ABSTRACT

Children and young adults with glioblastoma (GBM) have a median survival rate of only 12 to 15 months, and these GBMs are clinically and biologically distinct from histologically similar cancers in older adults. They are defined by highly specific mutations in the gene encoding the histone H3.3 variant H3F3A, occurring either at or close to key residues marked by methylation for regulation of transcription—K27 and G34. Here, we show that the cerebral hemisphere-specific G34 mutation drives a distinct expression signature through differential genomic binding of the K36 tri-methylation mark (H3K36me3). The transcriptional program induced recapitulates that of the developing forebrain, and involves numerous markers of stem-cell maintenance, cell-fate decisions, and self-renewal. Critically, H3F3A G34 mutations cause profound upregulation of MYCN, a potent oncogene that is causative of GBMs when expressed in the correct developmental context. This driving aberration is selectively targetable in this patient population through inhibiting kinases responsible for stabilization of the protein.

SIGNIFICANCE: We provide the mechanistic explanation for how the first histone gene mutation in human disease biology acts to deliver MYCN, a potent tumorigenic initiator, into a stem-cell compartment of the developing forebrain, selectively giving rise to incurable cerebral hemispheric GBM. Using synthetic lethal approaches to these mutant tumor cells provides a rational way to develop novel and highly selective treatment strategies.

INTRODUCTION

The clinical and molecular differences observed in glioblastoma (GBM) of children and young adults compared with the more common, histologically similar lesions in older adults is strongly suggestive of a distinct underlying biology (1). The identification of unique and highly specific mutations in the gene H3F3A, encoding the variant histone H3.3A in GBM of...
Distinct molecular and clinical correlates of \( H3F3A \) mutation subgroups. A, heatmap representing differential gene expression signatures between G34 versus K27, and G34 versus wild-type, pediatric GBM specimens identified by Paugh and colleagues (3). Top 100 differentially expressed genes are shown for each comparison. B, gene set enrichment analysis (GSEA) for differential gene expression signatures identified by Schwartzentruber and colleagues (2) versus those from Paugh and colleagues (3). Top, G34 versus K27: enrichment score (ES) = 0.833, \( P \) (family-wise error rate [FWER]) = 0.0, \( q \) (false discovery rate [FDR]) = 0.0. Bottom, G34 versus wild-type: ES = 0.94, FWER \( P = 0.0 \), FDR \( q = 0.0 \). C, heatmap representing differential gene expression signatures between G34 versus K27, and G34 versus wild-type, pediatric GBM specimens from (2). Top 100 differentially expressed genes are shown for each comparison. D, GSEA for differential gene expression signatures identified in (3) versus those in (2). Top, G34 versus K27: ES = 0.88, FWER \( P = 0.03 \), FDR \( q = 0.04 \). Bottom, G34 versus wild-type: ES = 0.90, FWER \( P = 0.0 \), FDR \( q = 0.0 \). E, hierarchical clustering of the integrated gene expression datasets, highlighting specific clusters of G34- and K27-mutant tumors, distinct from a more heterogeneous group of wild-type cases. G34V tumors are represented by asterisks. F, K-means consensus clustering finds the most stable number of subgroups to be 3, marked by \( H3F3A \) mutation status. G, K27- and G34-mutant pediatric GBM in our integrated dataset have distinct age incidence profiles, with K27 tumors peaking at 7 years in contrast to G34 at age 14. The 2 G34V tumors were diagnosed at age 14 and 20. H, Kaplan-Meier plot for overall survival of pediatric patients with GBM stratified by \( H3F3A \) status. K27-mutant tumors have significantly shorter survival than G34 (\( P = 0.0164 \), log-rank test). A single G34V case for which data were available had an overall survival of 1.4 years. wt, wild-type.

