Online Resource 1. Supplementary methods and results

Article: DNA methylation profiling classifies lateral ventricular neurocytomas with lipidized cells as central neurocytoma

Journal: Brain Tumor Pathology

Authors: Takeru Kuwabara, Yoshitaka Oda, Zen-ichi Tanei, Satoshi Tanikawa, Yui Moritsugu, Tomohito Yamauchi, Koki Ise, Yusuke Shirai, Marie Katou, Hirokazu Sugino, Lei Wang, Yusuke Fujiwara, Hiroaki Motegi, Yukitomo Ishi, Shigeru Yamaguchi, Masumi Tsuda, Kaishi Satomi, Junji Shibahara, Miki Fujimura, Shinya Tanaka

Corresponding author: Shinya Tanaka, Department of Cancer Pathology, Faculty of Medicine, Hokkaido University, Kita 15, Nishi 7, Kita-ku, Sapporo 060-8638, Japan. Tel: +81-11-706-7806, Fax: +81-11-706-5902, E-mail: tanaka@med.hokudai.ac.jp

Supplementary methods

Immunohistochemistry

Antibodies and dilutions are listed in Table S1. For all antibodies, heat-induced antigen retrieval was performed at high pH in PT Link (Agilent Technologies). Except for DLK1, sections were stained on an Autostainer Link 48 with the EnVision FLEX detection system (both Agilent Technologies). Sections of Cases 1 and 3 were stained at the Department of Cancer Pathology, Faculty of Medicine, Hokkaido University. GFAP, OLIG2, and Ki-67 of Case 2 were stained at Hokkaido University Hospital, and synaptophysin and NeuN of Case 2 were stained at the Department of Cancer Pathology. Perilipin-1 was stained at Hokkaido University Hospital for all cases.

Table S1. Primary antibodies used for immunohistochemistry

Antigen	Source (catalog no.)	Clone or type	Dilution	Cases
Synaptophysin	Agilent/Dako (M7315)	DAK-SYNAP, mouse monoclonal	1:100	1–3
NeuN	Merck Millipore (MAB377)	A60, mouse monoclonal	1:800	1–3
GFAP	Agilent/Dako	Rabbit polyclonal, ready-to-use	1:2	1, 3
GFAP	Agilent/Dako	Rabbit polyclonal	1:20,000	2
OLIG2	Immuno-Biological Laboratories (18953)	Rabbit polyclonal	1:500	1, 3
OLIG2	Immuno-Biological Laboratories (18953)	Rabbit polyclonal	1:200	2
Ki-67	Agilent/Dako (M7240)	MIB-1, mouse monoclonal	1:200	1–3
Perilipin-1	Cell Signaling Technology (3470)	Rabbit polyclonal	1:100	1–3
DLK1	Proteintech (10636-1-AP)	Rabbit polyclonal	1:1000	1–3

The ready-to-use GFAP reagent used for cases 1 and 3 was further diluted 1:2. DLK1 was stained manually for all cases at the Department of Cancer Pathology with EnVision FLEX reagents. Endogenous peroxidase was blocked with EnVision FLEX Peroxidase-Blocking Reagent for 5 min, sections were incubated with the primary antibody for 30 min and with EnVision FLEX/HRP for 30 min, developed for 10 min with EnVision FLEX DAB+ Chromogen diluted in EnVision FLEX Substrate Buffer, and counterstained with Carrazzi's hematoxylin. Pancreatic islet β cells, which express DLK1 [29, 30], served as the positive control.

