Research Article

Unraveling the Role of Musashi Signaling Pathway in Medulloblastoma: Insights from Differential Expression and Co-expression Network Analysis

Sina Nasr Esfahani[1], Bahareh Mazrouei[1], Fariba Dehghanian[1*], Mohammad Kazemi[2,3*]

[1]*Department of Cell and Molecular Biology & Microbiology, Faculty of Biological Science and Technology, University of Isfahan, HezarJarib Street, Isfahan 81746-73441, Iran*
[2]*Department of Genetics and Molecular Biology, Faculty of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran*
[3]*Reproductive Sciences and Sexual Health Research Center, Isfahan University of Medical Sciences, Isfahan, Iran*

Received 19 January 2026 Accepted 1 June 2026

Abstract

Medulloblastoma (MB) is the most common malignant pediatric brain tumor, characterized by molecular heterogeneity and aggressiveness. The RNA-binding protein Musashi-1 (*MSI1*) has been identified as a key regulator of MB tumorigenesis, promoting cancer stem cell phenotypes and treatment resistance. *MSI1* research in MB also includes its potential as a prognostic and diagnostic biomarker. In the current study, the role of *MSI1* in MB was highlighted using differential expression and weighted gene co-expression network analysis. In this study, the role of *MSI1* in MB was explored through bioinformatics analysis of RNA-Seq data of 93 MB samples (GEO datasets: GSE143940, GSE158413, GSE189919) and 266 normal cerebellum samples (GTEx). The DESeq2 package was used for differential expression analysis, and WGCNA was used to discover *MSI1*-associated co-expression networks. A total of 11,055 significantly differentially expressed genes were identified in MB, comprising 5,925 upregulated and 5,130 downregulated genes. WGCNA positioned *MSI1* in the MB-specific, non-preserved darkgrey module with hub genes *ANO2* (Log2FC: 4.80, p = 0) and *CACNA1F* (Log2FC: 2.18, p = 9.59E-65), implicating them in *MSI1*-driven oncogenic mechanisms. *MSI2*, on the other hand, was positioned in a preserved module, suggesting a more general function. These findings highlight MSI1's specificity for MB and its potential as a biomarker and therapeutic target. Upregulation of *ANO2* and *CACNA1F*, along with the role of *MSI1* in tumor aggressiveness, underscores the need for additional functional exploration to validate their therapeutic relevance. According to the systems biology approach, these two genes have a high correlation with *MSI1*; the gene selection criteria also included a gene that is a core gene (based on the widely used MCC, MNC, Degree, and EPC methods in CytoHubba analysis) and has significantly different expression depending on the DEG analysis threshold.

***Keywords*:** Medulloblastoma (MB), Musashi-1 (*MSI1*), Weighted Gene Co-expression Network Analysis (WGCNA), *ANO2*, *CACNA1F*

Introduction

Medulloblastoma (MB) is the most prevalent central nervous system malignant tumor in children (Baroni et al., 2021). MB accounts for 15% of all childhood brain tumors (Lannering et al., 2009). MB is a malignant embryonal tumor that occurs predominantly within the cerebellum or posterior fossa, characterized by its features, which include the rate of growth, tendency for dissemination into the leptomeninges, and profound neurological effects. Treatment of MB patients is typically a multidisciplinary approach that includes maximal resection of the tumor, cytotoxic chemotherapy, and craniospinal irradiation for individuals over three years of age (Gajjar and Robinson, 2014; Northcott et al., 2019). MB was recently divided into 12 molecular subtypes, including WNT, Sonic Hedgehog (SHH), Group 3, and Group 4, each with distinct genetic and epigenetic characteristics that underlie prognosis and treatment response. Group 4 is the most common MB subtype but the least biologically characterized (Cavalli et al., 2017; Louis et al., 2021). The Musashi-1 (*MSI1*) pathway has recently garnered significant interest as a central mediator of MB tumorigenesis and tumor growth (Kanemura et al., 2001). *MSI1* is a highly conserved RNA-binding protein mediating translation by binding to (G/A) U1-3(AGU) motifs within the 3'-UTR of target mRNAs. It is required for sustaining

* Corresponding authors' e-mail addresses:
fa.dehghanian@sci.ui.ac.ir
m_kazemi@med.mui.ac.ir

Licensed as Open Access CC-BY

neural stem cell self-renewal and preserving against differentiation through post-transcriptional modification of target messenger RNAs (mRNAs) (Kanemura et al., 2001; Richter et al., 1990). Through its interaction with the 3' untranslated regions (3' UTRs) of target mRNAs, *MSI1* regulates the expression of genes essential to oncogenic pathways, such as Notch, Hedgehog, and PI3K/AKT (das Chagas et al., 2020). It induces a stem-like state in tumor cells (Cambuli et al., 2015; das Chagas et al., 2020; Nakamura et al., 1994; Sakakibara et al., 1996). In MB, *MSI1* is also overexpressed in most SHH and Group 3 subgroups, thereby enhancing tumor cell proliferation, survival, and drug resistance (das Chagas et al., 2020; Kong et al., 2008). This stemness contributes to tumor initiation, metastasis, and recurrence; thus, *MSI1* is considered an accessible therapeutic target (Forouzanfar et al., 2020; Kharas and Lengner, 2017; Kudinov et al., 2017). *MSI1* regulates hundreds of targets through post-transcriptional mechanisms, influencing numerous cancer pathways, including NUMB/Notch, PTEN/mTOR, TGF/SMAD3, MYC, cMET, and others (Uren et al., 2015). *MSI1* regulates these processes when it interacts with the 3' UTR of target mRNAs, leading to changes in mRNA stability and/or translation (Kudinov et al., 2017).

Experimental suppression of *MSI1* by small interfering RNAs (siRNAs) or small-molecule inhibitors has demonstrated potential to suppress MB growth and induce sensitivity to traditional therapy in preclinical models (Zhao et al., 2022). Elevated *MSI1* expression in tumor tissue and peripheral blood correlates with advanced disease stages and poor clinical outcomes, suggesting its potential for non-invasive monitoring (Zheng et al., 2024). Lastly, knowledge of the interactions between *MSI1* and other molecular actors has the potential to enhance early diagnosis, individualize treatment, and improve patient survival in this aggressive disease.

In the current study, we highlighted the role of *MSI1* in MB. We conducted our work using bioinformatic data from MB patients and combined multiple datasets to develop meta-analytic data for sample size estimation. We used R packages such as DESeq2 and WGCNA to get expression patterns for *MSI1* in MB. Our findings showed that *MSI1* plays a significant role in MB aggressiveness and growth across various samples.

