The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma

Gang Wu1,15, Alexander K Diaz2,3,15, Barbara S Paugh2, Sherri L Rankin2, Bensheng Ju4, Yongjin Li1, Xiaoyan Zhu2, Chunxu Qu1, Xiang Chen1, Junyuan Zhang2, John Easton5, Michael Edmonson1, Xiaotu Ma1, Charles Lu6, Panduka Nagahawatte1, Erin Hedlund1, Michael Ruscetti1, Stanley Pounds7, Tong Lin7, Arzu Onar-Thomas7, Robert Huether1, Richard Kriwacki8, Matthew Parker1, Pankaj Gupta1, Jared Beckford1, Lei Wei9, Heather L Mulder5, Kristy Boggs5, Bhavin Vadodaria5, Donald Yergeau5, Jake C Russell2, Kerri Ochoa6, Robert S Fulton6, Lucinda L Fulton6, Chris Jones1,11,12, Frederick A Boop12, Alberto Broniscer13, Cynthia Wetmore13, Amar Gajjar13, Li Ding6, Elaine R Mardis6, Richard K Wilson6, Michael R Taylor4, James R Downing14, David W Ellison14, Jinghui Zhang1 & Suzanne J Baker2,3 for the St. Jude Children’s Research Hospital–Washington University Pediatric Cancer Genome Project

Pediatric high-grade glioma (HGG) is a devastating disease with a less than 20% survival rate 2 years after diagnosis1. We analyzed 127 pediatric HGGs, including diffuse intrinsic pontine gliomas (DIPGs) and non-brainstem HGGs (NBS-HGGs), by whole-genome, whole-exome and/or transcriptome sequencing. We identified recurrent somatic mutations in ACVR1 exclusively in DIPGs (32%), in addition to previously reported frequent somatic mutations in histone H3 genes, TP53 and ATRX, in both DIPGs and NBS-HGGs2–5. Structural variants generating fusion genes were found in 47% of DIPGs and NBS-HGGs, with recurrent fusions involving the neurotrophic receptor genes NTRK1, NTRK2 and NTRK3 in 40% of NBS-HGGs in infants. Mutations targeting receptor tyrosine kinase–RAS-PI3K signaling, histone modification or chromatin remodeling, and cell cycle regulation were found in 68%, 73% and 59% of pediatric HGGs, respectively, in both DIPGs and NBS-HGGs. This comprehensive analysis provides insights into the unique and shared pathways driving pediatric HGG within and outside the brainstem.

Although childhood and adult HGGs share related histopathological characteristics, adult HGGs arise predominantly in the cerebral cortex, whereas childhood HGGs more frequently involve a broader spectrum of locations. There are also substantial differences in the molecular features of pediatric and adult HGGs3,6–16. Histone H3 gene (H3F3A and HIST1H3B) mutations encoding p.Lys27Met are frequent in DIPGs, which arise in the brainstem almost exclusively in children, and in pediatric HGGs in midline structures such as the thalamus and cerebellum, whereas histone H3 gene mutations encoding p.Gly34Arg or p.Gly34Val occur in pediatric HGGs of the cerebral cortex3–5,17. In contrast, histone H3 gene mutations are extremely rare in adult HGGs3. HGGs arising in infants younger than 3 years of age have a better prognosis and a lower frequency of TP53 mutations, suggesting that there may be age-dependent subgroups of HGG, even within the pediatric population2. Thus, the selective pressures driving gliomagenesis in children vary with age and anatomical site.

To more comprehensively understand the pathways driving childhood glioma, we analyzed the genomic landscape of HGGs from 118 pediatric cases (127 tumors, 108 matched to germline DNA) consisting of 57 DIPGs and 70 NBS HGGs by whole-genome (n = 42), whole-exome (n = 80) and transcriptome (n = 75) sequencing (Supplementary Tables 1–9). A total of 39,590 sequence mutations, including single-nucleotide variations (SNVs) and small insertions or deletions (indels), and 2,039 structural variations were found by whole-genome sequencing, and an additional 2,600 sequence mutations and 138 structural variations were found by exome sequencing and transcriptome sequencing, respectively. Overall, the cohort showed a median background mutation rate of 9 × 10−7 mutations per base and a median of 22 structural

1Department of Computational Biology, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA. 2Department of Developmental Neurobiology, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA. 3Integrated Biomedical Sciences Program, University of Tennessee Health Science Center, Memphis, Tennessee, USA. 4Department of Chemical Biology and Therapeutics, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA. 5Department of Biostatistics, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA. 6Department of Structural Biology, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA. 7Biostatistics and Bioinformatics, Roswell Park Cancer Institute, Buffalo, New York, USA. 8Department of Oncology, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA. 9Biostatistics and Bioinformatics, Roswell Park Cancer Institute, Buffalo, New York, USA. 10Division of Cancer Therapeutics, Institute for Cancer Research, London, UK. 11Department of Surgery, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA. 12Department of Pathology, St. Jude Children’s Research Hospital, Memphis, Tennessee, USA. 13These authors contributed equally to this work. Correspondence should be addressed to S.J.B. (suzanne.baker@stjude.org) or Jinghui Zhang (jinghui.zhang@stjude.org).

Received 1 July 2013; accepted 6 March 2014; published online 6 April 2014; doi:10.1038/ng.2938
SMAD to total SMAD normalized to the empty vector control is shown below. With retroviruses expressing Flag-tagged wild-type ACVR1 or the indicated mutants and serum starved for 2 h. Quantification of the ratio of phosphorylated ACVR1 phenotype. Embryos injected with wild-type embryos exhibiting a dorsalized or ventralized embryos. The graph shows the percentage of embryos expressing the Arg258Gly (100%, Gly328Asp (83%, n = 23), Gly328Val (100%, n = 20) and Arg206His (100%, n = 27) mutants and the reduced severity of ventralization showing a decrease in the proportion of embryos with the V5 phenotype without the formation of dorsal structures for embryos expressing the Gly356Asp (89%, n = 27) and Gly328Val (49%, n = 24) mutants. Scale bar, 200 μm. (d) ACVR1 mutations drive increased levels of phosphorylated SMAD1/5 (pSMAD1/5) in primary astrocyte cultures. Protein blots are shown for lysates of primary astrocytes isolated from the brains of neonatal mice with conditional knockout of Tp53 transduced with retroviruses expressing Flag-tagged wild-type ACVR1 or the indicated mutants and serum starved for 2 h. Quantification of the ratio of phosphorylated SMAD to total SMAD normalized to the empty vector control is shown below.

**Figure 1** Recurrent genetic alterations in pediatric HGG. Genetic alterations detected in 19 genes, including ACVR1 and the gene most recurrently mutated in the pathways indicated on the left, are displayed. A diagonal white line indicates loss of the wild-type allele or a male case for ATRX, which is X linked. H3F3A (H3.3) and HIST1H3B (H3.1) mutations are grouped together in the category histone H3. Structural variants involving NTRK1, NTRK2 or NTRK3 and copy number variants (CNVs) of components of the CCND1, CCND2, CCND3, CDK4 or CDK6 G1 checkpoint complex are grouped together. Tumor subgroup (DIPG or NBS-HGG), location of NBS-HGG (midline versus cerebral hemisphere) and World Health Organization (WHO) tumor grade are indicated. White boxes for location or tumor grade indicate information not available. Mutations for 112 HGGs are shown. Four hypermutator samples and 11 samples for which only RNA-seq data were available were excluded from this summary. Data are shown in tabular form in Supplementary Table 9.