For the MIB-1 labeling index, one hot-spot area (1.6–2.4 mm²) was annotated on the whole-slide image of each case in QuPath (v0.7.0), and nuclei within it were detected by positive cell detection on the hematoxylin optical density image (pixel size 0.5 µm, smoothing sigma 2.5 µm, nucleus area 15–400 µm², hematoxylin optical density threshold 0.05). Nuclei with a mean DAB optical density above 0.25 were classified as positive. The detections were checked visually, and the same settings were used for all cases. The labeling

index was 2.1% (407 of 19,182 nuclei) in Case 1, 2.1% (224 of 10,559) in Case 2, and 1.9% (234 of 12,260) in Case 3.

crossNN classification

Two published crossNN models were obtained from Zenodo (record 14006255): the model trained on the Heidelberg reference cohort (Capper_et_al_NN; 91 methylation classes; MD5 369a4a4f116d3c97b7ed59f0907060f2) [2] and the pan-cancer model (pancan_devel_v5i_NN; 178 classes; MD5 4e84331d57d09ac3a27cba0c4ee85d90). Inference used the official implementation (https://gitlab.com/euskirchen-lab/crossNN; the repository has no version tags, so the version is given as commit 26e4db9 of June 17, 2025) [32]. As in a clinical evaluation of crossNN [5], IDAT files were processed individually with the preprocessIllumina function of minfi (v1.58.0) [1]. Probes with a detection P value > 0.01, the threshold given in the minfi documentation as indicating unreliable measurements [9], were treated as missing. Arrays were required to meet the quality criteria of the Heidelberg reference data, a median log2 intensity > 8 in both the methylated and unmethylated channels and a detection P value < 0.05 in ≥ 90% of probes [28]. Detection P values were calculated with the detectionP function of minfi, because the method was not specified but signals were obtained with minfi [28]. For replicate probes on the EPICv2 array, one probe per locus was selected with the hierarchy of Pearce et al. [19], in the following order. (1) Probes with fewer than three beads in either channel were excluded unless all replicates failed. (2) Probes with an ELBAR detection P value < 0.05 (see Unsupervised analysis) were preferred. (3) Probes were prioritized by the annotation of Peters et al. (Rep_results_by_NAME in Additional file 4), in the order “Superior probe”, “Best precision”, any annotation containing “Superior”, “precision”, and “Best”, or by Rep_results_by_SEQUENCE in the same order when all candidates were annotated “Insufficient evidence” [21]. (4) The probe with the lowest detection P value, or the highest mean bead count for ties at zero, was selected. (5) Remaining ties were resolved by manifest order. The same probes were used for crossNN, collapsed to their 450K-compatible identifiers, and beta values were binarized at 0.6 (> 0.6, +1; ≤ 0.6, −1; missing, 0), the empirically determined threshold used for the crossNN training data [32]. Confidence scores (class-standardized outputs of the linear layer converted to probabilities by the softmax function) were calculated with the official implementation, and scores > 0.4 were regarded as a valid classification for array data [32].

Three groups of public samples were processed identically as controls. The first comprised the primary CN profiled on the EPIC array (GSE288337, FFPE) [13], which met the quality criteria and were not used to train the models, after removal of seven duplicate profiles (90 samples, see Unsupervised analysis). The second comprised all 21 CN and 10 LIPN of the Heidelberg reference cohort (GSE90496) [2], which are part of the training data and served to check the input conversion. The third comprised two extraventricular neurocytomas (EVN, GSE152653) [15], both classified as EVN by the Heidelberg classifier v12.8 (calibrated scores 0.99999 and 0.99998).

For retraining, the IDAT files of the 20 sellar region neurocytomas (SELN) in GSE331022 (450K, n = 4, EPIC, n = 5, EPICv2, n = 11, all FFPE) [19] were obtained and processed as above. As in the crossNN study

[32], SELN classification was evaluated by fivefold cross-validation. The SELN were randomly ordered within each array type (seed 12345), concatenated in the order 450K, EPIC, and EPICv2, and assigned in turn to five folds of four samples. For each fold, a model was trained on the 2,801 samples (91 classes) of GSE90496 [2] and the 16 SELN outside the fold, and the four SELN of the fold were classified. A final model trained on the 2,801 samples and all 20 SELN (2,821 samples, 92 classes) was applied to the cases and the controls. Training followed the published crossNN settings (a bias-free linear layer, Adam with a learning rate and weight decay of 1×10^{-4}, batch size 32, 1,000 epochs, and random masking of the same 99.75% of input features for all samples in each mini-batch), and features that were constant across the training samples were removed as in the training script. Missing values were encoded as zero, as described in the crossNN paper and used by the official inference code, which encodes CpGs absent from the input (including those not on the EPICv2 array) as zero [32], rather than converted to −1 (unmethylated) as in the training script. EPIC- and EPICv2-profiled SELN lack some 450K probes, and encoding these as unmethylated could have taught the model platform-specific missingness as a SELN feature. The pan-cancer model was not retrained. Although its training data (pancan_devel_v5i.h5) are available in the same Zenodo record, distinguishing the cases from SELN is a question within CNS tumors, so retraining was limited to the Heidelberg-cohort model, and the pan-cancer model was used for inference only.