Materials and Methods

Data Acquisition

MB RNA-Seq data were downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) by using the National Center for Biotechnology Information (NCBI) (https://www.ncbi.nlm.nih.gov/). Three datasets were used: GSE143940 (52 samples), GSE158413 (14 samples), and GSE189919 (27 samples), totaling 93 MB samples. Additionally, normal cerebellum RNA-Seq samples (266 samples) were obtained from the Genotype-Tissue Expression (GTEx) database (https://gtexportal.org/home/) (gene_reads_v10_brain_cerebellar_hemisphere) to be used as a control group for comparative analyses. Datasets were downloaded in raw count formats to ensure consistency in downstream processing.

Data Preprocessing and Batch Effect Correction

To eliminate potential batch effects from merging multiple GEO datasets, the Surrogate Variable Analysis (sva) package in R (version 4.5) was used. The ComBat_seq function from the sva package (version 3.57.0) was used to adjust for batch effects, minimizing technical differences between datasets while preserving biological variability (Liu and Markatou, 2016).

Gene Filtering

In particular, genes with raw count values lower than 10 in more than half of the samples were discarded. Genes with low expression were filtered out to minimize unwanted noise and fluctuations, thereby increasing the reliability of downstream analyses. Furthermore, we were strict and retained only messenger RNAs (mRNAs) and long non-coding RNAs (lncRNAs) for the following analysis.

Outlier Detection and Removal

To ensure data quality, outlier samples were identified and removed based on principal component analysis (PCA) and hierarchical clustering. Samples exhibiting extreme deviations in gene expression profiles, as determined by PCA plots and clustering dendrograms, were excluded to prevent skewing of subsequent analyses (Karamizadeh et al., 2013; Kurita, 2021).

Differential Expression Analysis

The goal of differential gene expression analysis was to identify significant genes in the MB relative to the normal cerebellum. The DESeq2 (version 1.46.0) package in R (version 4.5) was used for this analysis. Raw count data were normalized using the variance-stabilizing transformation (VST) in DESeq2, and differentially expressed genes were identified using negative binomial generalized linear

models. Genes with an adjusted p-value < 0.05 (false discovery rate, FDR) and |log2 fold change| ≥ 0.58 were considered to be significantly differentially expressed (Li et al., 2022; Liu et al., 2021).

Weighted Gene Co-expression Network Analysis (WGCNA)

WGCNA was performed on MB and normal cerebellum samples using the WGCNA package (version 1.73) in R (version 4.5). A signed hybrid network (separate for normal and MB samples) was built, and its co-expression network was extracted. Soft-thresholding power was set based on the scale-free topology criterion ($R^2 > 0.80$) to convert gene expression matrices into adjacency matrices. Gene modules were detected by hierarchical clustering and dynamic tree cutting, with a minimum module size of 30 genes. Module eigengenes (MEs) were computed through the moduleEigengenes function. Modules whose MEs were highly correlated (correlation coefficient ≥ 0.75, corresponding to a dissimilarity threshold of 0.25, indicating 75% similarity) were merged to reduce redundancy (Langfelder and Horvath, 2008).

Module Preservation Analysis

To assess the robustness of co-expression modules across MB and normal cerebellum samples, module preservation analysis was conducted using the WGCNA package. The module preservation function was used to compute preservation statistics, including Zsummary and medianRank, to assess whether modules identified in the MB dataset were preserved in the normal cerebellum dataset. A Z-score > 10 was used as a threshold to indicate strong preservation, while scores between 2 and 10 indicated moderate preservation, and scores less than 2 were considered non-preserved modules (Nguyen and Zeng, 2025).

Network Visualization

Co-expression networks derived from WGCNA were visualized using Cytoscape (version 3.10.3). A correlation threshold (correlation coefficient ≥ 0.01) was applied to filter weak edges, ensuring that only strong co-expression relationships were visualized. The cytoHubba plugin was used to identify hub genes using MCC (Maximal Clique Centrality), MNC (Maximum Neighborhood Component), Degree, and EPC (Edge Percolated Component) in the selected module (Ma et al., 2021). The workflow of our analysis is demonstrated in Figure 1.



Figure 1. A WGCNA workflow that integrates gene expression and metadata from MB RNA-seq data in the GEO database.

Results

Identification and Filtering of Gene Expression Profiles in Medulloblastoma and Normal Cerebellum Samples

The merging of RNA-Seq data from 93 MB samples and 266 normal cerebellum samples identified a total of 22,969 common genes, with 4,050 long non-coding RNAs (lncRNAs) and 18,919 protein-coding genes. The dataset was subsequently further filtered to remove genes with raw counts of less than 10 in at least half of the samples; thus, 16,144 genes (including 1,148 lncRNAs and 14,996 protein-coding genes) were retained for subsequent analyses.

Principal Component Analysis and Hierarchical Clustering for Outlier Detection and Data Normalization in Medulloblastoma and Normal Cerebellum Samples

To study sample variability and identify outliers, PCA was performed. The initial PCA shows that MB and normal cerebellum samples were distinguished across the first principal component (PC1), which explained most of the variance in the system (Supplementary Figure 1A). However, a total of three MB and three normal cerebellum samples registered extreme values along the second principal component (PC2), hence indicating possible outliers (Supplementary Figure 1A). After the outliers were excluded, 90 MB and 263 normal cerebellum samples remained for DEG analyses. After outliers were removed, the samples clustered better, and PC1 distinguished deviant samples from normal ones (Supplementary Figure 1B). Moreover, using the variance stabilizing transformation (VST) method in DESeq2 to normalize the raw count data reduced technical variability, as illustrated by distinct distribution patterns of expression values before and after normalization in the box plots (Supplementary Figure 1C, D).

Second, hierarchical clustering identified four MB and six normal cerebellum samples as outliers due to unusual clustering patterns in the dendrograms (Supplementary Figure 2A). After their removal, 86 MB and 257 normal cerebellum samples were available for WGCNA analyses (Supplementary Figure 2B).

Identification and Characterization of Differentially Expressed Genes in Medulloblastoma Compared to Normal Cerebellum Using Gene Expression Profiling

The characterization of differentially expressed genes in MB versus normal cerebellum samples was performed to identify genes significantly deregulated in MB. With the cutoff values of an adjusted p-value < 0.05 and |log2 fold change| (|log2FC|) ≥ 0.58, 11,055 genes were identified as differentially expressed. Among these, 5,925 genes were upregulated, and 5,130 genes were downregulated in MB. A volcano plot was used to display the distribution of DEGs, highlighting the significantly expressed genes at cutoff values (Figure 2A).

