ARTICLE

Epigenomic alterations define lethal CIMP–positive ependymomas of infancy

S. C. Mack[1,2,3]*, H. Witt[4,5,6]*, R. M. Piro[6,7], L. Gu[6,8], S. Zuyderduyn[9], A. M. Stütz[6,10], X. Wang[1,2], M. Gallo[1], L. Garzia[1], K. Zayne[1], X. Zhang[11], V. Ramaswamy[1,2], N. Jäger[6,8], D. T. W. Jones[4,6], M. Sill[6,12], T. J. Pugh[13], M. Ryzhova[4,6], K. M. Wani[14], D. J. H. Shih[1,2], R. Head[1], M. Remke[1,2], S. D. Bailey[15,16], T. Zichner[6,10], C. C. Faria[1], M. Barszczyk[1,2], S. Stark[4,6], H. Seker-Cin[4,6], S. Hutter[4,6], P. Johann[4,6], S. Bender[4,6], V. Hovestadt[6,7], T. Tzaridis[4,6], A. M. Dubuc[1,2], P. A. Northcott[4,6], J. Peacock[1,2], K. C. Bertrand[1,2], S. Agnihotri[1], F. M. G. Cavalli[1], I. Clarke[1], K. Nethery-Brokx[1], C. L. Creasy[17], S. K. Verma[17], J. Koster[18], X. Wu[1], Y. Yao[1,2], T. Milde[5,6,19], P. Sin-Chan[1], J. Zuccaro[1], L. Lau[1], S. Pereira[1], P. Castelo-Branco[1], M. Hirst[20,21], M. A. Marra[21,22], S. S. Roberts[23], D. Fults[24], L. Massimi[25], Y. J. Cho[26], T. Van Meter[27], W. Grajkowska[28], B. Lach[29], A. E. Kulozik[5,6], A. von Deimling[6,30], O. Witt[5,6,19], S. W. Scherer[1], X. Fan[31,32], K. M. Muraszko[32], M. Kool[4,6], S. L. Pomeroy[13], N. Gupta[33], J. Phillips[34], A. Huang[1,35], U. Tabori[1,35], C. Hawkins[1,2], D. Malkin[36], P. N. Kongkham[1,2,3], W. A. Weiss[34], N. Jabado[37], J. T. Rutka[1,2,3], E. Bouffet[35], J. O. Korbel[10], M. Lupien[15,16,38], K. D. Aldape[14], G. D. Bader[9], R. Eils[6,8], P. Lichter[6,7], P. B. Dirks[1,2,3,39], S. M. Pfister[4,5,6], A. Korshunov[6,31,40] & M. D. Taylor[1,2,3]

Ependymomas are common childhood brain tumours that occur throughout the nervous system, but are most common in the paediatric hindbrain. Current standard therapy comprises surgery and radiation, but not cytotoxic chemotherapy as it does not further increase survival. Whole-genome and whole-exome sequencing of 47 hindbrain ependymomas reveals an extremely low mutation rate, and zero significant recurrent somatic single nucleotide variants. Although devoid of recurrent single nucleotide variants and focal copy number aberrations, poor-prognosis hindbrain ependymomas exhibit a CpG island methylator phenotype. Transcriptional silencing driven by CpG methylation converges exclusively on targets of the Polycomb repressive complex 2 which represses expression of differentiation genes through trimethylation of H3K27. CpG island methylator phenotype-positive hindbrain ependymomas are responsive to clinical drugs that target either DNA or H3K27 methylation both *in vitro* and *in vivo*. We conclude that epigenetic modifiers are the first rational therapeutic candidates for this deadly malignancy, which is epigenetically deregulated but genetically bland.

Ependymomas are malignancies that occur throughout the nervous system, but are more common in the hindbrain in children, as opposed to supratentorial and spinal cord tumours, which are more frequently diagnosed in adulthood. Despite being histologically identical, ependymomas from different regions of the nervous system are biologically and clinically distinct[1]. Current therapy for all ependymoma patients consists of maximal safe surgical resection, followed by radiation therapy[2]. Although adjuvant chemotherapy is routine for most children with

malignant brain tumours, it is not part of the current standard of care for ependymoma patients as multiple clinical trials have failed to show any survival benefit after cytotoxic chemotherapies[3]. Even at the time of disease recurrence, chemotherapy has not been shown to be effective for ependymomas; therefore, many children with recurrent ependymoma undergo a full but palliative second course of cranial irradiation[4]. Indeed, whereas treatment protocols for many other childhood malignancies have changed and improved in the past two decades, ependymoma

[1]Developmental & Stem Cell Biology Program, Arthur and Sonia Labatt Brain Tumour Research Centre, The Hospital for Sick Children, Toronto, Ontario M5G 1L7, Canada. [2]Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [3]Division of Neurosurgery, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [4]Division of Pediatric Neurooncology, German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany. [5]Department of Pediatric Oncology, Hematology and Immunology, University of Heidelberg, Heidelberg 69120, Germany. [6]German Cancer Consortium (DKTK), Heidelberg 69120, Germany. [7]Division of Molecular Genetics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany. [8]Division of Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany. [9]Department of Molecular Genetics, Banting and Best Department of Medical Research, The Donnelly Centre, University of Toronto, Toronto, Ontario M4N 1X8, Canada. [10]Genome Biology, European Molecular Biology, Laboratory Meyerhofstr. 1, Heidelberg 69117, Germany. [11]Department of Genetics, Norris Cotton Cancer Center, Dartmouth Medical School, Lebanon, New Hampshire 03756, USA. [12]Division of Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany. [13]Department of Neurology, Harvard Medical School, Children's Hospital Boston, MIT, Boston, Massachusetts 02115, USA. [14]Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. [15]Ontario Cancer Institute, Princess Margaret Cancer Centre–University Health Network, Toronto, Ontario M5G 1L7, Canada. [16]Ontario Institute for Cancer Research, Toronto, Ontario M5G 1L7, Canada. [17]Cancer Epigenetics Discovery Performance Unit, GlaxoSmithKline Pharmaceuticals, Collegeville, Pennsylvania 19426, USA. [18]Department of Oncogenomics, Academic Medical Center, Amsterdam 1105, The Netherlands. [19]CCU Pediatric Oncology, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany. [20]Centre for High-Throughput Biology, Department of Microbiology & Immunology, University of British Columbia, Vancouver, V6T 1Z4 British Columbia, Canada. [21]Canada's Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia V5Z 1L3, Canada. [22]Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia V6H 3N1, Canada. [23]Department of Pediatrics and National Capital Consortium, Uniformed Services University, Bethesda, Maryland 20814, USA. [24]Department of Neurosurgery, University of Utah School of Medicine, Salt Lake City, Utah 84132, USA. [25]Pediatric Neurosurgery, Catholic University Medical School, Gemelli Hospital, Rome 00168, Italy. [26]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, California 94305, USA. [27]Department of Pediatrics, Virginia Commonwealth, Richmond, Virginia 23298-0646, USA. [28]Department of Pathology, University of Warsaw, Children's Memorial Health Institute University of Warsaw, Warsaw 04-730, Poland. [29]Division of Anatomical Pathology, Department of Pathology and Molecular Medicine, McMaster University, Hamilton General Hospital, Hamilton, Ontario L8S 4K1, Canada. [30]Department of Neuropathology Ruprecht-Karls-University Heidelberg, Institute of Pathology, Heidelberg 69120, Germany. [31]University of Michigan Cell and Developmental Biology, Ann Arbor, Michigan 48109-2200, USA. [32]Department of Neurosurgery, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA. [33]Department of Neurosurgery, University of California San Francisco, San Francisco, California 94143-0112, USA. [34]Departments of Neurology, Pediatrics, and Neurosurgery, University of California, San Francisco, The Helen Diller Family Cancer Research Building, San Francisco, California 94158, USA. [35]Department of Neuro-oncology, The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada. [36]Department of Haematology and Oncology, The Hospital for Sick Children, Toronto, Ontario M5G 1X8, Canada. [37]Departments of Pediatrics and Human Genetics, McGill University and the McGill University Health Center Research Institute, Montreal, Quebec H3Z 2Z3, Canada. [38]Department of Medical Biophysics, University of Toronto, Toronto, Ontario M5G 1X8, Canada. [39]Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 1A8, Canada. [40]CCU Neuropathology, German Cancer Research Center (DKFZ), Heidelberg 69120, Germany.
*These authors contributed equally to this work.

therapy remains stagnant. The mechanisms underlying the chemo-resistance of ependymoma are not known.