Unsupervised analysis

IDAT files were processed individually with SeSAMe (v1.30.1) [33] following the pipeline of Pearce et al. [19]. After the default SeSAMe probe mask, channel inference for Infinium I probes, and nonlinear dye-bias correction (dyeBiasNL), detection P values were calculated with ELBAR (margin = 0.05, capMU = 3000, delta.beta = 0.2, n.windows = 500), and probes with $P \geq 0.05$ were masked before noob background correction. In samples in which ELBAR detected a dichotomous background distribution, noob was applied before ELBAR masking, as described by Pearce et al. [19]. This applied only to two GSE109379 samples. Probes with fewer than three beads in either channel were also masked [19], and replicate EPICv2 probes were selected as described above. All samples were required to meet the quality criteria described above. The GSE90496 reference classes (up to 30 samples per class, the cap used by Pearce et al. [19]) were CN, LIPN, the main differential diagnoses of CN [3, 11, 12, 31] (subependymoma of the supratentorial and posterior fossa compartments, supratentorial ependymoma with *ZFTA* (*RELA*) fusion, subependymal giant cell astrocytoma, dysembryoplastic neuroepithelial tumor, rosette-forming glioneuronal tumor, IDH-mutant and 1p/19q-codeleted oligodendroglioma), and non-neoplastic hemisphere, white matter, and cerebellum. Probes shared by all platforms were retained, probes on the sex chromosomes were removed, and only probes without missing values in any sample were used (no imputation). As in the Heidelberg classifier v12.8 [28], the 10,000 most variable probes were embedded with the umap function of uwot (v0.2.5) with default parameters (n_neighbors = 15, min_dist = 0.01, no principal component analysis, seed 12345) [18]. The embedding was repeated with n_neighbors = 5, 20, and 80, the values compared in the original UMAP report except for 320, which was too large for 345 samples [18]. UMAP was used for visualization only, and quantitative comparisons were made by correlation. Pearson correlations were calculated on beta values of

the 32,000 most variable probes without subtracting probe means, as in the analysis of the Heidelberg reference cohort [2]. Pairwise correlations among the GSE288337 CN served as a reference for the similarity expected among CN processed in a single batch. Because the GSE288337 cohort includes the cohort of Capper et al. [13], duplicate profiles were sought among the 118 CN of GSE288337 and GSE90496 with these correlations in the initial set of 352 samples. Pairs above the upper outer fence of all CN–CN correlations (third quartile + 3 × interquartile range, r = 0.722) [10] were regarded as repeated profiling of the same tumor. Seven pairs (r = 0.851–0.948), each of mutual nearest neighbors, were identified, six between GSE288337 and GSE90496 and one within GSE288337. The six GSE288337 samples and the sample with the higher GEO accession number of the pair within GSE288337 were excluded from all analyses, leaving 90 GSE288337 CN. To account for differences between FFPE and frozen material, a linear model of M values ($\log2((\beta + \varepsilon)/(1 - \beta + \varepsilon))$, $\varepsilon = 10^{-7}$ [19]) with methylation class and material as covariates was fitted to the GSE90496 samples with limma [24], and the estimated material effect was subtracted from all FFPE samples, including the SELN, the GSE288337 CN, and the cases.

