nature genetics

Somatic histone H3 alterations in pediatric diffuse intrinsic pontine gliomas and non-brainstem glioblastomas

Gang Wu[1,8], Alberto Broniscer[2,8], Troy A McEachron[3,8], Charles Lu[4], Barbara S Paugh[3], Jared Becksfort[5], Chunxu Qu[5], Li Ding[4], Robert Huether[1], Matthew Parker[1], Junyuan Zhang[3], Amar Gajjar[2], Michael A Dyer[3], Charles G Mullighan[6], Richard J Gilbertson[3], Elaine R Mardis[4], Richard K Wilson[4], James R Downing[6], David W Ellison[6], Jinghui Zhang[1] & Suzanne J Baker[3] for the St. Jude Children's Research Hospital–Washington University Pediatric Cancer Genome Project[7]

To identify somatic mutations in pediatric diffuse intrinsic pontine glioma (DIPG), we performed whole-genome sequencing of DNA from seven DIPGs and matched germline tissue and targeted sequencing of an additional 43 DIPGs and 36 non-brainstem pediatric glioblastomas (non-BS-PGs). We found that 78% of DIPGs and 22% of non-BS-PGs contained a mutation in *H3F3A*, encoding histone H3.3, or in the related *HIST1H3B*, encoding histone H3.1, that caused a p.Lys27Met amino acid substitution in each protein. An additional 14% of non-BS-PGs had somatic mutations in *H3F3A* causing a p.Gly34Arg alteration.

Diffuse intrinsic pontine glioma, an aggressive brainstem astrocytic tumor, arises almost exclusively in children, usually with histopathological diagnosis of glioblastoma, and has a long-term survival rate of less than 10% (ref. 1). To understand the molecular pathogenesis of DIPG, we performed whole-genome sequencing (WGS) of DNA from DIPGs and matched normal tissue from seven affected individuals. Tumors from four subjects had a recurrent somatic adenine-to-thymine transversion (c.83A>T) in *H3F3A* resulting in a substitution to methionine at lysine 27 (p.Lys27Met) in the histone H3.3 protein (**Supplementary Fig. 1**). In the tumor from a fifth affected individual, an analogous adenine-to-thymine transversion (c.83A>T) encoding a p.Lys27Met alteration was identified in the closely related *HIST1H3B* gene, which codes for the histone H3.1 isoform.

To determine the frequency of mutations in the histone H3 gene family, we performed targeted sequencing of the exons encoding Lys27 in all 16 genes that code for histone H3 isoforms in a validation cohort comprising 43 DIPGs as well as 36 non-BS-PGs (see **Supplementary Methods** and **Supplementary Tables 1** and **2**). Including the original seven individuals with DIPG analyzed by WGS, we found recurrent adenine-to-thymine transversions encoding p.Lys27Met alterations in *H3F3A* or *HIST1H3B* in 78% of DIPGs and 22% of non-BS-PGs (**Fig. 1** and **Table 1**). In addition, a new guanine-to-adenine transition resulting in a p.Gly34Arg alteration in *H3F3A* was identified in 5 of 36 (14%) non-BS-PGs but not in any of the 50 DIPGs analyzed. These three different H3 mutations were mutually exclusive. There was no evidence of loss of heterozygosity at the mutated loci using WGS or SNP arrays in the discovery cohort or using Sanger sequencing or SNP arrays in the validation cohort. For 32 DIPG samples and 6 non-BS-PG samples in which an *H3F3A* or *HIST1H3B* mutation was identified and for which matched normal DNA was available, the germline DNA was found to be wild type, verifying that these mutations were somatic in nature (**Supplementary Table 1**). Of note, the p.Lys27Met alteration was found in seven of eight DIPG samples obtained before therapy, indicating that this alteration was not secondary to therapy-induced mutagenesis. There were no Lys27 or Gly34 mutations in any of the other 14 genes encoding histone H3 isoforms.

The identified alterations in histone H3 seem to be exclusive to pediatric high-grade gliomas. We also detected a *H3F3A* mutation encoding the p.Lys27Met alteration in 1 of 9 pediatric anaplastic astrocytomas (grade 3), but no mutations in any of the 16 histone H3 genes were detected in other pediatric brain tumors, including 7 low-grade brainstem

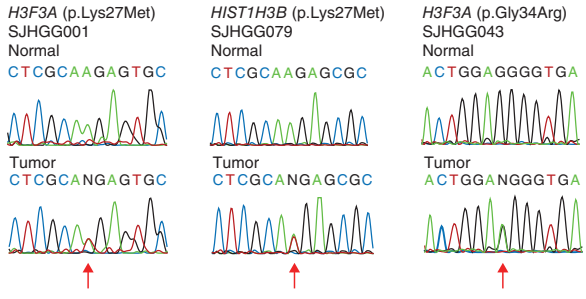

Figure 1 Recurrent somatic mutations in *H3F3A* and *HIST1H3B*. Sanger sequencing chromatograms showing representative *H3F3A* or *HIST1H3B* mutations encoding a p.Lys27Met substitution or a *H3F3A* mutation encoding a p.Gly34Arg substitution. Mutations are shown in the indicated tumor compared to matched normal DNA and are marked by arrows. N indicates the presence of both A and T for the p.Lys27Met alteration and both G and A for the p.Gly34Arg alteration. SJHGG001 and SJHGG079 are DIPGs, and SJHGG043 is a non-BS-PG.

