nature genetics

Recurrent somatic mutations in *ACVR1* in pediatric midline high-grade astrocytoma

Adam M Fontebasso[1,33], Simon Papillon-Cavanagh[2,33], Jeremy Schwartzentruber[3,33], Hamid Nikbakht[2], Noha Gerges[2], Pierre-Olivier Fiset[4], Denise Bechet[2], Damien Faury[2,5], Nicolas De Jay[2], Lori A Ramkissoon[6], Aoife Corcoran[6], David T W Jones[7], Dominik Sturm[7], Pascal Johann[7], Tadanori Tomita[8], Stewart Goldman[9], Mahmoud Nagib[10], Anne Bendel[11], Liliana Goumnerova[12,13], Daniel C Bowers[14], Jeffrey R Leonard[15], Joshua B Rubin[16], Tord Alden[8], Samuel Browd[17], J Russell Geyer[18], Sarah Leary[18], George Jallo[19], Kenneth Cohen[20], Nalin Gupta[21], Michael D Prados[21], Anne-Sophie Carret[22], Benjamin Ellezam[23], Louis Crevier[24], Almos Klekner[25], Laszlo Bognar[25], Peter Hauser[26], Miklos Garami[26], John Myseros[27], Zhifeng Dong[28], Peter M Siegel[28], Hayley Malkin[29], Azra H Ligon[6,13,30], Steffen Albrecht[4], Stefan M Pfister[7], Keith L Ligon[6,13,30,31,34], Jacek Majewski[2,34], Nada Jabado[1,2,5,34] & Mark W Kieran[13,29,32,34]

Pediatric midline high-grade astrocytomas (mHGAs) are incurable with few treatment targets identified. Most tumors harbor mutations encoding p.Lys27Met in histone H3 variants. In 40 treatment-naive mHGAs, 39 analyzed by whole-exome sequencing, we find additional somatic mutations specific to tumor location. Gain-of-function mutations in *ACVR1* occur in tumors of the pons in conjunction with histone H3.1 p.Lys27Met substitution, whereas *FGFR1* mutations or fusions occur in thalamic tumors associated with histone H3.3 p.Lys27Met substitution. Hyperactivation of the bone morphogenetic protein (BMP)-ACVR1 developmental pathway in mHGAs harboring *ACVR1* mutations led to increased levels of phosphorylated SMAD1, SMAD5 and SMAD8 and upregulation of BMP downstream early-response genes in tumor cells. Global DNA methylation profiles were significantly associated with the p.Lys27Met alteration, regardless of the mutant histone H3 variant and irrespective of tumor location, supporting the role of this substitution in driving the epigenetic phenotype. This work considerably expands the number of potential treatment targets and further justifies pretreatment biopsy in pediatric mHGA as a means to orient therapeutic efforts in this disease.

We and others recently identified recurrent mutations in *H3F3A*, encoding histone 3 variant 3 (H3.3), in 38% of pediatric supratentorial HGAs and mutations affecting histone H3.3 or H3.1 in ~80% of brainstem HGAs (diffuse intrinsic pontine gliomas, DIPGs)[1–3]. Mutations affecting the histone H3 tail that change the glycine at position 34 to either arginine or valine (p.Gly34Arg or p.Gly34Val), as well as mutations in genes affecting histone H3 post-translational modifications at lysine 36, predominate in cortical tumors, whereas lysine-to-methionine substitutions at residue 27 (p.Lys27Met) occur in midline tumors[1,2,4,5].

Herein we focus on the genomic (mutational spectrum and copy number alterations) and epigenetic (DNA methylation) landscape of treatment-naive pediatric midline (thalamus, cerebellum, spine and pons; non-cortical regions) high-grade astrocytomas (**Figs. 1–3**). These tumors are often surgically challenging or inoperable, and published studies have mainly used material collected after therapy and have provided limited genomic data other than structural alterations and mutational analysis of histone H3 variant genes and *TP53* (refs. 1,3,6,7). We analyzed 40 mHGAs, with whole-exome sequencing data for 39 tumors, including 25 biopsies from DIPG cases (11 previously reported[2,8]; Online Methods and **Supplementary Table 1**). The recurrent histone H3 gene mutation encoding p.Lys27Met was found in 37 of 40 cases (93%) and was distributed among three histone variant genes, including *H3F3A* (32/40 samples), *HIST1H3B* (4/39 samples) and *HIST1H3C* (1/39 samples). *HIST1H3C* also encodes the canonical histone H3.1 (as does *HIST1H3B*) and has not previously been reported to be mutated (**Fig. 1**). All five histone H3.1 alterations occurred in the pons of younger cases (**Supplementary Fig. 1**). Histone H3.3 p.Lys27Met alterations occurred in multiple midline locations, including the brainstem (19/25), thalamus (9/11) and rare locations for HGAs—the cerebellum, fourth ventricle and spinal cord (4/4). In contrast, only 1 histone H3.3 p.Lys27Met mutant was identified in 42 pediatric high-grade tumors located in cerebral hemispheres in a previously reported data set[2,5] (**Fig. 1** and **Supplementary Table 1**).

