Pre-Clinical Research Report

Journal of INTERNATIONAL MEDICAL RESEARCH

Journal of International Medical Research
49(6) 1–10

DOI: 10.1177/03000605211022210
journals.sagepub.com/home/imr
SAGE

Bioinformatics analysis of key biomarkers for retinoblastoma

Xin-mei Zhao[1], Yuan-Bin Li[2], Peng Sun[2], Ya-di Pu[3], Meng-jie shan[4,5] and Yuan-meng Zhang[6]

Abstract

Objective: To identify key genes involved in occurrence and development of retinoblastoma.
Methods: The microarray dataset, GSE5222, was downloaded from the gene expression omnibus (GEO) database. Differentially expressed genes (DEGs) between unilateral and bilateral retinoblastoma were identified and functional enrichment analysis performed. The protein–protein interaction (PPI) network was constructed and analysed by STRING and Cytoscape.
Results: DEGs were mainly associated with activation of cysteine-type endopeptidase activity involved in apoptotic process and small molecule catabolic process. Seven genes (WAS, GNB3, PTGER1, TACR1, GPR143, NPFF and CDKN2A) were identified as HUB genes.
Conclusion: Our research provides more understanding of the mechanisms of the disease at a molecular level and may help in the identification of novel biomarkers for retinoblastoma.

Keywords

bioinformatics, differentially expressed genes, unilateral retinoblastoma, bilateral retinoblastoma

Date received: 12 December 2020; accepted: 7 May 2021

[1]Ophthalmic Laboratory, The Affiliated Yantai Yuhuangding Hospital of Qingdao University, Yantai, Shandong, China
[2]Department of Ophthalmology, The Affiliated Yantai Yuhuangding Hospital of Qingdao University, Yantai, Shandong, China
[3]Qingdao University, Qingdao, Shandong, China
[4]Department of Plastic Surgery, Peking Union Medical College Hospital, Beijing, China
[5]Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China
[6]Department of Cardiology, The Third Medical Centre of Chinese PLA General Hospital, No. 69, Yongding Road, Hai Dian, Beijing, China

Corresponding author:
Yuan-Bin Li, Department of Ophthalmology, The Affiliated Yantai Yuhuangding Hospital of Qingdao University, 20 Yuhuangding East Road, Yantai, Shandong 264000, P.R. China.
Email: ytliyuanbin@163.com

Introduction

Although rare, retinoblastoma is the most common primary intraocular cancer of childhood and is associated with a high morbidity.[1] The global incidence rate is 1 in 16,000–18,000 live births and most cases are diagnosed before 2 years of age.[2,3] The tumour originates from embryonic retinal cells and there are two forms according to genetic invovlement.[3] The genetic form accounts for approximately 40% of all cases, typically affects both eyes (bilateral) with a median age of onset of 12 months.[3,4] The non-genetic form accounts for approximately 60% of all cases, with a median age at diagnosis of 2 years and is usually unilateral.[3,4] All children with the bilateral form, and 10–15% of children with the unilateral form, carry an RB1 gene mutation.[4] In China, there are approximately 1100 new cases of retinoblastoma each year scattered across 32 provinces.[2] In high income countries, the survival rate of children with this disease has been estimated to be over 95%, whereas in low-income countries it is approximately 30%.[2] Before 2005, the main treatment for retinoblastoma in China was enucleation, and most children had a poor prognosis and a fatal outcome.[2] However, a successful strategy has been developed in China whereby improved efficiency and use of common protocols has led to standardized classification and treatment of newly diagnosed patients which has led to increased survival.[2]

Early genetic analysis could assist in the screening of children with high risk of developing bilateral disease, which may be help protect the eyesight of children via risk-adapted follow-up and early treatment. Therefore, research into retinoblastoma genetics is extremely useful in the understanding and treatment of the disease. Microarray and bioinformatics analysis are now widely used to identify potential gene targets for the diagnosis and therapy of different cancers. In this study, we downloaded a microarray dataset and analysed gene expression to obtain differentially expressed genes (DEGs) between unilateral and bilateral retinoblastoma. Our objective was to identify candidate genes that may be useful as biomarkers for retinoblastoma.