Within each anatomical compartment (supratentorial/hindbrain/spinal), there is additional inter-tumoral heterogeneity in the form of well-documented molecular subtypes of ependymoma[5–7]. Ependymoma subtypes are clinically and functionally relevant, as rational therapies may only be effective in a single subtype of the disease[8]. Ependymomas are thought to arise from the regionally distinct radial glial cells. Differences between these radial glial cell populations are likely carried forward in the neoplasm, and may account for a portion of the observed heterogeneity[1,5]. Hindbrain ependymomas occur within the posterior fossa of the skull, and are clinically referred to as 'posterior fossa' (PF) ependymomas. There are two clear and distinct subtypes of PF ependymoma; one that occurs in older children and adults with very good prognosis (posterior fossa group B, or PFB), and another found predominantly in infants, which is associated with poor prognosis in spite of maximally aggressive therapy (posterior fossa group A, or PFA)[6,7].
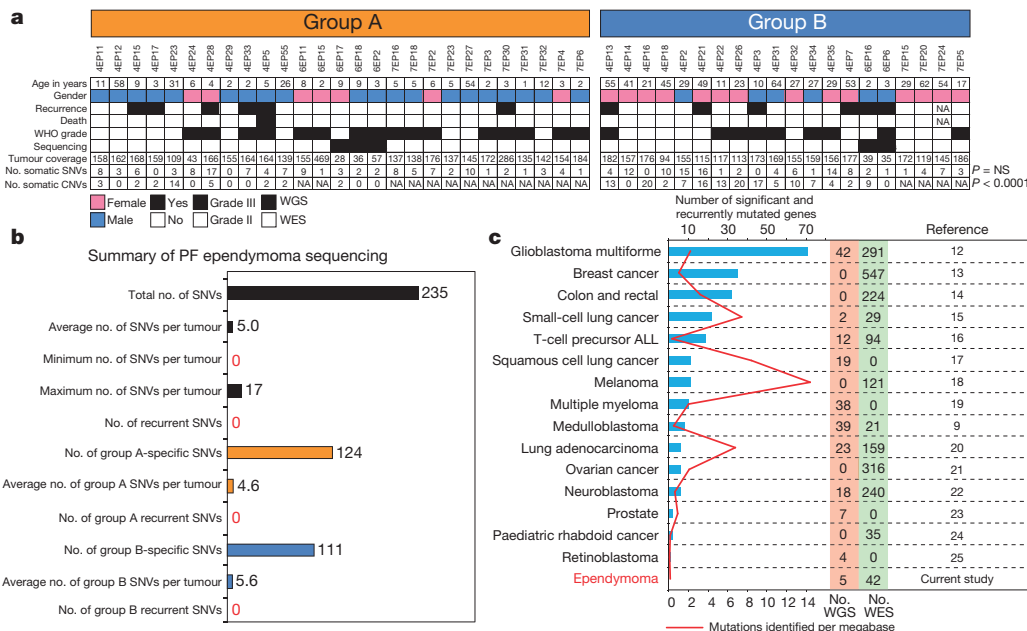
Ependymoma genome data

To uncover the biology of PF ependymomas, we undertook whole-genome sequencing of tumour and matching germline DNA from five ependymomas (3 PFA, 2 PFB), and whole-exome sequencing of an additional 42 PF ependymomas and their matching germline DNA (24 PFA and 18 PFB) (Fig. 1a, b, Supplementary Figs 1, 2 and Supplementary Tables 1, 2). Unlike some other childhood malignancies, the rate of somatic single nucleotide variants (SNVs) did not correlate significantly with the age at diagnosis (Supplementary Fig. 2)[9]. Further, the rate of somatic SNVs was extremely low in PF ependymomas, with an average of 5.0 somatic nonsynonymous SNVs per exome across the entire cohort (Fig. 1b), and low in both PFA (4.6 SNVs per tumour) and PFB ependymomas (5.6 SNVs per tumour, Supplementary Table 3). Perhaps the most surprising result was that there were zero significant recurrent mutations across the cohort of 47 PF ependymomas as detected by two different algorithms, MUTSIG[10] and MUSIC[11] (Fig. 1c and Supplementary Tables 3–6). Despite the absence of significant recurrent SNVs, PFB harboured frequent and recurrent large-scale copy number alterations (CNAs) indicative of chromosomal aneuploidy (Supplementary Fig. 3). Compared to other malignancies, PF ependymomas have a very low rate of SNVs per megabase, and the lowest number of recurrent significant SNVs, making PF ependymoma the first malignancy for which genome sequencing across a broad cohort ($n = 47$) has failed

to identify any significantly and recurrently mutated genes (Fig. 1c and Supplementary Tables 3–7)[9,12–25].

Ependymoma epigenome data

A number of other childhood nervous system malignancies, including medulloblastoma, retinoblastoma, glioblastoma, atypical teratoid/rhabdoid tumour, and neuroblastoma, have recently been demonstrated to harbour a paucity of recurrent mutations, with a significant proportion of the recurrent events converging on epigenetic mechanisms[9,22,24–33]. Owing to the absence of recurrent and significant SNVs and CNAs, we proposed that PFA ependymomas could be driven by epigenetic mechanisms. We studied DNA methylation patterns in a discovery cohort of 79 ependymomas using methyl-CpG-binding domain protein 2 (MBD2) recovery followed by hybridization to NimbleGen 385K CpG Island Promoter Plus microarrays (MBD2-chip). Unsupervised consensus clustering of CpG methylation profiles yielded three distinct subgroups, composed of supratentorial, PF, and mixed spinal/PF tumours, in a pattern highly similar to that yielded by unsupervised clustering of gene expression profiles (Fig. 2a and Supplementary Fig. 4)[6]. The group of pure PF tumours corresponds to PFA ependymomas, whereas the PFB ependymomas cluster with the spinal ependymomas. We validated our discovery cohort findings through study of a non-overlapping cohort of 48 PF ependymomas using an orthogonal technology (Illumina Infinium 450K methylation arrays). In these validation experiments, the DNA methylome of PFA ependymomas was very distinct from PFB tumours (Fig. 2b and Supplementary Fig. 5). Unsupervised clustering of CpG methylation signatures was very robust, supporting two major molecular subtypes, even after applying a number of distinct bioinformatic and biostatistical techniques (Supplementary Fig. 5). We conclude that PFA and PFB ependymomas have very distinct methylomes, and that epigenetic biomarkers could be used to develop a clinically relevant molecular classification of PF ependymomas. To this end, we identified three genes that exhibited increased CpG methylation in most PFA tumours, but not in PFB tumours (Supplementary Fig. 6). We determined the presence of CpG hypermethylation representing PFA tumours using a mass spectrometry-based technology (Sequenom) on our training cohort (Supplementary Figs 6, 7 and Supplementary Table 8). We were able to validate our Sequenom-based biomarker panel on an independent cohort of ependymomas using formalin-fixed, paraffin-embedded tissues to predict both progression-free and overall survival (Supplementary Figs 6, 7 and Supplementary Table 8). We conclude that division
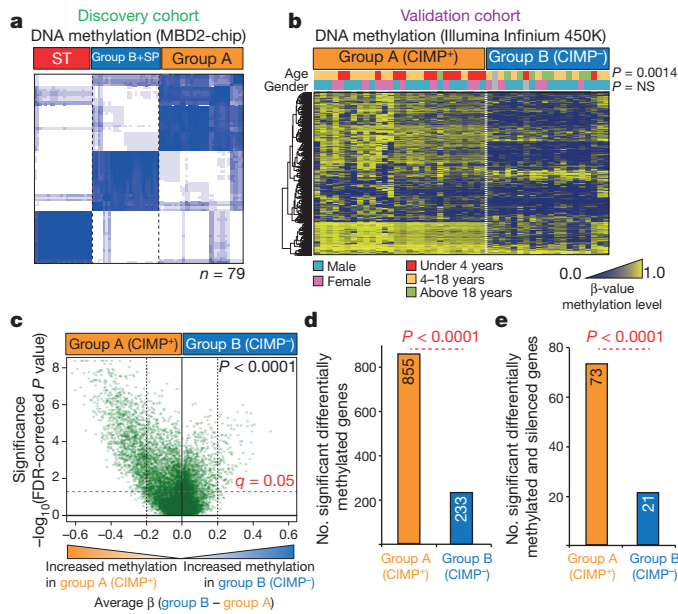

Figure 1 | Somatic SNVs are rare in the posterior fossa ependymoma genome. a, Summary of clinical and genomic details of PF ependymomas stratified according to group A and group B ependymoma (Wilcoxon rank-sum test). CNV, copy number variation; NA, not available; NS, not significant; WES, whole-exome sequencing; WGS, whole-genome sequencing. b, Bar graphs summarizing the numbers and frequencies of SNVs detected by whole-genome and whole-exome sequencing of PF ependymomas. c, Comparison of numbers of significant and recurrently mutated genes, and mutation rates, in several whole-genome and whole-exome sequencing studies of adult and paediatric cancers (false discovery rate (FDR) < 0.1).

Figure 2 | **DNA-methylation profiles suggest that group A ependymomas demonstrate a CpG island methylator phenotype. a**, Unsupervised hierarchical clustering of 79 ependymoma DNA-methylation profiles. SP, spinal cord; ST, supratentorial. **b**, Heatmap of 48 PF ependymoma DNA-methylation profiles. Group A and group B clinical differences were assessed using a two-sided Fisher's exact test. **c**, Volcano plot comparing the number of significant methylated CpG sites between group A and group B ($P < 0.05$, Wilcoxon rank-sum test, FDR-corrected). **d, e**, Differences in the number of methylated genes (**d**) and methylated and silenced genes (**e**) in group A versus group B ($P < 0.0001$, binomial distribution test).

of PFA from PFB ependymomas using a mass spectrometry-based biomarker should be feasible in a clinical setting.