Copy-number analysis

Samples on array types other than that of the 450K normal reference were analyzed with annotations created for both array types (CNV.create_anno in conumee2 [4]), using the probes shared by both (449,521 probes for 450K and EPIC and 390,564 for 450K and EPICv2). Because minfi loads each array type separately, the SELN were analyzed by array type. Samples with a noise parameter > 0.6 were excluded [4]. Arm-level log2 ratios were calculated as the mean of segment values weighted by the number of probes, as described for SeSAMe [33] (long arms only for chromosomes 13, 14, 15, 21, and 22). Gains and losses were called at ±0.1, the threshold applied to conumee segments by Eckhardt et al. [7], and focal changes were assessed by visual inspection of the profiles [19]. Because array signal intensities can shift with GC content [6], GC-rich chromosomes were examined. Chromosome 19 is the most gene-dense chromosome and has a high GC content [8], and chromosomes 16, 17, and 22 are, after chromosome 19, the richest in CpG islands [14]. The chromosome 19 log2 ratio (mean of both arms) was therefore correlated with the mean of 17q and 22q. Copy-number profiles generated by the Heidelberg classifier (Epignostix) were reviewed for arm-level changes, focal amplifications, and homozygous deletions.

Comparison of CN and LIPN

The 21 CN and 10 LIPN of the Heidelberg reference cohort (GSE90496, 450K, all FFPE) [2] were compared following Pearce et al. [19]. SeSAMe beta values were converted to M values ($\varepsilon = 10^{-7}$), and probes on the sex chromosomes and probes with missing values were removed (365,413 probes retained). A limma [24] model with tumor type as the only covariate (no intercept) was fitted, and the LIPN − CN contrast was tested. Pearce et al. also included tumor purity and array type, which were not available or not applicable here because all samples were profiled on the 450K array. CpGs with a Benjamini–Hochberg adjusted P value < 0.01 and $|\Delta\beta| > 0.20$ were considered differentially methylated [19]. DMRs were identified with DMRcate [20] using cpg.annotate (fdr = 0.01) and dmrcate (lambda = 1000, C = 2), as in Pearce et al. [19], with the default minimum of two CpGs, and DMRs with $|\Delta\beta| > 0.20$ (difference in the mean beta value of their CpGs)

were retained. Gene-set enrichment (GO and KEGG) was tested separately for regions hypomethylated and hypermethylated in LIPN with missMethyl (gometh and goregion) [17, 22], which accounts for the number of probes per gene, and terms with FDR < 0.05 were considered enriched.

Transcription factor motif enrichment was assessed with monaLisa (calcBinnedMotifEnrR) [16] using the 879 JASPAR2024 vertebrate CORE motifs [23]. All DMRs of each direction were resized around their centers to their median width, as in the monaLisa vignette. Sequences around all array CpGs without differential methylation (same width as the DMRs) were supplied as genome.regions without subsampling, and monaLisa sampled two candidate sequences per DMR from them (the default genome.oversample = 2) and selected the one closest in G+C composition, before weighting background sequences for GC and k-mer composition. Array CpGs have been used as the background in motif analyses of methylation array data [26, 27], and non-differential regions from the same assay have been used as the background with monaLisa [25]. The analysis was run once with a fixed seed (12345), as in the vignette, and motifs with an adjusted P value < 0.05 and log2 enrichment > 0 were considered enriched.

DMRs were examined in independent samples from the validation cohort of the same study (GSE109379, two CN and one LIPN, 450K) and in the GSE288337 CN (EPIC) [13]. For each sample, a LIPN score was calculated as the mean beta value of DMRs hypermethylated in LIPN minus that of DMRs hypomethylated in LIPN. CpGs annotated to 17 adipogenesis-related genes (*PPARG*, *CEBPA*, *CEBPB*, *CEBPD*, *PLIN1*, *PLIN4*, *FABP4*, *ADIPOQ*, *LPL*, *LEP*, *CIDEA*, *CIDEC*, *LIPE*, *PNPLA2*, *ZFPM1*, *NFATC1*, and *DLK1*), including *ZFPM1* and *NFATC1*, which have been described as part of an adipogenic program in LIPN [19], were examined in all samples, including the cases and the SELN. Genomic coordinates followed the hg19-based 450K annotation.