**RESULTS**

Initial evidence suggested a distinct gene expression signature associated with mutations at the K27 (lysine to methionine, K27M) versus G34 (glycine to either arginine, G34R, or valine, G34V) residues (2). We validated these data by identifying differential expression patterns for mutations with G34 versus K27 mutations in 2 independent datasets for which mutation data were either publicly available or were ascertained in our laboratory (refs. 2, 3; Fig. 1). In both instances, highly significant differential gene expression was noted between G34-mutant tumors and K27 or wild-type cases (Fig. 1A and C), which was consistent across the datasets as assessed by gene set enrichment analysis (GSEA; Figs. 1B and D) with enrichment scores (ES) of 0.833 to 0.943 and \( P \) (family-wise error rate; FWER) and \( q \) (false discovery rate; FDR) values of 0.0 to 0.04. Given the considerable overlap in gene expression signatures between studies, we subsequently utilized an integrated dataset (Supplementary Table S1), where hierarchical clustering resolved G34- and K27-mutant tumors from a more heterogeneous wild-type subgroup (Fig. 1E), confirmed by k-means consensus clustering (Fig. 1F). These subgroups also showed important clinical differences, as previously described (2), with K27-mutant tumors arising in younger children (peak age 7 years; \( P = 0.0312 \), t test; Fig. 1G) and having a worse clinical outcome (\( P = 0.0164 \), log-rank test; Fig. 1H) compared with G34 tumors (peak age 14 years) and \( H3F3A \) wild-type tumors. There were no significant transcriptional or clinicopathologic differences between G34R and G34V tumors, although a lack of samples of the latter (\( n = 2 \)) precludes robust analyses.
To understand the functional significance of H3F3A mutations in cerebral hemispheric tumors, we turned to a well-characterized (4) model of pediatric GBM, the KNS42 cell line, which was derived from a 16-year-old patient and harbors the G34V mutation (Fig. 2A). In contrast to the reported data in a single pediatric GBM sample with G34R (2), KNS42 cells did not show increased levels of total histone H3K36 trimethylation compared with a panel of H3F3A wild-type pediatric glioma cells (Fig. 2B, Supplementary Fig. S1). KNS42 cells harbor a nonsynonymous coding change of ATRX (Q891E) that appears in the single-nucleotide polymorphism databases (rs3088074), and Western blot analysis shows no diminution of protein levels. As ATRX is a known chaperone of histone H3.3 to the telomeres, a wild-type protein would not be expected to convey the alternative lengthening of telomeres (ALT) phenotype, as observed (Supplementary Fig. S2); however, this ought not play a significant role in gene transcription as deposition of H3.3 in euchromatin is carried out by alternative chaperones such as HIRA.

We conducted chromatin immunoprecipitation linked to next-generation whole genome sequencing (ChiP-Seq) for H3K36me3 to test the hypothesis that, rather than total H3K36me3, the G34V mutation may instead result in differential binding of the trimethyl mark throughout the genome. Compared with H3F3A wild-type SF188 pediatric GBM cells, H3K36me3 was found to be significantly differentially bound in KNS42 cells at 5,130 distinct regions of the genome corresponding to 156 genes (DESeq P < 0.05, overall fold change >2, contiguous median coverage >2; Supplementary Table S2). These observations were not due to differential gene amplification, as concurrent whole genome DNA sequencing showed that these bound genes were not found in regions on cell line-specific copy number alterations (Fig. 2C; Supplementary Fig. S3 and Supplementary Table S2). As trimethyl H3K36 is regarded as an activating mark for gene expression (5), we concurrently conducted ChiP-Seq for RNA polymerase II to produce a readout of transcriptional activity, and observed a significant correlation between H3K36me3 and RNA polymerase II binding for the 156 differentially bound genes (R² = 0.923, P < 0.0001; Fig. 2D). By integrating the H3K36me3 and RNA polymerase II data, we derived a ranked list of differentially trimethyl-bound and expressed genes (Fig. 2E).

To investigate the functions of the transcriptional programs targeted by this novel mechanism, we conducted gene ontology analysis of the differentially bound and expressed genes. These data revealed highly significant enrichment of the processes involved in forebrain and cortex development, as well as differentiation of neurons and regulation of cell proliferation (Fig. 2G). We identified a subset of 16 genes to be part of the core enrichment group showing significant overlap between G34-mutant pediatric GBM specimens and transcription driven by differential binding of H3K36me3 in KNS42 cells (Supplementary Table S3). By mapping the expression of these genes to published signatures of restricted spatiotemporal areas of brain development (6), we noted highly elevated levels at embryonic and early fetal time-points, which rapidly tailed off through mid-late fetal development and postnatal and adult periods (Fig. 2H).

Expression of the G34 core enrichment genes was particularly pronounced in the early fetal amygdala, inferior temporal cortex, and the caudal, medial, and lateral ganglionic eminences (Fig. 2H). Developmental expression patterns of G34 mutation-associated genes were in contrast to those observed with K27 mutation signatures derived from pediatric GBM specimens, which correlated with those of the embryonic upper rhombic lip, early-mid fetal thalamic, and cerebellar structures, and peaked during the mid-late fetal period (Supplementary Fig. S4).

