

Genome-wide investigation of intragenic DNA methylation identifies *ZMIZ1* gene as a prognostic marker in glioblastoma and multiple cancer types

Dimitrios Mathios^{1*}, Taeyoung Hwang^{2,3*}, Yuanxuan Xia¹, Jillian Phallen⁴, Yuan Rui^{2,3}, Alfred P. See⁵, Russell Maxwell^{1,6}, Zineb Belcaid¹, Joshua Casaos¹, Peter C. Burger⁷, Kerrie L. McDonald⁸, Gary L. Gallia¹, Leslie Cope⁴, Mihoko Kai⁶, Henry Brem¹, Drew M. Pardoll^{4,9,10}, Patrick Ha¹¹, Jordan J. Green^{1,2,3,4}, Victor E. Velculescu^{4,10}, Chetan Bettegowda^{1,4}, Chul-Kee Park^{1,12†} and Michael Lim^{1†}

¹Department of Neurosurgery, Johns Hopkins University School of Medicine, Baltimore, MD

²Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, MD

³Lieber Institute for Brain Development, Baltimore, MD

⁴Department of Oncology and Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine, Baltimore, MD

*D.M. and T.H. equally contributed to this work

†C.K.P. and M.L. shared co-senior authorship

Author contributions: Conceptualization: Mathios D, Hwang T, Park C-K, Lim M. Methodology: Mathios D, Hwang T, Phallen J, Xia Y, Rui Y, See AP, Maxwell R, Belcaid Z, Bettegowda C, Ha P, Park C-K. Software: Hwang T, McDonald KL, Cope L, Kai M, Ha P, Park C-K. Validation: Mathios D, Hwang T, Phallen J, Xia Y, Belcaid Z, Casaos J, Burger PC, Gallia GL, Lim M. Formal analysis: Mathios D, Hwang T, Phallen J, Xia Y, Park C-K. Investigation: Mathios D, Hwang T, Phallen J, Xia Y, Rui Y, See AP, Maxwell R, Belcaid Z, Casaos J, Burger PC, McDonald KL, Park C-K, Lim M. Resources: Mathios D, Hwang T, Rui Y, Casaos J, Brem H, Pardoll DM, Ha P, Green JJ, Velculescu VE, Bettegowda C, Park C-K, Lim M. Data curation: Mathios D, Hwang T, Phallen J, Xia Y, See AP, Park C-K. Writing—original draft: Mathios D, Hwang T, Phallen J, Xia Y, Park C-K, Lim M. Writing—review & editing: Mathios D, Hwang T, Xia Y, Rui Y, Park C-K, Lim M. Visualization: Mathios D, Hwang T, Phallen J, Xia Y. Supervision: McDonald KL, Gallia GL, Cope L, Kai M, Brem H, Pardoll DM, Ha P, Green JJ, Velculescu VE, Bettegowda C, Lim M. Project administration: Mathios D, Hwang T, McDonald KL, Cope L, Kai M. Funding acquisition: Park C-K, Lim M.

Additional Supporting Information may be found in the online version of this article.

Key words: *ZMIZ1*, intragenic DNA methylation, prognostic cancer marker, glioblastoma, TCGA

Abbreviations: BLCA: bladder urothelial carcinoma; BR6: 2,2'-disulfanediylbis(ethane-2,1-diyl) diacrylate; CGI: CpG island; ChIP: chromatin immunoprecipitation; E7: 1-(3-aminopropyl)-4-methylpiperazine; ENCODE: Encyclopedia of DNA Elements; G-CIMP: glioma CpG island methylator phenotype; HPF: high power field; IDH1/2: isocitrate dehydrogenase gene 1 or 2; KIRC: kidney renal cell carcinoma; MGMT: O⁶-alkylguanine DNA alkyltransferase; RPKM: reads per kilobase of transcript per million reads mapped; SB transposon: sleeping beauty transposon; SCC: squamous cell carcinoma; siRNA: small interfering ribonucleic acid; TCGA: The Cancer Genome Atlas; U87: human primary glioblastoma cell line; UCSC: University of California, Santa Cruz; WT: wild-type; *ZMIZ1*: zinc finger MIZ-type containing 1

Conflict of interest: DM, TH, PH, CB, CKP, ML are inventors on patent US 2016/0273050 A1 titled “Use of the *zmiz1* marker in directing treatment and predicting survival in cancer” submitted by Johns Hopkins. Also, the following authors have the following non competing COI. ML: Research support: Arbor, Biohaven, BMS, Accuray, DNatrix, Tocagen Consultant: Tocagen, SQZ Technologies, VBI. Nonresearch: Consultant—Stryker. DMP: Grants and Research Support: Bristol-Myers Squibb, Melanoma Research Alliance. Consultant: Five Prime Therapeutics, Aduro, Compugen, GlaxoSmithKline, Medimmune/AstraZeneca, Merck, Potenza Therapeutics, Sanofi, Tizona, DNatrix, Amgen, Rock Springs Capital, Immunomic Therapeutics, Janssen, Astellas, WindMill Therapeutics, Bayer. HB: Research Funding from NIH, Johns Hopkins University, Arbor Pharmaceuticals, Bristol-Myers Squibb and Acuity Bio Corp* and philanthropy, Consultant for AsclepiX Therapeutics, StemGen, InSightec, Accelerating Combination Therapies*, Camden Partners*, LikeMinds, Inc*, Galen Robotics, Inc.* and Nurami Medical* (*includes equity or options). PH: Advisory Board of Bayer/Loxo Oncology, Educational Fund/Research from Stryker/Xomed. VEV is a founder of Personal Genome Diagnostics, a member of its Scientific Advisory Board and Board of Directors, and owns Personal Genome Diagnostics stock, which are subject to certain restrictions under university policy. VEV is an advisor to Takeda Pharmaceuticals. Within the last 5 years, VEV has been an advisor to Daiichi Sankyo, Janssen Diagnostics and Ignity. The terms of these arrangements are managed by Johns Hopkins University in accordance with its conflict of interest policies. CB: Advisor for Depuy-Synthes.

