nature genetics

Genomic analysis of diffuse intrinsic pontine gliomas identifies three molecular subgroups and recurrent activating *ACVR1* mutations

Pawel Buczkowicz[1–3,24], Christine Hoeman[4,24], Patricia Rakopoulos[2,3], Sanja Pajovic[2], Louis Letourneau[5], Misko Dzamba[6], Andrew Morrison[2], Peter Lewis[7], Eric Bouffet[8], Ute Bartels[8], Jennifer Zuccaro[2], Sameer Agnihotri[2], Scott Ryall[2], Mark Barszczyk[2,3], Yevgen Chornenkyy[2,3], Mathieu Bourgey[5], Guillaume Bourque[5], Alexandre Montpetit[5], Francisco Cordero[4], Pedro Castelo-Branco[2], Joshua Mangerel[2], Uri Tabori[2,8], King Ching Ho[2], Annie Huang[2,8], Kathryn R Taylor[9], Alan Mackay[9], Anne E Bendel[10], Javad Nazarian[11], Jason R Fangusaro[12], Matthias A Karajannis[13], David Zagzag[13], Nicholas K Foreman[14], Andrew Donson[14], Julia V Hegert[15], Amy Smith[15], Jennifer Chan[16], Lucy Lafay-Cousin[16], Sandra Dunn[17], Juliette Hukin[17], Chris Dunham[17], Katrin Scheinemann[18], Jean Michaud[19], Shayna Zelcer[20], David Ramsay[20], Jason Cain[21], Cameron Brennan[22], Mark M Souweidane[22], Chris Jones[9], C David Allis[7], Michael Brudno[6,23], Oren Becher[4,25] & Cynthia Hawkins[1–3,25]

Diffuse intrinsic pontine glioma (DIPG) is a fatal brain cancer that arises in the brainstem of children, with no effective treatment and near 100% fatality. The failure of most therapies can be attributed to the delicate location of these tumors and to the selection of therapies on the basis of assumptions that DIPGs are molecularly similar to adult disease. Recent studies have unraveled the unique genetic makeup of this brain cancer, with nearly 80% found to harbor a p.Lys27Met histone H3.3 or p.Lys27Met histone H3.1 alteration. However, DIPGs are still thought of as one disease, with limited understanding of the genetic drivers of these tumors. To understand what drives DIPGs, we integrated whole-genome sequencing with methylation, expression and copy number profiling, discovering that DIPGs comprise three molecularly distinct subgroups (H3-K27M, silent and MYCN) and uncovering a new recurrent activating mutation affecting the activin receptor gene *ACVR1* in 20% of DIPGs. Mutations in *ACVR1* were constitutively activating, leading to SMAD phosphorylation and increased expression of the downstream activin signaling targets *ID1* and *ID2*. Our results highlight distinct molecular subgroups and novel therapeutic targets for this incurable pediatric cancer.

Brain tumors are the largest group of solid tumors and the leading cause of cancer-related deaths in childhood[1]. The most devastating of these is DIPG, which is almost universally fatal[2,3]. Despite collaborative efforts to improve treatments, survival has remained static over decades, and DIPGs are now the main cause of brain tumor–related death in children. Diagnosis of DIPG is based on a combination of clinical and radiological findings; a tissue biopsy is rarely acquired. Radiation is the mainstay of therapy but offers only symptom control, and, so far, chemotherapy has shown no benefit[4]. A potential contributor to the failure of DIPG clinical trials is the use of agents targeting the genetic alterations of adult glioblastomas (GBMs)[5]. A number of recent studies have reported differences at both the copy number and expression levels that distinguish pediatric DIPGs from both their adult and pediatric supratentorial GBM counterparts[6–9], indicating that they may be separate biological entities requiring their own therapeutic strategies. Recent identification of frequent histone H3 gene mutations (encoding p.Lys27Met) in DIPGs has suggested that both genetic and epigenetic mechanisms may be important drivers of these tumors[10–12]. However, the complete genetic and epigenetic landscapes of DIPG and the functional role that histone modifications may have in DIPG remain unknown. These data are critical for the development of better therapies for affected children.

We integrated deep sequencing analysis of 36 tumor-normal pairs (20 whole-genome sequencing (Illumina HiSeq 2000) and 16 whole-exome sequencing (Applied Biosystems SOLiD 5500xl)) with comprehensive methylation (28 DIPGs; Illumina Infinium450k methylation array), copy number (45 DIPGs; Affymetrix SNP6.0) and expression (35 DIPGs; Illumina HT-12 v4) data (**Supplementary Table 1**). All coding somatic single-nucleotide variants (SNVs) identified in the combined cohort are listed in **Supplementary Table 2**, including new mutations in *ACVR1*. We did not find mutations in *BRAF* or *IDH1* or structural rearrangements in *FGFR1* or *MYB* in our DIPG cohort. Verification of somatic alterations was conducted on all 36

A full list of author affiliations appears at the end of the paper.

Received 24 April 2013; accepted 5 March 2014; published online 6 April 2014; doi:10.1038/ng.2936

