Journal of Population Therapeutics & Clinical Pharmacology

RESEARCH ARTICLE
DOI: 10.53555/jptcp.v31i6.6813

DECIPHERING KEY HUB GENES IN MEDULLOBLASTOMA: INTEGRATIVE ANALYSIS OF TRANSCRIPTOMIC DATASETS REVEALS POTENTIAL BIOMARKERS AND THERAPEUTIC TARGETS

Muhammad Afzal[1], Zahid Iqbal Rajput[2], Sobia Mumtaz[3], Kashif Prince[4], Shahid Moin Qureshi[5], Hafiz Shehzad Muzammil[6], Muhammad Junaid[7], Dilshad Rashid[8], Umair Younas[9], Maryam Saddiqa[10], Muhammad Arif Rizwan[11]*

[1]Department of Biochemistry, Faisalabad Medical University, Faisalabad, Pakistan
[2]Department of Microbiology, Shaheed Benazir Bhutto University of Veterinary & Animal Sciences, Sakrand, Pakistan
[3]Livestock and Dairy Development Department Punjab, Pakistan
[4]Department of Medicine, Cholistan University of Veterinary & Animal Sciences, Bahawalpur, Pakistan
[5]Department of Human Anatomy/Histology & Embryology, Hebei North University, Hebei, P.R. China
[6]Department of Human Nutrition and Food Technology, Faculty of Allied Health Sciences, Superior University, Lahore, Pakistan
[7]Department of Dairy Technology, Faculty of Animal Production and Technology, University of Veterinary and Animal Sciences, Lahore, Pakistan
[8]Para Veterinary Institute Layyah, sub campus, University of Veterinary and Animal Science, Lahore, Pakistan
[9]Department of Livestock Management, Faculty of Animal Production and Technology (FAPT), Cholistan University of Veterinary and Animal Sciences, Bahawalpur, Pakistan
[10]Department of Livestock & Dairy Development Punjab, Pakistan
[11]*Institute of Continuing Education and Extension, Cholistan University of Veterinary and Animal Sciences, Bahawalpur, Pakistan

***Corresponding author:** Muhammad Arif Rizwan, Institute of Continuing Education and Extension, Cholistan University of Veterinary and Animal Sciences, Bahawalpur, Pakistan, Email address: marifrizwan@cuvas.edu.pk

Abstract

Introduction: Medulloblastoma (MB) is the most common malignant brain tumor in children, accounting for about 20% of all pediatric brain tumors. Originating in the cerebellum or posterior fossa, MB poses significant treatment challenges despite advances in surgical, radiation, and chemotherapeutic interventions. Understanding the molecular underpinnings of MB is crucial for developing targeted therapies and improving diagnostic tools.

Methods: This study utilized the GSE42656 dataset from the GEO database to identify differentially expressed genes (DEGs) and hub genes associated with MB. Using the limma package, we screened the top 50 DEGs and constructed a protein-protein interaction (PPI) network via the STRING database. Central nodes and potential key regulators were identified. Cytoscape software was

employed to determine hub genes using the degree method, and their diagnostic efficacy was evaluated using ROC curve analysis.
Results: The PPI network analysis highlighted central nodes such as GABRG2, STXBP1, and DLG4, suggesting their pivotal roles in MB. Notable genes like MAP4, RBFOX1, and NPTN, though less connected, were also significant. Clusters of interactions involved genes such as CLTA, PRKCE, and AASS, indicating potential pathways or complexes. Hub genes identified included AASS, CAMKB2, MAP4, and SLC12A5, all central due to their high connectivity. AASS showed high expression in MB samples, while MAP4 and SLC12A5 exhibited low expression. CAMKB2 displayed varying expression levels, indicating complex regulatory dynamics. ROC curve analysis demonstrated high diagnostic efficiency for these hub genes, with significant AUC values, highlighting their potential as biomarkers.
Conclusion: This study provides a comprehensive analysis of the genetic basis of MB, identifying key DEGs and hub genes with significant diagnostic potential. These findings contribute to understanding MB pathogenesis and underscore the importance of integrating genomic data with bioinformatics tools to uncover critical molecular mechanisms, offering promising targets for future research and therapeutic development.