Supplementary results: crossNN classification

All 21 CN and 10 LIPN of GSE90496, which are part of the training data, were correctly classified by both published models, confirming the input conversion.

Supplementary results: DLK1 immunohistochemistry

DLK1 was stained in all three cases. It was negative in lipidized cells and positive only in a small minority of tumor cells, with a granular cytoplasmic pattern. Because DLK1 is expressed in tumors of neuroectodermal origin, often with granular cytoplasmic staining [30], this focal staining probably reflects the neuroectodermal nature of the tumor cells rather than adipogenic commitment.

Supplementary results: global methylation

Global methylation was also lower in the 21 CN profiled on the 450K array (mean 0.452) than in the 30 non-neoplastic controls, all of which were profiled on the 450K array (mean 0.488, Welch's t-test, $P < 1 \times 10^{-15}$). The difference between CN and controls was therefore not explained by array type alone.

Supplementary results: copy number

A shallow decrease of chromosome 19 (log2 ratio ≤ −0.1) was found in two cases and in five of nine SELN

profiled on EPICv2, but in no SELN profiled on EPIC or 450K and in no reference CN or LIPN. The chromosome 19 log2 ratio correlated strongly with the mean of 17q and 22q (r = 0.78, n = 56, Online Resource 2, Supplementary Fig. 2). In Case 1, chromosome 16 also showed a shallow decrease (log2 ratio −0.133 for both arms), together with the decrease of chromosome 19 (−0.147). Chromosome 16, like chromosomes 17, 19, and 22, is rich in CpG islands [14], and these decreases were regarded as GC-related technical variation rather than true losses. The *FGFR3–TACC3* value of Case 1 (0.041) slightly exceeded the maximum of the reference CN (0.032), but this did not exceed the fluctuations around the region.

Supplementary results: comparison of CN and LIPN

In the reference cohort, 42,976 CpGs (34,150 hypomethylated and 8,826 hypermethylated in LIPN) differed between CN and LIPN. DMRcate identified 9,834 regions, of which 6,253 had $|\Delta\beta| > 0.20$ (5,057 hypomethylated and 1,196 hypermethylated in LIPN). The direction of the difference was reproduced in 5,960 of the 6,253 DMRs (95.3%) in the GSE109379 samples (correlation of Δβ, r = 0.89), and the GSE288337 CN were closer to the reference CN than to LIPN in 98.1% of DMRs.

DMRs hypermethylated in LIPN were enriched for developmental genes (multicellular organism development, FDR = 3.7×10^{-21}), transcription factor genes (DNA-binding transcription factor activity, RNA polymerase II-specific, FDR = 2.0×10^{-17}), and cadherin signaling (KEGG, FDR = 2.6×10^{-12}). DMRs hypomethylated in LIPN reached FDR < 0.05 only for two GO terms (cell periphery and structural constituent of skin epidermis) and one KEGG pathway (cornified envelope formation) at the region level. In the 5,057 DMRs hypomethylated in LIPN (median width 540 bp), 150 motifs were enriched. The top motifs were those of NFI (NFIA, NFIX, NFIC, and NFIB) and ZIC2 and E-box motifs of neural basic helix–loop–helix (bHLH) factors (e.g. ASCL2, HAND2, ATOH1, NEUROG2, ASCL1, and NEUROD2), and the myogenic bHLH factor MYOD1, which recognizes the same E-box, was also enriched. In the 1,196 DMRs hypermethylated in LIPN (median width 650 bp), 14 motifs, mainly of homeobox factors (e.g. LMX1A, LMX1B, VSX1, VSX2, and LHX6), reached only weak significance (adjusted P 0.034–0.045). PPARG::RXRA, PPARG, CEBPA, CEBPB, and CEBPD motifs were not enriched in either direction. Among C/EBP motifs, only one CEBPG matrix was enriched in DMRs hypomethylated in LIPN (adjusted P = 0.025).