Methods

Microarray database

All data used in this study were from publicly accessible databases and so ethical approval was not required. Using GPL96, Affymetrix HG-U133 GeneChip technology, we obtained gene expression profile for the microarray dataset, GSE5222, obtained from Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo). According to recommendations, the probe was transformed into an homologous gene symbol. GSE5222 included data from 10 sporadic unilateral retinoblastoma cases and 2 sporadic bilateral retinoblastoma cases.

Identification of DEGs

Differentially expressed genes (DEGs) between unilateral and bilateral retinoblastoma were identified using GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r/) an interactive tool in GEO. For the dataset GSE5222, statistical significance was defined at $P \leq 0.05$ and fold change ≥ 1.

GO and KEGG analysis of DEGs

Gene Ontology (GO) enrichment analysis is widely used in bioinformatics and is separated into three aspects of biology: biological processes (BP); cell components (CC); molecular functions (MF). Kyoto Encyclopaedia of Genes and Genomes (KEGG; https://www.kegg. jp/) is a public database resource for genome sequencing and other high-throughput experimental technologies for large molecular dataset

generation. Metascape (http://metascape.org/) is an online tool for functional annotations and visual analysis of DEGs.[5] In this study, the GO and KEGG analyses of DEGs were conducted using Metascape. Significance was defined at $P< 0.05$. In addition, the Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david. ncifcrf.gov) (version6.8), a database that stores comprehensive information about the functions of genes and proteins was also used to analyse the function of DEGS.[6]

Building and analysis of the protein–protein interaction (PPI) network

The Search Tool for the Retrieval of Interacting Genes (STRING, http://string-db.org) an online dataset, can predict and construct a PPI network after importing common DEGs identified in the GEO datasets.[7] The software, Cytoscape, was used to visualize and draw the PPI network to help identify HUB genes. The most important modules of the network map were identified by the Cytoscape plug-in, Molecular Complex Detection (MCODE, version 1.5.1; http://apps.cytoscape.org/apps/mcode).

Identification of HUB genes associated with retinoblastoma

The Comparative Toxicogenomics Database (CTD; http://ctdbase.org/) is used to investigate associations between genes and human diseases.[8] The relationship between gene products identified in our datasets and retinoblastoma were analysed using this database.

Identification of miRNA-gene pairs

TargetScan (www.targetscan.org) is an online database that can predict the biological targets of miRNAs.[9] In this study, TargetScan was used to identify the DE miRNAs that regulate the HUB genes.

Results

Identification of DEGs

Analysis of GSE5222 dataset identified 9165 DEGs between unilateral and bilateral retinoblastoma; these included 3533 up-regulated and 5632 down-regulated genes (Figure 1a). The heat maps of the GSE5222 dataset are shown in Figure 1b.

GO and KEGG analysis of DEGs

Functional enrichment analysis with Metascape showed that DEGs between unilateral and bilateral retinoblastoma were significantly 'enriched in activation of cysteine-type endopeptidase activity involved in apoptotic process', 'small molecule catabolic process' and 'toxoplasmosis' (Figure 2a b and c).

DAVID results of the GO analysis found that there were major variations in BP including: positive regulation of gene expression; response to lipopolysaccharide; activation of cysteine-type endopeptidase activity involved in apoptotic process (Figure 3a). Variations in MF of DEGs included: symporter activity; death effector domain binding; cysteine-type endopeptidase activity involved in apoptotic signalling pathway (Figure 3b). Variations in CC of DEGs included: integral component of plasma membrane; ripoptosome; CD95 death-inducing signalling complex (Figure 3c). Analysis of KEGG pathways indicated that the top pathways associated with DEGs were, pathways in cancer, p53 signalling pathway and Hepatitis B (Figure 3d).