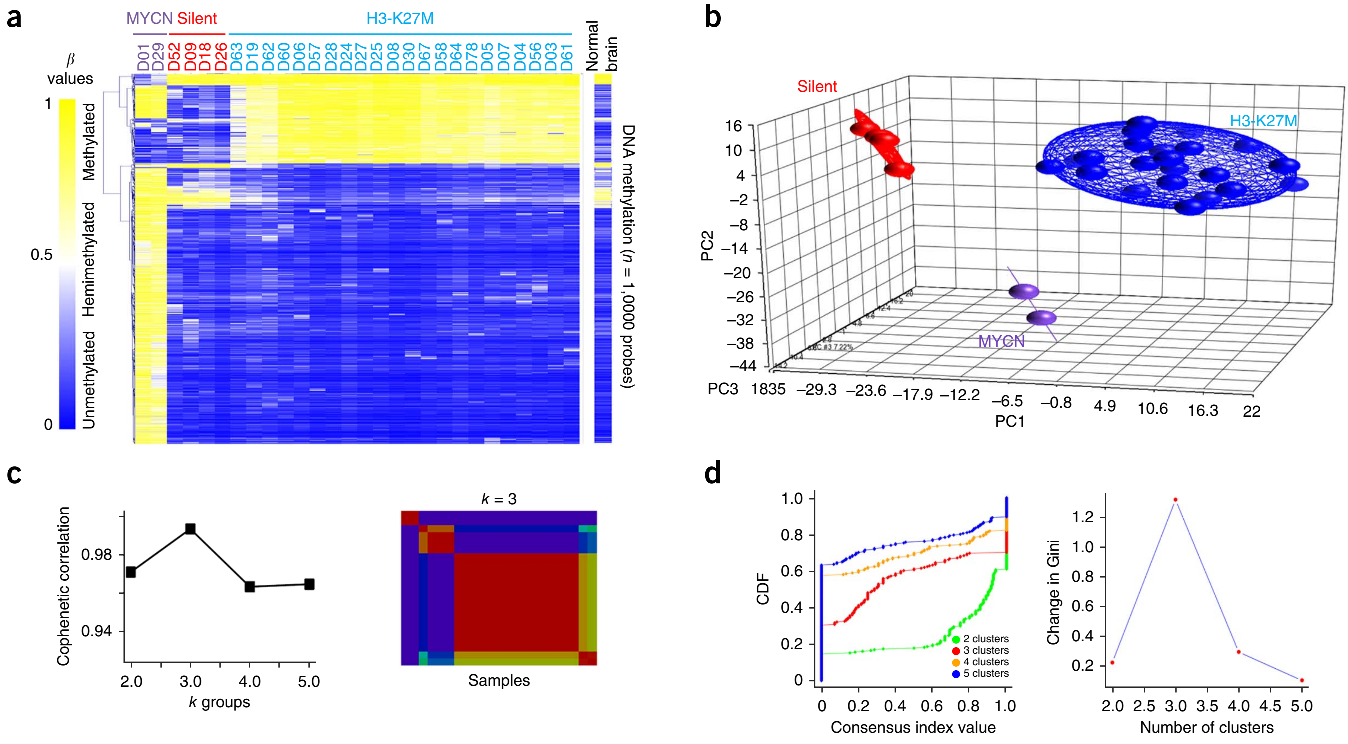

Figure 1 Methylation profiling identifies three molecular subgroups of DIPG. (**a**) Heat map of methylation levels in three DIPG subgroups identified by unsupervised hierarchical clustering. (**b**–**d**) Subrouping was supported by principal-components analysis (**b**), non-negative matrix factorization (cophenetic coefficient = 0.9934, *k* = 3) (**c**) and consensus clustering represented by a cumulative distribution function (CDF) and change in Gini coefficient (**d**). PC, principal component.

sample pairs using Fluidigm Array Ion Torrent chip sequencing (**Supplementary Table 3**). Further, for key SNVs, we also tested a validation cohort of an additional 25 tumors. This analysis shows that, although previously considered to be one disease, DIPG represents three distinct subgroups with different methylation, expression, copy number alteration (CNA) and mutational profiles.

Unsupervised subgrouping of DIPG cases on the basis of CpG island methylation (Online Methods) resulted in three distinct subgroups; 'MYCN', 'silent' and 'H3-K27M' (**Fig. 1a**). This subgrouping was supported by multiple analyses, including principal-components analysis (**Fig. 1b**), non-negative matrix factorization (**Fig. 1c**) and consensus clustering (**Fig. 1d**). Subgroup-specific differences were supported by the integration of mutation, structural, expression and clinical data (**Fig. 2**). The identification of DIPG subgroups will have key implications for the design of appropriate therapy for these tumors.

The MYCN subgroup had no recurrent mutations but was instead characterized by hypermethylation, high-grade histology and chromothripsis on chromosome 2p leading to recurrent high-level amplification of *MYCN* and *ID2* (**Table 1** and **Supplementary Figs. 1**–**3**). Tumors in this group overexpressed *MYCN* by 4-fold and 8-fold and *ID2* by 2.5-fold and 5-fold relative to the H3-K27M and silent groups, respectively. The top most overexpressed genes in the MYCN group included *FAP*, *HRSP12* and *DYX2*. Therapies aimed at targeting altered histone modifications would not be effective in this subgroup. Rather, children with this subtype of DIPG will potentially benefit from therapies targeting MYCN or possibly ID2. Tumors in the silent subgroup had silent genomes on the basis of both whole-genome sequencing, structural and SNP6.0 copy number analysis and had a lower mutation rate than tumors in the other two subgroups (median of 0.11 mutations per megabase (range of 0.02–0.19) in the silent cohort compared to 0.99 mutations per megabase (range of 0.19–24) in the other two groups; $P = 0.05$ (**Supplementary Table 4**)). All DIPGs with low-grade astrocytoma (LGA) histology were from this group, although, interestingly, there was no difference in overall survival in comparison with the other subgroups. Affected individuals in this group were diagnosed at a significantly younger age (4.81 ± 1.64 years for the children from the silent group versus 6.89 ± 2.62 years for those not in the silent group; $P = 0.04$). The p.Lys27Met histone H3 alteration was present in 44% of tumors (p.Lys27Met histone H3.3 in 33% and p.Lys27Met histone H3.1 in 11%), but there were no recurrent copy number changes as observed in the MYCN and H3-K27M subgroups. At the expression level, the silent subgroup showed overexpression of WNT pathway genes (**Supplementary Table 5**), as well as *MDM2*, *MSMP* and *ADAM33*, compared with the other two subgroups (**Table 1**). Notably, DIPGs in neither the MYCN nor the silent subgroup had receptor tyrosine kinase gene amplification, suggesting that the group of inhibitors targeting these kinases will be less effective in patients with these DIPG subtypes.

H3-K27M subgroup DIPGs were highly mutated in either histone H3.3 (*H3F3A*) or H3.1 (*HIST1H3B* and *HIST1H3C*). This group had highly unstable genomes (segmentation analysis of SNP6.0 data; 497 CNAs per genome versus 300 CNAs per genome in the silent group; $P = 0.04$). Alternative lengthening of telomeres (ALT; detected by telomere restriction fragment assay[13], the presence of C-circles[14] and/or telomere length from whole-genome sequencing data) is exclusively associated with the H3-K27M subgroup; only two of the tumors in this subgroup harbored *ATRX* mutations. ALT-positive cases were significantly older at the time of diagnosis ($P < 0.001$; **Supplementary Fig. 4**). Similarly, *PVT1*, *MYC* and *PDGFRA* gains or amplifications (**Fig. 2a**) and structural variants (**Supplementary Fig. 5**) were exclusive to this group. *TP53* mutations were enriched in this group (67.9% versus 33.3% in other groups; $P = 0.007$). Given the complexity and