Of the 472 CpGs annotated to the 17 adipogenesis-related genes, 160 were differentially methylated. All were hypomethylated in LIPN at *NFATC1* (n = 84), *ZFPM1* (41), *PNPLA2* (8), *PLIN4* (6), *ADIPOQ* (4), *LIPE* (3), *CIDEC* (1), and *PLIN1* (1), all were hypermethylated at *DLK1* (7) and *LEP* (2), and *CIDEA* (3) had both. Of the 39 DMRs at these genes, those at *NFATC1* (n = 21), *ZFPM1* (12), *PLIN4*, *PNPLA2*, *ADIPOQ*, and *LIPE* (one each) were hypomethylated in LIPN (Δβ −0.70 to −0.20), and the two at *DLK1* were hypermethylated (Δβ +0.30 and +0.31). No differentially methylated CpG or DMR was found at *PPARG*, *CEBPA*, *CEBPB*, *CEBPD*, *FABP4*, or *LPL*. The cases had LIPN scores (−0.44, −0.37, and −0.42) within the range of reference CN (GSE90496, −0.45 to −0.38, GSE288337, −0.44 to −0.29) and far from LIPN (+0.18 to +0.30), and they were closer to the reference CN than to LIPN in 88–89% of DMRs. SELN were intermediate (median −0.05, range −0.11 to 0.02). At the adipogenesis-related DMRs, the mean methylation of the cases was within the range of reference CN for every gene (e.g. *ZFPM1*, 0.81–0.86 vs median 0.81 in CN and 0.40 in LIPN,

NFATC1, 0.82–0.85 vs 0.84 and 0.44, *PNPLA2*, 0.70–0.72 vs 0.72 and 0.37, *DLK1*, 0.24–0.47 vs 0.41 and 0.72) (Fig. 4). At the level of single differentially methylated CpGs, 83.0–96.1% of the values of the cases were closer to the CN than to the LIPN mean, comparable to the independent GSE288337 CN (88.4%). SELN were also CN-like at these DMRs except for *DLK1*, which was hypermethylated as in LIPN.

References

1. Aryee MJ, Jaffe AE, Corrada-Bravo H et al (2014) Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 30:1363–1369
2. Capper D, Jones DTW, Sill M et al (2018) DNA methylation-based classification of central nervous system tumours. Nature 555:469–474
3. Chimatira R (2026) Rosette-forming glioneuronal tumor of the cerebellum with intratumoral hemorrhage: case report with radiologic–pathologic correlation in a resource-limited setting. Childs Nerv Syst 42:186
4. Daenekas B, Pérez E, Boniolo F et al (2024) Conumee 2.0: enhanced copy-number variation analysis from DNA methylation arrays for humans and mice. Bioinformatics 40:btae029
5. Dahnoun J, van Kempen LC, Koljenović S et al (2026) Clinical evaluation of the CrossNN DNA methylation classifier for central nervous system tumors. Brain Pathol e70137. https://doi.org/10.1111/bpa.70137
6. Diskin SJ, Li M, Hou C et al (2008) Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. Nucleic Acids Res 36:e126
7. Eckhardt A, Drexler R, Schoof M et al (2024) Mean global DNA methylation serves as independent prognostic marker in IDH-wildtype glioblastoma. Neuro Oncol 26:503–513
8. Grimwood J, Gordon LA, Olsen A et al (2004) The DNA sequence and biology of human chromosome 19. Nature 428:529–535
9. Hansen KD, Aryee M, Irizarry RA (2026) detectionP: detection p-values for all probed genomic positions. In: minfi reference manual, version 1.58.0. Bioconductor. https://bioconductor.org/packages/release/bioc/manuals/minfi/man/minfi.pdf. Accessed 29 October 2026
10. Hoaglin DC, Iglewicz B, Tukey JW (1986) Performance of some resistant rules for outlier labeling. J Am Stat Assoc 81:991–999
11. Kim H, Lee K, Shim YM et al (2023) Epigenetic alteration of H3K27me3 as a possible oncogenic mechanism of central neurocytoma. Lab Invest 103:100159
12. Koperek O, Gelpi E, Birner P et al (2004) Value and limits of immunohistochemistry in differential diagnosis of clear cell primary brain tumors. Acta Neuropathol 108:24–30
13. Krech M, Muench A, Teichmann D et al (2025) Outcome-associated factors in a molecularly defined cohort of central neurocytoma. Acta Neuropathol 149:61
14. Lander ES, Linton LM, Birren B et al (2001) Initial sequencing and analysis of the human genome. Nature 409:860–921
15. Lucas CG, Gupta R, Doo P et al (2020) Comprehensive analysis of diverse low-grade neuroepithelial tumors with FGFR1 alterations reveals a distinct molecular signature of rosette-forming glioneuronal tumor. Acta Neuropathol Commun 8:151
16. Machlab D, Burger L, Soneson C et al (2022) monaLisa: an R/Bioconductor package for identifying regulatory motifs. Bioinformatics 38:2624–2625
17. Maksimovic J, Oshlack A, Phipson B (2021) Gene set enrichment analysis for genome-wide DNA methylation data. Genome Biol 22:173
18. McInnes L, Healy J, Melville J (2018) UMAP: uniform manifold approximation and projection for dimension reduction. arXiv:1802.03426. https://doi.org/10.48550/arXiv.1802.03426
19. Pearce TM, Cimino PJ, Lucas CG et al (2026) Sellar region neurocytomas exhibit a CIMP and neuroendocrine-like epigenetic signature distinct from other intra-axial neurocytomas. Acta Neuropathol 152:20