PPI network construction and identification of HUB genes

The PPI network of DEGs was obtained via the STRING database (Figure 4a). The PPI network was analysed using MCODE to identify HUB genes (Figures 4b, c and d). Seven genes were identified as HUB genes

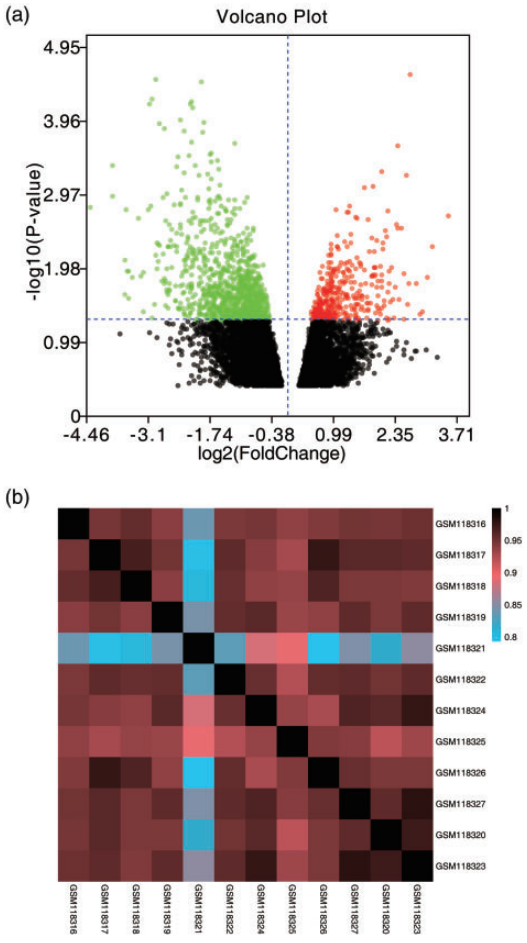

Figure 1. Expression of differentially expressed genes (DEGs) in the GSE5222 dataset. (a) Volcano plots of DEGs between unilateral retinoblastoma and bilateral retinoblastoma. (b) Heat maps of the DEGs between unilateral retinoblastoma and bilateral retinoblastoma. Dark red represents high expression and blue represents low expression.

(WAS, GNB3, PTGER1, TACR1, GPR143, NPFF and CDKN2A). Six genes were down-regulated (WAS, GNB3, PTGER1, TACR1, NPFF and CDKN2A) and one (GPR143) was up-regulated.

Identification of HUB genes related to cancer

The HUB gene-retinoblastoma interaction was investigated in CTD database. Figure 5 shows the results as radar diagrams.

Prediction of miRNAs that regulate HUB genes

The DE miRNAs that regulate the HUB genes identified in this study are shown in Table 1.

Discussion

Genetic factors are important in determining the prognosis of retinoblastoma. The retinoblastoma gene is a tumour suppressor

