK cell-related genes such as *KLRB1, NKG7*, and *GZMB*. Consequently, we defined T/NK_3 as NK cells. Specific marker genes *CD4* and *CD40LG* were expressed in T/NK_2, indicating that these clusters represented CD4 T cells (Figure 2D). Additionally, T/NK_1, T/NK_4, and T/NK_5 expressed *CD8A* and *CD8B*, specific markers of CD8 T cells. However, the T/NK_5 subcluster also showed an enrichment of NK cell marker genes, suggesting that T/NK_5 might be NKT cells (Figure 2D). Finally, T/NK_1 and T/NK_4 were confirmed as CD8 T cells (Figure 2D). The results indicated a marked increase in the abundance of NK cells in AMD. To elucidate a more refined landscape of T cell and NK cell subsets in AMD, the 2 CD8+ T cell subclusters were further characterized. FindMarkers differential expression analysis between T/NK_1 and T/NK_4 performed in Seurat revealed T/NK_1 exhibited significant upregulation of pro-inflammatory genes, including CXCL13, IGLC1, IFNG, and LIME1 (Figure 2E).

To explore potential disparities between the 2 CD8 T subclusters, we conducted pseudo-time and trajectory analysis. After mapping all cell labels (pseudotime, state, disease, and cell type) onto the pseudotime trajectory (Figure 2F), we identified 3 branch points and 7 states for the 2 CD8 T subclusters (Figure 2F). Gene expression changes along pseudotime revealed 3 distinct patterns associated with the differentiation trajectory of CD8 T cells (Figure 2G). All cells originated from the Naïve T-related pattern, characterized by high expression of LEF1, SELL, and CCR7. During the mid-to-late timeline, there was an enrichment of the memory T cell-related pattern, characterized by increased expression of CPR183, PRF1,
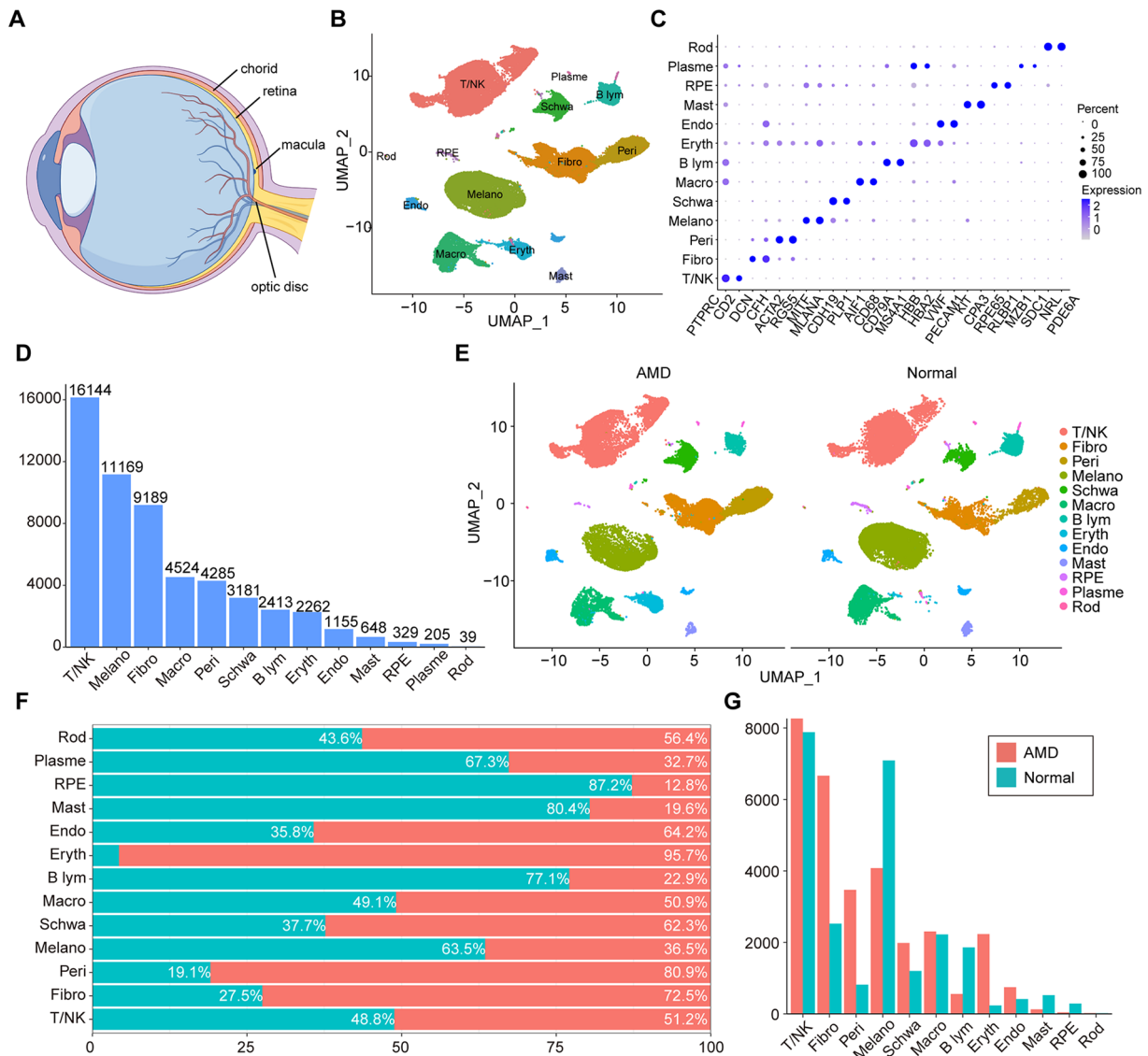
Figure 1. Single-cell RNA-seq profiling of different cell clusters in age-related macular degeneration (AMD) patients and health control: (A) eye anatomy; (B) uniform manifold approximation and projection (UMAP) visualization of all the single cells, (C) dot plot displaying the key marker gene of each cell cluster, (D) bar plot representing the cell number of each cell, (E) UMAP visualization of different originated from AMD and normal patients, (F) bar plot representing the relative contribution of cells originated from AMD and normal for each cell type, and (G) bar plot representing the number of cells originated from AMD and normal for each cell type.