Grant sponsor: Private Philanthropic Grants

DOI: 10.1002/ijc.32587

History: Received 12 Mar 2019; Accepted 4 Jul 2019; Online 2 Aug 2019

Correspondence to: Michael Lim, MD, Department of Neurosurgery, Johns Hopkins Hospital, 600 N. Wolfe Street, Phipps 123, Baltimore, MD 21287, USA, Tel.: +1-410-614-1627, Fax: +1-410-502-4954, E-mail: mlim3@jhmi.edu

⁵Department of Neurosurgery, Brigham and Women's Hospital, Harvard School of Medicine, Boston, MA

⁶Department of Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins University School of Medicine, Baltimore, MD

⁷Department of Neuropathology, Johns Hopkins University School of Medicine, Baltimore, MD

⁸Cure for Life Neuro-Oncology Group, Lowy Cancer Research Centre, Prince of Wales Clinical School, University of New South Wales, Sydney, NSW, Australia

⁹Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD

¹⁰Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD

¹¹Department of Otolaryngology-Head and Neck Surgery, Johns Hopkins University School of Medicine, Baltimore, MD

¹²Department of Neurosurgery, Seoul National University College of Medicine, Seoul, Republic of Korea

DNA methylation has long been recognized as a tumor-promoting factor when aberrantly regulated in the promoter region of genes. However, the effect of intragenic DNA methylation remains poorly understood on the clinical aspects of cancer. Here, we first evaluated the significance of intragenic DNA methylation for survival outcomes of cancer patients in a genome-wide manner. Glioblastoma patients with hypermethylated intragenic regions exhibited better survival than hypomethylated patients. Enrichment analyses of intragenic DNA methylation profiles with epigenetic signatures prioritized the intragenic DNA methylation of *ZMIZ1* as a possible glioblastoma prognostic marker that is independent of *MGMT* methylation in *IDH1* wild-type patients. This intragenic region harbored molecular signatures of alternative transcription across many cell types. Furthermore, we found that the intragenic region of *ZMIZ1* can serve as a molecular marker in multiple cancers including astrocytomas, bladder cancer and renal cell carcinoma according to DNA methylation status. Finally, *in vitro* and *in vivo* experiments uncovered the role of *ZMIZ1* as a driver of tumor cell migration. Altogether, our results identify *ZMIZ1* as a prognostic marker in cancer and highlight the clinical significance of intragenic methylation in cancer.

What's new?

Aberrant methylation of the promoter region of various genes has long been recognized as a tumor-promoting factor. But what about intragenic DNA? In this genome-wide study, the authors found that intragenic DNA methylation of the *ZMIZ1* gene can predict outcomes in multiple solid cancers, especially glioblastoma, astrocytoma, bladder cancer, and renal-cell carcinoma. Evidence from independent clinical cohorts, as well as from preclinical experiments, showed that intragenic *ZMIZ1* methylation affects overall survival in cancer patients, and tumor-cell migration *in vivo*. These results indicate that *ZMIZ1* may provide a useful prognostic biomarker for various cancers.

Introduction

The role of DNA methylation in cancer pathogenesis has been investigated in multiple cancer types. In particular, the methylation of CpG islands (CGIs) in the 5' promoter region can repress the transcription of tumor suppressor genes and lead to tumorigenesis.¹ Silencing such RNA expression can, in turn, change the activation status of several cancer-related pathways.^{2–7} Therefore, DNA methylation has been used as a biomarker for treatment response and prognosis in multiple cancers.^{8–10} For example, O⁶-alkylguanine DNA alkyltransferase (*MGMT*) methylation has been established as a biomarker for positive response to temozolomide and radiation in patients with glioblastoma.^{8–10} Recent work in gliomas has also identified another hypermethylation pattern at multiple loci that is associated with improved clinical outcomes^{11,12}: the Glioma CpG Island Methylator Phenotype (G-CIMP) is believed to result from mutations in the isocitrate dehydrogenase gene 1 or 2 (*IDH1/2*), well-known genetic drivers in gliomas.^{13,14}

Previous studies on prognostic methylation markers have mostly focused on methylation signatures in the 5' promoter

regions of known genes.^{3,15} However, recent genome-wide studies of DNA methylation show that tissue- and cell type-specific methylation occur within the body of a gene more frequently than in the 5' promoter gene region.^{16–19} Motivated by this increasing importance of intragenic DNA methylation, we sought to investigate its clinical importance in predicting cancer patient outcomes. We performed genome-wide evaluation of DNA methylation using The Cancer Genome Atlas (TCGA) and identified an intragenic prognostic marker for glioblastoma: zinc finger MIZ-type containing 1 (*ZMIZ1*). DNA methylation status in the gene body of *ZMIZ1* stratified patients by survival across multiple tumors in addition to glioblastoma. Furthermore, this intragenic region is associated with the expression of a truncated version of *ZMIZ1*. Preclinical experiments confirmed the role of *ZMIZ1* as a driver of tumor cell migration. Overall, our study provides evidence for *ZMIZ1* as a promising prognostic marker in cancer and reveals the clinical importance of intragenic DNA methylation in cancer.

Materials and Methods

Genome-wide DNA methylation measurements

Genome-wide methylation data for glioblastoma samples, measured by Illumina Infinium HumanMethylation450 BeadChip (henceforth “Infinium 450K array”), were obtained from TCGA and used as a discovery set to identify molecular markers. The methylation level for a given genomic site in a sample was described as “beta (β)” values ranging from 0 to 1. See Supporting Information Materials and Methods for more details.

Identification of intragenic molecular markers

We aimed to identify CGIs showing a strong bimodal distribution of beta values as potential molecular markers. First, sites within a gene body region with standard deviations >0.1 were selected. Second, for every selected site, samples were classified into two groups labeled as “L” and “H” through hierarchical clustering of beta values using complete linkage of Euclidean distance. Third, after the median of beta values for the two groups ($\tilde{\beta}_L$ and $\tilde{\beta}_H$) was calculated, the bimodality of the distribution was assured by three criteria (Supporting Information Fig. S1): (i) $|\tilde{\beta}_H - \tilde{\beta}_L| > 0.5$, (ii) at least 10% of samples are within $\tilde{\beta}_H \pm 0.1$ and $\tilde{\beta}_L \pm 0.1$, (iii) the average deviance, defined by $\frac{1}{j} \sum_{i=1}^j \min \{ |\tilde{\beta}_{ij} - \tilde{\beta}_H|, |\tilde{\beta}_{ij} - \tilde{\beta}_L| \}$ where j denotes a sample, is <0.1 for any given site i . Fourth, sites within a CGI associated with a 5' promoter of a reference transcript were excluded to remove possible 5' promoter-associated sites. Finally, CGIs supported by at least two sites were reported as candidate intragenic molecular markers.

Classification of samples into hypo or hypermethylated groups

In order to classify samples as hypermethylated or hypomethylated, M-values, defined as the ratio of the methylated-probe (M-probe) intensity to unmethylated-probe (UM-probe) intensity,²⁰ were fitted to the mixture model of two normal distributions by an Expectation–Maximization (EM) algorithm under the condition that the identified two peak values ($\tilde{\beta}_L$ and $\tilde{\beta}_H$) were used as means. If the posterior probability that a sample was hypermethylated was $<40\%$, then a sample was called “hypomethylated” while if the probability was $>60\%$, a sample was labeled as “hypermethylated.” Otherwise, a sample was excluded from the survival analysis (Supporting Information Fig. S2).