**Figure 2** ACVR1 mutations in DIPG activate BMP signaling. (a) Missense ACVR1 substitutions in DIPG were clustered in the GS or kinase domains. Each red circle indicates a DIPG sample carrying the specified alteration, and an asterisk indicates alterations previously found as germline mutations in individuals with FOP. The extracellular domain (EC) and transmembrane domain (TM) were not affected by mutations. (b) Human ACVR1 mutants ventralize zebrafish embryos. The graph shows the percentage of embryos exhibiting a dorsalized or ventralized phenotype. Embryos injected with wild-type ACVR1 mRNA (WT) showed a dorsalized phenotype C1–C4 (as shown in Supplementary Fig. 1), whereas embryos injected with mutant ACVR1 mRNA showed a ventralized phenotype (increasing severity from left to right). The Arg258Gly mutant had the least severe effect, resulting only in the V3–V4 ventralized phenotype, whereas the Gly328Val mutant had the most severe effect, with 90% of embryos showing the V5 ventralized phenotype. The number of embryos examined is shown on top. (c) Representative phenotype images of zebrafish embryos injected with the indicated ACVR1 mRNA. Untreated embryos expressing the Arg258Gly, Gly328Glu, Gly328Trp and Arg206His mutants have few to no dorsal structures, and embryos expressing the Gly356Asp and Gly328Val mutants are more severely affected. Treatment with LDN-193189 reversed the ventralization effects of mutant ACVR1, as demonstrated by the partial rescue of dorsal structures (e.g., the head) in embryos expressing the Arg258Gly (100%, n = 20), Gly328Asp (83%, n = 23), Gly328Trp (100%, n = 20) and Arg206His (100%, n = 27) mutants and the reduced severity of ventralization showing a decrease in the proportion of embryos with the V5 phenotype without the formation of dorsal structures for embryos expressing the Gly356Asp (89%, n = 27) and Gly328Val (49%, n = 24) mutants. Scale bar, 200 μm. (d) ACVR1 mutations drive increased levels of phosphorylated SMAD1/5 (pSMAD1/5) in primary astrocyte cultures. Protein blots are shown for lysates of primary astrocytes isolated from the brains of neonatal mice with conditional knockout of Tp53 transduced with retroviruses expressing Flag-tagged wild-type ACVR1 or the indicated mutants and serum starved for 2 h. Quantification of the ratio of phosphorylated SMAD to total SMAD normalized to the empty vector control is shown below.

Among recurrent mutations in pediatric HGG, the most frequently mutated gene not previously identified in cancer was ACVR1 (also known as ALK2) encoding a bone morphogenetic protein (BMP) type I receptor.
receptor (Figs. 1 and 2, and Supplementary Figs. 2 and 3). Clonal missense ACVR1 mutations were found exclusively in DIPGs (32%) and were significantly associated with younger age, longer survival time and the presence of HIST1H3B mutation encoding p.Lys27Met (P < 0.0000001) or with PIK3CA or PIK3R1 mutations (P < 0.005) (Figs. 1 and 2, Supplementary Fig. 3 and Supplementary Tables 4 and 5). Four of these somatic ACVR1 mutations were the same as germline mutations previously identified in the autosomal dominant syndrome fibrodysplasia ossificans progressiva (FOP), in which aberrant cellular differentiation drives progressive heterotopic ossifications18,19. All residues affected by mutation in DIPGs clustered around either the inhibitory glycine-serine-rich (GS) domain or the ATP-binding pocket of the kinase domain and would be expected to shift the kinase to an active conformation (Fig. 2 and Supplementary Fig. 3c)20. Indeed, mutations affecting these residues induced a weak gain of function20,21. A previous study showed that the ACVR1 mutation encoding p.Arg206His caused a ventilated phenotype in zebrafish embryos, an indicator of BMP pathway activation22. We tested all of the ACVR1 mutations found in DIPG using this assay. Zebrafish embryos injected with wild-type human ACVR1 mRNA displayed a mild dorsalized phenotype consistent with BMP pathway inhibition, whereas injection with mRNA for all six ACVR1 mutants, as demonstrated by the rescue of dorsal head structures (Fig. 2b,c and Supplementary Fig. 3d,e). A moderate dose of LDN-193189, a highly selective antagonist of the BMP pathway22,23, partially reversed the ventralization effects induced by ACVR1 mutants, as demonstrated by the rescue of dorsal head structures for the Arg258Gly, Gly328Glu, Gly328Trp and Arg260His mutants and the reduced severity of ventralization for the Gly356Asp and Gly328Val mutants (Fig. 2c). Expression of ACVR1 mutants in mouse primary astrocyte cultures caused increased levels of phosphorylated AKT (pAKT) and 44/p42 MAPK (p-44/p42 MAPK) in tumor relative to surrounding normal tissue. Scale bar, 50 μm.

Figure 3 Structural variants generate oncogenic chimeric NTRK fusion proteins. (a) All fusions included the C-terminal kinase domain from NTRK1, NTRK2 or NTRK3 (blue). N-terminal fusion partners include the tropomyosin domain (yellow) of TPM3, an actin-binding protein fused to NTRK1, and the BTB/POZ dimerization domain (gray) and the Kelch domain (orange) from the topoisomerase I-interacting protein BTBD1 or the pointed protein-protein interaction domain (purple) of the ETS transcription factor ETV6 fused to NTRK3. The N terminus of the actin-binding protein vinculin (light blue; VCL) was fused to NTRK2, and the N terminus of the ATP/GTP-binding protein AGBL4 (green) was fused to NTRK2. The functional carboxypeptidase domain of AGBL4 is not present in the fusion protein. For each fusion protein, the dashed red line shows the fusion point, with the amino acid positions of the N-terminal and C-terminal fusion partners at the breakpoint indicated. The number of amino acids in each full-length fusion protein is shown at the lower right of each fusion. (b) NTRK fusion proteins induce high-grade astrocytomas. Tumors induced by BTBD1-NTRK3 showed the frequent presence of giant cell glioblastoma. Immunohistochemical analysis of seven independent mice per construct shows pleomorphic tumor cells, expressing Flag-tagged TPM3-NTRK1 (top) or BTBD1-NTRK3 (bottom) and implanted into mouse brain. Representative hematoxylin and eosin staining (of seven independent mice per construct) shows pleomorphic tumor cells, many with features of astrocytic differentiation and high mitotic activity. Tumors induced by BTBD1-NTRK3 showed the frequent presence of giant cells reminiscent of giant cell glioblastoma. Immunohistochemical analysis showed expression of Flag-tagged NTRK fusion proteins and elevated levels of phosphorylated AKT (pAKT) and 44/p42 MAPK (p-44/p42 MAPK) in tumor relative to surrounding normal tissue. Scale bar, 50 μm.

Figure 4 Pediatric HGG alterations in histone modifiers or chromatin regulators. Alterations identified in NBS-HGG or DIPG are shown. Genetic alterations were identified in proteins that attach (writers; above) or remove (erasers; below) post-translational modifications of lysine residues (Lys4, Lys9, Lys18, Lys27, Lys36) in the tail of histone H3, as well as proteins involved in modifications of other histones or chromatin remodelers. Notably, there were no alterations in writers or erasers of Lys27, which is directly mutated at high frequency in DIPG and, to a lesser degree, in NBS-HGG. Ub, ubiquitination; ac, acetylation; ph, phosphorylation; H2, histone H2; H4, histone H4.
a downstream indication of active BMP signaling, with varying magnitude (Fig. 2d). LDN-193189 also effectively blocked signaling to phosphorylated SMAD1/5 downstream of mutant ACVR1 in primary astrocytes (Supplementary Fig. 3f).