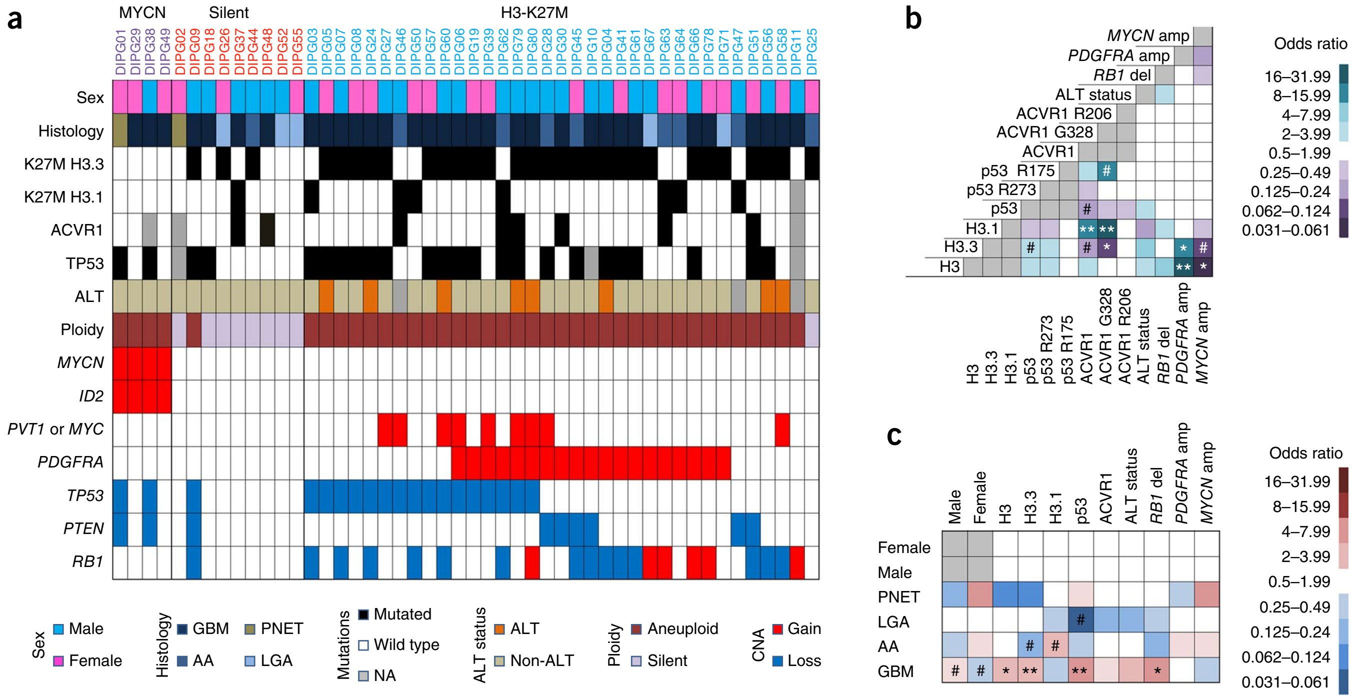

Figure 2 Molecular subgroups of DIPG share common clinical features and recurrent genomic events. (**a**) Clinical and genomic features such as sex, histology, frequency of recurrent mutations, ALT and CNAs are represented in a DIPG subgroup–specific manner. NA, not applicable; GBM, glioblastoma; PNET, primitive neuroectodermal tumor; AA, anaplastic astrocytoma; LGA, low-grade astrocytoma. (**b**) Probability of two mutational or structural features of DIPG co-occurring on the basis of OR suggests statistically significant association between p.Lys27Met histone H3.3 alteration and *PDGFRA* amplifications (OR = 8.0, $P = 0.0127$) and between p.Lys27Met histone H3.1 alteration and *ACVR1* mutations (OR = 15.8, $P < 0.001$). Amp, amplification; del, deletion. (**c**) Probability of mutations or structural events in DIPG occurring with a clinical feature such as sex or tumor histology on the basis of OR shows statistically significant correlation between *TP53* mutations and GBM histology (OR = 10.8, $P < 0.005$), among others. $^{\#}P < 0.05$, $^{*}P < 0.01$, $^{**}P < 0.001$.

heterogeneity of genetic changes for the H3-K27M subgroup, if therapies targeting the histone mutations become available, this subgroup would likely require multimodal therapies.

Certain mutations and structural variants were significantly more likely to co-occur—specifically, mutation encoding p.Lys27Met in histone H3.3 and *PDGFRA* amplifications (odds ratio (OR) = 8.0, $P = 0.0127$) and mutation encoding p.Lys27Met in histone H3.1 and *ACVR1* mutations (OR = 15.8, $P = 0.0004$) (**Fig. 2b**). Conversely, *MYCN* amplifications were statistically less likely to occur if a p.Lys27Met histone H3 alteration was present (OR = 0.019, $P = 0.0103$; **Fig. 2b**). Affected individuals carrying p.Lys27Met alterations in histone H3.3 were more likely to have GBM histology (OR = 5.3, $P = 0.0035$), whereas p.Lys27Met alterations in histone H3.1 were more likely to occur in cases with anaplastic astrocytoma histology (OR = 7.1, $P = 0.016$) and in younger cases (p.Lys27Met histone H3.1: 4.11 ± 2.03 years at diagnosis versus wild-type histone H3.1: 6.50 ± 3.50 years at diagnosis; $P = 0.040$) (**Fig. 2c**). *TP53* mutations were more likely to occur in cases with GBM histology regardless of p.Lys27Met status (OR = 8.3, $P = 0.0039$) and were unlikely to occur in cases with LGA histology (OR = 0.046, $P = 0.0040$) (**Fig. 2c**). A summary of the molecular and clinical features observed among cases in the MYCN, silent and H3-K27M DIPG subgroups is provided in **Table 1**.

After *H3F3A* and *TP53*, the next most frequently mutated gene in DIPG was *ACVR1* (encoding activin A receptor, type I), a new cancer gene. Mutations of *ACVR1* in four DIPGs (c.617G>A) resulted in a p.Arg206His substitution (**Fig. 3a**). One DIPG had a mutation affecting a neighboring codon (encoding p.Gln207Glu). Two DIPGs had a c.983G>A (p.Gly328Glu) mutation and five DIPGs in our cohort had a c.983G>T mutation that resulted in a p.Gly328Val substitution (**Fig. 3a**). In total, 20% of DIPGs had *ACVR1* mutations. *ACVR1* encodes the activin A (ALK2) receptor involved in BMP (bone morphogenetic protein) signaling[15]. The p.Arg206His substitution is in the GS (glycine-serine-rich) domain and is the most common alteration responsible for an autosomal dominant disease of the connective tissue, fibrodysplasia ossificans progressiva (FOP), in which endothelial cells are sensitized to BMP signaling, transitioning to osteoblasts[15]. The p.Gly328Glu and p.Gln207Glu substitutions are described as rare alterations in individuals with FOP[16,17]. The p.Gly328Val substitution has not previously been reported in individuals with FOP or in cancer. Mutations affecting the Gly328 residue map to the kinase domain and are spatially close to the GS domain[17], and, similar to the p.Arg206His substitution, they are thought to weaken the interaction of the GS domain with negative regulators of ACVR1 such as FKBP12, resulting in ligand-independent constitutive activation of the receptor. Levels of phosphorylated SMAD1 and SMAD5 (SMAD1/5) were higher in DIPG cases with *ACVR1* mutations compared to cases with wild-type *ACVR1* (**Fig. 3b**).

