의학박사 학위논문

중심성 신경세포종의 종양발생기전

Oncogenesis of central neurocytoma

2022년 8월

서울대학교 대학원
의학과 중개의학 전공

Tamrin Chowdhury

중심성 신경세포종의 종양발생기전

지도교수 박 철 기

이 논문을 의학박사 학위논문으로 제출함

2022년 4월

서울대학교 대학원
의학과 중개의학 전공
Tamrin Chowdhury

Tamrin Chowdhury의 의학박사 학위논문을 인준함

2022년 7월

위 원 장 _____ (인)

부위원장 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

위 원 _____ (인)

Abstract

Introduction: Central neurocytoma (CN) is one of the rarest occurring tumors of the central nervous system. Usually intraventricular in location, these tumors manifest in the younger adults, but older age cases also exist. Some research has been done in the past to identify the molecular and genomic characteristics driving the development and progression of CN, but no definite cause could be determined to date.

Methods: Multi-omics study, including whole-exome sequencing, bulk and single nuclei RNA sequencing, and methylation sequencing, was done to analyze the key genomic characteristics of CN. Immunohistochemistry (IHC) analysis was done to validate the genomic findings. Additionally, telomerase repeated amplification protocol (TRAP), c-circle, and telomerase restriction fragment (TRF) assays were done to determine the telomere maintenance mechanism of CN.

Results: FGFR3 hypomethylation leading to its overexpression was found to be a major event in the ontogeny of CN. This affected crucial downstream events like aberrant activity of PI3K-AKT and neuronal development pathways. Similarities of CN with radial glial cells were shown by gene markers. CN tumor cells were found to be in a dedifferentiated state in between normal radial glial and neuron cells.

Conclusion: It was postulated that the tumorigenesis of CN is due to dysregulation in the differentiation process from radial glial cells to neurons. Finally, this study exhibited the role of FGFR3 as one of the leading drivers in the tumorigenesis of CN.

Keyword : Central Neurocytoma, Radial glia, FGFR3, hypomethylation, PI3K-AKT

Student Number : 2018-31239

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ALT	Alternate lengthening of telomere
CN	Central neurocytoma
CNS	Central nervous system
DEG	Differential expressed gene
DIG	Digoxigenin
DMR	Differential methylated region
EPN	Ependymoma
FPKM	Fragments per kilobase of exon per million mapped fragments
GBM	Glioblastoma
IHC	Immunohistochemistry
LGG	Lower grade glioma
methyl-seq	Methylation sequencing
RNA-seq	RNA sequencing
RTA	Relative telomerase activity
SGZ	Sub granular zone
snRNA-seq	Single nuclei RNA sequencing
SVZ	Subventricular zone
TMB	Tumor mutation burden
TMM	Telomere maintenance mechanism
TRAP	Telomere repeat amplification protocol
TRF	Telomere length restriction fragmentation
WBC	White blood cell
WES	Whole exome sequencing

Chapter 1. Introduction

1.1. Study Background

Central neurocytoma (CN) is one type of neuroepithelial tumor of the central nervous system (CNS) that is located exclusively in the ventricles of the brain. CN is one of the rarest among all the other CNS tumors. Only about 0.1-0.5% of cases of primary brain tumors are reported to be CNs[1]. Due to its low incidence compared to other CNS tumors, a limited number of studies have been published on CN. CN usually occurs in the young adult group, but some older patient cases have also been reported. Patients with CN typically present with symptoms of hydrocephalus caused by its unique intraventricular location of origin that hinders cerebrospinal fluid circulation[2]. CN is classified as a WHO grade II tumor but is mostly considered a benign neoplasm. Surgical resection remains the mainstream management of CN, but good responses after radiotherapy or radiosurgery are also reported for small, recurrent, and residual tumors[3]. The prognosis of CN is usually excellent[4].

Central neurocytoma was first described as a distinct neoplasm of CNS in 1982 by Hassoun et al. [5]. A brief search in PubMed (https://pubmed.ncbi.nlm.nih.gov/) from 1982 to the present day revealed only about 300 published studies on this tumor (Figure 1). Among them, 267 were written in English, mostly including case reports or review-type studies. Only 17 of the studies included some molecular or sequencing analysis of the tumor, and to date, most of the established findings of CN are based on these handfuls of studies (Figure 1). The molecular studies conducted on CN helped establish

it as a new entity of CNS tumors showing a unique histological pattern of glial and neuronal differentiation of tumor cells[6-8]. Recent histological and cellular studies have deliberated on CN originating from the adult neural progenitor cells situated in the subventricular zone (SVZ) around the lateral ventricles or the subgranular zone (SGZ) in the hippocampus based on the similarity to the bipotential radial glial cells located in this regions[9]. Studies that included sequencing experiments mostly included SNP microarray or karyotyping studies of the tumor, which failed to discover any major genomic events responsible for the development and progression of CN[10-12]. Overexpressed genes and associated pathways linked with CN development in previously published studies are listed in Tables 1 and 2, respectively. But until now, the genetic elements behind CN tumorigenesis have been inconclusive.

In the present study, the latest multi-omics technology was used to establish a genetic profile for the intraventricular CN. Through whole-exome sequencing (WES) and bulk RNA sequencing (RNA-seq), it was confirmed that CNs do not contain any major repeated somatic mutations, gene fusions, and copy number alterations responsible for the tumorigenesis. However, upregulation of the PI3K-AKT pathway and alterations of several neuronal development pathways with enrichment of radial glial cell marker orchestrated by *FGFR3* upregulation were identified as some of the principal characteristics of the tumor. Further confirmation of the genetic characteristics of tumor cells was done through single nuclei RNA sequencing (snRNA-seq). Differential expression analysis and gene set enrichment analysis between radial glia, neuron, and CN tumor cell clusters confirmed the association of the PI3K-AKT pathway with the CN tumor cells and revealed additional pathways like axonal

guidance, neuronal migration, and neuron differentiation that might have played a key role in the development of CN. After methylation profiling with methylation sequencing (methyl-seq), it was concluded that hypomethylation in *FGFR3* CpG sites is the main causative operator of the *FGFR3* overexpression, which triggers downstream activation of the PI3K-AKT pathway that ultimately leads to CN development and progression.

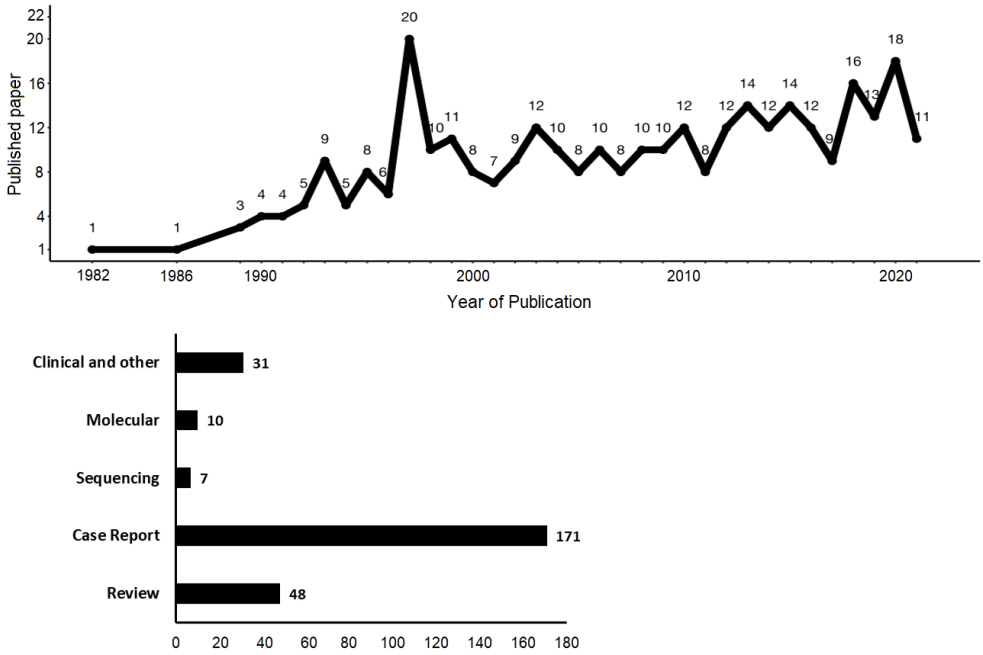

Figure 1. Published paper statistics on CN from 1982 to present. PubMed search of CN papers shows about 300 studies published on CN from 1982, among which most are case reports, reviews, and clinical studies. Molecular and sequencing-related studies on CN are the lowest.

Table 1. Overexpressed genes found in CN in previous studies

Gene name	Associated study
CALB2 *CHRDL2* *CRYAB* *FABP7* *QRFPR* *KISS1* *NEUROD4* *NTS* *RELN* *WNT4* *WNT11* *TTF1*	Vasiljevic et al, 2012
CHRNA3 *FGFR3* *FOXG1* *FZD1* *GABRB1* *GPR1* *JAG1* *NHLH2* *NR2E1* *NRG2* *PDGFD* *PIK3R3* *RXRA* *SOX2* *SOX11* *TCF4*	Sim et al, 2006
ADCYAP1 *AQP6* *BTG1* *MSTN* *IGF2* *NHLH1* *RC9* *RGS16* *SCGN* *SERPINF1* *SLIT1* *TCF4* *TOX3* *SOX4* *ZHX2*	Vasiljevic et al, 2012 & Sim et al, 2006

Table 2. Pathways linked with CN in previous studies

Pathway Name	Associated study
MARK4 human kinase	Sander et al, 2019
Notch signaling	
MAPK signaling	Vasiljevic et al, 2012
Calcium signaling	
Neurotrophin signaling	
Chemokine signaling	
WNT signaling	Vasiljevic et al, 2012 & Sim et al, 2006
IGF2 signaling pathway	Sim et al, 2006
PDGF receptor signaling	

1.2. Purpose of Research

1. Conducting a comprehensive genomic study of CN, including the latest sequencing technologies like WES, bulk, and single-cell RNA-seq, methyl seq, etc.
2. Corroborate the previously published genomic findings of CN through modern sequencing data.
3. Updating the knowledge gap regarding genomic characteristics responsible for CN development and progression and also tumor cell composition.
4. Identifying novel genetic and epigenetic features of CN tumorigenesis.

Chapter 2. Methods

2.1. Study design

In this study, the discovery cohort included 8 CN patient samples (Table 3), and the validation cohort included a tissue array of 14 CN tumor samples. For control, normal brain samples were used in both cohorts. WES and methyl-seq were conducted on 6 of the 8 samples of the discovery cohort samples. Bulk RNA-seq was done on 5, and snRNA-seq was done on 3 of the 8 samples. RNA-seq results were validated with the tissue array validation cohort. All the analysis results from the sequencing data were corroborated to reach a conceivable hypothesis behind CN tumorigenesis.

2.2. Sample collection

A total of 8 cases of fresh frozen tissue with histologically confirmed CN and matched peripheral blood samples were collected. 3 cases among these were collected as paired tumor and normal brain samples. Normal brain tissue was obtained from the frontal lobe located in the surgical corridor of the transcortical approach. Tumor and normal tissues were snap-frozen with liquid nitrogen immediately after tumor removal. White blood cells (WBC) were isolated from the blood by centrifuging the collected whole blood at 3500 rpm at 4°C. Snap frozen tissues and WBCs were then stored at -80°C. Additionally, formalin-fixed paraffin-embedded blocks of 14 cases of CN tumor samples and 2 normal brain samples were collected for validation and were made into a tissue array. Informed

consent was taken from all patients included, and the study was approved by the institutional review boards of Seoul National University Hospital (IRB No: H-1404-056-572).

2.3. DNA extraction

DNA was extracted from the frozen tumor tissue and WBC samples with Qiagen QIAamp DNA mini kit (Qiagen, Valencia, CA), following the manufacturer's protocol. Extracted DNAs were quantified with Nanodrop spectrophotometer, and then a minimum of 2μg DNA/sample were sent to Macrogen, Korea, for WES and methyl-seq.

2.4. RNA extraction

RNA was extracted from the frozen tumor tissue with RNeasy Lipid Tissue Mini Kit (Qiagen, Valencia, CA) following the manufacturer's protocol. The extracted RNA was quantified, quality checked with Nanodrop spectrophotometer (~1μg/sample) and then sent to Macrogen, Korea, for bulk RNA-seq.

2.5. Single nuclei extraction

Nuclei were isolated from frozen paired tumor and normal brain tissue separately using the 'Frankenstein' nuclei isolation protocol described previously[42]. Briefly, ~40mg of frozen tissue was homogenized in chilled Nuclei EZ Lysis Buffer (MilliporeSigma #NUC101), and then homogenate was filtered using a 70 μm-strainer mesh. The solution was centrifuged at 500 x g for 5 minutes at 4°C

in a benchtop centrifuge. Nuclei were resuspended in the EZ lysis buffer, centrifuged again, and equilibrated to nuclei wash/resuspension buffer (1x PBS, 1% BSA, 0.2U/μL RNase Inhibitor). Nuclei washing procedure was repeated three times and then stained with DAPI (10 μg/mL) or propidium iodide (PI). After isolation, nuclei were then sorted on a BD FACSAria II flow cytometer (Becton Dickinson) for sorting whole singlet nuclei and also to ensure that nuclei input was free of debris.

2.6. Whole exome sequencing and analysis

DNA samples were quality checked again by Bioanalyzer before preparing the sequencing library by random fragmentation of the DNA. Library preparation was done using SureSelectXT library prep kit. Sequencing was performed using the Illumina platform at Macrogen Korea. The generated BCL binary was then converted into raw FASTQ files utilizing the Illumina bcl2fastq package. Paired-end reads were mapped to the GRCh37 reference genome. The genome was aligned with BWA-MEM (v.0.7.15)[43]. The reads were sorted and indexed with samtools (v.1.6), and duplicate reads were marked with Picard (v.2.1.1) to reduce PCR duplication rate[44-45]. The reads proceeded base recalibration and indel-realignment process with GATK (v.3.8) for the further analysis[46]. Somatic mutations and indels were detected by MuTect2, and germline mutations were detected by HaplotypeCaller; both tools were provided from GATK (v.3.8). For somatic mutation, the mutation that passed the mutect filter was selected. Somatic and germline mutations were annotated with Cosmic 86, ExAC, and gnomAD by annovar[47-50]. To verify significant mutations, the total read depth cut off from 10 and the minor allele

frequency filter conditions are as follows: ExAC EAS < 0.01, gnomAD EAS < 0.01, and Korean < 0.01[51]. CNVs were detected with CNVkit and CoNIFER (v.0.2.2)[52-53]. After obtaining the RPKM with CoNIFER, the ZRPKM value was calculated, and then it transformed into tumor to normal log2 ratio was calculated to ascertain the amplification and deletion. Using CNVkit, CNV was checked both at the gene and segment level. It is defined as a deletion, if log2 value ≤ - 0.4, and an amplification if log2 value > 0.3. TMB was calculated as the number of mutations, which is located in the coding region, was divided by the length of the coding region of RefSeq genes (~ 30 Mb)[13].

2.7. Bulk RNA sequencing and analysis

After quality control of the samples, the sequencing library was prepared from the cDNA made from the RNA, followed by 5' and 3' adapter ligation. Library preparation was done using TruSeq standard mRNA LT sample prep kit for RNA-seq. RNA-seq was performed using the Illumina Novaseq 6000 at Macrogen Korea. The generated BCL binary was then converted into raw FASTQ files utilizing the Illumina bcl2fastq package. Transcriptome was aligned to the GRCh37 reference genome by STAR (v.2.6.0.a)[56]. The expected counts and FPKM were calculated by RSEM (v.1.3.1)[57]. Differential expression genes (DEGs) were calculated using the R package DESeq2[58]. If a gene has the absolute value of log2 ratio is more than equal to 2, adjusted p-value is less than 0.05, and the base mean is more than equal to 100, then it is considered a significant DEG. Gene ontology enrichment analysis was performed with gProfiler (https://biit.cs.ut.ee/gprofiler/gost)[59]. Furthermore, RNA fusion was

detected with STAR-Fusion (v.1.4.0)[60]. Reads were aligned with STAR (v.2.6.0.a). To identify the confident fusion, we made fusion criteria: (1) one of the fusion genes must be a protein-coding gene, (2) the spanning reads and junction reads should not be '0', and (3) the duplicates fusion, which has same read counts at the same position, were excluded.

2.8. Normal brain dataset

Sequencing reads and meta-data for normal postmortem human brains were downloaded through Synapse.org at accession syn12299750[54-55]. This dataset originated from postmortem tissue homogenates of dorsolateral prefrontal cortex gray matter approximating Brodmann area 46/9 in postnatal samples. RNA-seq libraries were constructed from high RNA quality samples using Illumina mRNA sequencing prep Kit following the manufacturer's protocol. The final cDNA libraries were sequenced by Illumina HiSeq 2000 with 100bp paired-end reads after multiple levels of quality controls.

2.9. Single nuclei RNA sequencing and analysis

For each sample, approximately 8,500 single nuclei were sorted directly into 25.1 μL of reverse transcription reagents from the 10x Genomics Single Cell 3′ Reagents kit (without enzyme). Libraries were prepared according to the manufacturer's instructions (10x Genomics) after undergoing 10x Chromium process. Prepared libraries were finally sequenced on the Next-seq (Illumina) at Psomagen Inc. (USA). The reads were aligned to GRCH37 reference

by cellranger (v.5.0.1)[61], provided by 10x Genomics. Aligned reads were run through the Seurat (v.4.0.1)[62] package basic pipeline in Rstudio (v.4.0.3). We filtered out cells that have the number of feature counts over 8500 or less than 200 and over 3 percent of mitochondrial counts for each sample. Six samples were integrated with Harmony (v.1.0)[63]. Cell type-specific marker was classified with scHCL (v.0.1.1)[17,] and well-known neural lineage markers were detected (Table 4). We performed gene ontology enrichment analysis with gProfiler to find differences between tumor and normal in radial glial cells and visualized tumor-specific markers with nebulosa[64]. We conducted pseudotime analysis using three major cell types - radial glial, tumor, and neuron cell - with monocle2[65-67].

2.10. DNA methyl sequencing and analysis

After quality control of the samples, the sequencing library was prepared by random fragmentation of the DNA, and library preparation was done using SureSelect Methyl-Seq library prep kit. Methyl-seq was performed using the Illumina platform at Macrogen Korea. The generated BCL binary was then converted into raw FASTQ files utilizing the Illumina bcl2fastq package. Methylome was aligned to the GRCh37 reference genome by Bowtie2 (v.2.2.7)[68]. Methylated and unmethylated reads were detected by Bismark (v.0.20.0)[69]. Methylation ratio is calculated as methylated reads in both strands (positive and negative) were divided by total reads. CpGs that intersected with the normal sample was extracted. The total number of intersected CpGs was 395,792. The average of each tumor and normal ratio for every CpGs were calculated. DMRs are called If the differences between normal and tumor average ratio is

more than 0.4. DMRs were defined as hypo-methylation if the average tumor ratio is less than 0.3 and hyper-methylation if it is greater than 0.7. A probe was considered CN specific if the mean of CN samples differs from the mean of the other CNS tumors by more than 0.4.