Keywords: Medulloblastoma; Core hub genes; ROC curve; GEO2R

Introduction

Medulloblastoma (MB) is the most common malignant brain tumor in children, accounting for approximately 20% of all pediatric brain tumors [1, 2]. It originates in the cerebellum or posterior fossa and is classified as a primitive neuroectodermal tumor. The genetic basis of MB is complex and involves multiple genetic and epigenetic alterations that drive tumor development and progression. Key genetic changes frequently observed in MB include mutations, copy number variations, and chromosomal rearrangements that affect critical signaling pathways such as the Sonic Hedgehog (SHH), WNT, and Notch pathways [3, 4]. These alterations can lead to uncontrolled cell proliferation, impaired differentiation, and resistance to apoptosis. Recent advances in genomic and transcriptomic profiling have identified several recurrent mutations in genes such as TP53, PTCH1, and MYC, which play pivotal roles in the tumorigenesis of MB [5, 6]. Additionally, epigenetic modifications, including DNA methylation and histone modifications, have been shown to contribute to the deregulation of gene expression in MB [7]. Understanding the genetic landscape of medulloblastoma is essential for the development of targeted therapies and personalized treatment strategies, which hold promise for improving patient outcomes and reducing treatment-related side effects. Despite advancements in surgical techniques, radiation therapy, and chemotherapy, the prognosis for MB patients remains variable, with long-term survival rates ranging from 50% to 70% [8]. The identification of molecular markers and therapeutic targets is crucial for improving diagnostic accuracy and developing targeted treatments.

Microarray technology has revolutionized the field of genomics by enabling the simultaneous analysis of thousands of genes [9]. This high-throughput approach provides comprehensive gene expression profiles that can be used to identify differentially expressed genes (DEGs) and understand complex biological processes [9]. The Gene Expression Omnibus (GEO) database, maintained by the National Center for Biotechnology Information (NCBI), is a publicly accessible repository that archives and freely distributes high-throughput gene expression data submitted by the scientific community [10]. It serves as a valuable resource for researchers aiming to explore gene expression patterns in various diseases [10].

In our study, we aimed to identify DEGs and hub genes associated with MB by utilizing the GSE42656 dataset from the GEO database. This dataset includes gene expression profiles from 5 BM patients and 8 normal samples. By analyzing this data, we sought to uncover key genetic components and interactions that may contribute to the pathogenesis of MB, thereby providing insights into potential diagnostic biomarkers and therapeutic targets.

Methodology

Gene expression data

The gene expression profile data for gene chip GSE42656 was retrieved from the Gene Expression Omnibus (GEO) database [10]. This dataset, provided by the Neuroscience and Trauma team at Barts and the London School of Medicine and Dentistry, includes 5 cases from GBM patients as the experimental group and 8 normal samples as the control group.

Raw data preprocessing and screening and integration of differentially expressed genes

The raw gene chip data underwent quality control, standardization, and log2 conversion using Affymetrix Expression Console and the RMA algorithm. To identify differentially expressed genes, the Limma package (Linear Models for Microarray Data) in the R software was utilized. Subsequently, the integration of differentially expressed genes from the two gene chips was performed using RobustRankAggreg.

PPI network construction and hub genes identification

The Search Tool for the Retrieval of Interacting Genes database (Version 10.0, http://string-db.org) was used to predict potential interactions between gene candidates at the protein level [11]. A combined score of >0.4 (medium confidence score) was considered significant. Additionally, Cytoscape software (Version 3.4.0, http://www.cytoscape.org/) was utilized for constructing the PPI network. Degree ≥20 was set as the cutoff criterion. The Molecular Complex Detection (MCODE) app was used to analyze PPI network modules [12], and MCODE scores >3 and the number of nodes >5 were set as cutoff criteria with the default parameters (Degree cutoff ≥2, Node score cutoff ≥2, K-core ≥2 and Max depth = 100). Finally, CytoHubba, a Cytoscape plugin, was utilized to explore PPI network hub genes; it provides a user-friendly interface to explore important nodes in biological networks and computes using eleven methods, of which MCC has a better performance in the PPI network.

ROC curve analysis

To analyze the diagnostic efficiency of hub genes, ROC curve analysis was performed using the SRplot tool. Expression data of the identified hub genes from both the experimental (MB patients) and control (normal samples) groups were organized and uploaded to SRplot. This web-based platform was then used to generate ROC curves for each hub gene, plotting the true positive rate (sensitivity) against the false positive rate (1-specificity) at various thresholds. The area under the curve (AUC) was calculated for each gene, providing a measure of its ability to discriminate between GBM patients and normal samples, with higher AUC values indicating better diagnostic performance. Optimal threshold values were determined to maximize diagnostic accuracy, balancing sensitivity and specificity. This comprehensive assessment facilitated by SRplot enabled the evaluation of the diagnostic potential of the identified hub genes.