To determine the biological consequence of *ACVR1* mutations *in vitro*, we transfected immortalized normal human astrocytes (iNHAs) with 3× Flag–tagged pCDH511b expression vectors containing Gly328Val ACVR1 and/or Lys27Met histone H3.3. Protein blot analysis showed endogenous expression of ACVR1 in all cells, with Flag-ACVR1 and higher total ACVR1 expression detected in transfected cells. To determine whether the ACVR1 mutant activated the BMP signaling pathway, we tested for phosphorylated SMAD1/5 by protein blot. Only cells expressing Gly328Val ACVR1 showed positivity for phosphorylated SMAD1/5 (**Fig. 3c**). To test downstream pathway activation, we performed quantitative PCR for *ID1* and *ID2*, which are known targets of BMP signaling[18–20]. Expression of either

Table 1 Clinical, genetic and epigenetic features of DIPG molecular subgroups

	MYCN	Silent	H3-K27M
Sex			
Male:female	1:3	2:1	1:1
Histology			
Grade II	0%	33%	0%
Grade III	0%	11%	26%
Grade IV	100%	56%	74%
Mutations			
p.Lys27Met histone H3.3	0%	35%	97%
ACVR1	0%	22%	18%
CpG methylation			
Median β value	75%	13%	15%
CNAs	*MYCN* amplification, chr. 2p chromothripsis	Few copy number changes, silent genome	*PVT1*/*MYC* and *PDGFRA* gain/amplification, *RB1* and *TP53* deletion
ALT phenotype			
ALT[a]	0%	0%	22%
Gene expression profiling	Overexpression of *DYX2, HSPR12, FAP*	Overexpression of *MDM2, MSMP, ADAM33*	Underexpression of *VAX2, HOXC6, HOXA9, HOXA7, DOXD4*
Age at diagnosis			
Median age	6.3 years	5 years	7 years

[a] Alternative lengthening of telomeres.

Lys27Met histone H3.3 or Gly328Val ACVR1 alone led to a twofold increase in *ID1* and *ID2* expression levels, whereas coexpression of these proteins resulted in a nearly fourfold increase in *ID1* and *ID2* expression levels, suggesting an additive effect (**Fig. 3d**). Gly328Val ACVR1 increased the growth rate of iNHAs compared to empty vector controls ($P = 0.0034$; **Fig. 3e**). An increase in the levels of phosphorylated SMAD1/5 was also seen by protein blot in postnatal day (P) 3 mouse brainstem progenitor cells from *Nes–tv-a*; *Tp53*^{fl} mice infected *in vitro* with RCAS (replication-competent avian sarcoma-leukosis virus long terminal repeat (LTR) with a splice acceptor) expressing Arg206His or Gly328Val mouse ACVR1 compared to wild-type ACVR1 ($P < 0.05$; data not shown). Brainstem progenitor cells infected with virus expressing mutant ACVR1 showed increased proliferation by BrdU (bromodeoxyuridine) incorporation ($P < 0.05$; **Fig. 3f**).

To assess the impact of Lys27Met histone H3.3 on gene expression and methylation patterns as well as its potential transformative capacity, we transfected iNHA cells with expression vectors encoding N-terminally Flag-tagged wild-type or Lys27Met histone H3.3 or with empty vector control. Clones expressing comparable amounts of wild-type and Lys27Met histone H3.3 protein, according to protein blot analysis (**Supplementary Fig. 6a**), that was verified by immunofluorescence staining to localize to the nucleus (**Supplementary Fig. 6b**) were selected and pooled to remove clonal variations in phenotype. Compared to controls, iNHA cells expressing Lys27Met histone H3.3 had a significantly reduced proliferation rate when grown adherently in DMEM ($P < 0.0001$; **Supplementary Fig. 6c**) and grew semi-adherently when seeded in neural stem cell (NSC) medium (**Supplementary Fig. 6e**). Compared to adherent iNHA cells expressing Lys27Met histone H3.3, the semi-adherent iNHA cells expressing Lys27Met histone H3.3 in NSC medium had higher expression of SOX2 by immunofluorescence staining with no marked changes in GFAP, nestin, TUJ1 or O4 levels (**Supplementary Fig. 7**). Alterations in both the expression (Illumina HT-12v4) and methylation (Illumina Infinium450k) profiles of the cells expressing Lys27Met histone H3.3 compared to both wild-type and empty vector controls were observed (**Supplementary Fig. 6f**). The top pathways perturbed by