The pattern of somatic mutations in specific genes or gene pathways showed striking features (**Fig. 1** and **Supplementary Fig. 1**). Recurrent somatic mutations in the activin A receptor, type I gene (*ACVR1*) occurred in 5 of 39 mHGAs and overlapped with p.Lys27Met alterations

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

Received 3 October 2013; accepted 14 March 2014; published online 6 April 2014; doi:10.1038/ng.2950

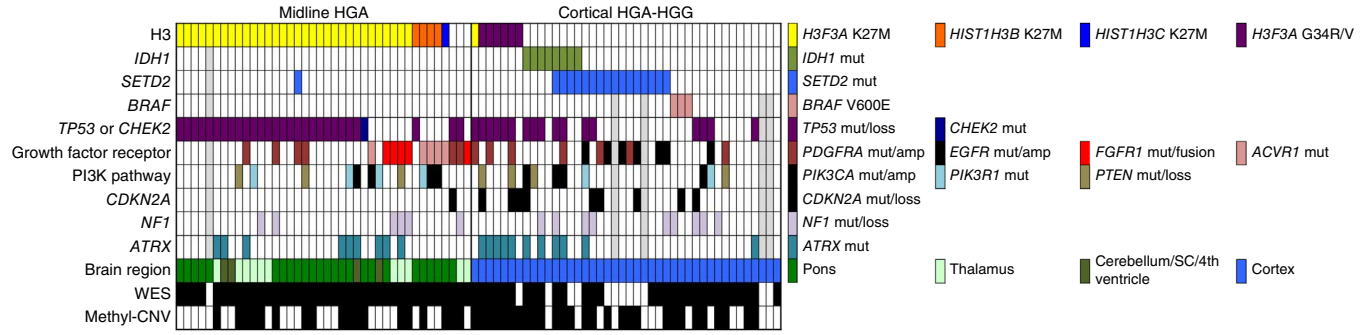

Figure 1 Genomic landscape of pediatric midline HGAs. Distribution of mutations and alterations in 40 pediatric midline high-grade astrocytomas (midline HGA) and 42 cortical high-grade astrocytomas and high-grade gliomas (cortical HGA-HGG) described in the study. Mutations (mut) were identified with whole-exome sequencing (WES) where available and are indicated by colored boxes. Amplifications (amp) and losses were identified using DNA methylation profile–derived copy number variant (methyl-CNV) analysis and are indicated where available by colored boxes. SC, spinal cord. Boxes in light gray indicate samples for which data are not available. Detailed information on tumor samples included herein can be found in **Supplementary Table 1**, with specific variants and transcript accessions presented in **Supplementary Table 6**.

(5/5), mainly in histone H3.1 (4/5) rather than histone H3.3 or wild-type histone H3 ($P = 0.0003$, Fisher's exact test; **Supplementary Fig. 1**). These specific amino acid residues in ACVR1 have previously been reported to be affected by germline mutations causing fibrodysplasia ossificans progressiva (FOP), an inherited musculoskeletal disease[9–14]. The p.Arg206His, p.Gly328Glu and p.Gly356Asp substitutions result in ligand-independent activation of the kinase, leading to an increase in BMP signaling and increased phosphorylation of SMAD1, SMAD5 and SMAD8 (SMAD1/5/8) in tissues[10], whereas the newly identified p.Gly328Val and p.Arg258Gly substitutions are predicted to exert gain-of-function effects similar to those for the previously described p.Arg258Ser, p.Gly328Glu, p.Gly328Trp and p.Gly328Arg alterations, on the basis of the physicochemical properties of the mutant residues and predicted protein structure (**Fig. 4** and **Supplementary Table 2**)[11,12]. ACVR1, also known as ALK2, is a type I receptor of the mammalian transforming growth factor (TGF)-β signaling family with critical developmental roles in the mouse embryo[13] and in early left-right patterning[14]. Investigation of ACVR1 pathway activation using immunohistochemical staining for phosphorylated SMAD1/5/8 in DIPG samples showed positive staining only in tumors with mutant *ACVR1* (4/4), including in DIPGs with the newly identified p.Gly328Val and p.Arg258Gly alterations (**Fig. 2**).

In addition, in the primary cell line DIPGIV, which carries histone H3.1 p.Lys27Met and the new ACVR1 substitution p.Gly328Val, we demonstrated increased levels of phosphorylated SMAD1/5/8 compared to normal human astrocytes (NHA cells) and significantly increased expression of genes containing BMP response elements—*ID1*, *ID2* and *ID3*—and *SNAI1* compared to KNS42, a cell line with the histone H3.3 p.Gly34Val alteration and wild-type ACVR1 (**Fig. 4**). These genes are early-response genes induced following active BMP2 signaling and represent SMAD1/5/8 downstream effectors[15,16].

We identified mutations in *FGFR1* in association with histone H3.3 p.Lys27Met alteration in 4 of 39 cases: 3 thalamic HGAs also having *NF1* mutations (previously reported in ref. 8) and 1 DIPG (**Fig. 1**). These *FGFR1* mutations affected hotspot residues in the tyrosine kinase domain of the FGFR1 receptor. They have been shown to lead to its constitutive activation in a subset of thalamic pilocytic astrocytomas, a grade I tumor that rarely progresses to higher-grade astrocytoma, and in NIH 3T3 cells[8]. Comparison with our data set of non-midline cortical HGA-HGG indicated that tumors carrying *ACVR1* or *FGFR1* mutations were exclusive to midline HGAs ($P = 0.0040$, Fisher's exact test; **Fig. 1** and **Supplementary Fig. 1b**).

TP53 mutations occurred in 28 of 40 samples, largely in combination with histone H3.3 p.Lys27Met alteration (25/28) as reported previously[1,2] and, to a much lesser extent, with histone H3.1 p.Lys27Met alteration (1/5) (**Fig. 1** and **Supplementary Fig. 1b**). One DIPG sample (mHGA1) had histone H3.3 p.Lys27Met alteration as well as a splicing mutation in *CHEK2*, a gene associated with Li-Fraumeni syndrome similar to *TP53*