2.11. Telomere repeat amplification protocol (TRAP) assay

The enzymatic activity of telomerase was measured using the TeloTAGGG Telomerase PCR ELISA PLUS kit (Roche) according to the manufacturer's protocol. Tumor tissues were homogenized in ice-cold lysis buffer using automill (Tokken). Briefly, after BCA protein quantification of the lysates, 10μg of proteins were incubated in a total volume of 50 μl reaction mixture at 25℃ for 30 min to allow the telomerase to add telomeric repeats to the end of the biotin-labeled primer. Consequently, PCR was conducted for 33 cycles of 94℃ for 30 sec, 50℃ for 30 sec, and 72℃ for 90 sec, followed by an additional extension time of 10min at 72℃ and holding at 4℃. The telomerase activity was measured at 450nm and the reference wavelength 690nm. Relative telomerase activity (RTA) of each sample was calculated according to the instruction of the TeloTAGGG Telomerase PCR-ELISA PLUS kit.

2.12. C-circle assay

Detection of C-circles was performed as previously described[72]. Briefly, 30 ng DNA was combined with 10 μl 2X Φ29

Buffer, 7.5 U Φ29 DNA polymerase (NEB), 0.2 mg/ml BSA, 0.1% (v/v) Tween 20, 1 mM each dATP, dGTP, and dTTP and incubated at 30°C for 4h and 8 h followed by 20 min at 70°C. Amplification products were deposited on a Hybond N+ nylon membrane (Bio-Rad) and developed using the TeloTAGGG Telomere Length Assay Kit (Roche). Chemiluminescent signals were visualized with the ChemiDoc XRS system (Bio-Rad).

2.13. Telomere length restriction fragmentation (TRF) assay

Telomere length was determined by Southern blot using TeloTAGGG Telomere Length Assay Kit (Roche) according to the manufacturer's protocol. Briefly, 1 μg DNA was digested with RsaI and Hinf I for O/N at 37 °C, then electrophoresed on 0.8% agarose gel at 50 V for 4 h, then transferred to a nylon membrane by Southern blotting. The blotting membrane was blocked and hybridized to a digoxigenin (DIG)-labeled probe specific for telomeric repeats for O/N. The washed blot was incubated with anti-DIG-alkaline phosphatase (1:10,00 dilution) for 30 min and developed using substrate in TeloTAGGG Telomere Length Assay kit (Roche). Then the chemiluminescent signal image was captured with the ChemiDoc XRS system (Bio-Rad). TeloTool version 1.3 was used for image analysis and telomere length calculation.

2.14. Immunohistochemistry (IHC) antibody

All the antibodies used for the IHC analysis of this study had

been purchased from Abcam Inc. (Cambridge, MA). *FGFR3* (ab10651, 1:250), *PIK3R3* (ab235234, 1:10) and *AKT1* (ab235958, 1:30) were used as PI3K-AKT pathway marker. *PAX6* (ab195045, 1:50), *SOX2* (ab92494, 1:20), *FABP7* (ab110099, 1:5) for radial glial cell, *SOX10* (ab180862, 1:200) for neuroepithelial cell, *EOMES* (ab23345, 1:20), *ASCL1* (ab213151, 1:50) for intermediate progenitor cell, *NEUROD1* (ab213725, 1:500) for immature neuron and *SYP* (ab32127, 1:2000) for mature neuron were used as neural development cell markers.

2.15. Visualization and statistics

Selected 10K probes are the highest variable probes that can classify the central nervous system tumors described previously (DKFZ). The methylation value corresponding to 10K probes was used to depict the tsne plot with Rtsne[70]. The circus plot is illustrated with Circa (http://omgenomics.com/circa/). Illustrated figures were created with the help of BioRender.com. All statistical analysis has been carried out using R. Wilcoxon test was used to compare the expression between normal and tumor data, and linear model was done to find out the correlation between expression and methylation.

2.16. Data availability

All sequencing files were deposited in the short read sequence archive (https://www.ncbi.nlm.nih.gov/sra) under BioProject number: PRJNA796513.

Table 3. Sample details of the discovery cohort

Sample No	Pathology	Location	Gender	Age	WES	Bulk RNA-seq	Methyl-seq	snRNA-seq
CN01	Central Neurocytoma	Intraventricular	Female	32	O	O	O	X
CN02	Central Neurocytoma	Intraventricular	Male	21	O	O	O	X
CN03	Central Neurocytoma	Intraventricular	Male	31	O	X	O	X
CN04	Central Neurocytoma	Intraventricular	Male	40	O	O	O	X
CN05	Atypical Central Neurocytoma	Intraventricular	Female	35	O	O	O	X
CN06	Atypical Central Neurocytoma	Intraventricular	Male	33	O	O	O	O
CN24	Central Neurocytoma	Intraventricular	Male	29	X	X	X	O
CN25	Central Neurocytoma	Intraventricular	Male	24	X	X	X	O

Table 4. List of gene markers used to annotate single cell clusters[15–17]

Cell Type	Gene Markers
Tumor cells	*CSMD1, NRXN3, PCDH9, ELN, GPR144*
Neuron	*MAP2, RBFOX3, SYT1, GAD1*
Radial glia	*TNC, SLC1A3, VIM, GFAP*
Oligodendrocyte	*MOG, MBP, MAG, CNP, CLDN11*
Oligodendrocyte-like	*MOG, MBP, MAG, CNP, CLDN11, PDGFRA, CSPG4*
Endothelial	*ABCB1, EBF1, CLDN5, FLT1, EPAS1, COBLL1*
Microglia	*ITGAM, PTPRC*
OPC	*PDGFRA, CSPG4*

Chapter 3. Results

3.1. Absence of major genetic alterations in CN

Multiplatform genome profiling of the 6 CN patient samples was performed to understand the landscape view of genome alterations (Figure 2). No recurrently mutated genes or fusion were found in CN samples. Detail of top somatic mutations are listed in Table 5, and detail of fusion genes are listed in Table 6. Copy number variation calling did not show any significant recurrent area of gain or loss in any of the chromosomes (Figure 3). Commonly observed genomic alterations previously reported in various CNS tumors, such as *IDH1, IDH2, TP53, NF1, SMARCB1, FUBP1,* and *ATRX* mutations, *PTEN* deletion, *EGFR* amplification, and 1p/19q deletion, were also checked and found to be absent in CN. Tumor mutation burden (TMB) was calculated for CN and compared with the previously published 3083 tumor datasets of 27 tumor types (Figure 4)[13]. The TMB in CN was much lower than glioblastoma (GBM) but closer to lower-grade gliomas (LGG); TMB in CN was 0.63, whereas GBM & LGG were 2.03 and 0.87 consequently.

Table 5. Top 5 somatic mutations found in each CN sample

Sample	chromosome	start	end	REF	ALT	variant type	Hugo symbol	AF
CN01	18	5891219	5891219	C	T	nonsynonymous SNV	*TMEM200C*	0.533
	11	70007298	70007298	T	C	nonsynonymous SNV	*ANO1*	0.482
	4	4239589	4239589	C	T	nonsynonymous SNV	*TMEM128*	0.088
	18	19995754	19995754	G	A	nonsynonymous SNV	*CTAGE1*	0.081
	18	30846896	30846896	C	T	nonsynonymous SNV	*CCDC178*	0.071
CN02	2	24262327	24262327	A	G	nonsynonymous SNV	*WDCP*	0.5
	1	42047002	42047002	G	A	nonsynonymous SNV	*HIVEP3*	0.489
	21	37595570	37595570	C	T	nonsynonymous SNV	*DOPEY2*	0.474
	3	78676548	78676548	G	A	synonymous SNV	*ROBO1*	0.464
	1	1.54E+08	1.54E+08	C	A	nonsynonymous SNV	*NUP210L*	0.286
CN03	7	56046059	56046059	G	A	synonymous SNV	*NIPSNAP2*	0.5
	17	10417440	10417440	C	T	nonsynonymous SNV	*MYH1*	0.447
	16	71976635	71976635	C	T	synonymous SNV	*PKD1L3*	0.431
	12	53183963	53183963	T	A	nonsynonymous SNV	*KRT3*	0.408
	19	36884927	36884931	TTCCA	-	frameshift deletion	*ZFP82*	0.393
CN04	6	96974257	96974257	C	T	synonymous SNV	*UFL1*	0.479
	17	73127185	73127185	G	C	nonsynonymous SNV	*NT5C*	0.475
	17	74072906	74072906	C	A	synonymous SNV	*GALR2*	0.465
	19	46443162	46443162	-	T	frameshift insertion	*NOVA2*	0.447

Sample	chromosome	start	end	REF	ALT	variant type	Hugo symbol	AF
	4	54880005	54880005	T	A	nonsynonymous SNV	*CHIC2*	0.403
CN05	10	88705381	88705381	A	G	nonsynonymous SNV	*MMRN2*	0.498
	11	6942738	6942738	C	T	nonsynonymous SNV	*OR2D3*	0.439
	14	75514751	75514751	G	A	synonymous SNV	*MLH3*	0.43
	11	211315	211315	C	T	nonsynonymous SNV	*RIC8A*	0.411
	4	39293435	39293435	G	C	nonsynonymous SNV	*RFC1*	0.34
CN06	X	85906137	85906137	G	A	nonsynonymous SNV	*DACH2*	0.904
	7	75172215	75172215	C	T	nonsynonymous SNV	*HIP1*	0.513
	1	41847770	41847770	C	T	synonymous SNV	*FOXO6*	0.5
	18	3879312	3879312	T	G	nonsynonymous SNV	*DLGAP1*	0.399
	9	1.4E+08	1.4E+08	A	C	nonsynonymous SNV	*MAN1B1*	0.362

Table 6. Gene fusion detected in CN samples

Sample no	Fusion Name	Junction Read Count	Spanning Frag Count	FFPM	Left Break	Left Break Entropy	Right Break	Right Break Entropy	Type
CN01	*SSBP3--DHCR24*	49	14	1.1201	GT	1.7465	AG	1.6402	Intrachromosomal
	Z83851.1--TCF20	6	3	0.16	GT	1.5058	AG	1.7465	Intrachromosomal
CN02	*NDUFV3--PKNOX1*	11	3	0.2338	GT	1.4716	AG	1.7819	Intrachromosomal
CN04	*RP11-166D18.1--CLSTN2*	27	7	0.5577	GT	1.5656	AG	1.7968	Intrachromosomal
CN05	*Z83851.1--TCF20*	8	2	0.179	GT	1.5058	AG	1.7465	Intrachromosomal
	NDUFV3--PKNOX1	6	1	0.1253	GT	1.4716	AG	1.7819	Intrachromosomal
CN06	*USP9Y--TTTY15*	8	1	0.1554	GT	1.8323	AG	1.9219	Intrachromosomal
	Z83851.1--TCF20	5	3	0.1381	GT	1.5058	AG	1.7465	Intrachromosomal
	AC093388.3--NAB1	4	2	0.1036	GT	1.8892	AG	1.9086	Intrachromosomal
	LRRC37A2--NSF	4	2	0.1036	GT	1.7819	AG	1.9656	Intrachromosomal
	AC007038.7--RPE	3	4	0.1209	GT	1.7968	AG	1.9656	Intrachromosomal
	VAX1--KIAA1598	3	3	0.1036	GT	1.8295	AG	1.8892	Intrachromosomal

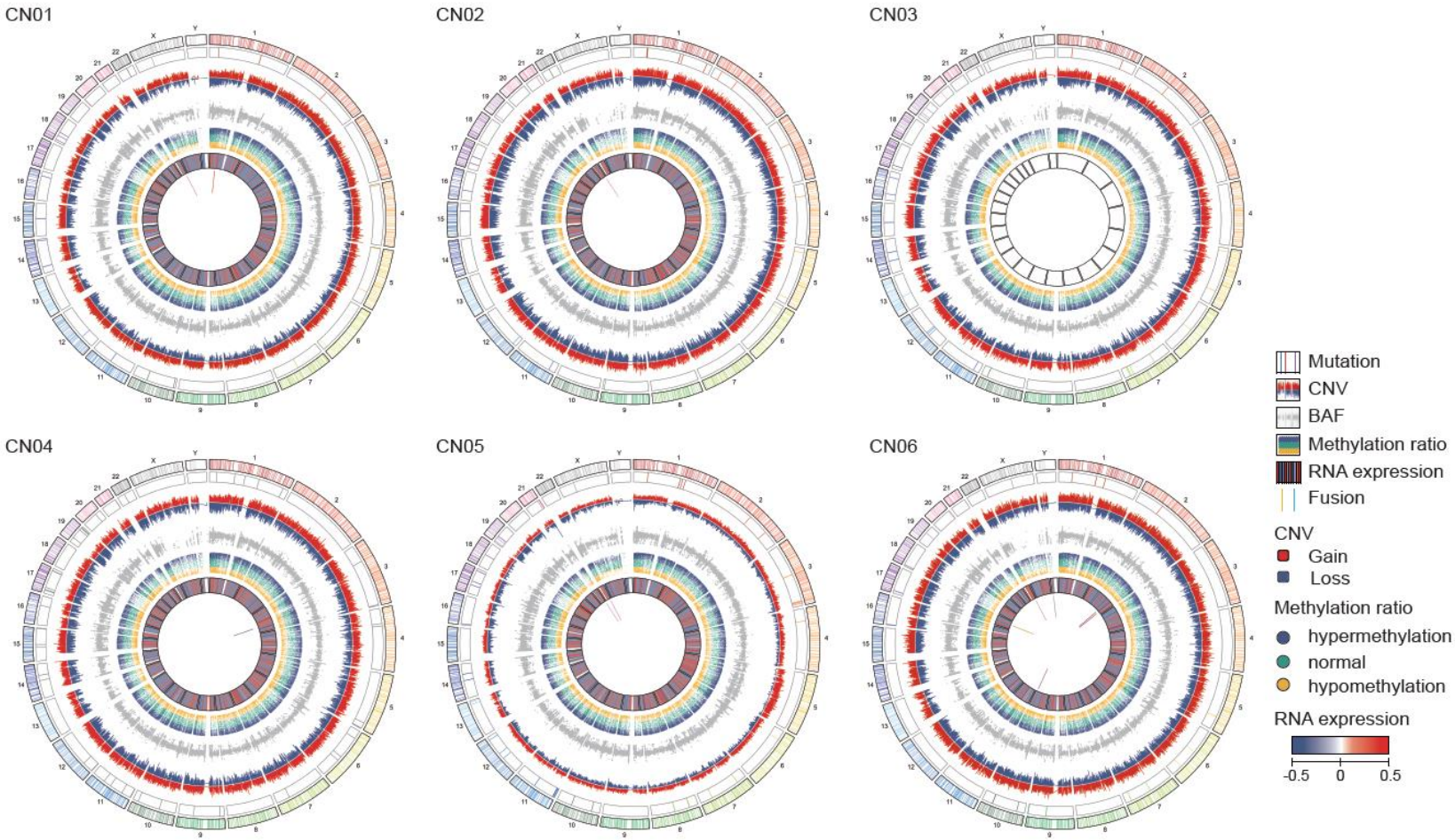
CN01
CN02
CN03
CN04
CN05
CN06
Mutation
CNV
BAF
Methylation ratio
RNA expression
Fusion
CNV
Gain
Loss
Methylation ratio
hypermethylation
normal
hypomethylation
RNA expression
-0.5
0
0.5

Figure 2. Genomic landscape view of 6 CN samples shows lack of significant driver somatic mutations. Circos plot of 6 CN samples showing their gene mutation, CNV, expression, fusion, and methylation statuses. No recurrent genomic alteration can be observed.

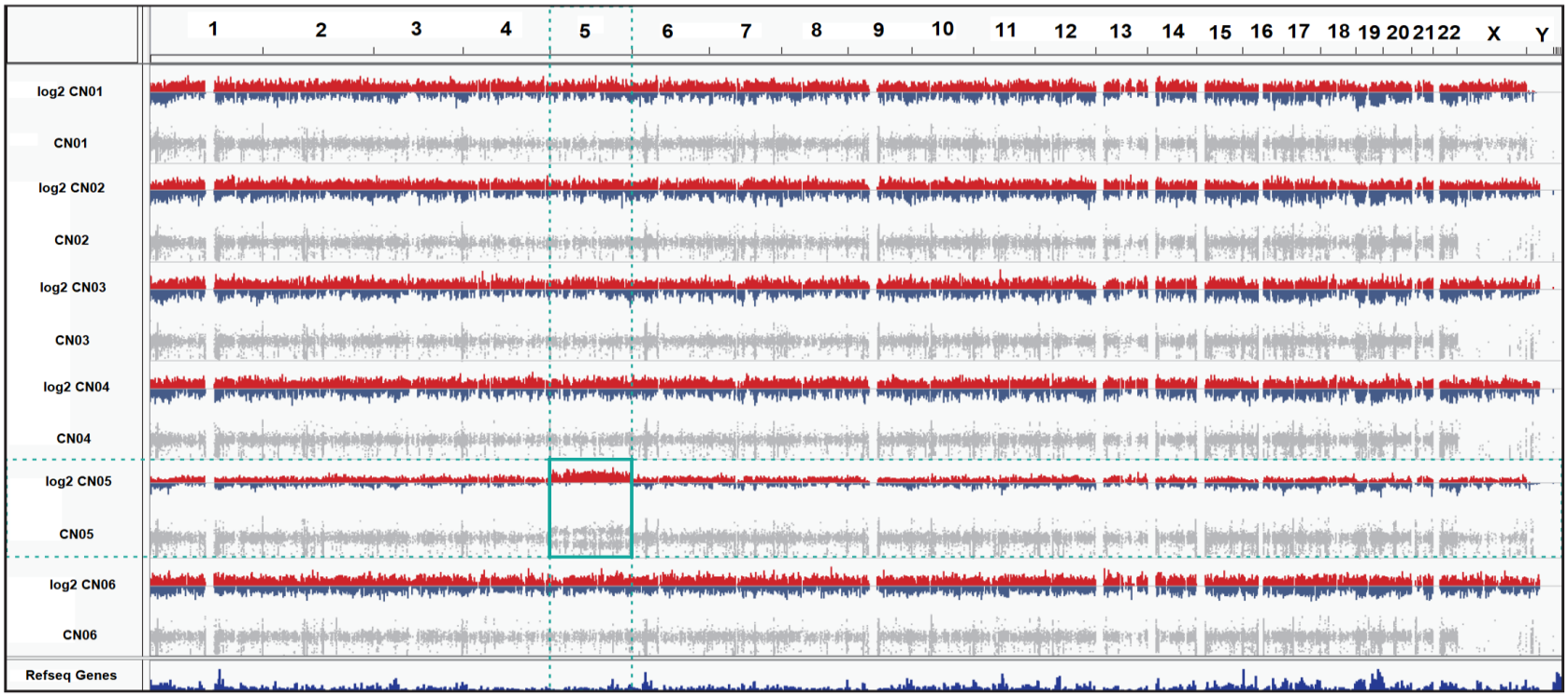

Figure 3. Copy number variation profile of 6 CN samples. CNV analysis of CN samples does not show any common region of loss or gain across chromosomes. CN05 sample shows copy number gain in chromosome 5. Others do not show any significant area of copy number gains or losses.