Specifically, the G34 mutation drives expression of numerous highly developmentally regulated transcription factors, including as an exemplar DLX6 (distal-less homeobox 6), which encodes a homeobox transcription factor that plays a role in neuronal differentiation in the developing forebrain (7). The highly significant differential H3K36me3 and RNA polymerase II binding observed by ChiP-Seq (Fig. 3A) was validated by ChiP-quantitative real-time PCR (qPCR) (Fig. 3B), and expression of DLX6 was noted to be significantly higher in G34 pediatric GBM samples than K27-mutant or wild-type tumors in the integrated gene expression datasets at the mRNA level (Fig. 3C), and at the protein level in a tissue microarray comprising 46 pediatric and young adult GBM cases (Fig. 3D and Supplementary Table S4). Other similarly validated forebrain development-associated transcription factor genes included ARX (8), DLX5 (7), FOXA1 (9), NRZI (10), POU5F1 (11), and SOX2, which is required for stem cell maintenance (Fig. 3E–H).

Strikingly, the most significant differentially bound and expressed gene in our G34-mutant KNS42 cells was MYCN (33-fold H3K36me3 compared with SF188, DESeq P = 7.94 × 10⁻⁶; 60-fold RNA Pol II, DESeq P = 1.59 × 10⁻⁶; Fig. 4A–D). Of note, a small number of H3F3A wild-type tumors also expressed high levels of MYCN, and were found to be MYCN gene amplified (Fig. 4C). However, amplification was not seen in G34-mutant tumors, which parallels observations in diffuse intrinsic pontine glioma where MYCN amplification was found in wild-type, but not K27-mutant, tumors (15). Transduction of the G34V mutation into normal human astrocytes (NHA) and transformed human fetal glial cells (SVG) conferred an approximately 2- to 3-fold increase in MYCN transcript levels over wild-type–transduced controls, validating these observations (Supplementary Fig. S13). H3F3A G34 mutation may therefore represent an alternative mechanism of enhancing expression levels of MYCN in pediatric GBM.

Targeting MYCN is an attractive therapeutic intervention in tumors harboring gene mutation such as neuroblastoma (16), and direct inhibition by siRNA knockdown in KNS42 cells reduced cell viability in proportion to the reduction of protein levels observed (Fig. 4E). Pharmacologic agents that
Figure 2. Differential binding of H3K36me3 in G34-mutant KNS42 cells drives pediatric GBM expression signatures. A, Sanger sequencing trace for KNS42 pediatric GBM cells reveals a heterozygous c.104G>T (pGly34Val) H3F3A mutation. B, Western blot analysis for mono-(me1), di-(me2), and tri-(me3) methylated histone H3 in G34-mutant KNS42 and wild-type (wt) pediatric glioma cell lines. Total H3 is used as an extracted histone loading control. C, Circos plot representing histone H3 in G34-mutant KNS42 and wild-type pediatric glioma cell lines. Total H3 is used as an extracted histone loading control. D, ChIP-Seq correlation plot of RNA polymerase II versus H3K36me3 for 65 differentially trimethyl-bound regions by ChIP-Seq in G34-mutant KNS42 cells. R² = 0.66, P < 0.0001. E, Heatmap representing a ranked list of differentially bound H3K36me3 and RNA polymerase II in G34V KNS42 versus wild-type SF188 cells, with top 20 genes listed. F, GSEA for preranked differentially bound genes identified in ChIP-Seq versus those in the integrated gene expression datasets. G, DAVID gene ontology analysis for preranked list of differentially bound genes identified in ChIP-Seq. Fold enrichment of processes are plotted and colored by FDR q value. H, Top, mean expression of the G34 core enrichment signature in a temporal gene expression dataset of human brain development. Period 1, embryonal; periods 2-7, fetal; periods 8-12, postnatal; periods 13-15, adulthood. Bottom, heatmap representing spatial differences in G34 core enrichment signature expression in structures within embryonic and early fetal development, with highest levels mapping to the ganglionic eminences and amygdala.
Figure 3. G34 induces a transcriptional program linked to forebrain development and self-renewal. A, ChIP-Seq of H3K36me3 and RNA polymerase II binding for G34-mutant KNS42 (blue) and wild-type (wt) SF188 cells (gray) for the DLX6 locus, which also encompasses the transcripts DLX5, DLX6-AS1, and DLX6-AS2. B, validation of ChIP-Seq data by ChIP-qPCR using specific primers targeting DLX6. Blue bars, KNS42, gray, SF188.***, P < 0.0001, t test. C, boxplot of DLX6 expression in the integrated pediatric GBM samples stratified by H3F3A status. Blue box, G34; green, K27; gray, wild-type.***, P < 0.001, ANDVA. D, top, immunohistochemistry for DLX6 protein in a G34-mutant pediatric GBM sample RMH2465. Bottom, barplot of DLX6 expression in a pediatric GBM tissue microarray stratified by H3F3A status. Blue bars, G34; green, K27; gray, wild-type.++, strong expression; +, moderate expression; −, negative.