Recurrent and clonal activating mutation of ACVR1 in 32% of DIPGs provides strong evidence that ACVR1 is an oncogenic driver in this disease. However, germline ACVR1 mutation in the genetic syndrome FOP is not associated with cancer predisposition, indicating that ACVR1 mutation likely provides a selective advantage in the presence of other critical mutations rather than driving tumor initiation. Consistent with this hypothesis, all six of the DIPG-associated ACVR1 mutants failed to render Tp53-null mouse astrocytes tumorigenic when implanted into the brain (data not shown). BMP signaling is associated with contrasting effects dependent on context, driving astrocytic differentiation or the proliferation of early hindbrain progenitors.

In tumors, BMP signaling induced the differentiation of medulloblastoma but drove either a differentiation or proliferation response in glioblastoma stem cells, related in part to the epigenetic state of the cell. In these context-dependent consequences likely drive the exclusive association between ACVR1 mutation and DIPG.

Principal-component analysis using the 1,001 most variable probe sets showed that HGG samples clustered by tumor location with no segregation of DIPGs by ACVR1 mutation status (Supplementary Fig. 3g). Genes involved in the regulation of immune system processes showed significantly different expression in DIPGs with and without ACVR1 mutation (Fisher’s exact test, P = 0.0002, false discovery rate (FDR) = 0.26%) (Supplementary Table 10).

Structural variants generating fusion genes were common and were identified in 47% of pediatric HGGs, in equal proportions of DIPGs and NBS-HGGs, from transcriptome and whole-genome sequencing data. Gene fusions involving the kinase domain of each of the three neurotrophin receptor (NTRK) genes and five different N-terminal fusion partners were identified in 4% of DIPGs and 10% of NBS-HGGs. Notably, 40% (4/10) of NBS-HGGs in children younger than 3 years old contained an NTRK fusion gene (Figs. 1 and 3, and Supplementary Tables 7 and 8). The NTRK receptors transduce a wide range of developmental signals in the nervous system, including the induction of neurite outgrowth, differentiation and neuronal survival or death. NTRK fusion genes have recently been identified at low frequency in low-grade pediatric astrocytomas as well as in adult glioblastomas. Two of the five NTRK fusions found in our cohort, TPM3-NTRK1 and ETV6-NTRK3, were previously identified in other cancer types and shown to be oncogenic.

To test the ability of NTRK fusion genes to drive glioma formation, we implanted Tp53-null primary mouse astrocytes transduced with retrovirus expressing Flag-tagged human TPM3-NTRK1 or BTBD1-NTRK3 into mouse brain. Both NTRK fusions induced high-grade astrocytomas with very short latency and complete penetrance (Fig. 3b and Supplementary Fig. 4). The resulting tumors showed elevated levels of phosphorylated AKT and p42/44 mitogen-activated protein kinase (MAPK), downstream indicators of phosphoinositide 3-kinase (PI3K) and MAPK pathway activation.

Although NTRK-activating fusions were specifically found at high frequency in tumors from infants, activation of receptor tyrosine kinase (RTK)-RAS-PI3K signaling through other mutations was frequent across the entire cohort, occurring in 69% of DIPGs and 67% of NBS-HGGs (Fig. 1, Online Methods and Supplementary Fig. 5). In contrast to previous reports detecting EGFRVIII expression in pediatric HGG, we only detected EGFRVIII expression in 1 of the 85 tumors analyzed by whole-genome sequencing or RNA sequencing (RNA-seq).

In addition to recurrent histone H3 gene, ATRX and SETD2 mutations reported previously, we also detected frequent mutations of other histone writers and erasers and of chromatin-remodeling genes (Fig. 4 and Supplementary Fig. 5). Interestingly, mutations affecting histone H3 writers were significantly more frequent in NBS-HGGs (P = 0.007), whereas frequencies of mutations affecting histone erasers were not significantly different in DIPGs and NBS-HGGs (P = 0.3). Although only ATRX mutations were highly recurrent, collectively, this group of genes involved in epigenetic regulation was targeted by mutation in 22% of DIPGs and 48% of NBS-HGGs, excluding histone H3 gene mutations (Fig. 4 and Supplementary Fig. 5).

These mutations were often concurrent with missense mutations in histone H3 genes. Indeed, including histone H3 gene mutations, 91% of DIPGs, 70% of midline NBS-HGGs and 48% of hemispheric NBS-HGGs contained mutations in histone genes and/or this subgroup of epigenetic regulators. p.Lys27Met histone H3.3 or H3.1 alteration results in a dominant loss of trimethylation at lysine 27 of histone H3 (H3K27me3) in the entire cellular pool of histone H3 proteins. Mutations that modulate H3K27me3 levels by targeting components of the Polycomb-repressive complex 2 (PRC2) that methylates H3K27 or the H3K27 demethylases KDM6A and KDM6B are found in other tumor types.

There were no such mutations across the entire HGG cohort, including in DIPGs with wild-type histone H3 genes, supporting the unique selective advantage of p.Lys27Met alteration in pediatric HGG (Fig. 4). Mutations in transcriptional regulators that affect the epigenetic landscape were also found, including focal amplifications of MYC and MYCN, which encode transcription factors that act to amplify the levels of expressed genes across the genome. Exome sequencing of DIPGs, 70% of midline NBS-HGGs and 48% of hemispheric NBS-HGGs contained mutations in histone genes and/or this subgroup of epigenetic regulators. p.Lys27Met histone H3.3 or H3.1 alteration results in a dominant loss of trimethylation at lysine 27 of histone H3 (H3K27me3) in the entire cellular pool of histone H3 proteins. Mutations that modulate H3K27me3 levels by targeting components of the Polycomb-repressive complex 2 (PRC2) that methylates H3K27 or the H3K27 demethylases KDM6A and KDM6B are found in other tumor types.

In addition to recurrent histone H3 gene, ATRX and SETD2 mutations reported previously, we also detected frequent mutations of other histone writers and erasers and of chromatin-remodeling genes (Fig. 4 and Supplementary Fig. 5). Interestingly, mutations affecting histone H3 writers were significantly more frequent in NBS-HGGs (P = 0.007), whereas frequencies of mutations affecting histone erasers were not significantly different in DIPGs and NBS-HGGs (P = 0.3). Although only ATRX mutations were highly recurrent, collectively, this group of genes involved in epigenetic regulation was targeted by mutation in 22% of DIPGs and 48% of NBS-HGGs, excluding histone H3 gene mutations (Fig. 4 and Supplementary Fig. 5).

These mutations were often concurrent with missense mutations in histone H3 genes. Indeed, including histone H3 gene mutations, 91% of DIPGs, 70% of midline NBS-HGGs and 48% of hemispheric NBS-HGGs contained mutations in histone genes and/or this subgroup of epigenetic regulators. p.Lys27Met histone H3.3 or H3.1 alteration results in a dominant loss of trimethylation at lysine 27 of histone H3 (H3K27me3) in the entire cellular pool of histone H3 proteins. Mutations that modulate H3K27me3 levels by targeting components of the Polycomb-repressive complex 2 (PRC2) that methylates H3K27 or the H3K27 demethylases KDM6A and KDM6B are found in other tumor types.

There were no such mutations across the entire HGG cohort, including in DIPGs with wild-type histone H3 genes, supporting the unique selective advantage of p.Lys27Met alteration in pediatric HGG (Fig. 4). Mutations in transcriptional regulators that affect the epigenetic landscape were also found, including focal amplifications of MYC and MYCN, which encode transcription factors that act to amplify the levels of expressed genes across the genome. Exome sequencing of DIPGs, 70% of midline NBS-HGGs and 48% of hemispheric NBS-HGGs contained mutations in histone genes and/or this subgroup of epigenetic regulators. p.Lys27Met histone H3.3 or H3.1 alteration results in a dominant loss of trimethylation at lysine 27 of histone H3 (H3K27me3) in the entire cellular pool of histone H3 proteins. Mutations that modulate H3K27me3 levels by targeting components of the Polycomb-repressive complex 2 (PRC2) that methylates H3K27 or the H3K27 demethylases KDM6A and KDM6B are found in other tumor types.