CST7, CD69, and IL15RA. The last stage was associated with the expression of NKG7, GZMA, IFNG, and CXCR3, representing the cytotoxic T cell-related pattern (Figure 2F and G). Given the T/NK_4 subcluster was mostly enriched in the latest timeline, we defined it as effector CD8 T cells, while the T/NK_1 cluster was labeled as memory/naïve CD8 T cells.

Differential disease states of NK cells reveal the heterogeneity of cellular communication levels and function characteristics

After identifying all subclusters of the T/NK cluster, we ultimately identified four clusters, including CD4T, CD8T, NK, and NKT cells. To further elucidate the underlying impact of the T/NK cluster in AMD, we conducted an analysis of cellular communication within all cell clusters. The number of cell interactions and the interaction strength among all cell clusters indicated that CD8T, NK, Macro, Endo, and NKT cells exhibited higher interaction strength among clusters (Figure 3A). Additionally, we performed a ligand-receptor pairs interaction analysis to gain insight into the differences between AMD and Normal groups. By comparing the information flow in the AMD condition with Normal, we identified some conserved and context-specific signaling pathways (sFigure 2A). Interestingly, the complement system was found to increase in AMD, identified as a key factor in the pathogenesis of AMD in a previous study.[27] Conversely, growth factor signaling, such as HGF, VEGF, IGF, and FGF, decreased in the AMD group compared with the Normal group. Subsequently, we visualized the differential number of interactions and interaction strength
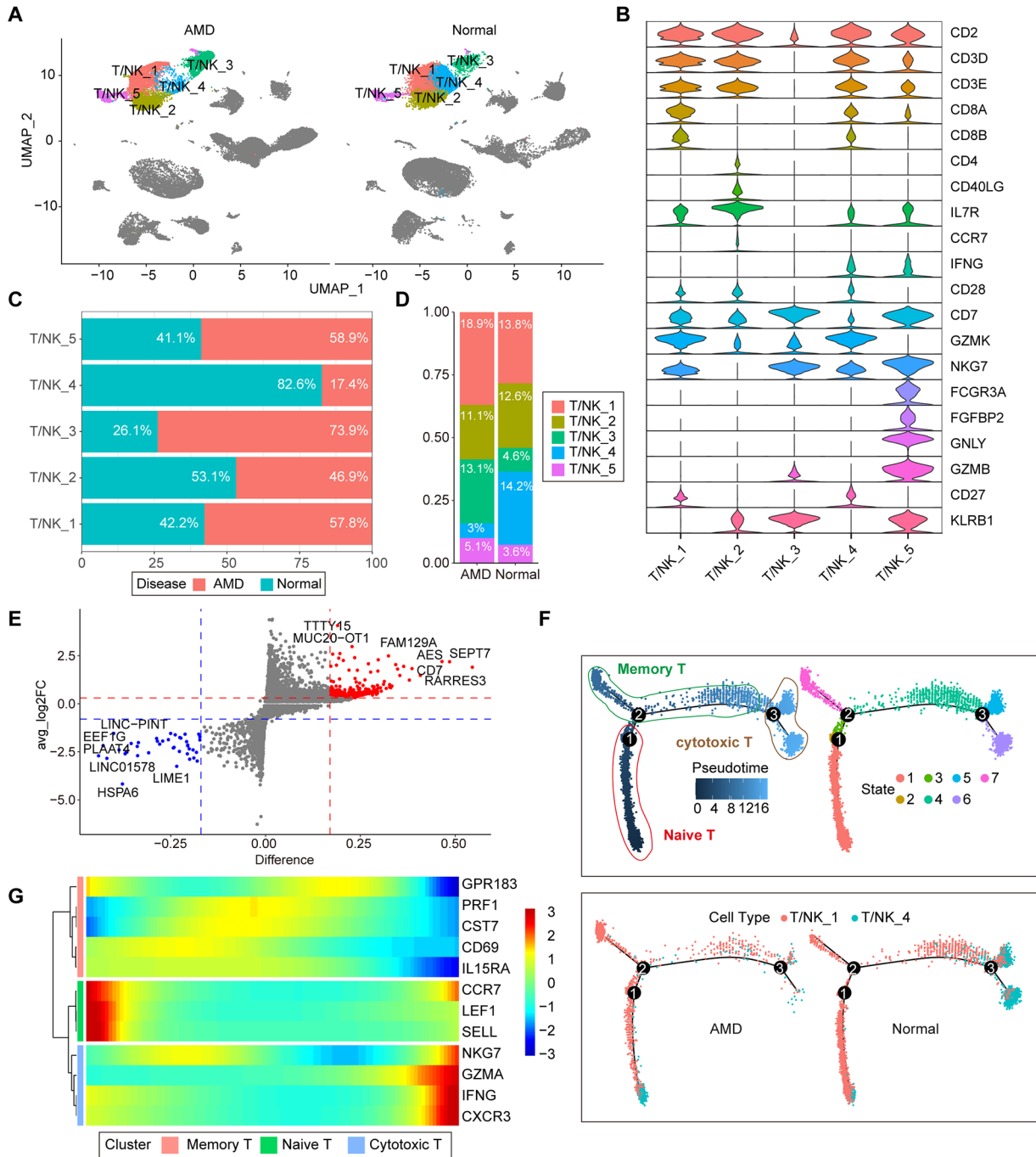
Figure 2. Characterization of the identities of sub-cluster in NK/T cell: (A) UMAP visualization of each NK/T cluster, (B) violin plot displaying the expression of key marker genes in each NK/T cluster, (C) bar plot representing the relative contribution of cells from AMD and normal for each, (D) bar plot representing the relative proportions of each cluster in AMD and normal, (E) volcano plot displaying the differential gene expression analysis between NK/T_1 and NK/T_4 cell, (F) differentiation trajectory of NK/T_1 and NK/T_4, with a color code for pseudo-time, state, disease original, and cell type, and (G) heatmap plot representing the expression of T cell specified marker genes along pseudo-time.

in the cell-cell communication network. The results showed a significant decrease in incoming and outgoing interaction strength in NK cells in the AMD group compared to the Normal group (Figure 3B and sFigure 2B).