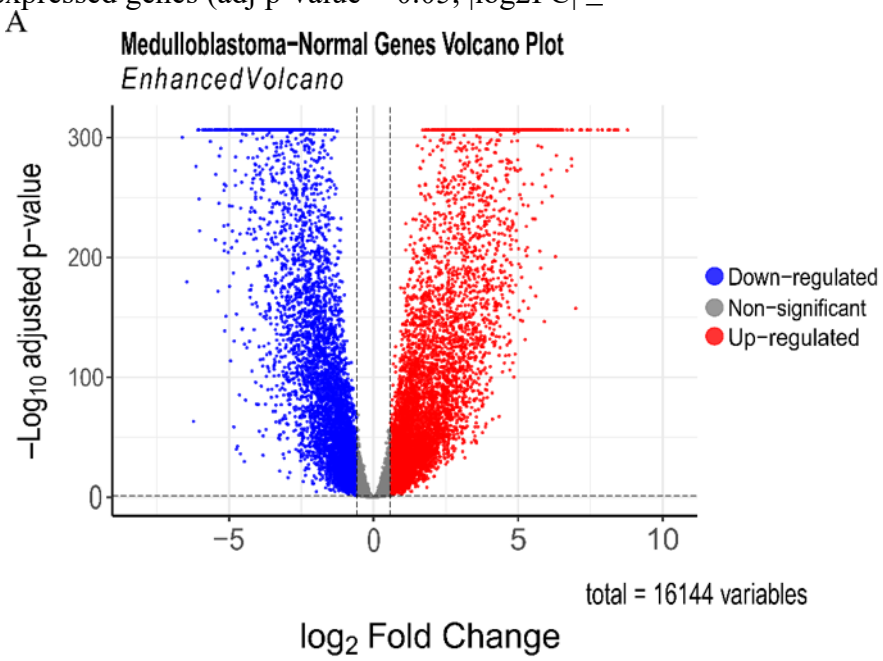
Construction and Analysis of Co-Expression Networks in Medulloblastoma and Normal Cerebellum Using WGCNA

WGCNA was used to construct co-expression networks from MB and normal cerebellum samples. Scale independence and mean connectivity analyses were used to select the optimal soft-thresholding powers, resulting in powers of 5 for MB samples and 4 for normal cerebellum samples, ensuring a scale-free topology ($R^2 > 0.80$) (Supplementary Figures 3A, 3 B). Initial WGCNA identified 69 co-expression modules in both MB and normal samples. After merging modules with highly correlated module eigengenes (threshold: 0.25), 43 modules were retained for MB and 36 for normal cerebellum samples (Figure 3A, B).

Module Preservation Analysis of Co-Expression Networks Reveals Differential Preservation of *MSI2*- and *MSI1*-Associated Modules in Medulloblastoma and Normal Cerebellum

Module-preservation analysis was performed to determine whether MB modules are preserved in normal cerebellum samples. The blue module (2863 genes), which had the Musashi-2 (*MSI2*) gene, was strongly preserved (Zsummary > 10) while the darkgrey module (214 genes), which had the *MSI1* gene, was not preserved (Zsummary < 2) (Figure 4A). These results suggest the blue module is a strong co-expression network preserved across conditions,

while the darkgrey module is MB-specific. There were 2,231 common genes between the blue module and significantly differentially expressed genes (adj p-value < 0.05, |log2FC| ≥ 0.58). Common genes between the darkgrey module and the significantly differentially expressed genes were 160 (Figure 2B, C).



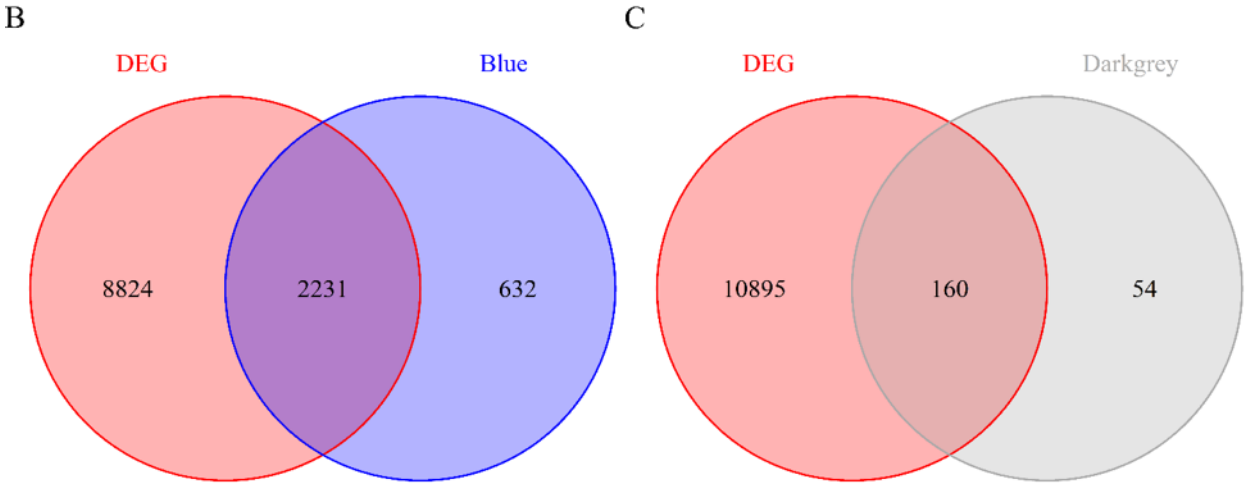


Figure 2. Analysis of differential gene expression, along with visualization and comparison of gene sets through overlapping genes from *MSI*-related modules and DEG analysis. (A) Volcano Plot. Every point symbolizes a gene. The x-axis represents log_2FC, while the y-axis represents $-log_{10}(padj)$. The dashed horizontal line marks the significance threshold (padj = 0.05), while the dashed vertical lines indicate the log_2FC cutoff (±0.58). Genes located in the upper-left and upper-right quadrants (exceeding both thresholds) are substantially downregulated or upregulated, respectively. RNA-seq data is normalized using DESeq2. Comparison of gene expression between two states (MB vs. control). Genes exhibiting an absolute log_2 fold change (log_2FC) ≥ 0.58 and an adjusted p-value (padj) < 0.05 are regarded as significantly differentially expressed (red: upregulated, blue: downregulated, gray: not significant). (B, C) Venn diagrams**.** Intersection of DEG lists from both comparisons (the *MSI* gene-related modules of WGCNA and DEG analysis). Every red circle indicates a collection of DEGs. Darkgrey module associated with *MSI1* and notable DEGs, and blue module related to *MSI2* and significant DEGs. The blue module had 2,231 shared genes (632 unique to the module, 8,824 unique DEGs), whereas the darkgrey module had 160 shared genes (54 unique to the module, 10,895 unique DEGs).