There was an enormous range in the complexity of the somatic mutations driving pediatric HGG (Supplementary Figs. 6 and 7). The ten NBS-HGGs in children under 3 years old showed significantly lower mutation rates than the rest of the cohort (P < 0.0001), suggesting that only a small number of driver mutations is required in tumors from this age group (Fig. 5 and Supplementary Fig. 7). ETV6-NTRK3 fusion was identified as one of only two non-silent alterations in SJHGG082, a...
glioblastoma from a 1-month-old subject. Notably, NTRK fusion genes, including two of the fusions found here, have been identified in multiple tumor types, including those from very young children such as congenital fibrosarcoma, as well as in papillary thyroid cancer, an adult disease. The high frequency of NTRK fusion genes in NBS-HGGs from children younger than 3 years of age, the paucity of additional mutations in these tumors and the rapid tumor onset in our experimental glioma model strongly suggest that these fusion genes are potent oncogenic drivers in early postnatal brain tumor development.

Four tumors, 3 with matched normal DNA, were scored as hypermutators, with an extremely high background mutation rate, including more than 800,000 somatic mutations in SJHGG11 in which biallelic germline PMS2 mutations were identified (Supplementary Fig. 7). Subject SJHGG003 carried a heterozygous germline PMS2 mutation and developed a grade IV hemispheric malignant glioneuronal tumor (MGNT) and a separate DIPG. Although both tumors independently acquired different somatic inactivating mutations in the remaining PMS2 allele, the basal mutation rate in the first tumor arising in this case, the MGNT (SJHGG003-D), was nearly 100-fold higher than in the DIPG that arose 2 years later (SJHGG003-A), demonstrating the potential range in the tumor mutation burden associated with inherited mutation in DNA mismatch repair genes (Fig. 5 and Supplementary Fig. 7). Hypermutator tumors were excluded from the evaluation of mutation frequency. Seven cases carried germline mutations in known cancer predisposition genes, including TP53, PMS2, MSH6 and NF1 (Supplementary Table 11).

Thirteen tumors (31%) analyzed by whole-genome sequencing had chromothripsis, identified by complex rearrangements with multiple interconnecting breakpoints corresponding to genomic segments with oscillating copy number states (Fig. 5, Supplementary Figs. 6 and 9, Supplementary Table 12 and Supplementary Note). In our cohort, chromothripsis resulted in oncogenic rearrangement, including BTT61-NTRK3 fusion and rearrangement or amplification of PDGFRα and EGFβ (Supplementary Fig. 8). Nearly half of all samples showing chromothripsis were collected before adjuvant therapy, indicating that the mechanism underlying their occurrence was, at least in some cases, independent of DNA-damaging therapeutics. SJHGG027_D, an NBS-HGG arising in a child with ataxia telangiectasia (A-T), had a relatively stable genome, despite a compromised DNA damage checkpoint due to the absence of functional ATM (Fig. 5 and Supplementary Fig. 7). Multiple subclones were identified in almost all HGG tumors. A founder clone or a descendant of a founder clone in the diagnosis tumor could seed the development of relapsed or autopsy tumor (Supplementary Fig. 10 and Supplementary Note).

We identified TERT promoter mutations in only 2% of DIPGs and 3% of NBS-HGGs, in strong contrast to the frequency of 86% identified in adult primary glioblastomas.

The genomic landscape of pediatric HGGs also includes frequent mutations in common cancer pathways, consistent with previous reports. TP53 mutations occurred in 42% of pediatric HGGs and were mutually exclusive with truncating mutations in the TP53-induced phosphate PPM1D, previously shown to impair the TP53-dependent G1 cell cycle checkpoint (Fig. 1). The G1 checkpoint regulators CCND1, CCND2, CCND3, CDK4 and CDK6 were predominantly amplified in DIPGs, whereas CDKN2A homozygous deletion was restricted to NBS-HGGs (Fig. 1). Taken together, mutations affecting cell cycle regulation, including the TP53 and RB pathways, were found in 59% of pediatric HGG (Fig. 1 and Supplementary Fig. 5).

This global view of the genetic landscape of pediatric HGG defines critical pathways driving a devastating spectrum of childhood brain tumors and identifies high-frequency mutations in potential therapeutic targets—ACVR1 in DIPGs and NTRK fusions in NBS-HGGs in infants.

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

Accession codes. All whole-genome sequencing, whole-exome sequencing and transcriptome data have been deposited at the European Genome-phenome Archive (EGA), which is hosted by the European Bioinformatics Institute (EBI), under accession EGAS0000100192.

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

ACKNOWLEDGMENTS

We thank the St. Jude Children’s Research Hospital Tissue Resource Facility and B. Gordon, M. Johnson, S. Brown and C. Calabrese in the St. Jude Small Animal Imaging Core for expert assistance with intracranial implantations. We thank E. Shore and A. Culbert (University of Pennsylvania) for helpful advice with antibody selection. This work was supported by the St. Jude Children’s Research Hospital—Washington University Pediatric Cancer Genome Project, by the American Lebanese and Syrian Associated Charities (ALSAC) of St. Jude Children’s Research Hospital, by grants from the US National Institutes of Health (P01 CA096832 to S.J.B., Jinghui Zhang and D.W.E. and R01 CA135554 to S.J.B.), by the Cure Starts Now Foundation and the Smile for Sophie Forever Foundation, and by Tyler’s Treehouse and Musicians Against Childhood Cancer.

