

RESEARCH

Open Access



Retinoblastoma gene expression profiling based on bioinformatics analysis

Jun Mao^{1†}, Mingzhi Lu^{2†}, Siduo Lu^{3†}, Yiqiao Xing², Xuejiao Xu¹, Ying Chen¹, Huirong Xu¹, Wei Zuo¹, Jingwen Zhou¹ and Wei Du^{1*}

Abstract

Background Retinoblastoma (RB) is frequently occurring malignant tumors that originate in the retina, and their exact cause and development mechanisms are yet to be fully comprehended. In this study, we identified possible biomarkers for RB and delved into the molecular mechanics linked with such markers.

Methods In this study GSE110811 and GSE24673 were analyzed. Weighted gene co-expression network analysis (WGCNA) was applied to screen modules and genes associated with RB. By overlapping RB-related module genes with differentially expressed genes (DEGs) between RB and control samples, differentially expressed retinoblastoma genes (DERBGs) were acquired. A gene ontology (GO) enrichment analysis and a kyoto encyclopedia of genes and genomes (KEGG) enrichment analysis were conducted to explore the functions of these DERBGs. To study the protein interactions of DERBGs, a protein–protein interaction (PPI) network was constructed. Hub DERBGs were screened using the least absolute shrinkage and selection operator (LASSO) regression analysis, as well as the random forest (RF) algorithm. Additionally, the diagnostic performance of RF and LASSO methods was evaluated using receiver operating characteristic (ROC) curves and single-gene gene set enrichment analysis (GSEA) was conducted to explore the potential molecular mechanisms involved with these Hub DERBGs. In addition, the competing endogenous RNA (ceRNA) regulatory network of Hub DERBGs was constructed.

Result About 133 DERBGs were found to be associated with RB. GO and KEGG enrichment analyses revealed that the important pathways of these DERBGs. Furthermore, the PPI network revealed 82 DERBGs interacting with each other. By RF and LASSO methods, *PDE8B*, *ESRRB*, and *SPRY2* were identified as Hub DERBGs in patients with RB. From the expression assessment of Hub DERBGs, it was found that the levels of expression of *PDE8B*, *ESRRB*, and *SPRY2* were significantly decreased in the tissues of RB tumors. Secondly, single-gene GSEA revealed a connection between these 3 Hub DERBGs and oocyte meiosis, cell cycle, and spliceosome. Finally, the ceRNA regulatory network revealed that hsa-miR-342-3p, hsa-miR-146b-5p, hsa-miR-665, and hsa-miR-188-5p may play a central role in the disease.

Conclusion Hub DERBGs may provide new insight into RB diagnosis and treatment based on the understanding of disease pathogenesis.

Keywords Retinoblastoma, Bioinformatics analysis, Biomarkers, ceRNA, Regulatory network

[†]Jun Mao, Mingzhi Lu and Siduo Lu are co-first authors

*Correspondence:

Wei Du

foxfeinia@163.com

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



Background

Retinoblastoma (RB), An ophthalmological common intraocular cancer, poses a serious threat to the vision and well-being of patients [1, 2]

Being a genetic disease, RB is caused by the deletion of the tumor suppressor gene, *BRI* [3]. With the advancement in gene detection technology, abnormal expression of genes such as *MYCN* has become an important factor in the progression of RB [4]. Due to the advancements in methods and innovation, RB treatment has undergone a noteworthy transformation from excision, radiotherapy, and intravenous chemotherapy to intra-arterial chemotherapy combined with local therapy [5–7].

While the pathogenesis of RB is unclear, the rate of eyeball removal in patients with advanced RB remains high, indicating the need for further treatment innovations. It is imperative to study the biological process (BP) and related potential mechanisms of RB to develop a new treatment strategy. Researchers have discussed the most important prognostic factors and potential mechanisms of RB through the use of existing data on RB and bioinformatics methods [8, 9]. For example, an analysis conducted by Wen et al. [10] identified two critical microRNA targets in RB: let-7a and let-7b by analyzing a variety of bioinformatics studies and identifying microRNA-target gene-transcription factor regulatory networks in RB. According to Gao et al. [11] the long noncoding RNA (lncRNA) MEG3 may play a role in tumor suppression in RB, and the activation of lnc00152 by Sp1 induces EMT through the miR-30d/SOX9/ZEB2 pathway and enhances the invasion and metastasis of RB cells through this pathway. The pathogenesis of cancer is extremely complex, but more research needs to be conducted on this topic.

A better comprehension of the genetic, environmental, as well as immune-regulatory factors of RB may provide important insights into its diagnosis and pathogenesis. Bioinformatics assessments have been used to diagnose many diseases, but enough assessments in RB have not been performed. This study aims to identify the biomarkers of RB prognosis using multiple bioinformatics-based datasets and explain its pathogenesis using functional enrichment analysis and the competing endogenous RNA (ceRNA) network. These findings may contribute to an additional understanding of the pathogenesis of RB and guide future research on this disease.

Materials and methods

Source of data

A total of two RB datasets, GSE110811 and GSE24673, were retrieved out of the Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>). GSE110811

(19 retinal tissue samples and 31 RB samples) and GSE24673 (2 cadaveric eye samples and 9 RB samples) were subsequently merged into a new dataset that contained 40 RB samples and 21 normal samples.

Weighted gene co-expression network analysis (WGCNA)

The gene-expression profiles from the GSE110811 and GSE24673 datasets were used to investigate the RB-associated module via the “WGCNA” R package (version 1.70-3) [12]. Initially, an adjacency matrix was constructed based on the formula for the adjacency matrix to describe the correlation strength between the nodes [13].

$$s_{ij} = |cor(x_i, x_j)| a_{ij} = s_{ij} \beta$$

where, i and j represent two different genes, and x_i and x_j represent expression scores of the corresponding genes. S_{ij} represents the correlation coefficient, and a_{ij} represents the strength of the correlation between i and j . In this research, the adjacency matrix was constructed with a scale-free topological index of 0.85 and an optimal soft-threshold power (β) of 12. The adjacency matrix was then converted into a matrix of topological overlap. Finally, hierarchical clustering trees were constructed by dynamically cutting trees (module sizes of 25) for identifying key modules through the introduction of genes with similar expression patterns into the same module.

Differential genes expression analysis

As a first step, differentially expressed genes (DEGs) between RB samples and normal samples in the merged dataset were identified using the “limma” R package (version 3.46.0), with an adjusted p value of <0.05 and a $|\log_2FC| > 1$ [14]. To obtain differentially expressed retinoblastoma genes (DERBGs), the DEGs were intersected with key module genes using the VennDiagram R package (version 1.6.20) [15]. To evaluate the potential functions of DERBGs, a Gene Ontology (GO) function and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses were conducted using the clusterProfiler R package [16–20].