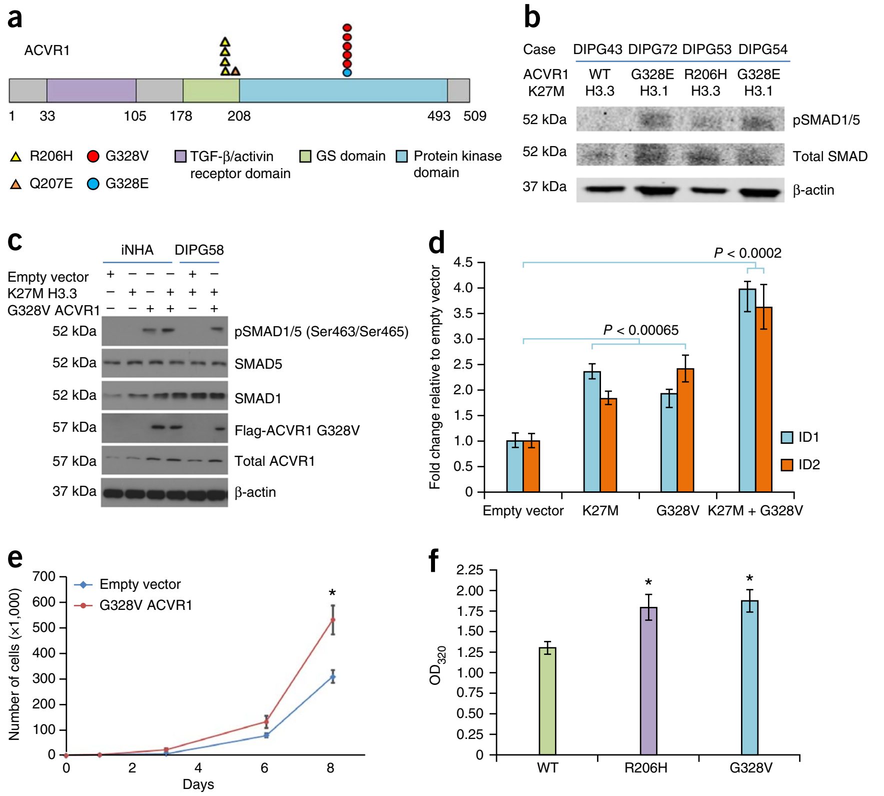

Figure 3 *ACVR1* mutations constitutively activate BMP signaling *in vitro* and in *ACVR1*-mutant DIPG. (**a**) Four mutations (encoding p.Arg206His, p.Gln207Glu, p.Gly328Gly and p.Gly328Val) were detected in 12 of 61 DIPG cases. The p.Arg206His and p.Gln207Glu alterations occur in the GS domain, and the Gly328 alterations occur in the protein kinase domain. (**b**) Human DIPGs with *ACVR1* mutations have increased levels of phosphorylated SMAD1/5 (pSMAD1/5) compared with DIPGs with wild-type (WT) *ACVR1*. (**c**) Protein blot showing increased levels of phosphorylated SMAD1/5 in *ACVR1*-mutant iNHA and DIPG58 cells transfected with construct expressing Gly328Val ACVR1 compared to control cells. (**d**) RT-PCR in iNHA cells transfected with empty vector or with vector expressing Lys27Met histone H3.3, Gly328Val ACVR1 or a combination of both shows increase in *ID1* and *ID2* gene expression compared to empty vector control. Error bars, s.d.; $n = 3$. (**e**) iNHA cells expressing mutant Gly328Val ACVR1 have an increased growth rate compared to empty vector controls. $^*P = 0.0034$. Error bars, s.d.; $n = 3$. (**f**) Compared to mouse brainstem progenitor cultures with wild-type *ACVR1*, those with mutant *ACVR1* have significantly higher BrdU incorporation, suggesting increased proliferation. $^*P < 0.05$. Error bars, s.e.m.; $n = 3$. OD_{320}, optical density at 320 nm.

Lys27Met histone H3.3 expression corresponded to molecular and cellular functions, especially an increase in cell-to-cell signaling ($P = 0.00882$, z score = 2.365) and a decrease in cell cycle progression ($P < 0.00001$, z score = −1.347). The top molecular and cellular functions affected by methylation changes in cells expressing mutant Lys27Met histone H3.3 were embryonic development ($P < 0.00001$), decreased cell growth and proliferation ($P = 0.00103$) and increased cell-to-cell signaling ($P < 0.00001$). Cells expressing Lys27Met histone H3.3 did not form colonies in soft agar. These cells showed reduced global levels of trimethylation at lysine 27 of histone H3 (H3K27me3) compared to controls (**Supplementary Fig. 6d**). Similarly, immunohistochemical staining of DIPGs showed lower global H3K27me3 levels in tumors positive for Lys27Met histone H3.3 compared to tumors with wild-type histone H3.3 (**Supplementary Fig. 6g,h**). Global levels of acetylation at lysine 27 of histone H3 (H3K27ac), trimethylation at lysine 4 of histone H3 (H3K4me3) and acetylation at lysine 9 of histone H3 (H3K9ac) were not altered in a mutation-dependent manner (data not shown).

Interestingly, expression of Lys27Met histone H3.3 also led to increased *ID1* and *ID2* expression yet did not induce phosphorylation of SMAD1/5, suggesting an alternate pathway for its action. Phosphorylated SMAD1/5 was also only observed in DIPG cases with *ACVR1* mutations and not in ones expressing the Lys27Met histone H3 mutant that had wild-type *ACVR1*. Our *in vitro* data suggest an additive effect of the combination of mutant *ACVR1* and *H3F3A* in increasing *ID1* and *ID2* levels. Further, a subset of DIPG cases with p.Lys27Met histone H3.3 alterations overexpressed *ID2*, and MYCN subgroup cases had genomic amplification of *ID2* with corresponding increased expression, suggesting a potentially common, ID2-related mechanism in DIPG tumorigenesis. ID2 may have a role in negatively regulating cell differentiation[21], and its expression has been associated with poor prognosis in other cancers[22–24]. Morphological changes and differential expression of stem cell markers, as well as global methylation and gene expression changes in iNHA cells expressing Lys27Met histone H3.3, suggest a cellular reprogramming event. However, further investigation into the effects of p.Lys27Met histone H3.3 alteration and *ACVR1* mutations in DIPG tumorigenesis is warranted.

Our results highlight the many pathways to tumorigenesis in DIPG. This complexity needs to be considered when designing new therapeutic approaches to improve outcome for affected children.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Whole-genome sequencing data are accessible through the European Genome-phenome Archive (accession EGAS00001000575). Methylation data are accessible through the Gene Expression Omnibus (accession GSE50022). Gene expression data are accessible through the Gene Expression Omnibus (accession GSE50021). SNP6.0 copy number data are accessible through the Gene Expression Omnibus (accession GSE50024).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We would like to thank all of the patients and families for donating tissue for this research. This work was supported by the Canadian Institutes of Health Research (CIHR, MOP 115004) and was funded in part by a Genome Canada/CIHR grant (cofunding from Genome BC, Génome Québec, CIHR-ICR (Institute for Cancer Research) and C17, through the Genome Canada/CIHR joint ATID Competition (project title: The Canadian Paediatric Cancer Genome Consortium (CPCGC): Translating Next-Generation Sequencing Technologies into Improved Therapies for High-Risk Childhood Cancer)). P.B. is a recipient of a CIHR Doctoral Frederick Banting and Charles Best Canada Graduate Scholarships award. O.B. is a Damon Runyon Clinical Investigator and is supported by the US Department of Defense and the Pediatric Brain Tumor Foundation. C.J., A. Mackay and K.R.T. acknowledge National Health Service (NHS) funding to the Biomedical Research Centre and support from the Stavros Niarchos Foundation. Sample collection for M.A.K. and D.Z. was supported in part by grant UL1TR000038 from the National Center for Research Resources, US National Institutes of Health and by grant 5P30CA016087-32 from the National Cancer Institute.