Cell lines and *in vitro* experiments

Human primary glioblastoma cell line (U87; U-87MG RRID: CVCL_0022) cells and LN18 (LN-18 RRID:CVCL_0392) cells were obtained from ATCC (Manassas, VA) and cultured in DMEM +10% FBS + 1% penicillin/streptomycin. Both cell lines were authenticated with short tandem repeat profiling and were

mycoplasma free. For *in vitro* experiment details, see Supporting Information Materials and Methods.

Nanoparticles formulation for *in vivo* experiments

We used two types of nanoparticles for the *in vivo* experiments. The first set of nanoparticles were synthesized by combining a mixture of patented mesoporous silica nanoparticles (Lemonex Inc., Seoul, Korea) with the small interfering ribonucleic acid (siRNA) of interest for 30 min prior to injection directly into the tumor. The second set of nanoparticles was designed by one of the authors (J.J.G.). These nanoparticles have a polymeric formulation. Polymer R647 was synthesized in a similar manner as in Kozielski *et al.*²¹ Briefly, 2,2'-disulfanediylbis(ethane-2,1-diyl) diacrylate (BR6) was copolymerized with 4-amino-1-butanol (S4) *via* a Michael addition reaction and end-capped with 1-(3-aminopropyl)-4-methylpiperazine (E7) to yield polymer BR6-S4-E7 (R647). R647-siRNA nanoparticles were formed by first dissolving the polymer and siRNA, respectively, in 25 mM sodium acetate; the two solutions were then mixed and nanoparticles were allowed to self-assemble for 10 min. Nanoparticles were lyophilized using sucrose as a cryoprotectant and stored at -80°C under desiccant. Prior to injections, nanoparticles were resuspended in deionized RNAase-, DNAase- and Protease-free water to a siRNA concentration of 5 μM (with a polymer to siRNA weight/weight ratio of 150) and a final sucrose concentration of 100 mg/ml. About 200 μl of the nanoparticle solution containing 1 nmol of siRNA was injected, taking care to position the needle to only have one point for both entry and exit to minimally disrupt the tumor.

In vivo knockdown of ZMIZ1 by siRNA in U87 tumor flank model of mouse

Six to eight weeks old athymic nude mice (NU/J) were purchased from the Jackson Laboratory and kept in appropriate animal facilities. All animal protocols were confirmed by the Johns Hopkins Institutional Animal Care and Use Committee. U87 cells that were grown *in vitro* were harvested and mixed with Matrigel Basement Membrane Matrix. About 100 μl of the mix containing 100,000 cells were implanted in the flank of each mouse.

Female Nu/J mice were implanted (Day 0) with 100,000 U87 cells to establish flank tumors. Mice were anesthetized with ketamine/xylazine (100/10 mg/kg) and injected with the Matrigel/U87 cell mix subcutaneously on the left flank. The tumor take rate was 100%. Mice were followed daily to assess tumor growth and were stratified on Day 5 (when tumor was on average 250 mm^3) by tumor volume into two different groups: one receiving ZMIZ1 siRNA nanoparticles and the other receiving control siRNA nanoparticles.

After stratification, the mice were injected intratumorally with the *in vivo* siRNA (control or ZMIZ1) on Days 5, 7, 10, 13 and 17. Specifically, S32836 ZMIZ1 *in vivo* siRNA (see Supporting Information Materials and Methods) and control-scrambled siRNA were used for intratumoral injection. A solution of 1 nmol of siRNA was mixed with 80 μl of the nanoparticle mix (InViVojectionTM

RNAi-nano Red, Lemonex) and was diluted with PBS to a total volume of 120 μ l. The mix was slowly injected into tumors. At the end of the experiment, the mice were sacrificed, and their tumors were either fixed with formalin, embedded in paraffin and were stained for hematoxylin and eosin and Ki67, or were snap-frozen for RNA extraction and RT-PCR analysis. See Supporting Information Materials and Methods for more details.

DNA methylation and RNA-seq datasets in nonglioma tumors. To investigate the clinical significance of the identified marker in other tumor types, we screened additional tumor types with a sufficient number of patients analyzed with the Infinium 450K array from TCGA. We here present data for lower-grade gliomas (343 samples), renal cell carcinoma (136 samples), bladder cancer (379 samples), esophageal cancer (202 samples) and skin cutaneous melanoma (452 samples).

TCGA Level 3 RNA-seq data were downloaded for all cancer types with a sufficient number of patients (bladder cancer = 107, breast cancer = 1,152, colon cancer = 191). Reads per kilobase of transcript per million reads mapped (RPKM) values for the first exon (chr10: 81003425–81,003,495: +, UCSC hg19) of the alternative *ZMIZ1* transcript (UCSC ID: uc001kag.2) were used to assess RNA expression associated with the identified molecular marker. A cut-off value of 1 standard deviation (SD) above the mean RPKM was used to stratify patients into the highly expressing uc001kag.2 group for all cancer types.

Data availability. The original TCGA data that support the findings of our study are available in the NCI GDC Data portal repository at the following URL: <https://portal.gdc.cancer.gov/repository>.

Results

The potential of intragenic DNA methylation for classifying glioblastoma, and its association with gene regulatory signatures

We first sought to identify genomic sites that exhibit differential methylation in glioblastoma from genome-wide DNA methylation profiles. We specifically focused on National Center for Biotechnology Information (NCBI) reference gene (RefSeq)-annotated CGI regions and identified 30,428 variable sites based on standard deviations of methylation levels across samples. Next, hierarchical clustering analysis was performed and revealed a subgroup of seven patients, mainly with *IDH1* mutations (6/7) or G-CIMP positive signatures (6/7; Supporting Information Figs. S3a and S3b). This subgroup of patients showed significantly improved overall survival compared to subgroups of non-G-CIMP or non-*IDH1*-mutated patients, confirming previously reported results.^{11,12}

Interestingly, we noted that the selected 30,428 sites were enriched in intragenic regions rather than 5' promoter regions: approximately 55% of the selected sites were intragenic while only 40% of the total CGI-associated sites were intergenic (Fisher's exact test $p < 10^{-15}$). We also observed that the intragenic sites alone

could identify the G-CIMP positive subgroup of glioblastoma patients (Figs. 1a and 1b). Moreover, these highly variable intragenic sites were significantly enriched with various epigenetic marks identified by the Encyclopedia of DNA Elements (ENCODE) project. For our analysis, we used data from both neural progenitors and other cell types. These enriched epigenetic markers included a regulatory signature (DNase-I hypersensitivity), two known intragenic histone markers (H3K36me3 and H3K79me2), an enhancer signature (H3K27ac), and a promoter/genic signature such as polymerase II (POLR2A)-binding sites (Fig. 1c).

Due to the strong contribution of *IDH1* in the G-CIMP phenotype, we excluded the *IDH1* mutated patient to rerun our analysis. We identified that the non-*IDH1* mutated glioblastoma patients form a hypermethylation cluster that shows a trend for improved survival compared to the nonhypermethylated glioblastoma patients ($p = 0.174$, Supporting Information Figs. S3c and S3d).