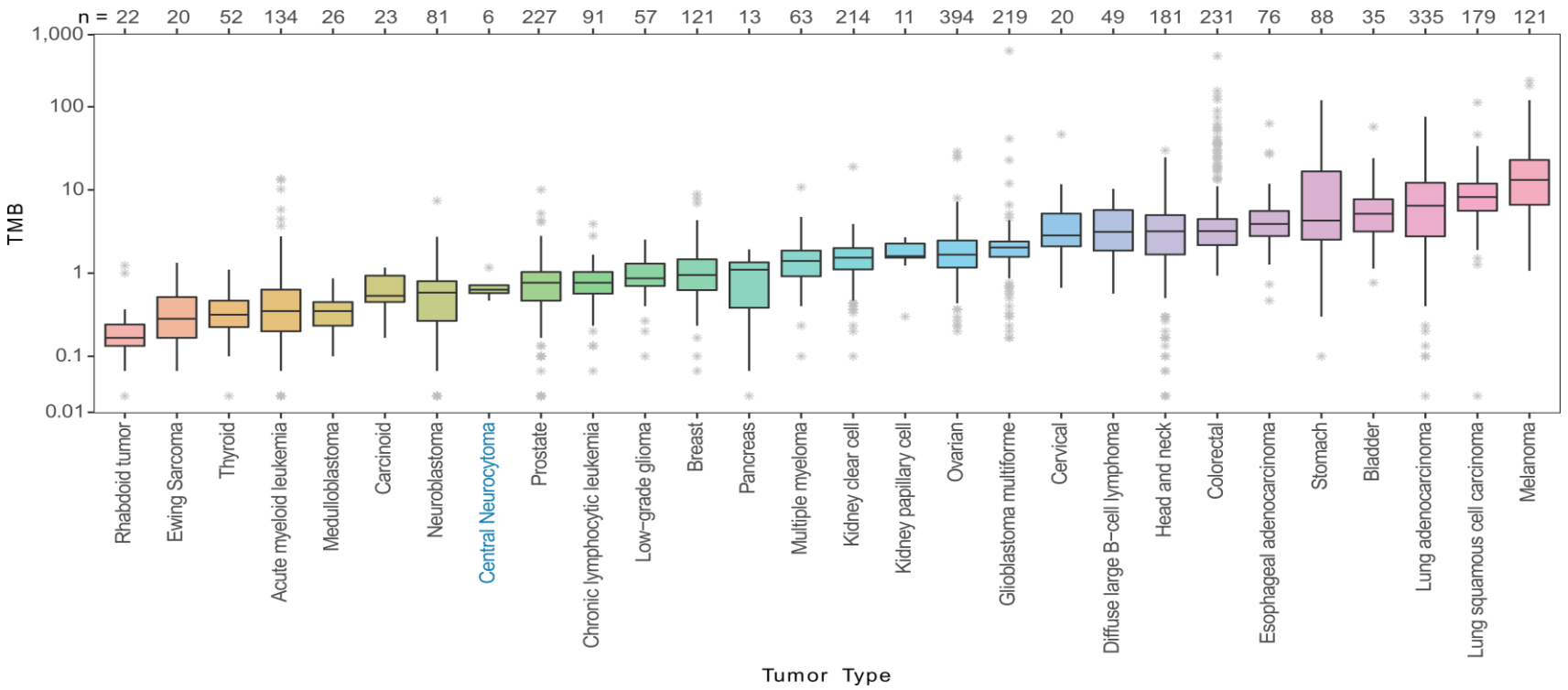

Figure 4. Tumor mutation burden of CN compared with other 27 tumor types. TMB of CN is lower than other CNS tumors such as lower-grade glioma and GBM.

3.2. Upregulation of PI3K-AKT pathway

Gene expression profiling of CNs and analysis of the differentially expressed genes (DEGs) compared with age-matched normal brain data revealed many significant genes were upregulated in CNs (Table 7). Pathway analysis with these genes showed that the PI3K-AKT pathway is the most significant aberrant pathway among the oncogenic pathways in CN (Figure 5). On closer inspection of PI3K-AKT pathway-related genes, *FGFR3*, *PIK3R3*, and *AKT1* genes were overexpressed in CN (Table 8). Additionally, the RNA expression of CN with those of previously published microarray data[11-12] was compared. Most of the upregulated genes in the present study were similarly upregulated in previous studies, including *FGFR3 and PIK3R3* (Table 9). The overexpression of *FGFR3*, *PIK3R3*, and *AKT1* genes in CN was confirmed at the protein level using IHC for an independent validation set of 14 CNs (Figure 6). However, any significant expression changes among downstream genes (*FOXO* and *GSK3*) of *AKT1* activation were not observed in CN except for the downstream genes related to *GSK3* (Figure 7). Functional annotation of the downstream genes related to *GSK3* classified these genes as metabolism, proliferation, survival, and neuronal function[14]. In CN, metabolism, proliferation, and survival-related genes were mostly upregulated, while neuronal function-related genes were downregulated (Figure 7). Additionally, gene sets related to neurodevelopment (Neuron differentiation, neural projection guidance, neural precursor proliferation, etc.) were checked, and it was found that most of them were downregulated in CN except for the neural precursor proliferation pathway (Figure 8).

Table 7. Top upregulated DEGs in CN

Gene	base Mean	log2 Fold Change	p-value	p-adj
CAPN6	352.79	12.29	0.00	0.00
EN1	124.98	11.76	0.00	0.00
GPR144	124.14	11.75	0.00	0.00
NEUROD4	317.05	11.37	0.00	0.00
KISS1	212.84	11.02	0.00	0.00
AC132217.4	627.93	10.87	0.00	0.00
FEZF1	239.95	10.82	0.00	0.00
CHRNA3	2116.44	10.76	0.00	0.00
FEZF1-AS1	432.08	10.55	0.00	0.00
CHRDL2	733.62	9.98	0.00	0.00
COL4A6	496.30	9.40	0.00	0.00
SCGN	899.53	9.23	0.00	0.00
IGF2	18315.29	9.18	0.00	0.00
SIX3	148.32	9.06	0.00	0.00
RORC	114.84	8.98	0.00	0.00
SCXA	369.37	6.80	0.00	0.00
TSPAN18	1088.27	6.65	0.00	0.00
RDH10	1063.00	6.53	0.00	0.00
HDC	204.02	6.51	0.00	0.00
SLC17A8	534.86	6.50	0.00	0.00
AOX1	238.22	6.46	0.00	0.00
CYTIP	110.31	6.46	0.00	0.00
RSPO4	341.93	6.27	0.00	0.00
SLFN11	868.88	6.21	0.00	0.00
PTGFR	285.78	6.16	0.00	0.00
NEUROG2	127.18	6.14	0.00	0.00
SUSD2	246.91	6.11	0.00	0.00
ADAMTS6	203.44	6.02	0.00	0.00
AQP6	267.01	6.01	0.00	0.00
QRFPR	134.58	5.98	0.00	0.00
PNMT	433.37	5.22	0.00	0.00
INSM1	433.26	5.20	0.00	0.00
PRLR	154.72	5.13	0.00	0.00
DRD2	175.95	5.11	0.00	0.00
COL20A1	1189.18	5.06	0.00	0.00
ACTC1	310.79	5.05	0.00	0.00
RARRES2	674.25	5.02	0.00	0.00
COL25A1	382.80	5.02	0.00	0.00

Gene	base Mean	log2 Fold Change	p-value	p-adj
GADD45G	391.86	5.02	0.00	0.00
NHLH2	317.05	5.01	0.00	0.00
MYO1B	2693.45	4.12	0.00	0.00
FGFR3	5348.90	4.10	0.00	0.00
PCDHGA2	444.80	4.10	0.00	0.00
MOB3C	330.11	4.08	0.00	0.00
PDGFD	212.27	4.08	0.00	0.00
RP11-572C15.6	101.57	4.08	0.00	0.00
ZHX2	1050.63	4.07	0.00	0.00
FAM110A	151.56	4.07	0.00	0.00
SLC12A1	145.04	4.05	0.00	0.00
PAX6	976.16	4.04	0.00	0.00
RXRA	2881.27	3.49	6.13E-110	4.76E-108
EPHA3	597.03	3.43	3.37E-44	6.84E-43
PLCB4	1235.28	3.26	6.50E-71	2.57E-69
FABP7	1280.13	3.11	3.86E-36	5.96E-35
CEBPB	366.3	3.02	1.68E-25	1.65E-24
TP53	150.47	2.94	5.52E-38	9.08E-37
PRKD2	290.31	2.87	1.30E-46	2.85E-45
CDON	367.33	2.78	5.68E-58	1.68E-56
INHBB	115.64	2.72	1.24E-18	8.34E-18
COL4A5	1345.99	2.62	2.32E-09	8.48E-09
PIK3R3	1414.26	2.59	2.72E-103	1.93E-101
SLIT1	6494.65	2.55	1.82E-34	2.63E-33
POU3F1	259.96	2.39	3.34E-18	2.20E-17
FOXO1	352.15	2.38	3.02E-15	1.68E-14
AKT1	2617.94	2.31	5.09E-40	9.01E-39
RERG	347.96	2.26	5.97E-18	3.87E-17
PLCB3	187.55	2.05	2.59E-13	1.28E-12
MPST	313.49	2.00	2.36E-19	1.65E-18

Table 8. List of upregulated pathways in CN

Pathway	adjusted_p_value	term_size	query_size	Top 5 DEGs in pathway
Ribosome	1.09E-18	152	327	*RPL5, RPLP0, RPL7, RPL6, RPL3*
AGE-RAGE signaling pathway in diabetic complications	7.75E-05	100	327	*COL4A6, STAT5A, PIK3R3, PLCB4, CDK4*
ECM-receptor interaction	0.000133	82	327	*COL4A6, THBS3, LAMB2, COL1A1, ITGB1*
Small cell lung cancer	0.000139	93	327	*COL4A6, RXRA, CASP9, PIK3R3, CDK4*
Pathways in cancer	0.000142	526	327	*COL4A6, STAT5A, RXRA, CASP9, PIK3R3*
PI3K-AKT signaling pathway	0.001127	351	327	***PIK3R3, AKT1, FGFR3***, *COL4A6, RXRA*
Hepatocellular carcinoma	0.006221	165	327	*PIK3R3, FZD1, GSTP1, CDK4, DDB2*
Platinum drug resistance	0.019269	72	327	*CASP9, PIK3R3, GSTP1, REV3L, AKT1*
Focal adhesion	0.024802	198	327	*COL4A6, PIK3R3, ILK, THBS3, PDGFD*
Central carbon metabolism in cancer	0.031126	64	327	*PIK3R3, PGAM2, HIF1A, SLC1A5, AKT1*
Human papillomavirus infection	0.031163	330	327	*COL4A6, PIK3R3, FZD1, JAG1, CDK4*

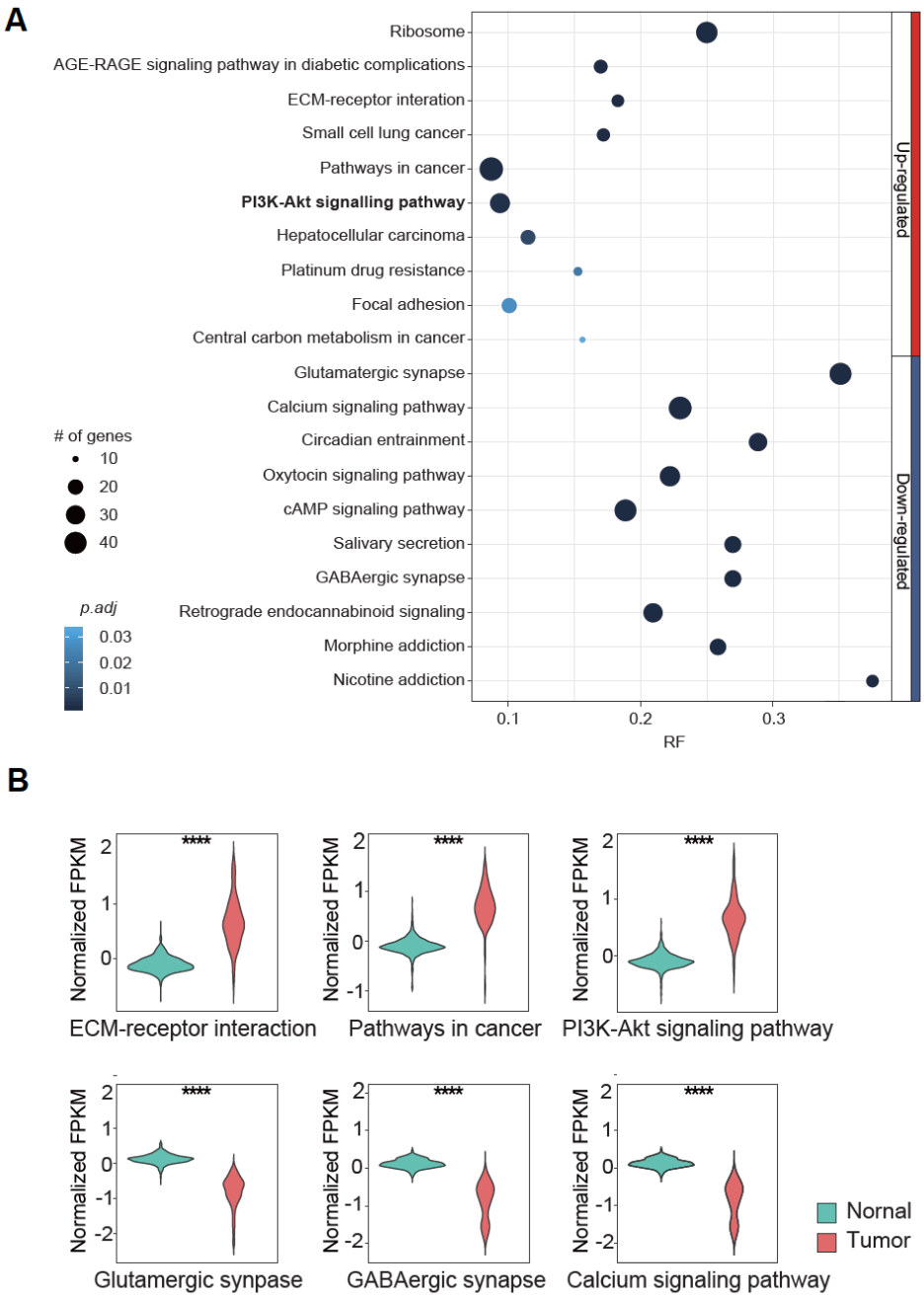

Figure 5. Pathway analysis of CN significant DEGs. A. Scatter plot of pathway analysis result shows PIK3-AKT pathway was significantly upregulated in CN, B. PI3K-AKT1 pathway-related genes were significantly overexpressed in CN compared to the normal brain samples (****: P-value<0.0001).

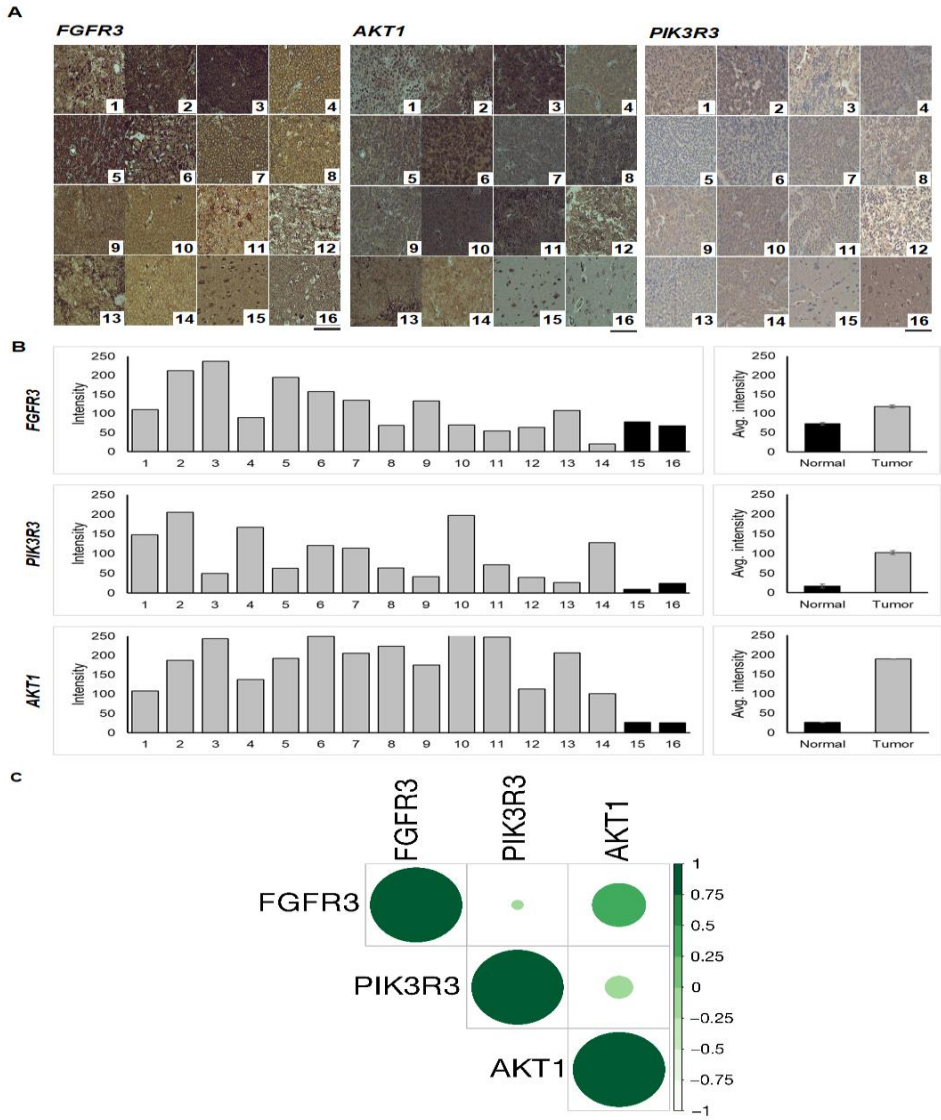

Figure 6. IHC validation of *FGFR3*, *PIK3R3,* and *AKT1* genes in CN. A. IHC pictures of *FGFR3, PIK3R3*, and *AKT1* genes. No 1-14 represents CN, and no 15-16 represents normal brain samples. All three were seen overexpressed in CN compared to normal brains. B. Quantification of IHC images of *FGFR3, PIK3R3*, and *AKT1* shows individual and average intensity of CN and normal brain samples confirming the findings of the IHC images. C. Correlation plot of *FGFR3, PIK3R3*, and *AKT1* IHC quantification scores shows a positive correlation of *FGFR3* and *PIK3R3* with *AKT1*.