AUTHOR CONTRIBUTIONS

S.J.B., Jinghui Zhang, A.K.D., B.S.P., J.E., L.D., E.R.M., R.K.W., M.R.T., J.R.D. and D.W.E. designed experiments or supervised research. A.G., A.B., C.W., F.A.B. and C.J. provided samples or clinical data. G.W., A.K.D., B.S.P., S.L.R., B.J., Y.L., X.Z., C.Q., X.C., Junyuan Zhang, J.E., M.E., X.M., C.L., P.N., E.H., M.R., S.P., T.L., A.O.-T., R.H., R.K., M.P., P.G., J.B., L.W., H.L.M., K.B., B.V., D.Y., J.C.R., K.O., R.S.F., I.L.F., L.D., E.R.M., R.K.W., M.R.T., J.R.D., D.W.E., Jinghui Zhang and S.J.B. performed experiments, analyzed data or prepared tables and figures. D.W.E. completed all pathological evaluations. S.J.B., Jinghui Zhang, G.W. and A.K.D. wrote the manuscript with contributions from D.W.E., J.R.D. and M.R.T.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Gottardo, N.G. & Gajjar, A. Chemotherapy for malignant brain tumors of childhood. J. Child Neurol. 23, 1149–1159 (2008).
2. Pollack, I.F. et al. Age and TP53 mutation frequency in childhood malignant gliomas: results in a multi-institutional cohort. Cancer Res. 61, 7404–7407 (2001).
3. Schwartzentruber, J. et al. Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature 482, 226–231 (2012).
4. Wu, G. et al. Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas. Nat. Genet. 44, 251–253 (2012).
5. 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).
6. Faury, D. et al. Molecular profiling identifies prognostic subgroups of pediatric glioblastoma and shows increased YB-1 expression in tumors. J. Clin. Oncol. 25, 1196–1208 (2007).
7. Gu, H.Q. et al. Genome-wide profiling using single-nucleotide polymorphism arrays identifies novel chromosomal imbalances in pediatric glioblastomas. Neuro-oncol. 12, 153–163 (2010).
8. 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).
9. 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).
10. Schiffman, J.D. et al. Oncogenic BRAF mutation with CDKN2A inactivation is characteristic of a subset of pediatric malignant astrocytomas. Cancer Res. 70, 512–519 (2010).
11. Puget, S. et al. Mesenchymal transition and PDGFRA amplification/mutation are key distinct oncogenic events in pediatric diffuse intrinsic pontine gliomas. PLoS ONE 7, e30933 (2012).
12. Zargooshi, 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).
13. Parsons, D.W. et al. An integrated genomic analysis of human glioblastoma multiforme. Science 321, 1807–1812 (2008).
14. Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature 455, 1061–1068 (2008).
15. Bax, D.A. et al. A distinct spectrum of copy number aberrations in pediatric high-grade gliomas. Clin. Cancer Res. 16, 3368–3377 (2010).
16. Barrow, J. et al. Homozygous loss of ADAM5A revealed by genome-wide analysis of pediatric high-grade glioma and diffuse intrinsic pontine gliomas. Neuro-oncol. 13, 212–222 (2011).
17. Sturm, D. et al. Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. Cancer Cell 22, 425–437 (2012).
18. Shore, E.M. & Kaplan, F.S. Role of altered signal transduction in heterotopic ossification and fibrodysplasia ossificans progressiva. Curr. Osteoporos. Rep. 9, 83–88 (2011).
19. 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).
20. Chaiukud, A. et al. Structure of the bone morphogenetic protein receptor ALK2 and implications for fibrodysplasia ossificans progressiva. J. Biol. Chem. 287, 36018–36027 (2012).
21. Bagarova, J. et al. Constitutively active ALK2 receptor mutants require type II receptor cooperation. Mol. Cell. Biol. 33, 2413–2424 (2013).
22. 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).
23. Cannon, J.E., Upton, P.D., Smith, J.C. & Morrell, N.W. Intersegmental vessel formation in zebrafish: requirement for VEGF but not BMP signalling revealed by selective and constitutively active ALK2 receptor mutants. Development 139, 3368–3377 (2010).
24. Bond, A.M., Bhalaria, O.G. & Kessler, J.A. The dynamic role of bone morphogenetic proteins in neural stem cell fate and maturation. Dev. Neurobiol. 72, 1068–1084 (2012).
25. Zhao, H., Ayrault, O., Zindy, F., Kim, J.H. & Roussel, M.F. Transcriptional down-regulation of Atoh1/MatH1 by bone morphogenetic proteins suppresses medulloblastoma development. Genes Dev. 22, 722–727 (2008).
26. Piccirillo, S.G. et al. Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. Nature 444, 761–765 (2006).
27. Lee, J. et al. Epigenetic-mediated dysfunction of the bone morphogenetic protein pathway inhibits differentiation of glioblastoma-initiating cells. Cancer Cell 13, 69–80 (2008).
28. Harel, L., Costa, B. & Fainzilber, M. On the death Trk. Dev. Neurobiol. 70, 298–303 (2010).
29. Thieue, C.J., Li, Z. & McKee, A.E. On Trk—the TrkB signal transduction pathway is an increasingly important target in cancer biology. Clin. Cancer Res. 15, 5962–5967 (2009).
30. Zhang, J. et al. Whole-genome sequencing identifies genetic alterations in pediatric low-grade gliomas. Nat. Genet. 45, 602–612 (2013).
31. Jones, D.T. et al. Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. Nat. Genet. 45, 927–932 (2013).
32. Frattini, V. et al. The integrated landscape of driver genomic alterations in glioblastoma. Nature 509, 111–116 (2014).
33. Lannon, C.L. & Sorensen, P.H. ET6V-NTRK3: a chimeric protein tyrosine kinase with transformation activity in multiple cell lineages. Semin. Cancer Biol. 15, 215–223 (2005).
34. Li, Y. et al. ET6V-NTRK3 fusion oncogene initiates breast cancer from committed mammary progenitors via activation of AP1 complex. Cancer Cell 12, 542–558 (2007).
35. Eguchi, M. et al. Fusion of ET6V to neurotrophin-3 receptor TRKc in acute myeloid leukemia with t(12;15)(p13;q25). Blood 93, 1355–1363 (1999).
36. Butti, M.G. et al. A sequence analysis of the genomic regions involved in the rearrangements between TPM3 and NTRK1 genes producing TRK oncogenes in papillary thyroid carcinomas. Genomics 28, 15–24 (1995).
37. Cetinbas, N. et al. Mutation of the salt bridge-forming residues in the ET6V-SAM domain interface blocks ET6V-NTRK3 induced cellular transformation. J. Biol. Chem. 288, 27940–27950 (2013).
38. Bax, D.A. et al. EGFRVIII deletion mutations in pediatric high-grade glioma and response to targeted therapy in pediatric glioma cell lines. Clin. Cancer Res. 15, 5763–5761 (2009).
39. Li, G. et al. Expression of epidermal growth factor variant III (EGFRVIII) in pediatric diffuse intrinsic pontine gliomas. J. Neurooncol. 108, 395–402 (2012).
40. 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).
41. Chan, K.M. et al. The histone H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene expression. Genes Dev. 27, 985–990 (2013).
42. Lewis, P.W. et al. Inhibition of PRG1–PRG9 (by a gain-of-function H3 mutation) found in pediatric glioblastoma. Science 340, 857–861 (2013).
43. Venneti, S. et al. Evaluation of histone 3 lysine 27 trimethylation (H3K27me3) and enhancer of zest 2 (EZH2) in pediatric glial and glioneuronal tumors shows decreased H3K27me3 in H3F3A K27M mutant glioblastomas. Brain Pathol. 23, 558–564 (2013).
44. Bender, S. et al. Reduced H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M mutant pediatric high-grade gliomas. Cancer Cell 24, 660–672 (2013).
45. Filligrabe, J., Kavanagh, E. & Joseph, B. Histone onco-modifications. Oncogene 30, 3391–3403 (2011).
46. Zhang, J. et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. Nature 481, 157–163 (2012).
47. Robinson, G. et al. Novel mutations target distinct subgroups of medulloblastoma. Nature 488, 43–48 (2012).
48. Dubuc, A.M. et al. Ablerrant patterns of H3K4 and H3K27 histone lysine methylation occur across subgroups in medulloblastoma. Acta Neuropathol. 125, 373–384 (2013).
49. Jones, D.T. et al. Dissecting the genomic complexity underlying medulloblastoma. Nature 488, 100–105 (2012).
50. Pugh, T.J. et al. Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. Nature 488, 106–110 (2012).
51. Nie, Z. et al. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. Cell 151, 68–79 (2012).
52. Lin, C.Y. et al. Transcriptional amplification in tumor cells with elevated c-Myc. Science 338, 438–442 (2012).
53. Killela, P.J. et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144, 27–40 (2011).
54. Kililea, P.J. et al. TERT promoter mutations occur frequently in gliomas and a subset of tumors derived from cells with low rates of self-renewal. Proc. Natl. Acad. Sci. USA 110, 6021–6026 (2013).
55. Klei, A. et al. Gain-of-function mutations of PPM1D/Wip1 impair the p53-dependent G1 checkpoint. J. Cell Biol. 201, 511–521 (2013).
ONLINE METHODS

Subject cohorts and sample details. HGGs (WHO grade III or IV) were requested from the St. Jude Children’s Research Hospital tissue resource core facility and from the Institute of Cancer Research/Royal Marsden Hospital with approval for genome sequence analysis in accordance with St. Jude Institutional Review Board (IRB) approval for the Pediatric Cancer Genome Project (PCGP) and the Clinical Research and Development Board of the Royal Marsden Hospital and the UK Children’s Cancer and Leukemia Group research ethics approval. Detailed clinicopathological and sequencing information are provided in Supplementary Table 1. There was a significant association between sex and tumor subtype, with 63% female DIPG cases and 63% male NBS–HGG cases (*P = 0.004*).