Given the significant alterations in the quantity of NK cells in AMD and the notable differences in cell interactions, we decided to focus on NK cells for the next phase of our research. We further identified the specific signaling changes of NK cells

in AMD compared to normal samples, with SEMA4 specifically present in AMD (Figure 3C). In addition, we performed differential expression analysis (DEA) and GSEA analysis of NK cells between AMD and Normal. A volcano plot was generated to visualize the top 10 upregulated and downregulated genes (Figure 3D). IL-related genes were significantly upregulated, while the HLA family was notably decreased in AMD. The differential genes underwent KEGG enrichment analysis,
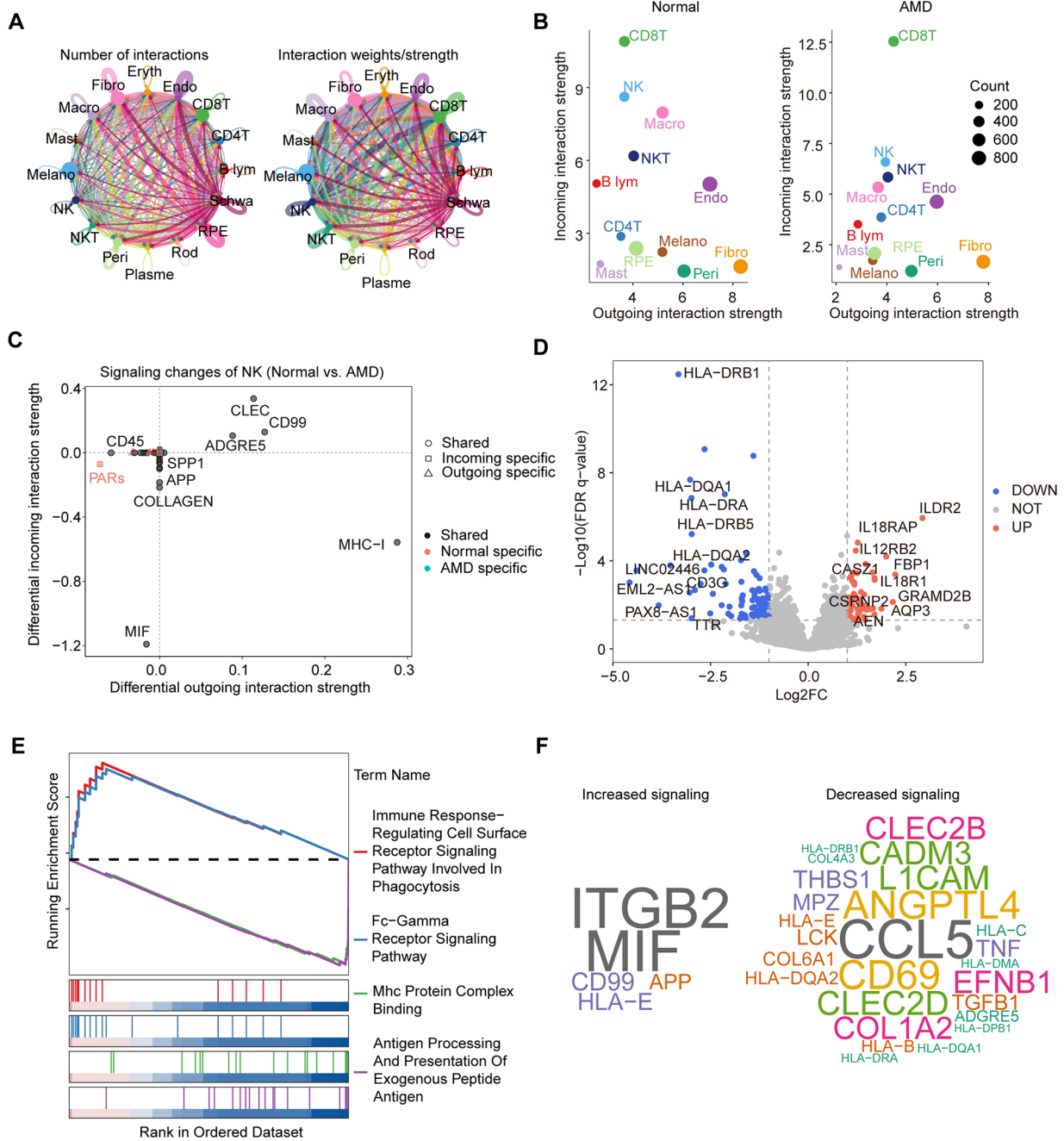
Figure 3. Differential disease states of NK cells reveal the heterogeneity of cellular communication levels and function characteristics: (A) analysis of cell-cell interaction in all cells, (B) dot plot representing the cell with significant changes in sending or receiving signals between AMD and normal, (C) dot plot representing the specific signaling changes of NK cell, (D) volcano plot of differential gene expression analysis of NK cell between AMD and normal, (E) gene set enrichment analysis (GSEA) for differential gene expression, and (F) wordcloud plot of down-regulated signaling ligand-receptor pairs between AMD and normal.