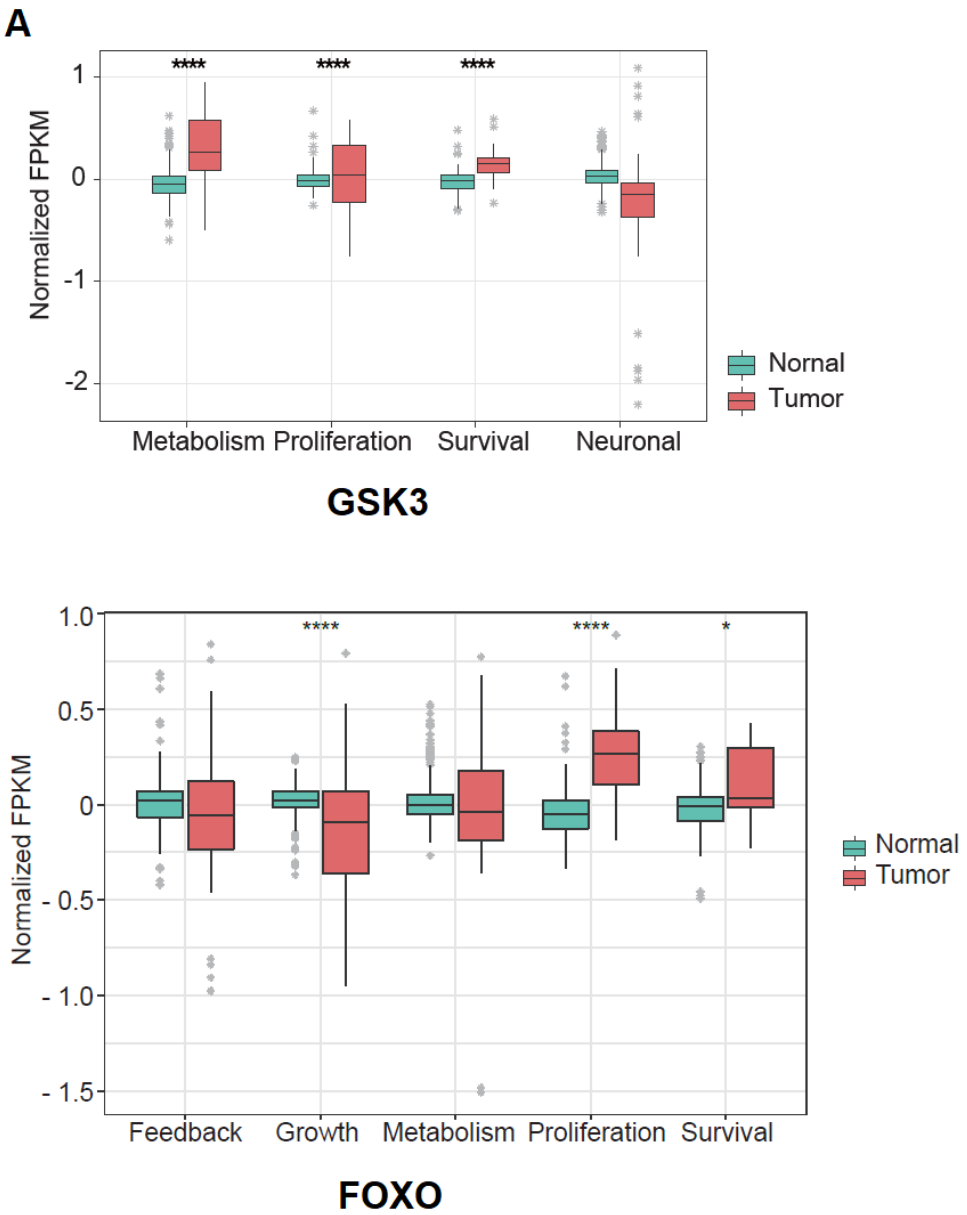

Figure 7. Downstream gene activation of the PIK3-AKT pathway. A. Box chart of collective gene expression of downstream genes of PIK3-AKT (*FOXO* and *GSK3*) shows significant upregulation of metabolism, proliferation, and survival function-related genes under the GSK3 gene.

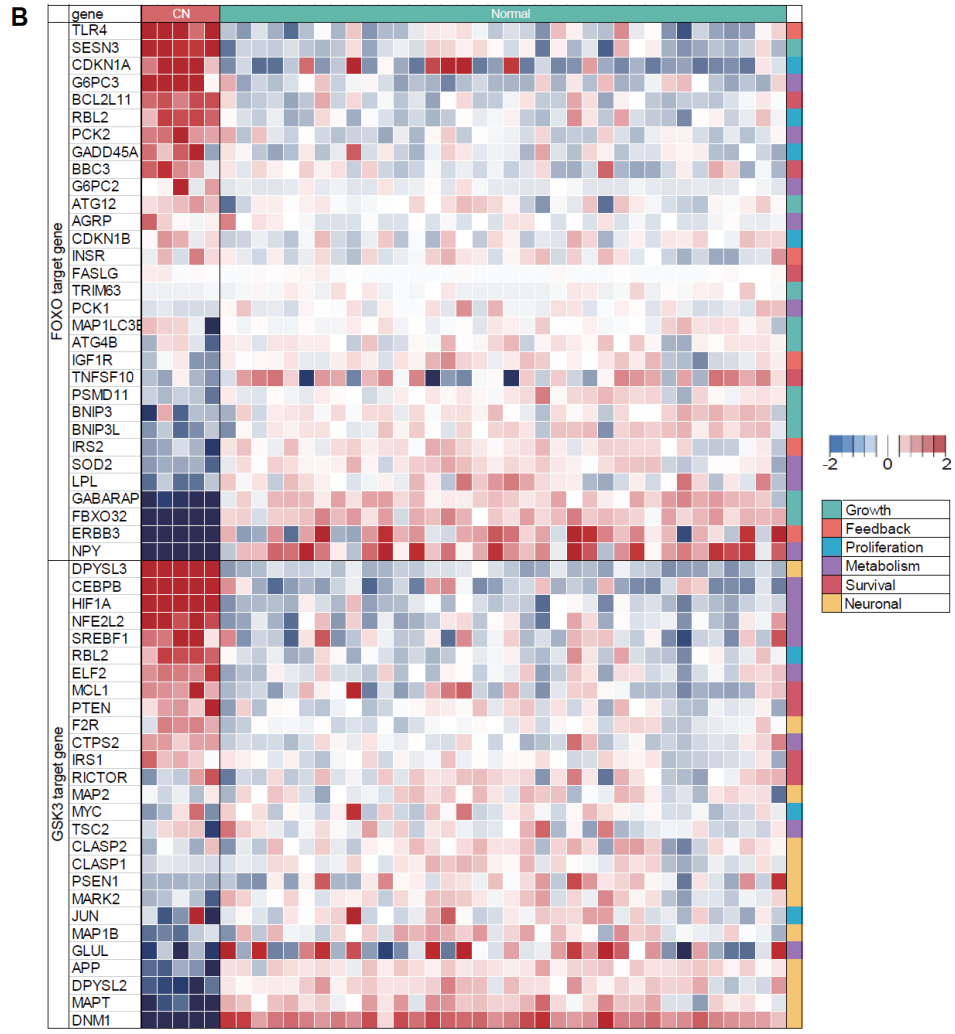

Figure 7 (Continued). Downstream gene activation of the PIK3-AKT pathway. B. Heatmap of individual genes downstream of *FOXO* and *GSK3* reflects the results found in the box plots.

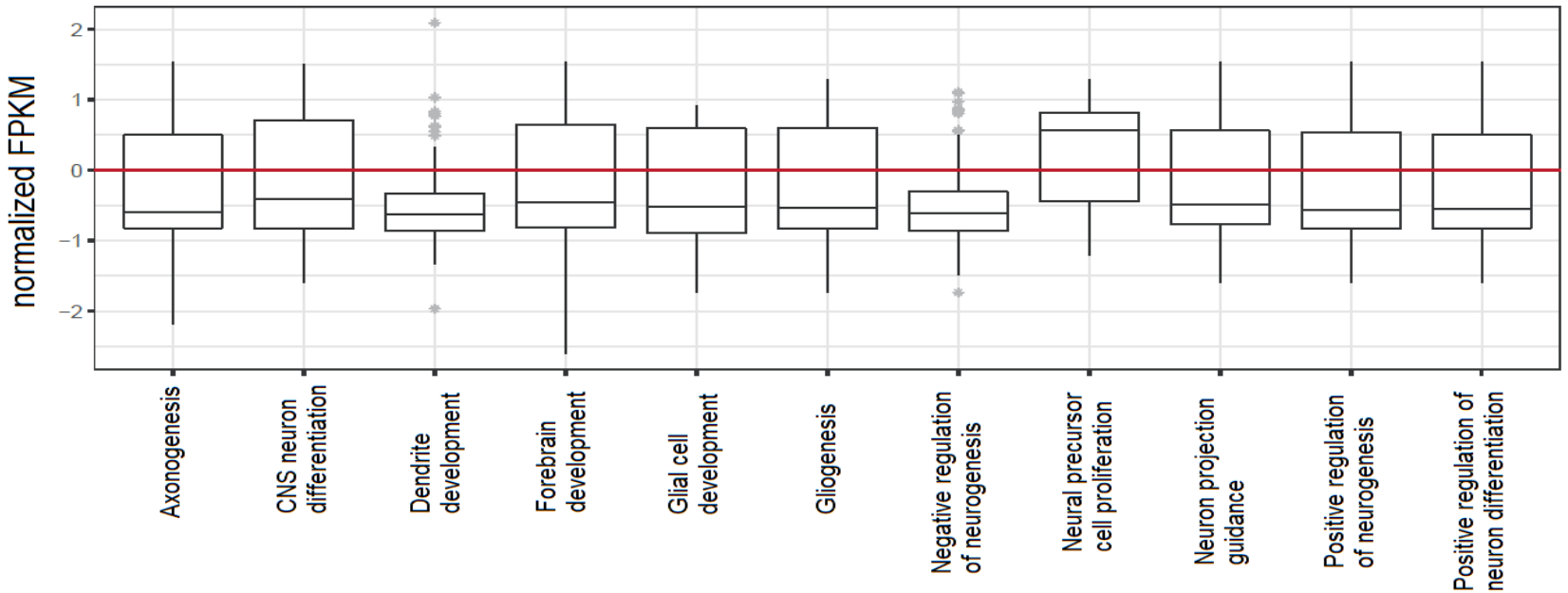

Figure 8. Neuronal development related pathways downregulated in CN. DEGs related to neuronal developmental pathways show downregulation of pathways such as CNS neuron differentiation, gliogenesis, neuron projection guidance, glial cell development, etc. Neural precursor cell proliferation pathway-related DEGs are upregulated.

Table 9. Comparison of this study results with previous published studies

Upregulated genes in CN reported in previous studies	Log2 Fold Change value in present study
CHRNA3	10.76323
FGFR3	4.104819
FZD1	4.142785
GABRB1	2.255745
GPR1	3.14667
JAG1	3.914759
NHLH2	5.005447
NR2E1	3.252535
PDGFD	4.081706
PIK3R3	2.599747
RXRA	3.496274
SOX11	3.459088
CALB2	5.966045
CHRDL2	9.981614
FABP7	3.112061
QRFPR	5.977095
KISS1	11.02432
NEUROD4	11.37481
RELN	3.410256
WNT4	4.814981
ADCYAP1	3.474997
AQP6	6.009566
BTG1	2.984412
MSTN	7.071411
IGF2	9.183429
NHLH1	8.504442
RGS16	5.935952
SCGN	9.232411
SERPINF1	3.238606
SLIT1	2.559653
TOX3	3.585637
SOX4	3.088334
ZHX2	4.074736

3.3. Radial glial cell gene markers enriched in CN

Taking cues from the fact that a set of neuronal function-related genes are downregulated, and neural precursor proliferation-related genes are upregulated in the CN, the expression of the neural lineage-related markers was investigated in search of CN cell of origin. The neural progenitor lineage can differentiate into either neural or glial-type cells depending on environmental cues. Radial glial cells have the capability to differentiate into mature neurons or glial cells, such as astrocytes or oligodendrocytes (Figure 9A). To identify the cell type of CN origin, well-known gene markers for each of these neural progenitor cell types previously reported in multiple studies were explored (Figure 9A, B, and Table 10), and the findings at the protein level were corroborated by IHC with a 14 CN-sample tissue array validation set (Figure 10). Collectively, at the RNA level, it was confirmed that CN showed lower expression of genes related to neuroepithelial cells, mature neurons, immature neurons, astrocytes, and oligodendrocytes but higher expression of radial glial cell and intermediate progenitor cell-related genes (Figure 9A). However, when delved into the gene expression patterns of each cell type, it was observed that CN exhibited more prominent enrichment of the radial glial cell signature than the intermediate progenitor cell signature. *SOX2, PAX6,* and *FABP7,* which were used as radial glial cell markers, showed significant analogous expression patterns at both the RNA and protein levels, whereas *EOMES* and *ASCL1*, which were used as markers of intermediate progenitor cells, showed different patterns (Figure 9B, and 10). *EOMES* was minimally overexpressed in CN samples compared to normal brains at the RNA level but not at the protein

level. This was not surprising considering the minimal RNA expression level difference (Figure 9B, and 10). On the other hand, *ASCL1* was considerably overexpressed in CN at the RNA level, but at the protein level, it was not significantly overexpressed (Figure 9B, and 10). In view of these results, it was determined that the CN origin is most likely to be radial glial cells rather than intermediate progenitor cells.

Table 10. List of marker genes used for neural precursor cells with reference

Marker gene	Neural progenitor cell type	Reference
SOX10	Neuroepithelial cell	PMID: 11641219, 33481357
PAX6	Radial glia	PMID: 19274100
SOX2	Radial glia	PMID: 26430216, 29568500
FABP7 (BLBP)	Radial glia	PMID: 17580100, 17428991
EOMES (TBR2)	Intermediate progenitor	PMID: 18385329, 35203375
ASCL1 (MASH1)	Intermediate progenitor	PMID: 21483754, 26421301
NEUROD1	Immature neuron	PMID: 19274100
SYP	Mature neuron	https://doi.org/10.1016/B978-0-323-44941-0.00001-1

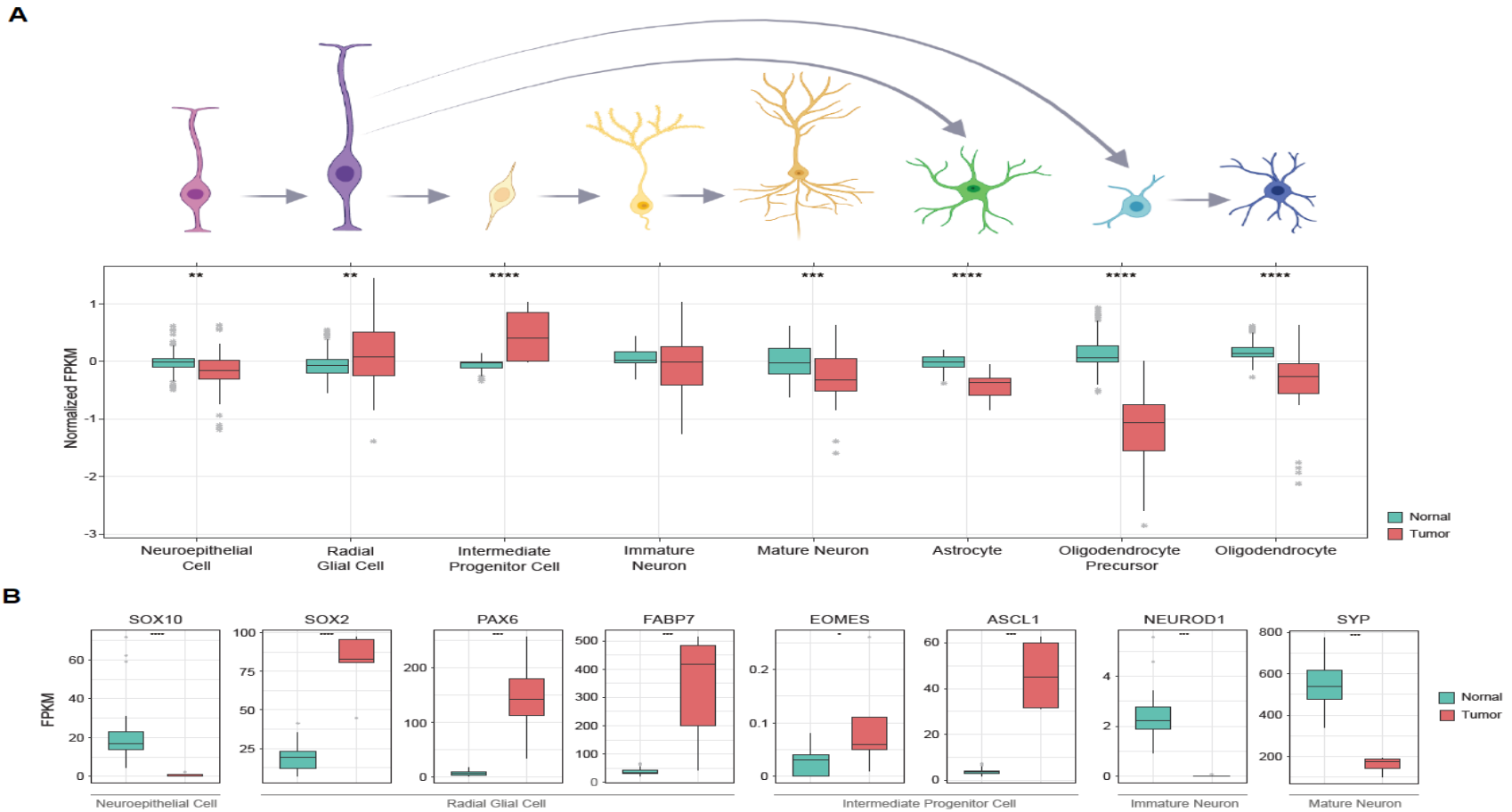
A
Normalized FPKM
Neuroepithelial Cell
Radial Glial Cell
Intermediate Progenitor Cell
Immature Neuron
Mature Neuron
Astrocyte
Oligodendrocyte Precursor
Oligodendrocyte
Nornal
Tumor
B
SOX10
SOX2
PAX6
FABP7
EOMES
ASCL1
NEUROD1
SYP
FPKM
Neuroepithelial Cell
Radial Glial Cell
Intermediate Progenitor Cell
Immature Neuron
Mature Neuron
Nornal
Tumor