Protein–protein interaction (PPI) network construction

To investigate if there are protein interactions between DERBGs, the Search tool to retrieve Interacting Genes and Proteins (STRING) website (<https://string-db.org>) was used to map a PPI network of these DERBGs. Further, the PPI network was visualized using Cytoscape, and the top ten DERBGs were identified using the maximal clique centrality (MCC) algorithm of Cytohubba [21].

Screening and validation of Hub DERBGs

Hub DERBGs were screened using the random forest (RF) method with the “Randomforest” R package (version 4.7-1) and the least absolute shrinkage and selection operator (LASSO) regression assessment with the “glmnet” R package (version 4.1-1) [21, 22]. Besides, the “pROC” R package (V 1.17.0.1) was used to evaluate the diagnostic performance of the RF and LASSO methods [23]. Following this, the Hub DERBG expression values were validated in the merged dataset ($p < 0.05$).

Single-gene gene set enrichment analysis (GSEA)

To explore the regulatory pathways and biological functions associated with these Hub DERBGs, the “clusterProfiler” R package (version 3.18.0) was used to perform the GSEA of each DERBG [16, 17]. An adjusted p value of < 0.05 was used to indicate significant thresholds for GSEA.

Construction of a ceRNA regulatory network

Differentially expressed microRNAs (DEmiRNAs) were identified in the GSE41321 dataset ($p < 0.05$ and $|\log_{2}FC| > 1$) using the “limma” R package (version 3.46.0) [14]. Meanwhile, the miRWalk DB (<http://mirwalk.umm.uni-heidelberg.de/>) was utilized for predicting miRNAs from the Hub DERBGs (binding $p > 0.95$ and energy < -15). To obtain the miRNA–messenger RNA (mRNA) relationship pair and common miRNAs, the predicted miRNAs were intersected with DEmiRNAs. Further, the “DEseq” R package (V1.34.0) was used to identify differentially expressed long noncoding RNAs (DElncRNAs) in the GSE125903 dataset ($p < 0.05$ and $|\log_{2}FC| > 1$) [24]. Also, based on the common miRNAs, the Starbase database (<http://starbase.sysu.edu.cn/>) was employed to predict lncRNAs (binding $p > 0.95$ and energy < -15) [25]. Finally, a ceRNA regulatory network was constructed by integrating the interactions between DEmiRNAs and DElncRNAs. To visualize the ceRNA network, the Cytoscape R package (V3.8.2) was used [21].

Statistical analysis

R software (<https://www.r-project.org/>, V4.0.3) was used to perform statistical analyses and to visually plot the results. Using Pearson correlation analysis, correlation coefficients and p values were calculated for the RB-related module and patients with RB. A p value of < 0.05 was considered statistically significant.

Results

Identification of the RB-associated modules and genes through WGCNA

A co-expression network was constructed by applying WGCNA to all samples and genes in the merged

dataset to identify RB-related modules and genes. Additional file 1 provides a comprehensive overview of the Principal Component Analysis (PCA) carried out on the gene expression datasets. The sample dendrogram in Fig. 1A indicates that the merged dataset does not contain any outliers. Following this, a scaleless network was constructed with a 12 soft-threshold power (β), and an index of scale-free topologies of 0.85 was set (Fig. 1B). Using a dynamic tree-cutting method to introduce genes with similar expression patterns into a single module (module size = 25), a hierarchical clustering tree with 8 modules was constructed (Figs. 1C and D). As shown in Fig. 1E and F, MEgreen has the highest correlation with RB ($Cor = 0.54$, $p = 1.6e-26$). Therefore, a total of 332 genes from the MEgreen module were analyzed in the following steps.

Identification of DERBGs in RB

The first step in identifying DERBGs associated with RB was to screen DEGs between RB samples and normal samples in the merged dataset. As illustrated in Fig. 2A and B, a total of 384 DEGs were identified in RB samples, of which 188 were downregulated and 196 were upregulated. Following this, 133 DERBGs were obtained for further analysis by intersecting DEGs with genes related to RB modules (Fig. 2C).

PPI network of DERBGs and functional analysis

A network map of the PPI network was developed using the STRING website to study the interactions among 133 DERBGs. It was possible to obtain a PPI network with 118 interactions and 82 nodes (Fig. 3A). To rank these DERBGs with interactions, the MCC algorithm of the Cytoscape software was used (Fig. 3B; *RDH8*, *RGR*, *CNGA1*, *ROM1*, *SAG*, *RHO*, *PAX6*, *RLBP1*, *CNGB1*, and *RDH12*). To explore the role of 82 DERBGs in BPs, GO and KEGG were constructed. The clustering coefficient was 0.267. As a result of GO enrichment analysis, these DERBGs were primarily involved in five terms. For example, in BPs, the light was perceived as a sensory stimulus, visual stimulus, and detectable stimulus; in addition, these DERBGs were primarily involved in five terms, such as photoreceptor outer segment, photoreceptor cell cilium, and 9+0 non-motile cilium in the cellular component (CC); in molecular function (MF), these DERBGs were mainly engaged in G-protein-coupled photoreceptor activity, photoreceptor activity, the activity of cell–cell adhesion mediators, and the activity of cell adhesion mediators (Fig. 3C and D). According to the KEGG pathways, these DERBGs are significantly associated with phototransduction, axon guidance, pathways of signal transmission