revealing a highly activated immune response in AMD. Conversely, MHC-related antigen processing signaling was suppressed in AMD compared to Normal (Figure 3E and sFigure 2C). Cellchat, with its capability to identify upregulated and downregulated signaling ligand-receptor pairs based on DEA, produced results consistent with GSEA, indicating an increased immune response in AMD (Figure 3F).

hdWGCNA revealed that NK cell is characterized by the turquoise module

We conducted hdWGCNA to investigate gene modules in NK cells and identified 9 modules (Figure 4A). Subsequently, we constructed the co-expression network for each gene module, highlighting the 3 hub genes (Figure 4B). The co-expression modules were visualized using the hub gene signature in the
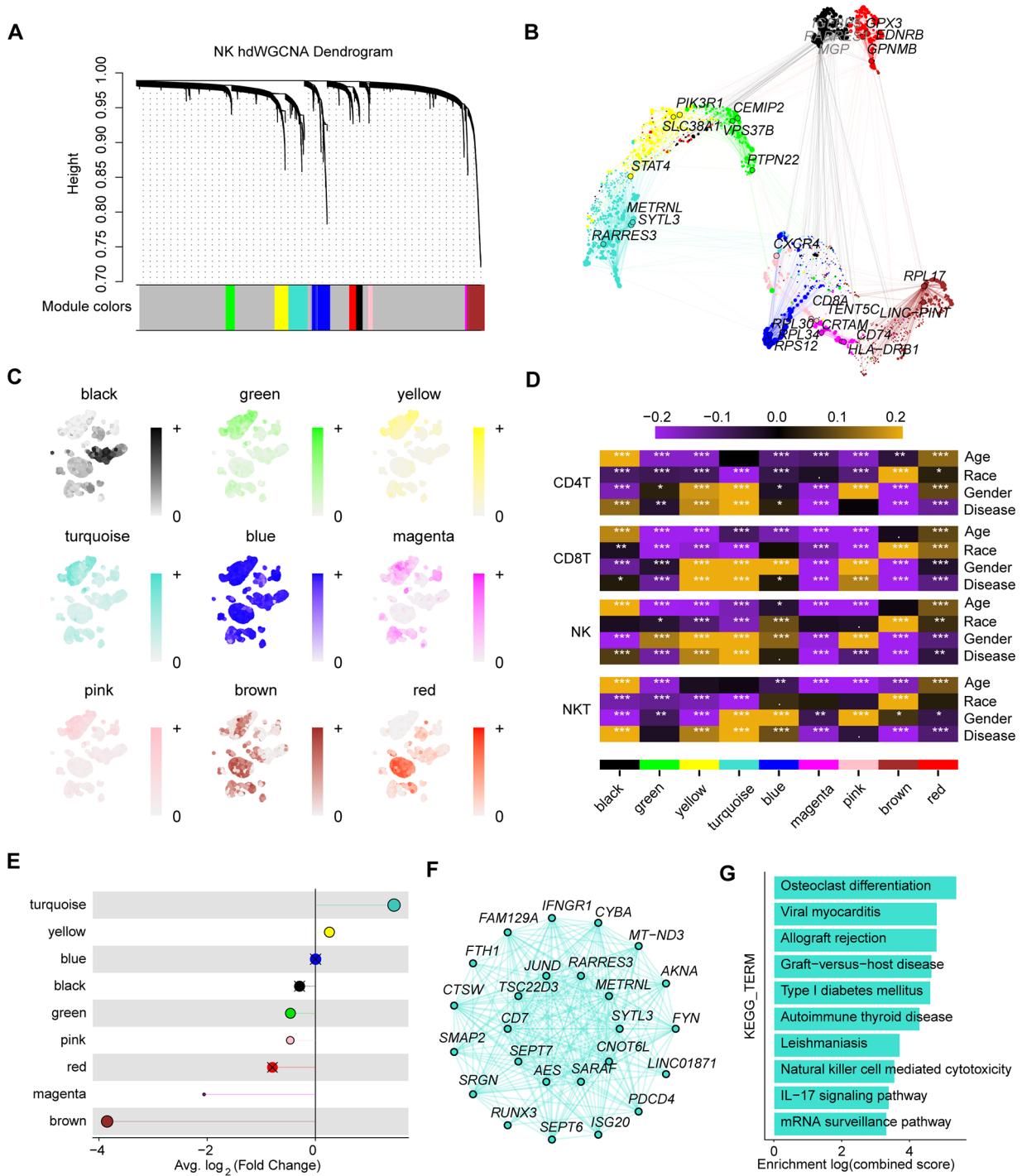
Figure 4. High dimensional weighted gene co-expression network analysis (hdWGCNA) revealed that the NK cell cluster is characterized by the turquoise module: (A) dendrogram was utilized to visualize all modules in the scale-free network, (B) UMAP visualization of top five genes from each module in co-expression networks, (C) UMAP plot showing the expression distribution of hub genes for each module across the NK cells, (D) heatmap plot representing correlations between different modules and different stage of age, race, gender or disease in different cell type, (E) lollipop plot showing the differential module eigengene analysis result of different modules, an "X" is placed over each point that does not reach statistical significance, (F) ModuleNetworkPlot showing the top 25 hub genes for the turquoise module, and (G) bar plot representing the KEGG enrichment results of genes in the turquoise module.