AUTHOR CONTRIBUTIONS

E.B., U.B., P.B., O.B. and C. Hawkins designed the study. P.B., C. Hoeman, F.C., P.R., S.P., A. Morrison, J.Z., S.A., S.R., M. Barszczyk, Y.C., P.C.-B., K.C.H. and J. Mangerel performed experiments. P.B., C. Hoeman, F.C., P.R., L.L., M.D., M. Bourgey, G.B. and A. Montpetit collected and analyzed data. O.B., C. Hawkins, C.J., K.R.T., A. Mackay, A.E.B., J.N., J.R.F., M.A.K., D.Z., N.K.F., A.D., J.V.H., A.S., J. Chan, L.L.-C., S.D., J.H., C.D., K.S., J. Michaud, S.Z., D.R., J. Cain, M.M.S., E.B., U.T. and U.B. provided reagents, tissue and mice. P.B., P.R., S.P., M.D., O.B. and C. Hawkins wrote the manuscript. P.L., C.B., C.D.A., M. Brudno, A.H. and U.T. gave technical support and conceptual advice. All authors approved the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Stiller, C.A. Population based survival rates for childhood cancer in Britain, 1980–1991. *Br. Med. J.* **309**, 1612–1616 (1994).
2. Freeman, C.R. & Perilongo, G. Chemotherapy for brain stem gliomas. *Childs Nerv. Syst.* **15**, 545–553 (1999).
3. Maria, B.L. *et al.* Brainstem glioma: I. Pathology, clinical features, and therapy. *J. Child Neurol.* **8**, 112–128 (1993).
4. Hargrave, D., Bartels, U. & Bouffet, E. Diffuse brainstem glioma in children: critical review of clinical trials. *Lancet Oncol.* **7**, 241–248 (2006).
5. Donaldson, S.S., Laningham, F. & Fisher, P.G. Advances toward an understanding of brainstem gliomas. *J. Clin. Oncol.* **24**, 1266–1272 (2006).
6. Buczkowicz, P. *et al.* Aurora kinase B is a potential therapeutic target in pediatric diffuse intrinsic pontine glioma. *Brain Pathol.* **23**, 244–253 (2013).
7. Paugh, B.S. *et al.* Genome-wide analyses identify recurrent amplifications of receptor tyrosine kinases and cell-cycle regulatory genes in diffuse intrinsic pontine glioma. *J. Clin. Oncol.* **29**, 3999–4006 (2011).
8. Paugh, B.S. *et al.* Integrated molecular genetic profiling of pediatric high-grade gliomas reveals key differences with the adult disease. *J. Clin. Oncol.* **28**, 3061–3068 (2010).
9. Zarghooni, M. *et al.* Whole-genome profiling of pediatric diffuse intrinsic pontine gliomas highlights platelet-derived growth factor receptor α and poly (ADP-ribose) polymerase as potential therapeutic targets. *J. Clin. Oncol.* **28**, 1337–1344 (2010).
10. Khuong-Quang, D.A. *et al.* K27M mutation in histone H3.3 defines clinically and biologically distinct subgroups of pediatric diffuse intrinsic pontine gliomas. *Acta Neuropathol.* **124**, 439–447 (2012).
11. Schwartzentruber, J. *et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**, 226–231 (2012).
12. Wu, G. *et al.* Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat. Genet.* **44**, 251–253 (2012).
13. Herbert, B.S., Shay, J.W. & Wright, W.E. Analysis of telomeres and telomerase. *Curr. Protoc. Cell Biol.* Chapter 18, Unit 18.6 (2003).
14. Henson, J.D. *et al.* DNA C-circles are specific and quantifiable markers of alternative-lengthening-of-telomeres activity. *Nat. Biotechnol.* **27**, 1181–1185 (2009).
15. Song, G.A. *et al.* Molecular consequences of the ACVR1^{R206H} mutation of fibrodysplasia ossificans progressiva. *J. Biol. Chem.* **285**, 22542–22553 (2010).
16. Kaplan, F.S. *et al.* Classic and atypical fibrodysplasia ossificans progressiva (FOP) phenotypes are caused by mutations in the bone morphogenetic protein (BMP) type I receptor ACVR1. *Hum. Mutat.* **30**, 379–390 (2009).
17. Petrie, K.A. *et al.* Novel mutations in ACVR1 result in atypical features in two fibrodysplasia ossificans progressiva patients. *PLoS ONE* **4**, e5005 (2009).
18. Kowanetz, M., Valcourt, U., Bergstrom, R., Heldin, C.H. & Moustakas, A. Id2 and Id3 define the potency of cell proliferation and differentiation responses to transforming growth factor β and bone morphogenetic protein. *Mol. Cell. Biol.* **24**, 4241–4254 (2004).
19. Kurooka, H., Nakahiro, T., Mori, K., Sano, K. & Yokota, Y. BMP signaling is responsible for serum-induced Id2 expression. *Biochem. Biophys. Res. Commun.* **420**, 281–287 (2012).
20. Shin, M. *et al.* Identification of a novel bone morphogenetic protein (BMP)-inducible transcript, BMP-inducible transcript-1, by utilizing the conserved BMP-responsive elements in the Id genes. *J. Bone Miner. Metab.* **31**, 34–43 (2013).

21. Chen, X.S., Zhang, Y.H., Cai, Q.Y. & Yao, Z.X. ID2: a negative transcription factor regulating oligodendroglia differentiation. *J. Neurosci. Res.* **90**, 925–932 (2012).
22. Liu, Z. *et al.* Overexpressed DNA-binding protein inhibitor 2 as an unfavorable prognosis factor promotes cell proliferation in nasopharyngeal carcinoma. *Acta Biochim. Biophys. Sin. (Shanghai)* **44**, 503–512 (2012).
23. Rollin, J. *et al.* The intracellular localization of ID2 expression has a predictive value in non small cell lung cancer. *PLoS ONE* **4**, e4158 (2009).
24. Wazir, U., Jiang, W.G., Sharma, A.K., Newbold, R.F. & Mokbel, K. The mRNA expression of inhibitors of DNA binding-1 and -2 is associated with advanced tumour stage and adverse clinical outcome in human breast cancer. *Anticancer Res.* **33**, 2179–2183 (2013).

[1]Division of Pathology, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada. [2]Arthur and Sonia Labatt Brain Tumour Research Centre, The Hospital for Sick Children, Toronto, Ontario, Canada. [3]Department of Laboratory Medicine and Pathobiology, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada. [4]Division of Pediatric Hematology/Oncology, Duke University Medical Center, Durham, North Carolina, USA. [5]Génome Québec Innovation Centre, McGill University, Montreal, Quebec, Canada. [6]Department of Computer Science, University of Toronto, Toronto, Ontario, Canada. [7]Laboratory of Chromatin Biology and Epigenetics, Rockefeller University, New York, New York, USA. [8]Division of Haematology/Oncology, The Hospital for Sick Children, Toronto, Ontario, Canada. [9]Division of Cancer Therapeutics, Institute of Cancer Research, London, UK. [10]Department of Pediatric Hematology-Oncology, Children's Hospitals and Clinics of Minnesota, Minneapolis, Minnesota, USA. [11]Center for Genetic Medicine, Children's National Medical Center, Washington, DC, USA. [12]Department of Pediatrics-Hematology, Ann & Robert H. Lurie Children's Hospital of Chicago, Chicago, Illinois, USA. [13]Division of Pediatric Hematology/Oncology and Neuropathology, New York University Cancer Institute, New York University Langone Medical Center, New York, New York, USA. [14]Department of Pediatrics, Children's Hospital Colorado, Denver, Colorado, USA. [15]Department of Pathology, Arnold Palmer Hospital for Children, Orlando, Florida, USA. [16]Clark H. Smith Brain Tumour Centre, Faculty of Medicine, University of Calgary, Calgary, Alberta, Canada. [17]Department of Pediatric Neurology and Neuro-Oncology, BC Children's Hospital, Vancouver, British Columbia, Canada. [18]Division of Hematology/Oncology, McMaster Children's Hospital, Hamilton, Ontario, Canada. [19]Department of Pathology and Laboratory Medicine, Children's Hospital of Eastern Ontario, Ottawa, Ontario, Canada. [20]Department of Hematology/Oncology, Children's Hospital London Health Sciences Centre, London, Ontario, Canada. [21]Monash Institute of Medical Research, Monash Medical Centre, Clayton, Victoria, Australia. [22]Department of Neurological Surgery, Weill Cornell Medical College and Memorial Sloan-Kettering Cancer Center, New York, New York, USA. [23]The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada. [24]These authors contributed equally to this work. [25]These authors jointly directed this work. Correspondence should be addressed to C. Hawkins (cynthia.hawkins@sickkids.ca) or O.B. (oren.becher@duke.edu).