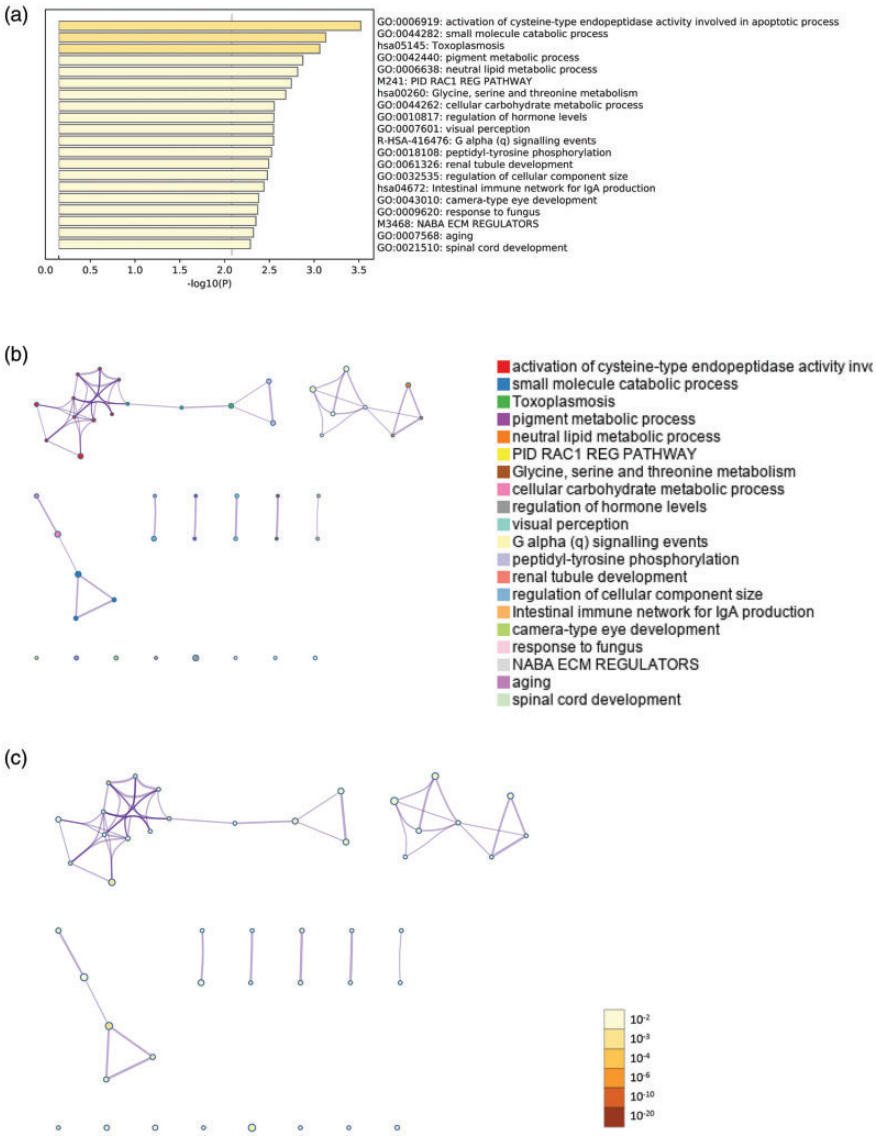

Figure 2. (a) Gene ontology (GO) analysis of common differentially expressed genes (DEGs) using Metascape. Heatmap of enriched terms across lists of input differentially expressed genes, coloured by *P*–values. (b) Network of enriched terms coloured by cluster identity, where nodes that share the same cluster identity are typically close to each other. (c) Network of enriched terms, coloured by *P* –values, where terms containing more genes tend to have more significant *P*–values.

gene that codes for the RB protein. Disease occurs from any mutation that inactivates both RB1 alleles.[2] All bilateral and multifocal unilateral forms are hereditary and all children with these forms, and 10–15% of children with the unilateral form, carry an RB1 gene mutation.[3]

In this study, we obtained GSE5222 gene expression file dataset from the GEO database. We identified 9165 DEGs between