CIMP phenotype in PFA ependymomas

We next compared the extent of promoter CpG methylation in PFA ependymomas to that of PFB ependymomas and found that PFA tumours have a much higher extent of CpG island methylation (Fig. 2c–e, Supplementary Figs 8, 9 and Supplementary Tables 9–14). In comparison to PFB ependymomas, PFA tumours have more methylated CpG sites (Fig. 2c), more genes with significant CpG methylation (Fig. 2d) and more genes that are transcriptionally silenced by CpG hypermethylation (Fig. 2e). We conclude that PFA ependymomas exhibit a 'CpG island methylator' or 'CIMP' phenotype, and suggest that PFA ependymomas be referred to as PFA CIMP-positive (PFA-CIMP$^+$) ependymomas, and PFB as PFB CIMP-negative (PFB-CIMP$^-$)[34] (Supplementary Table 15 and Supplementary Figs 8–12). To determine the mechanism by which CpG hypermethylation driving transcriptional silencing promotes the pathogenesis of PFA ependymoma, we performed a pathway analysis in our discovery cohort of PFA and PFB ependymomas studied by MBD2-chip (Fig. 3a and Supplementary Table 16). Although olfactory signalling was the only significant pathway enriched in PFB ependymomas, genes CpG methylated in PFA ependymoma showed a remarkable convergence on genes documented as silenced in embryonic stem (ES) cells by the Polycomb repressive complex 2 (PRC2). In our non-overlapping, independent validation data set studied by Illumina Infinium 450K arrays, we observed no significant pathways in the PFB ependymomas, whereas the PFA tumours exhibited the same convergence on gene targets that are silenced by PRC2 in ES cells (Fig. 3b and Supplementary Table 17). The PRC2 complex contains the histone methylase EZH2, which trimethylates H3K27, thereby driving gene silencing. Genes known to be required for differentiation and which are silenced by PRC2 have been documented to frequently undergo cancer-specific CpG methylation, and it is described that both DNA and histone methylation contribute to ongoing gene silencing in these cancers[35].
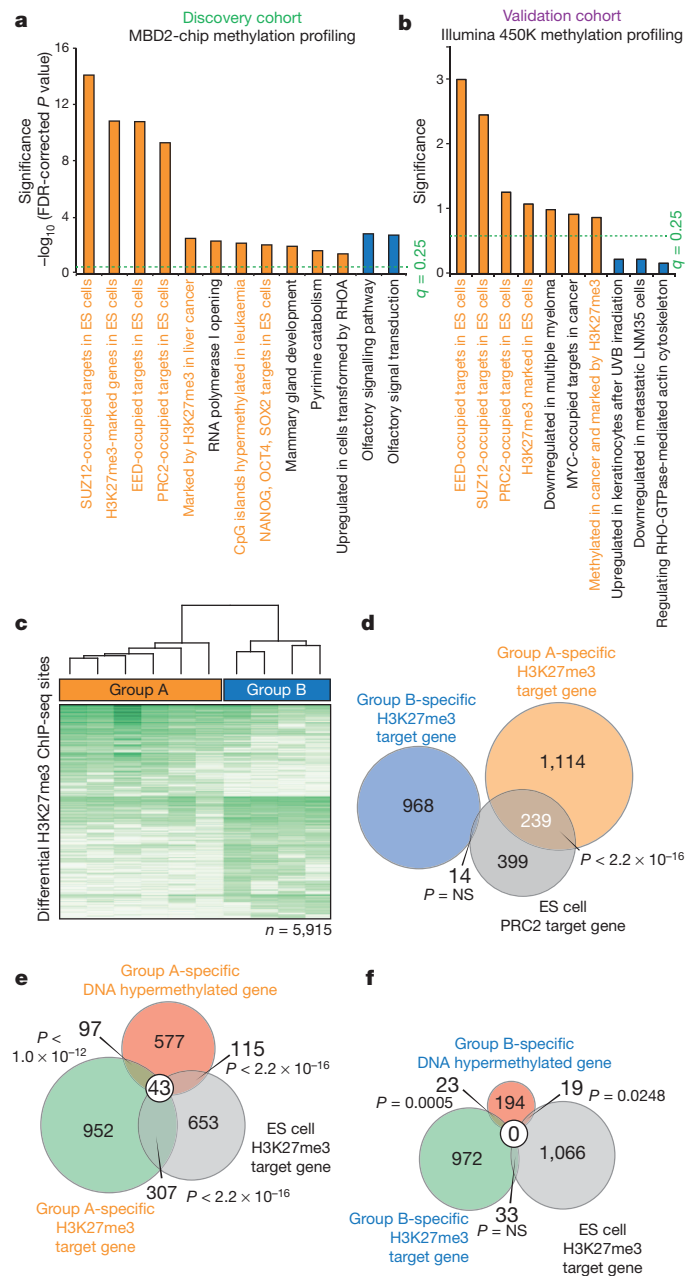

Figure 3 | **Group A (CIMP$^+$) and group B (CIMP$^-$) ependymomas are distinguished by CpG-hypermethylated and H3K27-trimethylated genes related to PRC2 occupancy in ES cells. a, b**, CpG-methylated pathways in group A (CIMP$^+$) and group B (CIMP$^-$) ependymomas in a discovery (**a**) and validation (**b**) cohort. **c**, Differential H3K27me3 binding sites distinguishing group A and group B ($P < 0.01$ (MACSv2.0), $P < 0.05$ (R:DiffBind)). **d, e**, Venn diagrams comparing group A and group B H3K27me3 genes with ES cell PRC2 genes (**d**) and group A H3K27me3 and DNA-hypermethylated genes with ES cell H3K27me3 genes (**e**). **f**, Group B H3K27me3 and DNA-hypermethylated genes with ES cell H3K27me3 genes (binomial distribution test)[46].

Convergence upon PRC2 targets

We next sought to validate these pathway findings by performing trimethylated H3K27 (H3K27me3) chromatin immunoprecipitation followed by massively parallel sequencing (ChIP-seq) in 11 primary PF ependymomas. Our findings demonstrate distinct H3K27me3 signatures in PFA-CIMP$^+$ versus PFB-CIMP$^-$ ependymomas (Fig. 3c, Supplementary Table 18 and Supplementary Figs 13, 14). Furthermore, the gene expression of H3K27me3 target genes can robustly stratify PFA-CIMP$^+$ from PFB-CIMP$^-$ tumours, thus highlighting the distinct epigenetic

differences between these subgroups (Supplementary Fig. 14). Examination of differential H3K27me3 targets demonstrated a convergence and significant overlap with PRC2 targets in ES cells observed exclusively in PFA-CIMP$^+$ tumours (Fig. 3d). Further, a significant proportion of shared PFA-CIMP$^+$ and ES cell H3K27me3 targets were CpG hypermethylated exclusively in PFA (CIMP$^+$) tumours, a pattern that was not detected in PFB-CIMP$^-$ ependymoma (Fig. 3e, f and Supplementary Tables 18, 19). We propose, therefore, that hyperactivity of the PRC2 complex leading to tumour suppressor gene silencing with subsequent gene silencing by DNA CpG hypermethylation contributes to the pathogenesis of PFA-CIMP$^+$ ependymoma.

We next sought to expand our global analysis of CpG methylation by performing whole-genome bisulphite sequencing in 3 PFAs, 3 PFBs, 3 fetal normal brains and 3 adult normal brains (Supplementary Tables 20–22). Here we observed the same patterns of increased CpG methylation at CpG islands occurring specifically in PFA-CIMP$^+$ tumours consistent with a CpG island methylator phenotype (Fig. 4a, c and Supplementary Fig. 15). In line with other solid tumours we identified additional cancer-specific epigenome patterns including hypomethylation of repetitive elements (long interspersed nuclear elements, short interspersed nuclear elements and long terminal repeats) restricted to PFA-CIMP$^+$ ependymoma and subgroup-specific partially methylated domains (Fig. 4b, d and Supplementary Figs 15, 16) These findings illustrate genome-wide DNA-methylation alterations in PFA-CIMP$^+$ ependymoma, concurrent with a silent genome exhibiting few CNAs and no significant and recurrent somatic SNVs.

CpG and histone methylation in tumour maintenance

Although our genomic and epigenomic data suggest that over-activity of the PRC2 complex, and/or subsequent promoter CpG hypermethylation, may be involved in driving the pathogenesis of PFA-CIMP$^+$ ependymomas, they do not address whether or not these mechanisms continue to be necessary for tumour maintenance, and would therefore constitute an effective target for therapy. Functional assessment of CpG and histone methylation in ependymoma is harshly limited by

the complete lack of established ependymoma cell lines, xenografts or transgenic mouse models[1,5,6]. To this end we established four short-term, patient-derived primary ependymoma cultures from two PFA-CIMP$^+$ tumours, and two childhood supratentorial ependymomas. We were unable to grow any PFB-CIMP$^-$ ependymomas in vitro. Treatment of PFA-CIMP$^+$ cultures with the DNA-demethylating agents 5-aza-2′-deoxycytidine (decitabine, here referred to as DAC) resulted in marked de-repression of gene sets enriched in EZH2 targets and known DNA-hypermethylated genes in other solid cancers (Fig. 4e, f). Furthermore, compared to supratentorial ependymoma primary cultures in vitro, DAC demonstrated significant anti-neoplastic effect on both PFA-CIMP$^+$ tumours at low-dose nanomolar levels (Fig. 4g). To model the effects of DAC on PFA-CIMP$^+$ cultures as early, and representative of the patient tumour as possible, we derived a passage zero ex vivo culture from a PFA-CIMP$^+$ metastasis. In this ex vivo culture we demonstrate significant impairment of neurosphere colony formation upon DNA-methylation blockade (Fig. 4h and Supplementary Fig. 17). Because DAC is FDA approved for the treatment of haematopoietic malignancies, we propose that it could be rapidly repurposed in a clinical trial for children with PFA-CIMP$^+$ ependymoma[36]. In addition, we observed additive effects on combining DAC and an FDA-approved histone deacetylase inhibitor (suberoylanilide hydroxamic acid, SAHA) against PFA-CIMP$^+$ ependymoma (Supplementary Fig. 18).