The study cohort comprised 127 tumors (57 DIPGs and 70 NBS–HGGs; 54 DIPGs and 54 NBS–HGGs with matching germline samples) from 118 individuals in 2 cohorts: a cohort for whole-genome sequencing (*n = 42, 20 DIPGs and 22 NBS–HGGs*) and a cohort for evaluating the frequency of abnormalities using whole-exome sequencing (*n = 80, 36 DIPGs and 44 NBS–HGGs*). Six tumors and their matched germline samples, including two hypermutator tumors (SJHGG003_D and SJHGG111_D) and four non-hypermutator tumors (SJHGG003_A, SJHGG008_A, SJHGG019_E and SJHGG022_D), were sequenced by both whole-genome and whole-exome sequencing. Among these tumors, 75 (31 DIPGs and 44 NBS–HGGs) were characterized by RNA-seq. In addition, 12 tumors (3 DIPGs and 9 NBS–HGGs) were characterized by RNA-seq only for structural variant discovery.

Tumor tissue was available from diagnosis and relapse in five cases (SJHGG019_E/S, SJHGG024_D/R, SJHGG033_D/R, SJHGG112_D/R and SJHGG115_D/R) and from diagnosis and autopsy in three cases (SJHGG003_D/A, SJHGG008_D/A and SJHGG093_D/A). One case developed two independent tumors, a hemispheric malignant glioneuronal tumor (SJHGG003_D) and a subsequent, independent DIPG (SJHGG003_A).

Histopathology was centrally reviewed by D.W.E., an experienced neuropathologist, and magnetic resonance imaging (MRI) of DIPG cases was reviewed by a pediatric neuro-oncologist (A.B.). DNA and RNA were extracted as previously described.

Whole-genome, whole-exome and transcriptome sequencing and analysis. Whole-genome sequencing, whole-exome sequencing and RNA-seq were performed as previously described. Paired-end sequencing was performed using the Illumina Genome Analyzer Ix or HiSeq platform with a 100-bp read length.

Whole-genome sequencing mapping, coverage and quality assessment, SNV and indel detection, tier annotation for sequence mutations, prediction of the deleterious effects of missense mutations and identification of LOH have been described previously. The reference human genome assembly GRCh37-lite was used for the mapping of all samples. Mapping statistics and coverage for each tumor on different sequencing platforms are summarized in Supplementary Table 2.

SNVs were classified into the following four tiers, as previously described: (i) tier 1: coding synonymous, nonsynonymous, splice-site and noncoding RNA variants; (ii) tier 2: conserved variants (cutoff: conservation score ≥ 500, on the basis of either the phastConsElements28way table or the phastConsElements17way table from the UCSC Genome Browser, and variants in regulatory regions annotated by UCSC annotation (regulatory annotations included are targetScan, ORegAnno, tfsbsConsites, vistaEnhancers, eponine, firstEF, L1 TAF1 Valid, Poly(A), switchDbTss, encodeUViennaRnaz, laminB1 and cpgIslandExt)); (iii) tier 3: variants in non-repeat-masked regions; and (iv) tier 4: the remaining SNVs.

All tier 1, tier 2 and tier 3 sequence mutations (including SNVs, indels and structural variations) discovered in non-hypermutator whole-genome sequencing samples were validated by a custom capture platform. The overall validation rate was 93%, with a median validation rate of 95% per sample. All tier 1 SNVs in the whole-exome sequencing cohort were also validated by custom capture. For all gene-coding indels found in whole-exome sequencing samples, we performed validation with the MiSeq platform, and the validation rate was 92% (167/182). Four non-hypermutator whole-genome sequencing samples (SJHGG003_A, SJHGG008_A, SJHGG019_E and SJHGG022_D) were subjected to exome sequencing, and the overlapping SNVs and indels were thus regarded as validated. Validated and high-quality variations for mutations in tiers 1–3 in non-hypermutator tumors are summarized in Supplementary Tables 3–5.

For two hypermutator whole-genome sequencing samples (SJHGG003_D and SJHGG111_D), exome sequencing served as validation. SNVs found in both whole-genome and whole-exome sequencing were regarded as having been validated. Validated and high-quality variations for tier 1 mutations in hypermutator tumors have been deposited at the PCGP Explorer website.

CNVs were identified by evaluating the difference in read depth between each tumor and its matching normal sample using a novel algorithm, CONSERTING (Copy Number Segmentation by Regression Tree in Next-Gen Sequencing; X.C., J. Wang, K. Roberts, M.P., S. Patel et al., unpublished data). The results are reported in Supplementary Table 6.

Structural variations in whole-genome sequencing data were analyzed using CREST and were annotated as previously described. Custom capture was used to validate somatic structural variations found in whole-genome sequencing samples. The results are reported in Supplementary Table 7.

Paired-end reads from RNA-seq were aligned to the following four database files using Burrows-Wheeler Aligner (BWA 0.5.5)36, (i) the human GRCh37-lite reference sequence, (ii) ReSeq, (iii) a sequence file representing all possible combinations of non-sequential pairs in ReSeq exons and (iv) the AceView database flat file downloaded from UCSC representing transcripts constructed from human ESTs. The mapping results from databases (ii)–(iv) were aligned to human reference genome coordinates. The final BAM file was constructed by selecting the best alignment in the four databases. Structural variation detection was carried out with CICERO, a novel algorithm that uses de novo assembly to identify structural variation in RNA-seq data (Y.L., B. Tang, M.R., J.E., K.B. et al., unpublished data). The structural variants detected in RNA-seq data were validated with MiSeq sequencing. Primer pairs were designed (with Primer3) to bracket the genomic regions containing putative structural variations. The structural variations found by RNA-seq are reported in Supplementary Table 8.

Microarray copy number and expression analysis. Affymetrix SNP60 arrays were analyzed as described. Candidate targets of focal amplification or deletion were identified from minimum common regions with copy number greater than 5 or less than 0.8. For DIPG samples, we also identified candidate targets of focal gain or loss as described. Briefly, we derived minimum common regions for recurrent focal gain (copy number >2.3) or recurrent focal deletion (copy number <1.7) found in at least two tumors or that were classified as a single focal gain or deletion. Regions associated with known CNVs were removed. All remaining regions with fewer than 60 genes were manually inspected for cancer- and/or glioma-related genes, and candidate targets of focal gain or deletion were selected. Focal amplifications of MYC, MYCN, PDGFRα, MET, CDK4, CDK6, CCND1, CCND2 and CCND3, as well as focal deletions of CDKN2A and CDKN2B, from SNP data were included in Figure 1 because many samples sequenced only the exome lacked these copy number data.

Affymetrix U133v2 expression array data were available for 71 HGGs, including 32 DIPG samples, 9 with ACVR1 mutation. Principal-component analysis was performed using GeneMaths with the top 1,000 most variable probes selected on the basis of median absolute deviation (MAD) score. Differentially expressed genes in DIPGs with and without ACVR1 mutation were selected with linear models (limma/R; P value <0.001), and enrichment of Gene Ontology (GO) terms among the selected genes was evaluated using the DAVID Bioinformatics Resource.