Figure 9. Neural precursor cell-related gene expression in CN and normal brain. A. Upper schematic figure showing various stages of neural precursor cells in the normal brain and their progression. The arrow shows the expected paths of the cells during normal neural differentiation. Lower bar charts show the combined gene expression patterns of these different neural developmental cell marker genes in CN and normal brain samples. Significant overexpression of genes related to radial glial and intermediate progenitor cells can be seen in CN (Asterisks represent the following P-value, * <0.05, **<0.01, ***<0.001, ****<0.0001), B. Individual marker gene expression patterns showing significant overexpression of all marker genes representing radial glial cell and intermediate progenitor cell in CN compared to normal brain (Asterisks represent the following P-value, * <0.05, **<0.01, ***<0.001, ****<0.0001)

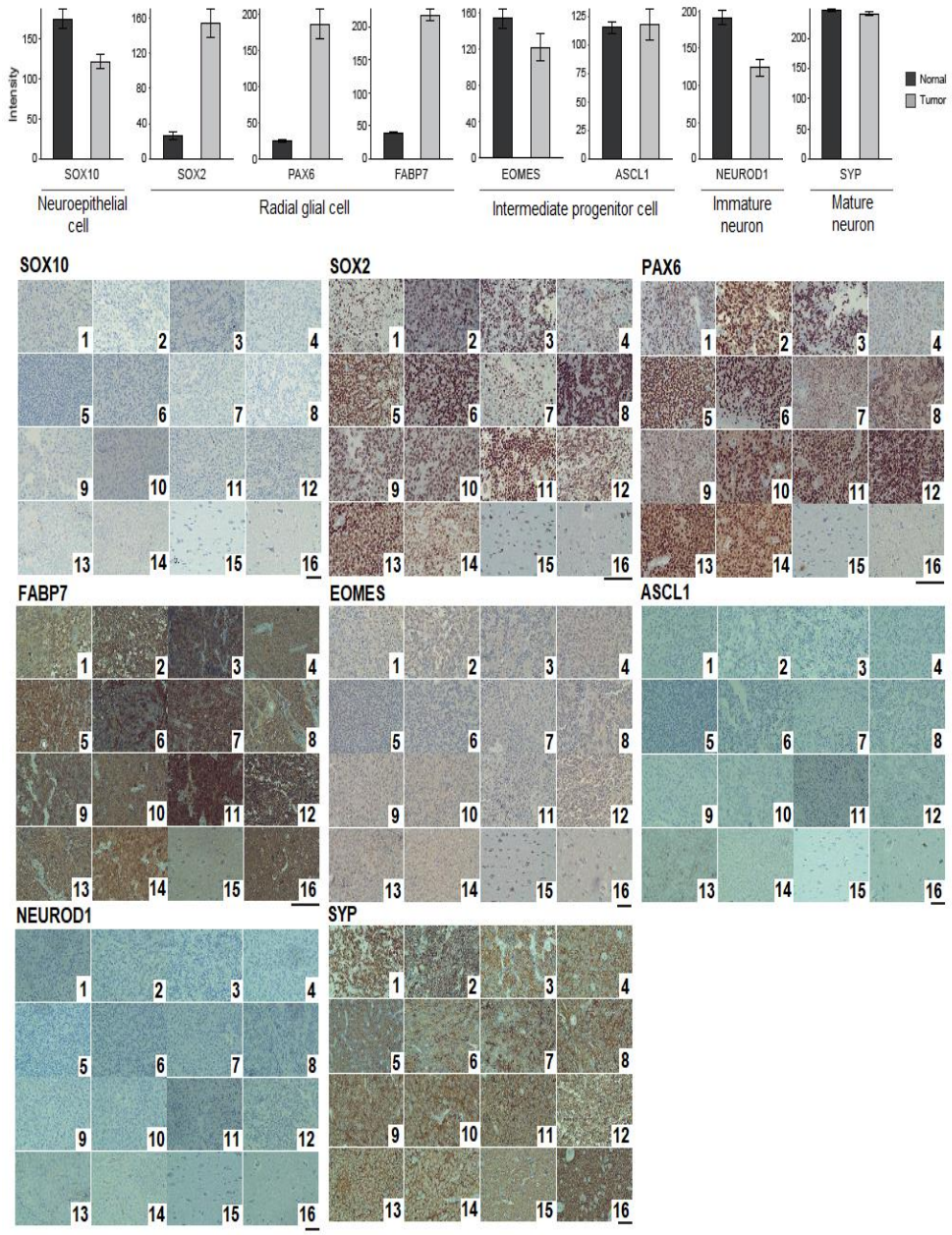

Figure 10. IHC validation of neural precursor cells in CN. Bar charts and IHC images of neural precursor cell marker genes showed radial glial cell markers *SOX2, PAX6,* and *FABP7* were overexpressed in CN, and intermediate progenitor cell markers *EOMES* and *ASCL1* were not overexpressed in CN. No 1-14 represents CN, and no 15-16 represents normal brain samples in the IHC images.

3.4. CN tumor cell features

To identify the cell population that constitutes CN, snRNA-seq was conducted with 3 CN tissues and paired normal brain tissue, obtained during the transcortical approach during resection. A total of 32,945 cells, which consisted of 17,122 normal cells and 15,823 tumor cells, were used in the analysis. Classification of the cell clusters was done based on previously published single-cell references from brain tissues and well-known markers (Table 4)[15-17]. As expected, the major cell clusters in the normal brain samples were recognized as radial glia (n=1138), oligodendrocyte (n=3334), and neuron (n=11484). Except for the major cluster of tumor cells (n=14663), many tumor cells overlapped with radial glia (n=552). The minor population of tumor cells was clustered with the neuron cell cluster (n=166), and the oligodendrocytes like cell cluster split into two groups, one consisting of normal cells (n=66) and the other enriched with tumor cells (n=130) (Figure 11). We did not find any oligodendrocyte like cells in the tumor sample.

The tumor cell population was further investigated to detect if there were any intratumoral heterogeneity present in CN. There were 6 clusters in the tumor samples altogether; among them, two clusters (clusters 0 and 2) had the majority of the tumor cell population. These two clusters were confirmed in all three tumor samples individually (Figure 12). The minor tumor cell clusters can be confirmed as sample-specific; these clusters can be either considered as sample bias or a sign of intratumoral heterogeneity. But due to the lack of enriched genes for each cluster and the limited number of cells per sample for these tumor clusters, this cannot be determined with confidence. Highly enriched genes for each cluster

were checked, and *FGFR3* was found to be enriched in both tumor and radial glial cells (Figure 13A). Genes previously confirmed via bulk RNA-seq as CN specific were also enriched in the snRNAseq data in CN tumor samples. Radial glial cell signature genes (*PAX6, FABP7*) and *FGFR3* expression were confirmed both in the radial glial and tumor cell clusters, whereas *PIK3R3* and *AKT1* were only expressed in the tumor cell cluster (Figure 13B). *SOX2* was less expressed in the single cell tumor clusters compared to the bulk RNA-seq data.

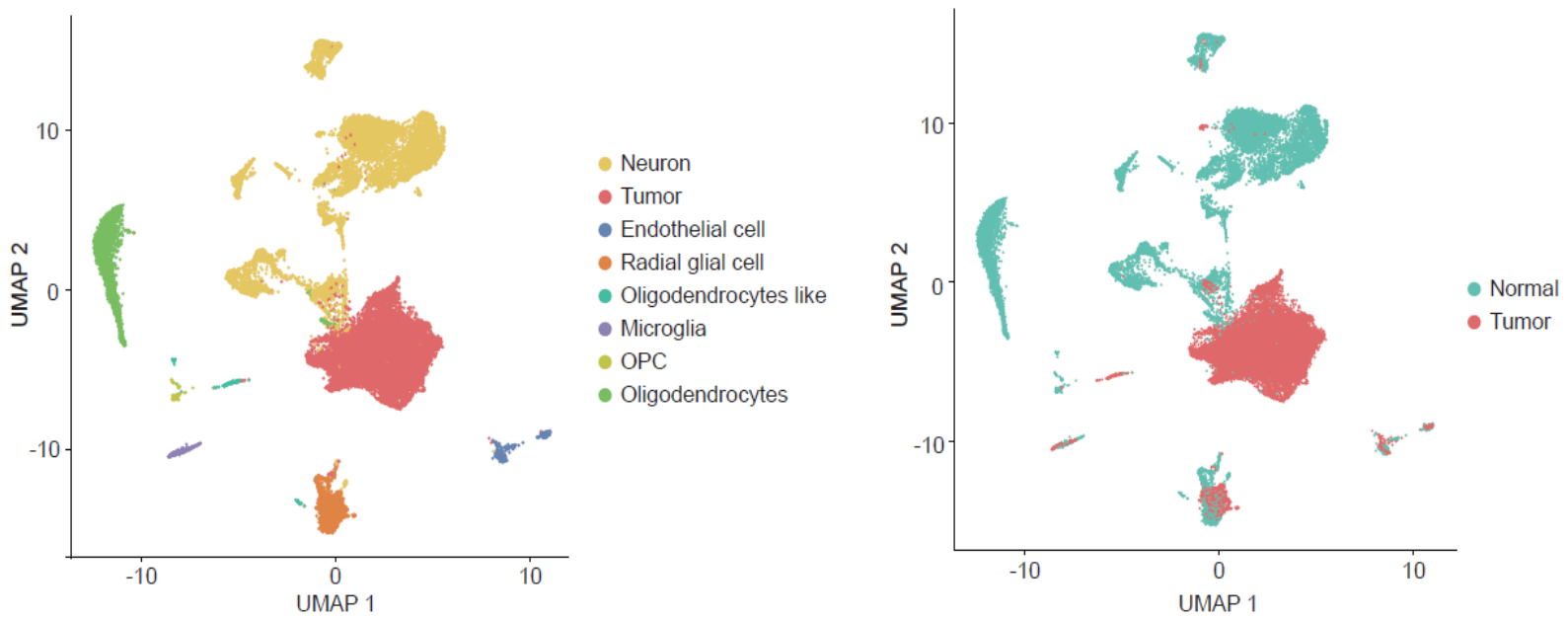

Figure 11. Cell populations identified in CN and normal brain. CN samples mainly consisted of tumor and radial glial cells, whereas normal brain samples consisted of radial glial cell, neuron, oligodendrocytes, OPC, microglia, and endothelial cells.

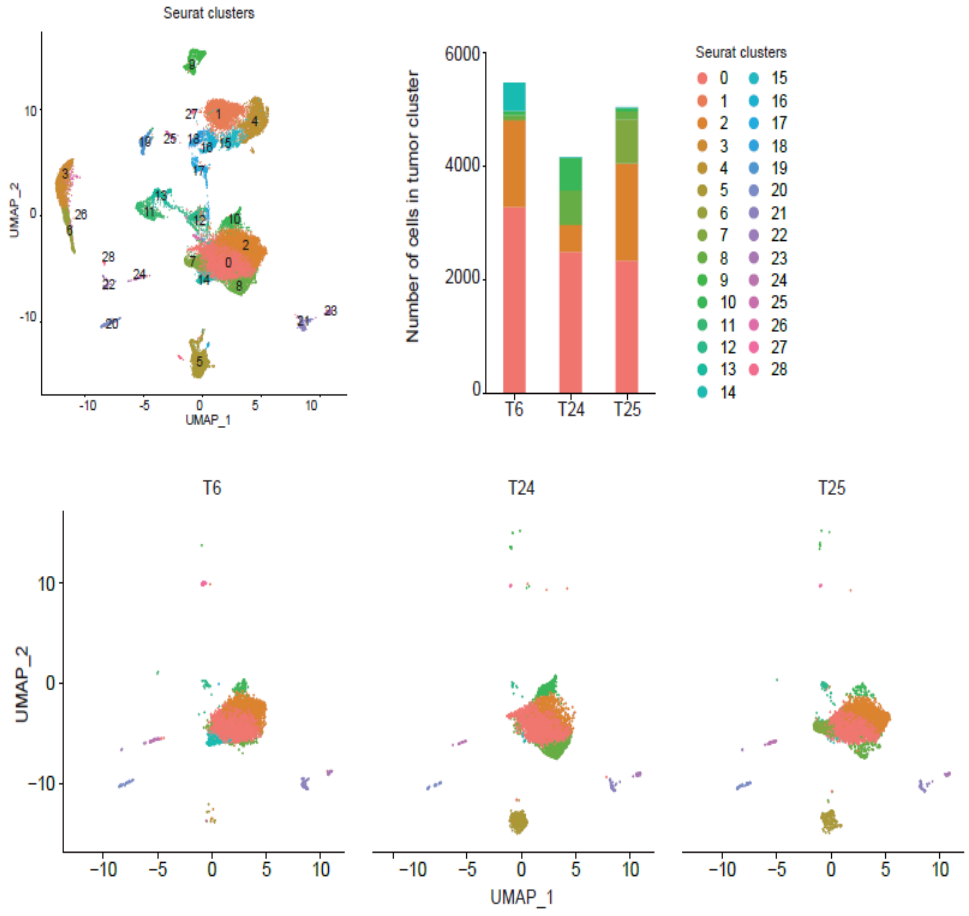

Figure 12. Individual clusters in each CN tumor sample. UMAP showing Seurat clusters of all cells based on snRNA-seq. Bar chart showing the number of cells in each tumor sample. Tumor cell-specific clusters (0, 2, 7, 8, 10, and 14) in 3 CN samples show that major clusters (0, 2, and 10) are common between the samples. Sample-specific clusters (7, 8, and 14) are minimal and more likely to represent sample-biased cells.

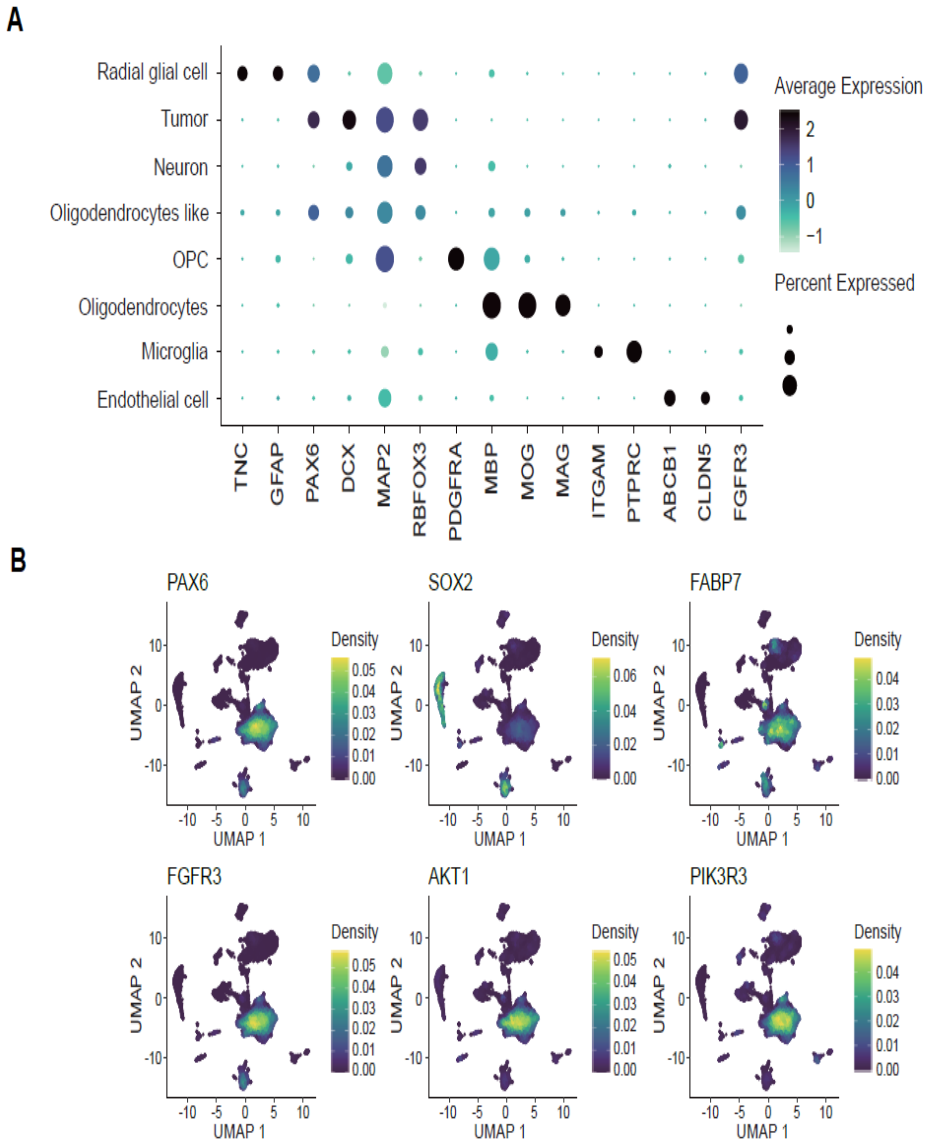

Figure 13. Enriched genes in cell populations identified in CN and normal brain. A. Dot plot for top marker genes in different cell clusters showed *FGFR3* was enriched in both tumor and radial glial cells. B. Bulk RNA-seq CN-specific genes were enriched in the tumor cell cluster in snRNA-seq data.

3.5. Dedifferentiated state of CN tumor cells in between radial glia and neuron

DEGs between the radial glial cell, neuron, and tumor cells in single cell clusters were investigated. Consistent with the bulk RNA-seq findings, the PI3K-AKT pathway was enriched in the snRNA-seq tumor cell cluster (Figure 14).