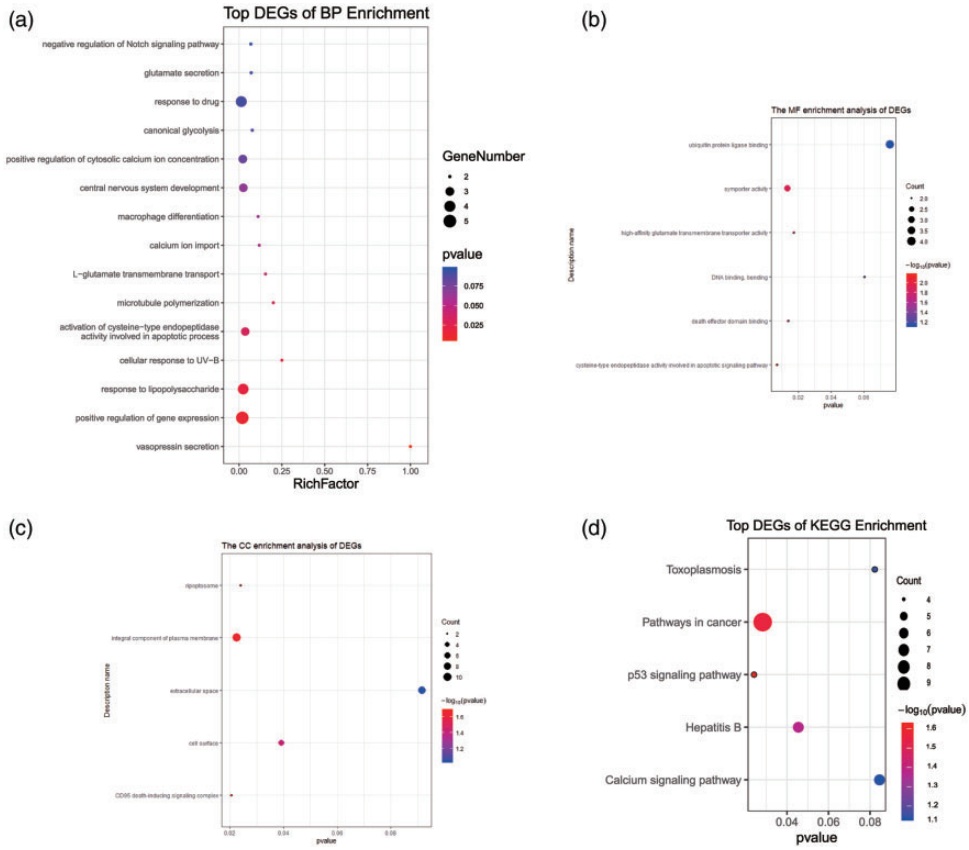

Figure 3. Enrichment analysis of differentially expressed genes (DEGs) by DAVID. Information relating to changes in Gene Ontology (GO) categories (a) biological processes (b) molecular function (c) cell component. (d) Kyoto Encyclopaedia of Genes and Genomes (KEGG) analyses of common DEGs.

unilateral and bilateral retinoblastoma. Based on DAVID analysis of DEGs 'positive regulation of gene expression' had the highest GO-BP enrichment score. This finding suggests that up-regulated genes affect the severity of retinoblastoma. Based on analysis of DEGs 'activation of cysteine-type endopeptidase activity involved in apoptotic process' had the highest GO-MF enrichment score. Apoptosis is an important biological process that affects the development of retinoblastoma. A previous study found that cell proliferation and apoptosis are linked in retinoblastoma and suggested that apoptosis in retinoblastoma is a caspase-3-independent pathway.[10]

KEGG analysis indicated that the DEGs were enriched in p53 signalling pathways and Hepatitis B. p53 is a tumour suppressor gene and mutations in this gene have been estimated to occur in more than 50% of all malignancies.[11] The protein encoded by this gene is a transcriptional factor, which controls the initiation of the cell cycle. When a cell is damaged and cannot be repaired, the p53 pathway plays a role in causing the cell to die by apoptosis.[12] The p53 gene normally slows down or monitors cell division, and abnormalities in this gene can lead to