DISCUSSION

Emerging evidence strongly suggests that pediatric GBMs with H3F3A mutations can be subclassified into distinct entities. Our data indicate key molecular and clinical differences between G34- and K27-mutant tumors, reflecting the anatomic specificity (K27 tumors restricted to the pons and thalamus and G34 to the cerebral hemispheres; ref. 15; Supplementary Table S4) and likely distinct developmental origins of these disease subgroups. Using the only known model of H3F3A-mutant cells to date, we propose that the gene expression signature associated with G34 mutation in pediatric GBM patient samples is likely driven by a genomic differential binding of the transcriptionally activating H3K36me3 mark.

Mapping these gene expression signatures to publicly available datasets of human brain development shows a strong overlap with the ganglionic eminences of the embryonic and early fetal periods. These structures represent a transiently proliferating cell mass of the fetal subventricular zone, are the source of distinct neuroglial progenitors (20), and are therefore strong candidates for the location of the cells of origin of cerebral hemispheric G34-driven pediatric GBM. As with other pediatric brain tumors (21, 22),
mutation-driven subgroups of GBM retain gene expression signatures related to discrete cell populations from which these distinct tumors may arise. In addition, this mutation-driven differential H3K36me3 binding leads to a significant upregulation of numerous genes associated with cell fate decisions. Thus, we have identified a transcriptional readout specific to KNS42 cells and not in a panel of 4 wild-type pediatric glioma cell lines. The most significant and selective hits were for MYCN and AURKA in KNS42 cells treated with individual oligonucleotides directed against AURKA for 48 and 96 hours. GAPDH is used as a loading control. Right, barplot showing effects on cell viability of KNS42 cells exposed to 0.1, 0.5, and 2 to 5 μmol/L VX-689 (triangle). GAPDH is used as a loading control.

**METHODS**

Primary Pediatric Glioblastoma Expression Profiling

Expression data from the Schwartzentruber and colleagues (ref. 2; GSE34824) and Paugh and colleagues (ref. 3, GSE19578) studies were retrieved from the Gene Expression Omnibus (www.ncbi.nlm.nih.gov/geo) and analyzed in GenePattern using a signal-to-noise metric. GSEA was implemented for testing of enrichment of gene lists. Pediatric GBM expression signatures were mapped to specific developmental stages and anatomic locations using a spatiotemporal gene expression dataset of human brain development in Kang and colleagues (ref. 6; GSE25129).

Tissue Microarrays

Immunohistochemistry for DLX6 (NBPI-85929, Novus Biologicals), SOX2 (EPR3131, Epitomics), and MYCN (#9405, Cell Signaling) was carried out on tissue microarrays consisting of 46 cases of GBM.

Figure 4. G34 H3K36me3 upregulates MYCN which is selectively targetable by kinases that destabilize the protein. A, ChIP-Seq of H3K36me3 and RNA polymerase II binding for G34-mutant KNS42 (blue) and wild-type (wt) SF188 cells (gray) for the MYCN locus, which also encompasses the MYCNOS transcript. B, validation of ChIP-Seq data by ChIP-qPCR using specific primers targeting MYCN; Blue bars, KNS42; gray, SF188 ex vivo. P < 0.0001, t test. C, boxplot of MYCN expression in the integrated pediatric GBM samples stratified by H3F3A status. Blue box, blue, gray, wild-type, ***, P < 0.01. ANOVA. Wild-type tumors with high mRNA expression were frequently amplified (Amp). D, top, immunohistochemistry for MYCN protein in a G34-mutant pediatric GBM sample. RMH2465 bottom, barplot of MYCN expression in a pediatric GBM tissue microarray stratified by H3F3A status. E, effects on cell viability of MYCN knockdown in KNS42 cells. Western blot analysis showing efficiency of reduction of MYCN by 3 individual siRNAs targeting MYCN. Positive control. **, t test versus control. ***, P < 0.001, t test versus control.
pediatric and young adult GBM ascertained for H3F3A mutation by Sanger sequencing.