Trajectory analysis with the three major cell populations of concern (tumor, radial glial cell, and neuron) was conducted to see if there is any pseudotime point difference in these populations considering cell differentiation status. Indeed, it was observed that the tumor cells were identified as in between the radial glial and neuron population, indicating that the CN tumor cells were more differentiated than the radial glial cell but not as much as the neuron (Figure 15). To find out the possible pathways that could play an active role driving radial glial cells to deviate from their natural differentiation course to become CN tumor cells, Gene Ontology analyses were done. Genes rarely expressed in tumor clusters but differentially expressed in the radial glial cell, and neuron clusters were investigated. As a result, it was observed that neuron-related genes (*CNTN1, FAT3, PTPRG, NTRK2*, etc.), annotated to categories such as nervous system development, neuron differentiation, generation of neurons, and neurogenesis, were downregulated in the tumor cluster cells (Figure 16).

Additionally, tumor and normal cells in radial glial cell cluster grouped together when depicted with UMAP with all cell types, but they clustered separately in pseudotime analysis. This indicated that there might be some differences between these cells despite

belonging to the broader radial glial cell class. In order to find the genes that are differently expressed in tumor and normal samples in the radial glial cell cluster, a principal component analysis was performed with only the radial glial cell type, and each cell was divided by PC2 (Figure 17A). As a result, it was possible to identify genes (*CNTN1, FAT3*, *EPHA3, EPHA5, NLGN1*, etc.) that contributed to PC2, which expressed differently in the radial glial cells between normal and tumor samples (Figure 17B). These results indicated that CN mainly consists of radial glial-like cells that deviated from its natural course of radial glial to neuron differentiation and developed into tumor cells with the dysregulation of several significant pathways essential for normal radial glial to neuron differentiation.

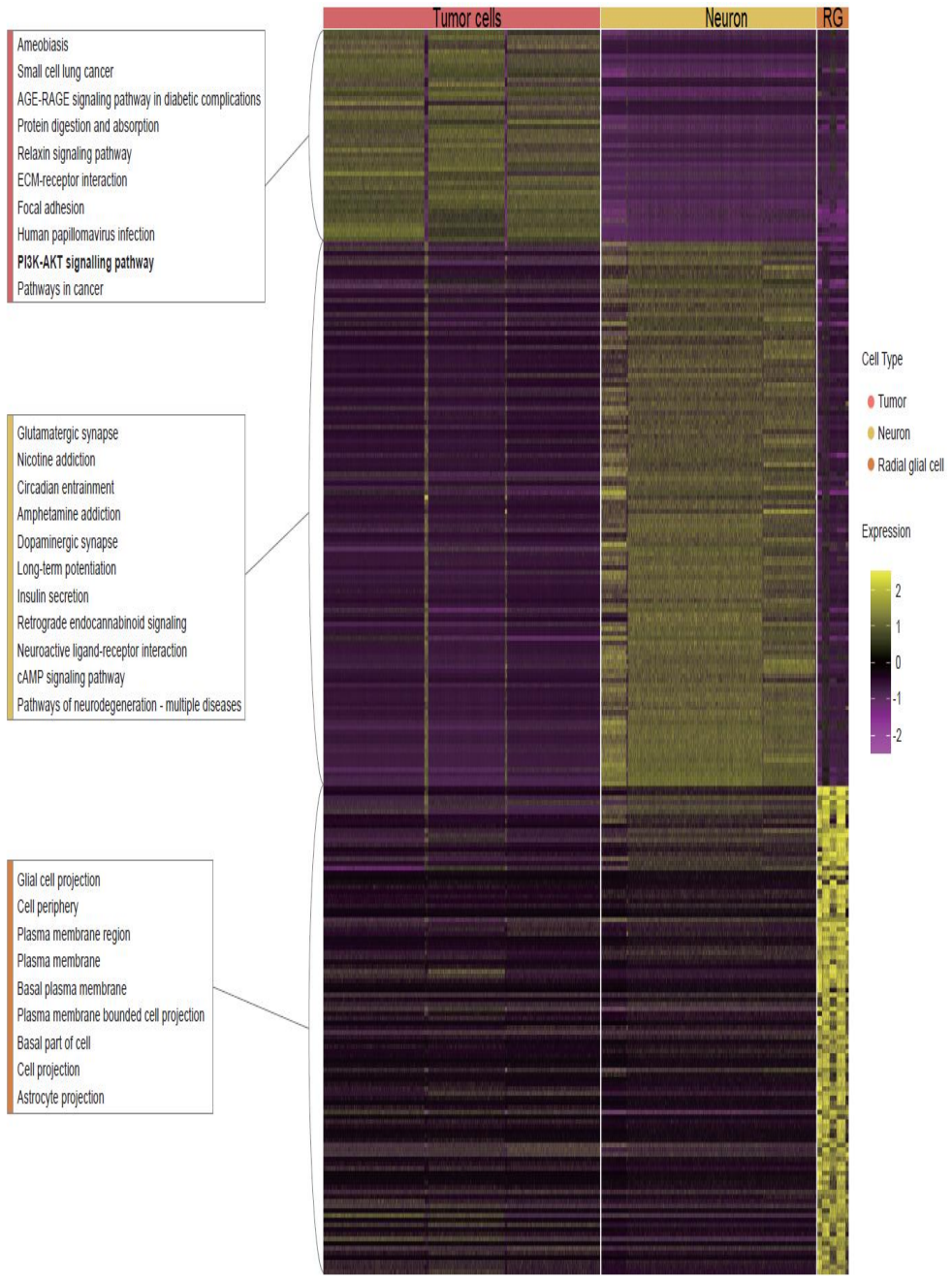

Figure 14. CN tumor cells show PI3K-AKT pathway enrichment. Gene ontology analysis with the tumor, neuron and radial glial cell populations from snRNA-seq data shows enrichment of the PI3K-AKT pathway in the CN tumor cell population, complimenting the bulk RNA-seq results.

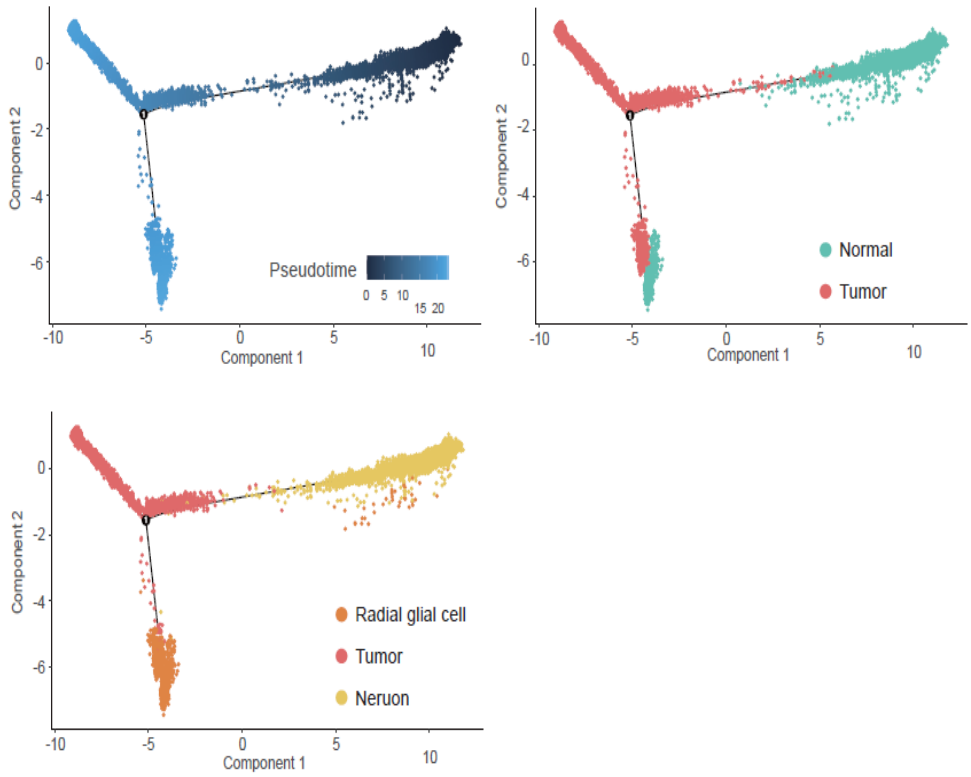

Figure 15. Trajectory analysis shows CN tumor cells in between radial glial and neuron cells. PCA plot with the pseudotime data shows CN tumor cells are more differentiated than radial glial cells but less differentiated than the neurons.

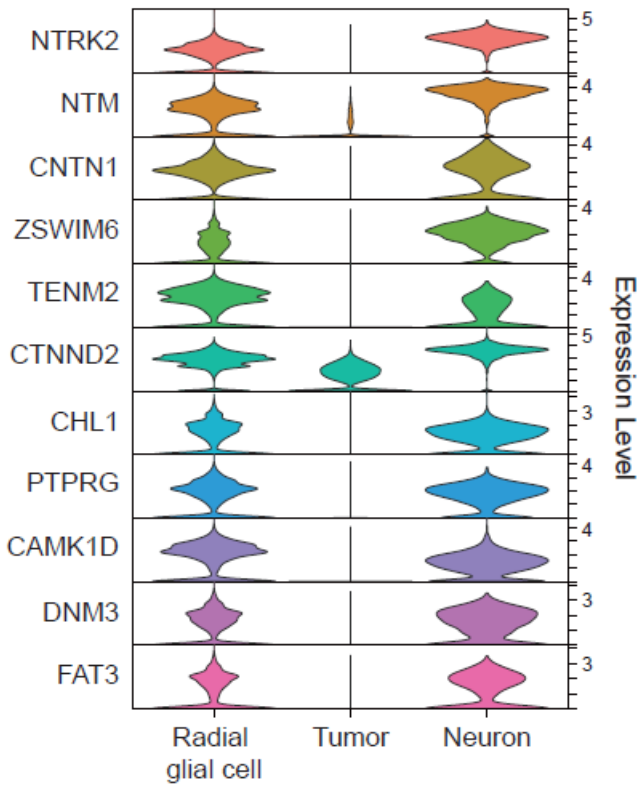

Figure 16. Genes essential for nervous system development, neuron differentiation are not expressed in CN tumor cells. Gene ontology analysis of tumor, neuron, and radial glial cells from snRNA-seq data shows genes such as *CNTN1, FAT3, PTPRG,* etc., are not expressed in the CN tumor cells.

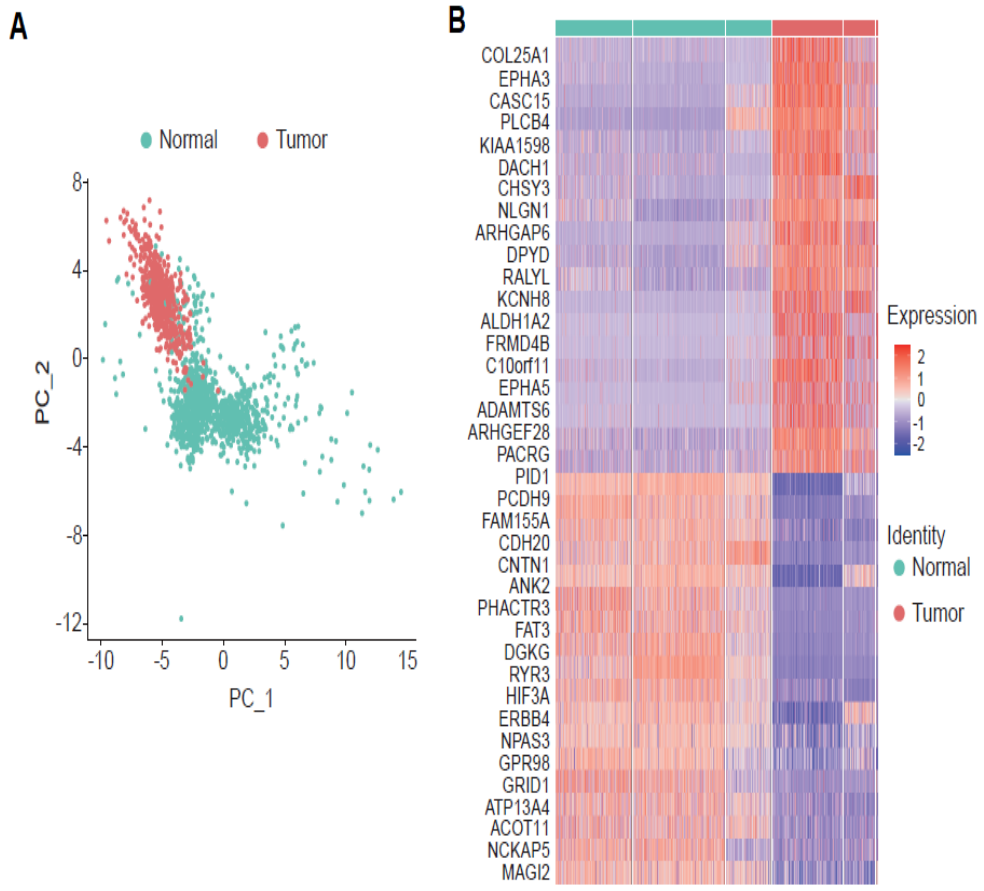

Figure 17. Radial glial cells from tumor and normal sample had fine difference. A. PCA shows separate clustering of tumor and normal sample cells within the radial glial cell cluster. B. Heatmap with key differential genes between normal and tumor sample radial glial cells.

3.6. Epigenetic characteristics of CN

The epigenetics profile of the CN was explored through methylation sequencing. To confirm the DNA methylation-based classification of the CN, 8,535 common probes from the targeted bisulfite sequencing data were matched with the most variable 10,000 core probes from 450K methyl array data which were selected to classify the CNS tumors by DFKZ classifier[18]. All the cases in this study were perfectly clustered with DKFZ CN groups in t-SNE analysis (Figure 18).

To verify whether the changes in methylation patterns affected the gene expression, the DNA methylation level of CpGs was investigated. If a DEG with more than 10 CpGs showed more than half differentially methylated regions (DMRs), it was selected. First, 167 DEGs were shortlisted with implemented DMR criteria; then, the top 10 genes were selected in order of the highest DMR ratio. *FGFR3* was the most differentially methylated gene (Figure 19A). Of the 46 probes in *FGFR3*, 42 CpGs were on the DKFZ highest variable 10K list, and as many as 76% of CpGs in *FGFR3* were hypo-methylated in CN (Figure 19B). Then the methylation level of the *FGFR3* genes within the DKFZ CNS tumors was checked to see if *FGFR3* hypomethylation is a CN-specific event. The 3 most variable probes (cg00525145, cg09145949, cg11777917) that differed by more than 0.8 in methylation ratios compared to the normal were selected (Figure 20A). Furthermore, methylation of the 3 most variable probes has a strong negative correlation with *FGFR3* expression (Figure 20B).

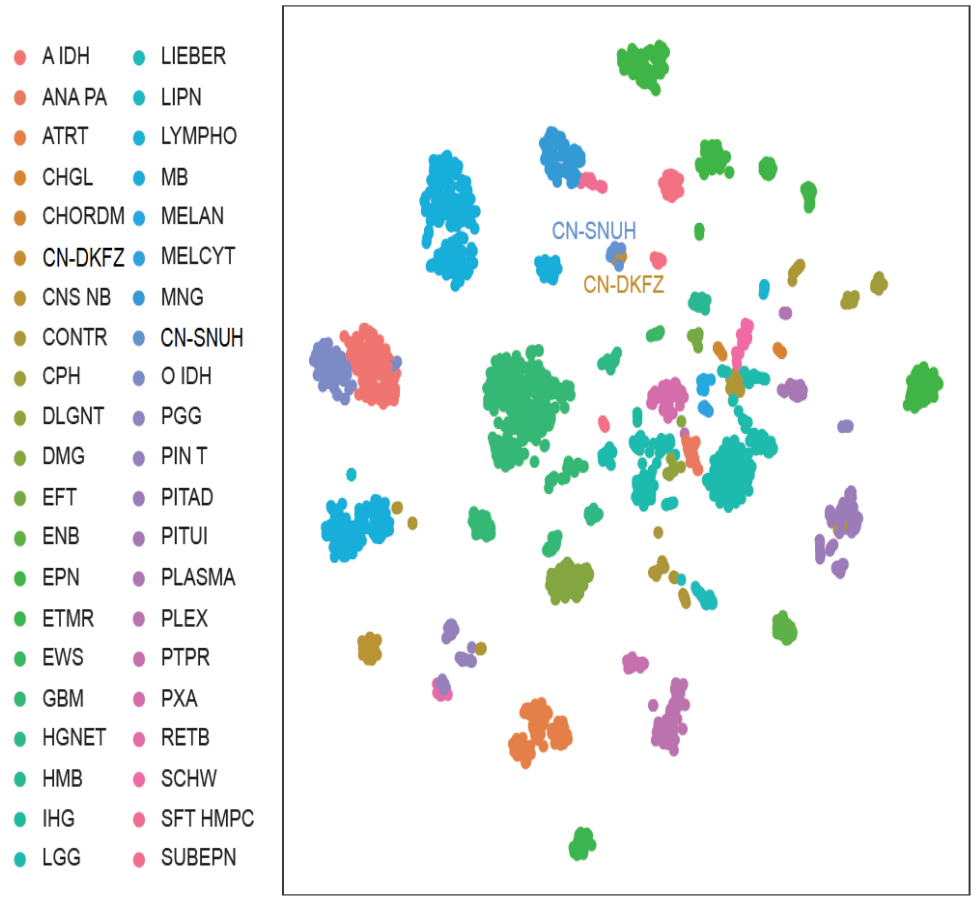

Figure 18. CN classification confirmation by methylation profiling. t-SNE clustering based on the methylation data showing CN samples from this study clustering exclusively with the public CNS data CN cluster confirming the accurate classification of the samples as CN. CN-SNUH represents samples from the current study (n=6), and CN-DKFZ represents samples from the comparing published methylation data (n=21)

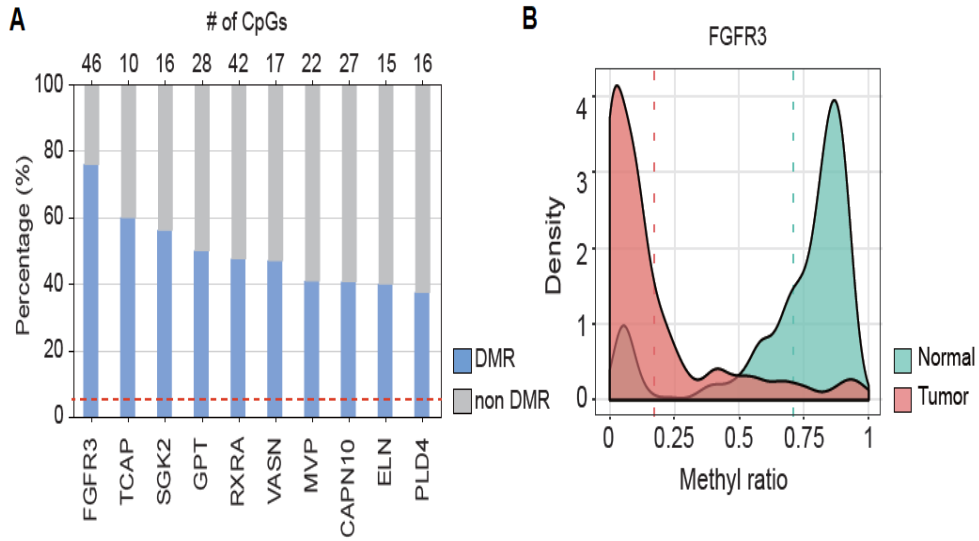

Figure 19. Significant methylation pattern of *FGFR3* in CN. A. Majority of *FGFR3* CpGs was DMRs. The *FGFR3* gene shows the highest percentage among the DMR/non-DMR ratio values, B. *FGFR3* shows hypomethylation compared to the normal brain samples in the methylation expression data.