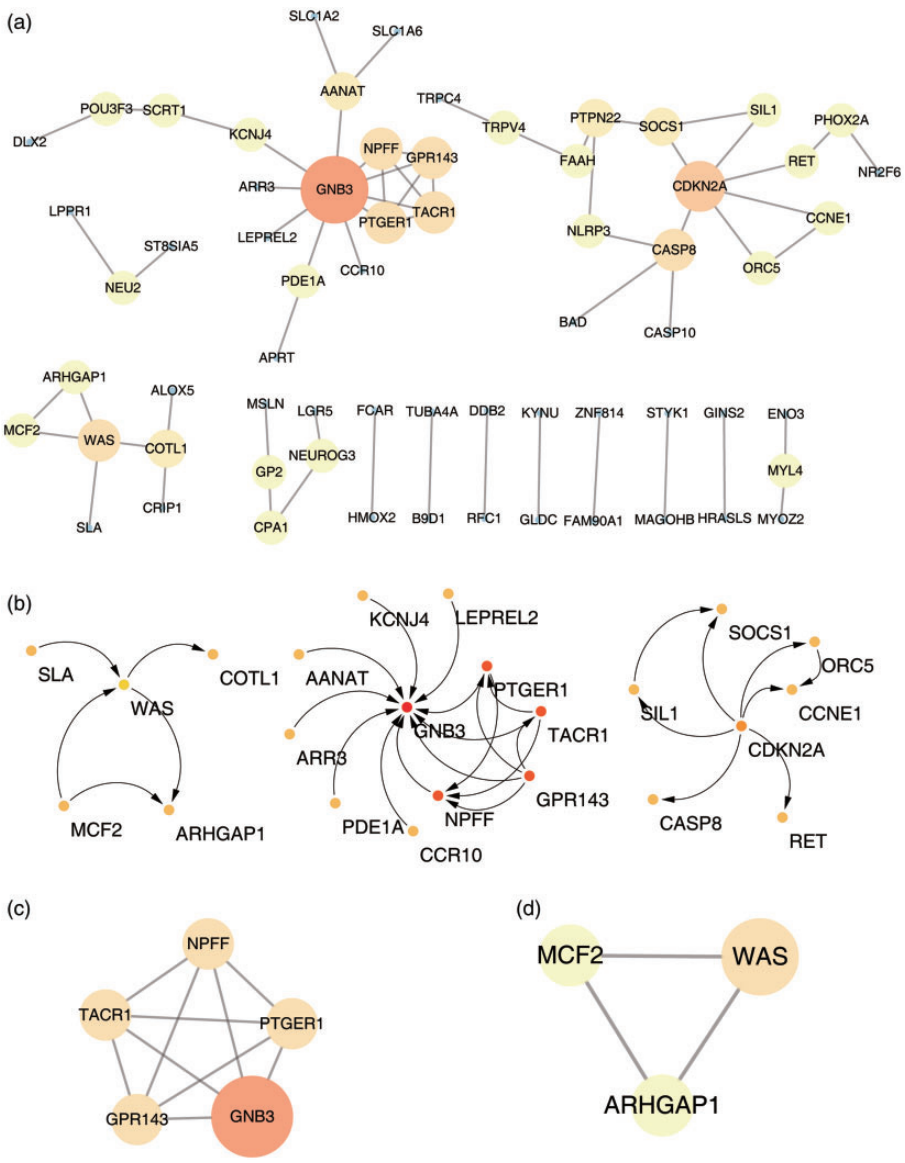

Figure 4. (a) The protein–protein interaction (PPI) network complex for the common differentially expressed genes (DEGs). (b) The top ten highly connected genes of the PPI network. (c) MCODE analysis of DEGs showed a module containing GNB3, GPR143, TACR1, NPFF and PTGER1. (d) MCODE analysis of DEGs also showed a second module containing WAS, MCF2, ARHGAP1.

cancer.[13] We suggest that the p53 pathway is an important factor involved in the formation of unilateral and bilateral retinoblastoma but this theory needs further verification.

Hepatitis B virus (HBV) infection is a global disease with high morbidity and after acute infection, HBV can exist persist for life.[14] While the mechanisms of carcinogenesis are unclear, the presence of HBV