Results

Screening of top 50 differentially expressed genes

The provided image depicts a protein-protein interaction (PPI) network for the top 50 differentially expressed genes from the GSE42656 dataset, identified using the limma package and constructed via the STRING database. Each node represents a protein encoded by a DEG, with edges indicating predicted interactions based on various evidence types. Central nodes such as GABRG2, STXBP1, and DLG4 suggest these proteins play pivotal roles in the network, potentially influencing key biological processes. Other notable genes include MAP4, RBFOX1, and NPTN, which, while less connected, may still be crucial in specific pathways. The network also highlights clusters of interactions, such as those involving CLTA, PRKCE, and AASS, indicating potential pathways or complexes.

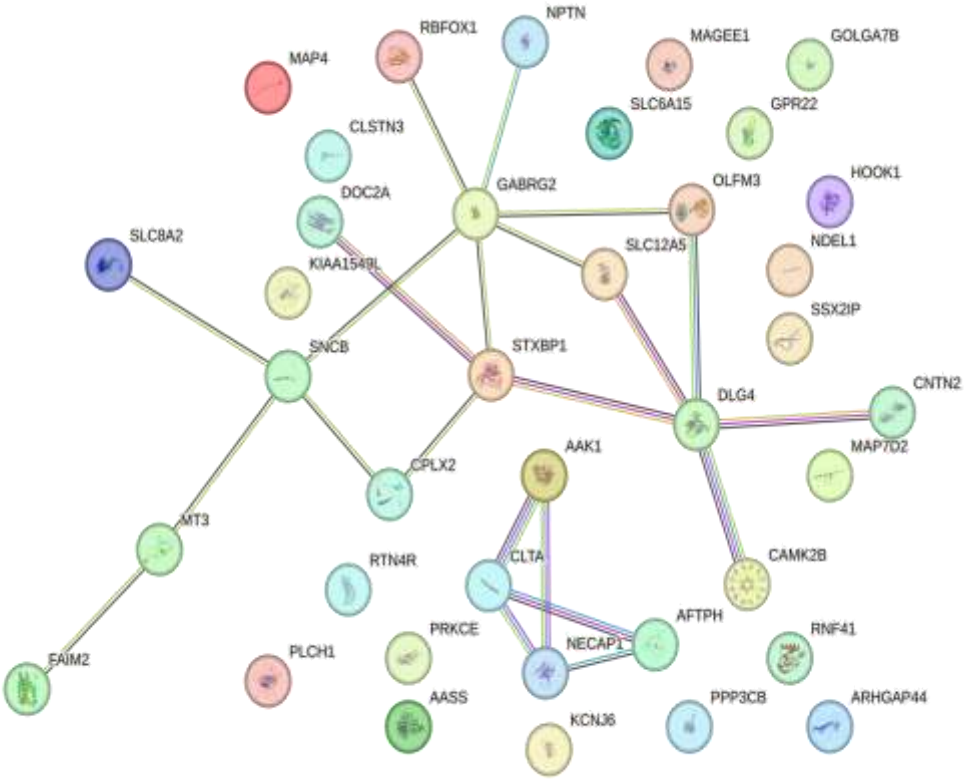

Figure 1: A PPI network of the top 50 differentially expressed genes in GSE42656 dataset

Identification of hub genes from the GSE42656 dataset

Following the construction of the PPI network for the top 50 differentially expressed genes from the GSE42656 dataset, the degree method in Cytoscape software was used to identify hub genes. This analysis revealed AASS, CAMKB2, MAP4, and SLC12A5 as key hub genes, indicating their central roles in the network due to their high number of connections. AASS is involved in lysine catabolism, highlighting its importance in metabolic regulation [13]. CAMKB2, a crucial player in calcium signaling pathways, suggests its role in neuronal function and plasticity [14]. MAP4 is essential for maintaining cellular structure and intracellular transport, emphasizing its significance in cell integrity [15]. SLC12A5, involved in maintaining chloride homeostasis in neurons, underscores its role in neuronal excitability and signaling [16]. Identifying these hub genes suggests they are central regulators in the studied condition, providing potential targets for further investigation and therapeutic intervention.