The tool compound 3-deazaneplanocin A (DZNep) is known to target the PRC2 complex and result in diminished trimethylation of H3K27 through degradation of PRC2 complex proteins[37]. Treatment of PFA-CIMP$^+$ ependymoma, but not supratentorial ependymoma, with DZNep is highly effective in the nanomolar range in vitro (IC$_{50}$ for E517, 95 nm; E520, 262 nm) (Fig. 5a). We also observed additive effects between DZNep and SAHA, and DZNep and DAC (Supplementary Fig. 18). Treatment of PFA-CIMP$^+$ ependymomas with DZNep, compared to controls, results in decreased expression of EZH2, decreased trimethylation of H3K27 and increased cleavage of PARP (Fig. 5b). In vivo treatment of established xenografts of human PFA-CIMP$^+$ ependymoma with DZNep using either a flank model (Fig. 5c), or an
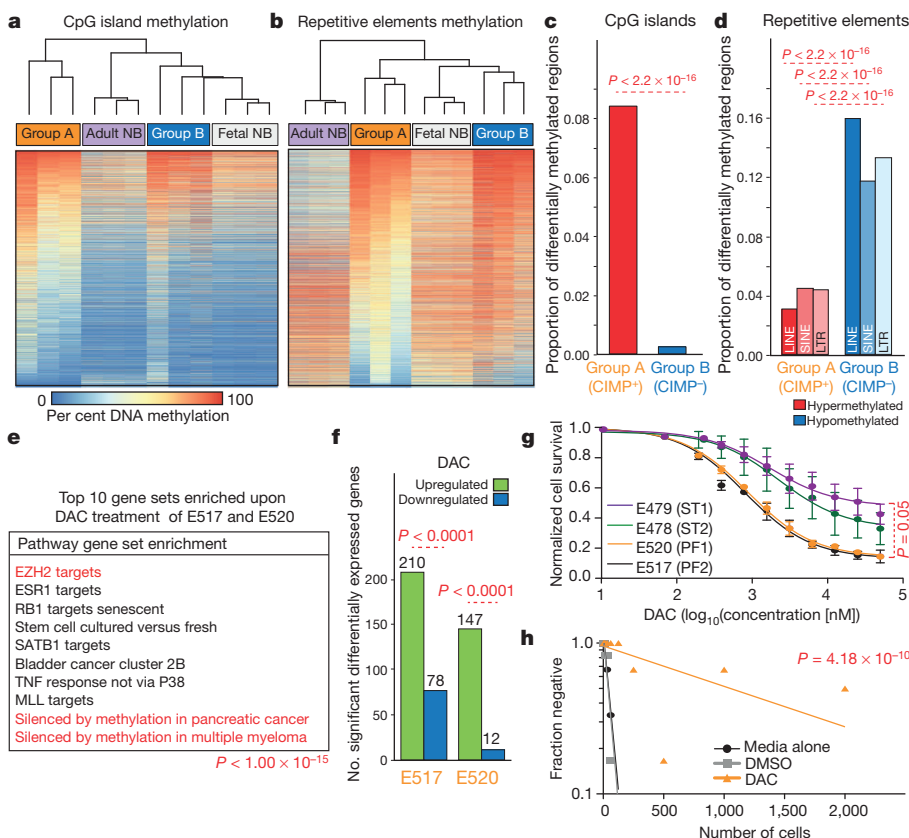

Figure 4 | Whole-genome bisulphite sequencing validates a CpG island methylator phenotype in group A ependymoma. a, b, Heatmap of DNA methylation at CpG islands (**a**) and repetitive regions (**b**) in group A (CIMP$^+$) versus group B (CIMP$^-$). NB, normal brain. **c, d,** Proportion of hypermethylated versus hypomethylated regions at CpG islands (**c**) and repetitive elements (**d**) in group A and group B ($P < 2.2 \times 10^{-16}$, binomial distribution test). LINE, long interspersed nuclear elements; LTR, long terminal repeats; SINE, short interspersed nuclear elements. **e, f,** Top 10 pathways upregulated (**e**) and differentially expressed (**f**) upon DAC treatment of E517-PF2 and E520-PF1 ($P < 0.0001$, binomial distribution test). **g,** Survival analysis of E478-ST2, E479-ST1, E517-PF2 and E520-PF1 cells treated for 7 days with DAC ($P = 0.05$, two-sided t-test, error bars, s.e.m.; technical ($n = 6$) over biological ($n = 2$)). **h,** Limiting dilution assay of zero passage group A cells treated for 2 weeks with DAC ($P = 4.18 \times 10^{-10}$, chi-squared test). DMSO, dimethylsulphoxide.
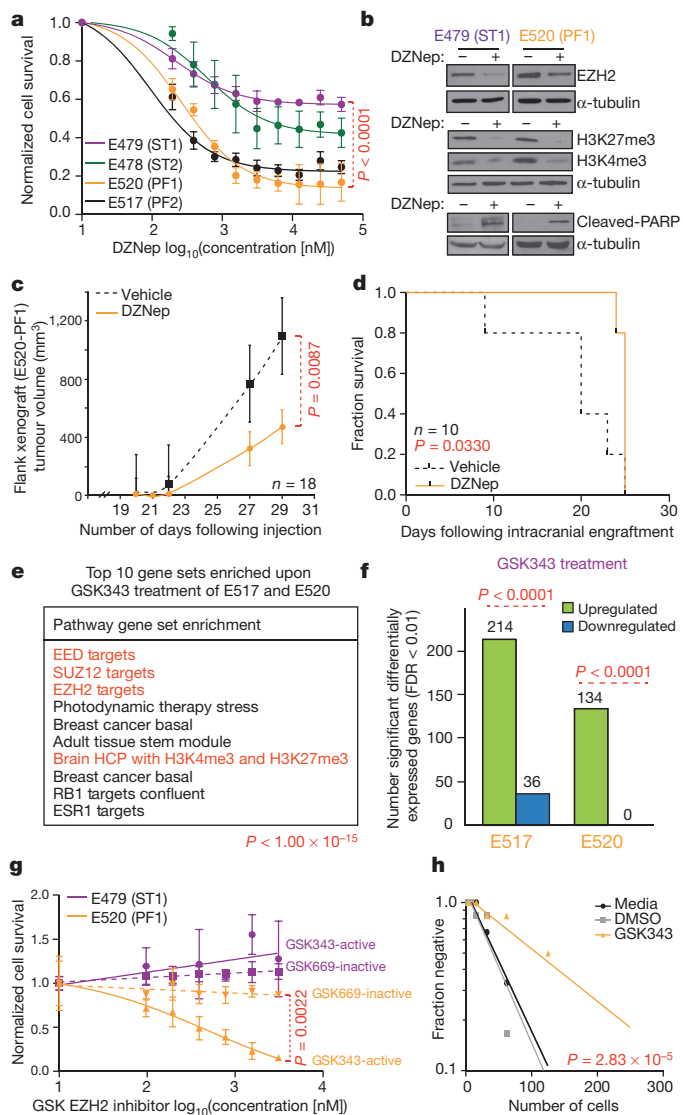
Figure 5 | Modulation of H3K27 methylation has anti-neoplastic effects against group A ependymoma. a, Survival of E479-ST1, E478-ST2, E520-PF1 and E517-PF2 cells treated for 7 days with DZNep ($P < 0.0001$, two-sided t-test, error bars = s.e.m., biological ($n = 3$)). **b**, EZH2, H3K27me3, H3K4me3 and cleaved-PARP protein expression in E520-PF1 and E479-ST1 cells treated for 7 days with DZNep (500 nM). **c**, E520-PF1 flank tumour volumes following treatment with DZNep ($P = 0.0087$, Wilcoxon-test, error bars = s.d.). **d**, Survival of E520-PF1 PF tumour-bearing mice treated with DZNep ($P = 0.033$, log-rank test). **e**, **f**, Top 10 pathways upregulated (**e**) and differentially expressed (**f**) upon treatment of E517 and E520 with GSK343. **g**, Cell proliferation of E520-PF1 and E479-ST1 cells treated for 11 days with GSK343 (active inhibitor) or GSK669 (inactive inhibitor) ($P = 0.0022$, two-sided t-test, error bars = s.e.m., technical ($n = 9$) over biological ($n = 3$)). **h**, Limiting dilution assay of passage zero group A cells treated for 2 weeks with GSK343 ($P = 2.83 \times 10^{-5}$, chi-square-test). HCP, high-CpG-density promoter.

orthotopic intracerebellar xenograft model (Fig. 5d) results in decreased tumour volume and improved survival. Furthermore, PFA-CIMP$^+$ ependymoma cells isolated from tumour xenografts treated *in vivo* with DZNep have a markedly reduced colony-forming ability compared to controls, suggesting that the compound targeted ependymoma cells with clonogenic or tumour-initiating potential (Supplementary Fig. 17). In addition, treatment with a recently published, and extremely potent, highly selective *S*-adenosyl-L-methionine competitive small molecular inhibitor of EZH2 (GSK343) results in significant de-repression of gene expression in PFA-CIMP$^+$ ependymoma including genes which are known targets of PRC2 in ES cells (Fig. 5e, f)[38]. We found that treatment

with GSK343 but not an inactive compound with the same molecular backbone (GSK669) resulted in diminished levels of H3K27me3, and had a potent antineoplastic effect against PFA-CIMP$^+$ ependymoma (Fig. 5g and Supplementary Fig. 17). These findings are further supported in a passage zero PFA-CIMP$^+$ *ex vivo* culture treated immediately with GSK343, which significantly impaired neurosphere colony formation (Fig. 5h). We therefore propose that ongoing hypermethylation of promoter CpG islands and H3K27 contribute to the maintenance of PFA-CIMP$^+$ ependymoma, and that targeting these epigenetic mechanisms represents the first identified rational targets for this chemotherapy-resistant, epigenetically dysregulated and genetically bland childhood cancer.

Discussion

A number of recent seminal publications have demonstrated that childhood nervous system tumours harbour very few recurrent genetic events, and that many of the recurrent events converge on genes important in epigenetic processes such as CpG hypermethylation, post-translational modification of histones, and even mutation of the actual histone genes[9,24–31]. Although mutation of *IDH1*, *IDH2*, *TET1*, *TET2* and/or *DNMT3A* has been documented in other types of cancer with a CIMP phenotype, we did not observe any such mutations in PF ependymomas[39,40] (Supplementary Table 23). Although a number of other paediatric malignancies have a very low incidence of recurrent somatic mutations, we are unaware of any other malignancies with zero significant recurrently mutated genes.

