Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Cancer Cell. Author manuscript; available in PMC 2018 October 09.

Published in final edited form as:

Cancer Cell. 2017 October 09; 32(4): 411–426.e11. doi:10.1016/j.ccell.2017.08.018.

Coordinated Splicing of Regulatory Detained Introns within Oncogenic Transcripts Creates an Exploitable Vulnerability in Malignant Glioma

Christian J. Braun[1,8], **Monica Stanciu**[1,8], **Paul L. Boutz**[1,8], **Jesse C. Patterson**[1], **David Calligaris**[2], **Fumi Higuchi**[3], **Rachit Neupane**[1], **Silvia Fenoglio**[1], **Daniel P. Cahill**[3], **Hiroaki Wakimoto**[3], **Nathalie Y. R. Agar**[2,4,5], **Michael B. Yaffe**[1,6,7], **Phillip A. Sharp**[1], **Michael T. Hemann**[1,9], and **Jacqueline A. Lees**[1,9,10]

[1]The David H. Koch Institute for Integrative Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2]Department of Neurosurgery, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA

[3]Department of Neurosurgery, Massachusetts General Hospital, Boston, MA 02144, USA

[4]Department of Cancer Biology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02215, USA

[5]Department of Radiology, Harvard Medical School, Boston, MA 02115

[6]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[7]Division of Acute Care Surgery, Trauma, and Critical Care, Department of Surgery, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, USA

Summary

Glioblastoma (GBM) is a devastating malignancy with few therapeutic options. We identify PRMT5 in an *in vivo* GBM shRNA screen and show that PRMT5 knockdown or inhibition potently suppresses *in vivo* GBM tumors, including patient-derived xenografts. Pathway analysis implicates splicing in cellular PRMT5 dependency, and we identify a biomarker that predicts

Contact information: Michael T. Hemann and Jacqueline A. Lees, David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, hemann@mit.edu and jalees@mit.edu.

[8]These authors contributed equally

[9]These authors are co-corresponding

[10]Lead contact.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Author contributions:

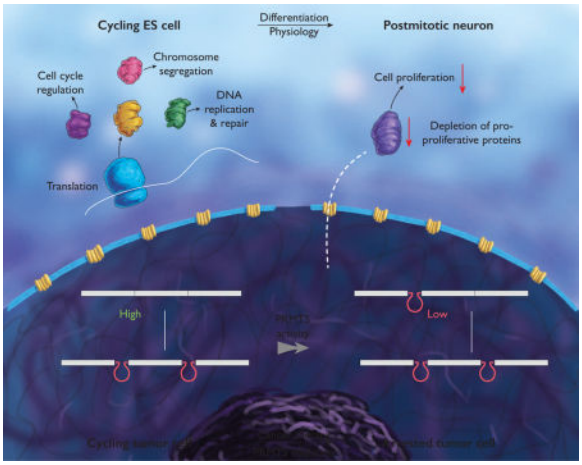
Conceptualization, C.J.B., M.S., M.H. and J.A.L.; Methodology, C.J.B., M.S., P.L.B., J.P., D.C., F.H., S. F., D.P.C., H.W., N.Y.R.A. and M.B.Y.; Software: P.L.B.; Formal Analysis: P.L.B. and J.P.; Investigation: C.J.B., M.S., P.L.B., D.C., F.H. and R.N.; Resources: F.H., D.P.C. and H.W.; Data curation: P.L.B.; Writing - original draft: C.J.B., M.S., P.L.B., M.T.H. and J.A.L; Writing -review and editing: C.J.B., M.S., P.L.B., P.A.S., M.T.H. and J.A.L; Supervision: D.P.C, H.W., N.Y.R.A., M.B.Y., P.A.S., M.T.H. and J.A.L; Funding Acquisition: P.A.S., M.T.H., J.A.L.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

sensitivity to PRMT5 inhibition. We find that PRMT5 deficiency primarily disrupts the removal of detained introns (DIs). This impaired DI-splicing affects proliferation genes, whose down-regulation coincides with cell cycle defects, senescence and/or apoptosis. We further show that DI-programs are evolutionarily conserved and operate during neurogenesis, suggesting that they represent a physiological regulatory mechanism. Collectively, these findings reveal a PRMT5-regulated DI splicing program as an exploitable cancer vulnerability.

Graphical Abstract

Braun et al. show that glioblastoma is selectively sensitive to the inhibition of PRMT5 and identify a predictive biomarker for this sensitivity. PRMT5 inhibition primarily disrupts the removal of detained introns, which results in the reduction of functional transcripts of mainly proliferation-associated genes.



Keywords

Splicing addiction; GBM; PRMT5; EPZ015666; biomarker

Introduction

Glioblastoma (GBM) is a highly malignant brain tumor with a grim clinical prognosis despite aggressive and multimodal treatment regimens (Stupp et al., 2005). With few exceptions, there is a dearth of targeted therapeutics to treat this disease (Lau et al., 2014). Moreover, there is a lack of well-defined GBM dependencies that represent possible therapeutic targets. *In vivo* loss of function screens offer an unbiased method to identify cellular processes that represent key vulnerabilities for cancer cells. These screens are particularly useful in identifying cancer dependencies that are not induced by mutation and, consequently, not revealed by tumor genome sequencing efforts (Gargiulo et al., 2014). However, high-throughput genetic approaches are challenging in GBM where the heterogeneity and consequent differential growth rates of patient derived xenograft (PDX) tumors causes high experimental noise.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Recent publications have suggested that aberrant RNA splicing is essential to the development and progression of certain malignancies (Dvinge et al., 2016). A putative regulator of this process is the arginine methyltransferase PRMT5. PRMT5 activity accounts for most of the symmetric dimethylation of arginine residues (SDMA) in mammalian cells (Stopa et al., 2015) with substrates acting in a wide variety of biological processes. For example, PRMT5 regulates transcription by targeting histones, nucleosome remodeling and co-repressor complexes, and numerous transcription factors. PRMT5 is thought to regulate splicing largely via its role as the enzymatic component of the methylosome, a multi-subunit complex that also contains PRMT5's obligate binding partner, MEP50, and the co-factor pICln (encoded by *CLNS1A* and herein referred to as such). The methylosome modifies specific Sm-proteins to facilitate small ribonucleoprotein (snRNP) assembly (Battle et al., 2006; Chari et al., 2008). Consistent with this role, complete *Prmt5* knockout (KO) is lethal to murine neural stem/progenitor cells and was reported to impair splicing (Bezzi et al., 2013). PRMT5 has recently emerged as a promising drug target, due to its frequent over-expression in a variety of malignancies (Stopa et al., 2015) as well as its synthetic lethal relationship with methylthioadenosine phosphorylase (*MTAP*) (Kryukov et al., 2016; Marjon et al., 2016; Mavrakis et al., 2016). Although PRMT5's potential as a cancer therapeutic target is widely appreciated, there is still considerable debate about which of its numerous activities underlie the PRMT5 dependency of tumor cells.

The realization of splicing as a targetable process in cancer will require answering important questions: if splicing is an essential process in all cells, is there a therapeutic window in which splicing inhibitors would specifically target cancer cells? Alternatively, are there targetable forms of RNA splicing that are preferentially relevant to cancer cells? The complexity of the problem is intensified by the ongoing discovery of new splicing modalities such as detained introns (DIs). DIs occur within otherwise completely spliced and polyadenylated transcripts but, in contrast to retained intron-containing transcripts, DI-containing pre-mRNAs remain in the nucleus (Boutz et al., 2015), where they are eventually post-transcriptionally spliced or degraded (Bresson et al., 2015; Yap et al., 2012). Notably, an increase in transcripts entering the DI pathway will result in reduced levels of productive coding mRNA. Intron retention has previously been observed to differentiate tumors from matched normal tissue (Dvinge and Bradley, 2015; Jacob and Smith, 2017).While there is extensive data showing that splicing processes are altered in cancer cells, little is known about how to target these processes in cancer patients.

In this study, we sought to develop a high-throughput screening approach to identify key GBM dependencies. We then aimed to use this information to elucidate the nature of these dependencies, to explore their druggability with inhibitors, to identify genetic factors influencing inhibitor sensitivity, and to define their role in normal brain development and biology.

Results

A pooled *in vivo* shRNA screen identifies PRMT5 as a key mediator of GBM growth

To establish the suitability of Gl261 for *in vivo* pooled shRNA screens, we generated mixed Gl261 populations in which a fraction of cells expressed Doxycycline (Dox) inducible

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

shRNA vectors targeting either KRAS, the driving oncogene of Gl261 cells (Newcomb and Zagzag, 2009), or Renilla luciferase (RLuc). Importantly, whereas transduced but uninduced cells only express the fluorophore Venus, upon Dox induction cells become positive for both Venus and dsRed, allowing assessment of both transduction and induction levels by flow cytometry. These cells were then injected into mouse brains to allow tumor engraftment, and shRNA expression was induced by oral Dox administration. Following mouse euthanasia and tumor dissociation, we assayed single-cell suspensions by flow cytometry to determine the proportion of transduced Venus$^+$ cells expressing the shRNAs (Venus$^+$ dsRed$^+$, Fig 1A). The results were highly reproducible between replicates and showed a strong selection against cells expressing KRAS shRNAs, compared to consistently high induction levels of control shRNAs (Fig 1B, S1A,B). The observed selective pressure against KRAS shRNA-expressing cells was sufficient to increase median survival (Fig 1C). We also tested the feasibility of positive selection by overexpressing O^6-methylguanine methyltransferase (MGMT), which confers resistance to Temozolomide (TMZ) (Hegi et al., 2005). We created mixed Gl261 populations in which a fraction of the cells overexpressed MGMT and found that the percentage of cells expressing MGMT was increased following TMZ exposure when compared to vehicle controls (Fig S1C,D). Thus, we can successfully score both negative and positive selection.

Another requirement for *in vivo* pooled screens is that neutral shRNAs are not subject to stochastic loss of library complexity or clonal outgrowth. To determine the library size that fulfilled these requirements, we performed limiting dilution experiments in which progressively smaller fractions of the Gl261 cells were labeled with GFP. We compared GFP$^+$ cell representation before and after formation of intracranial tumors and found that 0.2% GFP$^+$ cells were detected in all end stage tumors but 0.02% GFP$^+$ cells were not (Fig S1E), placing the maximal feasible *in vivo* library size between 500 and 5000 shRNAs (Fig S1F).

The contributions of epigenetic regulators to GBM growth and therapeutic response are becoming increasingly appreciated (Rodríguez-Paredes and Esteller, 2011). Thus, we conducted our screen using a pooled shRNA library that targeted 312 epigenetic regulators. We included two validated KRAS shRNAs to serve as positive controls, and the pool was transduced into cultured Gl261 cells. A fraction of the infected cells was kept as the input control, and the remainder cultured *in vitro* or immediately used for intracranial injection into 28 C57BL/6J recipients. Both *in vitro* and *in vivo* samples were divided into two cohorts and treated with TMZ or vehicle control. Mice were sacrificed at morbidity, and those closest to the median survival (n=4 for vehicle and n=5 for TMZ) were used to generate tumor cell DNA. The *in vitro* samples were harvested at the corresponding median survival times. We amplified the integrated shRNA loci and identified shRNA distributions by next generation sequencing (Table S1). Our shRNA enrichment and quantification protocol did not introduce significant amounts of stochastic variation (Fig S1G). Moreover, the majority of our library was preserved in the *in vivo* samples, suggesting that we successfully avoided bottlenecking (Fig 1D, S1H). Additionally, we found that our pre-validated, positive control KRAS shRNAs were selectively depleted (Fig S1I–N). We then scored the total library shRNA changes (Fig S1I–N, Table S2). The majority of shRNAs targeting PRMT5 were depleted under all four conditions. By mathematically combining the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

behavior of all shRNAs for each gene, we generated a gene level score and found that PRMT5 was the most significant hit (Fig 1E, S1O).

The enzymatic function of PRMT5 is essential for the proliferative fitness of GBM cells

To validate the importance of PRMT5 for tumor cell proliferation and/or viability, we conducted *in vitro* single-shRNA competition experiments in a set of human GBM cell lines using Dox-induced expression of either a RLuc control or three different PRMT5 shRNAs. The latter significantly reduced PRMT5 mRNA and protein levels (Fig S2A, B). PRMT5 depletion was rapidly selected against, as indicated by the outgrowth of shRNA-negative cells after Dox addition (Fig 2A–C, S2C). To confirm that these phenotypes were PRMT5-dependent, and not shRNA off-target effects, we overexpressed a shRNA-resistant PRMT5 cDNA. This restored PRMT5 expression in the PRMT5 knockdown cells (Fig 2D,E). It also fully restored symmetrically dimethylated histone 4 arginine 3 (H4R3me2s), a known PRMT5-mediated histone mark (Yang and Bedford, 2012) (Fig 2E, S2D). PRMT5 cDNA expression did not further increase H4R3me2s levels above the wild-type background suggesting that endogenous PRMT5 levels are saturating. Most importantly, the PRMT5 cDNA completely abolished the depletion of PRMT5 shRNA-expressing cells in our competition assay (Fig 2E). We also expressed a mutant PRMT5 lacking two residues (Y304 & Y307) critical for substrate interaction (Antonysamy et al., 2012). This mutant restored PRMT5 protein levels in the shRNA expressing cells, but did not rescue either loss of H4R3me2 dimethylation or competitive depletion of the cells (Fig 2F, S2E). Thus, PRMT5's enzymatic function is required for the persistence of GBM growth post-engraftment.

Inhibition of PRMT5 by EPZ015666 induces cell cycle defects and senescence

To further test the importance of PRMT5 for GBM growth, we used the PRMT5 inhibitor EPZ015666 (EPZ; Chan-Penebre et al., 2015). Treatment of U-87 MG cells with 10 μM EPZ yielded a dramatic reduction of SDMA marks (Fig 3A). EPZ treatment significantly suppressed the proliferation of several human GBM cell lines (Fig 3B). Studies in U-87 MG cells showed that this reflected loss of active S phase cells and accumulation of G_2/M populations (Fig 3C). EPZ did not induce apoptosis (judged by the absence of sub-G_1 cells, Fig 3C), even though U-87 MG cells possess wild-type *TP53*. Instead we saw induction of various senescence regulators (Fig S3) and most of the cells stained positive for senescence-associated (SA) β Gal (Fig 3D). Thus, PRMT5 inhibition results in potent suppression of glioma cell proliferation in *vitro*.

Genetic and pharmacological inhibition of PRMT5 impedes brain cancer growth *in vivo*

We next tested the validity of PRMT5 as a therapeutic target *in vivo*. First, we constitutively expressed either a control RLuc shRNA or two independent PRMT5 shRNAs, which both yield partial PRMT5 mRNA knockdown (Fig S4A). Both PRMT5 shRNAs conferred a significant extension of lifespan to mice bearing intracranial Gl261 tumors (Fig 4A). Since the PRMT5 shRNAs were constitutively expressed, the survival benefit of PRMT5 knockdown could result from impaired tumor cell engraftment or suppressed tumor growth. To directly test the latter mechanism, we transduced U-87 MG cells with viruses allowing Dox-inducible expression of control or PRMT5 shRNAs and injected these intracranially into immunocompromised mice, activating shRNA expression 5 days later. The shRNAs

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

were not expressed in the absence of Dox, based on the absence of dsRed signal (Fig 4B). In the presence of Dox, we observed significant selection against expression of the PRMT5 shRNA in end-stage tumors, in contrast to high induction rates of the control shRNA (Fig 4B). Thus, PRMT5 deficiency is highly deleterious to GBM tumors.

Next we then assessed the *in vivo* impact of EPZ. Twice-daily administration of EPZ by oral gavage for seven days did not cause weight loss, increased liver serum transaminase levels or decreased peripheral blood cell counts (Fig 4C–D). Having ruled out toxicity, we then tested oral EPZ treatment on tumors. First, we subcutaneously injected U-87 MG cells into nude mice, allowed tumor cell engraftment, and treated the mice with EPZ or vehicle for two one week periods. Longitudinal tumor monitoring showed a significantly lower tumor volume in treated, versus control, animals post-treatment (Fig 4E). Moreover, after 10 days of treatment, these tumors showed significant downregulation of KI67 and upregulation of Cleaved Caspase 3, as well as profound tumor fibrosis (Fig S4B–C). We next tested EPZ's effect on intracranial U-87 MG tumors and found that EPZ yielded a significant lifespan extension (Fig 4F). However, EPZ seemed to yield less impressive benefits for U-87 MG intracranial tumors, than for U-87 MG subcutaneous tumors. Moreover, EPZ provided no detectable survival benefit for intracranial Gl261 tumors (Fig S4D), in stark contrast to the anti-tumorigenic effects of PRMT5 shRNAs in the same model (Fig 4A). We therefore hypothesized that the blood brain barrier (BBB) limits EPZ's ability to access intracranial tumors. We used an in-silico algorithm to predict BBB permeability, and this analysis demarcated EPZ as a relatively poor BBB penetrator (Fig S4E). We then experimentally tested EPZ's BBB penetration efficiency by treating animals bearing intracranial U-87 MG tumors with EPZ at various times post-implantation and quantifying EPZ uptake by matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-MSI) of frozen brain sections (Fig 4G). EPZ was not detected within the normal brain tissues and only accumulated in tumors that were highly advanced (Fig 4G, S4F–I), where it successfully achieved significant inhibition of PRMT5 function (judged by loss of H4R3me2s; Fig S4J). Drug uptake was not simply proportional to tumor burden, but showed a dramatic increase between 14 and 21 days post injection (Fig 4H). The likely cause of this late stage-specific uptake is a breach of the BBB, due to development of leaky tumor vasculature upon neoangiogenesis. Accordingly, Evans Blue, a dye that is sequestered in the vasculature if the BBB remains intact (Leten et al., 2014), only accessed intracranial U-87 MG tumors at late stages (>21 days in this experiment; Fig 4I). Taken together, these data show that the BBB blocks uptake of EPZ, but the drug is able to access late stage tumors with a compromised BBB. Importantly, even short exposure to EPZ, late in tumor progression, is sufficient to achieve profound suppression of SDMA markers and significantly extend lifespan.