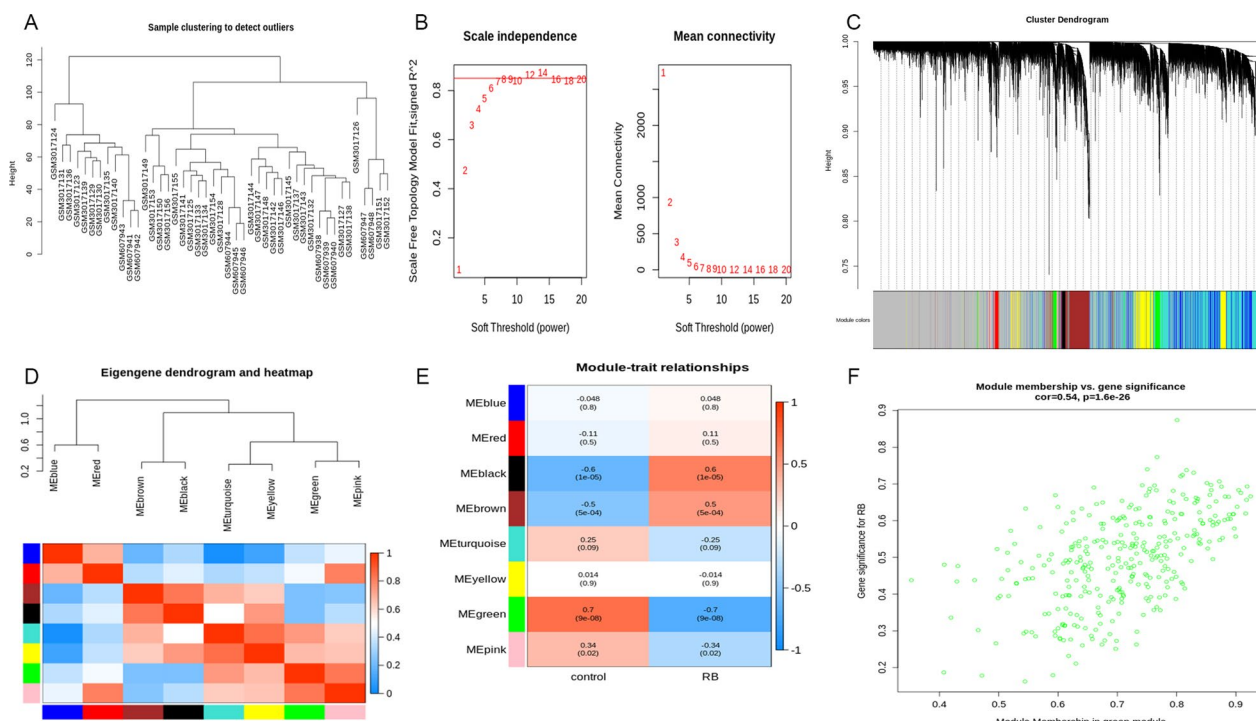


Fig. 1 Weighted gene co-expression network analysis results. **A** Sample dendrogram and trait heatmap; **B** Scale independence and mean connectivity; **C** Cluster dendrogram; **D** Eigengene dendrogram and eigengene heatmap; **E** Module trait relationships; **F** Scatter plot of genes in the green module

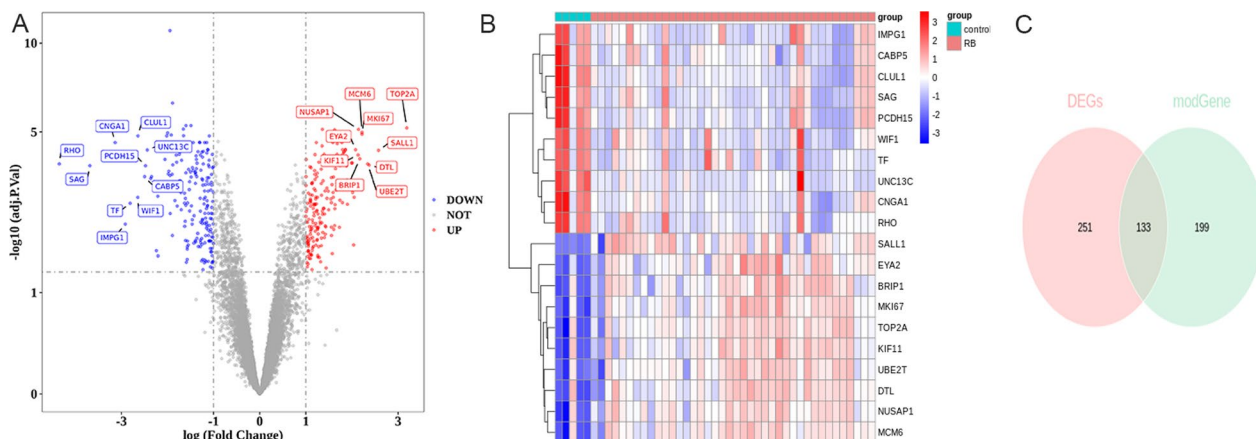


Fig. 2 Genome-wide analysis of gene expression of retinoblastoma (RB). **A** A volcano plot of differentially expressed genes (DEGs); **B** Heatmap of DEGs; **C** Venn diagram of DEGs detected among RB DEGs and RB module genes

regulating pluripotency, ferroptosis, and cocaine addiction pathways (Fig. 3E and F).

Screening and expression level validation of Hub DERBGs

The RF algorithm was used to identify the top ten DERBGs (*PDE8B*, *FBXO32*, *ESRRB*, *RDH8*, *TAOK3*, *SPRY2*, *MCUR1*, *CASZ1*, *CABP4*, and *SIK2*) from the merged

dataset for further validation and selection of Hub DERBGs with significantly characteristic value for classifying RB and normal samples (Fig. 4A and B). Additionally, three DERBGs were selected using the LASSO algorithm ($\lambda_{min}=0.06493903$) (Figs. 4C and D). By integrating the DERBGs screened by the RF and LASSO algorithms, a total of 7 DERBGs were identified,