Frequency of mutations in pathways. To summarize samples altered in RTK-RAS-P13K pathways, we identified the relevant genes in the Kyoto Encyclopedia of Genes and Genomes (KEGG), Ingenuity and the National Cancer Institute (NCI)–Nature Protein Interaction Database and limited analysis to the following genes that had at least one somatic mutation in our study cohort: (i) RTKs: PDGFRα, KIT, EGFR, MET, CSF1R, FGFR1, FGFR3, FLT1, IGF1R, INSR, NTRK1, NTRK2 and NTRK3; (ii) phosphatidylinositol (3,4,5)-trip phosphate (PI3P) regulation: PTEN, PIK3CA and PIK3R1; (iii) ligands: FGFR3, FGFR5, IGF1, PDGFRα and VEGF; and (iv) downstream effectors: PCLG2, PLEKHA2, YES1, SGK1, G6PC, GNl1, GNC11, MLST8, PHLP1, PKN2, PPP2R2D, PRKAA2, PRKCA, PRKCCZ, RAC1, RPS6, RPTOR, SGK3,
STK11, INPP5D, SOS1, AKT3, JAK1, KRAS, MAP2K1, TSC2, GAB2, PPP11L, NF1, BRAF, RASGRF2 and RASSF5. To summarize the samples altered in cell cycle regulation, we included the following genes: TP53, TP73, CCND1, CCND2, CDK1, CDK6, CDKN1B, CDKN2A, CDKN2B, RB1, CDK4, CCND3, RBL1 and CDC27. For DNA repair genes, we included the following: ATM, ATR, BRCA1, BRCA2, ERCC2, ERCC3, ERCC8, LIG4, RAD23A, XAB2, MSH6, LIG1, LIG3, LIG4, POLD1, POLE, RAD23A, RAD50, RAD54B, RUVBL2, SETX, PMS2 and UVRAG. To summarize samples altered in histone modifications and chromatin regulators, we limited our analysis to the following genes: (i) histone writers: MLL, MLL2, MLL3, MLL4, PDRM9, SMYD3, SETD1A, SETD2, SETD3 and ASH1L; (ii) histone eraser: JMJD1C, KDM2A, KDM3A, KDM3B, KDM4B, KDM5B, KDM5C and SIRT7; (iii) chromo- remodulators: ATRX, SMARC4, BRWD1, CBX4, CHD2, CHD4, CHD6, CHD7, CHD8 and RAD54B; and (iv) other writers: UHRF1, NCOA1, STK4, UBR2 and UBR5. Therefore, our estimates of the numbers of mutations affecting these pathways are conservative.

**TERT promoter mutation analysis.** Owing to the high GC content of TERT promoter regions, there was poor coverage (~3x on average) in whole-genome sequencing at the two recurrently mutated sites (chr. 5: 1,295,228 and 1,295,250). Therefore, a portion of the TERT promoter (hg19 coordinates, chr. 5: 1,295,151–1,295,347) was amplified by PCR from tumors and matched normal samples and sequenced by Sanger sequencing using the primers listed in **Supplementary Table 13** to check for two previously described TERT promoter SNVs. Sequence was analyzed using SNPDetector60, and manual review was performed using Consed61.

**Statistical analysis.** Association between ACVR1 mutation and age at diagnosis was analyzed by the Kruskal-Wallis test: H = 6.62, 1 degree of freedom, P = 0.010 (adjusted for ties). Association of ACVR1 mutation with the survival of individuals with DIPG was quantified by the log-rank test: χ² = 10.1, 1 degree of freedom, P = 0.00149. On the basis of a two-variable Cox model for individuals with DIPG where age at diagnosis (as a continuous variable) and ACVR1 mutation status were included simultaneously, it appeared that age at diagnosis was not significantly associated with survival within this cohort (P = 0.55), but ACVR1 mutation status was (P = 0.0046). The co-occurrence of ACVR1 mutations with other mutations was quantified by Fisher’s exact test: ACVR1 mutation co-occurrence with HIST1H2B mutation encoding p.Lys27Met, P = 0.000001 and ACVR1 mutation co-occurrence with PIK3CA and PIK3R1 mutations, P = 0.0049702.

The association between sex and tumor subtype (DIPG versus NBS-HGG) was quantified by Fisher’s exact test. The association between tier 1 mutation rate and tumors in children less than 3 years of age was quantified by Kruskal-Wallis test (H = 13.42, 1 degree of freedom, P < 0.0001). The association between mutations in histone H3 writers and NBS-HGG was quantified by Pearson χ² (by Monte Carlo), P = 0.0007.

To identify samples with extremely high and low mutation rates, we used least median squares (LMS) as a method for the robust estimation of the center of the data and outlier detection45. Basically, LMS identifies the shortest interval that covers at least 50% of the data. This interval represents the densest ‘bulk’ of the data, and outliers are detected by comparing the data values to a normal distribution with the same interquartile range. Thus, the LMS-identified interval will not be influenced by outliers, and it can be used to detect such outliers.

**Germline mutation analysis.** We identified germline variants as previously described31. In this study, we implemented additional filters, where we (i) only kept genes listed in Cancer Gene Consensus or genes involved in DNA damage or repair and (ii) only kept nonsense or splice-site SNVs and indels, except a few missense mutations for TP53.

**Zebrafish embryo injections.** cDNA for wild-type or mutant human ACVR1 was cloned into pCS2, and mRNA was generated using the mMESSAGE mMACHINE kit (Life Technologies). Zebrafish embryos of the AB strain were microinjected with approximately 20 pg of wild-type or mutant ACVR1 mRNA at the one-cell stage. Embryos were inspected at 24 hours post-fertilization (h.p.f.) and scored as ventralized (classes V1–V5) or dorsalized (classes C1–C4) on the basis of published criteria22,23. For chemical rescue, 2.5 µM LDN-193189 (Stemgent) was added to embryos at about 3 h.p.f.

**Primary astrocyte transduction and in vivo tumorigenesis assays.** A sequence encoding a Flag tag was incorporated by PCR immediately before the termination codon of cDNA encoding the full ORF of human wild-type or mutant ACVR1, TPM3-NTRK1 and BTBD1-NTRK2, and the resulting fragments were cloned into the retroviral vectors MSCV-ires-mCherry (MCl)64, which was modified by inserting an adaptor into a blunted EcoRI site to generate a Gateway cloning site, or MSCV-ires-Puro (MIP)65. Sequences of all constructs were verified. Retrovirus was produced by transfecting HEK 293T cells with these viral constructs along with helper plasmids as previously described66. HEK 293T cells were from the American Type Culture Collection (ATCC) and tested negative for mycoplasma.