Expression of hub genes among other differentially expressed genes

The heatmap in Figure 2 displays the expression levels of the identified hub genes among the top 50 differentially expressed genes from the GSE42656 dataset, using a color scale from red (high expression) to blue (low expression). Notably, AASS shows high expression, indicating its upregulation in the MB samples, while MAP4 and SLC12A5 exhibit low expression, suggesting downregulation I MB. CAMKB2 appears multiple times with varying expression levels, highlighting its complex regulatory dynamics. Hierarchical clustering reveals groups of genes with similar expression patterns, suggesting potential co-regulation or involvement in related pathways. Genes like PLCH1, NDEL1, and CELF4 have expression levels closer to the baseline, indicating more stable expression. In contrast, genes such as GABRG2 and STXBP1 show significant downregulation in MB (Figure 2).

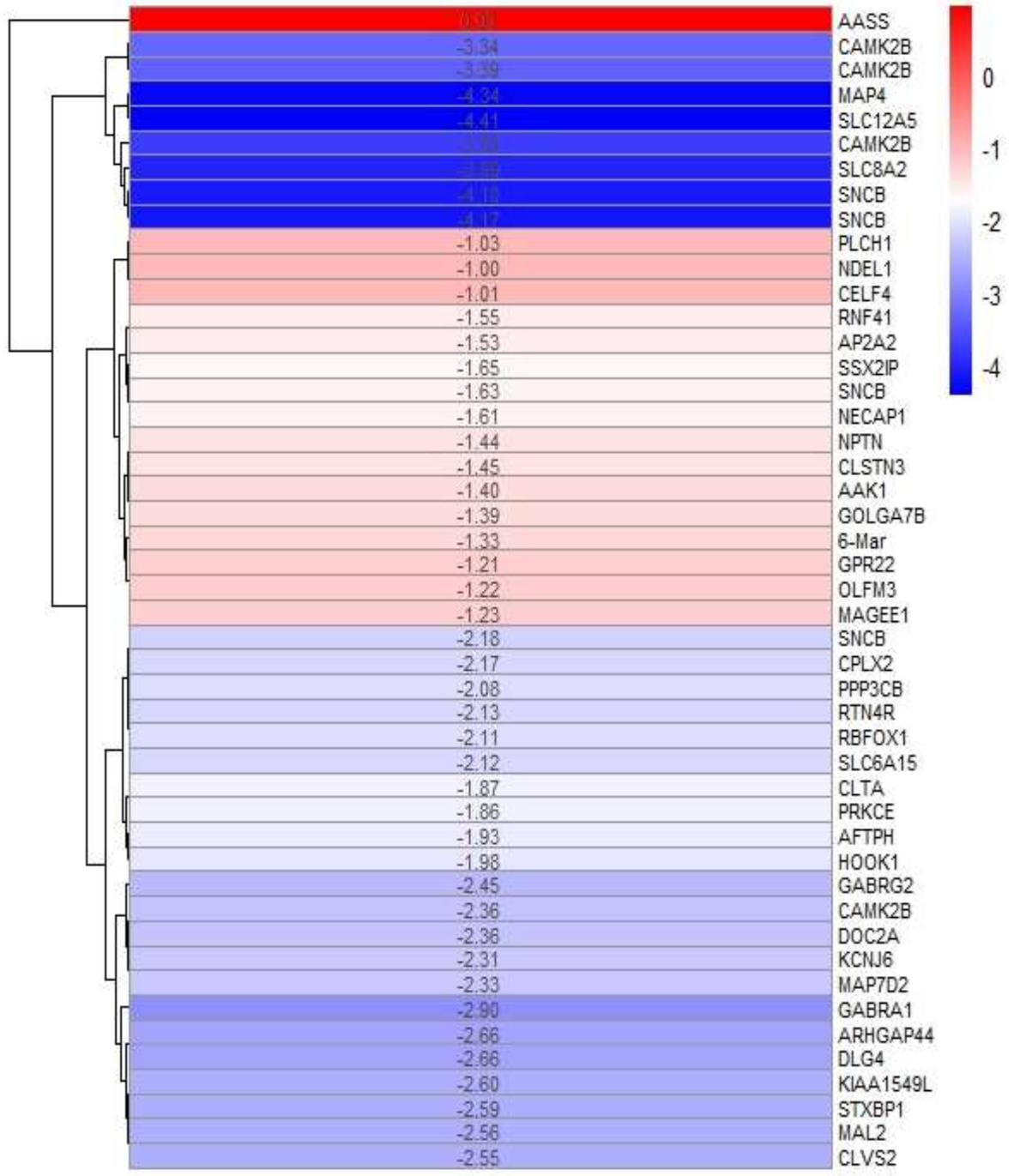

Figure 2: A heatmap depicting expression pattern of top 50 differentially expressed genes in MB.