Subgroups of patients with a CIMP phenotype have a better prognosis for some cancer histologies[34], but not others[41–45], suggesting that CIMP-positive tumours represent distinct subgroups of disease, but that the CIMP phenotype itself is not intrinsically benign or responsive to therapy. PFA-CIMP$^+$ ependymomas have a nearly normal genetic code and a very poor prognosis in comparison to aneuploid PFB (CIMP$^-$) ependymomas, for which five-year overall survival exceeds 95% of patients. Many of the cytotoxic chemotherapeutics currently used clinically function through promoting damage to the genomic DNA, which subsequently induces cancer cells with deranged and disorganized genomes to undergo apoptosis. In light of the nearly normal genetic code found in PFA-CIMP$^+$ ependymomas, perhaps it is not surprising that a therapy based on DNA damage has not shown efficacy in clinical trials.

Our data demonstrate hyperactivity of DNA CpG methylation and disparate PRC2 H3K27me3 signatures in poor-prognosis PF ependymomas that may be necessary for tumour maintenance. Drugs that target DNA CpG methylation, PRC2/EZH2, and/or histone deacetylase inhibitors represent the first rational strategies for therapy of this untreatable disease, and should be considered for testing in clinical trials for children with PFA-CIMP$^+$ ependymoma.

METHODS SUMMARY

Whole-genome and whole-exome sequencing of PF ependymoma. Sequencing was performed using Illumina technologies (HiSeq2000). Median DNA sequence coverage was 35-fold for the five whole-genome cases (range 28–57×), whereas median on-target coverage in the whole-exome cohort ($n = 42$) was 157-fold (range 43–469×). Exome capture was carried out initially with Agilent SureSelect V3 and for later samples Agilent SureSelect V4 (Human All Exon 50 Mb) in-solution reagents. Sequence data were aligned to the hg19 human reference genome assembly; uplicate and non-uniquely mapping reads were excluded from calling of somatic variants. A subset of sequence variants were validated using PCR and capillary sequencing.

DNA methylation profiling of PF ependymoma. MBD2 protein recovery was performed using MethylMiner (Invitrogen) followed by hybridization to NimbleGen 385K CpG Island Promoter Plus Arrays. Microarray data was quantile normalized using the R:LIMMA package. Log$_2$ ratios were then transformed into BATMAN methylation states[47]. Comparisons between subgroups were performed using a Wilcoxon rank-sum test, and P values were corrected for multiple testing using the Benjamini–Hochberg method.

Illumina Infinium 450K methylation analysis was performed on bisulphite-treated DNA (Zymo) at The Centre for Applied Genomics (Toronto). CpG probes $P < 0.05$ were removed from the analysis, arrays were then normalized using the R:BMIQ algorithm[42]. Methylation values were then exported as β-values. Methylation probes were then filtered to CpG sites, which mapped to promoters containing CpG islands. A Wilcoxon rank-sum test was used to identify the differentially methylated

CpG sites between group A (CIMP$^+$) and group B (CIMP$^-$). P values for differentially methylated CpG sites identified were then corrected for multiple testing using the Benjamini–Hochberg method. Significant differences between numbers of CpG sites, genes, or methylated and silenced genes, was calculated using a binomial distribution test. Gene expression data for all of these samples can be found in our previous publication[6].

Online Content Any additional Methods, Extended Data display items and Source Data are available in the online version of the paper; references unique to these sections appear only in the online paper

1. Taylor, M. D. et al. Radial glia cells are candidate stem cells of ependymoma. Cancer Cell 8, 323–335 (2005).
2. Merchant, T. E. et al. Conformal radiotherapy after surgery for paediatric ependymoma: a prospective study. Lancet Oncol. 10, 258–266 (2009).
3. Bouffet, E. & Foreman, N. Chemotherapy for intracranial ependymomas. Childs Nerv. Syst. 15, 563–570 (1999).
4. Bouffet, E. et al. Survival benefit for pediatric patients with recurrent ependymoma treated with reirradiation. Int. J. Radiat. Oncol. Biol. Phys. 83, 1541–1548 (2012).
5. Johnson, R. A. et al. Cross-species genomics matches driver mutations and cell compartments to model ependymoma. Nature 466, 632–636 (2010).
6. Witt, H. et al. Delineation of two clinically and molecularly distinct subgroups of posterior fossa ependymoma. Cancer Cell 20, 143–157 (2011).
7. Wani, K. et al. A prognostic gene expression signature in infratentorial ependymoma. Acta Neuropathol. 123, 727–738 (2012).
8. Atkinson, J. M. et al. An integrated in vitro and in vivo high-throughput screen identifies treatment leads for ependymoma. Cancer Cell 20, 384–399 (2011).
9. Jones, D. T. et al. Dissecting the genomic complexity underlying medulloblastoma. Nature 488, 100–105 (2012).
10. Lawrence, M. S. et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 499, 214–218 (2013).
11. Dees, N. D. et al. MuSiC: identifying mutational significance in cancer genomes. Genome Res. 22, 1589–1598 (2012).
12. Brennan, C. W. et al. The somatic genomic landscape of glioblastoma. Cell 155, 462–477 (2013).
13. Koboldt, D. C. et al. Comprehensive molecular portraits of human breast tumours. Nature 490, 61–70 (2012).
14. Muzny, D. M. et al. Comprehensive molecular characterization of human colon and rectal cancer. Nature 487, 330–337 (2012).
15. Peifer, M. et al. Integrative genomic analyses identify key somatic driver mutations of small-cell lung cancer. Nature Genetics 44, 1104–1110 (2012).
16. Zhang, J. et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature 481, 157–163 (2012).
17. Hammerman, P. S. et al. Comprehensive genomic characterization of squamous cell lung cancers. Nature 489, 519–525 (2012).
18. Hodis, E. et al. A landscape of driver mutations in melanoma. Cell 150, 251–263 (2012).
19. Chapman, M. A. et al. Initial genome sequencing and analysis of multiple myeloma. Nature 471, 467–472 (2011).
20. Imielinski, M. et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 150, 1107–1120 (2012).
21. Bell, D. et al. Integrated genomic analyses of ovarian carcinoma. Nature 474, 609–615 (2011).
22. Pugh, T. J. et al. The genetic landscape of high-risk neuroblastoma. Nature Genetics 45, 279–284 (2013).
23. Berger, M. F. et al. The genomic complexity of primary human prostate cancer. Nature 470, 214–220 (2011).
24. Lee, R. S. et al. A remarkably simple genome underlies highly malignant pediatric rhabdoid cancers. J. Clin. Invest. 122, 2983–2988 (2012).
25. Zhang, J. et al. A novel retinoblastoma therapy from genomic and epigenetic analyses. Nature 481, 329–334 (2012).
26. Northcott, P. A. et al. Subgroup-specific structural variation across 1,000 medulloblastoma genomes. Nature 488, 49–56 (2012).
27. Pugh, T. J. et al. Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature 488, 106–110 (2012).
28. Robinson, G. et al. Novel mutations target distinct subgroups of medulloblastoma. Nature 488, 43–48 (2012).
29. Schwartzentruber, J. et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature 482, 226–231 (2012).
30. Sausen, M. et al. Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. Nature Genet. 45, 12–17 (2013).
31. Wu, G. et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nature Genet. 44, 251–253 (2012).
32. Versteege, I. et al. Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. Nature 394, 203–206 (1998).
33. Northcott, P. A. et al. Multiple recurrent genetic events converge on control of histone lysine methylation in medulloblastoma. Nature Genet. 41, 465–472 (2009).
34. Noushmehr, H. et al. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell 17, 510–522 (2010).
35. Ohm, J. E. et al. A stem cell-like chromatin pattern may predispose tumor suppressor genes to DNA hypermethylation and heritable silencing. Nature Genet. 39, 237–242 (2007).
36. Kantarjian, H. et al. Decitabine improves patient outcomes in myelodysplastic syndromes: results of a phase III randomized study. Cancer 106, 1794–1803 (2006).
37. Tan, J. et al. Pharmacologic disruption of Polycomb-repressive complex 2-mediated gene repression selectively induces apoptosis in cancer cells. Genes Dev. 21, 1050–1063 (2007).
38. McCabe, M. T. et al. EZH2 inhibition as a therapeutic strategy for lymphoma with EZH2-activating mutations. Nature 492, 108–112 (2012).
39. Turcan, S. et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature 483, 479–483 (2012).
40. Williams, K. et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature 473, 343–348 (2011).
41. Fang, F. et al. Breast cancer methylomes establish an epigenomic foundation for metastasis. Sci. Transl. Med. 3, 75ra25 (2011).
42. Roman-Gomez, J. et al. Lack of CpG island methylator phenotype defines a clinical subtype of T-cell acute lymphoblastic leukemia associated with good prognosis. J. Clin. Oncol. 23, 7043–7049 (2005).
43. Zouridis, H. et al. Methylation subtypes and large-scale epigenetic alterations in gastric cancer. Sci. Transl. Med. 4, 156ra140 (2012).
44. Toyota, M. et al. CpG island methylator phenotype in colorectal cancer. Proc. Natl Acad. Sci. USA 96, 8681–8686 (1999).
45. Weisenberger, D. J. et al. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nature Genet. 38, 787–793 (2006).
46. Ben-Porath, I. et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nature Genet. 40, 499–507 (2008).
47. Down, T. A. et al. A Bayesian deconvolution strategy for immunoprecipitation-based DNA methylome analysis. Nature Biotechnol. 26, 779–785 (2008).