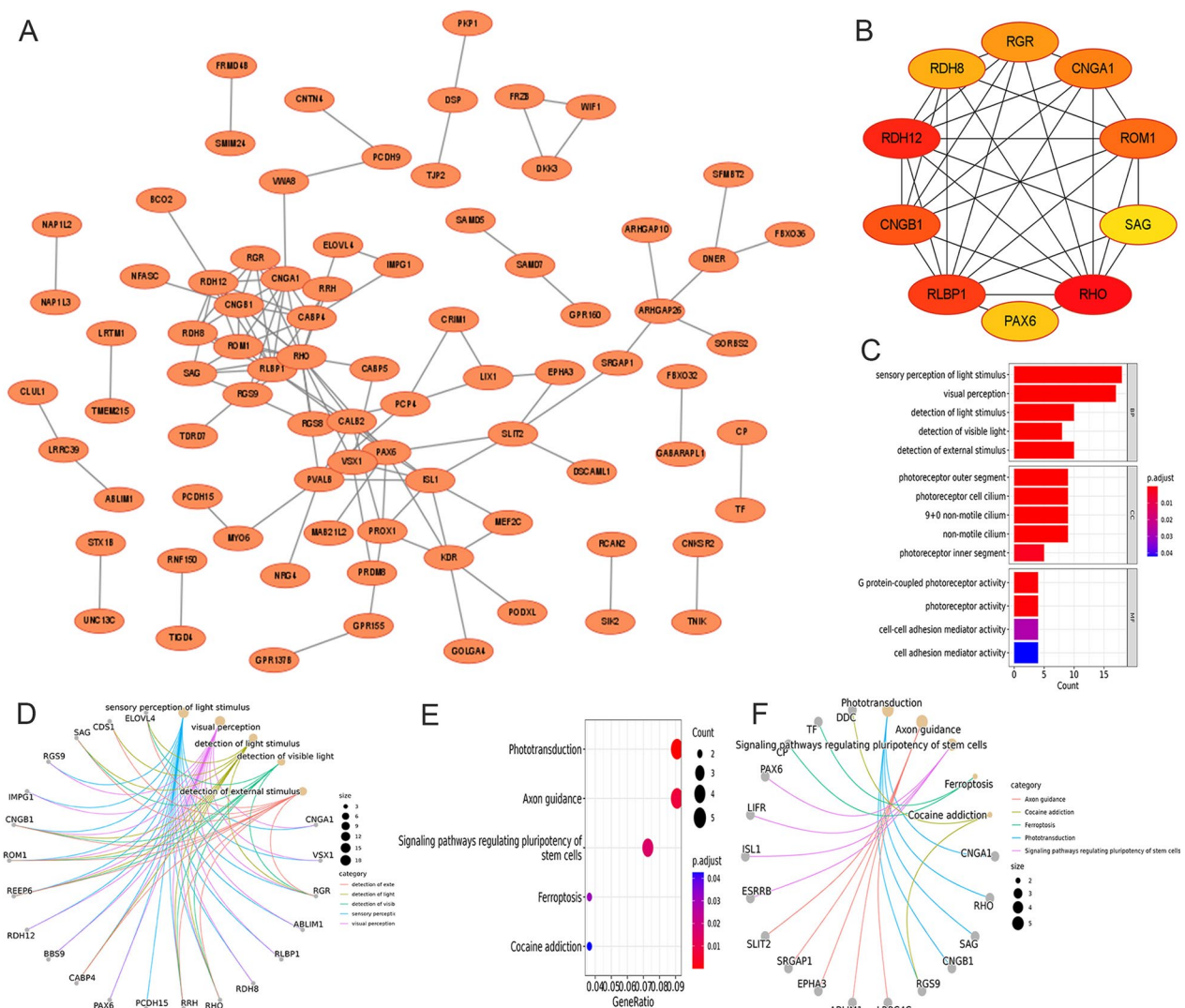


Fig. 3 A protein–protein interaction (PPI) network of differentially expressed retinoblastoma genes (DERBGs) and functional analysis. **A** PPI network; **B** The top 10 genes; **C** Gene Ontology (GO) enrichment of DERBGs; **D** Chord plot of GO enrichment; **E** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment of DERBGs; **F** Chord plot of the KEGG pathway enrichment

of which 3 (*PDE8B*, *ESRRB*, and *SPRY2*) were selected simultaneously by both algorithms (Fig. 4E). For classification and diagnostic purposes, those DERBGs were identified as Hun DERBGs. Further, the receiver operating characteristic (ROC) curve analysis and the confusion matrix heat map together demonstrate that RF and

LASSO algorithms can provide good diagnostic performance (Figs. 4F–I). The merged dataset was used to validate the levels of expression of three Hub DERBGs (*PDE8B*, *ESRRB*, and *SPRY2*). In RB tumor tissue, the levels of mRNA of *PDE8B*, *ESRRB*, and *SPRY2* were greatly reduced (Fig. 4J).

(See figure on next page.)

Fig. 4 Screening and expression level validation of Hub differentially expressed retinoblastoma genes (DERBGs). **A** Importance ranking of the top 10 DERBGs; **B** Random forest algorithm; **C** Lasso-logistic algorithm; **D** Lasso result graph; **E** Venn diagram of differentially expressed genes (DEGs) detected using the random forest algorithm and Lasso-logistic algorithm; **F** Confusion matrix heat map of the Lasso-logistic algorithm; **G** Receiver operating characteristic (ROC) curve analysis of the Lasso-logistic algorithm; **H** Confusion matrix heat map of the random forest algorithm; **I** ROC curve analysis of the random forest algorithm; **J** messenger RNA (mRNA) expression levels of *PDE8B*, *ESRRB*, and *SPRY2*

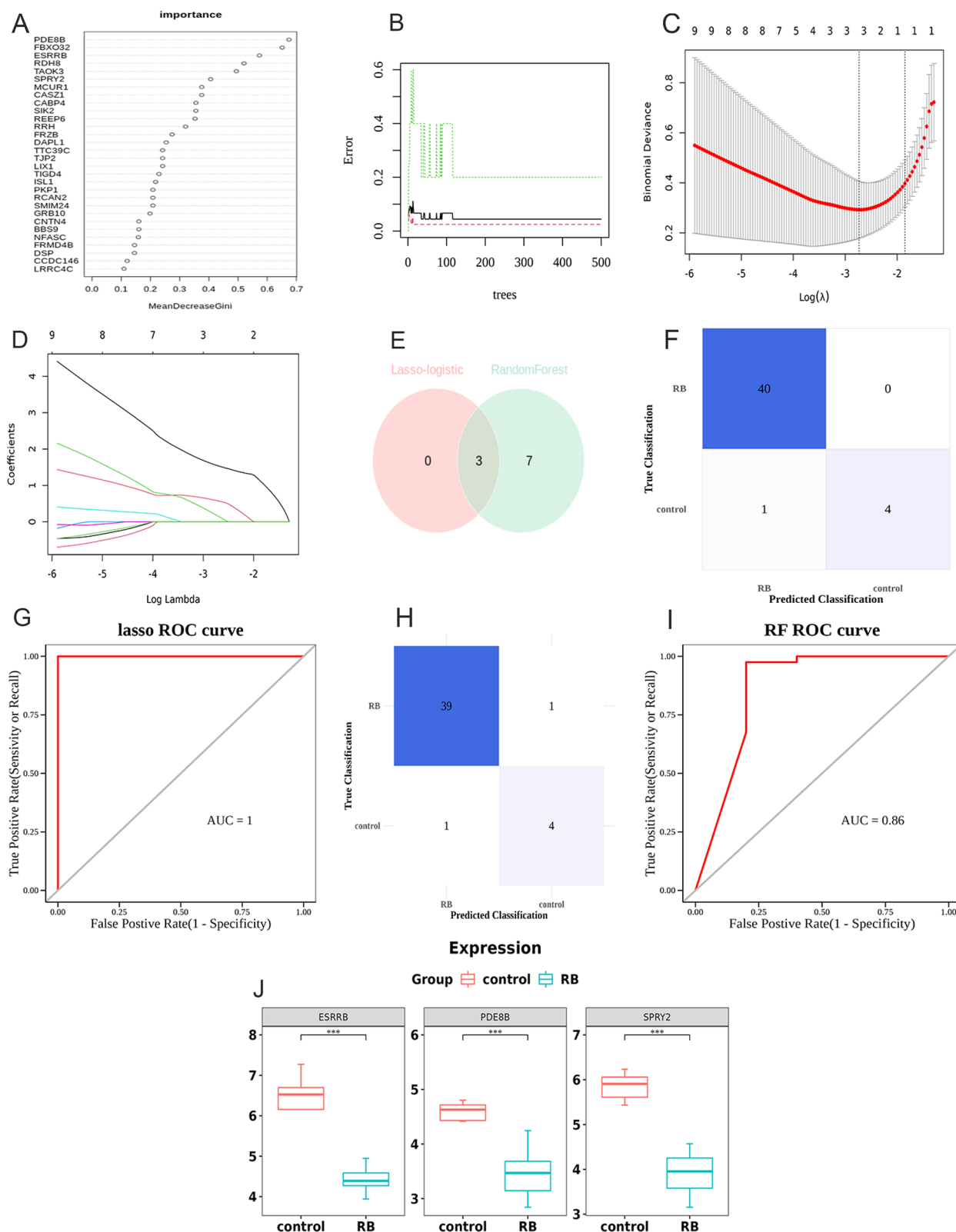


Fig. 4 (See legend on previous page.)