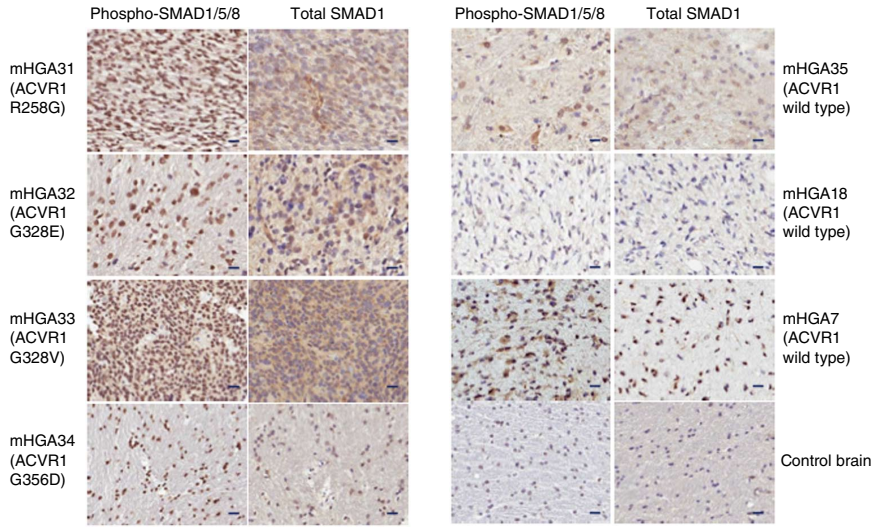

Figure 2 Increased levels of phosphorylated SMAD1/5/8 in *ACVR1*-mutant mHGAs. Immunohistochemical analysis of mHGAs harboring *ACVR1* mutations identified in this study ($n = 4$; left) demonstrate increased nuclear positivity of phosphorylated SMAD1/5/8 compared to mHGAs with wild-type *ACVR1* ($n = 3$) and a control brain sample (right), with total-SMAD1 staining shown in each case as a positive control. Scale bars, 20 μm (20× magnification). Clinicopathological and molecular characteristics of tumor samples are presented in **Supplementary Table 1**.

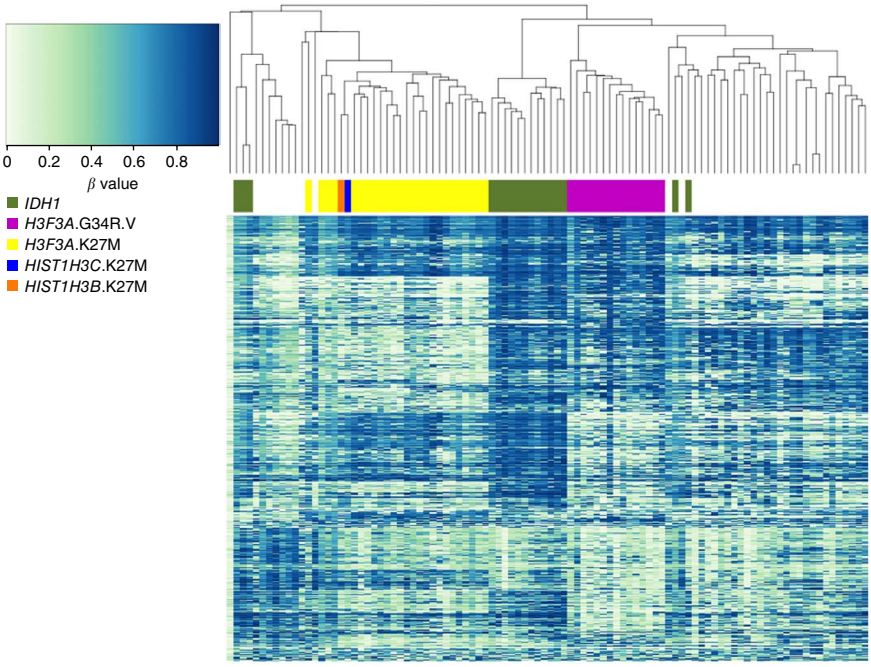

Figure 3 Clustering analysis of global DNA methylation profiles for 98 high-grade astrocytomas. Global DNA methylation clustering analysis of high-grade astrocytomas distributed across the brain demonstrates similar impact on epigenomic dysregulation caused by p.Lys27Met alteration regardless of age, brain location, associated mutations or the particular histone H3 variant affected. The top 10,000 most variable normalized methylation β values were used for UPGMA clustering with alterations of interest indicated. Robustness was assessed using multiscale bootstrapping (**Supplementary Fig. 5**). Detailed information regarding sample clinicopathological characteristics is included in **Supplementary Table 1**, with methylation-derived CNVs presented in **Supplementary Tables 4** and **5**.

(**Fig. 1** and **Supplementary Tables 1** and **6**). Three midline samples displayed no detectable mutation across *H3F3A*, *HIST1H3B* and *HIST1H3C*, even when sequenced at extremely high depth (average read depth across all tumors and genes of >20,000×; **Supplementary Table 3**). These samples carried alterations described in adult HGA in the form of combined *TP53* mutation, *PDGFRA* amplification and *CDKN2A* loss (mHGA35), *PDGFRA* and *EGFR* amplification and *NF1* loss (mHGA36), or the previously identified *FGFR1-TACC1* fusion[17,18] (mHGA37) (**Fig. 1** and **Supplementary Tables 1**, **3** and **4**).

We identified somatic *PDGFRA* mutations in a small subset of pediatric mHGAs (5/39; **Supplementary Table 1**). Notably, *PDGFRA* amplification occurred in 3 of 12 treatment-naive DIPGs (**Supplementary Fig. 2** and **Supplementary Table 1**). Interestingly, a cerebellar HGA sample with multiple biopsies taken from different anatomical loci showed *PDGFRA* amplification in only one tumor site, whereas similar somatic mutation and DNA methylation profiles were observed across all tumor sites (**Supplementary Fig. 3**). This finding supports the view that *PDGFRA* amplifications precede therapy but can be further promoted by it, as described in radiation-induced supratentorial HGA[19]. Mutations in other growth factor receptor and component genes described to have focal gains or losses in DIPG[6,7] or HGA (*MET*, *RB1* and *PARP1*) showed low incidence in our data set (**Fig. 1** and **Supplementary Tables 1** and **4**–**6**), whereas recurrent mutations in components of the phosphoinositide 3-kinase (PI3K) pathway, predicted to activate AKT signaling, were present in 10 of 39 mHGAs. With respect to alterations affecting telomere length, *ATRX* mutations were identified in nine midline tumors, exclusively in samples with histone H3.3 p.Lys27Met alteration and *TP53* mutation, and affected older children[1], whereas no *TERT* promoter mutations, assessed using targeted sequencing, were identified in mHGAs (**Supplementary Table 1**).