[1]Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA. [2]Department of Oncology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA. [3]Department of Developmental Neurobiology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA. [4]The Genome Institute, Washington University School of Medicine, Saint Louis, Missouri, USA. [5]Hartwell Center of Biotechnology and Bioinformatics, St. Jude Children's Research Hospital, Memphis, Tennessee, USA. [6]Department of Pathology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA. [7]A full list of consortium members is provided in the **Supplementary Note**. [8]These authors contributed equally to this work. Correspondence should be addressed to R.K.W. (rwilson@wustl.edu), J.R.D. (james.downing@stjude.org), Jinghui Zhang (jinghui.zhang@stjude.org) or S.J.B. (suzanne.baker@stjude.org).

Received 21 December 2011; accepted 9 January 2012; published online 29 January 2012; doi:10.1038/ng.1102

Table 1 Frequency of recurrent somatic mutations in DIPG and GBM

Gene	Amino acid change	DIPG[a] (%)	non-BS-PG[b] (%)
H3F3A	p.Lys27Met	30 (60)	7 (19)
H3F3A	p.Gly34Arg	0	5 (14)
HIST1H3B	p.Lys27Met	9 (18)	1 (3)
All H3		39 (78)	13 (36)

[a]For DIPGs, total $n = 50$. [b]For non-BS-PGs, total $n = 36$.

gliomas evaluated by targeted Sanger sequencing, 15 low-grade non-brainstem gliomas evaluated by WGS, and 38 medulloblastomas and 22 ependymomas, and mutations were also not detected in an additional 170 non–central nervous system pediatric tumors (**Supplementary Tables 1** and **3**). We also did not find any evidence of structural variations in the histone H3 loci in any of the tumor subtypes evaluated by WGS. Furthermore, there were no occurrences of germline polymorphisms affecting Lys27 or Gly34 in any of the histone H3 genes in the dbSNP 135 database, which includes 39,484,957 SNPs submitted by the 1000 Genomes Project, or in the National Heart, Lung, and Blood Institute (NHLBI) exome sequencing project (ESP) that contains 5,400 exomes (see URLs). In fact, no non-silent, coding germline variants were found in *H3F3A*, and only one encoded amino acid change (p.Gln20Glu) was identified in *HIST1H3B* in these two public SNP databases.

H3F3A is located on chromosome 1q, a region of large-scale chromosomal gain in more than 20% of both DIPGs and non-BS-PGs[2–8]; however, there was no significant correlation between the presence of *H3F3A* or *HIST1H3B* mutations and gain of chromosome 1q ($P = 0.7$). There was also no significant association between *H3F3A* or *HIST1H3B* mutation and amplification of receptor tyrosine kinase genes ($P = 0.7$) or cyclin D family genes, *CDK4* or *CDK6* ($P = 0.5$), alterations that occur in greater than 30% of DIPGs[2,4,6].

Although it is unclear exactly how these mutations affect histone H3 function, the Lys27 and Gly34 residues are located within the highly conserved N-terminal tail of the histone H3 protein, which influences the dynamic regulation of chromatin structure and accessibility. Alterations at these invariant residues, which are close to the site where the tail exits the globular histone core of the nucleosome, may affect nucleosome structure and function by affecting histone-DNA interactions, chromatin compaction or interactions with other effectors that bind to histones. Complex post-translational modifications of the histone tail have an important role in epigenetic regulation of gene expression by affecting both chromatin state and direct interactions between modified histones and transcriptional activator or repressor complexes[9]. It is notable that trimethylation of Lys27 in histone H3 (H3K27me3) is associated with the silencing of genes, especially those involved in developmental processes in pluripotent cells, and monomethylation (H3K27me) is associated with gene activation[10]. Acetylation and deacetylation of Lys27 of histone H3 are also highly regulated, with acetylated Lys27 being associated with transcriptionally active regions[11]. The replacement of Lys27 with methionine implies a loss of function, as it removes the ability to methylate or acetylate this position, which may affect the role of histone H3 in either repression or activation. However, the mutations are present in the heterozygous state in only two of a large family of histone H3 genes and always encode the same amino acid substitutions, suggesting a gain-of-function phenotype. Additional studies are required to determine the functional impact of the p.Lys27Met and p.Gly34Arg alterations.