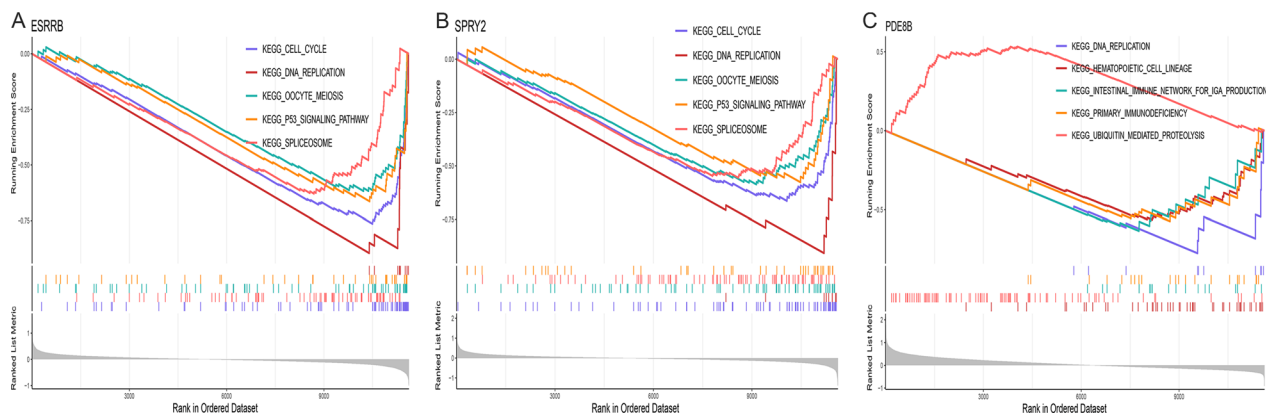


Fig. 5 Significantly enriched pathways of Hub differentially expressed retinoblastoma genes in retinoblastoma obtained by gene set enrichment analysis. **A** Enrichment plots for the five key pathways abnormally activated in estrogen-related receptor beta; **B** Enrichment plots for the five key pathways abnormally activated in Sprouty RTK signaling antagonist 2; **C** Enrichment plots for the five key pathways abnormally activated in phosphodiesterase 8B

Single-gene GSEA of Hub DERBGs in RB

A single-gene GSEA based on the KEGG gene sets was performed to determine the molecular mechanisms involved in Hub DERBGs in RB. As shown in Fig. 5, the top five KEGG pathways enriched by each Hub DERBG were identified. The estrogen-related receptor beta (*ESRRB*) gene was associated with cell cycle, spliceosome, and oocyte meiosis, as well as the p53 signaling pathway and DNA replication (Fig. 5A). Figure 5B shows the association between *SPRY2* and cell cycle, spliceosome, DNA replication, p53 signaling pathway, and meiosis of oocytes. There was a correlation between *PDE8B* and DNA replication, hematopoietic cell lineage, intestinal immune network for IgA production, primary immunodeficiency, and ubiquitin-mediated proteolysis (Fig. 5C).

Construction and assessment of a ceRNA regulatory network for Hub DERBGs in RB

To identify additional miRNAs and lncRNAs that may regulate the expression of Hub DERBGs, 54 DEMiRNAs were detected in the GSE41321 dataset, all of which were upregulated (Fig. 6A). Meanwhile, the miRWalk database was applied to predict miRNAs, and the predicted miRNAs were intersected with DEMiRNAs to determine the nine common miRNAs (hsa-miR-1225-5p, hsa-miR-1202, hsa-miR-342-3p, hsa-miR-146b-5p, hsa-miR-1207-5p, hsa-miR-892b, hsa-miR-665, hsa-miR-575, and hsa-miR-188-5p; Fig. 6B). Moreover, 83 DELncRNAs in the GSE125903 dataset were identified, of which 59 DELncRNAs were upregulated and 24 DELncRNAs were downregulated (Fig. 6C). In addition, based on these 9 common miRNAs, 13 lncRNAs (DLEU2, LINC00668, SNHG15, CRNDE, DLEU1, PTPRG-AS1, LINC00664, ENTPD3-AS1, EXTL3-AS1, LINC00963, SNHG7,

SNHG17, and LINC01134) and 4 miRNAs (hsa-miR-342-3p, hsa-miR-665, hsa-miR-185) were analyzed using the Starbase database (Fig. 6D). Finally, interactions between these miRNAs and lncRNAs were integrated to construct a ceRNA regulatory network, and the visualization of the network was carried out using Cytoscape (Fig. 6E).

Discussion

RB is one of the most serious eye diseases that can result in blindness, disability, and even death in infants [26–28]. Thus, the progression of novel strategies to diagnose and treat RB requires a detailed understanding of the mechanisms involved.

Additional file 2 illustrates the flow chart of the present study. This research aimed to identify RB DEGs and RB-associated gene modules using a systems biology approach called WGCNA. About 384 DEGs and nine module-clinical trait relationships significantly correlated with RB, which suggests that these module genes contribute significantly to the occurrence and progression of RB. From cross-DEG and RB-associated module genes, 133 DERBGs were obtained for further analysis. Further, functions and pathways involved in RB pathogenesis were examined. Fourteen GO terms and 5 KEGG pathways that were significantly enriched were identified. Among them, the notable ones are the pathways regulating the pluripotency of stem cells and ferroptosis. Certain cancers involve pathways of signal transmission that regulate stem cell pluripotency [29]. A new type of cell death characterized by distinct properties and recognizing functions that may be associated with physical conditions or different diseases, such as cancer, is called ferroptosis [30].