Identification of a truncated *ZMIZ1*-associated probe as an intragenic molecular marker for glioblastoma classification

The ability of the intragenic sites to discriminate survival status among glioblastoma patients through their association with regulatory signatures motivated us to identify specific intragenic markers. We first focused on CGIs harboring intragenic sites with strong bimodal distributions of beta values, based on the observation that glioblastoma patients tended to exhibit either hypo or hypermethylation at highly variable sites. Using a defined set of criteria (see Materials and Methods), 19 CGIs were identified with potential molecular markers (Supporting Information Table S1). Among the identified CGIs, we selected the ones associated with known RNA transcripts, resulting in only the *ZMIZ1* region. This region is near the genomic coordinate chr10: 81,002,109–81,003,687, which is not associated with the classical transcript of *ZMIZ1* in the RefSeq database²² but is related to an alternative transcript found in the UCSC mRNA database²³ (Fig. 2a). Consistently, the two variable probes in the intragenic region of *ZMIZ1* (*cg26654807* and *cg14371731*, as shown in Supporting Information Table S1) were associated with histone markers of transcriptional regulation across various cell lines of the ENCODE project (Supporting Information Table S2), suggesting that this region can function as an alternative promoter for transcription of the truncated *ZMIZ1* isoform. Finally, we chose the probe *cg26654807* instead of *cg14371731* as a representative molecular marker in the intragenic region of *ZMIZ1* because *cg26654807* was more associated with various transcription cofactors, implying a stronger regulatory potential (Fig. 2b).

In order to evaluate the prognostic ability of this molecular marker, the beta values of the *cg26654807* probe were modeled in a constrained normal mixture model (see Materials and Methods) to classify glioblastoma TCGA samples as hypermethylated or hypomethylated. In Kaplan–Meier survival analysis, hypermethylated samples exhibited a survival advantage over hypomethylated samples ($n = 117$, $p = 0.0094$, Fig. 2c). To control for confounding factors, univariate analyses were performed for age, gender, ethnic, extent of surgical resection, type of adjuvant treatment

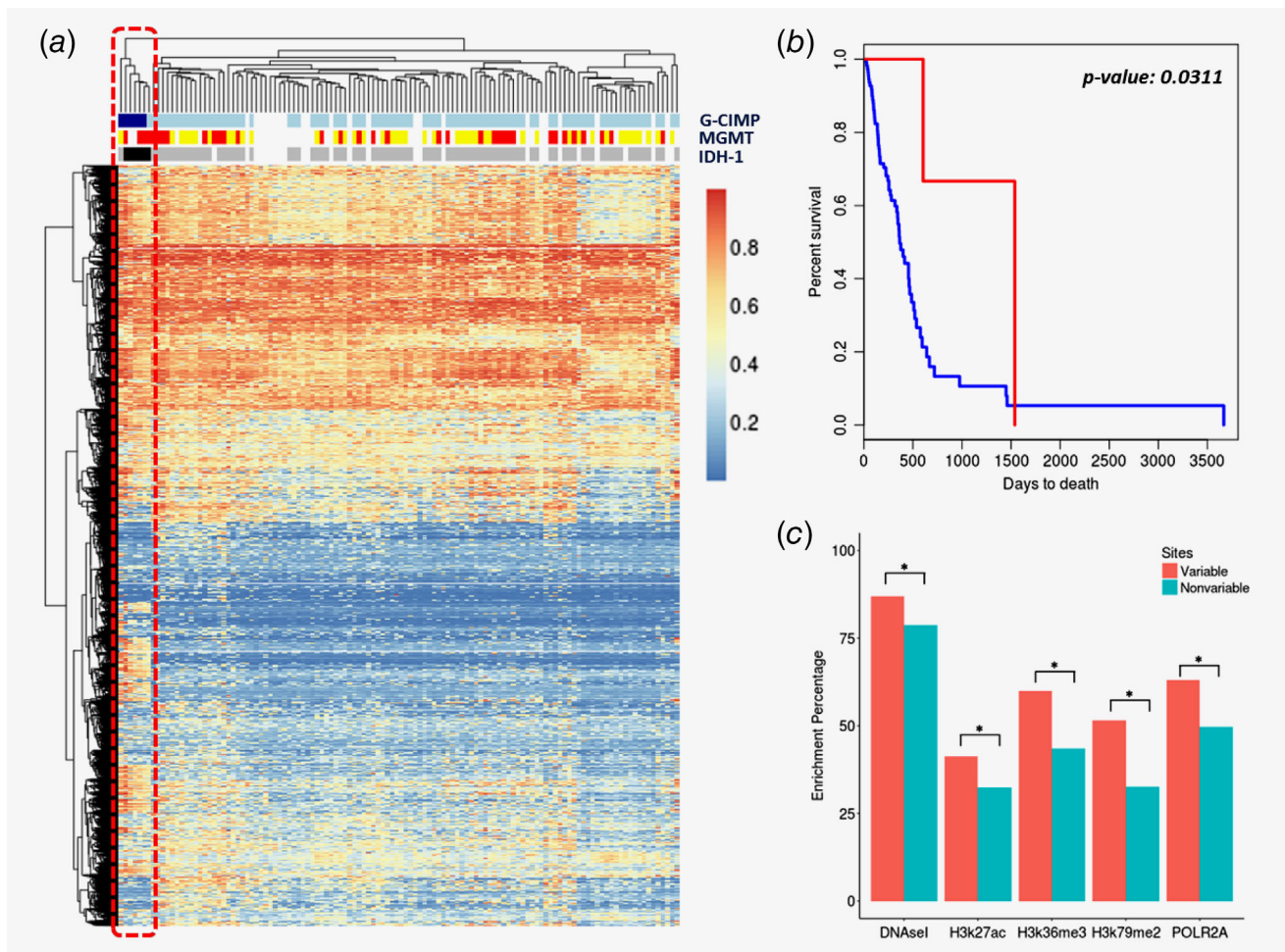


Figure 1. Classification of glioblastoma patients based on intragenic DNA methylation. (a) Hierarchical clustering of glioblastoma patients (columns) and the variable sites in DNA methylation level (rows) measured as beta values ranged from 0 to 1. The variable sites were selected based on standard deviation of beta values among the intragenic sites in the RefSeq-annotated CGI regions. The color gradients from blue to red correspond to beta values from 0 to 1. The known association of G-CIMP, the *MGMT* methylation and the *IDH1* mutation were marked as dark colors in indication bars. The identified subgroup of seven patients was denoted by the red square. (b) Kaplan–Meier survival analysis for two subgroups identified in (a) shows a significantly improved survival for patients with hypermethylation (red, corresponding to the red square in (a)) compared to hypomethylated patients (blue). (c) The majority of variable intragenic sites are significantly enriched with various epigenetic marks annotated by ENCODE, including DNase I hypersensitivity (regulator region), H3K27ac (enhancer), H3K36me3 (structural mark), H3K79me2 (structural mark) and Polymerase II binding sites. [Color figure can be viewed at wileyonlinelibrary.com]