PDX models have become the gold standard for compound testing in oncology (Hidalgo et al., 2014). Thus, we also examined the effect of EPZ on the growth of subcutaneous transplants of MGG8, a GBM PDX (Wakimoto et al., 2012). Continuous treatment with EPZ for 12 days, starting on day 9 after tumor cell transplantation, yielded cellular senescence (Fig S4K) and profound tumor suppression (Fig 4J). We therefore conclude that PRMT5 is a promising targetable vulnerability in GBM, and there is strong rationale for developing next generation PRMT5 inhibitors with increased BBB penetration.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

RNA processing and splicing pathways emerge as potential resistance mechanisms to PRMT5 inhibition

PRMT5 is known to influence a wide variety of biological processes. However, there is still considerable debate as to which of these underlie the PRMT5 dependency of tumor cells. We used two unbiased and complementary approaches to address this key question. First, we compared the responses of various cell lines to EPZ and then identified gene expression patterns that correlated with sensitivity or resistance. For this, we treated 25 tumor cell lines from various tissue origins in triplicate with 11 different EPZ concentrations (ranging from 0.5 μM to 500 μM) and assayed response 120 hr post-treatment (Fig 5A, S5A). We calculated absolute half-maximal inhibitory concentrations (IC_{50}), maximum drug effect (E_{max}) and drug activity scores to capture the efficacy and potency of the drug (Fig 5B). The highly reproducible responses identified both sensitive and resistant cell lines (Fig 5C), indicating that EPZ is not universally toxic. To identify the causative mechanism(s) for cellular drug resistance, we began by probing obvious candidates. First, we ruled out that drug-efflux pumps or drug metabolism were simply blocking PRMT5 inhibition by demonstrating that EPZ successfully reduced SDMA marks in highly resistant cell lines (Fig S5B). Second, we examined *MTAP*-status, because *MTAP* deletion is synthetic lethal with PRMT5 knockdown, although not with PRMT5 inhibition by EPZ (Kryukov et al., 2016; Marjon et al., 2016; Mavrakis et al., 2016). Consistent with this latter finding, we saw no correlation between *MTAP*-status and EPZ sensitivity (Fig S5C). Third, we established that neither *TP53* status, a common modulator of chemotherapeutic response, nor cell division rate were predictive of response (Fig S5C, D). Having ruled out these candidates, we undertook a systematic approach by examining transcriptome data available for 18 of our tested lines. We calculated Pearson correlation coefficients between the measured drug activity areas and individual gene expression levels across all 18 lines (Fig 5D, Table S3), and subjected the resulting rank ordered gene list to gene set enrichment analysis (GSEA). Gene sets associated with RNA-biology, RNA processing and splicing stood out in this analysis. These gene sets were strongly enriched at the bottom end of the ranked list (Fig 5E), suggesting that high expression of genes involved in RNA processing/splicing pathways correlates with resistance to PRMT5 inhibition.

We sought to identify a diagnostic gene signature that could predict relative responsiveness to EPZ. The splicing gene sets most strongly enriched by GSEA included snRNP components that are well-known targets of the PRMT5 methylosome. Given this finding, we created a curated list of 59 genes that included those encoding core components of the spliceosomal complex, the SMN complex, the PRMT5 methylosome, and co-factors that determine PRMT5 substrate specificity. The expression levels of some of these individual genes showed a reasonable correlation or anti-correlation with EPZ response (data not shown). However, we hypothesized that a combination of two gene expression profiles might better predict sensitivity, particularly if the gene products functionally interact. Thus, we calculated pairwise expression comparisons for all 59 genes to yield 1711 expression profiles, which were used to calculate Pearson significance p values for their correlation with drug activity scores across our 18 cell lines. 301 of these ratios outperformed the predictive value of either of the contributing single genes (Fig 5F). Amongst these, one of the top hits was the CLNS1A to RIOK1 (C/R) ratio, where high C/R correlated with high

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

EPZ sensitivity (Fig 5G and H). The C/R ratio had higher predictive value than CLNS1A (r = 0.04, p=0.87) or RIOK1 (r = −0.43, p=0.075) alone, and even PRMT5 itself (r = −0.45, p=0.06). Notably, CLNS1A and RIOK1 are mutually-exclusive PRMT5 co-factors that direct PRMT5's activity towards snRNP or ribosomal biogenesis, respectively (Guderian et al., 2011). The correlation between a high C/R ratio and increased EPZ sensitivity strongly suggests that sensitive tumor cells have limiting PRMT5 activity, which is preferentially directed towards the spliceosome.

To directly test the diagnostic value of the C/R ratio, we used gene expression data for previously untested cell lines to calculate their C/R ratios and thus predict their responsiveness (Fig S5E). We selected 11 lines and experimentally determined their EPZ activity scores, revealing a strong correlation with these predictions (Fig 5I). Furthermore, knockdown experiments in two different cell lines showed that these factors both have determinative roles. Specifically, CLNS1A knockdown (Fig S5F) increased sensitivity to EPZ (Fig 5J and S5G), suggesting that CLNS1A downregulation causes less PRMT5 activity to be directed towards spliceosome biogenesis, thereby reducing splicing capacity (Fig 5H). Accordingly, RIOK1 knockdown (Fig S5F) decreased EPZ sensitivity (Fig 5J and S5G), presumably by shifting PRMT5 activity towards spliceosome biogenesis. Collectively, our data suggest that EPZ-sensitive cells have lower expression levels of splicing-associated genes, and maintain a higher C/R ratio to direct limiting PRMT5 activity towards spliceosomal functions. Essentially, these cells appear to be addicted to splicing.

Pharmacological PRMT5 inhibition primarily reduces the expression of productive gene isoforms by promoting intron detention

As a parallel approach to determine how PRMT5 inhibition achieves tumor suppression, we compared the gene expression patterns of U-87 MG cells treated with EPZ or vehicle for 72 hr. Using standard approaches to detect changes in the total levels of gene transcripts irrespective of splicing variants (counting reads falling within exons), we found that many expressed genes exhibited significant [Posterior Probability of Differential Expression (PPDE) > 0.95; FDR-adjusted p value < 0.05] downregulation (11.5%) or upregulation (10.1%) in response to EPZ (Fig 6A, S6A–C). We then screened for EPZ-associated changes in splicing patterns. Unexpectedly, PRMT5 inhibition had little influence on the relative representation of canonical alternative splicing (AS) isoforms (Fig 6B, S6D; significant events shown in color).

In stark contrast to other AS events, we found that the splicing of detained introns (DIs) was profoundly perturbed by PRMT5 inhibition. *In toto*, we identified 3801 DIs distributed across 2408 individual genes in our samples. Of these, 1619 (43%) reached statistical significance (FDR-adjusted p value <0.05) for differential splicing following EPZ treatment. Importantly, the vast majority of these DIs (1557, or 96%) were significantly upregulated following PRMT5 inhibition (Fig 6B, S6D). Thus, the dominant effect of PRMT5 inhibition on splicing is to decrease the splicing efficiency of many DIs (see Fig S6E, F).

Transcripts with DIs are trapped in the nucleus and thus remain untranslated until their eventual decay (Boutz et al., 2015). Consequently, increased levels of DI-containing transcripts result in reduced levels of productive isoforms and thus encoded proteins. Despite

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

this fact, exon reads from DI-containing transcripts are typically included in the measured transcript pool of a gene. To assess the level of transcripts productive for protein synthesis for each gene, we corrected our data by removing DI-isoforms from the gene expression calculation (Fig 6C). This identified 750 genes whose coding isoforms were significantly downregulated by EPZ (PPDE< 0.95; Fig 6D, left pie chart). A less stringent cutoff (DI-containing genes with a consensus coding isoform drop ≥ 10%, Fig S6G) identified an additional 597 genes that were downregulated at the coding isoform level (Fig 6D, right pie chart; Fig S6H). This raised the total number of downregulated DI-containing genes to 1347, representing more than 55% of all DI-containing genes. In contrast, total transcript quantification without correction for isoform composition resulted in only 171 of the 2408 DI-containing genes meeting the statistical cutoff for up- or down-regulation upon PRMT5 inhibition (Fig S6H, leftmost pie chart). This underscores the importance of considering the contributions of productive isoforms for gene expression quantification, and unveils a profound effect of DI regulation on productive mRNA levels. The changes observed with the mitotic spindle regulator kinase AURKB illustrate this point: EPZ decreased the total AURKB transcript levels by 30%, but after removing non-productive DI-containing transcripts, the productive coding isoform was actually decreased by more than 80% (Fig 6E, F).

To establish the robustness of this PRMT5-DI connection, we assayed additional representative genes for DI-inclusion, and also transcript and protein levels. qRT-PCR analyses of both AURKB and the translation initiation factor EIF4E confirmed that the levels of unproductive DI-containing transcripts were increased relative to the total transcript pool in response to either pharmacological or genetic PRMT5 deficiency (Fig S6I–L). Importantly, this resulted in reduced levels of both AURKB and EIF4E proteins (Fig 6G, Fig S6I–K). We also detected an increased ratio of DI-containing transcripts to total transcripts, and/or a reduction in protein levels for RAF1, TRIO, DNA2, RPS6KB1, FBXO18, SOS1, and STAT5B (Fig S6M, N). Additionally, we observed an increased ratio of DI-transcript to total transcript for AURKB, EIF4E and RAF1 *in vivo* following EPZ treatment of PDX tumors (Fig S6O). Notably, single-gene *in vitro* competition assays showed selection against cells harboring knockdown of AURKB or another DI-regulated gene, FGFR1 (Fig S6P). Thus, for two different genes within our DI-regulated set, depletion of even a single gene is sufficient to confer a proliferative disadvantage to GBM cells.

A prior study reported that genetic KO of *Prmt5* in murine neuronal progenitor cells (NPCs) resulted in several hundred AS changes and also noted a general accumulation of reads within introns without further investigation of their identity (Bezzi et al., 2013). This raised the question of whether genetic loss of *Prmt5* also perturbed DI splicing. To address this, we quantified significant splicing differences affecting canonical AS events versus DIs within the murine NPC *Prmt5* KO data set. The results of this analysis were consistent with EPZ-treated human GBM cells. While ~200 canonical AS events were significantly altered in the KO cells compared to wild-type controls, these affected events represented only a small percentage (1–5%) of all AS variants detected in these cells. In contrast, and again consistent with the effects of the PRMT5 inhibitor, DI splicing was profoundly disrupted in the *Prmt5* KO cells with 1755 DIs (or 38.4%) being affected (Fig S6Q). Taken together, our analyses show a predominant upregulation of DIs *in vitro* and *in vivo* in response to depletion of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

PRMT5 via pharmacological inhibition, siRNA-knockdown or genetic ablation that is conserved across different species.

The PRMT5-DI axis affects genetic regulators of proliferation

DIs are frequently phylogenetically conserved and exhibit cell-type and context-dependent changes in splicing efficiency, suggesting that they are regulated events rather than splicing errors (Boutz et al., 2015). Indeed, many DIs affected by PRMT5 depletion are already observed at significant levels in untreated human U-87 MG and wild-type murine NPC cells, indicating that PRMT5 inhibition acts to increase levels of existing DIs rather than causing *de novo* splicing changes. These findings raised the possibility that DI-containing genes are regulated as part of a coherent biological program. To address this question, we identified the functional categories of DI-containing genes within our complete data set from both control and drug-treated cells. This analysis revealed a highly significant enrichment for genes involved in cell cycle progression, chromosome segregation, RNA biology and DNA repair (Fig 6H). Moreover, DI-regulated genes in the murine NPC data set were similarly enriched in proliferation-associated gene sets (Fig S6Q). This strongly suggests that DIs play a role in the coordinate regulation of specific functional gene expression programs, which is conserved across species.

It is important to note that PRMT5 inhibition also alters the expression levels of non-DI containing genes. To gain further insight into the broader functional consequences of PRMT5 inhibition, we also performed pathway enrichment analysis for all genes downregulated by EPZ treatment of U-87 MG cells i.e., both DI-containing genes, after correcting for isoform-specific effects, and non-DI containing genes (Fig 6I). We found gene sets associated with cell cycle, DNA replication, and chromosomal segregation to be significantly downregulated upon PRMT5 inhibition, suggesting that PRMT5 inhibition affects DI-containing and non DI-containing genes that function in a coherent gene expression program strongly associated with cell cycle regulation (Fig S6R). *Prmt5* KO in murine NPCs yielded similar results (Fig S6Q). Moreover, qRT-PCR analysis of 18 different proliferation genes confirmed coordinated downregulation in response to EPZ treatment (Fig S6C). In contrast to the downregulated genes, few upregulated genes were affected by DI inclusion, and therefore the vast majority of gene expression changes (93%) were identical whether we corrected for productive isoforms or not (Fig S6S). Pathway enrichment analysis of upregulated genes identified programs associated with cell and organ development, not proliferation (Fig S6S). Taken together, our data indicate that depleting PRMT5 activity causes dramatic down-regulation of cell cycle regulators through two distinct mechanisms: it promotes intron detention and thereby reduces the levels of the productive coding isoforms for a large fraction of DI-containing genes and it reduces the mRNA levels of many other genes either indirectly, or through a DI-independent mechanism. The overwhelming impact on cell cycle genes is consistent with the anti-proliferative effects of PRMT5 suppression.

Weak 5′ splice sites are a frequent feature of PRMT5-dependent introns but are insufficient to determine selectivity

Our data indicate that PRMT5 inhibition has significant selectivity for a subset of DIs, rather than affecting splicing in general. This raises questions about the underlying mechanistic

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

bases for this specificity. Thus, we asked whether DIs that are upregulated upon EPZ treatment exhibit genomic features that distinguish them from DIs with constant or decreased levels. We examined 5′ and 3′ splice sites because these tend to be weaker for DIs than other introns (Boutz, et al. 2015), and weak 5′ splice sites have previously been reported to be associated with PRMT5-dependent splicing changes (Bezzi et al., 2013; Sanchez et al., 2010). On a population basis, DIs upregulated by either PRMT5 inhibition or *Prmt5* KO were found to have significantly weaker mean 5′ splice site scores than DIs that were not upregulated (Fig 7A). Despite this general tendency, there was not an absolute correlation between 5′ splice site strength and sensitivity to PRMT5 deficiency. Specifically, some individual DIs possessed weak 5′ splice site scores without being upregulated by PRMT5 deficiency, and some of the PRMT5-sensitive DIs had stronger 5′ splice sites than unaffected DIs (Fig 7A, S7A). Additionally, canonical AS events such as cassette exons were rarely affected by PRMT5 deficiency even though they often had similar 5′ splice site scores to PRMT5-affected DIs (Fig S7B). Thus, despite being a frequent feature of PRMT5-dependent DIs, weak 5′ splice sites scores are neither necessary nor sufficient to define PRMT5 dependency of individual introns. We looked more specifically at the sequence features of the 5′ splice sites, asking whether the co-occurrences of specific nucleotides were enriched or depleted in DIs that are increased by PRMT5 deficiency versus DIs that are not (Fig 7B, C and S7C). The most striking difference was an underrepresentation of 5′ splice splices bearing Gs at both the −1 and +5 positions. Notably, these sequence features are conserved between PRMT5-affected DIs in human GBM cells and murine NPCs suggesting that they are under positive selective pressure. We also considered 3′ splice site strength, and found that on a population basis this was either not significantly different (for *Prmt5* KO), or modestly, but significantly, higher (for PRMT5 inhibition) for PRMT5-depedent DIs (Fig 7A). This suggests that 3′ splice site strength is not a major determinant of PRMT5 sensitivity, but does not rule out the possibility that weak 3′ splice sites may contribute to the sensitivity of individual DIs, and/or that other genomic features, such as the branchpoint region or flanking exonic or intronic splicing enhancers that lack consensus sequences and are therefore difficult to define computationally, might also play key roles.