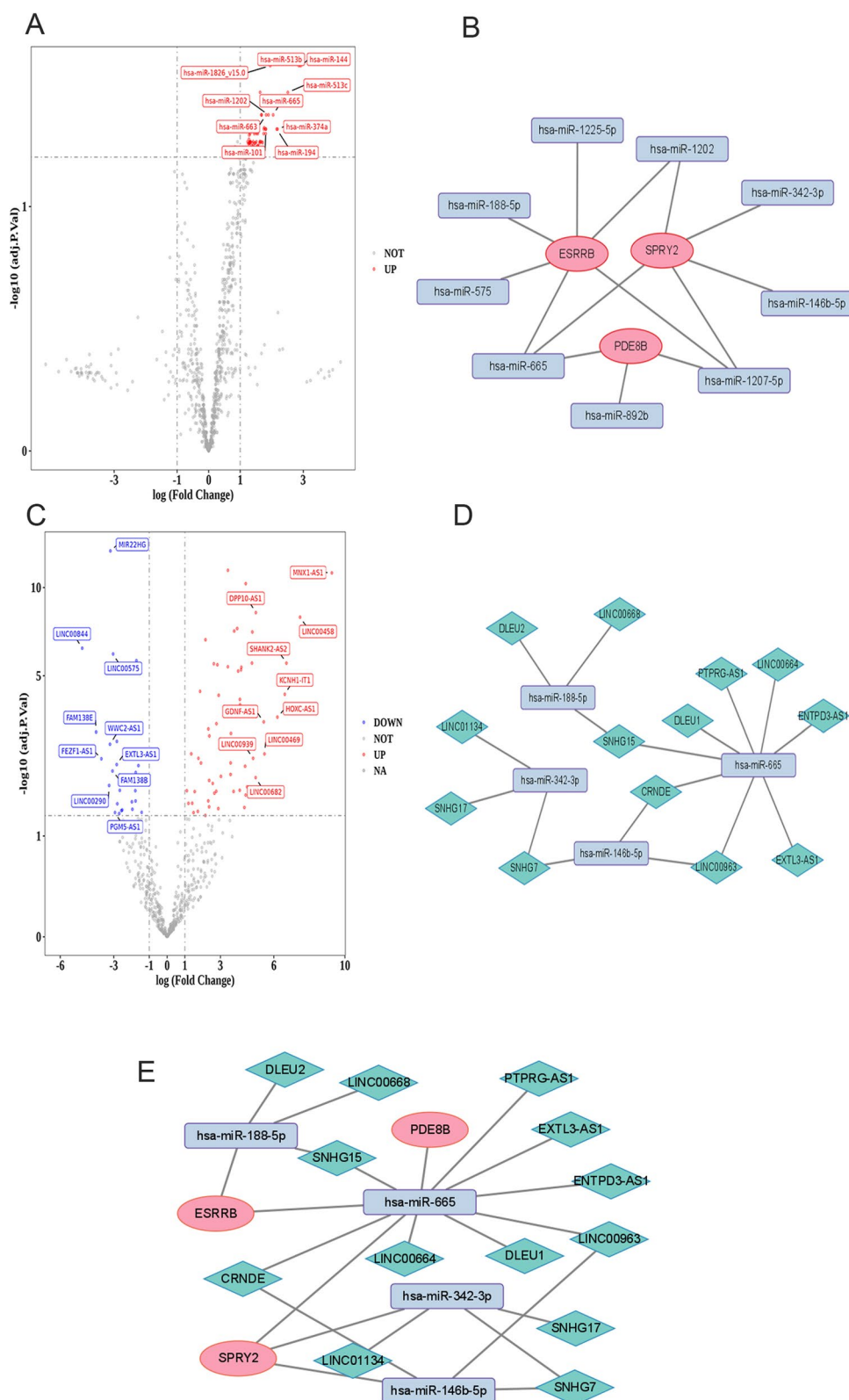


Fig. 6 ceRNA regulatory network of Hub differentially expressed retinoblastoma genes in retinoblastoma. **A** Volcano plot of differentially expressed miRNAs (DEmiRNAs) in the GSE41321 dataset; **B** DEmiRNAs network; **C** Volcano plot of differentially expressed long noncoding RNAs (DElncRNAs) in the GSE125903 dataset; **D** mi-LncRNAs network; **E** ceRNA network

PPI networks of DEGs were built. *RDH8*, *RGR*, *CNGA1*, *ROM1*, *SAG*, *RHO*, *PAX6*, *RLBP1*, *CNGB1*, and *RDH12* were identified as the ten most important Hub genes identified from PPI analysis. A previous bioinformatics study on RB also showed that *ROM1*, *CNGB1*, and *RDH12* may have a role in predicting the progress of RB, which is consistent with the findings of this research [31]. Previous studies have shown that *SAG* is a prospective target that could further be explored as a potential candidate in therapy and may further assist in understanding the mechanism of RB [32]. *SAG* is related to photoreceptors, which are the “cell of origin” in RB. The proteins might likely participate in unidentified pathways in RB. Interestingly, *RGR*, *CNGA1*, and *RLBP1* play an important role in retinitis pigmentosa. Whether they also play an important role in RB needs further investigation [33]. It can be inferred that the progression of RB may be significantly influenced by these genes.

The RF and Lasso logistics diagnostic models revealed the top three genes with the highest score degree, namely *PDE8B*, *ESRRB*, and *SPRY2*. *PDE8B* (phosphodiesterase 8B) is a gene encoding an enzyme that catalyzes the hydrolysis of a secondary messenger molecule, cAMP, by cyclic nucleotide phosphodiesterase (PDE). In addition, *SPRY2* (Sprouty RTK Signaling Antagonist 2) encodes a protein that belongs to the Sprouty family [34–37]. Outcomes from GSEA were enriched for p53 SIGNALING PATHWAY and SPLICEOSOME, in agreement with previous findings, indicating that the gene plays an important role in RB initiation and development [38, 39].

In an extensive range of processes, *ESRRB*, a protein-coding gene, plays an important role in the cell cycle, spliceosomes, and oocyte meiosis, as well as the p53 signaling pathway and DNA replication [40, 41]. Among them, HEMATOPOIETIC CELL LINEAGE, PRIMARY IMMUNODEFICIENCY, and UBIQUITIN-MEDIATED PROTEOLYSIS are strongly associated with the mechanism of cancer [42–45]. Nevertheless, the role of *ESRRB* in the progression of RB remains unknown.

Competitive endogenous RNA networks elucidate the mechanisms of RNA interactions that serve as key players in numerous biological processes. Although the precise mechanisms are yet to be fully understood, it is evident that these noncoding RNAs assume distinct functions in RB development. For example, study has shown that part of lncRNA DANCR can increase tumor aggressiveness [46]. The other study also showed that the lncRNA UCA1 promotes carboplatin resistance in RB cells by acting as a ceRNA for miR-206 [47]. Our study has revealed novel regulatory networks that were not previously reported. The regulatory analysis of the ceRNA network presents an opportunity for further validation and investigation through gene overexpression and knockout, as well as the use of double-luciferase reporter analysis.