(chemotherapy and radiation), intensity of treatment, G-CIMP status, *IDH1* mutation status and *MGMT* methylation status. When we adjusted for significant factors from these univariate analyses ($p < 0.05$), *ZMIZ1* methylation status trended toward significance ($p = 0.066$) while *IDH1* mutation ($p = 0.929$) and *MGMT* methylation status ($p = 0.990$) were not significant (Supporting Information Table S3). In addition, considering that *IDH1* mutation is known to cause the G-CIMP phenotype¹⁴ and confer a survival benefit in glioblastoma patients, we evaluated our *cg26654807* marker with *IDH1* wild-type (WT) cases ($n = 85$). There, hypermethylated cases had improved survival compared to hypomethylated cases ($n = 85$, $p = 0.0257$, Fig. 2d). Further analyses with glioblastoma patients showed significant improvement in terms of progression to death after radiation treatment and

progression to death after chemotherapy ($n = 76$, $p = 0.0061$ and $n = 51$, $p = 0.014$, respectively, Supporting Information Fig. S4).

We also validated our molecular marker, *cg26654807*, with a glioblastoma patient cohort independent from TCGA. A group of adult German glioblastoma patients (age ≥ 18)²⁴ was used to validate our results. The samples were labeled as “hypomethylated” if the beta value of *cg26654807* is < 0.4 , while “hypermethylated” was assigned to the beta value > 0.6 as was done for the TCGA patients. A significant survival improvement was confirmed for hypermethylated patients ($n = 38$, $p = 0.0015$, Fig. 2e). The survival differences became moderate when we considered the samples only with *IDH1* WT ($n = 33$, $p = 0.0878$, Fig. 2f) presumably due to a much smaller number of patients compared to the TCGA patients.

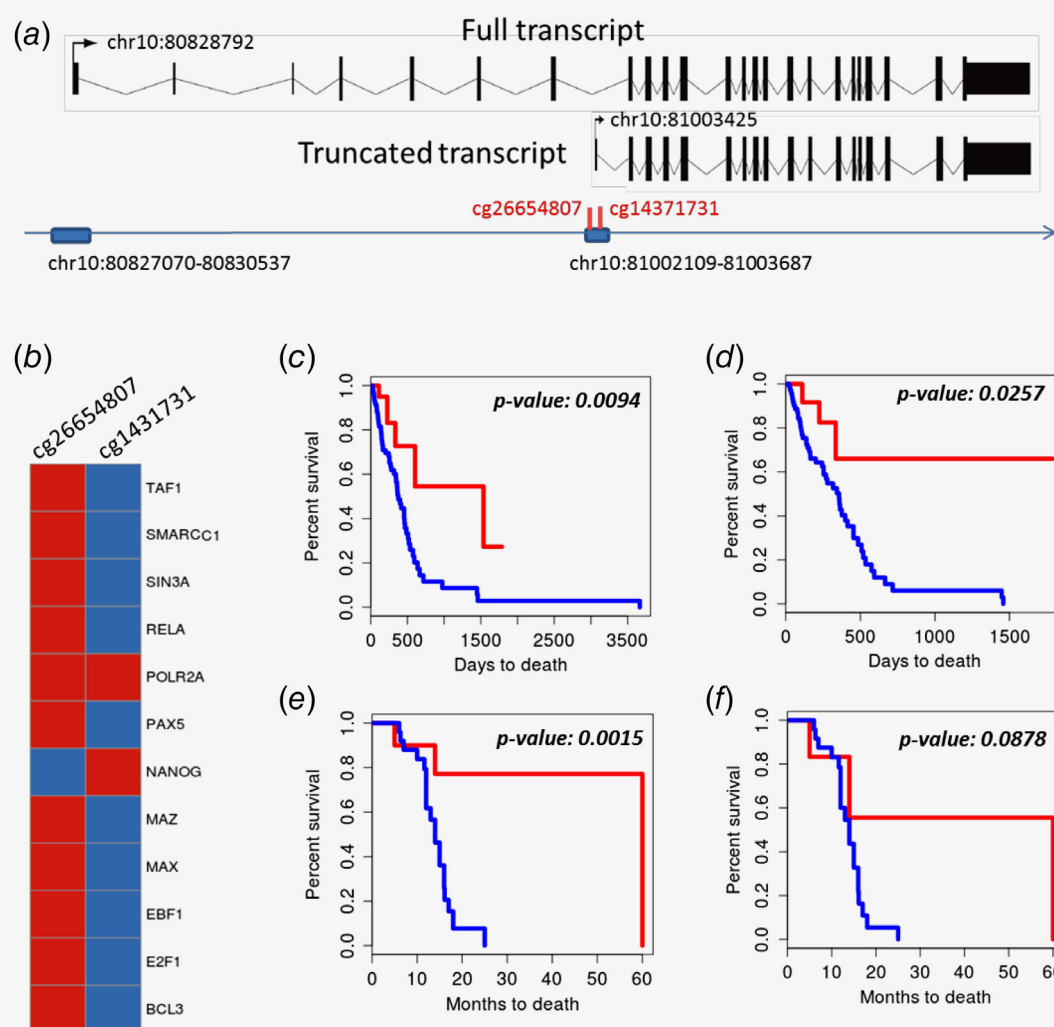


Figure 2. ZMIZ1-associated molecular marker for glioblastoma classification. (a) The schematic diagram of *ZMIZ1* gene (chr10:80,828,792–81,076,285, UCSC hg19) depicting two alternative transcripts of *ZMIZ1* and highlighting the full transcript (UCSC ID: uc001kaf.2) and the truncated isoform (UCSC ID: uc001kag.2). Two CpG islands (denoted by blue rectangles on the arrow with the genomic coordinates) span the promoter region of the two transcripts. In the CGI associated with the truncated transcript, we identified two probes (two red bars) showing a bimodal distribution of beta values in the TCGA glioblastoma samples. (b) The two probes in (a) are assessed by their associations with transcription factors using ENCODE chromatin immunoprecipitation (ChIP)-seq data (red: significant association, blue: no association). (c, d) Kaplan–Meier survival analyses were performed with TCGA samples including *IDH1* mutation (c) and excluding *IDH1* mutation samples (d). Red denotes the group of hypermethylated cases while blue indicates the hypomethylated case based on the probe *cg26654807*. (e, f) Kaplan–Meier survival analyses were performed with independent German samples including *IDH1* mutation (e) and excluding *IDH1* mutation samples (f). Color coding is the same as in c, d: red is a hypermethylated group, blue is a hypomethylated group. Sample sizes for (c)–(f) are (c) total = 117, red = 20, blue = 97; (d) total = 85, red = 12, blue = 73; (e) total = 38, red = 10, blue = 28; (f) total = 33, red = 6, blue = 27. [Color figure can be viewed at wileyonlinelibrary.com]