20. Peters TJ, Buckley MJ, Statham AL et al (2015) De novo identification of differentially methylated regions in the human genome. Epigenetics Chromatin 8:6
21. Peters TJ, Meyer B, Ryan L et al (2024) Characterisation and reproducibility of the HumanMethylationEPIC v2.0 BeadChip for DNA methylation profiling. BMC Genomics 25:251
22. Phipson B, Maksimovic J, Oshlack A (2016) missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. Bioinformatics 32:286–288
23. Rauluseviciute I, Riudavets-Puig R, Blanc-Mathieu R et al (2024) JASPAR 2024: 20th anniversary of the open-access database of transcription factor binding profiles. Nucleic Acids Res 52:D174–D182
24. Ritchie ME, Phipson B, Wu D et al (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47
25. Sadoughi B, Hernández-Rojas R, Hamou H et al (2026) Non-invasive measures of DNA methylation capture molecular aging in wild capuchin monkeys. Mol Ecol 35:e70468
26. Scala G, Federico A, Greco D (2021) CpGmotifs: a tool to discover DNA motifs associated to CpG methylation events. BMC Bioinformatics 22:278
27. Scala G, Federico A, Palumbo D et al (2020) DNA sequence context as a marker of CpG methylation instability in normal and cancer tissues. Sci Rep 10:1721
28. Sill M, Schrimpf D, Patel A et al (2026) Advancing CNS tumor diagnostics with expanded DNA methylation-based classification. Cancer Cell 44:340–354
29. Tornehave D, Jansen P, Teisner B et al (1993) Fetal antigen 1 (FA1) in the human pancreas: cell type expression, topological and quantitative variations during development. Anat Embryol (Berl) 187:335–341
30. Tornehave D, Jensen CH, Teisner B, Larsson LI (1996) FA1 immunoreactivity in endocrine tumours and during development of the human fetal pancreas; negative correlation with glucagon expression. Histochem Cell Biol 106:535–542
31. Xiao X, Zhou J, Wang J et al (2017) Clinical, radiological, pathological and prognostic aspects of intraventricular oligodendroglioma: comparison with central neurocytoma. J Neurooncol 135:57–65
32. Yuan D, Jugas R, Pokorna P et al (2025) crossNN is an explainable framework for cross-platform DNA methylation-based classification of tumors. Nat Cancer 6:1283–1294
33. Zhou W, Triche TJ Jr, Laird PW, Shen H (2018) SeSAMe: reducing artifactual detection of DNA methylation by Infinium BeadChips in genomic deletions. Nucleic Acids Res 46:e123