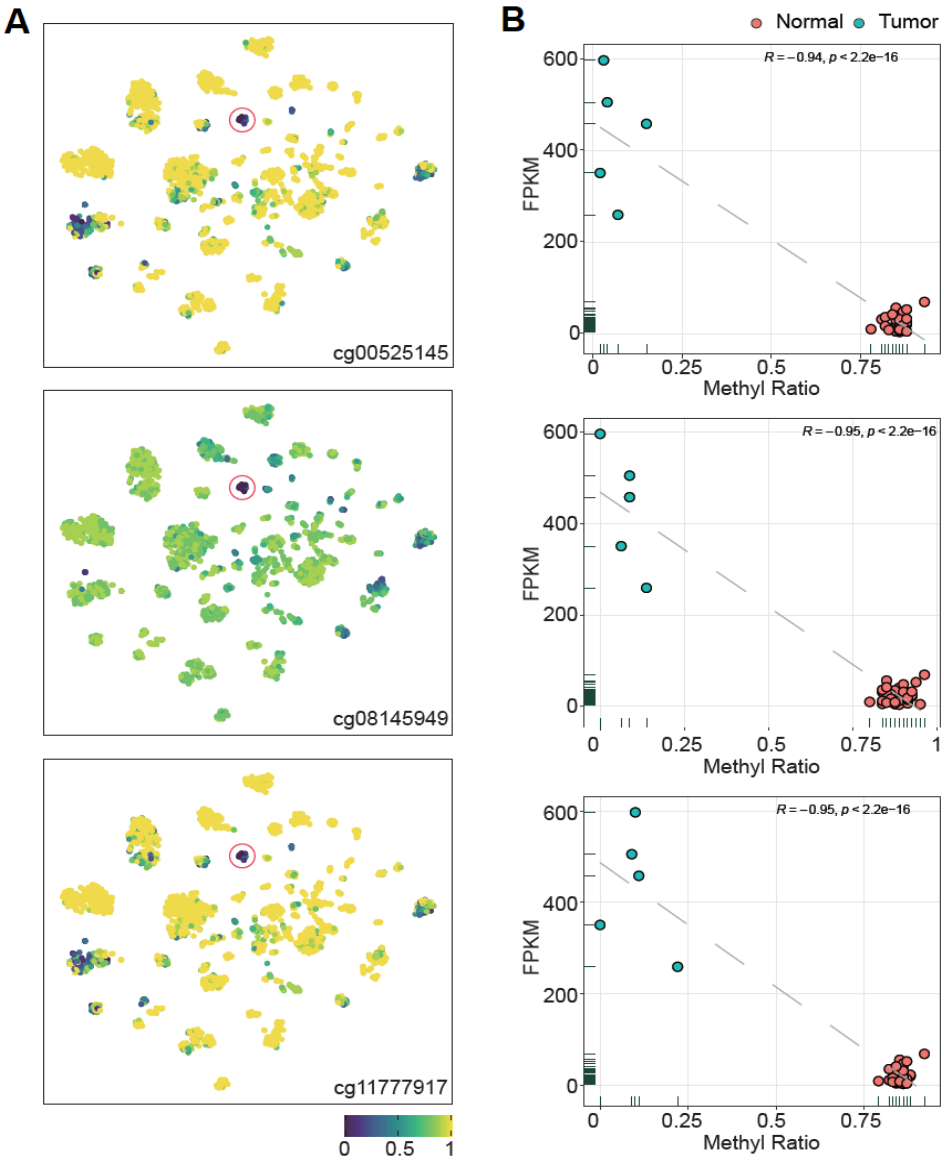

Figure 20. CN specific top 3 CpG probes of *FGFR3*. A. *FGFR3* specific CpG probes (n=3) showing significant hypomethylation compared to other CNS tumors and normal brain samples. CN samples indicated by red circles contain 6 samples from the current study and 21 samples obtained from the published DFKZ CNS methylation data. B. Correlation plot of the methylation ratios with FPKM values shows a significant negative correlation of all three *FGFR3* specific probes.

3.7. Telomere maintenance mechanism in CN

None of the CN cases in the present study had any mutations in *TERT* promoter or *ATRX,* which are frequently observed in gliomas. To evaluate the active telomere maintenance mechanism (TMM) in CN, first, the *TERT* expression was checked only to find that none of the CN samples expressed *TERT* (Figure 21A). Telomerase activity was measured in CN tissues by PCR-ELISA, and all the CN tissues showed very low telomerase activity (Figure 21B). In addition, all the CN samples were negative for c-circle, implying there is no alternative lengthening of telomeres (ALT) phenomenon in CN (Figure 21C). The measurement of mean telomere length of CN by southern blot result ranges from 9 to 13kb on average (Figure 21D). The detection range for telomere length was between 20-30kb on average. The telomere length in tumors did not vary much according to the age of the patients and also showed an intertumoral homogenous pattern (Table 11). In previous studies, telomere length for the normal brain has been reported as more than 10kb compared to other astroglial brain tumors (<10kb)[19]. Taken together, the TMM in CN is not actively operated, implying its benign nature of biological behavior.

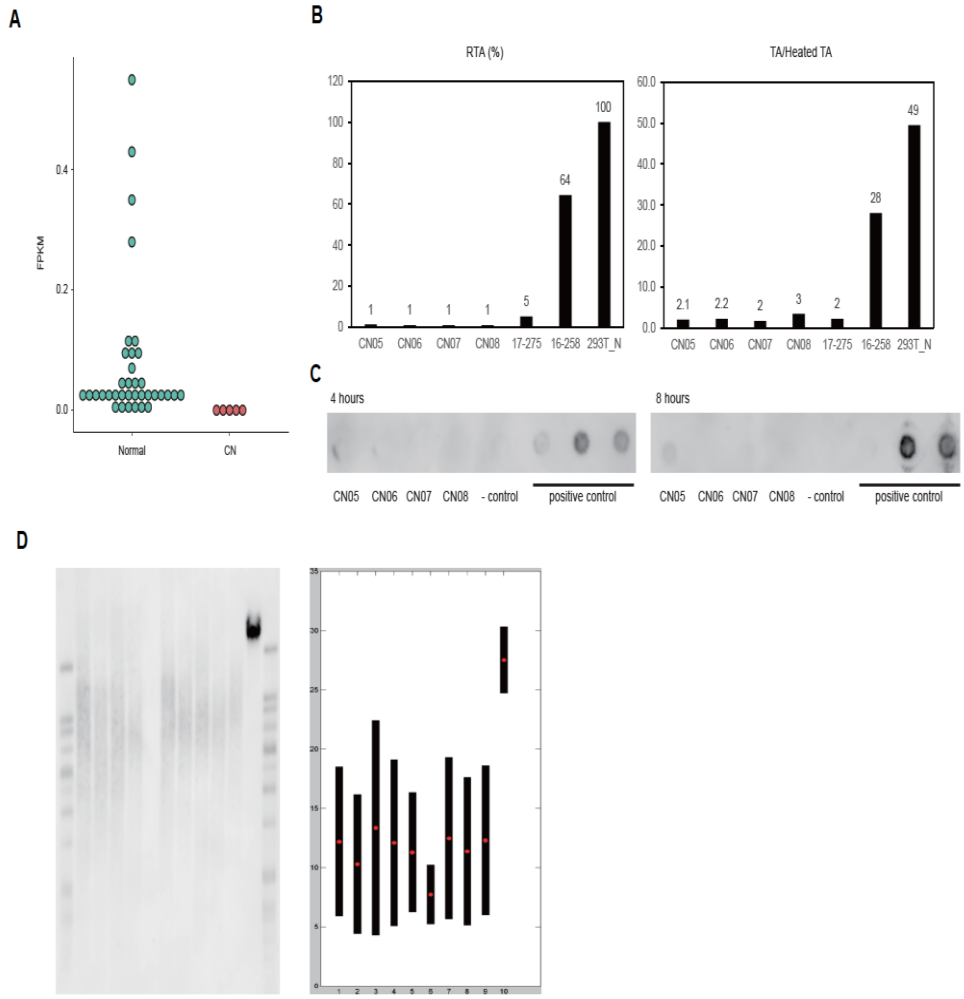

Figure 21. Telomere maintenance mechanism of CN. A. *TERT* gene expression was absent in all the CN samples compared to normal brain samples. B. TRAP assay showing minimal telomerase activity in CN compared to the positive control. C. C-circle assay showing negative results for alternate telomere activity (ALT). D. TRF assay showing a mean telomere length between 9-13 kb in CN.

Table 11. Telomere length measurement of CN

Sample name	CN05	CN06	CN24	CN25	Cut Control	Uncut Control
Max Intensity	10.7	7.07	8.05	6.73	8.98	28.28
Mean TRF	11.28	7.71	12.47	11.37	12.29	27.51
SD TRF	4.28	2.11	5.8	5.3	5.34	2.39
Range TRF	17.11	8.46	23.18	21.19	21.35	9.56
Fit Quality	86.4	83.68	72.7	64.92	73.75	93.59

3.8. Potential ontogeny of CN tumorigenesis

Combining the results of this study, tumorigenesis of CN is thought to be originated from the aberrant radial glial cells in the SVZ harboring overexpression of *FGFR3* caused by hypomethylation of its CpG sites. Overexpression of *FGFR3* in these cells might have perturbed the abnormal neural development pathways and activated the oncogenic PI3K-AKT pathway along with downregulation of several essential pathways in normal neuronal differentiation and migration (Figure 22). Downregulation of *CNTN1, FAT3* genes in CN tumor cells which are essential for normal CNS development and neuron differentiation, also seemed to aid the tumorigenesis process. The culmination of all these events leads the radial glial-like cells to have deviated from their natural course of differentiation and migration into mature neurons and instead end up inside the ventricles in a limbo where they develop to become CN. With these supporting results, it can be affirmed that *FGFR3* plays a crucial role in CN tumorigenesis.

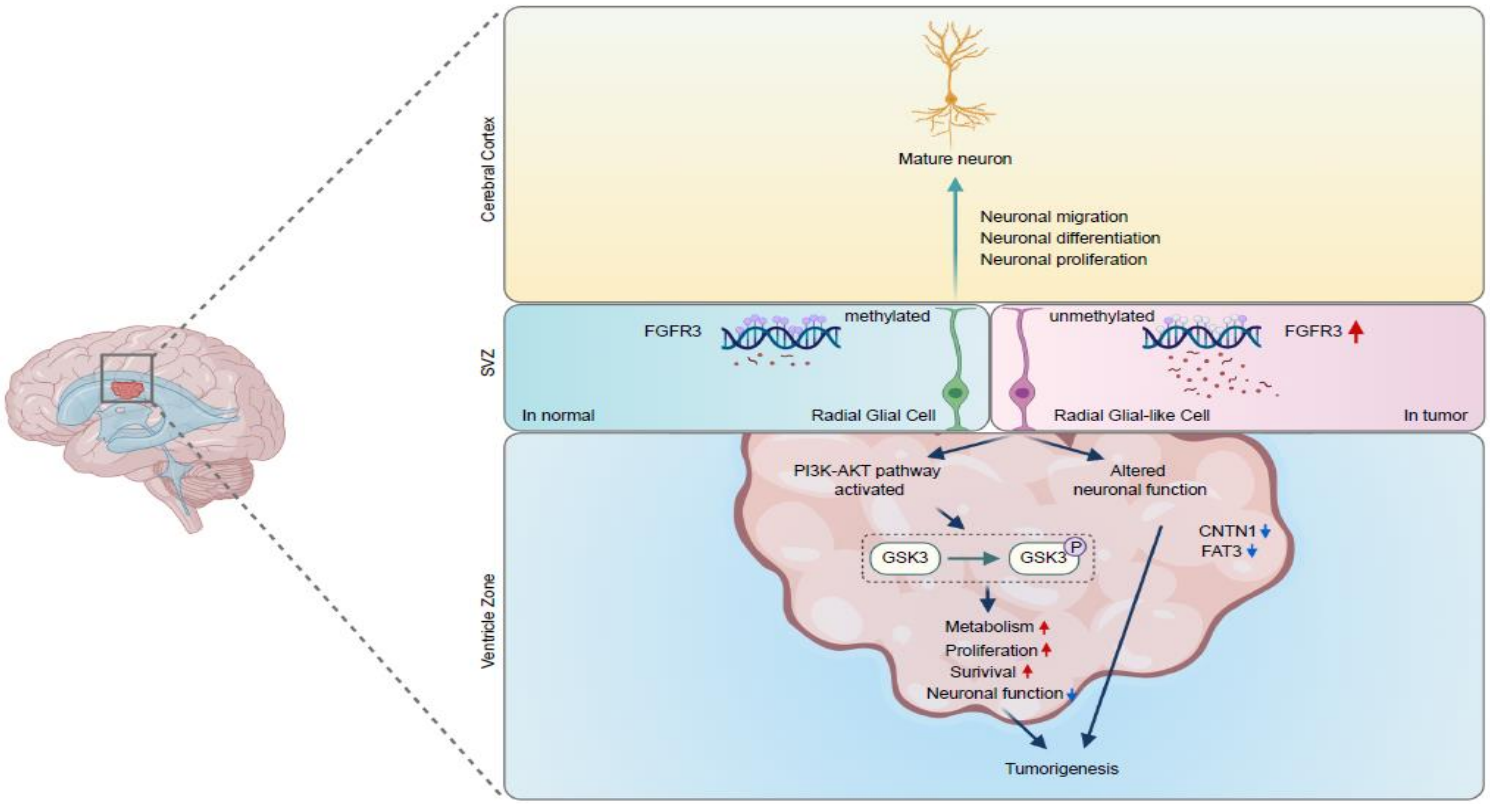
Cerebral Cortex
Mature neuron
Neuronal migration
Neuronal differentiation
Neuronal proliferation
SVZ
FGFR3
methylated
In normal
Radial Glial Cell
unmethylated
FGFR3
Radial Glial-like Cell
In tumor
Ventricle Zone
PI3K-AKT pathway activated
Altered neuronal function
GSK3
GSK3
P
CNTN1
FAT3
Metabolism
Proliferation
Surivival
Neuronal function
Tumorigenesis

Figure 22. Schematic representation of CN tumorigenesis. The Middle left, and upper panels show the normal course for radial glial cell differentiation into mature neurons. The middle right and lower panel show the potential course for CN tumorigenesis where *FGFR3* hypomethylation driven *FGFR3* overexpressed radial glial-like cells proceed off course inside the ventricle as a result of PIK3-AKT activation and altered neuronal functions such as differentiation, and migration.

Chapter 4. Discussion

Central neurocytoma is a unique entity of CNS tumor in terms of its typical intraventricular location and unique histopathological characteristics of both glial and neuronal origin. CN was included in the WHO classification of CNS tumors as a separate entity in 1993[8] and still holds its position in the present 2021 WHO classification under the category of glioneuronal and neuronal tumors[20-21]. The biological behavior of CN is reported to be mostly benign, but some aggressive types also exists[22-23]. The origin of CN was first proposed in 1991 by von Deimling et al. mentioning it is from the bipotential progenitor cells of the subventricular zone in the adult brain[9]. Later studies showed the dual characteristics of both neuronal and glial cells in the in vitro primary culture of CN[6,24]. However, due to the rarity of these tumors, the research related to CN is limited, especially when it comes to comprehensive genomic studies using the latest technology. In this study, the genomic profiling of CN was explored with next-generation sequencing technologies such as WES, bulk and single nuclei RNA-seq, and Methyl-seq to understand the key elements behind CN development.

It was confirmed that CN did not harbor any recurrent somatic mutations or copy number variant regions, or gene fusions driving its tumorigenesis. This result is consistent with the failure to find a universal driver gene of CN in previous studies[10-12]. Next, focusing on the transcriptomics data and gene ontology pathway analysis with the DEGs in the present study, it was identified that the altered PI3K-AKT pathway as a consistent candidate for CN tumorigenesis. Top DEGs included *FGFR3*, *PIK3R3*, and *AKT1*, all of

which play a major role in the activation of the PI3K-AKT pathway and oncogenesis[25-26]. Among the downstream genes of the PI3K-AKT pathway, genes related to GSK3, notably neuronal function-related genes, were all significantly downregulated, which is consistent with the previously published studies[27-28]. Aberrant activities of the neural developmental pathways were also found, and pathway analysis also revealed the upregulation of the neural precursor cell proliferation pathway. After checking the marker genes expression related to the neuronal cell stages, it was confirmed and validated that radial glial cell-related marker genes *PAX6*, *SOX10,* and *FABP7*, were significantly upregulated in the CN tissue samples. Studies with CN tumor spheres showed phenotypic similarity of CN tumor cells with radial glial cells and neural progenitor cells, suggesting that CN cells might have originated from the radial glial cells located in the adult SVZ and SGZ[11,29].