PRMT5 is involved in the assembly of snRNP complexes (Chari et al., 2008). Moreover, the underrepresentation of Gs at positions −1 and +5 of the 5′ splice site might be expected to weaken basepairing with the snRNAs, which is critical for spliceosome assembly and catalysis. Given these issues, we hypothesized that PRMT5 sensitivity might result from a reduction in the biogenesis or steady-state levels of snRNPs. Consistent with the known role of PRMT5, we found that EPZ treatment caused rapid loss of SDMA marks of the Sm proteins SNRPB and SNRPD3 (Fig S7D, E). However, this had no detectable impact on either the levels or nuclear/cytoplasmic localization of spliceosomal snRNAs (Fig S7F). Thus, the ability of PRMT5 deficiency to impair the splicing of DIs is not a simple consequence of impaired biogenesis of individual snRNP complexes.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

PRMT5-associated intron detention represents a regulated biological process engaged during cellular differentiation

Our finding that PRMT5 deficiency primarily affects DIs that are enriched within proliferation regulators in both mouse and human cells argues for the existence of a coordinated and regulated splicing program under the control of PRMT5. We hypothesized that such a program might be engaged during terminal differentiation. To test this, we turned to a published RNA-seq data set tracking *in vitro* neurogenesis from mouse embryonic stem cells to >95% mature glutamatergic neurons (Hubbard et al., 2013). We analyzed PRMT5 mRNA levels and found that these declined progressively during the neuronal differentiation process (Fig S8A). We then tracked genome-wide changes in the ratios of coding isoforms to DI-containing isoforms at each time point and identified four distinct clusters (Fig 8, Table S6). Cluster 4 showed increased intron detention and consequent reduction of productive coding transcripts as the cells proceeded through the differentiation process. Importantly, this gene cluster exhibited a strong functional enrichment for cell cycle regulators. The remaining three gene clusters displayed significant enrichments for gene sets that were unrelated to cell cycle or proliferation (Fig 8, S8B). Intriguingly, cluster 3 showed a reduction in the level of DIs and consequent increases in the levels of coding isoforms as the cells moved towards differentiation and the affected genes were strongly associated with neuronal identity (such as axonal and synaptic genes).

We then directly compared the DI-containing genes in this differentiation data set with those identified in our analysis of the *Prmt5* KO in murine NPCs. In particular, we focused on the group of DIs that were significantly upregulated by *Prmt5* loss in NPCs (herein referred to as PRMT5-regulated DIs) and asked how these were distributed across the four differentiation clusters. We found that PRMT5-regulated DIs mapped to all four clusters, but only exhibited statistically significant enrichment in the proliferation-associated Cluster 4.

Collectively, these data suggest that the splicing of sets of DIs is differentially regulated during neuronal differentiation, thereby cooperating with transcriptional changes to coordinate cell cycle exit and terminal differentiation. Specifically, steady reduction in PRMT5 levels from ES cells to mature neurons is mirrored by suppression of proliferation genes via upregulation of their constituent DIs. In concert, DI levels are reduced in genes involved in the neuronal differentiation program, enabling their expression. We hypothesize that the strong upregulation of PRMT5 mRNA levels in malignant gliomas (Fig S8C) is phenocopying the high levels of PRMT5 mRNA in stem cells to enable their proliferative capacity.

Discussion

PRMT5 has recently come into focus as a highly promising cancer target due to its overexpression in many cancer types and the finding that its depletion is tumor suppressive in many of these settings (Stopa et al., 2015; Yang and Bedford, 2012). This appreciation has created strong impetus to understand the underlying bases of PRMT5's oncogenic activity. Our data strongly suggests that splicing regulation is a, if not the, major component of PRMT5's oncogenic activity. Specifically, our broad and unbiased tumor cell line screen led to the conclusion that limited splicing capacity creates a vulnerability for PRMT5 inhibition.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Indeed, we believe that PRMT5 inhibitor-sensitive cells are essentially addicted to splicing. This conclusion is reinforced by the finding that a high CLNS1A/RIOK1 ratio predicts PRMT5 inhibitor sensitivity and that knockdown of these factors can render cells more sensitive (CLNS1A) or more resistant (RIOK1) to EPZ treatment. This suggests that PRMT5's activity is highly skewed towards spliceosomal targets in sensitive cells and that inhibition of PRMT5 diminishes splicing capacity in a manner that impedes tumor maintenance and progression. Importantly, our data argue that the C/R ratio can serve as a molecular biomarker to predict sensitivity to PRMT5 inhibition. A major challenge in cancer therapy is the high frequency with which targeted therapeutics fail to demonstrate efficacy in clinical trials, despite promising pre-clinical results (Braun and Hemann, 2016). It is possible that the C/R ratio can be used to predict patient response to PRMT5 inhibitors.

Although there is a widespread perception that introns within polyadenylated transcripts are simply a consequence of splicing errors, recent studies have raised the possibility that such introns are regulatory units (Boutz et al., 2015; Braunschweig et al., 2014). We consider the PRMT5-DI program to be a *bona fide* physiological mechanism. First, DIs are seen at significant levels in normal cells, i.e. without inhibition or mutation of splicing regulators and they are well conserved between mouse and human. Second, our data show that DIs are highly enriched in proliferation regulators and changes in DI inclusion result in the coordinate downregulation of cell cycle genes, which accounts for the proliferative impairment of PRMT5-deficient tumors. Finally, during normal neuronal differentiation, PRMT5 levels are progressively reduced and the inclusion of DIs within many genes including proliferation regulators increases, resulting in reduced expression levels, as the cells move towards the post-mitotic state. This suggests that the frequent upregulation of PRMT5 in tumor cells acts to co-opt the developmental PRMT5-DI axis to enhance DI splicing and enable tumor cell proliferation.

How is PRMT5 deficiency selectively targeted to particular DIs while rarely affecting other classes of introns? Given PRMT5's known role in Sm protein methylation and snRNP assembly, our initial hypothesis was that PRMT5 deficiency would deplete the pools of snRNAs, which in turn might negatively, and selectively, impact the splicing of introns with the "weakest" recognition sequences. However, we do not detect any significant reduction in the levels or nuclear accumulation of snRNAs. This does not rule out the possibility that PRMT5 deficiency somehow alters the assembly or function of the mature spliceosome. Indeed, we favor the idea that PRMT5 loss disrupts the biogenesis and/or recycling of tri-snRNP complexes that are critical for the splicing process, perhaps via changes at Cajal bodies, of which several constituent components possess, or recognize, SDMA marks (Boisvert et al., 2002). However, extensive additional analyses will be required to test this hypothesis. Our data clearly identify weak 5′ splice sites as a frequent feature of PRMT5-regulated DIs. Additionally, we determined that there is a general under-representation of 5′ splice sites with Gs at both the −1 and +5 positions in PRMT5-sensitive introns. However, these features are not sufficient to explain PRMT5 dependency. Many DIs, and also most cassette exons, are immune to the effects of PRMT5 deficiency despite exhibiting these features. Thus, we hypothesize that other splicing regulators expressed at different developmental stages will work in concert with the weak 5′ splice sites to achieve complex splicing patterns such as the ones observed during neuronal differentiation.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

It is becoming apparent that core spliceosome components are frequently mutated in cancer and that splicing factors hold great promise as therapeutic targets (Lee and Abdel-Wahab, 2016). It is tempting to speculate that all tumor cells require a mechanism to circumvent the regulatory processes that restrain efficient splicing in non-proliferative, post-mitotic cells. Upregulation of PRMT5 is one way to achieve this goal, potentially as a cause or consequence of acquiring a more stem cell-like state. We anticipate that other mutations will impinge on DI regulation, in a similar manner to PRMT5 upregulation. These will offer additional or alternate therapeutic targets, beyond PRMT5, to broaden the treatment options for GBM, as well as other tumor types.

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Jacqueline A. Lees (jalees@mit.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

In vivo tumor models—Gl261 and U-87 MG cells were injected into the left cerebral hemispheres of 7 to 9 week old female C57BL/6J (The Jackson Laboratory, 000664) or female NCr nude mice (Taconic, NCRNU-F) performed in close analogy to previously described methods (Ozawa and James, 2010). Subcutaneous injections of U-87 MG and MGG8 cells lines were performed by carefully lifting up the flank skin of NCr nude mice and injecting a sterile cell/PBS suspension (10^6 cells in 100 μl) into the subcutaneous space. Mice were observed daily for the development of tumors. Treatment was initiated when roughly 50% of mice had developed palpable subcutaneous tumors. For experiments with multiple treatment groups, mice were randomly distributed across experimental cohorts. Dimensions of subcutaneous tumors were estimated by repetitive caliper measurements. Tumor volume was calculated with the following formula: Volume = length × width × height × 3.14/6. All mouse experiments were approved by MIT's committee on animal care (CAC) prior to execution. We confirm that all experiments conform to the relevant regulatory standards.

Cell lines—Gl261 murine glioma cells were a kind gift from David Zagzag (New York University) and were cultured in DMEM complete medium (90% DMEM/10%FBS). U-87 MG (ATCC HTB-14), T98G (ATCC CRL-1690), PANC-1 (ATCC CRL-1469), U-2 OS (ATCC HTB-96), HCT-116 (CCL-247), Foreskin fibroblasts CCD-1112Sk (ATCC CRL-2429), A172 (ATCC CRL-1620), LN229 (ATCC CRL-2611), C2BBe1 (ATCC CRL-2102), LN-18 (CRL-2610) and U138 (ATCC HTB-16) cells were acquired from ATCC and grown in DMEM complete medium. HEK293T cells were a kind gift from Piyush Gupta (Massachusetts Institute of Technology) and were propagated in DMEM complete medium. NGP, ACN, KELLY, SKNAS, LAN-1 and BE(2)-C cells were gifts from Alexandre Puissant and Kimberly Stegmaier (Dana-Farber Cancer Institute) and were kept in DMEM complete medium. MCF-7 (ATCC HTB-22) cells were acquired from ATCC and kept in EMEM media supplemented with 10% FBS and 0.01 mg/l bovine insulin (Sigma, I6634). WM2664 (ATCC CRL-1676), SKMEL2 (ATCC HTB-68) and MEWO (ATCC

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

HTB-65) cells were acquired from ATCC and cultured in EMEM medium supplemented with 10% FBS. ZR-75-1 (ATCC CRL-1500) cells were acquired from ATCC and kept in RPMI supplemented with 10% FBS. T47D (ATCC HTB-133) cells were acquired from ATCC and kept in RPMI supplemented with 10% FBS and 0.2 U/mL bovine insulin (Sigma, I6634). OVCAR-8 cells were a kind gift from Madeleine Oudin and Frank Gertler (Massachusetts Institute of Technology) and cultured in DMEM complete media. CAL-33 cells were acquired from DSMZ (ACC 447) and cultured in DMEM complete media. HT-29 cells were acquired from ATCC (ATCC HTB-38) and kept in McCoy's 5a Medium Modified supplemented with 10% FBS, HepG2 were acquired from the David H. Koch Institute core cell line collection bank and grown in EMEM media supplemented with 10% FBS. NCIH2009, SKLU1, NCIH460, A549, NCIH358 and Calu6 cells were kept in RPMI media with 10% FBS and were a kind gift of Jens Köhler and Pasi A. Jänne (Dana-Farber Cancer Institute). The human patient-derived xenograft line MGG8 was derived from a patient tumor at Massachusetts General Hospital (Wakimoto et al., 2012) and grown in Neurobasal medium (Life Technologies) supplemented with N2/B27, penicillin/streptomycin (Life Technologies), GlutaMAX (Life Technologies), recombinant human EGF (20 ng/mL, R & D systems), and recombinant human FGF2 (20 ng/mL, R & D systems).

METHOD DETAILS

In vivo competition assays—Mixed populations of tumor cells were transplanted intracranially as described above. ShRNA expression was then induced by Doxycycline administration at for all mice within the same experiment in parallel at the indicated times post tumor cell injection. Upon display of signs of morbidity, mice were euthanized and decapitated. Tumor tissue was liberated from the brain followed by preparation of a single-cell suspension employing a brain tumor dissociation kit (Miltenyi, #130-095-942) in combination with the gentleMACS dissociator (Miltenyi) according to the manufacturer's instructions. Flow cytometry was conducted using BD Fortessa or BD LSR-II machines. Raw data was analyzed using FlowJo software v10 (Tree Star). The ratio of induced over non-induced cells was calculated by the following formula: $(dsRed2^{+})/(mVenus^{+}, dsRed2^{-})$.

ShRNA screen—GFP dilution assays were used as a surrogate for shRNA library sizes in order to establish a feasible library complexity for *in vivo* screens in the Gl261 model. In detail, Gl261 cells transduced at different percentages with a retroviral vector expressing GFP were injected into the brain of C57BL/6J mice. Upon onset of morbidity, mice were euthanized and tumors were dissociated. Percentage of GFP^{+} cells was estimated by flow cytometry. The relationship between the percentage of GFP^{+} cells in a mixed population and the modeled library size is: Library size = $1/GFP^{+}$ %.

The shRNA library was acquired from Transomic technologies and consisted of 2059 shRNAs targeting 311 epigenetics-related genes in a MSCV/LTR/mIR30/PGK/Puromycin/IRES/green fluorescent protein (pMLP) retroviral vector. The library was then transferred into a MSCV/LTR/mIR30/SV40/green fluorescent protein (MLS) retroviral vector by EcoRI/XhoI liberation of the shRNA encoding regions and two shRNAs targeting KRAS were spiked in. Gl261 cells were then infected at low infection rates to avoid multiple infections per cell. Successfully transduced cells were later enriched by fluorescence-

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

activated cell sorting (FACS). After a brief *in vitro* expansion, cells were either injected into syngeneic C57BL/6J mice, distributed into 9 culture dishes (3 × *in vitro* vehicle control, 3 × *in vitro* low-dose temozolomide, 3 × *in vitro* high-dose temozolomide) or subjected to genomic DNA extraction for the t=0 days control. *In vitro* culture conditions were kept so that at least a 250 × shRNA representation was conserved at all times. Half of the mice were treated with either a single i.p. injection of temozolomide (Sigma) at 50 mg/kg BW on day 8 after cell injection or with the corresponding vehicle. On the same day, tissue culture cells were either treated with vehicle, a low dose of temozolomide (7.5 μM) or a high dose of temozolomide (10 μM). Mice were examined daily for clinical signs of tumor onset and were euthanized once they had reached established euthanasia criteria. *In vitro* samples were harvested when roughly 50% of the mice of the corresponding treatment or control mouse cohort had been euthanized. Tissue was flash-frozen in liquid nitrogen. Genomic DNA from both the tumors and the tissue culture samples was extracted using Trizol reagent (Sigma) followed by ethanol precipitation. Integrated shRNA viral vectors were then amplified from genomic DNA using HotStarTaq DNA Polymerase (Qiagen, # 203205).

The DNA integrity and quantity was determined on the Agilent 2100 Bioanalyzer (Agilent), followed by pooling of the multiplexed DNA fragments and sequencing with a HiSeq sequencer (Illumina). Non-normalized count numbers per shRNA were calculated by aligning the sequencing reads to the input shRNA library allowing for a single mismatch. The counts tables were then used as an input for the edgeR based shRNA-seq bioinformatical tool for pooled shRNA screen analysis (Dai et al., 2014).

Cell line screen statistical analysis

Cell line dosing and creation of dose-response-curves: The half maximal inhibitory concentration (IC_{50}) and the maximum effect (Emax) were calculated using the Dr. Fit software package (Di Veroli et al., 2015). Drug activity areas were calculated by summing up the individual differences between a measured activity at a certain drug dose and a fixed reference point of no activity similar to an approach taken in the Cancer Cell Line Encyclopedia (CCLE, (Barretina et al., 2012)). Therefore, a drug activity area of 0 corresponds to an inactive drug, whereas an activity area of 11 corresponds to a drug with a complete activity at any concentration. *TP53* gene mutation status was derived from literature, *MTAP* status was both derived from literature (where available) and tested by qRT-PCR.

Gene set enrichment analysis: For pathway enrichment analysis of cell lines with published transcriptome data, activity areas were Pearson-correlated to the level of individual gene transcripts across cell lines. Gene-centric RMA-normalized mRNA expression data was acquired from CCLE (http://www.broadinstitute.org/ccle, (Barretina et al., 2012)). The resulting ranked Pearson-correlation-score/gene transcript matrix was then used as an input for the pre-ranked list tool of gene set enrichment analysis (GSEA) (Subramanian et al., 2005). Enrichment scores were calculated for the following Molecular Signatures Database (MSigDB) collections: (1) CP:KEGG: KEGG gene sets (http://www.genome.jp/kegg/pathway.html) and (2) CP:REACTOME: Reactome gene sets (http://www.reactome.org/). Highlighted REACTOME splicing/mRNA processing data sets:

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

MRNA_PROCESSING,
PROCESSING_OF_CAPPED_INTRON_CONTAINING_PRE_MRNA,
REACTOME_MRNA_SPLICING, REACTOME_MRNA_CAPPING,
REACTOME_MRNA_SPLICING_MINOR_PATHWAY.

EPZ sensitivity biomarker derivation: We first created a curated list of 59 genes with core components of the spliceosome complex, the SMN complex, the PRMT5 methylosome and PRMT5 cofactors. We next calculated pairwise expression comparisons of these 59 genes yielding 1711 expression profiles. These expression profiles were then used to calculate Pearson significance p values for their correlation the drug activity scores across 18 cell lines.

Functional enrichment analysis for gene level changes and DIs—Functional enrichment analysis was performed utilizing either the publicly available DAVID functional annotation tool (https://david.ncifcrf.gov/) or GSEA gene set enrichment analysis (http://software.broadinstitute.org/gsea/index.jsp). Gene set enrichment scores were calculated for the following gene sets:

CP:REACTOME: Reactome gene sets (http://www.reactome.org/) and

CP:KEGG: KEGG gene sets (http://www.genome.jp/kegg/pathway.html).

Functional enrichment maps were generated by processing DAVID output files with the Enrichment map tool (Merico et al., 2010) for Cytoscape 3.4.0 (Shannon et al., 2003).