Author Contributions S.C.M., H.W., M.D.T., A.K. and S.M.P. conceived and led the study. S.C.M. performed sample preparation, data acquisition and bioinformatic analysis related to DNA methylation, ChIP-seq and transcriptional profiling experiments. S.C.M. also performed western blot analysis and contributed to in vitro and in vivo experiments. H.W. led whole-genome sequencing, whole-exome sequencing and whole-genome bisulphite sequencing efforts, with analytical support from R.M.P., L. Gu., N. Jäger, D.T.W.J., S.S., S.H., T.Z., A.M.S., T.J.P., M.S., H.S.-C., T.T., V.H. and J.O.K. S.Z. and G.D.B. developed and applied pathway analysis methods for all data sets in this study. K.N.-B., M.G., L. Garzia, K.Z., X. Wang, M.B., S.B., P.J., X. Wu, K.C.B., T.M., J.Z., P.S.-C., C.C.F., P.C.-B., Y.Y. and S.A. performed in vitro, in vivo and molecular/biochemical experiments central to the project. DNA methylation and ChIP-seq bioinformatic analysis and experimental design were supported by X.Z., V.R., P.N.K., A.M.D., P.A.N., D.J.H.S., J.P., M. Remke, F.M.G.C., L.L., S.P., S.W.S., S.D.B., M.G., J.K., I.C., R.H. and P.B.D., developed and characterized ependymoma primary cultures, and assisted with in vitro and in vivo experiments. C.L.C. and S.K.V. developed the EZH2 compounds and provided inhibitors for experimentation. S.S.R., L.M., Y.J.C., T.V.M., W.G., B.L., M. Ryzhova, A.K., N.G., J.P., K.W. and K.D.A. provided patient samples and clinical details that made this study possible. M.H., M.A.M., M.L., D.F., A.E.K., A.v.D., O.W., D.M., X.F., K.M.M., M.S.L.P., E.B., W.A.W., A.H., U.T., C.H., J.T.R., N. Jabado, J.O.K., R.E., P.L., G.D.B., K.D.A., P.B.D. and S.M.P. provided valuable input regarding study design, data analysis and interpretation of results. M.D.T., S.C.M., H.W., A.K. and S.M.P. wrote the manuscript. M.D.T. and S.M.P. provided financial and technical infrastructure and oversaw the study. M.D.T. and A.K. are joint senior authors and project co-leaders.

Author Information Illumina 450K CpG Methylation array data, NimbleGen 385K CpG Island Plus array data, and ChIP-seq data have been deposited at the Gene Expression Omnibus (GEO: http://www.ncbi.nlm.nih.gov/geo/) as a GEO super-series under the accession number GSE43353. Whole-genome and whole-exome sequencing data have been deposited in the European Genome-Phenome Archive (EGA: https://www.ebi.ac.uk/ega/) hosted by the EBI, under the accession number EGAS00001000443. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.K. (Andrey.Korshunov@med.uni-heidelberg.de) or M.D.T. (mdtaylor@sickkids.ca).

METHODS

Patients and tumour samples. Tumour samples, clinical information and animal studies were processed in approval with local ethics board from both the institutions as described previously[6]. Informed consent was obtained from all patients, as described previously[6]. No patient underwent chemotherapy or radiotherapy before the surgical removal of the primary tumour. This study included only primary samples for analysis, and further excluded WHO grade I histological variants of ependymoma. Detailed clinical description of patient characteristics is shown for the sequencing cohort in Supplementary Table 2 and for the methylation cohort in Supplementary Table 8. Tumour subgrouping was based on gene expression profiling or immunohistochemical analysis as described previously[6]. At least 80% of tumour cell content was estimated in all tumour samples of the sequencing cohort by staining cryosections (~5 μm thick) of each sample with haematoxylin and eosin. Diagnoses were confirmed by histopathological assessment by at least two neuropathologists, including a central pathology review that used the 2007 WHO classification for central nervous system tumours.

DNA library preparation and Illumina sequencing. Tumour and control samples were individually processed, in every case thorough histological examination proved that each tumour consisted of >80% tumour cells, in most cases it was >95%. DNA from tumour and control samples (blood) was prepared and sequenced individually. The Agilent SureSelect Human All Exon 50 Mb target enrichment kit (v3 initially, switched to v4 subsequently) was used to capture all human exons for deep sequencing, using the vendor's protocol v2.0.1. The SureSelect Human All Exon Kit targets regions of 50 megabases (Mb) in total size, which is approximately 1.7% of the human genome. In brief, 3 μg of genomic DNA was sheared with a Covaris S2 to a mean size of 150 base pairs (bp). 500 ng of library was hybridized for 24 h at 65 °C with the SureSelect baits. The captured fragments of the tumour samples and controls were sequenced in 105-bp single-end mode on an Illumina HiSeq2000 deep sequencing instrument (based on Illumina, Inc., v3 sequencing chemistry). Median coverage of whole-exome-sequenced tumour samples was 157-fold (range 43–469×) and for control samples (blood DNA) 146-fold (range 80–222×). In addition, whole-genome libraries (before the exome hybridization step) were sequenced three lanes each in paired-end 105-bp mode on the HiSeq2000, as described by Jones et al.[9].

To increase the coverage of the samples for whole-exome sequencing we used the following strategy: exome capture was initially carried out with Agilent SureSelect (Human All Exon 50 Mb) in-solution reagents using the default Illumina adapters (without barcode). To introduce Illumina Multiplex barcodes into the existing libraries at a later stage, 15 ng final exome-enriched library (without barcode) was used as a template in a 50 μl PCR reaction. The Herculase II Fusion enzyme (Agilent) was used together with the NEBNext Universal PCR primer for Illumina and NEBNext Index primer (NEB no. E7335S) with the following conditions: the initial denaturation step for 2 min at 98 °C was followed by 4 cycles of 30 s 98 °C, 30 s 57 °C, 1 min 72 °C, and a final 10 min at 72 °C step. 6–7 barcoded samples were then sequenced on the Hiseq2000 in 2 × 100-bp paired-end mode.

DNA sequence data processing. For each sequencing lane, read pairs were mapped to the human reference genome (hg19, NCBI build 37.1, downloaded from the UCSC genome browser at http://genome.ucsc.edu/) using BWA version 0.5.9-r16 with default parameters and maximum insert size set to 1 kilobase (kb). We used SAMtools to generate a chromosomal coordinate-sorted BAM file. Post-processing of the aligned reads included merging of lane-level data and removal of duplicate read pairs per sequencing library using Picard tools (version picard-1.48, http://picard.sourceforge.net). Lane, library and sample information was captured in the read group tag in the header of the merged final BAM file. Only uniquely aligned reads (minimum mapping quality of 1) were considered for downstream mutation analysis. Coverage calculations following duplicate removal considered all informative bases of the reference genome (excluding Ns, where N indicates that neither A, C, G or T DNA base could be accurately called). A mean Phred-scaled base quality of at least 25 across the length of the read was required. For target capture sequencing, only bases of reads overlapping the targets ± 100 bases were considered for coverage calculations. Sequencing statistics are given in Supplementary Tables 1 and 2.

SNV detection. Our analysis pipeline for SNV detection integrates publicly available tools with custom in-house software and applies several filtering and annotation steps. SNV calling is based on SAMtools mpileup and bcftools (version 0.1.17), using parameter adjustments to allow calling of somatic variants. Default settings of bcftools are designed for diploid samples, but owing to tumour heterogeneity, polyploidy and normal cell contamination, tumour genomes often have a significantly lower mutant allele frequency than that seen in normal diploid genomes. Therefore, somatic SNVs are often not called by standard tools designed for detection of single nucleotide polymorphisms, for example, in population studies such as the 1000 Genomes Project (http://www.1000genomes.org/). Initial SNV candidates were identified by using SAMtools mpileup for each tumour

sample, considering only reads with a minimum mapping quality of 30 and bases with a minimum base quality of 13, after application of the extended base alignment quality (BAQ) model. BAQ is the Phred-scaled probability of a read base being misaligned, and it is designed to reduce false SNV calls caused by misalignments. After the pileup of high-quality bases at each position of the input BAM file, bcftools applies the prior and performs the actual SNV calling. We changed the default probability of calling a variant if $P(\text{ref}|D) < 0.5$ to 1.0, which results in all positions containing at least one high-quality non-reference base to be reported as a variant. Therefore, this initial set of SNV candidates contains a high fraction of false positive calls, but ensures that true somatic mutations with low allele frequency (well below the expected 50% allele frequency) are reported. This initial SNV call set is then subjected to various filters. SNVs covered by fewer than three reads in the tumour and control samples, with somatic allele frequency <10%, or with only one read supporting the variant, were excluded. In addition, a minimum of 10 high-quality reads available at the corresponding position in the control sample were required, in order to be able to distinguish somatic from germline variants. Local sequence context can lead to incorrect base calls, but typically involves reads sequenced from a single strand only. Thus, if the variant call was supported by reads from only one strand, the ± 10 bases around the SNV were automatically screened for Illumina-specific error profiles and excluded if a profile was matched. For all tumour SNV calls the pipeline generates a pileup of the bases in the normal sample considering only uniquely mapping reads. SNV calls were categorized as germline or somatic according to whether there was evidence for the same event at the same locus in the BAM file of the tumour-matched control sample. Filtered calls were annotated with RefSeq gene annotations, dbSNP build 135 and variants from the 1000 Genomes project. Calls matching the position of known dbSNP (up to version 131) or known 1000 Genome variants were excluded from the high-confidence somatic call set (calls matching the position of dbSNP version >131 but not the position of 1000 Genome variants were retained because cancer-relevant somatic mutations, such as several TP53 mutations, have been included in more recent dbSNP versions). In addition, we filtered out SNVs that were found in at least 1% of the control samples or at least 1% of a set of 162 unrelated controls from other studies, because they constitute likely unannotated, naturally occurring SNPs and/or false positives, for example, artefacts related to sequencing and mapping. The pipeline integrates Annovar (http://www.openbioinformatics.org/annovar/) to determine whether the observed amino acid change has synonymous, nonsynonymous, nonsense or splice-site-changing properties on the encoded protein. Variants were further annotated with genes listed in the Cancer Gene Census (http://www.sanger.ac.uk/genetics/CGP/Census/) and entries from the Catalogue of Somatic Mutations version 57 (COSMIC, http://www.sanger.ac.uk/genetics/CGP/cosmic/), in addition to the full RefSeq gene summary, full gene name and genomic size. A subset of sequence variants and indels were validated by capillary sequencing by Sanger using purified PCR products. Primer sequences are available upon request.