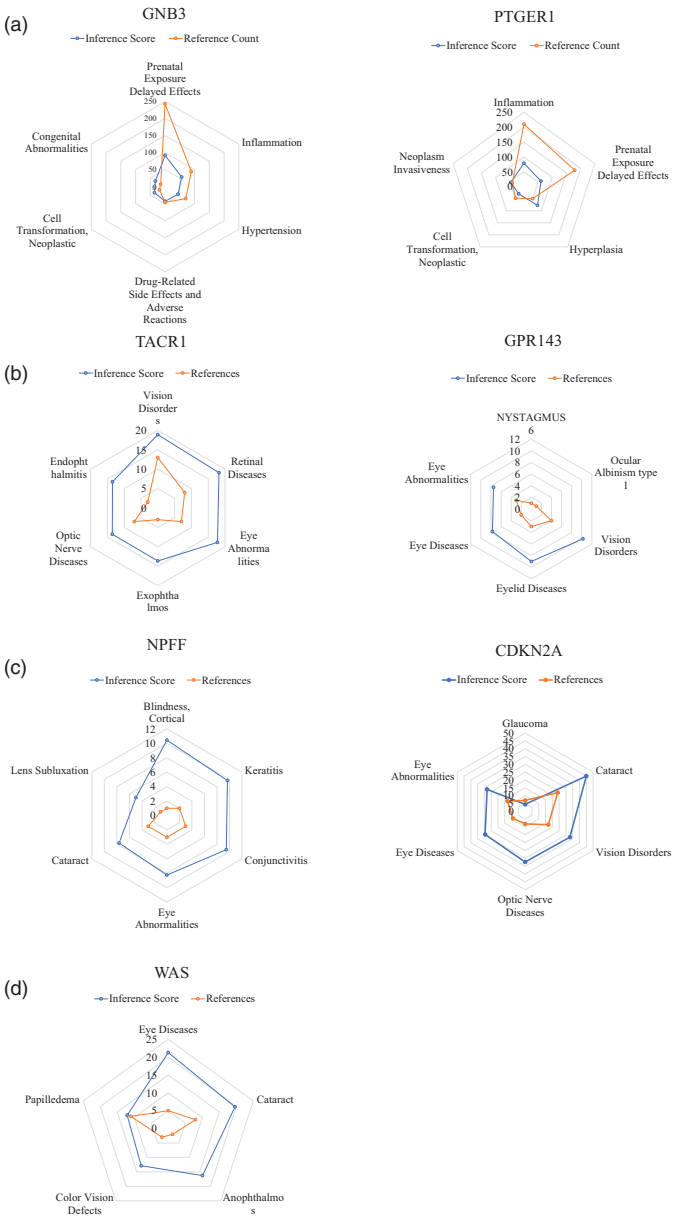

Figure 5. Relationship between unilateral retinoblastoma and bilateral retinoblastoma as related to common differentially expressed genes (DEGs) based on the Comparative Toxicogenomics Database (CTD) database.

may influence the development and progression of retinoblastoma. Indeed, one study found that RB1 mutation was related to the occurrence of liver cancer.[15] We are not aware of any other studies that have found an association between HBV and retinoblastoma. Moreover, our KEGG analysis was based on sequencing files. Therefore, we suggest further research is required to corroborate our findings.

Table 1. Summary of miRNAs that regulate HUB genes.

HUB Gene	Predicted miRNA	HUB Gene	Predicted miRNA
GNB3	hsa-miR-521	NPFF	hsa-miR-3166
	hsa-miR-4436b-5p		hsa-miR-1273h-3p
	hsa-miR-6509-3p		hsa-miR-4771
PTGER1	hsa-miR-1306-3p	CDKN2A	hsa-miR-1307-3p
	hsa-miR-5580-5p		hsa-miR-4268
	hsa-miR-4498		hsa-miR-760
TACR1	hsa-miR-613	WAS	hsa-miR-4436b-3p
	hsa-miR-206		hsa-miR-4632-5p
	hsa-miR-1-3p		hsa-miR-7843-5p
GPR143	hsa-miR-711		
	hsa-miR-4638-3p		
	hsa-miR-378b		

Seven HUB genes were identified (WAS, GNB3, PTGER1, TACR1, GPR143, NPFF and CDKN2A). Six of these genes were down-regulated but G-protein coupled receptor 143 (GPR143) was up-regulated. GPR143 is a receptor for tyrosine, L-DOPA and dopamine.[16] GPR143 binds to L-DOPA, then stimulates Ca2+ influx in the cytoplasm, and increases the secretion of neurotrophic factor SERPINF1 (*GPR137B protein (human) - STRING interaction network [string-db.org]*). This signal transduction is ligand-dependent and occurs in melanocytes through a G (q)-mediated pathway.[17] Its activity is mediated by G proteins which activate the phosphoinositide signalling pathway. GPR143 also acts as an intracellular G protein-coupled receptor involved in the biogenesis, organization and transportation of melanosomes.[18] The gene has been linked to eye disorders in other studies. For example, one study found that GPR143 mutations caused developmental eye defects in X-linked ocular albinism type 1.[16]

Six genes were downregulated (WAS, GNB3, PTGER1, TACR1, NPFF and CDKN2A). WAS encodes for the Wiskott-Aldrich syndrome protein which is involved in innate and adaptive immunity via regulation of actin cytoskeleton-dependent cellular processes.[19] The GNB3 gene encodes for Guanine nucleotide-binding protein which participates as a modulator or transducer in several transmembrane signalling systems.[20] PTGER1 gene codes for a G protein-coupled receptor (GPCR) of the rhodopsin-like receptor family. The TACR1 gene encodes the receptor for the tachykinin neuropeptide substance P, also referred to as neurokinin 1 and it may be related to G proteins.[21] NPFF gene encodes for Neuropeptide FF peptides. The CDKN2A gene provides instructions for making several proteins, of which the most well-studied are the p16(INK4A) and the p14(ARF) proteins.

In conclusion, we downloaded and analysed the GSE5222 dataset using GEO2R tools to search for DEGs between unilateral and bilateral retinoblastoma. DEGs were analysed to identify genes that may play an important role between unilateral and bilateral retinoblastoma. Our research provides more understanding of the mechanisms of the disease at a molecular level and may help in the identification of novel biomarkers for the diagnosis and treatment of retinoblastoma.