Functional enrichment analysis of genes affected by DIs in U-87 MG cells: We first identified genes harboring DIs and then ran DAVID functional enrichment using all expressed genes as a background control.

Functional enrichment analysis of gene level/coding isoform changes in U-87 MG cells following EPZ treatment: We first identified genes downregulated upon PRMT5 inhibition with EPZ in the corresponding stringency categories. We then ran DAVID functional enrichment using all expressed genes as a background control.

Functional enrichment analysis of genes affected by DI increase upon ES cell differentiation: We first identified genes harboring DIs and belonging to the indicated behavioral cluster. We next performed DAVID functional enrichment analysis using all genes expressed as a background control.

Nuclear/cytoplasmic fractionation and Northern blots—U-87 MG cells treated for 72 hr with DMSO or EPZ were harvested by trypsinization, washed and counted, and 5 million cells per sample were pelleted by centrifugation. Nuclear and cytoplasmic fractions were isolated using the NE-PER reagents (Thermo Fisher) according to manufacturer's directions, except that the initial incubation step was performed for 8 min rather than the recommended 10 min. Nuclear and cytoplasmic RNA was extracted according to manufacturer's directions using TRIzol-LS (Thermo Fisher) added directly to the final

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

lysates. Northern blotting protocols were adapted from published sources (Rio, 2014; Wu et al., 2013). Following purification of RNA, nuclear and cytoplasmic cell equivalents from 500,000 (for Sybr Gold staining) or 50,000 (for Northern blots) cells were loaded on a 10% acrylamide-urea/TBE gel. The gel was then stained with SYBR Gold (Thermo Fisher) and imaged, followed by transfer to a Hybond-N+ membrane (GE Healthcare). The membrane was crosslinked using a Stratalinker UV crosslinker (Stratagene), using the autocrosslink setting (120000 μJ for 1 min at 254 nM). The membrane was then blocked overnight at 42° C with rotation in DIG Easy-Hyb buffer (Roche) containing 1% w/v blocking reagent (Roche). Antisense DNA oligonucleotide probes were labeled at the 3′ end using terminal transferase (NEB) in the presence of digoxigenin-11-UTP (Sigma-Aldrich) for 10 min, then precipitated. After blocking, membranes were then incubated overnight at 42° C with 1nM each probe. The membrane was washed in 2X SSC/0.1% SDS (20X SSC: 3M NaCl, 0.3M trisodium citrate), 1X SSC/0.1% SDS, 0.1X SSC each for 15 min at 42° C. The membrane was then equilibrated in DIG buffer (0.1 M Maleic acid, 0.15 M NaCl, pH 7.5) and blocked for 30 min at room temperature in blocking buffer (DIG buffer containing 1% w/v Roche blocking reagent). The membrane was next incubated for 30 min at room temperature with 1:2000 dilution of alkaline-phosphatase conjugated anti-Digoxigenin Fab fragments (Roche) in blocking buffer. After 4X-10 min washes in DIG buffer + 0.3% Tween 20, the membrane was equilibrated in detection buffer (0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5), then submerged in CDP-STAR reagent (Roche) for detection. For quantification, a SYBR Gold-stained gel containing three independently fractionated replicates was scanned and quantified using a Typhoon Trio with ImageQuant software (GE Healthcare) and the gel was aligned with the Northern blot to identify the correct bands representing each snRNA.

Western Blot—Proteins were extracted in RIPA buffer (Boston Bioproducts, #BP-115DG) and separated by SDS-PAGE. They were then transferred onto PVDF membranes and probed with antibodies against Symmetric Di-Methyl Arginine Motif/SDMA (Cell Signaling, #13222, MultiMab rabbit monoclonal antibody mix, 1:1000), HSP90 (BD, #610418, Clone 68, 1:10,000), PRMT5 (Cell Signaling, #2252S, 1:1000), GAPDH (Santa Cruz, #sc-365062, Clone G-9, 1:1000), SOS1 (Cell Signaling, #5890, 1:1000), STAT5B (Santa Cruz, #sc-1656, 1:200), RAF1 (Cell Signaling, #9422, 1:1000), AURKB/AIM1 (BD, #611082, Clone 6, 1:1000), SNRPB (Sigma, #HPA003482, 1:200), SNRPD3 (Sigma, #HPA001170, 1:200), SART3/TIP110 (Bethyl, #A301-521A, 1:10,000), PRP3 (MBL, #D171-3, 1:1000), SNRNP40 (MBL, #RN096PW, 1:1000), anti-H4R3me2s (Abcam, #5823, 1:500) and EIF4E (Cell Signaling, #9742, 1:1000). Proteins of interest were detected with HRP-conjugated α-Rabbit (Cell Signaling, #7074, 1:3000), Rat (Cell Signaling, #7077,1:2000) and α-Mouse (Cell Signaling, #7076, 1:2000) antibodies and visualized with the Pierce ECL Western blotting substrate (Thermo Scientific) using the ImageQuant LAS 4000 imaging system (GE) or film exposure.

Cell cycle staining and flow cytometry—For cell cycle analysis, 5-ethynyl-2′-deoxyuridine (EdU) was purchase from Life Technologies (#E10187), suspended in DMSO and incorporated into cells at a concentration of 10 μM for 30 min. Cells were then harvested by trypsinization, washed with PBS and fixed with ethanol. They were then permeabilized with ice-cold 0.25% TritonX100 in PBS for 15 min on ice, rinsed with 1%

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

BSA dissolved in PBS and stained with a Click-it reaction cocktail (Life Technologies, # C10269) according to the manufacturer's instructions. Cells were then washed with ice-cold 0.25% Tritonx100 in PBS and stained with DAPI (Life Technologies, #D1306) at a final concentration of 1 μg/mL. *In vitro* competition assays were performed by noting the percentage transduced GFP+ on the day 0 and day 12 of doxycycline treatment. Flow cytometry was conducted using BD Fortessa, BD Canto, BD Celesta or BD LSR-II machines. Raw data was analyzed using FlowJo software v10 (FlowJo LLC).

Cell proliferation assays—CellTrace Far Red Cell Proliferation Kit (Life technologies, # C34564) was used for *in vitro* cell proliferation studies according to the manufacture's manual. In brief, tissue culture cells were fluorescence dye stained for 20 min at 37°C in a tissue culture incubator, washed multiple times with serum containing sterile tissue culture medium and then incubated in the presence of EPZ or vehicle for multiple days. On the stated days, cells were trypsinized, washed and then subjected to flow cytometry (channel: APC-Cy7). The mean fluorescence of cells was noted and later normalized to the initial mean fluorescence at 24 hr after CellTrace labeling. The relative amount of proliferation for every time point was then estimated by calculating the reciprocal of the mean fluorescence fraction.

Cell line doubling time analysis—Cell line doubling time information was derived from the NCI60 screen and can be found at https://dtp.cancer.gov/discovery_development/nci-60/publications.htm.

Senescence-associated (SA) β-Galactosidase Staining—Cells were fixed in 0.5% gluteraldehyde in PBS for 15 min at room temperature and then washed in 1mM MgCl2 in PBS. They were then stained in β-galactosidase staining solution (5 mM potassium ferrocyanide, 5mM potassium ferricyanide, 1 mg/ml X-Gal (Affymetrix, #10077) in PBS/1mM MgCl2), washed with water and overlaid with a 25% glycerol/water solution. Cells were imaged on an EVOS microscope (Life Technologies).

Vector generation—The following retroviral shRNA constructs were used: MSCV/LTR/miR30/SV40/GFP (MLS), constitutive pRSF91-GFP-miRE and Doxycycline-inducible Tre/dsRed/miRE/PGK/Venus/IRES/rtTA3 (RTREVIR). pRSF91-GFP-miRE was a kind gift of Adrian Schwarzer and Axel Schambach (MH Hannover, Germany), RTREVIR was a gift of Johannes Zuber (IMP Vienna, Austria). ShRNA target sequences are listed separately. A mouse MGMT cDNA vector was purchased from Origene (clone ID: MC202089), the protein coding DNA sequence was PCR amplified and ligated into the restriction-enzyme linearized retroviral vector pMIG. pMIG was a gift from William Hahn (Addgene plasmid # 9044). A DNA string encoding for a codon-optimized PRMT5 coding sequence (CDS) was synthesized (Gene Art Project, Thermo Fisher Scientific), PCR-amplified and cloned into the restriction enzyme linearized lentiviral vector pLJM1- by Gibson assembly to create pLJM1-PRMT5. pLJM1-EGFP was a gift from David Sabatini GFP (Addgene plasmid # 19319). A mutant PRMT5 CDS (p.Tyr304del, p.Tyr307del) lentiviral expression construct was created by PCR-based site-directed mutagenesis of pLJM1-PRMT5.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

***In vitro* and *in vivo* Doxycycline treatment—**Doxycycline powder was purchased from Clontech (#631311) and resuspended in water. Tissue culture cells were treated with a final concentration ranging from 1 μg/mL to 10 μg/mL (depending on the induction to toxicity profile of individual cell lines). Doxycycline was delivered *in vivo* by drinking water (1 mg/mL + 2% sucrose) or solid food pellet (purchased from Bio-Serv, #S3888) supplementation.

In vitro EPZ sensitivity cell line screen—The PRMT5 inhibitor EPZ015666 was purchased from Selleck Chemicals (#S7748) and DC Chemicals (#DC8012). Compound integrity was controlled for every new batch by both Nuclear Magnetic Resonance (NMR) Spectroscopy and Matrix-assisted laser desorption/ionization (MALDI). EPZ was dissolved in DMSO. A 12-point EPZ dilution series (including the no-drug DMSO control point) was plated in a 96-well plate 2-fold dilution steps, starting at 500 mM. 100 nl of the compound solutions were pin-transferred (V&P Scientific, CA, pin tool mounted onto Tecan Freedom Evo 150 MCA96 head, Tecan, CA) into 96-well plates and incubated for 120h. Cells were seeded in 96-plates on the day prior to drug treatment. Seeding cell numbers were adjusted for different cell lines so that vehicle-treated wells reached sub-confluence 120 hr after vehicle treatment. The IncuCyte zoom microscope (Essen Bioscience) was used to estimate well confluence.

In vitro shRNA competition assays

To score influence factors on EPZ sensitivity: Tumor cells were transduced with the constitutive shRNA vector pRSF91-GFP-miRE. Mixed populations of transduced and untransduced tumor cells were then seeded out (day 0) and treated with 10 μM EPZ or DMSO (vehicle). Percentages of transduced cells were estimated by FACS 6 days after drug treatment. Fold-changes were calculated by: Fold change = $GFP\%_{EPZ}/GFP\%_{Vehicle}$.

To score depletion following shRNA knockdown: Tumor cells were transduced with the Doxycycline-inducible shRNA vector RTREVIR. Mixed populations of transduced and untransduced tumor cells were then seeded out and treated with Doxycycline at concentrations stated above or vehicle (day 0). Cells were then cultured in the presence of vehicle or Doxycycline for the indicated time periods. Relative percentages of transduced cells were either estimated by FACS repetitively over the course of the experiment or at a single endpoint. The transduced $Venus^+$ fraction of cells was then normalized for every time point to the fraction of $Venus^+$ cells on day 0.

In vivo EPZ and Temozolomide drug treatment—The PRMT5 inhibitor EPZ was purchased from Selleck Chemicals (#S7748) and DC Chemicals (#DC8012). For in vivo treatment studies, EPZ was dissolved in 0.5% methyl-cellulose in water (10 mg/mL) and administered to mice every 12 hr by oral gavage (dose: 100 mg EPZ per kg body weight) for the indicated times. 0.5% methyl-cellulose in water administered per oral gavage was used as a negative control. Temozolomide was purchased from Sigma-Aldrich and administered by intraperitoneal injection at a dose of 50 mg/kg body weight.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Histological stains and immunohistochemistry

Masson's Trichrome staining: Tumors were fixed were fixed in 10% neutral buffered formalin, embedded in paraffin and sectioned at 4 M. After deparaffinization, slides were placed onto an Artisan Special Stainer (Dako-Agilent Technologies) and stained with a Masson's Trichrome stain kit. Slides were then scanned using a Leica Aperio AT2 Digital Slide Scanner.

H&E stain: Tumors were fixed in were fixed in 10% neutral buffered formalin, processed to paraffin and sectioned at 4 M. After deparaffinization, these were then stained consecutively with Harris Acidified Hematoxylin and Eosin, dehydrated and then coverslipped.

KI67, Cleaved Caspase 3 and H4R2me2s staining of tissue sections: Tumor bearing brains were fixed in 10% neutral buffered formalin, embedded in paraffin and sectioned at 5 M (brain sections) or 4 M (subcutaneous tumor sections). After deparaffinization and antigen retrieval, sections were quenched with H_2O_2 (Thermo Fisher, #TA125H2O2q), blocked with Mouse Block M (Biocare RBM961L) then stained with anti-H4R3me2s (Abcam, #5823, 1:500), anti-KI67 (#550609, 1:50, BD) or anti-Cleaved Caspase 3 (#9664, 1:2000, Cell Signaling) antibodies followed by an incubation with the ImmPRESS HRP Anti-Rabbit IgG Polymer Detection Kit (Vector Laboratories, #MP-7405, undiluted, for H4R2me2s) or Mouse-on-Mouse HRP-Polymer (Biocare Medical, MM620, for KI67) or Rabbit-on-Rodent HRP-Polymer (Biocare Medical, RMR622, for Cleaved Caspase 3) followed by Quanto DAB visualization (Thermo Fisher, #TA-125-QHDX) using the Thermo Autostainer 360. Slides were then scanned using a Leica Aperio AT2 Digital Slide Scanner and staining was quantified using Aperio ImageScope Software v12.3.0.5056. For H4R3me2s, the average nuclear staining intensity of the tumor was normalized to the average staining intensity of an area of normal, non-tumor bearing adjacent area of the brain. For Cleaved Caspase 3 and KI67, positive nuclei were counted in a blinded fashion and the resulting counts were normalized to the counted tumor area.

Tumor senescence stains: Following euthanasia, tumors were placed in disposable embedding molds, covered with OCT (Tissue-Tek), carefully frozen by partial submersion in Isopropanol-covered dry ice and cryosectioned. Slides were then fixed in 0.5% gluteraldehyde in PBS for 15 min at room temperature and washed in 1 mM MgCl2 in PBS. They were then stained in β-galactosidase staining solution (5 mM potassium ferrocyanide, 5 mM potassium ferricyanide, 1 mg/ml X-Gal (Affymetrix, #10077) in PBS/1mM $MgCl_2$), washed with water and overlaid with a 25% glycerol/water solution. Slides were then imaged on an EVOS microscope (Life Technologies).

In vitro siRNA treatment—SiRNA transfections were conducted using the Lipofectamine RNAiMAX Transfection Reagent (Qiagen) according to the manufacturer's instructions.

Matrix-assisted laser desorption ionization mass spectrometry imaging (MALDI-MSI)—The tissue mimetic model was prepared by transferring 100 μL aliquots of mouse brain tissue homogenates into 7 pre-weighed 1.5 mL Eppendorf tubes for tissue

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

weight estimation. The appropriate concentrations of EPZ standard were then added in each tube at final concentrations of 50000, 10000, 5000, 1000, 500, 100, and 50 ng of drug per g of tissue homogenate. Based on the previously published work (Groseclose and Castellino, 2013), we next developed a homemade 3D printed model later used as negative mold. This negative mold was filled with a 40% gelatin solution, and placed at −80°C for 2 hr. Once frozen, the gelatin containing negative mold was placed into a Microm HM550 cryostat (Thermo Scientific) with the chamber chilled at −20°C. The frozen gelatin cylinder was removed from the negative mold and replaced with 100 μL of the spiked tissue homogenates. The tissue model was then placed at −80°C for 24 hr followed by 12 μm cryosectioning using a Microm HM550 cryostat with both the chamber and the specimen holder pre-chilled at −20°C. The resulting sections were thaw-mounted onto ITO-coated microscopic slides (Bruker Daltonics) for MALDI-MSI and onto optical slides for hematoxylin and eosin (H/E) staining. The mounted sections were then dried for 15 min in a desiccator prior to matrix application. 2,5-dihydroxybenzoic acid matrix (2,5-DHB, 160 mg/mL solution in methanol/0.1% TFA 70:30 vol/vol) was deposited using a TM-sprayer (HTX imaging) using the following conditions: flow rate, 180 μl/min; spray nozzle velocity, 1,200 mm/min; spray nozzle temperature, 75°C; nitrogen gas pressure, 10 psi; track spacing, 2 mm; number of passes, 4. Mass spectra of mouse brain tissue and tissue mimetic sections were acquired using a SolariX XR Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR) (9.4 T) (Bruker Daltonics). MALDI-MSI experiments were acquired with a pixel step size for the surface raster set to 125 μm in FlexImaging 4.0 software (Bruker Daltonics). The analyses were performed in positive ion mode by continuous accumulation of selected ions (CASI) in a mass range comprised between m/z 380-620 and a laser intensity set to 20%. Each mass spectrum is the sum of 250 laser shots randomized over 10 positions within the same spot (25 shots/position) at a laser frequency of 1000 Hz. The MALDI images were displayed using FlexImaging 4.0 and SCILS Lab 2016 (SCILS GmbH) after total ion current (TIC) normalization. The permeability of EPZ through the blood vessel was visualized following the signal of the drug (m/z 304.203 ± 0.001) and a biomarker of vasculature (i.e. heme at m/z 616.177 ± 0.001) as previously described (Liu et al., 2013).