Lysine-to-methionine substitutions in histone variants at residue Lys27 have recently been shown to inhibit SET domain–containing histone methyltransferases[20], possibly accounting for the specific DNA methylation pattern we observed in non-brainstem tumors with the histone H3.3 p.Lys27Met alteration[4]. When clustering our samples by global DNA methylation, all samples with a given histone H3 alteration (absent or affecting Lys27 or Gly34) clustered together but did not group on the basis of tumor location within the brain, the particular histone H3 gene mutated, and additional partner mutations or structural alterations identified (**Fig. 3** and **Supplementary Figs. 4** and **5**). The global epigenetic profile is thus strongly associated with alteration of the histone H3 variant mark. Interestingly, *TP53* alteration was associated with increased broad copy number changes as we previously showed[2] (average of 32 events per sample), indicating a level of genomic instability. The number of copy number changes was significantly higher in samples with the histone H3.3 p.Gly34Arg or p.Gly34Val alteration than in other groups ($P = 0.02997$; **Supplementary Table 7**), in keeping with global hypomethylation identified in these tumors[4].

The gain-of-function alterations in three growth factor receptor genes—*ACVR1*, *FGFR1* and *PDGFRA*—associate with histone H3 p.Lys27Met variants in midline HGA. These mutations are not seen concurrently, and *ACVR1* and *FGFR1* mutations are mutually exclusive with *TP53* alterations and privilege specific locations within the midline of the brain. *ACVR1* mutations do not seem to correlate with differential survival in patients with HGA, although no definite conclusions can be drawn because of our limited sample size (**Supplementary Fig. 6**). The lack of reported central nervous system (CNS) tumor development in humans with FOP or *Acvr1* (*Alk2*) mouse models suggests that aberrant activation of this pathway is not sufficient for tumorigenesis[10] and that it may act in concert with histone H3 p.Lys27Met substitution and other alterations we identified in the PI3K pathway to induce tumorigenesis (**Fig. 1**). Interestingly, *ACVR1* is mainly expressed at embryonic day (E) 14 in the cortex of mouse embryos and at very low levels in the brainstem, and aberrantly active ACVR1 leads to increased ventralization of zebrafish embryos[11,12,21]. Aberrant ACVR1 signaling in the brainstem, a midline structure, may specify a patterning defect in DIPG, as this gene is involved in left-right patterning development. FGFR signaling regulates neural progenitor maintenance and the development of the ventral midbrain[22]. Similar to *ACVR1* mutations, the *FGFR1* gain-of-function mutations we identify only lead to grade I astrocytomas if present on their own in a tumor[8] and are commonly found in association with *NF1* and histone H3.3 p.Lys27Met alterations in midline HGA. Mutations affecting histone H3 variants in pediatric HGAs mirror *IDH* gene mutations in their requirement for additional alterations to potentially induce HGA. Mutations encoding p.Gly34Arg and

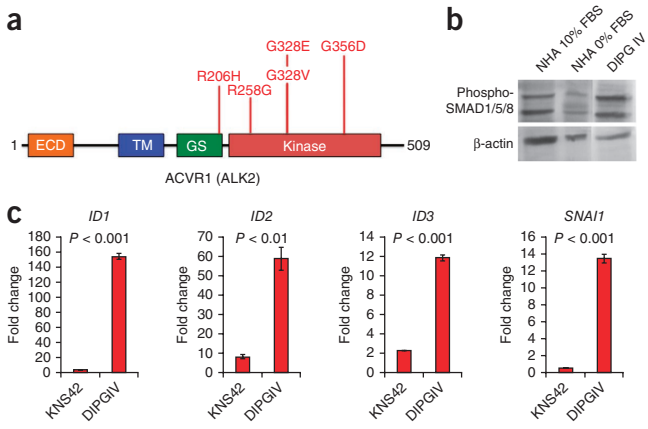

Figure 4 Mutations identified in *ACVR1* are associated with activation of downstream SMAD signaling pathways. (**a**) Distribution of alterations identified in ACVR1 ($n = 5$) demonstrating the impact of amino acid substitutions on the kinase domain of the protein. ECD, extracellular domain; TM, transmembrane domain; GS, glycine-serine-rich domain. (**b**) Immunoblotting analysis of phosphorylated SMAD1/5/8 levels in NHA cells with wild-type *ACVR1* grown in 10% FBS or in NHA and DIPGIV cells (*ACVR1* mutation encoding p.Gly328Val) starved for 1 h in medium free of serum and growth factors (0%). (**c**) Quantitative PCR (qPCR) analysis of the expression of the downstream BMP effectors *ID1*, *ID2*, *ID3* and *SNAI1* in *ACVR1*-mutant DIPGIV cells and in the glioblastoma (GBM) cell line KNS42 with wild-type *ACVR1*. Values represented are fold changes calculated using the $2^{-\Delta\Delta C_t}$ method, normalized to *ACTB* expression in calibrator NHA cells. *P* values were calculated using two-tailed *t* tests for significance, with error bars representing s.d. from two technical replicates.

p.Gly34Val have been found in *H3F3A* and not in other histone genes thus far and invariably associate with *TP53* and *ATRX* mutations in the cortex. In mHGA, mutations encoding p.Lys27Met arise in different histone H3 variants depending on age and tumor location. They associate with *TP53* mutations or with activated neurodevelopmental growth factor receptor pathways through distinct hits in *ACVR1*, *FGFR1* or *PDGFRA* to achieve tumorigenesis. These alterations in growth factor receptors and in members of the PI3K pathway offer previously unforeseen therapeutic possibilities in a deadly cancer, while the observed level of genomic instability calls for caution in the choice of adjuvant therapy whenever possible. Notably, we show that small-needle pretherapy biopsies can reliably identify the mutational landscape in HGA. This technique will allow for the tailoring of available therapies to the results obtained from stereotactic biopsy in children affected by this fatal brain tumor while effort continues to be made to target p.Lys27Met histone alterations.