Functional significance of ZMIZ1 in glioblastoma

The prognostic potential of the *ZMIZ1* locus suggests that *ZMIZ1* may have a key biological role in glioblastoma. We first investigated the *in vitro* effect of *ZMIZ1* knock-down or overexpression in glioblastoma cell lines. Specifically, we surveyed the effect of *ZMIZ1* in migration and viability of the established U87 and LN18 cell lines (Figs. 3a–3d and Supporting Information Figs. S5a–S5c). Two siRNAs were used to differentiate knockdown effects of the full

transcript and alternative truncated transcript (see Supporting Information Materials and Methods and Fig. S6). *ZMIZ1* knockdown led to a decrease in migration while *ZMIZ1* ectopic expression (both full and truncated transcript) led to a significant increase in the migratory potential of glioblastoma cells (Fig. 3d and Supporting Information Fig. S5c). However, there was little effect on apoptosis (Fig. 3a and Supporting Information Fig. S5a), proliferation (Fig. 3b and Supporting Information Fig. S5b) or viability (Fig. 3c).

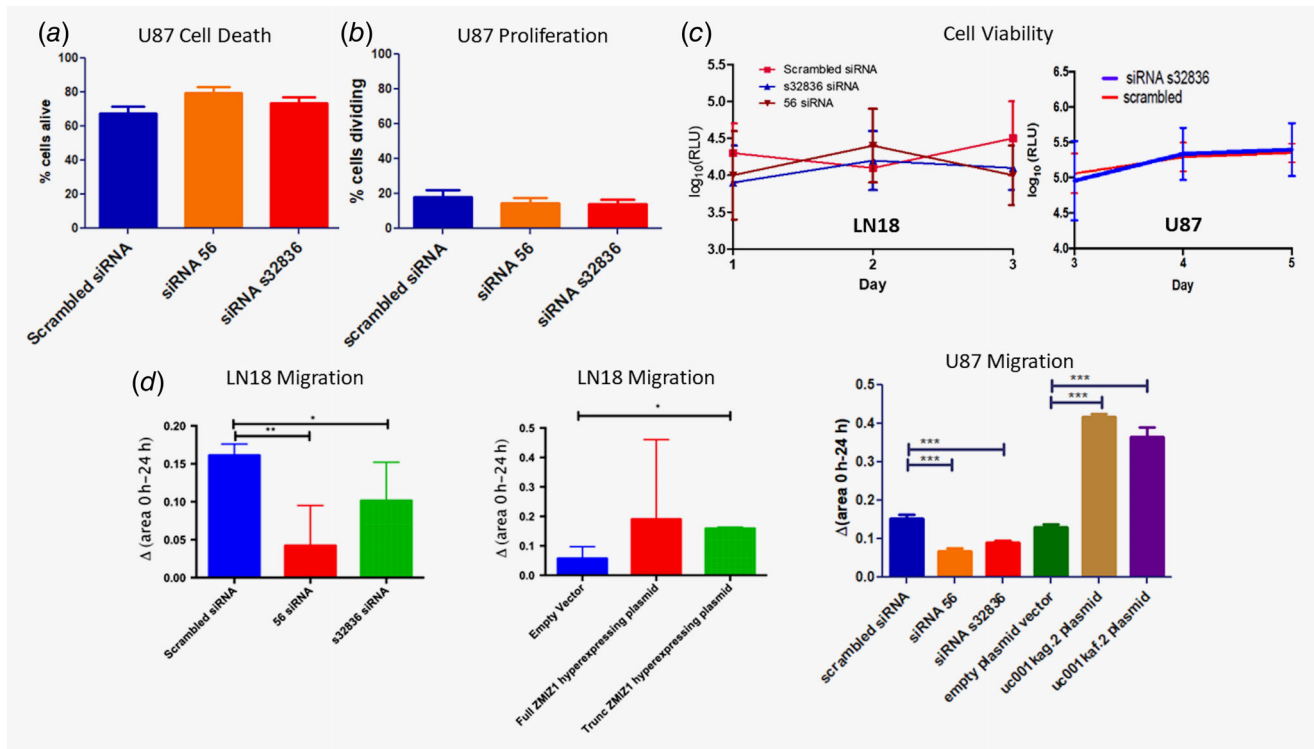


Figure 3. *In vitro* function of *ZMIZ1* gene in glioblastoma cell lines. (a–c) *In vitro* cell death (a), proliferation (b) and viability (c) assays with U87 cell line. All results here are reported after single treatment with siRNA and observation of the cells in a 24-hr culture. Knocking down *ZMIZ1* transcripts were mediated by siRNA: 56 for truncated transcript, s32836 for full transcript and scrambled for control. The cell death assay was done by measuring the percent of cells that were annexin V negative, PI negative after dual staining ($n = 3$, error bars denote SD). The proliferation of cells was measured by the dilution of carboxylfluorescein succinimidylester dye ($n = 3$, error bars denote SD). The viability assay was performed after mixing luminescence dye in the cell culture and reading photon intensity emission with a luminometer ($n = 12$, error bars denote SD). (d) The migration of cells is reported as area covered in the scratch from the creation of the “wound” (time 0) until its closure (24 hr). Overexpression of *ZMIZ1* was achieved by transfection of plasmid vector (either the full transcript [uc001kaf.2] or the truncated transcript [uc001kag.2]; $n = 6$, error bars denote SD). All experiments were performed at least twice. [Color figure can be viewed at wileyonlinelibrary.com]

Next, we tested whether *in vivo* siRNA knockdown of *ZMIZ1* affects tumor growth. We used a U87 flank tumor model since U87 shows high expression of *ZMIZ1* (see Materials and Methods). While control mice (scrambled siRNA-treated mice) showed exponential tumor growth, *ZMIZ1* siRNA-treated mice had decreased rates of tumor growth (Fig. 4a, $p < 0.0001$). These results confirmed the positive effect of *ZMIZ1* expression on tumor growth *in vivo*, suggesting the importance of *ZMIZ1* regulation for glioblastoma. We further analyzed U87 flank tumors injected with either *ZMIZ1* siRNA or scrambled siRNA to identify the level of knockdown achieved from delivery of our nanoparticle vector (Fig. 4b) and elucidate a potential *in vivo* mechanism for the effect of *ZMIZ1* on tumor growth. Ki67 index was used as a measure of *in vivo* proliferation potential and, in accordance with our *in vitro* data, *ZMIZ1* knockdown had no effect on proliferation potential ($p = 0.18$, average Scrambled vs. *ZMIZ1* siRNA: 0.95 vs. 0.96, Fig. 4c, Supporting Information Fig. S7a). To explain the differences in tumor sizes observed *in vivo*, we calculated cell density by counting number of cells per high power field (HPF) in the two studied groups.