Small insertion and deletion (indel) detection. Small insertions and deletions were identified with SAMtools and bcftools. The indel discovery pipeline is similar to the SNV pipeline (as described above), but using default bcftools parameters, to reduce the known high false positive rate associated with current indel detection methods for deep sequencing data. To call an indel a germline event, we only required one indel-supporting read in the matching normal sample, again to reduce the high fraction of false-positive somatic indel calls. Calls overlapping simple repeat or microsatellite regions were excluded as such regions are commonly observed to yield false-positive calls. Annotation of indels was identical to SNV annotation.

Computation of recurrently mutated genes. To search for genes mutated at significant frequency, we applied the MutSig algorithm, a method that corrects for background mutation rate and gene length. Details can be found at Broad CGA tools website (http://www.broadinstitute.org/cancer/cga/mutsig) including previously published studies[10].

Identification of rearrangements and generating of Circos plots. Structural rearrangements, namely deletions, tandem duplications, inversions and translocations, were detected using DELLY, which is based on paired-end mapping. The structural rearrangement calls were filtered using the corresponding ependymoma germline samples, germline data of additional medulloblastoma samples[9], and phase I 1000 Genomes Project (http://1000genomes.org) genome data to exclude germline structural variants as well as rearrangement calls caused by mapping artefacts. We only considered those rearrangements for further analysis, which were present in at most 0.5% of the 1000 Genomes Project samples assessed and not in the additional germline samples. Two rearrangement calls were considered to be identical, hence constituting a likely germline variant if they displayed an overlap in terms of genomic coordinates with their end coordinates differing by less than 5 kb. Furthermore, rearrangement calls with less than 10 supporting pairs

as well as supporting pairs with average mapping quality less than 20 were excluded for further analysis. The circular whole-genome plots were generated using Circos.

Identification of pathways affected by SNVs. A pathway association test was used identify groups of functionally related genes that contained a greater than expected number of SNVs in one or the other ependymoma subtype. For each gene set, SNVs were stratified by subtype (PFA versus PFB). The number of SNVs in all genes and the number of SNVs observed in a given gene set were totalled for each subtype. A Fisher's exact test was performed with the null hypothesis that the frequency of SNVs in a given gene set was equal in the two subtypes. To correct for multiple testing, gene names were randomly shuffled and the analysis repeated to obtain a null P value for a given gene set. Randomization was done 10,000 times and, for each gene set, the percentage of null P values that were the same or lower than that obtained from the actual data was used as an estimate of the FDR.

Generation of copy number profiles from Illumina 450K methylation data. Low-resolution (450K probes) copy number variations were detected from the 450K Infinium methylation array in a custom approach using the sum of both methylated and unmethylated signals. Probes found to be highly variant in the six normal cerebellum samples were excluded from the analysis according to the following criteria: removal of probes not within the 0.05 and 0.85 quantile of median summed values or over the 0.8 quantile of the median absolute deviation. Log-ratios of samples to the median value of control samples were calculated, and sample noisiness was determined as the median absolute deviation of adjacent probes. Probes were then combined by joining 20 adjacent probes, and resulting genomic windows less than 100 kb in size were iteratively merged with adjacent windows of smaller size. Windows of more than 5 Mb were excluded from analysis, resulting in a total of 8,654 windows throughout the genome. For each window, the median probe value was calculated and shifted to minimize the median absolute deviation from all windows to zero for every sample. Segmentation was performed by applying the circular binary algorithm.

MBD2-assisted recovery and sample preparation. Genomic DNA was isolated according to previous methods[6]. DNA (6 μg) was immunoprecipitated using the MBD2 protein and quantified using the Qubit fluorometer (Invitrogen). Enrichment was assessed by quantitative PCR for positive controls: *RASSF1A*, *DLK1*, *H19* and negative controls: *ACTB* and *GAPDH*. Bound and unbound fractions from the MBD2 pulldown were whole-genome-amplified (SIGMA-WGA2) in triplicate, pooled, quantified using Qubit, and subjected to another round of quantitative PCR for the above control targets. DNA was sent to NimbleGen to be hybridized to Nimblegen 385K CpG Island Promoter Plus arrays, in which 'Immunoprecipitated-IP' fraction was labelled with Cy5 and 'unbound' fraction with Cy3.

Methylation analysis of MBD2-chip data. Microarray data was quantile normalized using the LIMMA Bioconductor package. Log$_2$ ratios were then imported into Agilent Genomics Workbench (Agilent Technologies), following which the BATMAN algorithm was used to infer the methylation statuses associated with each probe[47]. Mean methylation states were calculated for probes within a 1,000-bp window and termed a region of interest (ROI). ROIs were then filtered to those with greater than 4 probes and mapped to autosomal chromosomes. ROIs exhibiting a standard deviation greater than 0.65 were used for subgroup assignment as described below. Comparisons between subgroups were performed using a Wilcoxon rank-sum test, and P values were corrected for multiple testing using the Benjamini–Hochberg method. For comparisons of DNA methylation and other factors in this manuscript a Wilcoxon test was used and corrected for multiple testing, such that no assumptions were made regarding the normality of the data distributions.

Illumina Infinium 450K methylation sample preparation and data analysis. Genomic DNA was isolated according to previous methods[6]. DNA (1 μg) was used for bisulphite treatment (Qiagen, EpiTect plus) with the use of DNA protect buffer, particularly in the case of DNA from formalin-fixed paraffin-embedded tissue. Bisulphite-treated DNA was then quantified using spectrophotometry (Nanodrop). >500 ng was sent to The Centre for Applied Genomics (TCAG, Toronto) for hybridization to Illumina 450K Methylation Arrays. Array pre-processing was performed using GenomeStudio (Illumina) with background subtraction adjustment applied. Arrays were also normalized using the BMIQ method, which produced the same finding of a group A-specific CIMP. Methylation values were then exported as β-values (estimates of actual CpG methylation levels). Probes that overlapped with known single nucleotide polymorphisms, which mapped to chromosomes X and Y, and were Illumina control probes, were removed from the analysis. Methylation probes were then filtered to CpG sites, which mapped to promoters containing CpG islands. A Wilcoxon rank-sum test (Mann–Whitney) was used to identify the differentially methylated CpG sites between group A (CIMP$^+$) and group B (CIMP$^-$). P values for differentially methylated CpG sites identified were then corrected for multiple testing using the Benjamini–Hochberg method. Significant differences between numbers of CpG sites, genes or methylated and silenced genes, was calculated using a binomial distribution test. Methylated

and silenced genes were identified in two ways: (1) by identifying genes which were methylated and downregulated following comparison between group A and B using a Wilcoxon rank-sum test; or (2) by performing a Pearson correlation between the methylation status of a CpG site with the corresponding downstream gene. Methylated and silenced genes (within the same tumour) were identified by genes demonstrating significant and preferential methylation in a particular subgroup and evidence of downregulation as compared to a collection of normal brain samples. Gene expression data for these samples can be found in our previous publication[6].

Subgroup analysis of gene expression and methylation data. To detect robust sample clusters from the gene expression data (Affymetrix Exon 1.0ST) we performed hierarchical clustering using the top 1,000 varying probes as described previously in Witt *et al.*[6]. For clustering of MBD2-chip data we performed consensus hierarchical clustering with agglomerative average linkage as our method for consensus clustering. (R package: ConsensusClusterPlus). The change in area under the cumulative distribution function curve was used to identify the principal number of subgroups for a given clustering method. Silhouette analysis was performed to evaluate cluster representation of samples (R package: cluster). To evaluate the concordance between gene expression and DNA-methylation subgroup stratification we calculated the Rand index, with the significance assessed by permutation of sample labels and computing the Rand index over 10,000 iterations in order to generate a null distribution. Illumina 450K methylation data was clustered using the probes exhibiting a standard deviation of >0.2 as described previously. A variety of consensus clustering methods was performed: K-means, non-negative matrix factorization, hierarchical clustering and self-organizing maps were used. The distance metric used in the case of K-means was Euclidean, whereas a Pearson correlation was used for all other methods. Principal component analysis was performed within Partek Genomics Suite (Partek Inc.) to compare group A (CIMP$^+$) and group B (CIMP$^-$) posterior fossa subtypes with the same genes or CpG sites used for consensus hierarchical clustering and consensus K-means clustering, respectively.

H3K27me3 and EZH2 ChIP-seq profiling and analysis in PF ependymoma samples. 10–20 mg of fresh–frozen primary tumour samples was homogenized in 1% formaldehyde and allowed to incubate at room temperature for 6–10 min. Crosslinking was stopped with the addition of 125 mM of glycine, and samples were washed twice with ice-cold PBS containing 1% BSA and 10% FBS. Samples were then sonicated to ~200-bp fragments using a Bioruptor (Diagenode). The chromatin immunoprecipitation was then performed using 5 mg of EZH2 antibody (no. 39875-Active Motif) or H3K27me3 antibody (C15410069-Diagenode) overnight at 4 °C as described previously[48]. DNA was quantified using PicoGreen (Invitrogen) and libraries were prepared using NEBNext ChIP-seq Illumina Sequencing library preparation kit (NEB). Samples were barcoded (NEB Next Barcodes) and pooled in equimolar amounts such that up to 6 samples could be sequenced by paired-end Illumina HiSeq 2000 sequencing (Illumina).