URLs. FASTX-Toolkit, http://hannonlab.cshl.edu/fastx_toolkit/; Genome Analysis Toolkit (GATK), http://www.broadinstitute.org/gsa/wiki/; Picard, http://picard.sourceforge.net/; SAMtools, http://samtools.sourceforge.net/; BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi; Ensembl, http://useast.ensembl.org/index.html; UniProt, http://www.uniprot.org/; UCSC Genome Browser, http://genome.ucsc.edu/; dbSNP, http://www.ncbi.nlm.nih.gov/SNP/.

METHODS

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

Accession codes. Whole-exome sequencing data can be accessed through the European Genome-phenome Archive (EGA) under accession EGAS00001000720. DNA methylation data can be accessed through the Gene Expression Omnibus (GEO) under accession GSE55712.

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

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude toward all staff at the McGill University and Génome Québec Innovation Centre for excellent technical expertise, library preparation and sequencing. The authors are very grateful to J.-J. Lebrun (McGill University) for primer sequences and materials for SMAD signaling studies. This work was performed within the context of the I-CHANGE (International Childhood Astrocytoma Integrated Genomics and Epigenomics) Consortium and was supported by funding from Genome Canada, Génome Québec, the Institute for Cancer Research of the Canadian Institutes for Health Research (CIHR), McGill University and the Montreal Children's Hospital Foundation. This work was also supported by Hungarian Scientific Research Fund (OTKA) contract T-04639, National Research and Development Fund (NKFP) contract 1A/002/2004 (P.H. and M.G.) and TÁMOP-4.2.2A-11/1/KONV-2012-0025 (A.K. and L.B.). N.J. is a member of the Penny Cole laboratory and the recipient of a Chercheur Clinicien Senior Award. J. Majewski holds a Canada Research Chair (tier 2). L.G., K.L.L. and M.W.K. are supported by NCI P01CA142536. We acknowledge the support of the Zach Carson DIPG Fund at the Dana-Farber Cancer Institute (DFCI), the Ellie Kavalieros Fund (DFCI), the Mikey Czech Foundation, the Prayer From Maria Foundation, the Hope for Caroline Fund (DFCI), the Ryan Harvey DIPG Fund (DFCI), the Stop&Shop Pediatric Brain Tumor Program (DFCI) and the Pediatric Brain Tumor Clinical and Research Fund (DFCI). A.M.F. is supported by a studentship from CIHR, as well as by an award from the CIHR Systems Biology Training Program at McGill University. D.B. is supported by a studentship from the T.D. Trust/Montreal Children's Hospital Foundation, and N. Gerges is supported by a studentship from the Cedars Cancer Institute. N.D.J. is supported by an award from the McGill Integrated Cancer Research Training Program.

AUTHOR CONTRIBUTIONS

A.M.F., N. Gerges, P.-O.F., D.B., D.F., L.A.R., A.C., A.H.L., S.A. and Z.D. performed experiments. A.M.F., S.P.-C., J.S., H.N., N.D.J., A.H.L., S.A., Z.D. and P.M.S. analyzed the data and produced figures and tables. D.T.W.J., D.S., P.J., T.T., S.G., M.N., A.B., L.G., D.C.B., J.R.L., J.B.R., T.A., S.B., J.R.G., G.J., K.C., N. Gupta, M.D.P., A.-S.C., B.E., L.C., A.K., L.B., P.H., M.G., J. Myseros, H.M., S.A. and S.M.P. provided tissue samples. K.L.L., J. Majewski, N.J. and M.W.K. provided project leadership and designed the study. All authors contributed to the final manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. 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).
2. Schwartzentruber, J. *et al.* Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. *Nature* **482**, 226–231 (2012).
3. Wu, G. *et al.* Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. *Nat. Genet.* **44**, 251–253 (2012).
4. Sturm, D. *et al.* Hotspot mutations in *H3F3A* and *IDH1* define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* **22**, 425–437 (2012).
5. Fontebasso, A.M. *et al.* Mutations in *SETD2* and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas. *Acta Neuropathol.* **125**, 659–669 (2013).
6. 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).
7. 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).
8. Jones, D.T. *et al.* Recurrent somatic alterations of *FGFR1* and *NTRK2* in pilocytic astrocytoma. *Nat. Genet.* **45**, 927–932 (2013).
9. Shore, E.M. *et al.* A recurrent mutation in the BMP type I receptor *ACVR1* causes inherited and sporadic fibrodysplasia ossificans progressiva. *Nat. Genet.* **38**, 525–527 (2006).
10. Fukuda, T. *et al.* Generation of a mouse with conditionally activated signaling through the BMP receptor, ALK2. *Genesis* **44**, 159–167 (2006).