Mouse experiments were approved by the St. Jude Children’s Research Hospital Institutional Animal Care and Use Committee and are in compliance with national and institutional guidelines. Tp53-null primary mouse astrocytes were isolated from the cortex or brainstem of postnatal day 2 GFAP-cre; Tp53−/− mice67,68 and transduced with retroviruses as previously described69. Primary astrocyte cultures isolated from multiple mice were pooled for retroviral transductions to control for potential variation among primary cultures. For tumorigenesis studies, Tp53-null primary astrocytes at passage one were transduced with retroviruses generated from MIC vectors expressing human wild-type or mutant ACVR1, TPM3-NTRK1 or BTBD1-NTRK2 and implanted into female immuno- deficient mice at 6–16 weeks of age as described70. When mice displayed brain tumor symptoms, tumors were dissected and fixed in 4% paraformaldehyde in PBS at 4 °C overnight. Tumors were then processed, embedded in paraffin, and cut into 5-µm sections. Hematoxylin and eosin–stained sections from all collected tumors were evaluated by a clinical neuropathologist (D.W.E.). Immunohistochemistry was performed with microwave antigen retrieval, primary antibodies against phosphorylated AKT (Ser473) (Cell Signaling Technology, 9271; 1:50 dilution), phosphorylated p42/44 MAPK (Thr202/Tyr204) (Cell Signaling Technology, 4270; 1:400 dilution) or Flag (Sigma-Aldrich, F1804; 1:100 dilution), biotinylated secondary antibo- dies and horseradish peroxidase (HRP)- conjugated streptavidin (Elite ABC, Vector Labs), detected with NovaRED substrate (Vector Labs) and counter- stained with hematoxylin (Vector Labs). For TPM3-NTRK1, BTBD1-NTRK2 and empty vector controls, seven mice were implanted for each construct. For empty vector control, wild-type ACVR1 or Gly328Glu, Arg258Glu or Arg206His ACVR1, seven mice were implanted for each construct, and, for Gly356Asp, six mice were implanted. A few of the immuno-deficient mice were euthanized when they became ill without showing brain tumor symptoms (one expressing Gly328Glu ACVR1 at 97 d after implantation, one expressing Arg258Glu ACVR1 at 217 d after implantation, 2 expressing Gly356Asp at 139 and 190 d after implantation, and one expressing wild-type ACVR1 at 212 d after implantation). The remaining mice were euthanized between 219 and 222 d after implantation. Brains were formalin fixed and paraffin embedded, sectioned, stained with hematoxylin and eosin–stained sections in the area surrounding the implantation site and by immunohistochemistry for the mCherry marker expressed as part of the bicis- tronic message with ACVR1. None of the mice showed brain tumor symptoms, and no histological evidence of ACVR1-driven tumor formation was detected in the entire cohort.

**Protein blotting.** For protein blotting, Tp53-null primary mouse astrocytes were isolated from neonatal brainstem and transduced at passage one with MIP retroviruses expressing wild-type or mutant ACVR1. After 48 h, cells were selected with 2.5 µg/ml puromycin for 48 h. For serum starvation, cells were washed twice with PBS and then incubated for 2 h in medium without serum or growth factors. For treatment with LDN-193189, cells were incubated in standard astrocyte growth medium69 with vehicle (DMSO) or LDN-193189 (1 µM). Cells were lysed in RIPA buffer with protease and phosphatase inhibi- tors (Roche). Protein (20 µg) was resolved on 4–12% NuPAGE Bis-Tris gels, transferred to nitrocellulose membranes and detected with antibodies targeting phosphatase SMAD1/5 (Cell Signaling Technology, 9516; 1:1,000 dilution),...
total SMAD1/5/8 (Santa Cruz Biotechnology, sc-6031; 1:500 dilution), phosphorylated p38MAPK (Invitrogen, 44-684G; 1:1,000 dilution), total p38MAPK (Cell Signaling Technology, 9212; 1:1,000 dilution), Flag (Sigma-Aldrich, F1804; 1:1,000 dilution) and tubulin (Santa Cruz Biotechnology, sc-23948; 1:1,000 dilution). After incubation with the appropriate HRP-conjugated secondary antibody, membranes were exposed to chemiluminescent substrate (SuperSignal West Dura, Thermo Scientific). Images were obtained with the Odyssey Imaging System (LI-COR Biosciences).

RT-PCR validation of NTRK fusions. Random-primed cDNA was generated by reverse transcription from tumor RNA and used for PCR to identify the fusion gene of interest. PCR products were analyzed by Sanger sequencing. Primers are listed in Supplementary Table 13.

Statistical evaluation of chromothripsis in HGG tumors analyzed by whole-genome sequencing. Chromothripsis was described as localized chromosome shattering and repair occurring in a single event. The initial criterion was oscillation between 2 main CNV states53 and was found in 15 HGG tumors in this study. Most recently, Korbel and Campbell71 proposed four potential criteria for assessing chromothripsis: (i) clustering of breakpoints; (ii) randomness of DNA fragment joints; (iii) randomness of DNA fragment order; and (iv) ability to walk the derivative chromosome. Because randomness of DNA fragment order (criterion (iii)) was not entirely valid, even in Korbel and Campbell’s own analysis, we decided not to evaluate this feature. For the 13 tumors in Supplementary Table 12, we performed Bartlett’s goodness-of-fit test for exponential distribution to assess whether the distribution of structural variation breakpoints in each tumor departed from the null hypothesis of random distribution. A significant departure from random distribution supports clustering of structural variation breakpoints. To evaluate whether there was any bias in the DNA fragment joints categorized by structural variation type (i.e., deletion, tandem duplication, head-to-head rearrangement or tail-to-tail rearrangement), we applied the goodness-of-fit test separately for inter- and intrachromosomal events with a minimum of five structural variations. A significant P value suggests biased fragment joints, which would not support chromothripsis. When both inter- and intrachromosomal data were available, we reported the lower P value to represent a more conservative assessment of the random distribution for DNA fragment joints.

Details of tumor purity and tumor heterogeneity estimations and of tumor evolution analysis are included in the Supplementary Note.

56. Torchia, E.C., Boyd, K., Reilig, J.E., Qu, C. & Baker, S.J. EWS/FLI-1 induces rapid onset of myeloid/erythroid leukemia in mice. Mol. Cell. Biol. 27, 7918–7934 (2007).
57. Zhang, J. et al. A novel retinoblastoma therapy from genomic and epigenetic analyses. Nature 481, 329–334 (2012).
58. Wang, J. et al. CRESTMaps somatic structural variation in cancer genomes with base-pair resolution. Nat. Methods 8, 652–654 (2011).
59. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009).
60. Zhang, J. et al. SNPidetector; a software tool for sensitive and accurate SNP detection. PLOS Comput. Biol. 1, e53 (2005).
61. Gordon, D., Abajian, C. & Green, P. Consed: a graphical tool for sequence finishing. Genome Res. 8, 195–202 (1998).
62. Roussenov, P. Least Median Squares Regression. J. Am. Stat. Assoc. 79, 871–880 (1984).
63. Holmfeldt, L. et al. The genomic landscape of hypodiploid acute lymphoblastic leukemia. Nat. Genet. 45, 242–252 (2013).
64. Szymczak, A.L. et al. Correction of multi-gene deficiency in vivo using a single ‘self-cleaving’ 2A peptide-based retroviral vector. Nat. Biotechnol. 22, 589–594 (2004).
65. Kawauchi, D. et al. A mouse model of the most aggressive subgroup of human medulloblastoma. Cancer Cell 21, 168–180 (2012).
66. Persans, D.A. et al. Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. Blood 93, 488–499 (1999).
67. Bajenaru, M.L. et al. Astrocyte-specific inactivation of the neurofibromatosis 1 gene (NF1) is insufficient for astrocytoma formation. Mol. Cell. Biol. 22, 5100–5113 (2002).
68. Jonkers, J. et al. Synergistic tumor suppressor activity of BRCA2 and p53 in a conditional mouse model for breast cancer. Nat. Genet. 29, 418–425 (2001).
69. Fraser, M.M. et al. Pten loss causes hypertrophy and increased proliferation of astrocytes in vivo. Cancer Res. 64, 7773–7779 (2004).
70. Endersby, R., Zhu, X., Hay, N., Ellison, D.W. & Baker, S.J. Nonredundant functions for Akt isoforms in astrocyte growth and gliomagenesis in an orthotopic transplantation model. Cancer Res. 71, 4106–4116 (2011).
71. Korbel, J.O. & Campbell, P.J. Criteria for inference of chromothripsis in cancer genomes. Cell 152, 1226–1236 (2013).