It was also noticed that pathways related to neurodevelopment, such as neuron development, neuron differentiation, generation of neurons, and neurogenesis, were downregulated in CN. In a previous study, it had been established that neurogenesis occurs in the SVZ and SGZ of the brain[30]. During this process, neuroepithelial cells get differentiated from radial glial cells. These cells can be differentiated into both neuronal cells and glial cells depending upon the environmental cues. Radial glial cells in the SVZ follow the rostral migratory system and migrate towards the olfactory bulb and ultimately to the frontal cortex to give rise to either neurons or glial cells[31]. Besides, radial glial cells in the SGZ follow a short path from the dentate gyrus towards the hippocampus[31]. The neurodevelopmental pathways play a crucial role during these migration processes to guide the cells toward the right

path[32]. Downregulation in these pathways might have caused the radial glial cells to deviate from their original course and migrate more centrally into the ventricles. This hypothesis was inspected through the snRNA-seq analysis in the current study, where genes related to neuronal migration and axon guidance (*CNTN1, PTPRG, NTRK2*, *FAT3,* etc.) were clearly observed to be more upregulated in the neuron cell clusters compared to the tumor cell clusters. *CNTN1* genes have been previously reported to play an important role during neuronal migration, and the PTPRG-CNTN signaling was indicated as a critical mechanism for normal neuronal developmental process[33-34]. Interestingly, a single alteration in the *FAT3* gene was also shown to be sufficient to cause fundamental changes that drive CNS evolution in a previous study[73]. Furthermore, this analysis identified several genes that were differentially expressed in the radial glial cells in between the normal and tumor samples, among which Eph receptor-related genes like *EPHA3* and *EPHA5* were expressed in radial glial cells of the tumor sample but not in radial glial cells of the normal brain. This is a compelling finding because the role of the Eph receptor family has been well associated with tumorigenesis in previous research, and crosstalk between the Eph receptor and PI3K-AKT signaling has also been reported[35]. Whether the overexpression of Eph receptor-related genes plays any role in the activation of the PI3K-AKT pathway and tumorigenesis in CN in addition to *FGFR3* is another exciting scope that can be explored in the future.

In previous studies, various oncogenic signaling pathways had been linked with CN development (Table 2). Although this study corroborated the overexpressed genes found in CN in previous studies, pathway analysis revealed different scenarios compared to

the previous studies. PI3K-AKT signaling was established as a major driver pathway in CN tumorigenesis in the current study. It was interesting that previous studies had reported MAPK signaling and WNT signaling as potential drivers of CN because often, in the oncogenic process, crosstalk of MAPK and WNT signaling occurs with PI3K-AKT signaling. Whether this type of complex pathway network is also in play for CN tumorigenesis needs to be further evaluated in the future.

The key finding of the present study is CN-specific hypomethylation of 3 CpG sites of *FGFR3*. This epigenetic change might be the main reason behind the overexpression of *FGFR3* at the transcriptomics level in CN. In one previous study, hypermethylation of death-associated protein kinase promoter had been identified in CN[71]. Single copy gene hypomethylation correlated with expression level has been reported in various types of cancers previously, and this process has been conjectured as one of the possible causes of cancer development initiation[36]. The *FGFR3* overexpression initiates a series of cascades, including abnormal neuronal precursor cell proliferation, angiogenesis, tumor cell migration, differentiation, survival, and also activating pathways like PI3K-AKT. The activation of *FGFR3* pathway has also been reported to exert a potent impact on cortical and hippocampal lamination, brain size, neuronal differentiation, and axonal pathfinding[37-38]. Additionally, FGFR gene family alterations have been reported frequently in low-grade neuroepithelial tumors[39]. Moreover, *FGFR3* upregulation was also found in previous CN studies[11]. In this study, CN tumorigenesis was found to be initiated by epigenetic changes rather than genetic aberrations. Interestingly another well-known intraventricular CNS tumor, ependymoma (EPN), of which subtypes also lack genetic

mutation, is also linked with epigenetic changes as its driver of oncogeneis[40-41]. But unlike CN, EPN is a heterogeneous tumor divided by various subcategories according to a recent single-cell RNA-seq study[40]. It is reported that one of the subcategories of EPN driven by *C11orf95-RELA* or *C11orf95-YAP* fusions were shown to have enriched in radial glial-like cells with aberrant neurodevelopmental pathways and *FGFR3* overexpression followed by epigenetic changes[40]. Although evidence of any such fusions in CN was not found in the present study, the similarities of epigenetically influenced *FGFR3* overexpression in both EPN and CN are intriguing and might uncover new mechanisms of brain tumor development with future research. Furthermore, promising results had been showed in controlling EPN tumor cell proliferation with FGFR targeted therapies in vitro[40]. This corroborative evidence from the EPN offers the potential implication of anti-*FGFR3* therapy for CN.

Based on the discoveries of this study, CN tumorigenesis is thought to originate from the aberrant radial glial cells in the SVZ, in which *FGFR3* overexpression is caused by hypomethylation of CpG sites in the gene emerges as an important driver. The overexpression of *FGFR3* in these cells might perturb abnormal neural development pathways and activate the oncogenic PI3K-AKT pathway, in addition to downregulating several essential pathways involved in normal neuronal differentiation and migration (Figure 22). The culmination of all these events leads radial glial-like cells to deviate from their natural course of differentiation and migration in which they become mature neurons and instead end up inside ventricles in a state from which they develop into CN. With these supporting results, it can be affirmed that *FGFR3* plays a crucial role

in CN tumorigenesis.

Despite its limited sample size, this study adds a significant amount of new data to enhance current knowledge of the very rare CNS tumor CN. It also indicates multiple possible future research directions, including but not limited to whole genome studies to further explore the absence of genomic drivers in CN, transcriptomic and single-cell studies with increased CN sample numbers, functional validation of *FGFR3* as the key driver in CN, etc. Furthermore, the specific mechanism by which the hypomethylation of *FGFR3* drives radial glial cells to deviate from their natural differentiation course to instead give rise to CN inside ventricles will require further exploration via large-scale methylation studies of CN in the future.

Chapter 5. Conclusion

To conclude, this study established an epigenetic cause of tumorigenesis without gene mutation in CN and identified the PI3K-AKT pathway as a key oncogenic pathway. The hypomethylation of the *FGFR3* promoter and its overexpression have the promising potential to be considered a biomarker and treatment target in CN. Although this study provides significant results to uncover the CN tumorigenesis process, it still had some limitations when it comes to sampling size. Future large-scale genomic and epigenomic studies with increased sample numbers will be advantageous in identifying other potential causes that might aid CN development and progression.

References

1. Choudhari KA, Kaliaperumal C, Jain A, Sarkar C, Soo MY, Rades D, et al. Central neurocytoma: a multi-disciplinary review. Br J Neurosurg. 2009;23(6):585-95.
2. Yang I, Ung N, Chung LK, Nagasawa DT, Thill K, Park J, et al. Clinical manifestations of central neurocytoma. Neurosurgery clinics of North America. 2015;26(1):5-10.
3. Lee SJ, Bui TT, Chen CH, Lagman C, Chung LK, Sidhu S, et al. Central Neurocytoma: A Review of Clinical Management and Histopathologic Features. Brain tumor research and treatment. 2016;4(2):49-57.
4. Kim CY, Kim DG, Joo JD, and Kim YH. Clinical outcome and quality of life after treatment of patients with central neurocytoma. Neurosurgery clinics of North America. 2015;26(1):83-90.
5. Hassoun J, Gambarelli D, Grisoli F, Pellet W, Salamon G, Pellissier JF, et al. Central neurocytoma. An electron-microscopic study of two cases. Acta Neuropathol. 1982;56(2):151-6.
6. Ishiuchi S, and Tamura M. Central neurocytoma: an immunohistochemical, ultrastructural and cell culture study. Acta Neuropathol. 1997;94(5):425-35.
7. Tsuchida T, Matsumoto M, Shirayama Y, Imahori T, Kasai H, and Kawamoto K. Neuronal and glial characteristics of central neurocytoma: electron microscopical analysis of two cases. Acta Neuropathol. 1996;91(6):573-7.
8. Kim DG, and Park CK. Central neurocytoma: establishment of the disease entity. Neurosurgery clinics of North America. 2015;26(1):1-4.
9. von Deimling A, Kleihues P, Saremaslani P, Yasargil MG, Spoerri O, Sudhof TC, et al. Histogenesis and differentiation potential of central neurocytomas. Lab Invest. 1991;64(4):585-91.
10. Sander C, Wallenborn M, Brandt VP, Ahnert P, Reuschel V, Eisenloffel C, et al. Central neurocytoma: SNP array analyses, subtel FISH, and review of the literature. Pathol Res Pract.

2019;215(7):152397.
11. Sim FJ, Keyoung HM, Goldman JE, Kim DK, Jung HW, Roy NS, et al. neurocytoma is a tumor of adult neuronal progenitor cells. J Neurosci. 2006;26(48):12544-55.
12. Vasiljevic A, Champier J, Figarella-Branger D, Wierinckx A, Jouvet A, and Fevre-Montange M. Molecular characterization of central neurocytomas: potential markers for tumor typing and progression. Neuropathology. 2013;33(2):149-61.
13. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature. 2013;499(7457):214-8.
14. Manning BD, and Cantley LC. AKT/PKB signaling: navigating downstream. Cell. 2007;129(7):1261-74.
15. Velmeshev D, Schirmer L, Jung D, Haeussler M, Perez Y, Mayer S, et al. Single-cell genomics identifies cell type-specific molecular changes in autism. Science. 2019;364(6441):685-9.
16. Couturier CP, Ayyadhury S, Le PU, Nadaf J, Monlong J, Riva G, et al. Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. Nat Commun. 2020;11(1):3406.
17. Han X, Zhou Z, Fei L, Sun H, Wang R, Chen Y, et al. Construction of a human cell landscape at single-cell level. Nature. 2020;581(7808):303-9.
18. Capper D, Jones DTW, Sill M, Hovestadt V, Schrimpf D, Sturm D, et al. DNA methylation-based classification of central nervous system tumours. Nature. 2018;555(7697):469-74.
19. La Torre D, Conti A, Aguennouz MH, De Pasquale MG, Romeo S, Angileri FF, et al. Telomere length modulation in human astroglial brain tumors. PLoS One. 2013;8(5):e64296.
20. Louis DN, Perry A, Wesseling P, Brat DJ, Cree IA, Figarella-Branger D, et al. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol. 2021;23(8):1231-51.
21. Louis DN, Perry A, Reifenberger G, von Deimling A,

Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol. 2016;131(6):803-20.
22. Sun S, Malicki DM, Levy ML, and Crawford JR. Atypical central neurocytoma with aggressive features in a child. BMJ Case Rep. 2020;13(6).
23. Goyal S, Kataria T, Gupta D, Dhyani A, Mohapatra I, and Narang KS. Atypical central neurocytoma with leptomeningeal dissemination: a case report. J Egypt Natl Canc Inst. 2020;32(1):23.
24. Ishiuchi S, Nakazato Y, Iino M, Ozawa S, Tamura M, and Ohye C. In vitro neuronal and glial production and differentiation of human central neurocytoma cells. J Neurosci Res. 1998;51(4):526-35.
25. Jiang N, Dai Q, Su X, Fu J, Feng X, and Peng J. Role of PI3K/AKT pathway in cancer: the framework of malignant behavior. Mol Biol Rep. 2020;47(6):4587-629.
26. Li L, Zhang S, Li H, and Chou H. FGFR3 promotes the growth and malignancy of melanoma by influencing EMT and the phosphorylation of ERK, AKT, and EGFR. BMC Cancer. 2019;19(1):963.
27. Hur EM, and Zhou FQ. GSK3 signalling in neural development. Nat Rev Neurosci. 2010;11(8):539-51.
28. Salcedo-Tello P, Ortiz-Matamoros A, and Arias C. GSK3 Function in the Brain during Development, Neuronal Plasticity, and Neurodegeneration. Int J Alzheimers Dis. 2011;2011:189728.
29. Shin HY, Kim JW, Paek SH, and Kim DG. The characteristics of neuronal stem cells of central neurocytoma. Neurosurgery clinics of North America. 2015;26(1):31-6.
30. Ming GL, and Song H. Adult neurogenesis in the mammalian brain: significant answers and significant questions. Neuron. 2011;70(4):687-702.
31. Ghashghaei HT, Lai C, and Anton ES. Neuronal migration in the adult brain: are we there yet? Nat Rev Neurosci. 2007;8(2):141-51.
32. Russell SA, and Bashaw GJ. Axon guidance pathways and the control of gene expression. Dev Dyn. 2018;247(4):571-80.

33. Chen YA, Lu IL, and Tsai JW. Contactin-1/F3 Regulates Neuronal Migration and Morphogenesis Through Modulating RhoA Activity. Front Mol Neurosci. 2018;11:422.
34. Nikolaienko RM, Hammel M, Dubreuil V, Zalmai R, Hall DR, Mehzabeen N, et al. Structural Basis for Interactions Between Contactin Family Members and Protein-tyrosine Phosphatase Receptor Type G in Neural Tissues. J Biol Chem. 2016;291(41):21335-49.
35. Anderton M, van der Meulen E, Blumenthal MJ, and Schafer G. The Role of the Eph Receptor Family in Tumorigenesis. Cancers (Basel). 2021;13(2).
36. Hoffmann MJ, and Schulz WA. Causes and consequences of DNA hypomethylation in human cancer. Biochem Cell Biol. 2005;83(3):296-321.
37. Inglis-Broadgate SL, Thomson RE, Pellicano F, Tartaglia MA, Pontikis CC, Cooper JD, et al. FGFR3 regulates brain size by controlling progenitor cell proliferation and apoptosis during embryonic development. Dev Biol. 2005;279(1):73-85.
38. Huang J-Y, Krebs BB, Miskus ML, Russell ML, Duffy EP, Graf JM, et al. Enhanced FGFR3 activity in post-mitotic principal neurons during brain development results in cortical dysplasia and axon miswiring. bioRxiv. 2020.
39. Bale TA. FGFR- gene family alterations in low-grade neuroepithelial tumors. Acta Neuropathol Commun. 2020;8(1):21.
40. Gojo J, Englinger B, Jiang L, Hubner JM, Shaw ML, Hack OA, et al. Single-Cell RNA-Seq Reveals Cellular Hierarchies and Impaired Developmental Trajectories in Pediatric Ependymoma. Cancer Cell. 2020;38(1):44-59 e9.
41. Versteeg R. Cancer: Tumours outside the mutation box. Nature. 2014;506(7489):438-9.
42. Jaffe AE, Hoeppner DJ, Saito T, Blanpain L, Ukaigwe J, Burke EE, et al. Profiling gene expression in the human dentate gyrus granule cell layer reveals insights into schizophrenia and its genetic risk. Nat Neurosci. 2020;23(4):510-9.
43. Li H, and Durbin R. Fast and accurate short read alignment

with Burrows-Wheeler transform. Bioinformatics. 2009;25(14):1754-60.
44. Picard toolkit. Broad Institute, GitHub repository. 2019.
45. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9.
46. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20(9):1297-303.
47. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Res. 2019;47(D1):D941-D7.
48. Karczewski KJ, Weisburd B, Thomas B, Solomonson M, Ruderfer DM, Kavanagh D, et al. The ExAC browser: displaying reference data information from over 60 000 exomes. Nucleic Acids Res. 2017;45(D1):D840-D5.
49. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alfoldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581(7809):434-43.
50. Wang K, Li M, and Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010;38(16):e164.
51. Kwak SH, Chae J, Choi S, Kim MJ, Choi M, Chae JH, et al. Findings of a 1303 Korean whole-exome sequencing study. Exp Mol Med. 2017;49(7):e356.
52. Talevich E, Shain AH, Botton T, and Bastian BC. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. PLoS Comput Biol. 2016;12(4):e1004873.
53. O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature. 2012;485(7397):246-50.
54. Jaffe AE, Straub RE, Shin JH, Tao R, Gao Y, Collado-Torres L, et al. Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. Nat Neurosci.

2018;21(8):1117-25.
55. BrainSeq: Neurogenomics to Drive Novel Target Discovery for Neuropsychiatric Disorders. Neuron. 2015;88(6):1078-83.
56. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
57. Li B, and Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323.
58. Love MI, Huber W, and Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
59. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. 2019;47(W1):W191-W8.
60. Haas BJ, Dobin A, Li B, Stransky N, Pochet N, and Regev A. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. Genome Biol. 2019;20(1):213.
61. Zheng GX, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. Nat Commun. 2017;8:14049.
62. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, 3rd, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573-87 e29.
63. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019;16(12):1289-96.
64. Alquicira-Hernandez J, and Powell JE. Nebulosa recovers single cell gene expression signals by kernel density estimation. Bioinformatics. 2021.
65. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol.

2014;32(4):381-6.
66. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods. 2017;14(10):979-82.
67. Qiu X, Hill A, Packer J, Lin D, Ma YA, and Trapnell C. Single-cell mRNA quantification and differential analysis with Census. Nat Methods. 2017;14(3):309-15.
68. Langmead B, and Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9(4):357-9.
69. Krueger F, and Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics. 2011;27(11):1571-2.
70. Van Der Maaten L. Accelerating t-SNE using tree-based algorithms. The Journal of Machine Learning Research. 2014;15(1):3221-45.
71. Chung CL. Differential hypermethylation of death associated protein kinase promoter in central neurocytoma and oligodendroglioma. Biomed Res Int. 2014;2014:506458
72. Henson JD. The C-Circle assay for alternative-lengthening of telomeres activity. Methods.2017,2;114:74-84.
73. Deans, M. R. *et al.* Control of neuronal morphology by the atypical cadherin Fat3. *Neuron* **71**, 820-832, doi:10.1016/j.neuron.2011.06.026 (2011).

국문초록

배경: 중심성 신경세포종은 일반적으로 젊은 성인에게서 많이 발생하는 중추신경계에서 가장 드물게 발생하는 종양 중 하나이다. 중심성 신경세포종의 발생과 분화에 관련하여 통합적인 분자 및 유전자 특성 연구는 현재까지 부재하며 종양발생기전은 확실하게 밝혀지지 않았다.

방법: 중심성 신경세포종의 종양발생기전을 분석하기 위해 동결조직을 이용하여 whole-exome sequencing, bulk RNA sequencing, single nuclei RNA sequencing, methylation sequencing을 시행하여 포괄적인 다중 오믹스 분석(Multi-Omics Analysis)을 진행하였다. 발견된 특정 유전자 발현을 검증하기 위해 별도의 파라핀고정 조직을 이용하여 면역 조직 화학 분석을 시행하였다.

결과: 중심성 신경세포종의 종양 발생 과정에는 *FGFR3* 저메틸화 및 과발현이 주요 인자임을 규명하였고 이는 PI3K/AKT 신호전달 경로 활성화를 유발함을 확인하였다. 또한 중심성 신경세포종과 방사형아교세포의 유전학적 유사성을 분석하여 중심성 신경세포종의 기원이 방사형아교세포의 탈분화 상태에 있는 세포임을 확인하였다.

결론: 중심성 신경세포종의 발생은 방사형아교세포의 뉴런으로의 분화 과정 중 조절 장애에 의한 것으로 추정되었으며 종양 발생 및 진행을 주도하는 요인 중 하나로서 FGFR3의 과별현이 중요함을 규명하였다.

주요어 : 중심성 신경세포종, 방사형아교세포, *FGFR3*, PI3K-AKT

학 번 : 2018-31239