11. Chaikuad, A. *et al.* Structure of the bone morphogenetic protein receptor ALK2 and implications for fibrodysplasia ossificans progressiva. *J. Biol. Chem.* **287**, 36990–36998 (2012).
12. Fukuda, T. *et al.* A unique mutation of ALK2, G356D, found in a patient with fibrodysplasia ossificans progressiva is a moderately activated BMP type I receptor. *Biochem. Biophys. Res. Commun.* **377**, 905–909 (2008).
13. Mishina, Y., Crombie, R., Bradley, A. & Behringer, R.R. Multiple roles for activin-like kinase-2 signaling during mouse embryogenesis. *Dev. Biol.* **213**, 314–326 (1999).
14. Kishigami, S. *et al.* BMP signaling through ACVRI is required for left-right patterning in the early mouse embryo. *Dev. Biol.* **276**, 185–193 (2004).
15. Miyazono, K., Maeda, S. & Imamura, T. BMP receptor signaling: transcriptional targets, regulation of signals, and signaling cross-talk. *Cytokine Growth Factor Rev.* **16**, 251–263 (2005).
16. Korchynskyi, O. & ten Dijke, P. Identification and functional characterization of distinct critically important bone morphogenetic protein–specific response elements in the *Id1* promoter. *J. Biol. Chem.* **277**, 4883–4891 (2002).
17. Singh, D. *et al.* Transforming fusions of *FGFR* and *TACC* genes in human glioblastoma. *Science* **337**, 1231–1235 (2012).
18. Zhang, J. *et al.* Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. *Nat. Genet.* **45**, 602–612 (2013).
19. 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).
20. Lewis, P.W. *et al.* Inhibition of PRC2 activity by a gain-of-function H3 H3 mutation found in pediatric glioblastoma. *Science* **340**, 857–861 (2013).
21. Shen, Q. *et al.* The fibrodysplasia ossificans progressiva R206H ACVR1 mutation activates BMP-independent chondrogenesis and zebrafish embryo ventralization. *J. Clin. Invest.* **119**, 3462–3472 (2009).
22. Lahti, L., Peltopuro, P., Piepponen, T.P. & Partanen, J. Cell-autonomous FGF signaling regulates anteroposterior patterning and neuronal differentiation in the mesodiencephalic dopaminergic progenitor domain. *Development* **139**, 894–905 (2012).

[1]Division of Experimental Medicine, Montreal Children's Hospital, McGill University and McGill University Health Centre, Montreal, Quebec, Canada. [2]Department of Human Genetics, McGill University, Montreal, Quebec, Canada. [3]Wellcome Trust Sanger Institute, Hinxton, UK. [4]Department of Pathology, Montreal Children's Hospital, McGill University Health Centre, Montreal, Quebec, Canada. [5]Department of Pediatrics, McGill University, Montreal, Quebec, Canada. [6]Department of Medical Oncology, Center for Molecular Oncologic Pathology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [7]Division of Pediatric Neuro-oncology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [8]Department of Neurological Surgery, Ann & Robert H. Lurie Children's Hospital of Chicago, Northwestern University, Chicago, Illinois, USA. [9]Department of Pediatrics Hematology-Oncology, Ann & Robert H. Lurie Children's Hospital of Chicago, Northwestern University, Chicago, Illinois, USA. [10]Clinical Pediatric Neurosurgical Oncology, Department of Surgery, Boston Children's Hospital, Boston, Massachusetts, USA. [11]Department of Pediatric Hematology-Oncology, Children's Hospitals and Clinics of Minnesota, Minneapolis, Minnesota, USA. [12]Department of Neurosurgery, Boston Children's Hospital, Boston, Massachusetts, USA. [13]Harvard Medical School, Boston, Massachusetts, USA. [14]Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, Texas, USA. [15]Department of Neurological Surgery, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA. [16]Department of Pediatrics, Hematology-Oncology, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA. [17]Department of Neurosurgery, Seattle Children's Hospital, Seattle, Washington, USA. [18]Cancer and Blood Disorders Center, Seattle Children's Hospital, Seattle, Washington, USA. [19]Department of Neurological Surgery, Johns Hopkins Hospital, Baltimore, Maryland, USA. [20]Department of Pediatric Hematology-Oncology, The Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, Baltimore, Maryland, USA. [21]Department of Neurological Surgery, University of California, San Francisco, San Francisco, California, USA. [22]Department of Hematology-Oncology, Centre Hospitalier Universitaire (CHU) Sainte-Justine, Université de Montréal, Montreal, Quebec, Canada. [23]Department of Pathology, Centre Hospitalier Universitaire (CHU) Sainte-Justine, Université de Montréal, Montreal, Quebec, Canada. [24]Department of Neurosurgery, Centre Hospitalier Universitaire (CHU) Sainte-Justine, Université de Montréal, Montreal, Quebec, Canada. [25]Department of Neurosurgery, Medical and Health Science Center, University of Debrecen, Debrecen, Hungary. [26]2nd Department of Paediatrics, Semmelweis University, Budapest, Hungary. [27]Department of Neurosurgery, Children's National Medical Center, Washington, DC, USA. [28]Rosalind and Morris Goodman Cancer Research Centre, McGill University, Montreal, Quebec, Canada. [29]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [30]Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts, USA. [31]Department of Pathology, Boston Children's Hospital, Boston, Massachusetts, USA. [32]Division of Pediatric Hematology/Oncology, Boston Children's Hospital, Boston, Massachusetts, USA. [33]These authors contributed equally to this work. [34]These authors jointly directed this work. Correspondence should be addressed to K.L.L. (keith_ligon@dfci.harvard.edu), J. Majewski (jacek.majewski@mcgill.ca), N.J. (nada.jabado@mcgill.ca) or M.W.K. (mark_kieran@dfci.harvard.edu).

ONLINE METHODS

Patient samples and consent. All samples were obtained with informed consent after approval of the institutional review boards (IRBs) of the respective hospitals they were treated in and were independently reviewed by senior pediatric neuropathologists (S.A. and K.L.L.) according to World Health Organization (WHO) guidelines. Samples were obtained from the Montreal Children's Hospital (Montreal, McGill University Health Centre), Boston Children's Hospital (Boston, Harvard University), the Brain Tumor Toronto Bank (BTTB; Toronto, University Health Network) and from collaborators in Hungary, in addition to previously published mHGA samples[2,5] with sequencing data (n = 11) and with previously published DNA methylation data (n = 89) for a total of 98 tumors included herein for DNA methylation and copy number variant (CNV) analysis[4,5]. DIPG biopsy samples obtained before therapy were from Dana-Farber Cancer Institute protocol 10-321 (n = 12), a prospective phase II biopsy study of newly diagnosed DIPG. The protocol is IRB approved through the Dana-Farber Harvard Cancer Center IRB, FDA IND 111,882, http://ClinicalTrials.gov/ identifier NCT01182350, and has local institutional approval at all participating sites; informed consent was obtained from all parents. Additional pediatric midline HGA samples were from needle biopsies or partial resections before treatment (n = 28). Sequencing and clinical data for this cohort are presented in **Supplementary Table 1**, and methylation-derived CNVs in genes of interest are presented in **Supplementary Tables 4** and **5**.