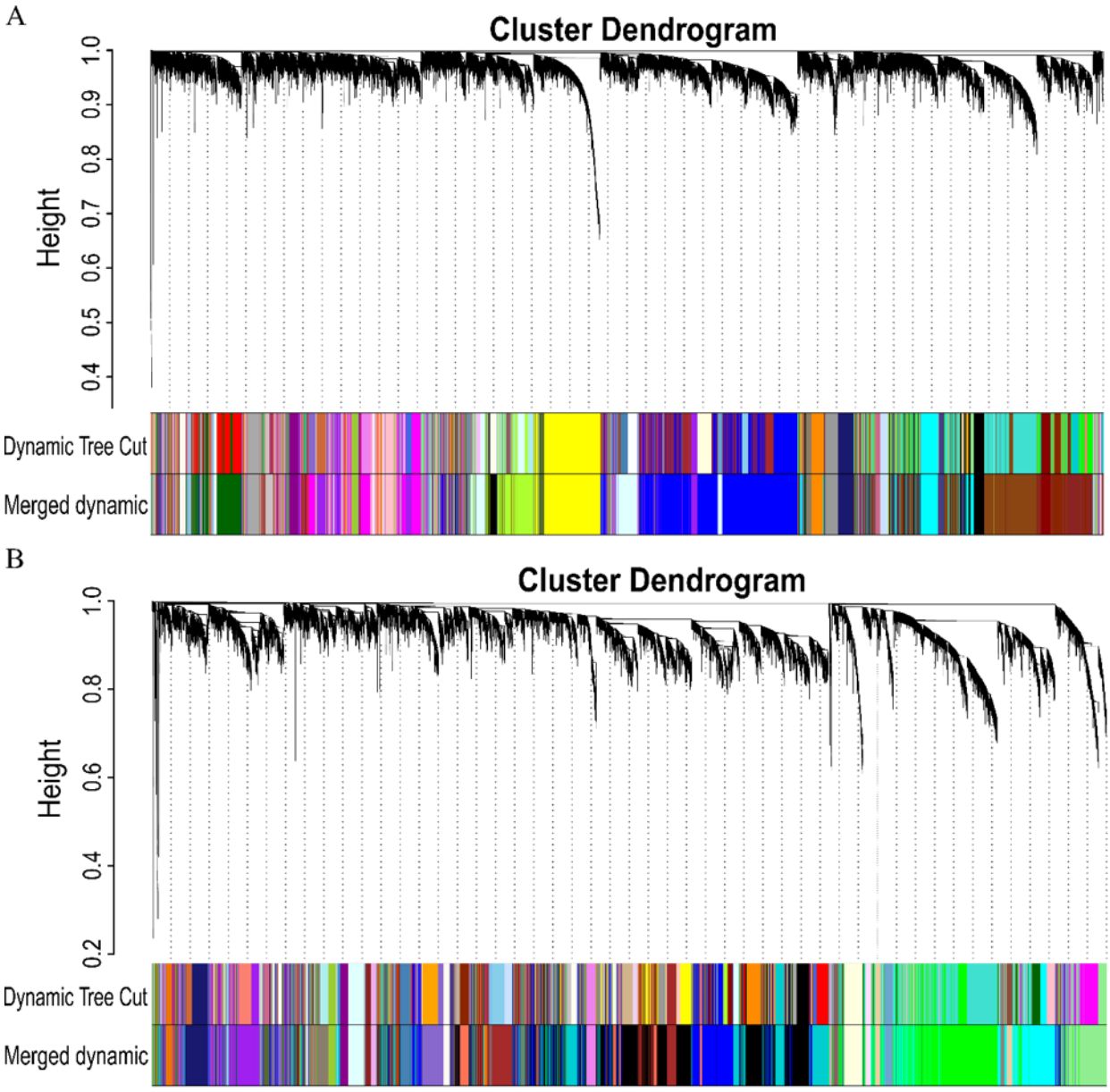


Figure 3. Cluster dendrogram for gene co-expression module identification and merging. (A) Hierarchical clustering of MB modules and module colors before and after merging (threshold: 0.25), resulting in 43 modules. (B) Dendrogram of normal cerebellum modules and module colors before and after merging, resulting in 36 modules. A dendrogram is constructed using the topological overlap matrix (TOM) dissimilarity (1 – TOM) to cluster genes based on their co-expression patterns. The height axis represents the dissimilarity measure. Each leaf corresponds to a single gene. The Dynamic Tree Cut algorithm is applied to identify initial modules, which are shown as colored bands below the dendrogram. Module eigengene clustering and merging to merge highly similar modules, module eigengenes (MEs; first principal component of each module) are calculated. Modules whose eigengenes cluster below the threshold (threshold: 0.25) are merged into a single module. The merged module assignments are shown in the final color bar below the gene dendrogram, with previously distinct but similar modules receiving the same color. This process reduces module redundancy and improves biological interpretability.