Of note, mutations were found affecting more than one histone H3 isoform. H3.1 and H3.2 are replication-dependent histones, synthesized and incorporated into nucleosomes during S phase, whereas H3.3 is a replication-independent isoform that is selectively enriched within actively transcribed genes and transcription factor–binding sites and at telomeres, where its incorporation can affect telomere stabilization[12]. The higher frequency of *H3F3A* mutations suggests that disruption of the specialized functions of H3.3 may provide a specific selective advantage. However, the recurrent and mutually exclusive monoallelic mutations encoding p.Lys27Met in either *H3F3A* or *HIST1H3B* suggest that a similar gain-of-function effect is possible, despite the differences in regulation of H3.3 and H3.1. It is possible that differential transcriptional or post-translational regulation of *H3F3A* and *HIST1H3B* relative to other histone H3 isoforms may in part explain why these two genes are selectively targeted for mutation in the cells of origin for DIPG and non-BS-PG.

Given the complex regulatory roles of histone H3, these mutations could potentially affect epigenetic regulation of gene expression, selective regulation of developmental genes or telomere length or stability. Although mutations in genes encoding post-translational modifiers of histones have been identified in a number of different cancers[13], this is the first report to our knowledge of somatic mutations in histone H3. Perhaps consistent with a role for the Lys27 residue in histone H3 in regulating expression of genes associated with development, histone H3 mutations were not identified in exome sequence data from adult glioblastomas[14]. This indicates that these mutations confer a selective advantage in the unique context of the developing brain and highlights the significant difference in the underlying biology of gliomagenesis in children and adults.

URLs. dbSNP 135, http://www.ncbi.nlm.nih.gov/projects/SNP/; NHLBI ESP, http://evs.gs.washington.edu/EVS/; European Genome-phenome Archive, http://www.ebi.ac.uk/ega/.

Accession numbers. WGS data have been deposited at the European Genome-phenome Archive (EGA), which is hosted by the European Bioinformatics Institute (EBI), under EGAS0000100192.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank X. Zhu and A. Diaz for assistance with PCR reactions, J. Partridge for helpful discussions, the Hartwell Center of Biotechnology and Bioinformatics at St. Jude Children's Research Hospital and Beckman Coulter Genomics. This work was supported by the St. Jude Children's Research Hospital–Washington University Pediatric Cancer Genome Project and the American Lebanese Syrian Associated Charities (ALSAC) of St. Jude Children's Research Hospital and by grants from the US National Institutes of Health (NIH) (CA096832), the Sydney Schlobohm Chair of Research from the National Brain Tumor Society, the Cure Starts Now Foundation, the Smile for Sophie Forever Foundation, Tyler's Treehouse Foundation, the Musicians Against Childhood Cancer and the Noyes Brain Tumor Foundation.

AUTHOR CONTRIBUTIONS

S.J.B., T.A.M., B.S.P., C.Q., J.R.D., E.R.M. and R.K.W. designed the experiments. A.B. and A.G. provided samples and clinical data. D.W.E. performed histopathological analyses. D.W.E., M.A.D., C.G.M., R.J.G. and J.R.D. provided data from other tumor types. G.W. and C.L. analyzed the whole-genome sequence data. T.A.M., B.S.P. and Junyuan Zhang performed validation experiments. J.B. analyzed the Sanger sequencing data for the validation cohort. R.H. performed structural modeling for the mutations. L.D. and Jinghui Zhang supervised the data analysis. G.W., M.P., Jinghui Zhang and S.J.B. prepared the tables and figures. J.R.D., Jinghui Zhang and S.J.B. wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

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

1. Broniscer, A. & Gajjar, A. *Oncologist* **9**, 197–206 (2004).
2. Paugh, B.S. *et al. J. Clin. Oncol.* **29**, 3999–4006 (2011).

3. Paugh, B.S. *et al. J. Clin. Oncol.* **28**, 3061–3068 (2010).
4. Barrow, J. *et al. Neuro. Oncol.* **13**, 212–222 (2011).
5. Bax, D.A. *et al. Clin. Cancer Res.* **16**, 3368–3377 (2010).
6. Zarghooni, M. *et al. J. Clin. Oncol.* **28**, 1337–1344 (2010).
7. Qu, H.Q. *et al. Neuro. Oncol.* **12**, 153–163 (2010).
8. Warren, K.E. *et al. Neuro. Oncol.* published online (7 November 2011), doi:10.1093/neuonc/nor190.
9. Campos, E.I. & Reinberg, D. *Annu. Rev. Genet.* **43**, 559–599 (2009).
10. Barski, A. *et al. Cell* **129**, 823–837 (2007).
11. Reynolds, N. *et al. EMBO J.* published online (2 December 2011), doi:10.1038/emboj.2011.431.
12. Goldberg, A.D. *et al. Cell* **140**, 678–691 (2010).
13. Füllgrabe, J., Kavanagh, E. & Joseph, B. *Oncogene* **30**, 3391–3403 (2011).
14. Parsons, D.W. *et al. Science* **321**, 1807–1812 (2008).