Evans blue blood brain barrier analysis—Evans blue dye (Sigma, #E2129) was dissolved freshly in a sterile 0.9% NaCl/H_2O solution at a concentration of 20 mg/mL and sterile filtered. 100 μl of Evans blue solution was then injected into the lateral tail-vein of mice followed by an incubation of approximately 1 hr. Animals were then anesthetized, transcardially perfused with heparinized PBS solution and decapitated. The brain was carefully removed, weighted and snap-frozen in liquid nitrogen. The frozen tissue was then powderized under liquid-nitrogen protection in a standard mortar followed by careful resuspension in 1 mL 50% Trichloroacetic acid (TCA, Alfa Aesar, #A11156) in 0.9% NaCl. The mixture was then transferred to a gentleMACS C Tube (Miltenyi) and dissociated with the gentleMACS tissue dissociator (Miltenyi). Solid components were spun out, the supernatant was harvested and diluted at a ratio of 1:4 with 100% ethanol. Light emission at 680 nm after excitation at 620 nm was recorded with a plate reader (Tecan) and compared to a standard dilution series of Evans blue dye. Light emission was visualized with the Odyssey LTX imaging system (Li-cor).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Non-terminal mouse bleeding and blood tests—Peripheral blood of mice was collected in serum and EDTA-coated microvette tubes (Sarstedt). Subsequent blood tests were performed by the MIT Division of Comparative Medicine (DCM) comparative pathology laboratory. Displayed reference ranges were adjusted to species, gender and age.

Total RNA purification, cDNA synthesis and quantitative Real-Time PCR (qRT-PCR)—Total RNA was isolated employing the NucleoSpin RNA kit (Machery-Nagel). For gene expression quantification, equal amounts of cDNA were synthesized using the PrimeScript RT reagent Kit with gDNA Eraser (Takara) and mixed with the Fast SYBR Green Mastermix (Applied Biosystems). Experiments were performed using a StepOnePlus Real-Time PCR System (Applied Biosystems) and StepOne Software v2.2.2 (Applied Biosystems) was utilized for data processing. Murine or human GAPDH was amplified as an internal loading control. The number of cycles required to cross a fluorescence threshold (CT-value) was noted for each transcript and normalized to the internal control GAPDH (for total transcript level studies) or to the corresponding total transcript (for DI level studies). For DI-specific qRT-PCR assays, one or two primer pairs spanning the DI and one of the adjacent exons were employed. If multiple DI-specific primer pairs were employed for the same DI, relative abundance changes were estimated and normalized to the corresponding total transcript level followed by mathematical averaging.

RNA-seq—Total RNA from U-87 MG cells treated with EPZ at a concentration of 10 μM for 72 hr was isolated using the NucleoSpin RNA kit (Machery-Nagel) followed by RNA integrity and quantity determination on the Agilent 2100 Bioanalyzer (Agilent). The TruSeq RNA sample preparation kit (Illumina) was used according to manufacturer's instructions for a PolyA-based mRNA enrichment. The fragmented mRNA samples were subjected to cDNA synthesis and library generation using the TruSeq RNA sample preparation kit (Illumina). Sequencing was performed with a NextSeq sequencer (Illumina). Embryonic stem cell differentiation data is publicly available from the Sequence Read Archive (SRA) under the accession number PRJNA185305. *Prmt5* KO data in NPCs is publicly available from GEO under the accession number GSE45285.

Mapping reads and junctions, determination and quantitation of alternative splicing events: Raw RNA-seq reads were mapped using STAR aligner version 2.4.1d (Dobin et al., 2013) with the parameters:

U-87 MG human glioblastoma cell line, mapped to Hg19 with genome annotation based on Gencode v19: STAR –runMode alignReads –runThreadN 2 –genomeDir hg19 –twopassMode Basic –sjdbOverhang 74 –outSAMtype BAM SortedByCoordinate –outFilterMultimapNmax 20 –outFilterMismatchNmax 999 –outFilterMismatchNoverLmax 0.04 –alignIntronMin 70 –alignIntronMax 500000 –alignMatesGapMax 500000 –alignSJoverhangMin 8 –alignSJDBoverhangMin 1 –outSAMstrandField intronMotif –outFilterType BySJout.

Identification and quantitation of detained introns: For additional details, see (Boutz et al., 2015). To identify detained introns, reads were first mapped using Bowtie 1.0.1 (Langmead et al., 2009) with the parameters:

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

bowtie -n 1 -l 30 -e 70 -t -p 2 -m 1 -S -X 500

Mapped reads were then filtered using Bedtools 2.25.0 (Quinlan and Hall, 2010) remove reads overlapping expressed repeats from the UCSC genome browser RNA repeats and Gencode v19 or Ensembl Mm9 annotations and coding exons annotated in the Gencode v19 annotations. Reads that spanned an exon-intron junction were required to have a minimum overlap of 10 nucleotides within the intron. Paired-end fragments where both reads mapped to the same intron, were counted as a single read. The set of intronic coordinates derived from mapped reads as described above was processed to remove first and last introns, overlapping alternative introns (e.g. the intron spanning a skipped cassette exon, proximal 5′ and distal 3′), introns containing a known polyA site (derived from the Gencode v19 metadata "PolyA site" coordinates for human data and 3′ sequencing-derived polyA sites (Almada et al., 2013) for mouse), and overlapping introns within complex splicing units using an interval scheduling algorithm. Using DESeq (Anders and Huber, 2010), we normalized intronic read counts in order to account for depth of coverage and data dispersion across replicates. For each biological replicate, we generated an in silico replicate under a null model to use as a baseline for enrichment analysis, as follows. For each gene, the sum of normalized intronic read counts was then used to allocate reads to individual introns under a null model. First, an "effective length" for each intron was derived using 100mer mapability tracks (Roderic Guigo Lab, CRG, Barcelona; (ENCODE Project Consortium, 2011)) from the UCSC Genome Browser. Positions of non-unique 100mer alignments were summed and then subtracted from the corresponding intron's length. A variance stabilizing transform based on the square root of intron effective length adjusted by RNA-seq read length (√(L-d) where L=mapability adjusted intron length, d = RNA-seq read length) was then used to weight individual introns. The sum of normalized intronic reads per gene in each RNA-seq replicate was then partitioned and allocated to each intron proportional to its weight. This results in two in silico null model replicates, one corresponding to each RNA-seq replicate. Differential analysis using DESeq was then used to determine introns enriched in read coverage (observed read counts) compared to the in silico null model replicates (expected counts, if all introns within a gene are present at equal levels) using an FDR adjusted p value threshold of 0.01 and fold change threshold of 2 with a required minimum of 20 reads per intron from all replicates combined to eliminate low-end noise. Individual detained intron differences in splicing efficiency were measured using DEXSeq (Anders et al., 2012), with the custom annotation described above as input. A genome track file can be found in the supplement allowing to display DIs across the U-87 MG transcriptome.

Mouse *in vitro* neurogenesis data set, mapped to Mm9 with genome annotation based on UCSC mm9: STAR –runMode alignReads –runThreadN 2 –genomeDir UCSC_mm9 –twopassMode Basic –sjdbOverhang 49 –outSAMtype BAM SortedByCoordinate –outFilterMultimapNmax 20 –outFilterMismatchNmax 999 –outFilterMismatchNoverLmax 0.06 –alignIntronMin 70 –alignIntronMax 500000 –alignMatesGapMax 500000 –alignSJoverhangMin 8 –alignSJDBoverhangMin 1 –outSAMstrandField intronMotif –outFilterType BySJout Mapped splice junctions detected in all samples were combined and processed using custom Python scripts to filter junctions representing ~< 1% of transcripts, and to define alternative splicing categories.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Alternative and constitutive classifications were performed using custom Python scripts, and are agnostic with regard to existing annotations other than known gene boundaries. If no overlapping introns exist for a given intron, it is assigned to the constitutive class. The subgroups containing overlapping introns are assigned a splicing classification if the start and end coordinates of all of the constituent introns fall into a pattern representing a known splice type (cassette, mutually exclusive, alternative 5′ splice site, alternative 3′ splice site). The introns within a classified subgroup are designated according to their position within the splicing event:

USMXEa: Intron upstream of the first of two mutually-exclusive exons (exon a)

DSMXEa: Intron downstream of exon a (filtered due to overlap with exon b)

MXEMI: Intron between first and second mutually-exclusive exons

USMXEb: Intron upstream of the second of two mutually-exclusive exons (exon b, filtered due to overlap with exon a)

DSMXEb: Intron downstream of exon b

MXEBS: Intron skipping both mutually exclusive exons (filtered due to overlap with exons)

CAS_US: Intron upstream of cassette exon

CAS_DS: Intron downstream of cassette exon

CAS_Skipped: Intron skipping cassette exon (filtered due to overlap with exon)

DUAL: Intron with alternative 5′ and 3′ splice sites (grouped with complex)

A3P: Alternative 3′ splice site, proximal

A3D: Alternative 3′ splice site, distal (filtered due to overlap with exon)

A5P: Alternative 5′ splice site, proximal (filtered due to overlap with exon)

A5D: Alternative 5′ splice site, distal

T3P: Tandem (NAGNAG) alternative 3′ splice site, proximal (grouped with Alt 3′ ss)

T3D: Tandem (NAGNAG) alternative 3′ splice site, distal (filtered due to overlap with exon)

T5P: Tandem (NGUNGU) alternative 5′ splice site, proximal (filtered due to overlap with exon)

T5D: Tandem (NGUNGU)) alternative 5′ splice site, distal (grouped with Alt 5′ ss)

Complex: Intron part of alternative event that cannot be categorized into only one of the primary classes (mutually exclusive, cassette, alt 5′ss, alt 3′ss)

Constitutive: No alternative splice site usage

To quantify differences in alternative splicing events between DMSO controls and EPZ treatment, splicing events determined from mapped junctions as described above were

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

converted into event-specific gff3 annotation files compatible with MISO (Katz et al., 2010). Because MISO performs pairwise-comparisons, we first determined which biological replicates were most correlated for splicing consistency and determined the most consistent pairings by Wilcoxon rank-sum comparison. We then performed three separate MISO analyses per alternative splicing type (Cassette exons, alternative 3′ or tandem 3′ splice site, alternative 5′ or tandem 5′ splice site, and mutually-exclusive exons) using these pairs of replicates: DMSO_r1 paired with EPZ_r2 (D1E2), DMSO_r2 with EPZ_r3 (D2R3), and DMSO_r3 with EPZ_r1 (D3E1). To be considered significant, 2 of the 3 replicate comparisons were required to move in the same direction, both with a Bayes factor of >= 5.

Genome-wide transcript annotation: We used custom Python scripts to derive a complete transcriptome annotation based on Hg19 (Gencode v19) gene start and end boundaries and the location of polyA sites for data derived from the U-87 MG human glioblastoma cell line, and Mm9 (Mus_musculus_NCBI_build37.1) for the neuronal differentiation data, the genomic locations of detained introns as determined above, and the genomic locations of all splice junctions. For comparisons of isoforms containing detained introns to the consensus coding sequence, the annotation for each gene was set to identify the consensus isoform, i.e. the junctions defining each exon are the most frequently used junction detected within all of the combined samples. Each detained intron within a gene was then assigned to an additional transcript.

Gene and isoform expression: Gene-level and isoform-level quantification was performed using Rsem version 1.2.26 (Li and Dewey, 2011) and EBSeq (Leng et al., 2013) using the custom genomic annotation described above, as well as Gencode v19 annotation with all three biological replicates each for DMSO control and EPZ-treated cells with the parameters:

> rsem-calculate-expression –forward-prob 0.5 –output-genome-bam -p 2 –paired-end Genome_files/gencode_hg19

Followed by read-count matrix assembly and standard EBSeq differential expression analysis. Gene-level and isoform-level quantification was also performed using DESeq (Anders and Huber, 2010) to obtain independent validation as well as to derive an adjusted p value for use in gene expression volcano plots. The $\log_2$ fold-changes (DMSO vs. EPZ) for all genes were highly correlated between the two different annotations (r=0.88 using EBSeq, r=0.76 using DESeq)

Gene expression differences at the total gene level were considered significant if the posterior probability of differential expression (PPDE) as determined by EBSeq > 0.95, or if EBSeq did not output a statistical significance measure but both annotations using DESeq were significant at an adjusted p value of < 0.05 and changing in the same direction, gene expression was considered significantly different. See Table S4 for details.

To determine significant differences in detained intron versus consensus coding transcripts, all genes with a detained intron isoform that changed in abundance at a PPDE of >0.95, or in which the isoform significance was not measured or not significant but for which the change in detained intron isoforms was >= 10%, were defined as either “DIUp” if the detained

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

intron levels increased upon EPZ treatment, “DIDown” if they decreased upon EPZ treatment, or “DINoChange” if they did not significantly change upon drug treatment. See Table S5 for details.

Unsupervised clustering of biological replicates: Total gene level expression as calculated in transcripts per million (TPM) by EBSeq was used to calculate the $\log_2$-fold change in average TPM for the 3 biological replicates of each treatment (DMSO & EPZ). Using the EBSeq statistical significance for each gene, genes with PPDE > 0.95 and $\log_2$-fold change >= 0.5 were used to identify the most significantly changed gene set, resulting in a set of 928 genes. The per-gene Z score was calculated for each biological replicate, and these scores were plotted and clustered using the Heatmap.2 function in R with hierarchical clustering by Euclidean distance.

Population-based (lower stringency) significance measure for differential coding isoform expression: To avoid false-negative classification of genes as not having differential coding isoform expression due to replicate noise or expression changes below the individual isoform significance level, we increased the sensitivity of the statistical assessment using a population-based method, as is often employed in the context of microRNA targets. Taking each group of genes, we generated a cumulative distribution function (CDF) plot (see Fig S5F) using the EBSeq output for the $\log_2$-fold change in the consensus coding isoform of each gene in the group. Significance of population distribution shifts were determined by a two-sided Kolmogorov-Smirnov (KS) test. Two groups of DI-containing genes for which the coding isoforms did not meet the individual isoform statistical significance level (DI unchanged/Coding unchanged/Gene unchanged, light blue; and DI up/Coding unchanged/Gene unchanged, green) both showed statistically significant population based shifts toward lower coding isoform expression in EPZ-treated cells. To conservatively pick the most down-regulated genes from this population, we required them to exhibit a decrease of >= 10% in coding isoform percentage, resulting in 597 lower-stringency genes in this group.

Splice site strength: Relative splice site strengths were determined using maximum entropy modeling with the program MaxEnt (Yeo and Burge, 2004) using human splice sites as the training set. Motifs surrounding the 5′ splice sites of each set of DIs were derived using Weblogo (http://weblogo.berkeley.edu).

The di-nucleotide frequency matrices analysis was designed to allow for the identification of not only nucleotides that diverge from the consensus in absolute frequency, but also of pairs of nucleotides at two different positions in a motif (here, the 9-nucleotide consensus 5′ splice site) that are significantly enriched or depleted even if the absolute frequency of each member of the pair is insignificantly different between two conditions. Such di-nucleotide variations can be conceived of as altering the interaction of the sequence with a second oligonucleotide (in this case, the U1 or U6 snRNAs) engaged in a base-pairing interaction in which the it is the sum of all base-paired nucleotides, rather than the individual base-pairs, that determine the strength of the interaction. For each intron, the 9 nucleotides spanning the −3 to +6 positions of the 5′ splice site (the invariant GT positions being +1 and +2) were extracted using samtools (Li et al., 2009) and divided into lists according to whether the splice site usage was unaffected, spliced more, or spliced less in EPZ-treated or *Prmt5*-KO

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

cells. Using a custom python script, a matrix was compiled such that for each instance of all four nucleotides in all nine positions, the number of times each nucleotide is present at every other position in the 5′ splice site of the introns in that set was recorded. These raw numbers were then normalized to give the frequency of each specific di-nucleotide pairing within each set. The differences in this frequency were determined between unchanged and spliced more, and unchanged and spliced less, for each di-nucleotide pair. A Z-score was calculated for each of the differences in frequency, and those exhibiting a Z-score >= 2 or <= −2 were considered significant. The di-nucleotide pair with the highest (absolute value) Z-score at each position-pair was selected and the results plotted as a pseudo-correlation matrix, using the R package corrplot (https://CRAN.R-project.org/package=corrplot).

Detained intron isoform clustering during neurogenesis: Detained intron containing and consensus isoforms were quantified across all time points as described above. The percentage of consensus isoform for each gene across the time course was normalized by Z-score and the clustering of genes by these Z-scores was performed using k-means clustering in R (kmeans function in{stats}), with the optimal number of clusters (here we used 4) determined by within group sum of squares.

The Cancer Genome Atlas (TCGA) PRMT5 expression analysis—Expression levels in the form of RSEM-determined normalized transcript levels for PRMT5 were obtained from the Broad Institute FireHose browser (http://firebrowse.org/) and the measurements of statistical significance in the expression level distributions between normal brain, lower-grade gliomas, and glioblastoma multiforme were performed in R using Wilcoxon rank-sum tests.

***In silico* assessment of blood brain barrier permeability—**The SDF chemical data files for compounds were obtained from https://cactus.nci.nih.gov/chemical/structure or from https://pubchem.ncbi.nlm.nih.gov/. These were then used to predict the blood brain barrier (BBB) penetration score using a published tool (Liu et al., 2014).