Declaration of conflicting interests

The authors declare that there are no conflicts of interest.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the Ophthalmic Laboratory Research Fund of Yantai Yuhuangding Hospital (no. Kj0184).

ORCID iDs

Meng-jie shan https://orcid.org/0000-0002-3451-1103
Yuan-meng Zhang https://orcid.org/0000-0003-0236-0635

References

1. Rao R and Honavar SG. Retinoblastoma. *Indian J Pediatr* 2017; 84: 937–944.
2. Dimaras H, Corson TW, Cobrinik D, et al. Retinoblastoma. *Nat Rev Dis Primers* 2015; 1: 15021
3. Aerts I, Lumbroso-Le Rouic L, Gauthier-Villars M, et al. [Retinoblastoma update]. *Arch Pediatr* 2016; 23: 112–116.
4. Melamud A, Palekar R and Singh A. Retinoblastoma. *Am Fam Physician* 2006; 73: 1039–1044.
5. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019; 10: 1523.
6. Huang DW, Sherman BT, Tan Q, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *GenomeBiol* 2007; 8: R183.
7. Zhao C and Sahni S. String correction using the Damerau-Levenshtein distance. *BMC Bioinformatics* 2019; 20: 277.
8. Davis AP, Grondin CJ, Johnson RJ, et al. The Comparative Toxicogenomics Database: update 2019. *Nucleic Acids Res* 2019; 47: D948–D954.
9. Zhu KP, Zhang CL, Ma XL, et al. Analyzing the Interactions of mRNAs and ncRNAs to Predict Competing Endogenous RNA Networks in Osteosarcoma Chemo-Resistance. *Mol Ther* 2019; 27: 518–530.
10. Sitorus RS, Gumay S and Van Der Valk P. The apoptotic paradox in retinoblastoma. *Ann N Y Acad Sci* 2009; 1171: 77–86.
11. Duffy MJ, Synnott NC and Crown J. Mutant p53 as a target for cancer treatment. *Eur J Cancer* 2017; 83: 258–265.
12. Lacroix M, Riscal R, Arena G, et al. Metabolic functions of the tumor suppressor p53: Implications in normal physiology, metabolic disorders, and cancer. *Mol Metab* 2020; 33: 2–22.
13. Whibley C, Pharoah PD and Hollstein M. p53 polymorphisms: cancer implications. *Nat Rev Cancer* 2009; 9: 95–107.
14. Seeger C, Mason WS. Molecular biology of hepatitis B virus infection. *Virology* 2015; 479–480: 672–686.
15. Ahn SM, Jang SJ, Shim JH, et al. Genomic portrait of resectable hepatocellular carcinomas: implications of RB1 and FGF19 aberrations for patient stratification. *Hepatology* 2014; 60: 1972–1982.
16. De Filippo E, Schiedel AC and Manga P. Interaction between G Protein-Coupled Receptor 143 and Tyrosinase: Implications for Understanding Ocular Albinism Type 1. *J Invest Dermatol* 2017; 137: 457–465.
17. Palmisano I, Bagnato P, Palmigiano A, et al. The ocular albinism type 1 protein, an intracellular G protein-coupled receptor, regulates melanosome transport in pigment cells. *Hum Mol Genet* 2017; 26: 3028–3029.
18. Innamorati G, Piccirillo R, Bagnato P, et al. The melanosomal/lysosomal protein OA1 has properties of a G protein-coupled receptor. *Pigment Cell Res* 2006; 19: 125–135.
19. McEvoy JD and Dyer MA. Genetic and Epigenetic Discoveries in Human Retinoblastoma. *Crit Rev Oncog* 2015; 20: 217–225.
20. Krishnan R, Boddapati N and Mahalingam S. Interplay between human nucleolar GNL1 and RPS20 is critical to modulate cell proliferation. *Sci Rep* 2018; 8: 11421.
21. Schank JR, Tapocik JD, Barbier E, et al. Tacr1 gene variation and neurokinin 1 receptor expression is associated with antagonist efficacy in genetically selected alcohol-preferring rats. *Biol Psychiatry* 2013; 73: 774–781.