Diagnostic efficacy of the hub genes

The ROC curves in Figure 3 in the provided image illustrate the diagnostic efficiency of the hub genes AASS, CAMKB2, MAP4, and SLC12A5 in MB). Each curve approaches the upper left corner, indicating high sensitivity and specificity. The area under the curve (AUC) values for all four genes is high, signifying excellent diagnostic accuracy. This suggests that these hub genes are highly effective biomarkers for distinguishing MB from non-MB cases. The strong diagnostic performance of these genes highlights their potential utility in clinical settings, offering reliable tools for MB diagnosis and potentially guiding treatment strategies.

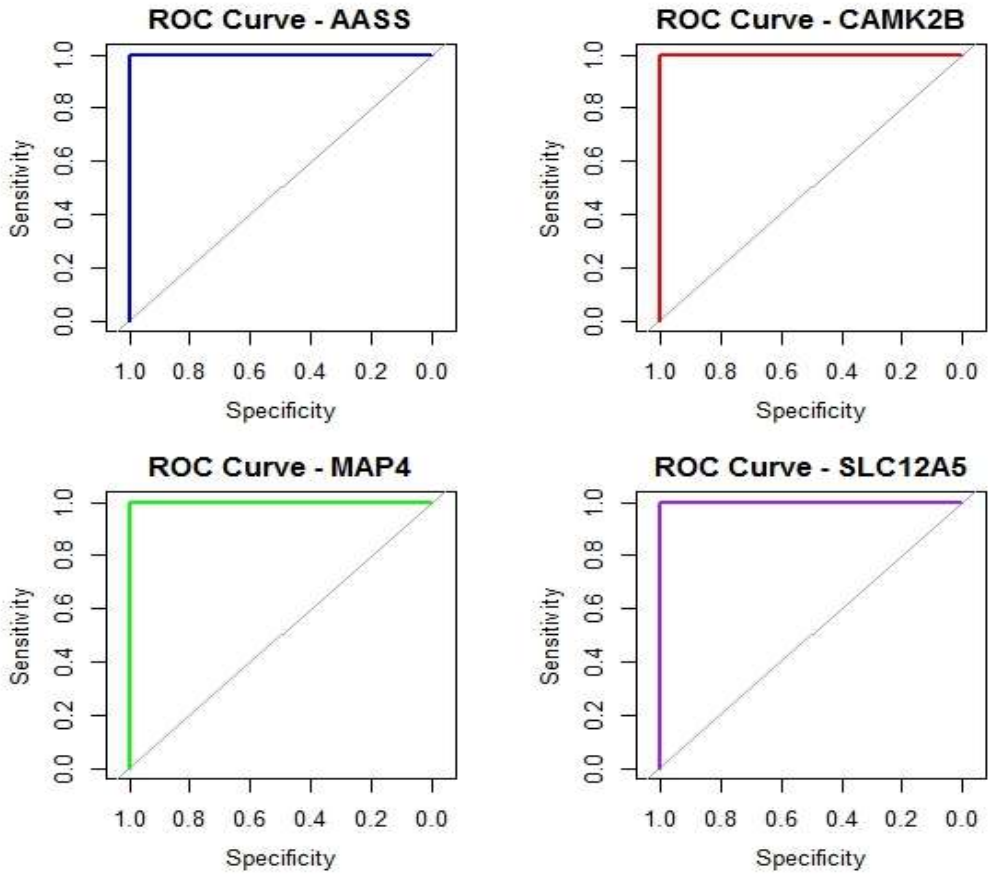

Figure 3: ROC curves of hub genes depicting their diagnostic abilities.

Discussion

Medulloblastoma (MB) is the most common malignant brain tumor in children, accounting for approximately 20% of all pediatric brain tumors [17]. Originating in the cerebellum or posterior fossa, MB is classified as a primitive neuroectodermal tumor [17]. Despite advancements in surgical techniques, radiation therapy, and chemotherapy, the prognosis for MB patients remains variable, with long-term survival rates ranging from 50% to 70%. Understanding the genetic and molecular underpinnings of MB is crucial for the development of targeted therapies and improved diagnostic tools.