QUANTIFICATION AND STATISTICAL ANALYSIS

GraphPad PRISM 7, R 2.15.3, R 3.2.3 and Python 2.7.2 software packages were used to perform statistical analyses. Error bars represent standard error of the mean (SEM), unless otherwise stated. Statistical tests employed are specified in figure legends. The log-rank (Mantel-Cox) test was used for the survival curves analyses.

DATA AND SOFTWARE AVAILABILITY

U-87 MG RNA-seq raw data has been deposited to GEO (accession number: GSE93813).

It can be accessed at:

https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=svknwwwofrgdnuh&acc=GSE93813.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Acknowledgments

MIT Ludwig Center for Cancer Research, David H. Koch Institute Frontier Research program, Kathy and Curt Marble Cancer Research Fund, and NIH/NCI grants PO1-CA42063, R01GM034277 and P30-CA14051. Fellowships: Mildred-Scheel fellowship/German Cancer Foundation (C.J.B.), MIT School of Science Fellowship in Cancer Research (M.S.) and American Cancer Society (P.L.B.). We thank the Koch Institute Swanson Biotechnology Center for technical support and colleagues for input on the manuscript. The results published here are in part based upon data generated by the TCGA research network.

References

Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature. 2013; 499:360–363. [PubMed: 23792564]

Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol. 2010; 11:R106. [PubMed: 20979621]

Anders S, Reyes A, Huber W. Detecting differential usage of exons from RNA-seq data. Genome Res. 2012; 22:2008–2017. [PubMed: 22722343]

Antonysamy S, Bonday Z, Campbell RM, Doyle B, Druzina Z, Gheyi T, Han B, Jungheim LN, Qian Y, Rauch C, et al. Crystal structure of the human PRMT5:MEP50 complex. Proc Natl Acad Sci U S A. 2012; 109:17960–17965. [PubMed: 23071334]

Barretina J, Caponigro G, Stransky N, Venkatesan K, Margolin AA, Kim S, Wilson CJ, Lehár J, Kryukov GV, Sonkin D, et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature. 2012; 483:603–607. [PubMed: 22460905]

Battle, DJ., Kasim, M., Yong, J., Lotti, F., Lau, CK., Mouaikel, J., Zhang, Z., Han, K., Wan, L., Dreyfuss, G. Cold Spring Harbor Symposia on Quantitative Biology. 2006. The SMN complex: An assembly machine for RNPs; p. 313-320.

Bezzi M, Teo SX, Muller J, Mok WC, Sahu SK, Vardy LA, Bonday ZQ, Guccione E. Regulation of constitutive and alternative splicing by PRMT5 reveals a role for Mdm4 pre-mRNA in sensing defects in the spliceosomal machinery. Genes Dev. 2013; 27:1903–1916. [PubMed: 24013503]

Boisvert FM, Côté J, Boulanger MC, Cléroux P, Bachand F, Autexier C, Richard S. Symmetrical dimethylarginine methylation is required for the localization of SMN in Cajal bodies and pre-mRNA splicing. J Cell Biol. 2002; 159:957–969. [PubMed: 12486110]

Boutz PL, Bhutkar A, Sharp PA. Detained introns are a novel, widespread class of post-transcriptionally spliced introns. Genes Dev. 2015; 29:63–80. [PubMed: 25561496]

Braun CJ, Hemann MT. Rewiring the solid tumor epigenome for cancer therapy. Expert Rev Anticancer Ther. 2016; 16:977–987. [PubMed: 27410491]

Braunschweig U, Barbosa-Morais NL, Pan Q, Nachman EN, Alipanahi B, Gonatopoulos-Pournatzis T, Frey B, Irimia M, Blencowe BJ. Widespread intron retention in mammals functionally tunes transcriptomes. Genome Res. 2014; 24:1774–1786. [PubMed: 25258385]

Bresson SM, Hunter OV, Hunter AC, Conrad NK. Canonical Poly(A) Polymerase Activity Promotes the Decay of a Wide Variety of Mammalian Nuclear RNAs. PLoS Genet. 2015; 11:e1005610. [PubMed: 26484760]

Chan-Penebre E, Kuplast KG, Majer CR, Boriack-Sjodin PA, Wigle TJ, Johnston LD, Rioux N, Munchhof MJ, Jin L, Jacques SL, et al. A selective inhibitor of PRMT5 with in vivo and in vitro potency in MCL models. Nat Chem Biol. 2015; 11:432–437. [PubMed: 25915199]

Chari A, Golas MM, Klingenhäger M, Neuenkirchen N, Sander B, Englbrecht C, Sickmann A, Stark H, Fischer U. An Assembly Chaperone Collaborates with the SMN Complex to Generate Spliceosomal SnRNPs. Cell. 2008; 135:497–509. [PubMed: 18984161]

Dai Z, Sheridan JM, Gearing LJ, Moore DL, Su S, Dickins RA, Blewitt ME, Ritchie ME. shRNA-seq data analysis with edgeR. 2014 F1000Research 3.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013; 29:15–21. [PubMed: 23104886]

Dvinge H, Bradley RK. Widespread intron retention diversifies most cancer transcriptomes. Genome Med. 2015; 7:45. [PubMed: 26113877]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. Nat Rev Cancer. 2016; 16:413–430. [PubMed: 27282250]

ENCODE Project Consortium. A User's Guide to the Encyclopedia of DNA Elements (ENCODE). PLoS Biol. 2011; 9:e1001046. [PubMed: 21526222]

Gargiulo G, Serresi M, Cesaroni M, Hulsman D, van Lohuizen M. In vivo shRNA screens in solid tumors. Nat Protoc. 2014; 9:2880–2902. [PubMed: 25411954]

Groseclose MR, Castellino S. A Mimetic Tissue Model for the Quantification of Drug Distributions by MALDI Imaging Mass Spectrometry. Anal Chem. 2013; 85:10099–10106. [PubMed: 24024735]

Guderian G, Peter C, Wiesner J, Sickmann A, Schulze-Osthoff K, Fischer U, Grimmler M. RioK1, a New Interactor of Protein Arginine Methyltransferase 5 (PRMT5), Competes with pICln for Binding and Modulates PRMT5 Complex Composition and Substrate Specificity. J Biol Chem. 2011; 286:1976–1986. [PubMed: 21081503]

Hegi ME, Diserens AC, Gorlia T, Hamou MF, de Tribolet N, Weller M, Kros JM, Hainfellner JA, Mason W, Mariani L, et al. MGMT Gene Silencing and Benefit from Temozolomide in Glioblastoma. N Engl J Med. 2005; 352:997–1003. [PubMed: 15758010]

Hidalgo M, Amant F, Biankin AV, Budinská E, Byrne AT, Caldas C, Clarke RB, de Jong S, Jonkers J, Mælandsmo GM, et al. Patient-derived xenograft models: an emerging platform for translational cancer research. Cancer Discov. 2014; 4:998–1013. [PubMed: 25185190]

Hubbard KS, Gut IM, Lyman ME, McNutt PM. Longitudinal RNA sequencing of the deep transcriptome during neurogenesis of cortical glutamatergic neurons from murine ESCs. F1000Research. 2013; 2:35. [PubMed: 24358889]

Jacob AG, Smith CWJ. Intron retention as a component of regulated gene expression programs. Hum Genet. 2017

Katz Y, Wang ET, Airoldi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat Methods. 2010; 7:1009–1015. [PubMed: 21057496]

Kryukov GV, Wilson FH, Ruth JR, Paulk J, Tsherniak A, Marlow SE, Vazquez F, Weir BA, Fitzgerald ME, Tanaka M, et al. MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells. Science (80-). 2016; 351:1214–1218.

Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10:R25. [PubMed: 19261174]

Lau D, Magill ST, Aghi MK. Molecularly targeted therapies for recurrent glioblastoma: current and future targets. Neurosurg Focus. 2014; 37:E15.

Lee SCW, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. Nat Med. 2016; 22:976–986. [PubMed: 27603132]

Leng N, Dawson JA, Thomson JA, Ruotti V, Rissman AI, Smits BMG, Haag JD, Gould MN, Stewart RM, Kendziorski C. EBSeq: An empirical Bayes hierarchical model for inference in RNA-seq experiments. Bioinformatics. 2013; 29:1035–1043. [PubMed: 23428641]

Leten C, Struys T, Dresselaers T, Himmelreich U. In vivo and ex vivo assessment of the blood brain barrier integrity in different glioblastoma animal models. J Neurooncol. 2014; 119:297–306. [PubMed: 24990826]

Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011; 12:323. [PubMed: 21816040]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–2079. [PubMed: 19505943]

Liu H, Wang L, Lv M, Pei R, Li P, Pei Z, Wang Y, Su W, Xie XQ. AlzPlatform: an Alzheimer's disease domain-specific chemogenomics knowledgebase for polypharmacology and target identification research. J Chem Inf Model. 2014; 54:1050–1060. [PubMed: 24597646]

Liu X, Ide JL, Norton I, Marchionni MA, Ebling MC, Wang LY, Davis E, Sauvageot CM, Kesari S, Kellersberger KA, et al. Molecular imaging of drug transit through the blood-brain barrier with MALDI mass spectrometry imaging. Sci Rep. 2013; 3:2859. [PubMed: 24091529]

Marjon K, Cameron MJ, Quang P, Clasquin MF, Mandley E, Kunii K, McVay M, Choe S, Kernytsky A, Gross S, et al. MTAP Deletions in Cancer Create Vulnerability to Targeting of the MAT2A/PRMT5/RIOK1 Axis. Cell Rep. 2016; 15:574–587. [PubMed: 27068473]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Mavrakis KJ, McDonald ER, Schlabach MR, Billy E, Hoffman GR, deWeck A, Ruddy DA, Venkatesan K, Yu J, McAllister G, et al. Disordered methionine metabolism in MTAP/CDKN2A-deleted cancers leads to dependence on PRMT5. Science. 2016; 351:1208–1213. [PubMed: 26912361]

Merico D, Isserlin R, Stueker O, Emili A, Bader GD. Enrichment Map: A Network-Based Method for Gene-Set Enrichment Visualization and Interpretation. PLoS One. 2010; 5:e13984. [PubMed: 21085593]

Newcomb EW, Zagzag D. Abstract - SpringerLink. CNS Cancer. 2009

Ozawa T, James CD. Establishing intracranial brain tumor xenografts with subsequent analysis of tumor growth and response to therapy using bioluminescence imaging. J Vis Exp. 2010:e1986.

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010; 26:841–842. [PubMed: 20110278]

Rio DC. Northern Blots for Small RNAs and MicroRNAs. Cold Spring Harb Protoc. 2014; 2014 pdb.prot080838-prot080838.

Rodríguez-Paredes M, Esteller M. Cancer epigenetics reaches mainstream oncology. Nat Med. 2011; 17:330–339. [PubMed: 21386836]

Sanchez SE, Petrillo E, Beckwith EJ, Zhang X, Rugnone ML, Hernando CE, Cuevas JC, Godoy Herz MA, Depetris-Chauvin A, Simpson CG, et al. A methyl transferase links the circadian clock to the regulation of alternative splicing. Nature. 2010; 468:112–116. [PubMed: 20962777]

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13:2498–2504. [PubMed: 14597658]

Stopa N, Krebs JE, Shechter D. The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. Cell Mol Life Sci. 2015; 72:2041–2059. [PubMed: 25662273]

Stupp R, Mason WP, van den Bent MJ, Weller M, Fisher B, Taphoorn MJB, Belanger K, Brandes AA, Marosi C, Bogdahn U, et al. Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma. N Engl J Med. 2005; 352:987–996. [PubMed: 15758009]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545–15550. [PubMed: 16199517]

Di Veroli GY, Fornari C, Goldlust I, Mills G, Koh SB, Bramhall JL, Richards FM, Jodrell DI. An automated fitting procedure and software for dose-response curves with multiphasic features. Sci Rep. 2015; 5:14701. [PubMed: 26424192]

Wakimoto H, Mohapatra G, Kanai R, Curry WT, Yip S, Nitta M, Patel AP, Barnard ZR, Stemmer-Rachamimov AO, Louis DN, et al. Maintenance of primary tumor phenotype and genotype in glioblastoma stem cells. Neuro Oncol. 2012; 14:132–144. [PubMed: 22067563]

Wu J, Huang HY, Hopper AK. A rapid and sensitive non-radioactive method applicable for genome-wide analysis of *Saccharomyces cerevisiae* genes involved in small RNA biology. Yeast. 2013; 30:119–128. [PubMed: 23417998]

Yang Y, Bedford MT. Protein arginine methyltransferases and cancer. Nat Rev Cancer. 2012; 13:37–50. [PubMed: 23235912]

Yap K, Lim ZQ, Khandelia P, Friedman B, Makeyev EV. Coordinated regulation of neuronal mRNA steady-state levels through developmentally controlled intron retention. Genes Dev. 2012; 26:1209–1223. [PubMed: 22661231]

Yeo G, Burge CB. Maximum Entropy Modeling of Short Sequence Motifs with Applications to RNA Splicing Signals. J Comput Biol. 2004; 11:377–394. [PubMed: 15285897]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Significance

PRMT5 inhibitors are an emerging class of targeted cancer therapeutics. However, even highly promising molecularly-oriented agents have frequently yielded disappointing results in clinical trials due to a lack of predictive biomarkers enabling pre-selection of patients most likely to respond to therapy. Here, we overcome this constraint by establishing that tumor cells sensitive to the lead PRMT5 inhibitor EPZ015666 are ‘addicted’ to a specific form of splicing, detained introns. Sensitive cells exhibit an altered expression ratio of two PRMT5 co-factors, which skews PRMT5 methylation activity towards core components of the spliceosome and acts as a predictive biomarker for EPZ015666 sensitivity. We thus define a broad and targetable splicing program, which may be exploitable for therapy far beyond PRMT5 inhibition.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- DI splicing is a regulated pathway promoting proliferation gene expression
- GBMs assume control of the DI pathway, creating an exploitable vulnerability
- PRMT5 inhibition induces DI inclusion and yields potent anti-tumor activity
- CLNS1A/RIOK1 ratio is a predictive biomarker for PRMT5 inhibitor sensitivity

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

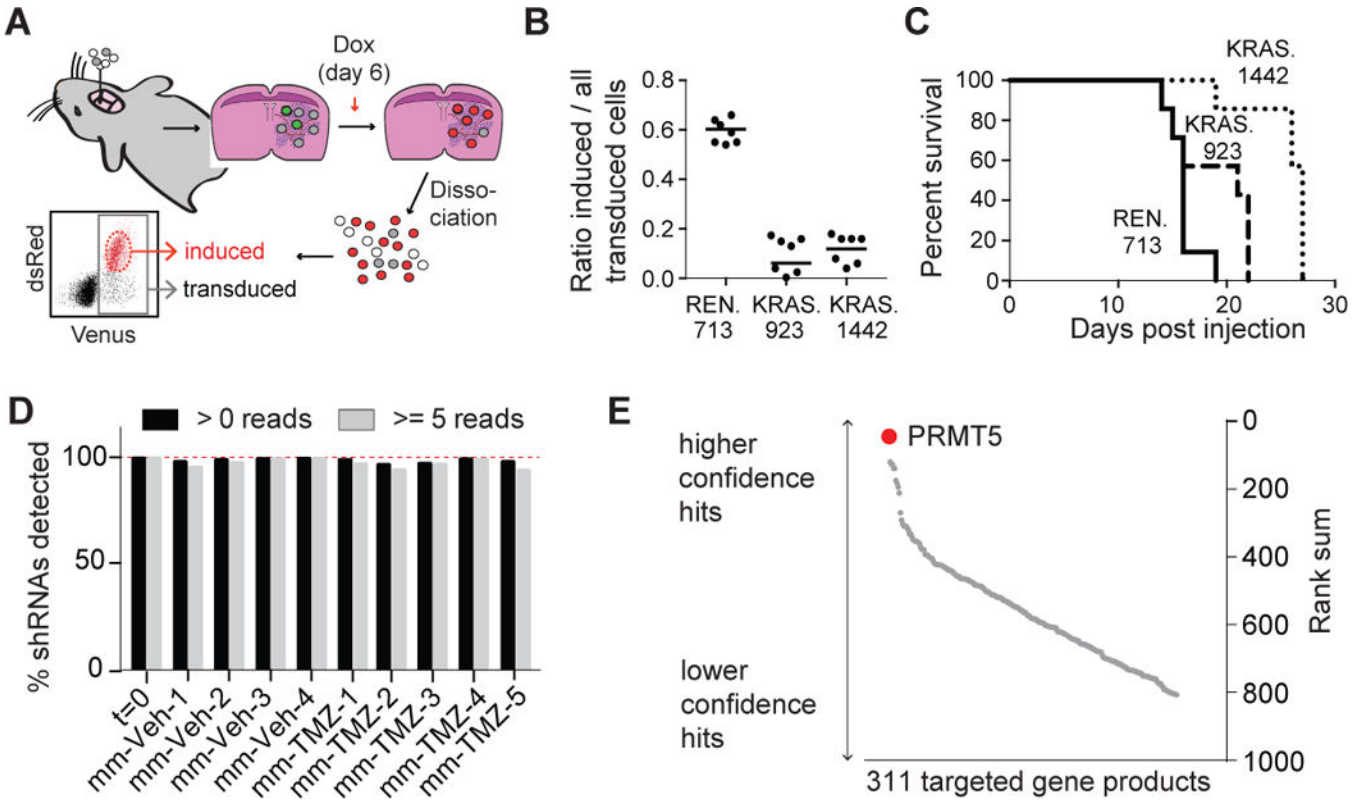


Figure 1. RNAi screen identifies PRMT5 as mediator of GBM growth

(**A**) Schematic of *in vivo* competition assays in intracranial GBMs. (**B**) shRNA representation at harvesting in an intracranial competition assay of Gl261 cells with shRNA expression induction 6 days post cell implantation into C57BL/6J mice, n=7 per condition). 1-way ANOVA p<0.0001, Dunnett's multiple comparisons tests: adj. p value = 0.0001 for REN.713 vs. KRAS.923 and REN.713 vs. KRAS.1442. Horizontal lines represent mean. (**C**) Survival plot of mice in (B), Log-rank (Mantel-Cox) p=0.0839 for REN.713 vs. KRAS.923 and p=0.0005 for REN.713 vs. KRAS.1442. (**D**) Percentage of shRNAs normalized to expected library input across indicated samples. (**E**) Rank sum waterfall plot across all shRNA sub-screens. See also Figure S1/Table S1/Table S2.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