ONLINE METHODS

Patients and samples. Biological material and clinical data were gathered for 74 DIPG samples, including normal brain and/or peripheral blood samples, if available. DIPGs were diagnosed by a neuroradiologist on the basis of magnetic resonance imaging (MRI). A contrasting lesion with diffuse involvement of at least 50% of the pons was required for DIPG diagnosis. All patient material was collected after receiving informed consent and was approved by the institutional review board of contributing centers. Samples for 20 of these patients represented pretreatment samples (2 non-treated patients from autopsy), and 54 samples represented post-treatment autopsy samples. The median age of diagnosis was 6.37 years, with a median survival time of 10.4 months.

Whole-genome sequencing alignment and structural variants. The entire genomes of 20 DIPG tumor and matched normal DNA pairs were sequenced using next-generation Illumina technologies. DNA samples were paired-end sequenced on an Illumina HiSeq 2000 using DNA fragments of 259–333 bp in size with 100-bp read length. Mean coverage was 35–67× for tumor samples and 28–74× for normal samples. The Burrows-Wheeler Aligner (BWA) tool was used to map all reads to GRCh37/hg19 using default parameters[25]. PCR duplicates were removed from alignments using Picard 1.77 (**Supplementary Table 6**). Structural variants were identified using PRISM 1.1.6 and PRISM CTX 1.0.1 (ref. 26). Recurrent structural variants (**Supplementary Table 7**) were found after subtracting common structural variants[27] and all structural variants found in normal tissue and were visually validated using the DNAC algorithm and SNP6.0 data in Partek Genomics Suite (v.6.6) as well as whole-genome sequencing data in Savant (v2.0.3)[28] and the Integrated Genomics Viewer (v2.2). Prediction of structural variants in chromothriptic regions was performed by discordant read-pair clustering. Discordant read pairs (fragment size $> 3 \times \sigma + \mu$) were initially selected by greedy clustering using a sliding window of $15 \times \sigma$ and were refined by clipped-read mappings, where σ is the s.d. of the insert size and μ is the average insert size (**Supplementary Fig. 8**). Localization of a chromothriptic event was predicted using a hidden Markov model (HMM) informed by depth of coverage and discordant read pairs. Maximum-likelihood copy counts of chromothriptic regions were determined as previously described[29].

Exome capture was carried out with the SOLiD platform (Applied Biosystems). Sequence data were aligned to the GRCh37/hg19 human reference genome assembly; duplicates and reads with non-unique mapping were excluded. We validated 151 sequence variants in 128 genes using PCR amplification by Fluidigm arrays and Ion Torrent chips (Life Technologies). These SNVs were selected for validation on the basis of variant allele frequency from whole-genome sequencing and/or whole-exome sequencing and being found to be significantly altered upon analysis of the combined cohort with MutSig (Broad Institute), an algorithm testing whether the observed mutations in a gene are not simply a consequence of random background mutation processes (**Supplementary Table 6**).

Methylation profiling. Comprehensive methylation profiling of 28 DIPG samples at the Microarray Centre (University Health Network, Toronto, Ontario, Canada) used the Illumina Infinium450k array. Bisulfite conversion was carried out using the EZ DNA Methylation kit (Zymo Research) according to the manufacturer's specifications. Methylation profiling of iNHA cells transfected with vectors expressing Lys27Met and wild-type histone H3.3 was also carried out on this array. DNA was extracted using the DNeasy kit (Qiagen) according to the manufacturer's protocols. Subgrouping and clustering on the basis of differential CpG probe methylation was performed and validated using multiple programs and algorithms, including non-negative matrix factorization (GenePattern, Broad Institute), consensus hierarchical clustering (GenePattern, Broad Institute; ConsensusClusterPlus, Bioconductor/R; MultiExperiment Viewer, Dana-Farber Cancer Institute), *k*-means clustering (GenePattern, Broad Institute; ConsensusClusterPlus, Bioconductor/R), silhouette clustering (ConsensusClusterPlus, Bioconductor/R) and significance analysis of microarrays (SAM; MultiExperiment Viewer, Dana-Farber Cancer Institute). Subgroups were validated for significance by SigClust (Bioconductor/R).

Gene expression profiling. Expression profiling was conducted on 35 DIPG samples and 10 normal brain controls at the Microarray Centre (University Health Network, Toronto, Ontario, Canada) using the DASL protocol for the Illumina HT-12 V4 BeadChip array. Differential gene expression in iNHA cells transfected with vector expressing Lys27Met or wild-type histone H3.3 was also assessed on this array. RNA was extracted using the QIAShredder and RNeasy kit (Qiagen) according to the manufacturer's specifications. RNA quality was assessed using the Bioanalyzer 2100 (Agilent Technologies). RNA Integrity Number (RIN) ranged from 1.3 to 8.2. Microarray data were normalized in Partek Genomics Suite v6.6 using per-probe median-centered quantile normalization. Data analysis was conducted on $\log_2$-transformed data or fold change data (relative to non-neoplastic brain). Genes differentially expressed among the subgroups were identified using one-way ANOVA, and significance was corrected for false discovery rate.

Copy number analysis. Copy number analysis was conducted for 40 DIPG samples using SNP6.0 (Affymetrix). Digestion, labeling and hybridization of DNA were performed by the Centre for Applied Genomics at The Hospital for Sick Children or at the Microarray Centre at the University Health Network. CEL data were analyzed for CNAs using the segmentation tool and HMM in Partek Genomics Suite v6.6 and Genotyping Console 4.1 (GTC4.1; Affymetrix) as previously described[9].

***In vitro* modeling of Lys27Met histone H3.3 and Gly328Val ACVR1.** iNHA cells[30] were grown in DMEM supplemented with 10% FBS (Invitrogen). Cells were tested and found to be negative for mycoplasma. Cells were passaged at 80% confluence. Clones expressing wild-type and mutant (Lys27Met) histone H3.3 (*H3F3A*) were created by site-direction mutagenesis and PCR using pCMV SPORT6 plasmid (Addgene) containing mouse *H3f3a* cDNA. Expression of Flag-tagged wild-type and Lys27Met H3F3A was established by cloning the corresponding cDNAs into p3xFLAG-CMV-10 vectors (Sigma-Aldrich) (see **Supplementary Table 8** for primer sequences). Clones were confirmed by bidirectional Sanger sequencing. Cells were transfected using FuGENE 6 transfection reagent (Promega) according to the manufacturer's instructions.