We found significant differences in the cell density of the studied groups ($p < 0.001$) with the average cell density in the Scrambled siRNA group estimated at 135 cells/HPF vs. *ZMIZ1* siRNA group estimated at 171 cells/HPF (Fig. 4d, Supporting Information Fig. S7b). This increased cell density in the absence of proliferation rate changes implies that knockdown of *ZMIZ1* leads to reduced tumor cell migration (Fig. 5).

ZMIZ1 as a molecular marker in multiple tumor types

To further assess the prognostic significance of *ZMIZ1* as a molecular biomarker, we analyzed additional TCGA tumors with adequate numbers of patients whose DNA methylation of the probe cg26654807 or RNA expression of the truncated isoform were assessed with the Infinium 450k array or RNA-sequencing (see Materials and Methods).

Low-grade glioma. We extended our analysis of *ZMIZ1* methylation status to low-grade glioma (LGGs) in adult patients (Grade II and III). Astrocytomas, but not oligodendrogliomas, showed survival differences in *ZMIZ1* hypermethylated cases compared to

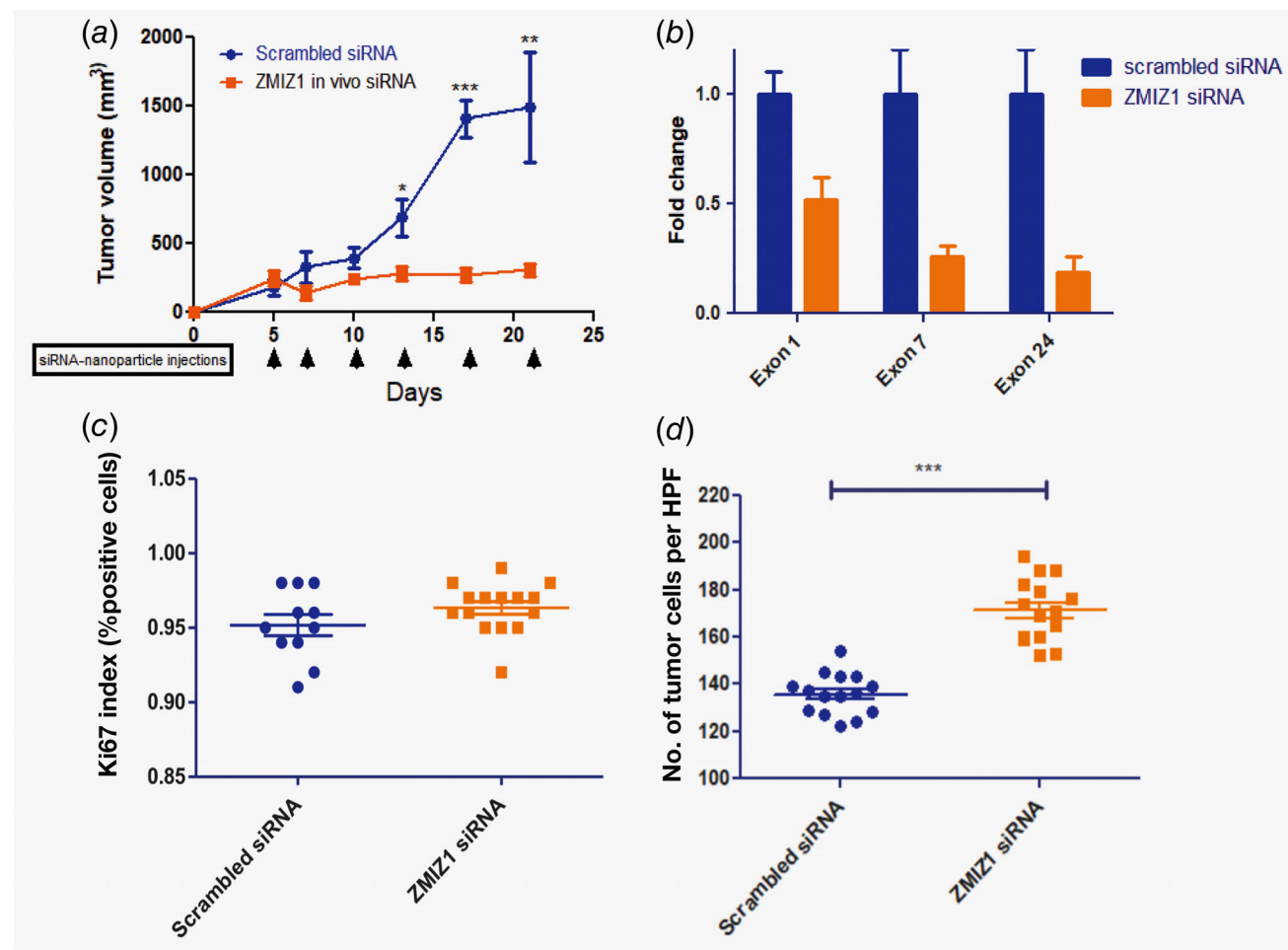


Figure 4. *In vivo* effect of ZMIZ1 knockdown and mechanism of function in glioblastoma. (a) After implanting U87 human glioma cells in athymic/nude mice, we monitored tumor growth after injection of either scrambled siRNA–nanoparticle complex or ZMIZ1 siRNA–nanoparticle complex intratumorally three times weekly for 2 weeks. Flank tumor size calculation was based on measuring three orthogonal diameters of the tumor by calipers. The blue line denotes the scrambled siRNA group while the orange line the 532836 siRNA treated group. $n = 10$, Error bars denote SD. (b) RT-PCR analysis of the mouse tumors was performed to assess percent of knockdown of ZMIZ1 transcript *in vivo*. On average 50% knockdown was observed at the first exon of the transcript and on average 80% knockdown close to the functional domain of the ZMIZ1 protein. $n = 9$, Error bars denote upper and lower limit of the observed values. (c) Multiple areas of maximal viable tumor were compared in scrambled siRNA group and ZMIZ1 siRNA group with regards to the number of ki67 positive cells/HPF and no differences were observed in the proliferation rate of these tumors (in agreement with the *in vitro* data). $n = 15$, Error bars denote SD. (d) Significant difference ($p < 0.0001$) was observed in the cell density of the two groups of mice, with ZMIZ1 siRNA treated mice having higher cell density compared to the scrambled siRNA tumors. $n = 15$, Error bars denote SD. [Color figure can be viewed at wileyonlinelibrary.com]

hypomethylated cases ($p < 0.001$; Figs. 6a and 6c). Multivariate analysis including age, sex, grade, ZMIZ1 methylation status and IDH1 mutation, showed ZMIZ1 hypermethylation was an independent marker of improved survival in low grade astrocytomas. ZMIZ1 reach statistical significance ($p = 0.015$) whereas IDH1 status did not reach statistical significance ($p_{IDH1} = 0.329$).