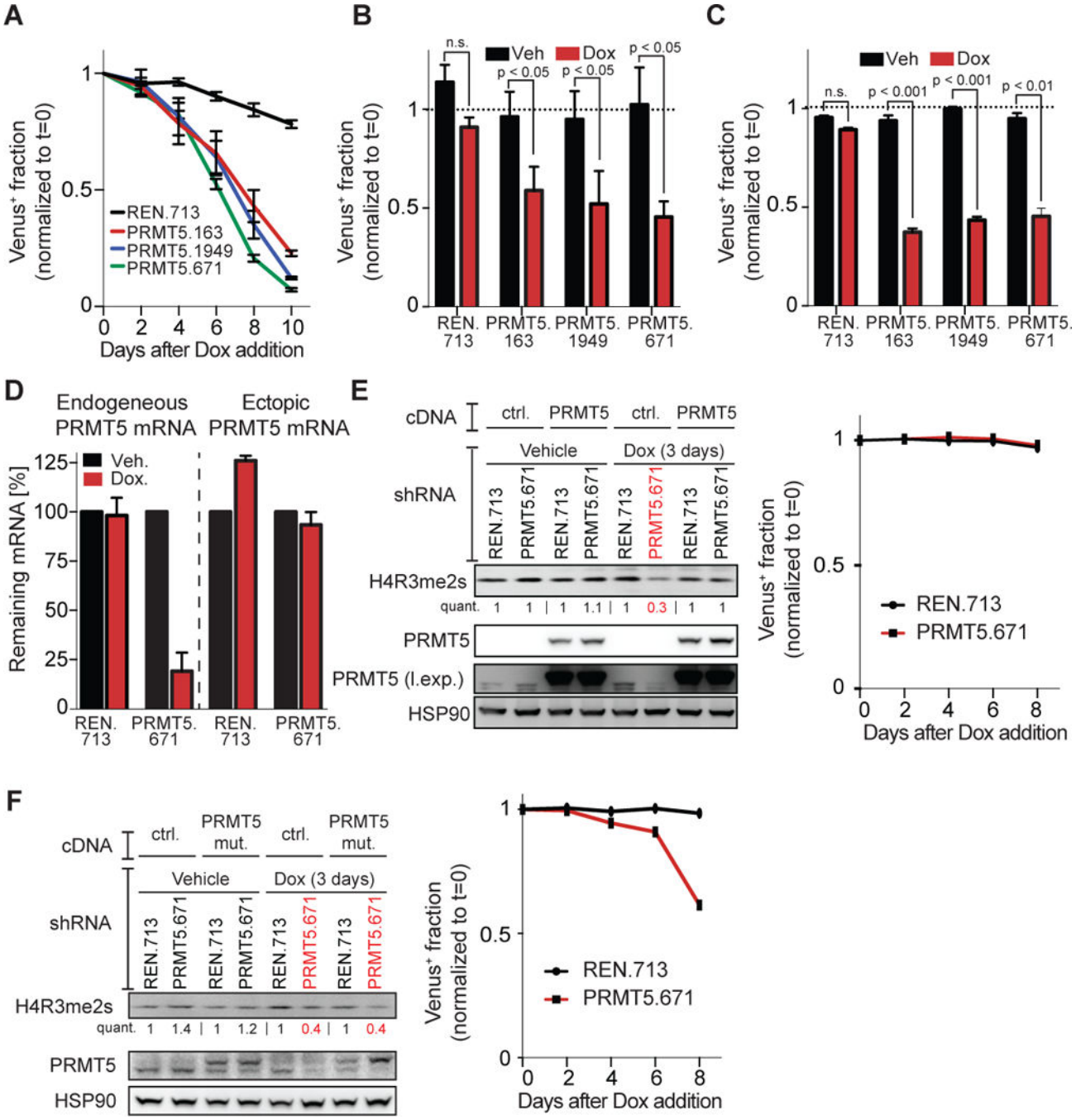


Figure 2. Loss of PRMT5 impairs cellular fitness in a methyltransferase dependent manner

(**A**) *In vitro* competition assay in U-87 MG cells (n=3) normalized to pre-shRNA induction levels (t=0). (**B,C**) Competition assay using the U138 (**B**) or the A172 (**C**) cell line, n=3/line, assayed 12 days post shRNA induction with Dox. Significance estimated by paired t-tests. (**D**) Changes in endogenous and ectopic PRMT5 mRNA levels 2 days post shRNA induction. (**E**) U-87 MG cells with constitutive expression of wild-type, shRNA resistant PRMT5 cDNA and Dox induced expression of shRNAs as indicated, assayed by immunoblot (I. exp. = increased exposure) and competition assay. (**F**) U-87 MG cells with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

constitutive expression of PRMT5 mutant cDNA and Dox induced expression of shRNAs as indicated, assayed by immunoblot (left) and competition assay (right). Error bars = +/− SEM. See also Figure S2.

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

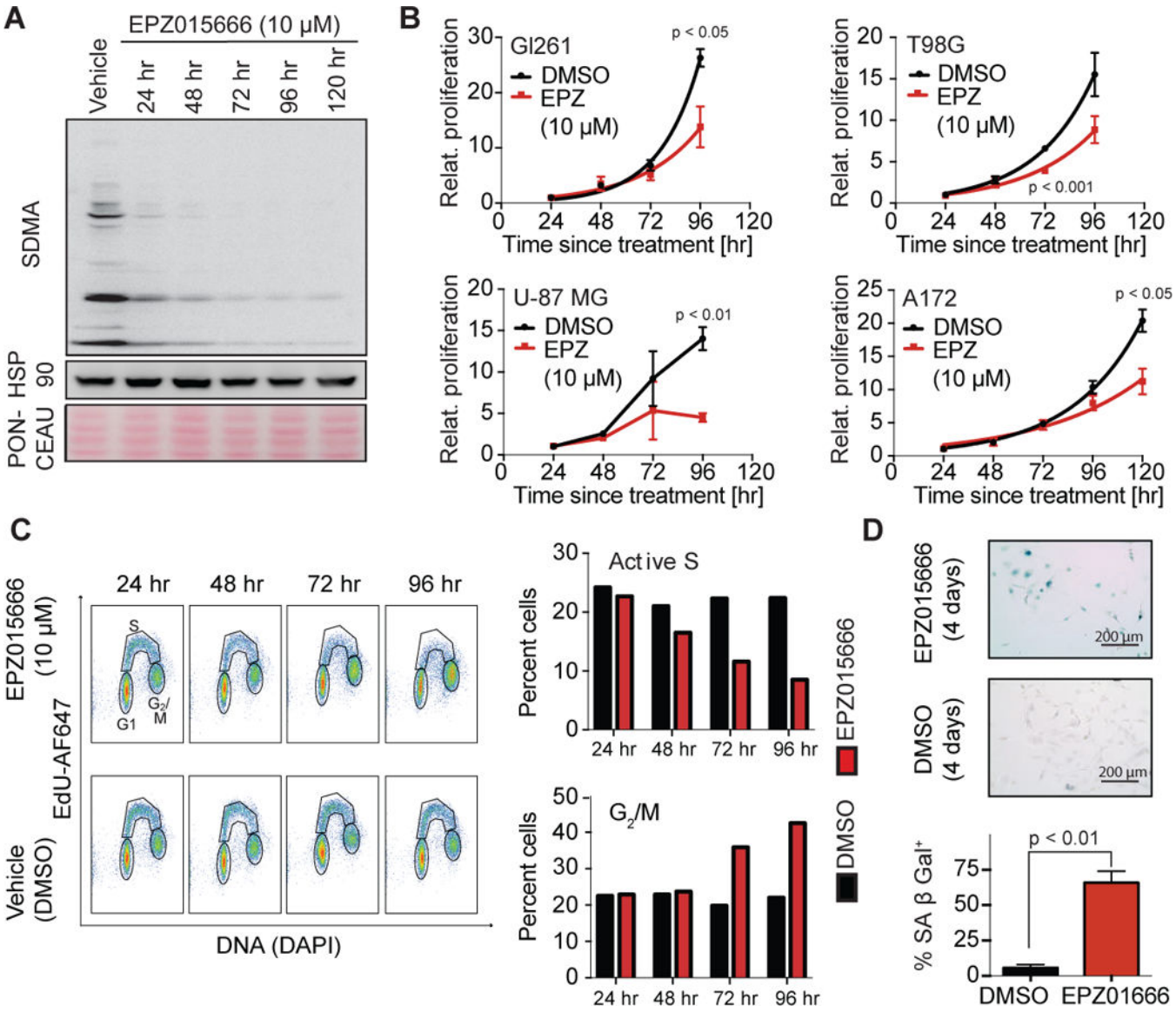


Figure 3. PRMT5 inhibition promotes G_2/M arrest and senescence

(A) Immunoblot of SDMA over time for U-87 MG cells treated with 10 μM EPZ. (**B**) Relative proliferation of indicated GBM cell lines in response to DMSO control or 10 μM EPZ. Significance estimated by unpaired t-tests. (**C**, **D**) FACS-based cell cycle profiles (**C**, left) with quantification (**C**, right) and senescence-associated (SA) β-galactosidase staining (**D**) of U-87 MG cells treated with DMSO or 10 μM EPZ. Significance estimated by unpaired t-test. Error bars = +/− SEM. See also Figure S3.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

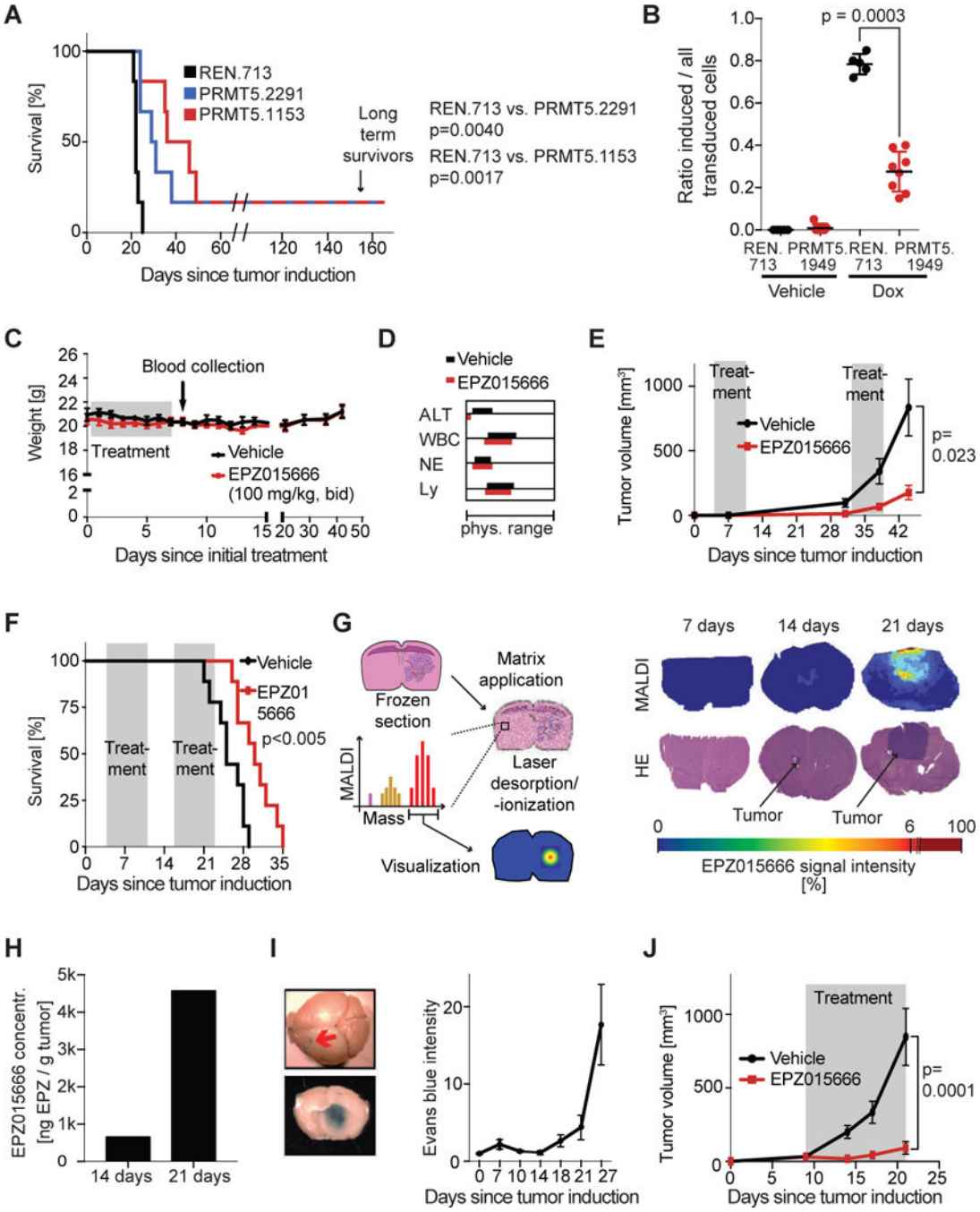


Figure 4. Genetic or pharmacological inhibition of PRMT5 has strong anti-tumor effects

(**A**) Survival plot of C57BL/6J mice following intracranial implantation of Gl261 cells constitutively expressing the indicated shRNA constructs (n=6 per condition). Significance estimated with Log-Rank (Mantel-Cox) test. (**B**) Intracranial competition assay of U-87 MG cells in nude mice, shRNA induction 5 days post injection. n=5 for REN.713, n=8 for PRMT5.1949. Significance estimated with KS-test. Horizontal lines represent mean. **(C, D)** Weight (**C**), and blood (**D**) profiles of non-tumor bearing C57BL/6J mice dosed with oral EPZ or vehicle. ALT = alanine aminotransferase, WBC = white blood cells, NE =

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

neutrophils, Ly = lymphocytes. (**E**) Tumor volume over time in mice with subcutaneous (s.c.) U-87 MG tumors treated with vehicle (n=24 tumors/12 mice) or EPZ (n=22 tumors/11 mice). Significance estimated with Mann-Whitney test at the final time point. (**F**) Survival of mice with intracranial U-87 MG tumors, treated with EPZ (n=9) or vehicle (n=9) for indicated times. Significance: Log-Rank (Mantel-Cox test). (**G, H**) MALDI-MSI workflow (**G**, left) and MALDI signal intensity/H&E brain sections of intracranial U-87 MG tumors in nude mice (**G**, right), and quantification of EPZ concentration (**H**). (**I**) Evan's blue dye extravasation, macroscopic example of dye extravasation with red arrow depicting tumor injection site (left) and Evan's blue dye intensity in brain samples of nude mice at indicated days post U-87 MG cell injection (right). (**J**) Volumes of s.c. MGG8 PDX tumors treated with vehicle/EPZ at indicated times, n=10. Significance estimated with Mann-Whitney test at final time point. Error bars: +/− SEM. See also Figure S4.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