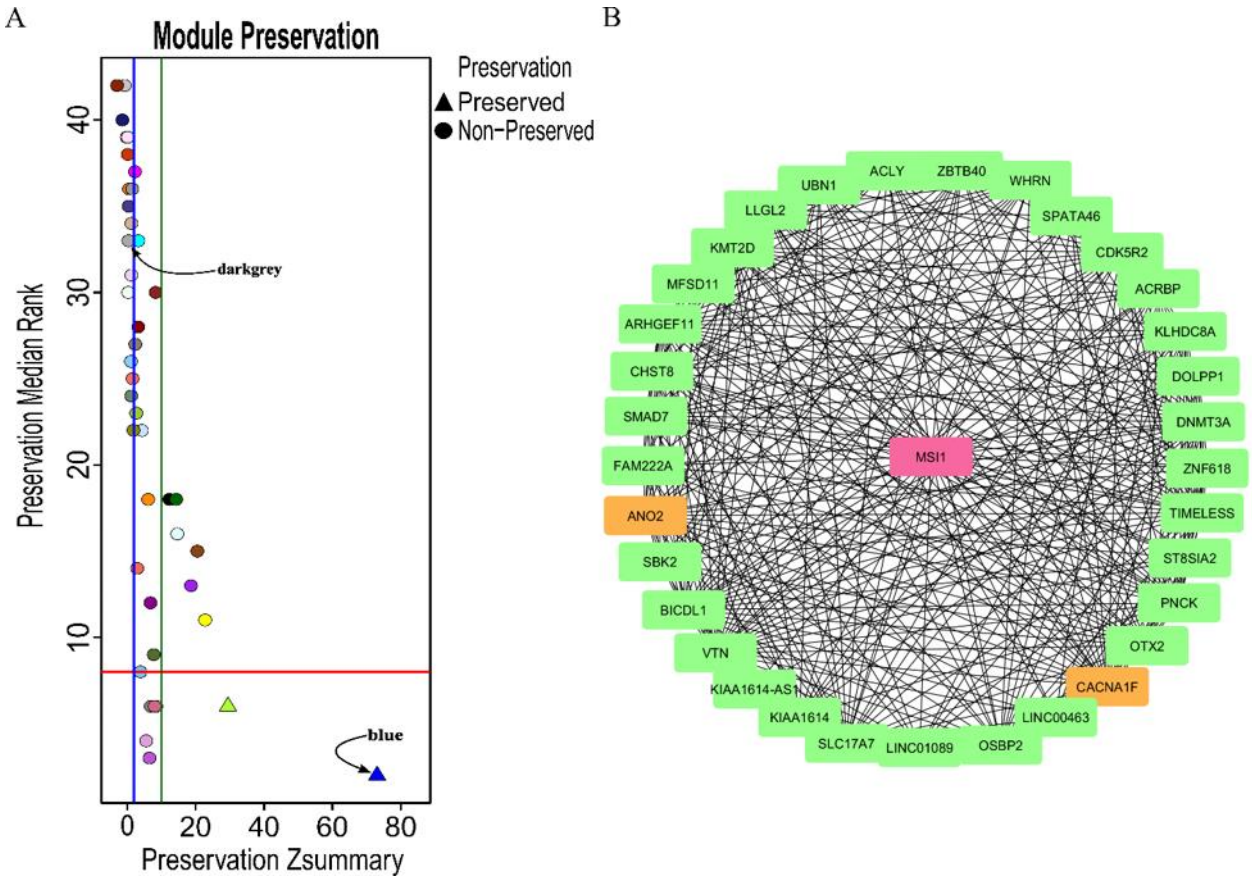


Figure 4. Module preservation analysis and Cytoscape-based network reconstruction of darkgrey module. (A) Module preservation analysis. Every point in this diagram signifies a module. Modules represented by triangles are preserved modules, while modules depicted by circles are non-preserved modules. The blue and green dotted lines represent the Zsummary thresholds of 2 and 10, while the red dotted line marks the MedianRank threshold of 8. Modules exhibiting a Zsummay greater than 10 and a MedianRank lower than eight were deemed preserved. Plot of Z summary and Median rank scores for MB modules in normal cerebellum samples, with the blue module (*MSI2*) showing strong preservation and the darkgrey module (*MSI1*) showing no preservation. (B) Cytoscape network reconstruction. A representative co-expression network for a module of interest (darkgrey module) is visualized using Cytoscape, with nodes representing genes and edges indicating co-expression correlations (correlation coefficient ≥ 0.01). The hub genes *ANO2* and *CACNA1F* are highlighted, indicating strong connectivity with *MSI1*.

Identification of *ANO2* and *CACNA1F* as Key Hub Genes in the Medulloblastoma-Specific Darkgrey Module through Co-Expression Network Analysis and Differential Expression Profiling

The darkgrey module, which contained *MSI1*, was selected for further analysis because it was specific to MB and related to the Musashi pathway. Cytoscape was used to visualize the co-expression network of the darkgrey module. Two genes, *ANO2* and *CACNA1F*, were identified as common hub genes in all four metrics and exhibited strong co-expression with *MSI1*, indicating high connectivity and a crucial role in the darkgrey module. Furthermore, differential expression analysis revealed that *ANO2* exhibited a log2 fold change (log2FC) of 4.80 with an adjusted p-value of 0, while *CACNA1F* showed a log2FC of 2.18 with an adjusted p-value of $9.59 \times 10^{(-65)}$. These genes were illustrated within the darkgrey network, showing their central roles in the MB-specific network (Figure 4B). On the other hand, *MSI2* was identified as a member of the preserved blue module in Musashi pathway analysis, indicating preserved function in MB and the normal cerebellum. Alternatively, *MSI1*, a member of the non-preserved darkgrey module, was specific to MB. The hub genes *ANO2* and *CACNA1F*, members of the darkgrey module, showed significant co-expression with *MSI1* and may therefore be involved in the Musashi pathway and MB progression.

Discussion

The findings from this study emphasize the tumorigenic role of *MSI1* in MB, particularly given its MB-specific molecular profile, as observed in the WGCNA network. This finding that the identification of *MSI1* in the non-preserved darkgrey module is related explicitly to MB underscores its significance as both a biomarker and a target for therapy for this deadly childhood brain tumor. On the other hand, *MSI2* in the preserved blue module, due to its grouping in the preserved part, suggests a wider preserved function in both the MB and normal cerebellum, which may reflect a less specific role of the gene in MB tumorigenesis. These findings align with previous studies that report elevated *MSI1* expression in MB, particularly within the SHH and Group 3 subgroups, and suggest a role for *MSI1* in tumor cell growth, survival, and treatment resistance (das Chagas et al., 2020; Kong et al., 2008). Through post-transcriptional processes, *MSI1* regulates hundreds of targets, including numerous cancer pathways, such as NUMB/Notch, PTEN/mTOR, TGF/SMAD3, MYC, cMET, and many others (Uren et al., 2015).