original UMAP distribution. Notably, the green, yellow, and turquoise modules were enriched in the T/NK cluster, with the turquoise module showing the most specificity for NK cells (Figure 4C). To assess the relationship with AMD, module trait correlation was performed for each cell cluster, revealing a

significant positive correlation between the turquoise module and AMD in NK cells (Figure 4D). Differential module eigengene analysis corroborated this finding, indicating a significant positive association between the turquoise module and AMD (Figure 4E). In light of these findings, we conducted a more
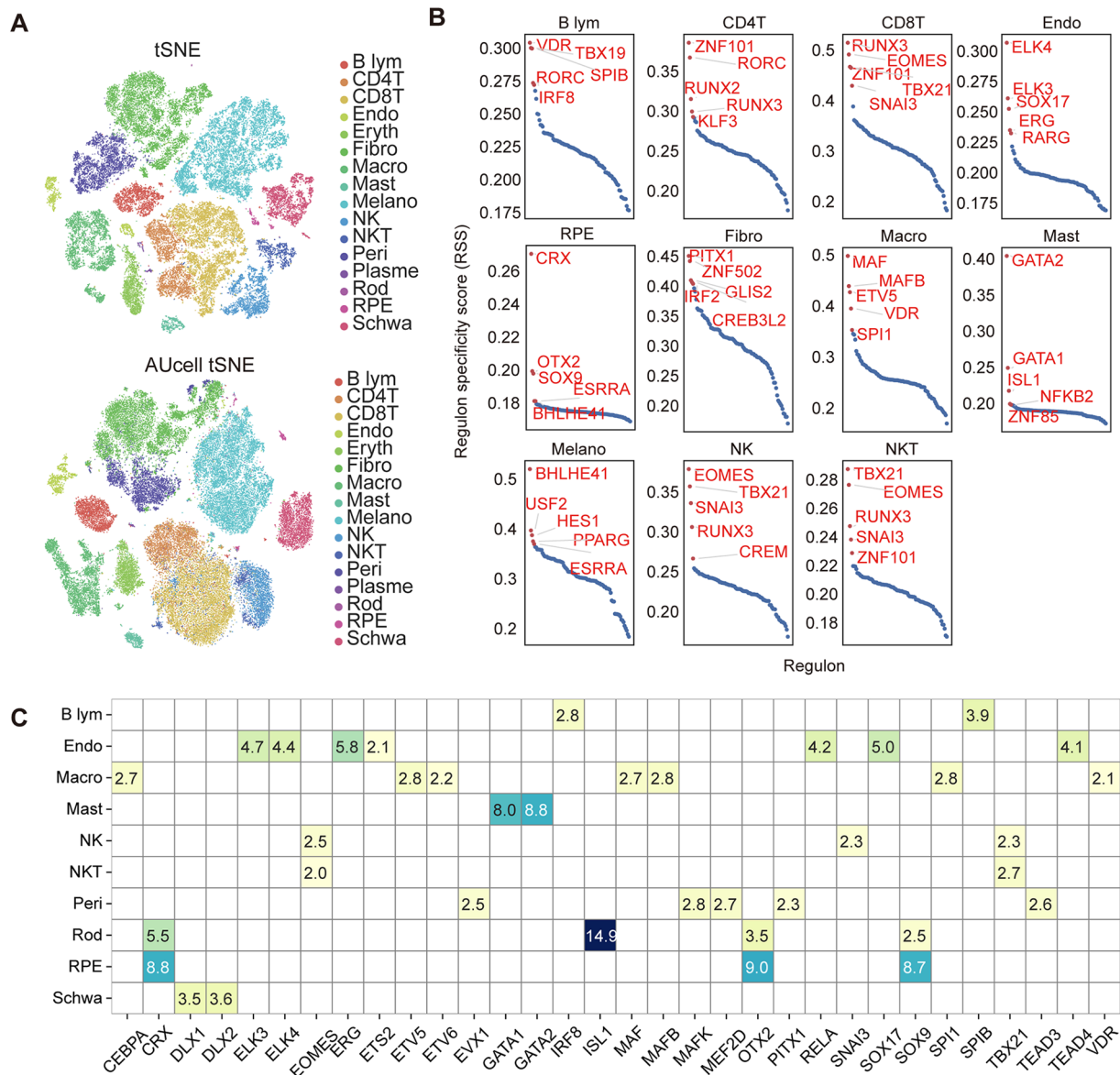
Figure 5. Single-cell regulatory network inference and clustering (SCENIC) analysis revealed the key transcription factor in NK cells: (A) t-Distributed Stochastic Neighbor Embedding (tSNE) plot showing the different cells based on raw data (top) and regulon activity (AUcell), (B) dot plot showing regulon specificity scores (RSS) in each cell type, and (C) heatmap plot showing cell type-specific regulators.

in-depth exploration of the co-expression network within the turquoise module. The display of the top 25 hub genes and pathway enrichment highlighted a significant focus on inflammation-related pathways, as indicated by KEGG pathway enrichment (Figure 4F and G).

SCENIC analysis revealed the key transcription factor in NK cells

We next utilized SCENIC to unravel the gene regulatory network (GRN) across all clusters, followed by hierarchical clustering based on regulon activity. The results revealed a striking similarity in the GRN between NK and NKT cells (Figure 5A). Additionally, we employed the RSS to assess regulons and ranked them within each cluster (Figure 5B). NK cells were thus selected for further study. Notably, TBX21,

EOMES, and SNAI3 were identified as distinctive transcription factors in the GRN of NK cells (Figure 5C). The regulatory networks of TBX21, EOMES, and SNAI3 were further validated using iRegulon (sFigure 3A-C). Consistent with the above regulon results, TBX21, EOMES, and SNAI3 were identified as the master regulators in AMD.