Whole-exome DNA sequencing. Standard genomic DNA extraction methods were performed according to described company protocols (Qiagen). Paired-end library preparations were carried out using the Nextera Rapid Capture Exome kit according to instructions from the manufacturer (Illumina) from 50 ng of total starting genomic DNA. Sequencing was performed in rapid-run mode with 100-bp paired-end reads on an Illumina HiSeq 2000. We removed adaptor sequences, quality trimmed reads using the FASTX-Toolkit and then used a custom script to ensure that only read pairs with both mates present were subsequently used. Reads were aligned to hg19 with Burrows-Wheeler Aligner (BWA) 0.5.9 (ref. 23), and indel realignment was performed using the Genome Analysis Toolkit (GATK)[24]. Duplicate reads were then marked using Picard and excluded from downstream analyses. We assessed coverage of consensus coding sequence (CCDS) bases using GATK, which showed that the majority of samples had >92% of CCDS bases covered by at least 10 reads and >88% of CCDS bases covered by at least 20 reads.

For each sample, SNVs and short indels were called using SAMtools mpileup[25] with the extended base alignment quality (BAQ) adjustment (–E) and were then quality filtered to require at least 20% of reads supporting each variant call. Variants were annotated using both ANNOVAR[26] and custom scripts to identify whether they affected protein-coding sequence and whether they had previously been seen in the 1000 Genomes Project data set (November 2011), the National Heart, Lung, and Blood Institute (NHLBI) Grand Opportunity (GO) exomes or in approximately 1,000 exomes previously sequenced at our center. Variants in candidate genes of interest described herein in the midline HGA cohort (n = 39) are detailed in **Supplementary Table 6**.

MiSeq targeted high-depth DNA sequencing of *H3F3A*, *HIST1H3B* and *HIST1H3C*. Genomic DNA from midline HGA samples was used for high-depth sequencing of the *H3F3A*, *HIST1H3B* and *HIST1H3C* genes to investigate the frequency of reads encoding p.Lys27Met in samples, notably those wild type by whole-exome sequencing and high-resolution melting assays. Midline HGAs (n = 24) were sequenced using the MiSeq sequencing platform (Illumina) with an average coverage of >20,000× of the analogous p.Lys27Met base change across the three histone variants (more specifically, for the *H3F3A* gene, the average coverage was >12,000×, for the *HIST1H3B* gene, the coverage was >12,000×, and, for the *HIST1H3C* gene, the coverage was >35,000×). Reads were mapped to the reference genome (human hg19) using the BWA genome aligner[23]. Alignment files were fed to the mpileup tool from the SAMtools package[25] to find all the variations without any filter applied by the conventional variant callers. An in-house parser program was developed to extract different variations at the desired positions (in this case, affecting Lys27) in the mapped paired reads covering the histone genes *H3F3A*, *HIST1H3B* and *HIST1H3C*. The results are provided in **Supplementary Table 3**.

RNA sequencing. RNA was extracted from case tumor mHGA37 using the Qiagen RNeasy Lipid Tissue Mini kit according to instructions from the manufacturer. Library preparation was performed with rRNA depletion methods according to instruction from the manufacturer (Epicentre) to achieve greater coverage of mRNA and other long noncoding transcripts. Paired-end sequencing was performed on the Illumina HiSeq 2000 platform.

RNA sequencing fusion analysis. RNA sequencing FASTQ files were used for fusion analysis with the deFuse software algorithm[27] according to indicated settings. Algorithmic output was then analyzed for high-confidence fusion transcripts, which were then reconstructed *in silico* using online bioinformatics tools and databases, including BLAST, Ensembl, UniProt and the UCSC Genome Browser, to assess impact on putative fusion proteins and compare them with existing, previously described fusions[8,17].

***TERT* promoter mutation sequencing.** Characterized mutations in the *TERT* promoter, C228T, and C250T variants with G>A nucleotide substitutions at genomic positions 1,295,228 and 1,295,250 (hg19), respectively[28], were sequenced using the Sanger method in midline samples (n = 14) and cortical samples (n = 10) using the following cycling conditions: 96 °C for 1 min, 96 °C for 10 s, 60 °C for 5 s, 72 °C for 1 s and 72 °C for 30 s, repeated for 33 cycles. Primer sequences are detailed in **Supplementary Table 8**.

DNA methylation analysis. Extracted tumor DNA was analyzed for genome-wide DNA methylation patterns using the HumanMethylation450 BeadChip platform according to instructions from the manufacturer (Illumina) and analyzed as described in refs. 4,5. From the selection of probes on the array, we removed probes from sex chromosomes (chromosomes X and Y) as well as those located at sites with documented SNPs (according to dbSNP). Methylation values were normalized using the Subset-quantile Within Array Normalization (SWAN) procedure provided in the R package minfi[29]. We performed hierarchical clustering using the 10,000 most variable sites. Distance was assessed using $d = 1 - r$, where r is the Pearson product-moment coefficient. Clustering was performed using average linkage (UPGMA) and was validated for the robustness of the procedure via multiscale resampling (1,000 iterations) using the R package pvclust[30] (**Supplementary Fig. 5**).