Differential expression analysis revealed marked molecular dysregulation in MB. Two of the hub genes, *ANO2* and *CACNA1F*, in the darkgrey module are upregulated (log2FC = 4.80 and 2.18) and are strongly co-expressed with *MSI1*, suggesting their relevance to *MSI1*-driven oncogenic pathways. *ANO2*, a calcium-activated chloride channel, and *CACNA1F*, a voltage-gated calcium channel, may play roles in regulating ionic signaling, which is recognized as crucial for the growth and survival of cancer cells and their interactions with stromal cells. The absence of the darkgrey module in normal cerebellum samples highlights the specificity of *MSI1*-associated networks to MB. This suggests that *MSI1*, along with its co-expressed genes such as *ANO2* and *CACNA1F*, may contribute to molecular mechanisms unique to MB, particularly within Group 4—the most prevalent yet least characterized subtype (Cavalli et al., 2017; Louis et al., 2021). Although the role of *MSI1* in MB has already been studied, the main aim of this study is to investigate the role of the Musashi signaling pathway (of which *MSI1* is a central component) in MB. The therapeutic potential of targeting *MSI1* appears quite significant. Preclinical findings suggest that suppressing *MSI1*, either through siRNAs or small-molecule inhibitors, can reduce tumor growth and enhance the response to conventional therapies (Zhao et al., 2022). The upregulation of *ANO2* and *CACNA1F* in MB suggests that these genes could serve as additional therapeutic targets or, at the very least, useful biomarkers for monitoring disease progression. Notably, increased *MSI1* expression is associated with later disease stages and poorer clinical outcomes.

Limitations of the study include reliance on publicly available datasets, which may introduce variability in sample collection or sequencing protocols. While batch-effect correction using ComBat_seq has reduced some technical variability, residual confounding may remain. Additionally, the study's focus on RNA-Seq data limits its ability to assess protein-level interactions, which are relevant to confirming the post-transcriptional regulatory role of *MSI1*. Future studies should encompass proteomic investigations and functional assays to confirm the roles of *ANO2* and *CACNA1F* in *MSI1*-mediated MB pathogenesis. Furthermore, given that appropriate experimental tests can improve the validity of studies, it is recommended that future studies conduct experiments to confirm the roles and activities of known genes and markers using bioinformatics analysis.

Conclusions

In conclusion, this study highlights the critical role of *MSI1* in MB through its specific expression in the non-preserved darkgrey module and its co-expression with hub genes *ANO2* and *CACNA1F*. These results enhance our comprehension of the molecular complexity of MB and highlight the promise of *MSI1* as both a biomarker and a therapeutic target. By integrating advanced bioinformatics with experimental validation, future research can build on these insights to develop personalized therapies, ultimately improving outcomes for patients with MB.

Author contributions

SNE: Writing original draft, writing review and editing; Methodology, Investigation, Formal analysis, Data curation; BM: Writing original draft, writing, Methodology, review and editing; FD: Writing review and editing, Methodology, Investigation; MK: Writing review and editing.

Conflict of Interests

The authors declare that they have no conflict of interest.

References

Baroni M., Guardia G. D., Lei X., Kosti A., Qiao M., Landry T., et al. (2021) The RNA-Binding protein

musashi1 regulates a network of cell cycle genes in group 4 medulloblastoma. Cells 11:56.

Cambuli F., Correa B., Rezza A., Burns S., Qiao M., Uren P., et al. (2015) A mouse model of targeted Musashi1 expression in whole intestinal epithelium suggests regulatory roles in cell cycle and stemness. Stem cells 33:3621–3634.

Cavalli F. M., Remke M., Rampasek L., Peacock J., Shih D. J., Luu B., et al. (2017) Intertumoral heterogeneity within medulloblastoma subgroups. Cancer cell 31:737–754. e736.

das Chagas P. F., Baroni M., Brassesco M. S. and Tone L. G. (2020) Interplay between the RNA binding-protein Musashi and developmental signaling pathways. The Journal of Gene Medicine 22:e3136.

Forouzanfar M., Lachinani L., Dormiani K., Nasr-Esfahani M. H., Gure A. O. and Ghaedi K. (2020) Intracellular functions of RNA-binding protein, Musashi1, in stem and cancer cells. Stem Cell Research & Therapy 11:193.

Gajjar A. J. and Robinson G. W. (2014) Medulloblastoma—translating discoveries from the bench to the bedside. Nature reviews Clinical oncology 11:714–722.

Kanemura Y., Yamasaki M., Mori K., Fujikawa H., Hayashi H., Nakano A., et al. (2001) Musashi1, an evolutionarily conserved neural RNA-binding protein, is a versatile marker of human glioma cells in determining their cellular origin, malignancy, and proliferative activity. Differentiation 68:141–152.

Karamizadeh S., Abdullah S. M., Manaf A. A., Zamani M. and Hooman A. (2013) An overview of principal component analysis. Journal of signal and information processing 4:173.

Kharas M. G. and Lengner C. J. (2017) Stem cells, cancer, and MUSASHI in blood and guts. Trends in cancer 3:347–356.

Kong D.-S., Kim M. H., Park W.-Y., Suh Y.-L., Lee J.-I., Park K., et al. (2008) The progression of gliomas is associated with cancer stem cell phenotype. Oncology reports 19:639–643.

Kudinov A. E., Karanicolas J., Golemis E. A. and Boumber Y. (2017) Musashi RNA-binding proteins as cancer drivers and novel therapeutic targets. Clinical Cancer Research 23:2143–2153.

Kurita T. 2021. Principal component analysis (PCA). *In* Computer vision: a reference guide. Springer. 1013–1016.

Langfelder P. and Horvath S. (2008) WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics 9:559.

Lannering B., Sandström P. E., Holm S., Lundgren J., Pfeifer S., Samuelsson U., et al. (2009) Classification, incidence and survival analyses of children with CNS tumours diagnosed in Sweden 1984–2005. Acta paediatrica 98:1620–1627.

Li D., Zand M. S., Dye T. D., Goniewicz M. L., Rahman I. and Xie Z. (2022) An evaluation of RNA-seq differential analysis methods. PloS one 17:e0264246.

Liu Q. and Markatou M. (2016) Evaluation of methods in removing batch effects on RNA-seq data. Infect Dis Transl Med 2:3–9.

Liu S., Wang Z., Zhu R., Wang F., Cheng Y. and Liu Y. (2021) Three differential expression analysis methods for RNA sequencing: limma, EdgeR, DESeq2. JoVE (Journal of Visualized Experiments):e62528.

Louis D. N., Perry A., Wesseling P., Brat D. J., Cree I. A., Figarella-Branger D., et al. (2021) The 2021 WHO classification of tumors of the central nervous system: a summary. Neuro-oncology 23:1231–1251.