Identification and validation of a predictive model for AMD based on hub TF of NK cell

We conducted further investigations to investigate the relationships between NK cells and AMD, aiming to gain insights into the underlying mechanisms driving AMD. Our focus was on predicting AMD, and we constructed a predictive model using the hub TFs of NK cells by integrating the hub genes of the turquoise module from hdWGCNA and the key transcription
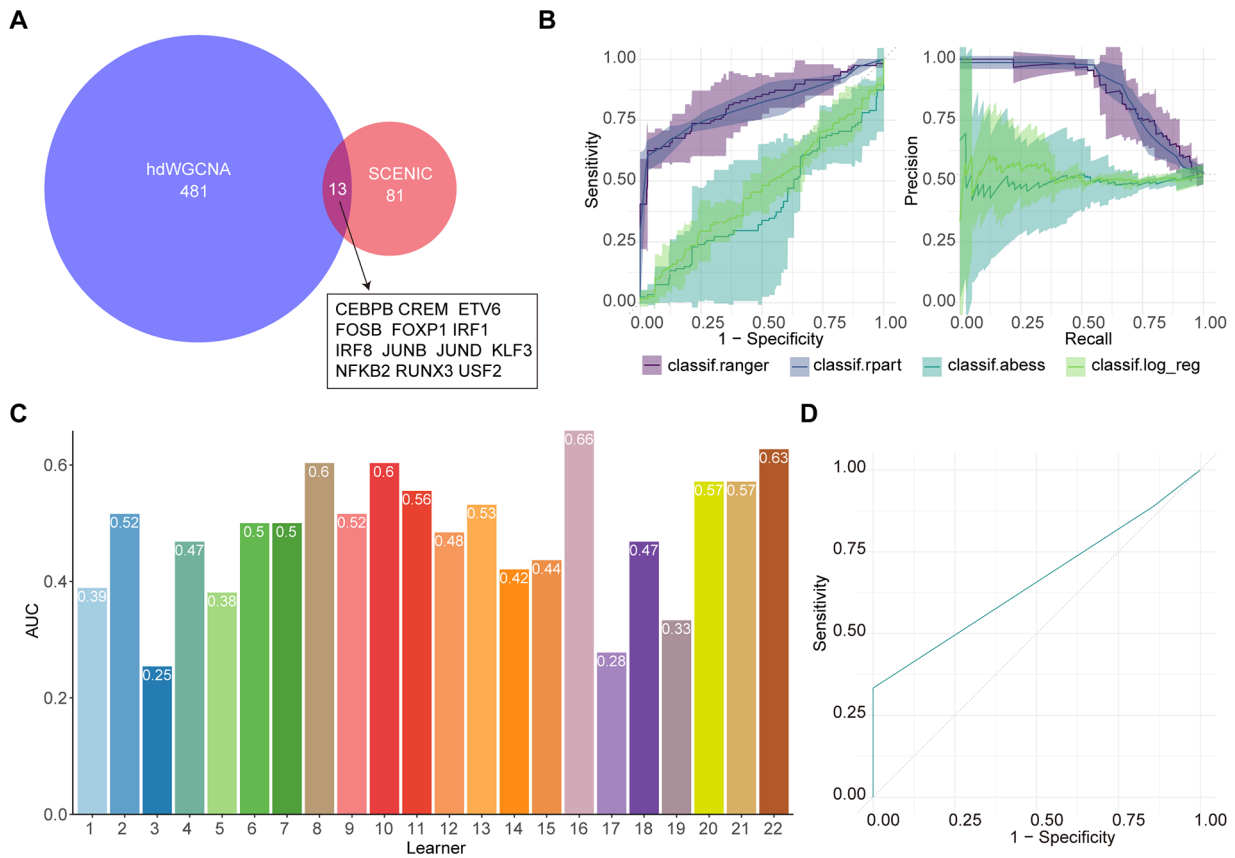
Figure 6. Identification and validation of a predictive model for AMD based on hub genes of NK cells: (A) Venn plot displaying the thirteen overlapping key regulators identified by merging the genes in hdWGCNA turquoise module and genes in SCENIC in NK cell, (B) receiver operating characteristic curve (ROC) and precision-recall curve of four feature selection in the training data set (GSE135092 and GSE29801), (C) area under curve (AUC) of total 22 random forest learners performed in the testing set (GSE136661), and (D) ROC of learners 16 in the testing set.

Table 1. The AUC and ACC of each learner.

LEARNER_ID	RESAMPLING_ID	ITERS	AUC	ACC
Random forest	CV	22	0.857	0.784
Decision trees	CV	22	0.725	0.798
Fast best-subset selection	CV	22	0.460	0.491
Logistic regression	CV	22	0.328	0.370

Abbreviations: ACC, accuracy; AUC, area under curve; CV, cross validation.

factors in NK cells from SCENIC. A Venn plot was generated to illustrate the overlap between the key genes from hdW-GCNA and SCENIC, leading to the selection of 13 genes (*CEBPB, CREM, ETV6, FOSB, FOXP1, IRF1, IRF8, JUNB, JUND, KLF3, NFKB2, RUNX3, USF2*) (Figure 6A). Subsequently, we utilized the function of nested resampling in the mlr3 package to select a subset of TFs for unbiased performance estimates. The training set was obtained by combining data from GSE135092 and GSE29801 after correcting for batch effects (sFigure 4A), with 220 samples enrolled in the training set. The distribution of the 13 hub genes in the training set also indicated a certain degree of correlation among some hub genes (sFigure 4B). Therefore, we employed three machine learning algorithms—random forest (RF), decision trees (DT), and ABESS—for feature selection to reduce bias. Twenty-two-fold cross-validation was applied for resampling, and ROC and precision-recall curves were plotted (Figure 6B). The area under the curve (AUC) and accuracy (ACC) of all resampling results for each learner was calculated, revealing that the RF learner exhibited the best performance among all learners (Table 1). Further validation of the 22 RF learners in the testing set from GSE136661 indicated that learner 16 yielded the best performance (AUC = 0.66), identifying *CREM, FOXP1, IRF1, NFKB2,* and *USF2* as hub genes (Figure 6C and D). In summary, our findings suggest that the expression of the identified 5 hub TF genes may serve as a potential predictor for AMD.