The following are some of the highlights of this study. First, multiple datasets were used in this study, which resulted in stronger outcomes than those observed in previous studies. Second, two diagnostic models have been developed in this study for the first time before the selection of final key genes by crossover. From a comparative analytical perspective, this is worthy of further investigation. In conclusion, the follow-up analysis in this study was complete, and the constructed ceRNA could be used as a basis for further research. Nevertheless, a limitation of this study is that it was not possible to validate the results by Quantitative Real Time Polymerase Chain Reaction (qRT-PCR) due to the lack of clinical samples.

Conclusions

To summarize, this study identified key genetic components and the functional pathways that may contribute to the progression of RB. In this study, Hub genes and pathways were identified that may contribute to a better understanding of the mechanisms underlying RB pathogenesis. Bioinformatics methods were used to construct a regulatory network for ceRNA related to RB. As well as identifying potential prognostic biomarkers, a deeper understanding of the development of RB tumors has been achieved. In the future, more experimental studies are required to validate the underlying biological regulatory mechanisms involved.

Abbreviations

RB	Retinoblastoma
BP	Biological process
CC	Cellular component
MF	Molecular function
ceRNA	Competing endogenous RNA
GEO	Gene expression omnibus
DEGs	Differentially expressed genes
DERBGs	Differentially expressed retinoblastoma genes
KEGG	Kyoto encyclopedia of genes and genomes
PPI	Protein–protein interaction
STRING	Search tool to retrieve interacting genes and proteins
MCC	Maximal clique centrality
RF	Random forest
LASSO	Least absolute shrinkage and selection operator
DEmiRNAs	Differentially expressed micro RNAs
DElncRNAs	Differentially expressed long noncoding RNAs
PCA	Principal component analysis
qRT-PCR	Quantitative real time polymerase chain reaction
ROC	Receiver operating characteristic

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12920-023-01537-4>.

Additional file 1. Principal Component Analysis (PCA) of Gene Expression Datasets.

Additional file 2. Flow Chart of the Bioinformatics Analysis.

Acknowledgements

Not applicable.

Author contributions

JM, ML, and SL analyzed and interpreted the data and were the major contributors in writing the manuscript. They contributed equally to this work. WD, as the corresponding author, were primarily responsible for experimental design and revision of the paper. YX, XX, YC, and HX contributed to data acquisition and the rough draft. YX, WZ and JZ made a substantial contribution to the revision of the manuscript. All authors read and approved the final manuscript.

Funding

The study did not receive any government and non-government funding. No financial disclosures. The authors declare that there are no conflicts of interest with this work.

Availability of data and materials

The datasets analyzed during the current study are available in the "GSE110811 and GSE24673", (<http://www.ncbi.nlm.nih.gov/geo/>).

Declarations**Ethics approval and consent to participate**

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹The Eighth Affiliated Hospital of Sun Yat-sen University, Shenzheng, China.

²Aier Eye Hospital of Wuhan University, Wuhan, China. ³The First Affiliated Hospital of Kunming Medical University, Yunnan, China.

Received: 21 February 2023 Accepted: 9 May 2023

Published online: 13 May 2023

References

- Ortiz M, Dunkel I. Retinoblastoma. *J Child Neurol*. 2016;31(2):227–36.
- Yang J, Dang Y, Zhu Y, Zhang C. Diffuse anterior retinoblastoma: current concepts. *Onco Targets Ther*. 2015;8:1815–21.
- Knudson A. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA*. 1971;68(4):820–3.
- Rushlow D, Mol B, Kennett J, Yee S, Pajovic S, Thériault B, Prigoda-Lee N, Spencer C, Dimaras H, Corson T, et al. Characterisation of retinoblastomas without RB1 mutations: genomic, gene expression, and clinical studies. *Lancet Oncol*. 2013;14(4):327–34.
- Abramson D, Daniels A, Marr B, Francis J, Brodie S, Dunkel I, Gobin Y. Intra-arterial chemotherapy (ophthalmic artery chemosurgery) for group D retinoblastoma. *PLoS ONE*. 2016;11(1):e0146582.
- Shields C, Manjandavida F, Lally S, Pieretti G, Arepalli S, Caywood E, Jabbour P, Shields J. Intra-arterial chemotherapy for retinoblastoma in 70 eyes: outcomes based on the international classification of retinoblastoma. *Ophthalmology*. 2014;121(7):1453–60.
- Abramson D, Fabius A, Francis J, Marr B, Dunkel I, Brodie S, Escuder A, Gobin Y. Ophthalmic artery chemosurgery for eyes with advanced retinoblastoma. *Ophthalmic Genet*. 2017;38(1):16–21.
- Wu N, Chen S, Luo Q, Jiang Z, Wang X, Li Y, Qiu J, Yu K, Yang Y, Zhuang J. Kruppel-like factor 2 acts as a tumor suppressor in human retinoblastoma. *Exp Eye Res*. 2022;216: 108955.
- Shi K, Zhu X, Wu J, Chen Y, Zhang J, Sun X. Centromere protein E as a novel biomarker and potential therapeutic target for retinoblastoma. *Bioengineered*. 2021;12(1):5950–70.
- Wen Y, Zhu M, Zhang X, Xiao H, Wei Y, Zhao P. Integrated analysis of multiple bioinformatics studies to identify microRNA-target gene-transcription factor regulatory networks in retinoblastoma. *Transl Cancer Res*. 2022;11(7):2225–37.
- Gao Y, Luo X, Zhang J. Sp1-mediated up-regulation of Inc00152 promotes invasion and metastasis of retinoblastoma cells via the miR-30d/SOX9/ZEB2 pathway. *Cell Oncol (Dordr)*. 2021;44(1):61–76.
- Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinf*. 2008;9:559.
- Tian Z, He W, Tang J, Liao X, Yang Q, Wu Y, Wu G. Identification of important modules and biomarkers in breast cancer based on WGCNA. *Onco Targets Ther*. 2020;13:6805–17.
- Ritchie M, Phipson B, Wu D, Hu Y, Law C, Shi W, Smyth G. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7): e47.
- Chen H, Boutros P. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinf*. 2011;12:35.
- Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Cambridge (Mass))*. 2021;2(3):100141.
- Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf*. 2013;14:7.
- Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
- Kanehisa M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci Publ Protein Soc*. 2019;28(11):1947–51.
- Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res*. 2023;51:D587–92.
- Shannon P, Markiel A, Ozier O, Baliga N, Wang J, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13(11):2498–504.
- Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33(1):1–22.
- Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J, Müller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinf*. 2011;12:77.
- Bao Y, Wang L, Shi L, Yun F, Liu X, Chen Y, Chen C, Ren Y, Jia Y. Transcriptome profiling revealed multiple genes and ECM-receptor interaction pathways that may be associated with breast cancer. *Cell Mol Biol Lett*. 2019;24:38.
- Li J, Liu S, Zhou H, Qu L, Yang J. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 2014;42:D92–97.
- Ellsworth R. The practical management of retinoblastoma. *Trans Am Ophthalmol Soc*. 1969;67:462–534.
- Abramson D, Beaverson K, Chang S, Dunkel I, McCormick B. Outcome following initial external beam radiotherapy in patients with Reese-Ellsworth group Vb retinoblastoma. *Arch Ophthalmol (Chicago, Ill:1960)*. 2004;122(9):1316–23.
- Scott I, Murray T, Feuer W, Van Quill K, Markoe A, Ling S, Roth D, O'Brien J. External beam radiotherapy in retinoblastoma: tumor control and comparison of 2 techniques. *Arch Ophthalmol (Chicago, Ill:1960)*. 1999;117(6):766–70.
- Yang L, Shi P, Zhao G, Xu J, Peng W, Zhang J, Zhang G, Wang X, Dong Z, Chen F, et al. Targeting cancer stem cell pathways for cancer therapy. *Signal Transduct Target Ther*. 2020;5(1):8.
- Mou Y, Wang J, Wu J, He D, Zhang C, Duan C, Li B. Ferroptosis, a new form of cell death: opportunities and challenges in cancer. *J Hematol Oncol*. 2019;12(1):34.
- Cao M, Wang S, Zou J, Wang W. Bioinformatics analyses of retinoblastoma reveal the retinoblastoma progression subtypes. *PeerJ*. 2020;8: e8873.
- Naru J, Aggarwal R, Mohanty A, Singh U, Bansal D, Kakkar N, Agnihotri N. Identification of differentially expressed proteins in retinoblastoma tumors using mass spectrometry-based comparative proteomic approach. *J Proteomics*. 2017;159:77–91.
- Wang Q, Chen Q, Zhao K, Wang L, Wang L, Traboulsi E. Update on the molecular genetics of retinitis pigmentosa. *Ophthalmic Genet*. 2001;22(3):133–54.
- Milillo A, La Carpia F, Costanzi S, D'Urbano V, Martini M, Lanuti P, Vischini G, Larocca L, Marchisio M, Miscia S, et al. A SPY2 mutation leading to