Ma H., He Z., Chen J., Zhang X. and Song P. (2021) Identifying of biomarkers associated with gastric cancer based on 11 topological analysis methods of CytoHubba. Scientific reports 11:1331.

Nakamura M., Okano H., Blendy J. A. and Montell C. (1994) Musashi, a neural RNA-binding protein required for Drosophila adult external sensory organ development. Neuron 13:67–81.

Nguyen P. and Zeng E. (2025) A protocol for weighted gene co-expression network analysis with module preservation and functional enrichment analysis for tumor and normal transcriptomic data. Bio-protocol 15.

Northcott P., Robinson G., Kratz C., Mabbott D., Pomeroy S., Clifford S., et al. (2019) Medulloblastoma Nat Rev Dis Primers 5 (1): 11.

Richter K., Good P. and Dawid I. (1990) A developmentally regulated, nervous system-specific gene in Xenopus encodes a putative RNA-binding protein. The New biologist 2:556–565.

Sakakibara S.-i., Imai T., Hamaguchi K., Okabe M., Aruga J., Nakajima K., et al. (1996) Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. Developmental biology 176:230–242.

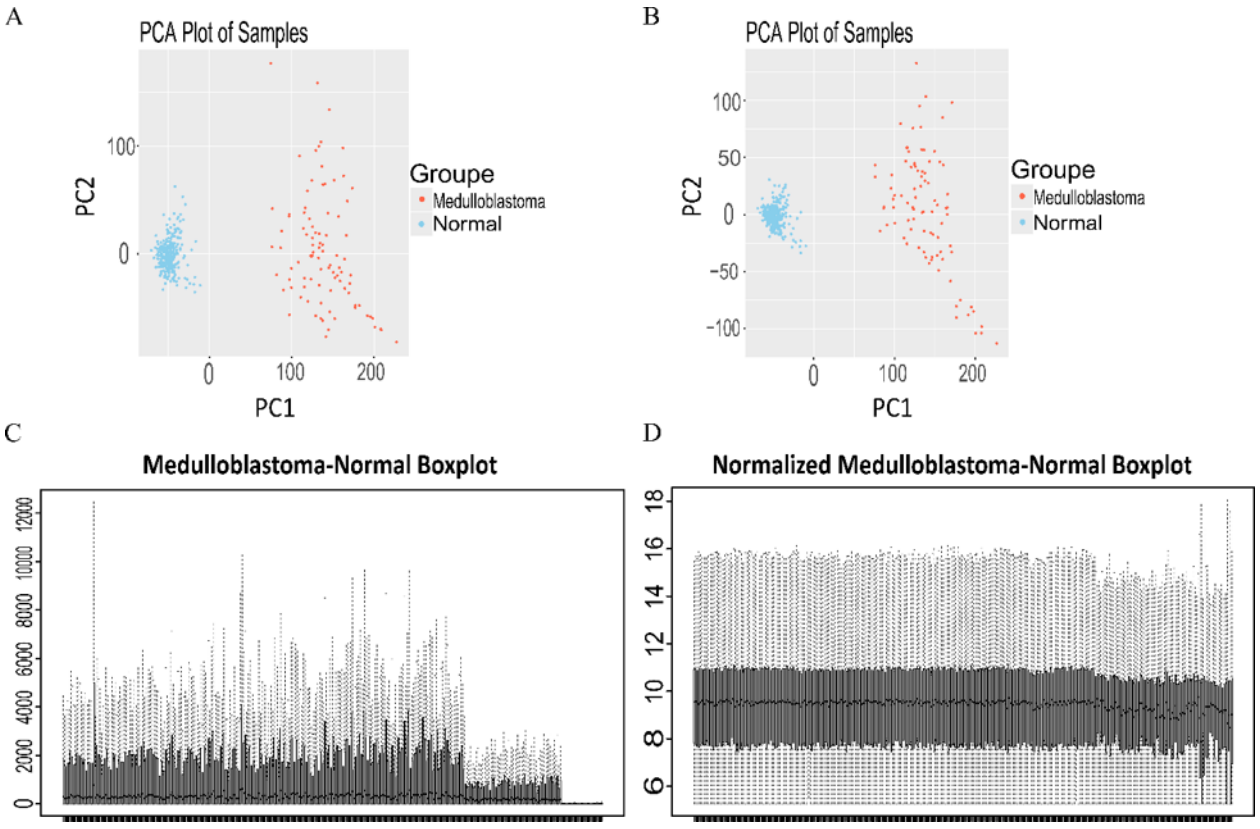
Uren P. J., Vo D. T., de Araujo P. R., Pötschke R., Burns S. C., Bahrami-Samani E., et al. (2015) RNA-binding protein Musashi1 is a central regulator of adhesion pathways in glioblastoma. Molecular and cellular biology 35:2965–2978.

Zhao F., Li Z., Dong Z., Wang Z., Guo P., Zhang D., et al. (2022) Exploring the potential of exosome-related LncRNA pairs as predictors for immune microenvironment, survival outcome, and microbiotain landscape in esophageal squamous cell carcinoma. Frontiers in immunology 13:918154.