**Cell Line Analysis**

Pediatric GBM KNS42 cells were obtained from the JCRB (Japan Cancer Research Resources) cell bank. Pediatric SF188 cells were kindly provided by Dr. Daphne Haas-Kogan (University of California San Francisco, San Francisco, CA), and UW479, Res259, and Res186 were kindly provided by Dr. Michael Bobola (University of Washington, Seattle, WA). All cells have been extensively characterized previously (4), and were authenticated by short tandem repeat (STR) profiling. Western blot analysis was carried out for total histone H3 (ab97968, Abcam), as well as H3K36 trimethylation (ab9050, Abcam), dimethylation (ab9049, Abcam), and monomethylation (ab9050, Abcam), after histone extraction using a histone purification minikit (ActiveMotif), and quantitated by scanning on the Storm 860 Molecular Imager (GE Healthcare) and analyzed using ImageQuant software (GE Healthcare). Additional Western blots for MYCN (#9405, Cell Signaling), ATRX (sc-15408, Santa Cruz), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH; #2118, Cell Signaling) were carried out according to standard procedures.

**Chromatin Immunoprecipitation**

Chromatin immunoprecipitation (ChIP) was carried out employing antibodies against H3K36me3 and RNA polymerase II using the HistonePath and TranscriptionPath assays by ActiveMotif. Whole genome sequencing was carried out using an Illumina HiSeq2000 instrument with more than 30-fold coverage. Validation of active regions was carried out by ChIP-quantitative PCR (qPCR).

**siRNA Screening and Validation**

siRNA screening was carried out on a library of 714 human kinases using Dharmacon SMARTpools (Dharmacon), with cell viability estimated via a highly sensitive luminescent assay measuring cellular ATP levels (CellTiter-Glo; Promega). Z-scores were calculated using the median absolute deviation of all effects in each cell line. Individual ON-TARGETplus oligonucleotides for validation were obtained from Dharmacon and knockdown validated by Western blot analysis for AURKA (#4718, Cell Signaling), ATRX (sc-15408, Santa Cruz), and glyceraldehyde-3-phosphate dehydrogenase (GAPDH; #2118, Cell Signaling) were carried out according to standard procedures.

**Disclosure of Potential Conflicts of Interest**

L. Bjerke, Alan Mackay, M. Nandhabalan, A. Burford, A. Jury, S. Popov, D.A. Bax, D. Carvalho, K.R. Taylor, M. Vinci, I. Bajrami, C.J. Lord, D. Hargrave, P. Workman, and C. Jones are employees of The Institute of Cancer Research, which has a commercial interest in AURKA and CHK1 inhibitors. No potential conflicts of interest were disclosed by the other authors.

**Authors’ Contributions**

**Conception and design:** L. Bjerke, A. Mackay, M. Nandhabalan, D. Bax, D. Hargrave, P. Workman, C. Jones

**Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.):** L. Bjerke, A. Mackay, M. Nandhabalan, A. Burford, A. Jury, S. Popov, D.A. Bax, D. Carvalho, K.R. Taylor, M. Vinci, I. Bajrami, C.J. Lord, D. Hargrave, P. Workman, C. Jones

**Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis):** L. Bjerke, A. Mackay, M. Nandhabalan, A. Burford, A. Jury, S. Popov, D. Bax, K.R. Taylor, M. Vinci, I.M. McGonnell, D. Hargrave, A. Ashworth, P. Workman, C. Jones

**Writing, review, and/or revision of the manuscript:** L. Bjerke, A. Mackay, M. Nandhabalan, I.M. McGonnell, R.M. Reis, D. Hargrave, A. Ashworth, P. Workman, C. Jones

**Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases):** L. Bjerke, M. Nandhabalan, I. Bajrami

**Study supervision:** P. Workman, C. Jones

**Acknowledgments**

The authors acknowledge NHS funding to the National Institute for Health Research Biomedical Research Centre.

**Grant Support**

This work is supported by Cancer Research UK, the Wellcome Trust, the Samantha Dickson Brain Tumour Trust, and The Stravros Niarchos Foundation.

**REFERENCES**


Histone H3.3 Mutations Drive Pediatric Glioblastoma through Upregulation of MYCN

Lynn Bjerke, Alan Mackay, Meera Nandhabalan, et al.

Cancer Discovery  Published OnlineFirst March 28, 2013.