Discussion

In the present study, single-cell analysis revealed different cell types in the choroid of macula from AMD and Normal samples, especially the T/NK cluster. Subsequently, we analyzed the five subclusters of T/NK cells, identifying CD4T, effector CD8 T cells, naïve/memory CD8T, NK, and NKT cell populations. Notably, there was a significant upregulation in the proportion and number of NK cells in AMD. To gain further insights, we employed CellChat to analyze cell-cell communication pathways between AMD and Normal. Our findings indicated a significant alteration in the interaction strength of NK cells. Building upon these results, we established a link between NK cells and the occurrence and development of AMD. Moreover, we conducted hdWGCNA on NK cells in AMD, revealing that the turquoise module was significantly associated with AMD in NK cells. The hub genes within this module were enriched in pathways related to immune regulation and inflammatory response. Additionally, SCENIC analysis on NK cells in AMD unveiled master transcription factors crucial in regulating NK cell function, subsequently playing a key role in AMD. Combining the hub genes from the turquoise module and the transcription factors from the regulatory network of NK cells, we identified 13 candidate transcription factors with a crucial regulatory effect on the occurrence of AMD. Subsequently, employing 4 machine learning techniques, we filtered these transcription factors to identify those with the potential to predict the prevalence of AMD. Following validation in an independent dataset, we ultimately identified 5 key transcription factors. These findings offer novel insights into the genesis process and biological importance of NK cells in AMD.

In AMD, NK cells exhibited a notable increase in both proportion and number. Concurrently, we observed a significant decrease in HLA class II molecules, including HLA-DR and HLA-DQ, in NK cells. HLA-DR-positive NK cells demonstrated functional activity and the ability to present antigens to T cells, indicating a protective effect in infectious diseases or chronic inflammation.[28] Interestingly, some interleukin (IL) receptors related genes, such as ILDR2, IL18RAP, and IL12RB2, showed a significant increase in NK cells in AMD. IL-18, known to enhance NK activity by binding to IL-18R,[29] and the increased expression of IL-12R on NK cells were associated with elevated levels of NK cell-derived cytokines.[30] Pathway enrichment analysis further supported these findings. In summary, this study highlights the crucial role of NK cells in AMD, showcasing highly functional activity implicated in the pathologic process through excessive cytokine secretion and incorrectly directed cytotoxicity.

In the present study, based on machine learning approaches, we identified *CREM, FOXP1, IRF1, NFKB2,* and *USF2* as key genes. Previous research has explored the roles of *FOXP1, IRF1,* and *NFKB2* in the development and progression of AMD.[31-33] *FOXP1* exhibited expression in the retinal pigment epithelium (RPE), and its knockout in RPE led to reduced vascular endothelial growth factor (VEGF), contributing to choroidal neovascularization in AMD.[31] *IRF1* has been implicated in the etiology of AMD, with IL-27 increasing the expression of complement factor H, a key player in AMD development, through the activation of the STAT1-IRF-1 pathway in RPE.[32] Furthermore, the H haplogroup of Mitochondrial DNA (mtDNA) was protective against AMD, and NFKB2 showed differential expression between the H and J haplogroups.[33] Collectively, FOXP1, IRF1, and NFKB2 emerged as crucial players in AMD development. Moreover, our study has uncovered two hitherto undocumented TFs associated with AMD. CREM emerged as the top differentially expressed TF in melphalan-induced vascular toxicity.[34] Notably, CREM-1 demonstrated protective effects against retinal cell apoptosis induced by light exposure.[35] Considering that vascular injury and cell apoptosis are pivotal pathological stages in AMD, the implication is that CREM may exert influence in the context of AMD.[36] Additionally, USF2 exhibited high expression during embryonic eye development, suggesting a potential role in eye development.[37] Overall, the identification of these predictive genes might represent a breakthrough in the pathogenesis and development of AMD.

This study acknowledges certain limitations that warrant attention in future research. Firstly, the limited sample size in the machine learning analysis may impact the accuracy and reliability of the results. Expanding the sample size could enhance the robustness of the findings. Secondly, the selected TFs of NK cells in AMD should be subject to further experimental exploration to elucidate their molecular mechanisms and assess their therapeutic potential.

Conclusion

In conclusion, our study provides valuable insights into the intricate relationship between NK cells and AMD. The substantial increase in NK cell proportions among AMD patients, along with altered cell interactions and communication strength, underscores the pivotal role of NK cells in the pathogenesis of AMD. Utilizing hdWGCNA and SCENIC, we unveiled hub module genes and identified key TFs associated with NK cells in AMD. The predictive model, integrating hub genes and key TFs, revealed 5 potential TFs that could serve as predictors for the development of AMD. This comprehensive approach enhances our understanding of the regulatory mechanisms within NK cells during AMD. The identified TFs, particularly CREM, FOXP1, IRF1, NFKB2, and USF2, open new avenues for exploring molecular interventions in the context of AMD.

Author Contributions

Ethics Approval

ORCID iD

Haihua Zheng 🆔 https://orcid.org/0009-0009-2059-4305

Data Availability Statement

The dataset used in this study is available in the GEO database (https://www.ncbi.nlm.nih.gov/gds).