The genetic basis of MB involves numerous genetic and epigenetic alterations that drive tumor development and progression [18]. Key genetic changes frequently observed in MB include mutations, copy number variations, and chromosomal rearrangements that affect critical signaling pathways such as the Sonic Hedgehog (SHH), WNT, and Notch pathways [19]. These alterations lead to uncontrolled cell proliferation, impaired differentiation, and resistance to apoptosis [19]. Recent genomic and transcriptomic profiling studies have identified several recurrent mutations in genes such as TP53, PTCH1, and MYC, which play pivotal roles in MB tumorigenesis [20]. Additionally, epigenetic modifications, including DNA methylation and histone modifications, contribute to the deregulation of gene expression in MB [20].

In this study, we utilized the GSE42656 dataset from the GEO database to identify differentially expressed genes (DEGs) and hub genes associated with MB. Using the limma package, we identified the top 50 DEGs and constructed a protein-protein interaction (PPI) network via the STRING database. Using Cytoscape software and the degree method, we identified AASS, CAMKB2, MAP4, and SLC12A5 as key hub genes, indicating their central roles in the network due to their high number of connections. AASS, involved in lysine catabolism, highlights its importance in metabolic regulation [21]. CAMKB2, a crucial player in calcium signaling pathways, suggests its role in neuronal function and plasticity [22]. MAP4 is essential for maintaining cellular structure and

intracellular transport, emphasizing its significance in cell integrity [15]. SLC12A5, involved in maintaining chloride homeostasis in neurons, underscores its role in neuronal excitability and signaling [23]. Identifying these hub genes suggests they are central regulators in the studied condition, providing potential targets for further investigation and therapeutic intervention.

Our findings were compared to reported studies to contextualize our results. Previous research has identified the SHH, WNT, and Notch pathways as critical in MB pathogenesis, aligning with our identification of hub genes that influence these pathways [24]. Studies have shown that TP53, PTCH1, and MYC mutations are recurrent in MB, which complements our identification of central nodes such as GABRG2 and STXBP1 that may interact with these mutated genes. The identified hub genes in our study, including AASS and CAMKB2, have not been extensively reported in MB studies, highlighting the novelty and potential significance of our findings.

The diagnostic efficacy of the identified hub genes was evaluated using ROC curve analysis. The high area under the curve (AUC) values for AASS, CAMKB2, MAP4, and SLC12A5 suggest excellent diagnostic accuracy, indicating that these hub genes are highly effective biomarkers for distinguishing MB from non-MB cases. This strong diagnostic performance underscores their potential utility in clinical settings, offering reliable tools for MB diagnosis and potentially guiding treatment strategies.

Conclusion

In conclusion, our study provides a comprehensive analysis of the genetic basis of MB, identifying key DEGs and hub genes with significant diagnostic potential. These findings contribute to the understanding of MB pathogenesis and highlight novel targets for future research and therapeutic development. Our results underscore the importance of integrating genomic data with bioinformatics tools to uncover critical molecular mechanisms in cancer.

Conflict of interest

None

Acknowledgement

None

References

[1] Liang L, Borlase S, Aiken C, Felton K, Hogg A, van Landeghem F, Klonisch T, Eisenstat DD and Werbowetski-Ogilvie TE. Primary Pediatric Brain Tumors of the Posterior Fossa: Part II A Comprehensive Overview of Medulloblastoma. Development of the Cerebellum from Molecular Aspects to Diseases 2023; 421-455.

[2] Atallah O, Al-Qaraghuli AK, Shaker NM, Akar NM, Abdulsada AM and Hoz SS. Pediatric Brain Tumors. In: editors. Pediatric Neurosurgery: In Multiple-Choice Questions. Springer; 2024. p. 163-176.

[3] Menyhárt O and Győrffy B. Principles of tumorigenesis and emerging molecular drivers of SHH-activated medulloblastomas. Annals of clinical and translational neurology 2019; 6: 990-1005.

[4] Lospinoso Severini L, Ghirga F, Bufalieri F, Quaglio D, Infante P and Di Marcotullio L. The SHH/GLI signaling pathway: a therapeutic target for medulloblastoma. Expert opinion on therapeutic targets 2020; 24: 1159-1181.

[5] Riobo-Del Galdo NA, Lara Montero A and Wertheimer EV. Role of hedgehog signaling in breast cancer: pathogenesis and therapeutics. Cells 2019; 8: 375.