To generate stable transfectants, iNHA cells were cotransfected with a DNA fragment containing the hygromycin gene. Stably transfected clones were selected with 300 μg/ml hygromycin and were pooled to eliminate clonal variations in phenotype. For the isolation of DNA and RNA used for expression and methylation arrays, respectively, cells were collected at passages 4 and 15 at about 80% confluence. Total cell pellet for each sample was divided into two parts, one used for the isolation of genomic DNA and the other used for the isolation of RNA. Genomic DNA was isolated using the DNeasy Blood and Tissue kit (Qiagen), and bisulfite conversion was carried out using the EZ DNA Methylation kit. Total RNA was isolated by RNeasy kit. iNHA cells from passages 4 and 15 were analyzed on methylation (Infinium 450k) and expression (Illumina HT-12 v4) arrays by similar methods to those described for patient samples.

ACVR1 mutants were created from human *ACVR1* cDNA using the QuikChange Site-Directed Mutagenesis kit (Stratagene) (see **Supplementary Table 8** for primer sequences). The c.983G>T mutation was generated resulting in the p.Gly328Val amino acid change. The sequence encoding the Gly328Val ACVR1 mutant was then subcloned into the pCDH511b vector (System Biosciences) in frame with sequence for 3× Flag tag at the 3′ end of the cDNA with pCDH511b, pCDH511b-Gly328Val-ACVR1, pCDH511b-Lys27Met-H3.3 or pCDH511b-Gly328Val-ACVR1 and pCDH511b-Lys27Met-H3.3. Cells were harvested 72 h after transfection, and protein and RNA were extracted for protein blot and quantitative PCR analyses, respectively. Protein blotting was conducted using antibodies to phosphorylated SMAD1/5 (1:1,000 dilution; Cell Signaling Technology, 9516), total SMAD1 (1:1,000 dilution; Cell Signaling Technology, 6944), Flag (1:1,000 dilution; Sigma, F3165) and ACVR1 (1:1,000 dilution; Cell Signaling Technology, 4398). Reverse transcription was conducted using the RevertAid First-Strand cDNA Synthesis kit (Thermo Scientific). See **Supplementary Table 8** for RT-PCR primer pairs.

Cell counting and cell proliferation. Cell counting on iNHA cells expressing Lys27Met histone H3.3 and Gly328Val ACVR1 as well as control iNHA cells was performed in triplicate using the automated Vi-Cell Viability Analyzer from Beckman-Coulter. Cell proliferation was measured by BrdU incorporation. In brief, *Nes–tv-a*; $Tp53^{fl}$ mouse brainstem progenitors were plated in a 96-well flat-bottom plate (10 × 10^3 cells/well) in triplicate and incubated with virus expressing wild-type, Arg206His or Gly328Val ACVR1 for 72 h. Subsequently, cells were pulsed with BrdU and treated according to the manufacturer's protocol. Absorbance was measured using a Molecular Devices Versa Max Tunable Microplate Reader.

Detection of alternative lengthening of telomeres. The ALT phenotype was determined by telomere restriction fragment (TRF) assay or the presence of C-circles. TRF assays were performed using the TeloTAGGG Telomere Length Assay kit (Roche). Briefly, 1.5 µg of sample DNA was used per lane. The presence of long telomere length (<3 to >50 kb) was considered indicative of the ALT phenotype. Positive and negative controls included with the kit were run in tandem with sample DNA. C-circle assays were performed as previously described[14]. Briefly, amplification of C-circles was conducted on 16 ng of sample DNA, which was incubated for 8 h at 30 °C with master mix containing 5 units of ϕ29 polymerase. Quantitative PCR (qPCR) was employed to detect the presence of C-circles. ϕ29 polymerase–amplified product and non-amplified DNA (2 ng) were run in triplicate on a LightCycler 480 (Roche). qPCR was conducted using the following conditions: 95 °C for 15 min followed by 35 cycles of 95 °C for 15 s and 54 °C for 2 min. Samples with a delta mean C_t value greater than 0.2 (when comparing mean triplicate C_t values of ϕ29 polymerase–amplified and non-amplified DNA) were considered to be positive for C-circles. GMA47 ALT-positive fibroblasts were used as a positive control, and HeLa cervical cancer cells were used as an ALT-negative control.

Immunohistochemistry. Available surgical and autopsy material was immunohistochemically stained with selected antibodies against H3K27me3 (1:50 dilution; Millipore, 07-449), H3K27ac (1:1,000 dilution; Abcam, ab4729), H3K4me3 (1:1,000 dilution; Cell Signaling Technology, 9727) and H3K9ac (1:1,000 dilution; Abcam, ab4441). Sections (5 µm) of DIPG tissue microarrays (TMAs) constructed in our laboratory or tissue sections were cut from paraffin blocks and mounted on positively charged slides. Sections were baked overnight at 60 °C. Wax was removed by several xylene washes followed by tissue hydration by immersion in decreasing concentrations of ethanol in distilled water. Tissue sections were heat treated in citrate buffer for the purpose of antigen retrieval. Immunodetection was performed with the automated Benchmark XT stainer (Ventana) using the Ultraview Universal DAB Detection kit (Ventana). Slides were counterstained with the Hematoxylin II kit (Ventana).

Statistical analysis. Statistical analysis was performed on GraphPad Prism 5 software or SPSS v21 (IBM). OR values were calculated as previously shown[31]. Two-group comparisons were analyzed by two-sided Fisher's exact test. Continuous-scale data were calculated by unpaired two-tailed Student's *t* test. *P* values ≤0.05 were considered significant. Cox proportional hazards model and significance testing ($\alpha = 0.05$) based on the Wald test was used for multivariate analysis.

25. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
26. Jiang, Y., Wang, Y. & Brudno, M. PRISM: pair-read informed split-read mapping for base-pair level detection of insertion, deletion and structural variants. *Bioinformatics* **28**, 2576–2583 (2012).
27. Mills, R.E. *et al.* Natural genetic variation caused by small insertions and deletions in the human genome. *Genome Res.* **21**, 830–839 (2011).
28. Fiume, M. *et al.* Savant Genome Browser 2: visualization and analysis for population-scale genomics. *Nucleic Acids Res.* **40**, W615–W621 (2012).
29. Medvedev, P., Fiume, M., Dzamba, M., Smith, T. & Brudno, M. Detecting copy number variation with mated short reads. *Genome Res.* **20**, 1613–1622 (2010).
30. Sasai, K. *et al.* O6-methylguanine-DNA methyltransferase is downregulated in transformed astrocyte cells: implications for anti-glioma therapies. *Mol. Cancer* **6**, 36 (2007).
31. Bland, J.M. & Altman, D.G. Statistics notes. The odds ratio. *Br. Med. J.* **320**, 1468 (2000).

doi:10.1038/ng.2936