ChIP-seq reads were aligned using the BWA algorithm with removal of redundant reads (Picard Algorithm) likely to represent ChIP-seq PCR library artefacts, yielding uniquely mapped ChIP-seq reads. Peaks were identified using MACS (version 2) with a P value cutoff of 0.01. Differential peaks were identified using the R: Bioconductor DiffBind package ($P < 0.05$) and annotated to the nearest gene ±5 kb using Cistrome (http://cistrome.org/ap/root). Overlap analysis between H3K27me3 genes or EZH2 target genes was assessed statistically using a binomial distribution test. Unsupervised consensus clustering of H3K27me3-predicted target genes was performed using the top 1,000 genes exhibiting the greatest standard deviation. Supervised analysis of predicted H3K27me3 target genes was also performed using significant analysis of microarrays (SAM) with an FDR cutoff of 0.01.

Sequenom analysis of ependymoma samples. Validation of gene methylation was performed using Sequenom Mass Spectrometry. Primers were designed using Sequenom: EpiDesigner and tested on bisulphite-treated universally methylated DNA (Invitrogen) by standard PCR (Qiagen) followed by Sanger Sequencing. For bisulphite-treated tumour samples, following PCR amplification, amplicons were sent to Genome Quebec for quantification using Sequenom Mass Spectrometry.

Subgroup stratification of ependymoma samples in a validation cohort. Sequenom primers were designed to three highly methylated genes in group A (CIMP$^+$), *PKP1*, *CRIP1*, *CYP26C1*, as selected by CpG coverage and PCR efficiency. PCR amplification was performed in a training data set consisting of the samples, which were analysed by Illumina 450K methylation arrays. These three methylated genes were used to train a classification model using the Prediction Analysis for Microarrays algorithm. Class prediction was performed on a non-overlapping cohort of 82 samples collected from The Hospital for Sick Children, Children's Hospital Boston, University of Michigan, and the MD Anderson Cancer Center. Posterior probabilities corresponding to Group A (CIMP$^+$) or Group B (CIMP$^-$) were calculated for each sample, and an odds ratio >2-fold (probability group A/probability group B) for either subgroup was used to classify tumours. Survival was graphed throughout the manuscript using Kaplan–Meier curves and assessed statistically using a log-rank test.

Pathway analysis of DNA-methylation data. A pathway association test was used to identify groups of functionally related genes that contained a greater than expected number of methylation events in one or the other ependymoma subtype. A gene was considered to be methylated if all profiled sites within 1,000 bp upstream of the transcription start site showed a variance of no more than 0.1 and a mean score of >0.5. For each gene set, methylation events were classified by subtype. The number of methylation events in all genes and the number of methylation events observed in a given gene set were totalled for each subtype. A Fisher's exact test was performed with the null hypothesis that the frequency of methylation events in a given gene set was equal in the two subtypes. To correct for multiple testing, gene names were randomly shuffled and the analysis repeated to obtain a null P value for a given gene set. Randomization was done 10,000 times and, for each gene set, the percentage of null P values that were the same or lower than that obtained from the actual data was used as an estimate of the false discovery rate.

Whole-genome bisulphite sequencing, DNA preparation and differentially methylated region (DMR) analysis. To prepare strand-specific MethylC-seq libraries, adaptor-ligated DNA fragments with insert lengths of 200–250 bp were bisulphite-converted using the EZ DNA Methylation kit (Zymo Research). After PCR amplification in six parallel reactions using the FastStart High Fidelity PCR kit (Roche), library aliquots were pooled per sample and sequenced using the Illumina HiSeq 2000 platform. This yielded an average of 513 million ($\pm$ 102 million (s.d.)) 101-bp paired-end reads per sample.

For analysis of DMR enrichment in specific genomic sites, we first extracted genomic features from UCSC genome browser. Then the percentage of total genomic CpGs for each genomic feature was calculated as a background value. Thereafter, the percentage of total hypermethylated/hypomethylated CpGs in each genomic feature was calculated on the basis of the DMR list. The enrichment fold change was then set as the ratio between the two percentages above. To test the significance of the enrichment/depletion, we randomly permuted the CpGs from all DMRs in the whole genome for 10,000 times and used Fisher's exact test to determine the significance of the difference between the observed and simulated results.

Whole-genome bisulphite sequencing DMR and partially methylated domain (PMD) calling. Whole-genome bisulphate sequencing data was mapped to hg19 using BSMAP (version 2.74). The potential duplications were removed afterwards using Picard tools. BisSNP (version 0.82.2) was then used to detect and remove SNPs and CpGs with potential technical biases before DMR calling. BSmooth was used to smooth bisulphite sequencing data and call candidate DMRs as described previously[49]. PMDs were detected using MethylSeekR.

Ependymoma short-term primary cell culture and *in vitro* drug treatment. Primary ependymoma cells were isolated from patients and cultured on Laminin (Sigma)-coated plates in Neurobasal media (Invitrogen) consisting of N2 (Invitrogen), B27 (Invitrogen), glutamine (Invitrogen), BSA (Sigma), heparin (Sigma), human EGF (Invitrogen) and human basic FGF (Invitrogen). Media was replenished every other day while leaving ~50% conditioned media to encourage continued cell proliferation. Cell viability assays were performed in 96 wells using an Alamar Blue stain (Invitrogen) or MTS Aqueous One (Promega) according to manufacturer's instructions. DAC (Sigma) was dissolved to a stock concentration of 2 mM in PBS and stored in aliquots at $-20\,^{\circ}$C. DAC was prepared fresh and added to treatment media on a daily basis at the appropriate final concentration, for a total of 7 days. DZNep (Cayman Chemical) was dissolved to a stock concentration of 25 mM in DMSO and stored in aliquots at $-20\,^{\circ}$C. DZNep treatments were performed every other day along with replenishment of cell culture medium for a total of 7 days. GSK343 (active compound) and GSK669 (inactive compound) were dissolved in DMSO and used to treat cells at varying concentrations with media replenishment every other day for a total of 11 days.

Gene expression profiling of DAC- and GSK343-treated cultures. Primary cell cultures were treated for 5 days in DAC (500 nM) or GSK343 (500 nM), following

which RNA was isolated using the Trizol (Invitrogen) method. RNA libraries were prepared according to the manufacturer's recommendations and hybridized to Affymetrix Gene 1.0ST arrays. The RMA method with quantile normalization was used for gene expression array normalization. Differentially expressed genes were detected using significance of microarray analysis (FDR < 0.01).

Western blot analysis. Ependymoma cell cultures were lysed in PLC lysis buffer containing deoxycholate, with sonication to facilitate the release of nuclear histones. SDS–PAGE analysis was performed in a 12% gel, loading 20 μg of protein, as quantified by BCA (Pierce). Membranes were blocked with 5% BSA (Roche) diluted in TBST. Western blot antibodies were used at the following concentrations in overnight incubations (2% BSA): EZH2 (Abcam: ab110646, 1:5,000), H3K27me3 (Cell Signaling: no. 9733, 1:5,000), H3K4me3 (Cell Signaling: no. 9751, 1:5,000), cleaved PARP (Cell Signaling: no. 5625, 1:1,000) and α-tubulin (Cell Signaling: no. 2148, 1:20,000). Secondary antibodies were used at a concentration of 1:5,000 for all primary antibodies, and 1:20,000 for α-tubulin.

Flank injections and *in vivo* treatments of immunodeficient mice. For all animal studies, following engraftment of tumour cells, mice were then randomly assigned a treatment of vehicle versus treatment to control for assignment biases and other confounding factors. 50,000 E520-PF1 ependymoma cells were injected subcutaneously into flanks of 5–8-week-old female immunodeficient NOD-SCID gamma mice. Tumours were allowed to develop for 7 days until either visible or palpable. DZNep or vehicle (Sigma: Cremophor) was administered 3 consecutive days a week at a dosage of 3 mg per kg per day via intraperitoneal injections. Tumours were monitored and measured continuously using a caliper. Experimental end point was determined when tumours reached 15 mm in size. Final tumour volumes were determined using caliper measurements. Investigators were blinded during measurement of tumour volumes. A comparison between tumour volumes of DZNep- versus vehicle-treated mice was calculated using a Wilcoxon rank-sum test. For all animal studies, adequate sample sizes were chosen such that any result could be appropriately evaluated statistically using a two-sided nonparametric test. For flank xenograft experiments this entailed a Wilcoxon rank-sum test, and for intracranial experiments this involved a log-rank test.

Cerebellar xenografts and *in vivo* treatments of immunodeficient mice. 10,000 cells were xenografted by stereotactic injection into PFs of female immunodeficient NOD-SCID gamma mice of 5–8 weeks old. Tumours were allowed to develop for 7 days, following which DZNep (3 mg per kg day) or vehicle (Sigma: Cremophor) was administered by intraperitoneal injection. Mice were treated according to the same protocol for flank-tumour-bearing mice (above). Survival of mice was visualized using a Kaplan–Meier curve and quantified using a log-rank test.

Limiting dilution assays (LDAs) of primary ependymoma patient samples or ependymoma xenografts. Cells from a lung metastasis resection or tumour xenograft were dissociated according to previously published protocols[50]. LDAs were performed in a 96-well plate format. LDAs from xenografts were not treated with inhibitors but monitored for neurosphere colony formation. For primary patient samples serial dilutions of cells were performed to reach cell doses of 2,000 cells per well at the highest dose and 4 cells per well at the lowest dose. A total of 10 cell doses were tested with 6 technical replicates per dose. Cells were treated with selected compounds about 5 h post-surgery. GSK343 was used at a concentration of 3 μM, and DAC was used at 0.5 μM. Fresh media and drugs were added to the cells after 7 days. Wells were scored for sphere formation on day 14. Statistical analysis was performed with the Extreme Limiting Dilution Analysis web-based software.

48. Lupien, M. *et al.* FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970 (2008).
49. Hansen, K. D. *et al.* Increased methylation variation in epigenetic domains across cancer types. *Nature Genet.* **43**, 768–775 (2011).
50. Gallo, M. *et al.* A tumorigenic MLL-homeobox network in human glioblastoma stem cells. *Cancer Res.* **73**, 417–427 (2013).