Copy number variant detection. To assay CNVs in our samples, we used a methylation-based method and controls implemented in the R/Bioconductor packages CopyNumber450k and CopyNumber450kData, respectively. CNV analysis for copy number gains and losses in previously described genes of interest is presented in **Supplementary Tables 4** and **5**, respectively. Gross genomic aberrations were assessed using UCSC Genome Browser banding and our in-house algorithm described above. Bands covered with >90% significantly amplified or deleted segments were counted as 'abnormal' and summed for each sample. Samples were grouped together by mutation type, and a t test was performed to assess statistically significant differences in aberration count between mutation subgroups. CNV analysis for broad areas of genomic instability is included in **Supplementary Table 7**.

Cell lines and protein blotting. hTert-immortalized NHAs were obtained from A. Guha (Labatt Brain Tumour Centre) and were grown in DMEM supplemented with 10% FBS as previously described. The DIPGIV primary cell line (a kind gift from M. Monje, Stanford University) and KNS42 cells (purchased from the Japanese Collection of Research Bioresources (JCRB) Cell Bank) were grown as previously described[31]. Mycoplasma-tested cell lines were starved of serum and growth factors for 1 h before protein extraction. Proteins were extracted in Tris-NaCl-EDTA lysis buffer as previously described[32], and blotting was performed for phosphorylated SMAD1/5/8 (Cell Signaling Technology, 9511; 1:500 dilution in 5% BSA solution) on total lysates from NHA cells and from DIPGIV cells as previously described[12]. β-actin (13E5, Cell Signaling Technology, 4970; 1:1,000 dilution in 5% BSA solution) was used as a loading control.

Quantitative PCR for BMP target genes. qPCR was performed to assess the levels of activity downstream of the ACVR1 (ALK2) receptor in total RNA

extracted from DIPGIV (H3.1 Lys27Met, ACVR1 Gly328Val), KNS42 (H3.3 Gly34Val, ACVR1 wild type) and NHA cells grown in DMEM supplemented with 10% FBS using *ID1*, *ID2*, *ID3* and *SNAI1* with the primer sequences detailed in **Supplementary Table 8**. Briefly, total RNA was extracted using the miRNeasy mini kit (Qiagen) according to the manufacturer's instructions, with purity and integrity assessed using Nanodrop (Thermo Fisher) and Experion (Bio-Rad) methodologies. RNA (100 ng) was used for reverse transcription with iScript RT Supermix (Bio-Rad) following the manufacturer's instructions. RT-PCR was run on a LightCycler 96 (Roche) with the SsoFast Evagreen SuperMix kit (Bio-Rad). Cycling conditions were 95 °C for 30 s followed by 40 cycles of 95 °C for 5 s and 60 °C for 20 s. Fold change values were calculated using the $2^{-\Delta\Delta C_t}$ method with *ACTB* expression and NHA cells used as the calibrator.

Immunohistochemistry. Immunohistochemistry was performed on formalin-fixed, paraffin-embedded slides from case with (n = 4) or without (n = 3) *ACVR1* mutation and control normal brain (n = 1) to assess phosphorylation of SMAD1/5/8 downstream of the ACVR1 (ALK2) receptor (reviewed in ref. 33). Immunohistochemistry was carried out as previously described[34] using antibodies to phosphorylated SMAD1/5/8 (Cell Signaling Technology, 9511; 1:50 dilution) and SMAD1 (Invitrogen, 38-5400; 1:100 dilution). Immunohistochemistry processing and imaging were performed blinded to *ACVR1* mutation status with representative images presented in **Figure 2**.

FISH for *PDGFRA* amplification. FISH was performed as previously described in refs. 35,36. Briefly, FISH was performed using 4-μm tissue sections from a subset of midline HGAs, with a BAC probe directed against the *PDGFRA* genomic locus in chromosomal region 4q12 (RP11-231c18; green) and a probe directed against 4p11.2-4q11.1 (CEN4; red) as a control to visualize chromosome 4. Scoring is included in **Supplementary Table 1**, and representative images are shown in **Supplementary Figure 2**.

23. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
24. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
25. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
26. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
27. McPherson, A. *et al.* deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. *PLoS Comput. Biol.* **7**, e1001138 (2011).
28. Horn, S. *et al. TERT* promoter mutations in familial and sporadic melanoma. *Science* **339**, 959–961 (2013).
29. Maksimovic, J., Gordon, L. & Oshlack, A. SWAN: subset-quantile within array normalization for Illumina Infinium HumanMethylation450 BeadChips. *Genome Biol.* **13**, R44 (2012).
30. Suzuki, R. & Shimodaira, H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**, 1540–1542 (2006).
31. Monje, M. *et al.* Hedgehog-responsive candidate cell of origin for diffuse intrinsic pontine glioma. *Proc. Natl. Acad. Sci. USA* **108**, 4453–4458 (2011).
32. Siegel, P.M., Ryan, E.D., Cardiff, R.D. & Muller, W.J. Elevated expression of activated forms of Neu/ErbB-2 and ErbB-3 are involved in the induction of mammary tumors in transgenic mice: implications for human breast cancer. *EMBO J.* **18**, 2149–2164 (1999).
33. Schmierer, B. & Hill, C.S. TGFβ-SMAD signal transduction: molecular specificity and functional flexibility. *Nat. Rev. Mol. Cell Biol.* **8**, 970–982 (2007).
34. Kleinman, C.L. *et al.* Fusion of *TTYH1* with the C19MC microRNA cluster drives expression of a brain-specific *DNMT3B* isoform in the embryonal brain tumor ETMR. *Nat. Genet.* **46**, 39–44 (2014).
35. Ramkissoon, L.A. *et al.* Genomic analysis of diffuse pediatric low-grade gliomas identifies recurrent oncogenic truncating rearrangements in the transcription factor *MYBL1*. *Proc. Natl. Acad. Sci. USA* **110**, 8188–8193 (2013).
36. Firestein, R. *et al. CDK8* is a colorectal cancer oncogene that regulates β-catenin activity. *Nature* **455**, 547–551 (2008).