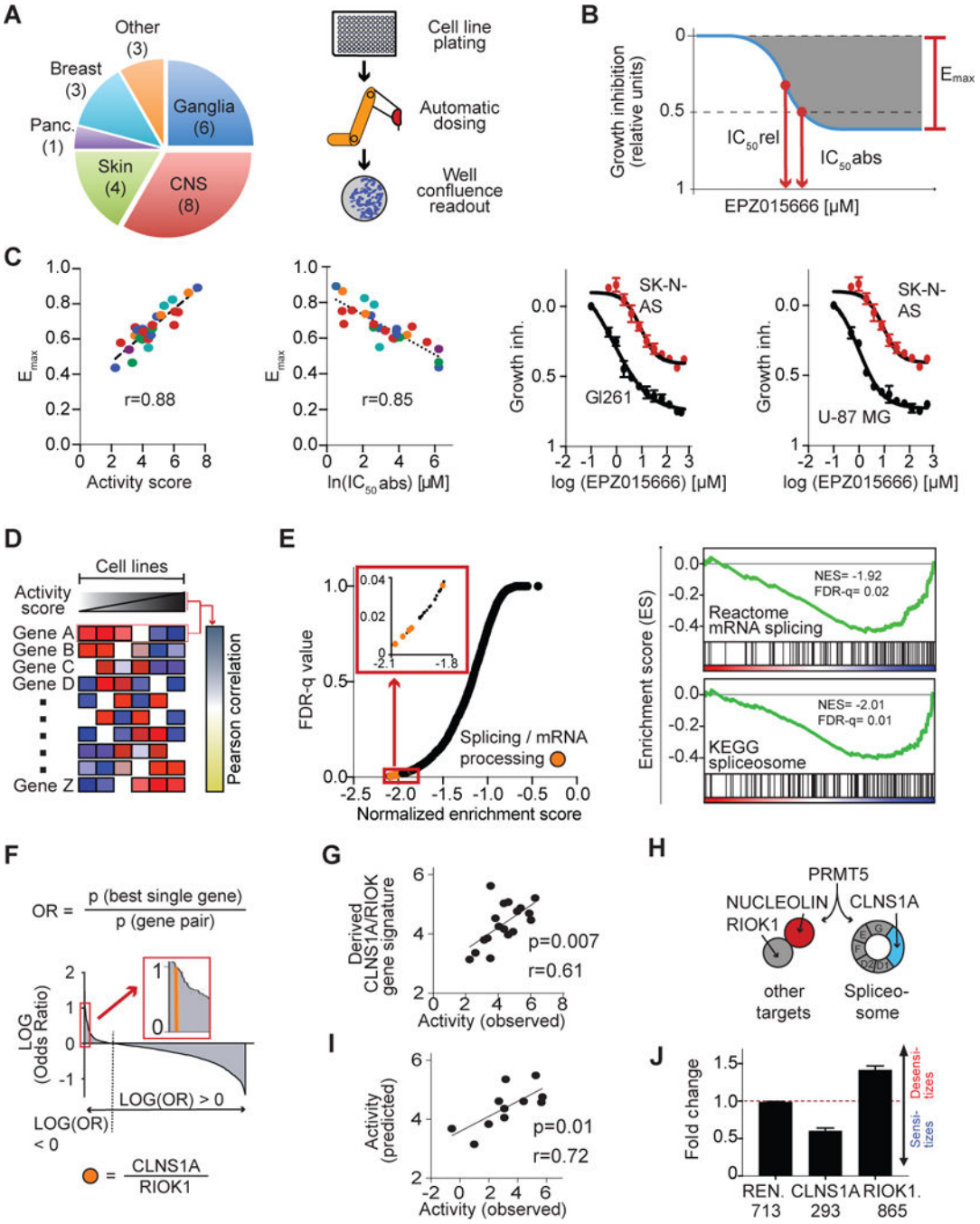


Figure 5. Biomarker predicts EPZ015666 sensitivity

(A) Tissue origin and numbers of cell lines used in screen (pie chart) and schematics of EPZ dosing pipeline. (**B**) Cell line sensitivity parameters. (**C**) Linear regressions of indicated cell line drug response (left two graphs). Color indicates tissue origin, as in (A). EPZ dose curves of indicated cell lines (right two graphs). (**D**) Correlation scheme linking cell line EPZ sensitivity to gene expression. (**E**) GSEA FDR-q values and normalized enrichment scores for Reactome gene sets (left) and examples of GSEA-derived enrichment plots (right). (**F**) Log odds ratios of the p value for the single gene of each pair most correlated with activity

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

score over the p value of the paired metagene. (**G**) Correlation between C/R ratio-based biomarker signature and measured activity scores. (**H**) Schematic of competition between CLNS1A and RIOK1 for PRMT5. (**I**) Correlation between predicted and measured activity scores. (**J**) U-87 MG competition assay with indicated shRNAs after 6 days vehicle/10 μM EPZ treatment. Ratio: % $shRNA^+$ cells in drug over vehicle treated cells. Error bars = +/− SEM. See also Figure S5/Table S3.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

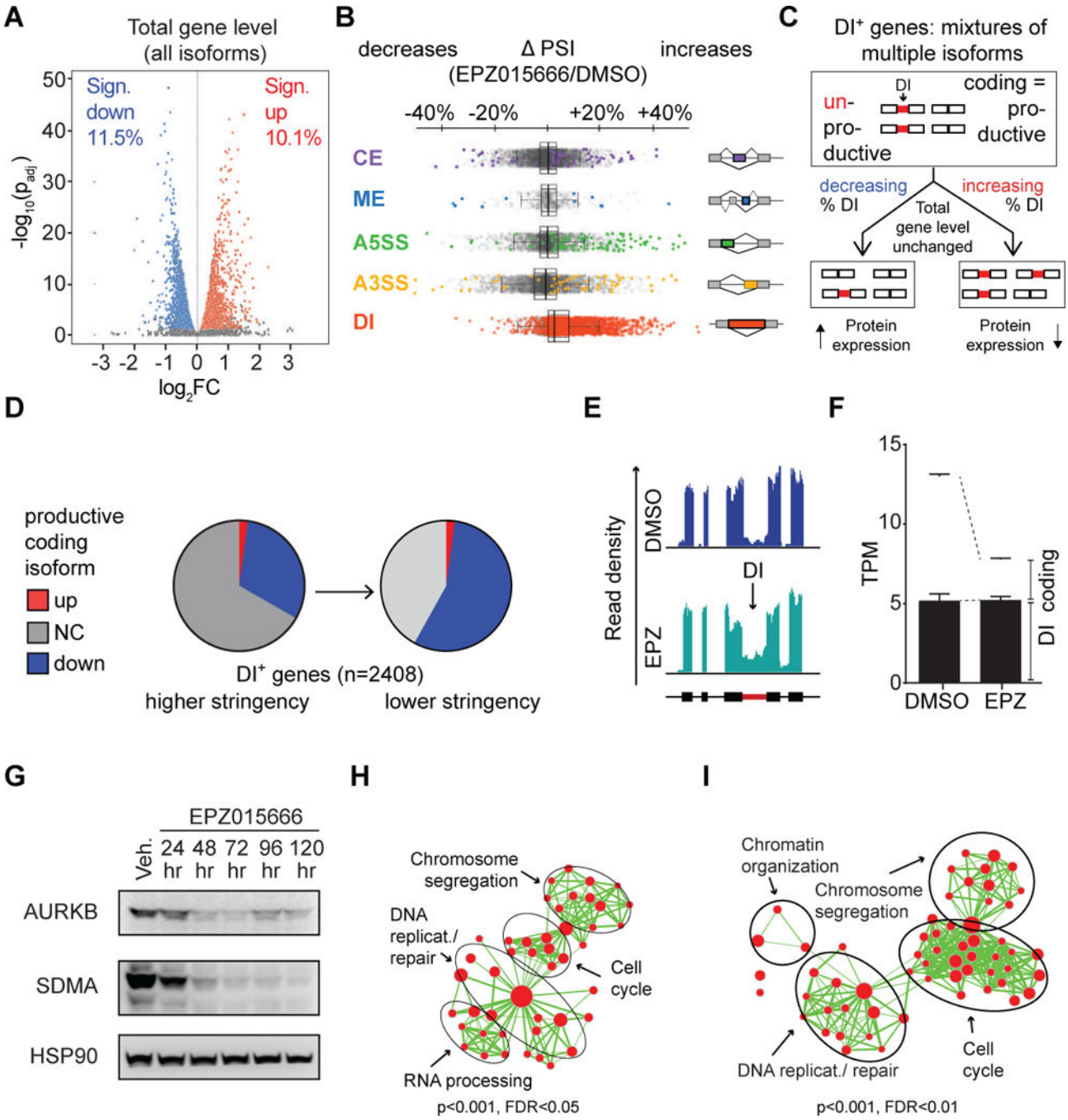


Figure 6. Increased intron detention dominates splicing changes upon PRMT5 inhibition

(**A**) Log_2-fold total gene level changes in EPZ-treated cells versus controls. (**B**) Change in percent spliced in (PSI) between EPZ and control (CE = cassette exon; ME = mutually exclusive exon; A5SS and A3SS = alternative 5′ and 3′ splice sites respectively; and DI = detained intron). Vertical lines within boxes = median, edges of boxes = 1st and 4th quartiles, whiskers = min/max values, outliers (>/< 1.5 × interquartile range) excluded. (**C**) Transcript pool of a given gene is depicted as a mix of DI negative (coding, productive) and DI positive (non-coding, unproductive) isoforms. (**D**) Proportion of significantly changed DI

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

$^{+}$ genes following subtraction of unproductive transcripts. (**E**) Read density in AURKB across DI and flanking exons, normalized to mapped reads/sample. (**F**) Transcripts per million (TPM)-normalized AURKB expression in control vs. EPZ treatment, isoform composition indicated, error bars = +/− SEM. (**G**) Immunoblot of indicated proteins in U-87 MG cells following EPZ treatment. (**H, I**) Network maps of enriched gene sets in all DI-containing genes (**H**), and all genes down-regulated following EPZ treatment (**I**). Nodes: significantly enriched gene set, node size: proportional to the gene number, width of edges: degree of gene overlap between nodes. Functionally related gene sets labeled. See also Figure S6/Table S4/Table S5/Data S1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

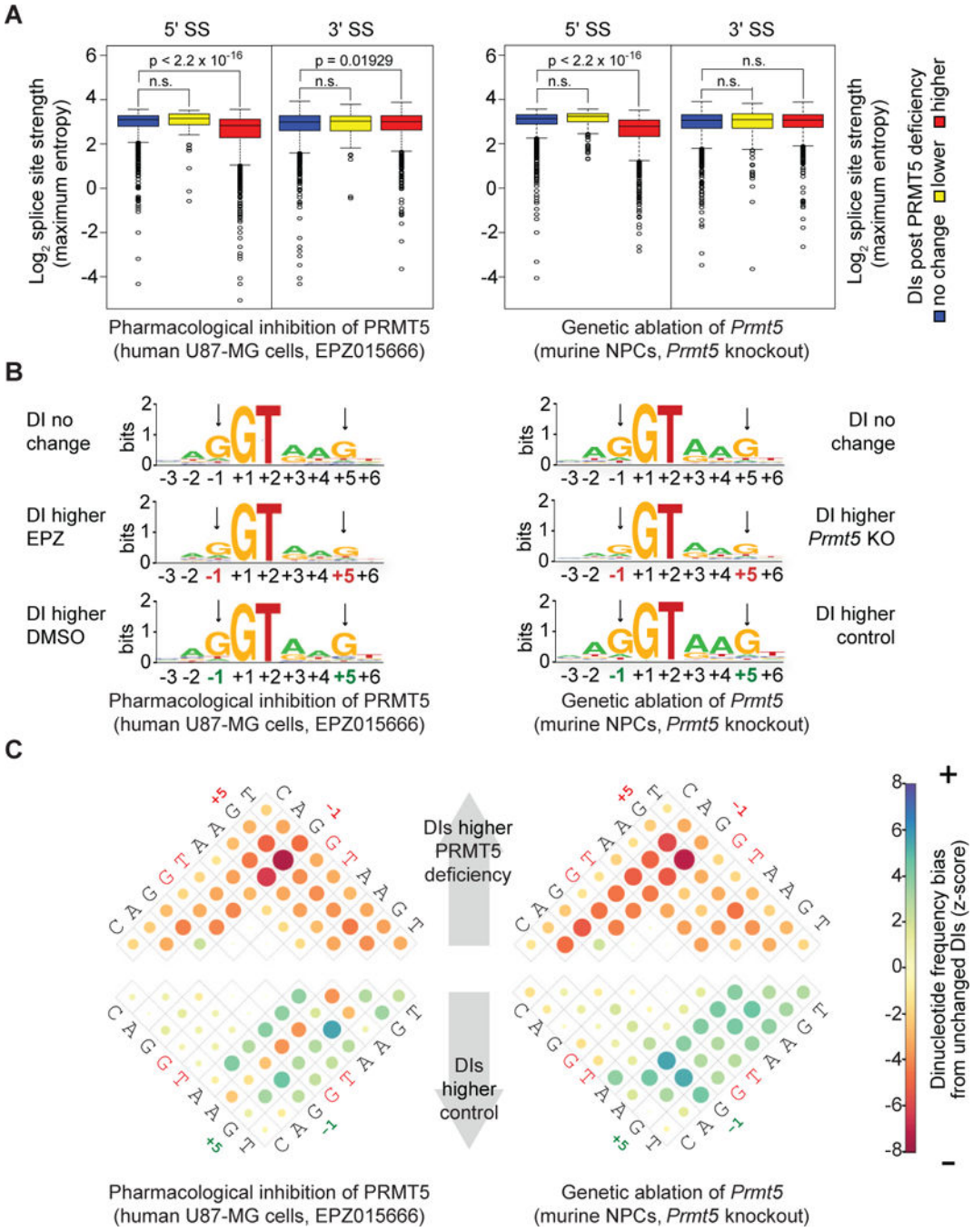


Figure 7. 5′ splice site di-nucleotide frequency differs between PRMT5 sensitive and insensitive introns

(**A**) Relative 5′ and 3′ splice site strengths of the indicated DI sets. Horizontal lines within boxes = median, edges of boxes = 1st and 4th quartile, whiskers = min/max values, outliers excluded. (**B**) Sequence logo of the 5′ splice sites of the indicated sets of DIs. (**C**) Di-nucleotide frequency biases between unchanged and changed DIs in indicated DI sets. See also Figure S7.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

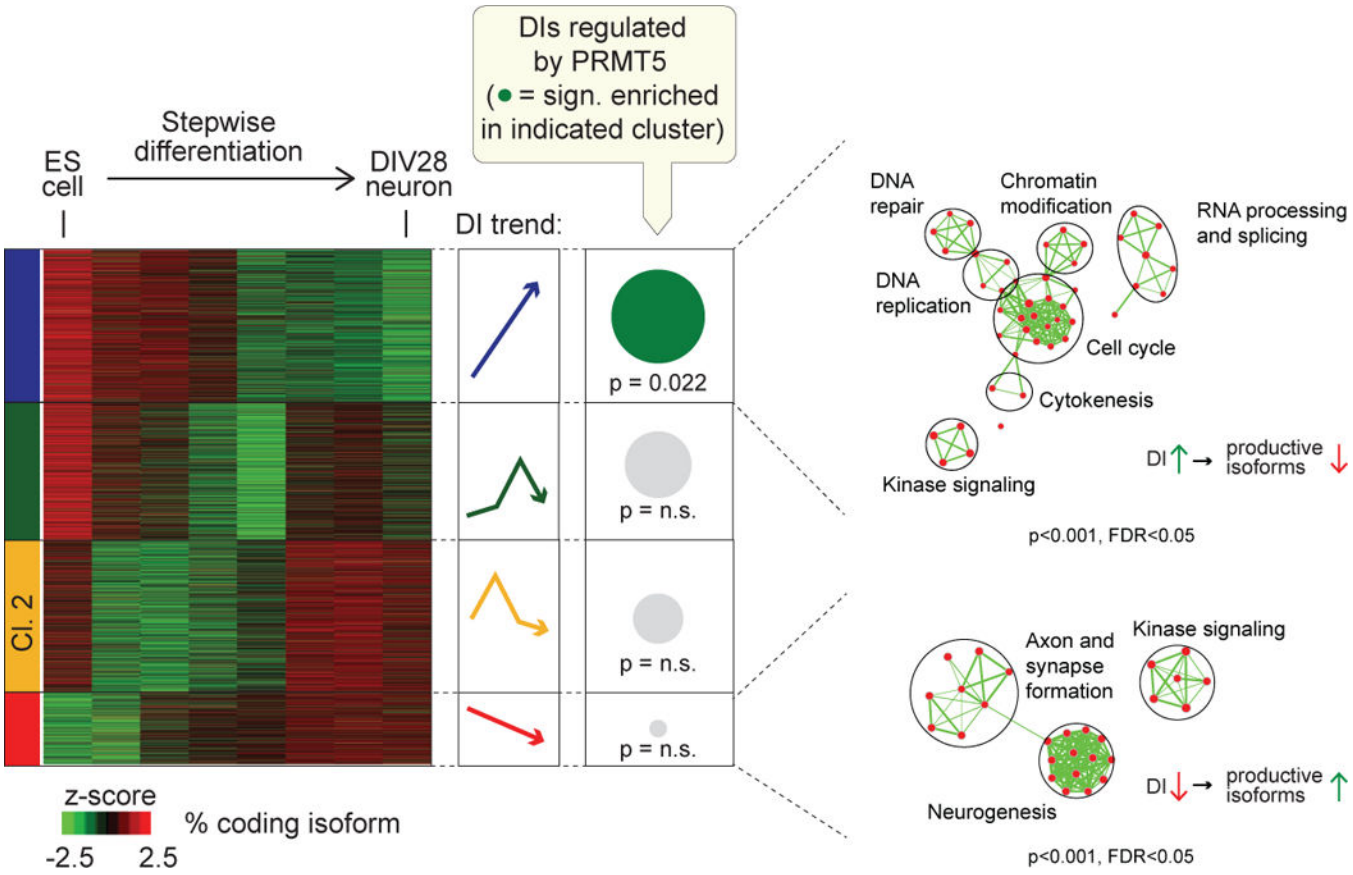


Figure 8. Detained introns are coordinately regulated during neurogenesis

Clustered, Z-score normalized coding isoform percentages. Enrichment of DIs upregulated upon *Prmt5* KO within each cluster is shown as a circle. Circle size: proportional to number of DIs, Significance: Bonferroni-corrected hypergeometric p value. DI expression trends indicated schematically. Network maps display significantly enriched gene sets within the indicated cluster. Functionally related gene sets labeled. See also Figure S8/Table S6.