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. Wong WL, Su X, Li X, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health*. 2014;2:e106-e116.
2. Handa JT, Bowes Rickman C, Dick AD, et al. A systems biology approach towards understanding and treating non-neovascular age-related macular degeneration. *Nat Commun*. 2019;10:3347.
3. Lotery A, Trump D. Progress in defining the molecular biology of age related macular degeneration. *Hum Genet*. 2007;122:219-236.
4. Chen Y, Bedell M, Zhang K. Age-related macular degeneration: genetic and environmental factors of disease. *Mol Interv*. 2010;10:271-281.
5. Donoso LA, Kim D, Frost A, Callahan A, Hageman G. The role of inflammation in the pathogenesis of age-related macular degeneration. *Surv Ophthalmol*. 2006;51:137-152.
6. Ng PQ, Saint-Geniez M, Kim LA, Shu DY. Divergent metabolomic signatures of TGFβ2 and TNFα in the induction of retinal epithelial-mesenchymal transition. *Metabolites*. 2023;13:213. doi:10.3390/metabo13020213
7. Smith W, Assink J, Klein R, et al. Risk factors for age-related macular degeneration: pooled findings from three continents. *Ophthalmology*. 2001;108:697-704.
8. Seddon JM, Ajani UA, Mitchell BD. Familial aggregation of age-related maculopathy. *Am J Ophthalmol*. 1997;123:199-206.
9. Allingham MJ, Loksztejn A, Cousins SW, Mettu PS. Immunological aspects of age-related macular degeneration. *Adv Exp Med Biol*. 2021;1256:143-189.
10. Zhang Y, Wong WT. Innate immunity in age-related macular degeneration. *Adv Exp Med Biol*. 2021;1256:121-141.
11. Scholl HP, Charbel Issa P, Walier M, et al. Systemic complement activation in age-related macular degeneration. *PLoS One*. 2008;3:e2593.
12. Combadière C, Feumi C, Raoul W, et al. CX3CR1-dependent subretinal microglia cell accumulation is associated with cardinal features of age-related macular degeneration. *J Clin Invest*ig. 2007;117:2920-2928.
13. Montaldo E, Del Zotto G, Della Chiesa M, et al. Human NK cell receptors/markers: a tool to analyze NK cell development, subsets and function. *Cytometry A*. 2013;83:702-713.
14. Goverdhan SV, Khakoo SI, Gaston H, Chen X, Lotery AJ. Age-related macular degeneration is associated with the HLA-Cw*0701 genotype and the natural killer cell receptor A haplotype. *Investig Ophthalmol Vis Sci*. 2008;49:5077-5082.
15. Zeng Y, Yin X, Chen C, Xing Y. Identification of diagnostic biomarkers and their correlation with immune infiltration in age-related macular degeneration. *Diagnostics*. 2021;11:1079. doi:10.3390/diagnostics11061079
16. Rosmaraki EE, Douagi I, Roth C, et al. Identification of committed NK cell progenitors in adult murine bone marrow. *Eur J Immunol*. 2001;31:1900-1909.
17. DeHart SL, Heikens MJ, Tsai S. Jagged2 promotes the development of natural killer cells and the establishment of functional natural killer cell lines. *Blood*. 2005;105:3521-3527.
18. Jaleco AC, Neves H, Hooijberg E, et al. Differential effects of Notch ligands Delta-1 and jagged-1 in human lymphoid differentiation. *J Exp Med*. 2001;194:991-1002.
19. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184:3573-3587.e29.
20. Jin S, Plikus M, Nie Q. CellChat for systematic analysis of cell-cell communication from single-cell and spatially resolved transcriptomics. *BioRxiv*. 2023. doi:10.1101/2023.11.05.565674
21. Murphy AE, Skene NG. A balanced measure shows superior performance of pseudobulk methods in single-cell RNA-sequencing analysis. *Nat Commun*. 2022;13:7851.
22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
23. Qiu X, Hill A, Packer J, et al. Single-cell mRNA quantification and differential analysis with census. *Nat Methods*. 2017;14:309-315.
24. Morabito S, Reese F, Rahimzadeh N, Miyoshi E, Swarup V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. *Cell Rep Methods*. 2023;3:100498.
25. Aibar S, González-Blas CB, Moerman T, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. 2017;14:1083-1086.
26. Lang M, Binder M, Richter J, et al. Mlr3: a modern object-oriented machine learning framework in R. *J Open Source Softw*. 2019;4:1903.
27. Armento A, Ueffing M, Clark SJ. The complement system in age-related macular degeneration. *Cell Mol Life Sci*. 2021;78:4487-4505.
28. Erokhina SA, Streltsova MA, Kanevskiy LM, et al. HLA-DR-expressing NK cells: effective killers suspected for antigen presentation. *J Leukoc Biol*. 2021;109:327-337.
29. Hyodo Y, Matsui K, Hayashi N, et al. IL-18 up-regulates perforin-mediated NK activity without increasing perforin messenger RNA expression by binding to constitutively expressed IL-18 receptor. *J Immunol*. 1999;162:1662-1668.
30. Parihar R, Dierksheide J, Hu Y, Carson WE. IL-12 enhances the natural killer cell cytokine response to ab-coated tumor cells. *J Clin Invest*ig. 2002;110:983-992.
31. Yan M, Li J, Yan L, Li X, Chen JG. Transcription factor foxp1 is essential for the induction of choroidal neovascularization. *Eye Vis*. 2022;9:10.
32. Amadi-Obi A, Yu CR, Dambuza I, et al. Interleukin 27 induces the expression of complement factor H (CFH) in the retina. *PLoS One*. 2012;7:e45801.
33. Atilano SR, Malik D, Chwa M, et al. Mitochondrial DNA variants can mediate methylation status of inflammation, angiogenesis and signaling genes. *Hum Mol Genet*. 2015;24:4491-4503.
34. Yang Y, Xing Y, Liang C, et al. In search of underlying mechanisms and potential drugs of melphalan-induced vascular toxicity through retinal endothelial cells using bioinformatics approach. *Tumour Biol*. 2016;37:6709-6718.
35. Xu Y, Yu S, Shu Q, et al. Upregulation of CREM-1 relates to retinal ganglion cells apoptosis after light-induced damage in vivo. *J Mol Neurosci*. 2014;52:331-338.
36. van Lookeren Campagne M, LeCouter J, Yaspan BL, Ye W. Mechanisms of age-related macular degeneration and therapeutic opportunities. *J Pathol*. 2014;232:151-164.
37. Fujimi TJ, Aruga J. Upstream stimulatory factors, USF1 and USF2 are differentially expressed during Xenopus embryonic development. *Gene Expr Patterns*. 2008;8:376-381.