- MAPK/ERK pathway inhibition is associated with an autosomal dominant form of IgA nephropathy. *Eur J Hum Genet EJHG*. 2015;23(12):1673–8.
35. Lim J, Wong E, Ong S, Yusoff P, Low B, Guy G. Sprouty proteins are targeted to membrane ruffles upon growth factor receptor tyrosine kinase activation. Identification of a novel translocation domain. *J Biol Chem*. 2000;275(42):32837–45.
 36. Appenzeller S, Schirmacher A, Halfter H, Bäumer S, Pendziwiat M, Timmerman V, De Jonghe P, Fekete K, Stögbauer F, Lüdemann P, et al. Autosomal-dominant striatal degeneration is caused by a mutation in the phosphodiesterase 8B gene. *Am J Hum Genet*. 2010;86(1):83–7.
 37. Horvath A, Giatzakis C, Tsang K, Greene E, Osorio P, Boikos S, Libè R, Patronas Y, Robinson-White A, Remmers E, et al. A cAMP-specific phosphodiesterase (PDE8B) that is mutated in adrenal hyperplasia is expressed widely in human and mouse tissues: a novel PDE8B isoform in human adrenal cortex. *Eur J Hum Genet EJHG*. 2008;16(10):1245–53.
 38. Romani A, Zauli E, Zauli G, AlMesfer S, Al-Swailem S, Voltan R. MDM2 inhibitors-mediated disruption of mitochondrial metabolism: a novel therapeutic strategy for retinoblastoma. *Front Oncol*. 2022;12:1000677.
 39. Tu J, Huo Z, Yu Y, Zhu D, Xu A, Huang M, Hu R, Wang R, Gingold J, Chen Y, et al. Hereditary retinoblastoma iPSC model reveals aberrant spliceosome function driving bone malignancies. *Proc Natl Acad Sci USA*. 2022;119(16): e2117857119.
 40. Collin R, Kalay E, Tariq M, Peters T, van der Zwaag B, Venselaar H, Oostrik J, Lee K, Ahmed Z, Caylan R, et al. Mutations of ESRRB encoding estrogen-related receptor beta cause autosomal-recessive nonsyndromic hearing impairment DFNB35. *Am J Hum Genet*. 2008;82(1):125–38.
 41. Bombail V, Collins F, Brown P, Saunders P. Modulation of ER alpha transcriptional activity by the orphan nuclear receptor ERR beta and evidence for differential effects of long- and short-form splice variants. *Mol Cell Endocrinol*. 2010;314(1):53–61.
 42. Hu T, Ghazaryan S, Sy C, Wiedmeyer C, Chang V, Wu L. Concomitant inactivation of Rb and E2f8 in hematopoietic stem cells synergizes to induce severe anemia. *Blood*. 2012;119(19):4532–42.
 43. Özyörük D, Güzelkücü Z, Metin A, Emir S, Yazal Erdem A, Kacar D, Koca Yozgat A, Aker C, Çakmakçı S, Özdemir S, et al. Clinical profile and outcomes of primary immunodeficiency and malignancy in childhood at a tertiary oncology center in developing country. *Pediatr Hematol Oncol*. 2022;39(7):600–12.
 44. Fragkandrea I, Nixon J, Panagopoulou P. Signs and symptoms of childhood cancer: a guide for early recognition. *Am Fam Physician*. 2013;88(3):185–92.
 45. Wang Y, Zheng Z, Zhang J, Wang Y, Kong R, Liu J, Zhang Y, Deng H, Du X, Ke Y. A novel retinoblastoma protein (RB) E3 ubiquitin ligase (NRBE3) promotes RB degradation and is transcriptionally regulated by E2F1 transcription factor. *J Biol Chem*. 2015;290(47):28200–13.
 46. Wang J, Yang Y, Li K. Long noncoding RNA DANCR aggravates retinoblastoma through miR-34c and miR-613 by targeting MMP-9. *J Cell Physiol*. 2018;233(10):6986–95.
 47. Wang N, Fan H, Fu S, Li S, Zhou B, Jin Q, You Z. Long noncoding RNA UCA1 promotes carboplatin resistance in retinoblastoma cells by acting as a ceRNA of miR-206. *Am J Cancer Res*. 2022;12(5):2160–72.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