[6] Sursal T, Ronecker JS, Dicpinigaitis AJ, Mohan AL, Tobias ME, Gandhi CD and Jhanwar-Uniyal M. Molecular stratification of medulloblastoma: clinical outcomes and therapeutic interventions. Anticancer research 2022; 42: 2225-2239.

[7] Szczepanek J, Skorupa M, Jarkiewicz-Tretyn J, Cybulski C and Tretyn A. Harnessing epigenetics for breast cancer therapy: The role of DNA methylation, histone modifications, and MicroRNA. International journal of molecular sciences 2023; 24: 7235.
[8] Güth U, Elfgen C, Montagna G and Schmid SM. Long-term survival and cure in distant metastatic breast cancer. Oncology 2019; 97: 82-93.
[9] Govindarajan R, Duraiyan J, Kaliyappan K and Palanisamy M. Microarray and its applications. J Pharm Bioallied Sci 2012; 4: 0975-7406.
[10] Clough E and Barrett T. The Gene Expression Omnibus Database. Methods Mol Biol 2016; 3578-3579_3575.
[11] Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, Jensen LJ and von Mering C. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 2021; 49: D605-D612.
[12] Bader GD and Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC bioinformatics 2003; 4: 1471-2105.
[13] Leandro J, Khamrui S, Suebsuwong C, Chen PJ, Secor C, Dodatko T, Yu C, Sanchez R, DeVita RJ, Houten SM and Lazarus MB. Characterization and structure of the human lysine-2-oxoglutarate reductase domain, a novel therapeutic target for treatment of glutaric aciduria type 1. Open Biol 2022; 12: 21.
[14] Fang L, Wu J, Lin Q and Willis WD. Calcium-calmodulin-dependent protein kinase II contributes to spinal cord central sensitization. J Neurosci 2002; 22: 4196-4204.
[15] Nabti I, Reddy BJN, Rezgui R, Wang W, Gross SP and Shubeita GT. The ubiquitous microtubule-associated protein 4 (MAP4) controls organelle distribution by regulating the activity of the kinesin motor. Proc Natl Acad Sci U S A 2022; 119: 3.
[16] Rahmati N, Hoebeek FE, Peter S and De Zeeuw CI. Chloride Homeostasis in Neurons With Special Emphasis on the Olivocerebellar System: Differential Roles for Transporters and Channels. Front Cell Neurosci 2018; 12:
[17] Thorbinson C and Kilday JP. Childhood Malignant Brain Tumors: Balancing the Bench and Bedside. Cancers 2021; 13:
[18] Herceg Z and Hainaut P. Genetic and epigenetic alterations as biomarkers for cancer detection, diagnosis and prognosis. Mol Oncol 2007; 1: 26-41.
[19] Menyhárt O and Győrffy B. Principles of tumorigenesis and emerging molecular drivers of SHH-activated medulloblastomas. Ann Clin Transl Neurol 2019; 6: 990-1005.
[20] Luo Z, Dong X, Yu J, Xia Y, Berry KP, Rao R, Xu L, Xue P, Chen T, Lin Y, Huang G, Li H, Zhou W and Lu QR. Genomic and Transcriptomic Analyses Reveals ZNF124 as a Critical Regulator in Highly Aggressive Medulloblastomas. Front Cell Dev Biol 2021; 9:
[21] Galili G, Tang G, Zhu X and Gakiere B. Lysine catabolism: a stress and development super-regulated metabolic pathway. Curr Opin Plant Biol 2001; 4: 261-266.
[22] Nicole O and Pacary E. CaMKIIβ in Neuronal Development and Plasticity: An Emerging Candidate in Brain Diseases. Int J Mol Sci 2020; 21:
[23] McMoneagle E, Zhou J, Zhang S, Huang W, Josiah SS, Ding K, Wang Y and Zhang J. Neuronal K+-Cl- cotransporter KCC2 as a promising drug target for epilepsy treatment. Acta Pharmacologica Sinica 2024; 45: 1-22.
[24] Amjad E, Asnaashari S, Jahanban-Esfahlan A and Sokouti B. The role of MAPK, notch and Wnt signaling pathways in papillary thyroid cancer: Evidence from a systematic review and meta-analyzing microarray datasets employing bioinformatics knowledge and literature. Biochemistry and Biophysics Reports 2024; 37: 101606.