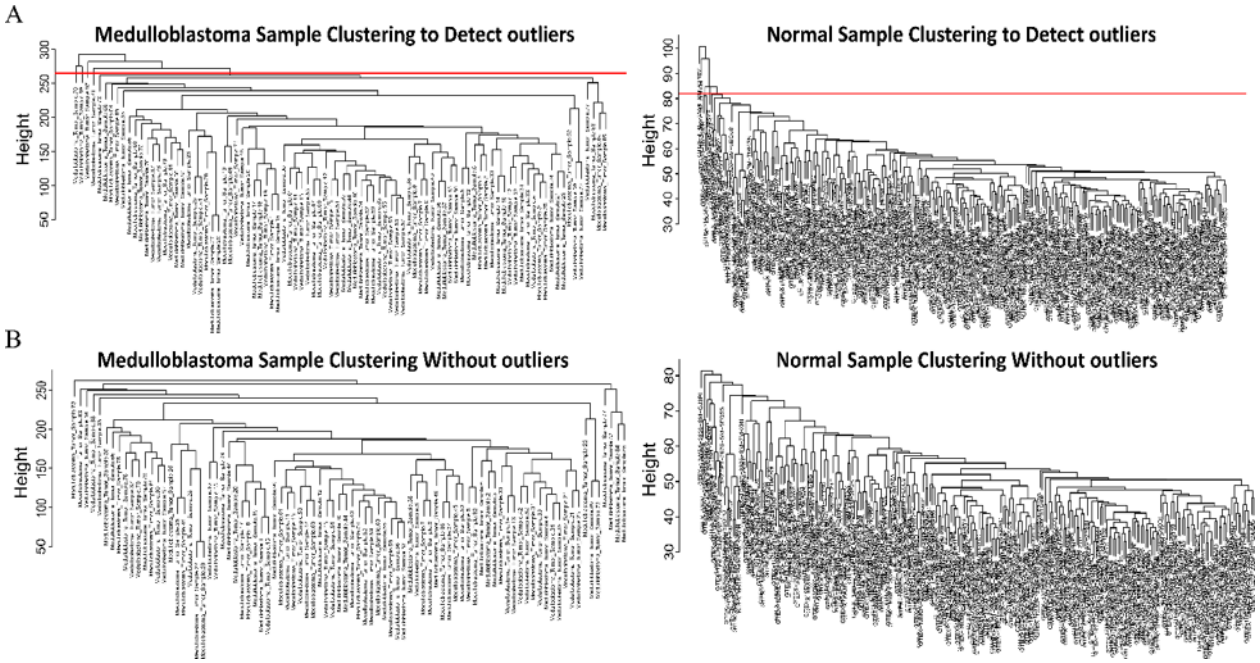
Zheng X., Chen X., Hu X., Chen L., Mi N., Zhong Q., et al. (2024) Downregulated BMP–Smad1/5/8 signaling causes emphysema via dysfunction of alveolar type II epithelial cells. The Journal of Pathology 262:320–333.

Open Access Statement:
This is an open access article distributed under the Creative Commons Attribution License (CC-BY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

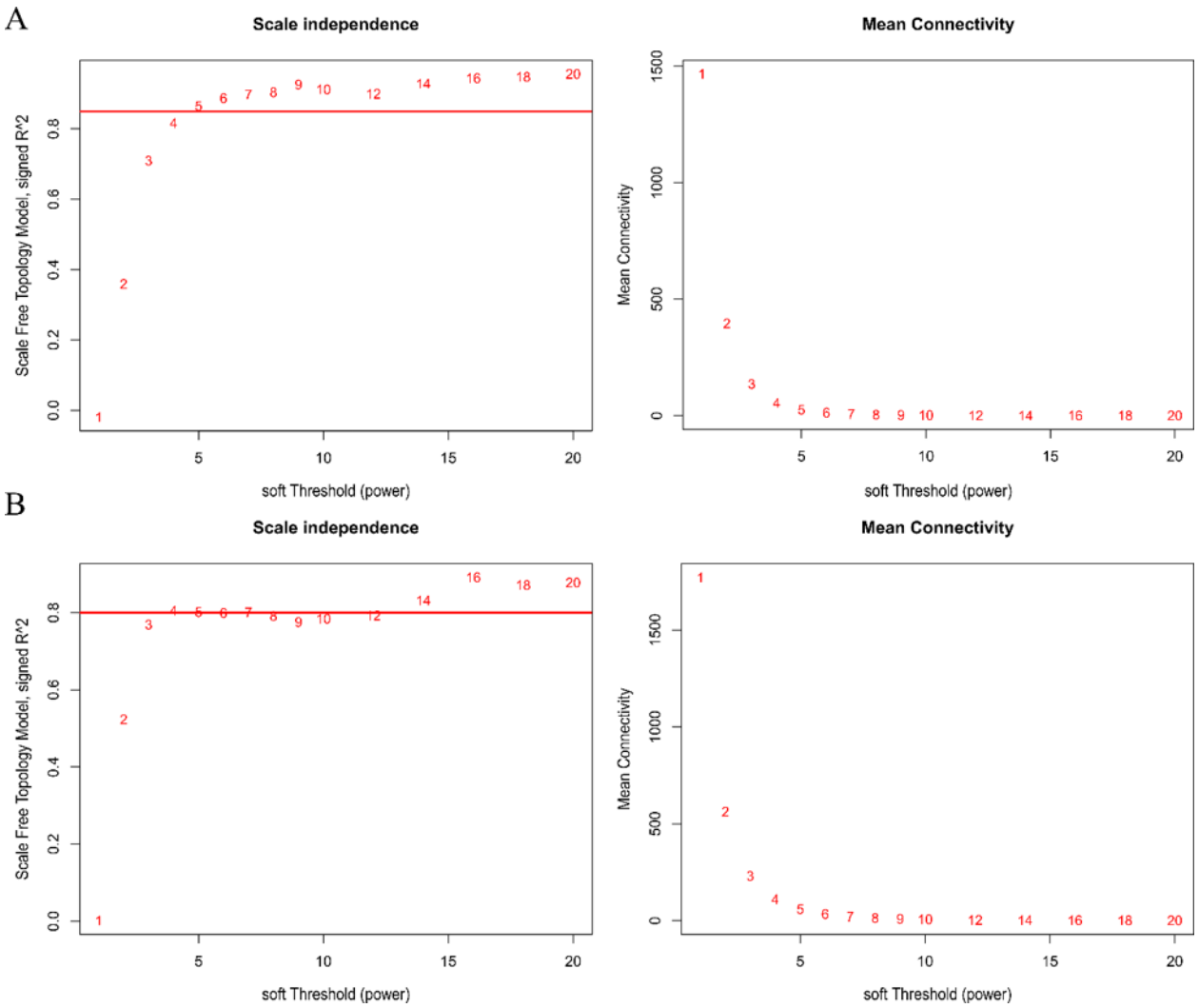
Supplementary Figures



Supplementary Figure 1. PCA and Normalization Results. (A) PCA plot before removing outliers showing separation of medulloblastoma and normal samples along PC1 and highlighting outliers along PC2. (B) PCA plot after outlier removal, showing improved sample clustering. (C, D) Box plots comparing the distribution of expression values before (C) and after (D) normalization, demonstrating reduced technical variability.



Supplementary Figure 2. Hierarchical Clustering Dendrograms. (A) Dendrogram of medulloblastoma and normal samples before outlier removal, highlighting four medulloblastoma and six normal cerebellum divergent samples. (B) Dendrogram of medulloblastoma and normal samples after outlier removal, showing cohesive clustering of remaining samples.



Supplementary Figure 3. WGCNA Soft-Thresholding Power Selection. (A) Scale independence and mean connectivity plots for medulloblastoma samples, with power = 5 selected. (B) Scale independence and mean connectivity plots for normal cerebellum samples, with power = 4 selected.