Bladder urothelial carcinoma, kidney renal cell carcinoma, colorectal cancer and skin cutaneous melanoma. Bladder urothelial carcinoma (BLCA) samples demonstrated a bimodal distribution for the cg26654807 probe ($n = 434$). For survival analyses, we used 379 samples after removing Stage I samples and Hispanic samples in order to focus on high stage BLCA. The hypermethylated

group showed improved survival over the hypomethylated group ($p = 0.042$; Figs. 5b and 5c). Even after multivariate analysis using age, gender, ethnic and tumor stage, ZMIZ1 methylation remained an independent predictor of survival ($p = 0.028$; Fig. 6b). Kidney renal cell carcinoma (KIRC) samples ($n = 480$) exhibited unimodal distribution of DNA methylation for the cg26654807 probe. Here, we categorized KIRC samples into two groups by using simple thresholds: “hypermethylated” samples had DNA methylation levels of at least the average + 1 standard deviation while “hypomethylated” samples had levels at most the average – 1 SD. After removing Asian samples to avoid a sparse group ($n = 2$), univariate and multivariate analyses of hypermethylated ZMIZ1 cases in KIRC demonstrated significant differences in survival ($p = 0.043, 0.003$,

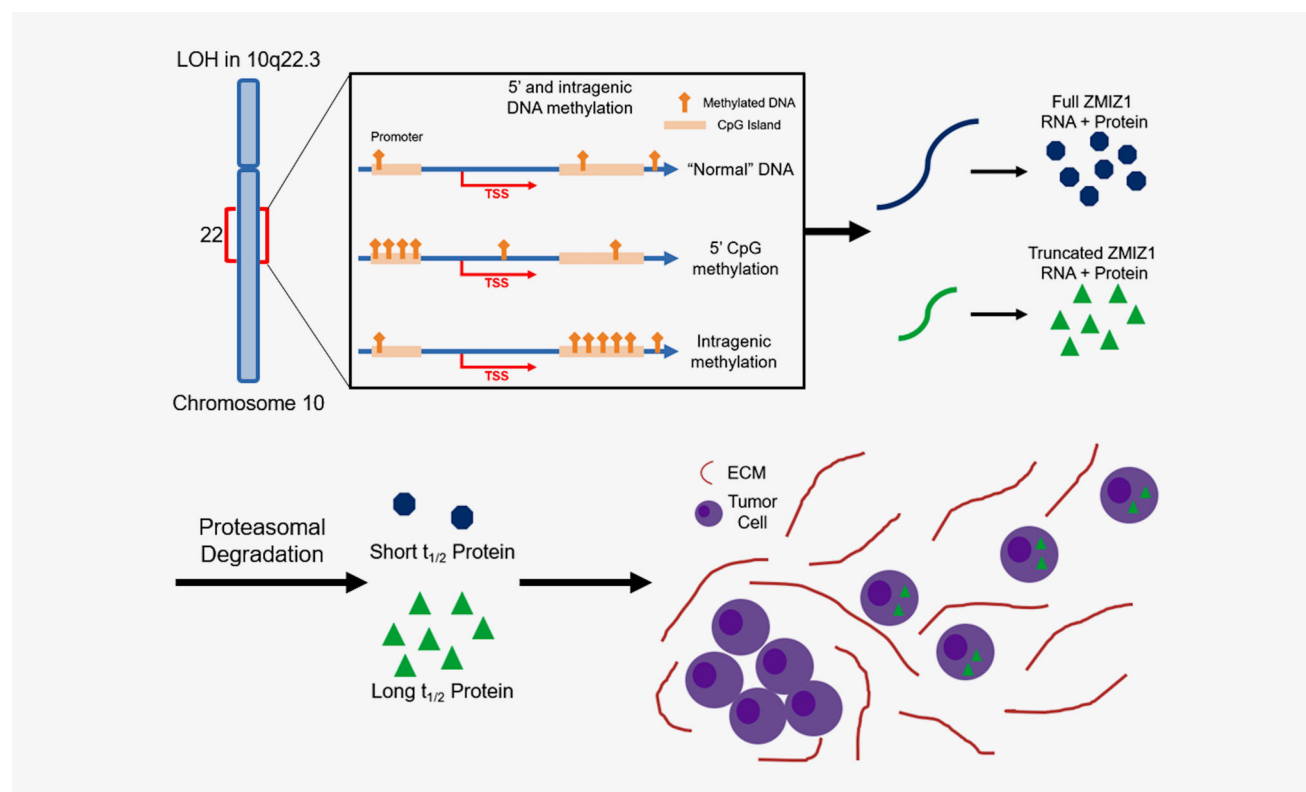


Figure 5. Proposed mechanism of action for effect of *ZMIZ1* intragenic methylation on cancer. The *ZMIZ1* gene is located on the chromosomal region 10q22 which in glioblastomas exhibits loss of heterozygosity in the great majority of the samples. Intragenic methylation in this region can delete the function of the second allele of *ZMIZ1*. When the intragenic locus remains unmethylated, the truncated *ZMIZ1* RNA transcript (green) is upregulated. This translates to a more stable, truncated protein (green triangles) that undergoes proteasomal degradation at a slower rate. Cancer cells that overexpress this version of the protein display increased migration, leading to the worse clinical cancer outcomes observed in analyses of multiple clinical genomic databases. [Color figure can be viewed at wileyonlinelibrary.com]

respectively, Fig. 6c, Supporting Information Fig. S8a). Skin cutaneous melanoma (SKCM) showed a similar pattern but without significance ($p = 0.199$, Supporting Information Fig. S8b).

RNA expression analysis in glioblastoma, breast cancer, BLCA and colon adenocarcinoma. Since we found that intragenic DNA methylation associated with alternative transcripts may serve as a marker of survival, we explored the relationship between DNA methylation and the expression level of the associated truncated transcript in glioblastoma. To investigate this, we collected glioblastoma samples from TCGA that have both RNA-seq and DNA methylation data. Using 58 TCGA glioblastoma samples, we calculated the Spearman correlation coefficient between the expression level of the truncated transcript and the DNA methylation level (beta): $\rho \approx -0.06$ with no statistical significance ($p = 0.6714$). However, when we compared the DNA methylation level based on the *cg26654807* probe to the proportion of truncated *ZMIZ1* transcript relative to total amount of *ZMIZ1* transcripts (sum of full transcript and truncated transcript), there was a significant negative correlation ($\rho \approx -0.3$ with $p = 0.0236$). Therefore, the survival advantage we observed in hypermethylated samples may be caused by a reduced relative level of truncated transcript that is known to be more stable than the full transcript.

Since *ZMIZ1* is known to affect androgen receptors, we attempted to evaluate the effect of *ZMIZ1* in breast cancer and prostate cancer cohorts. Since most breast cancer (BRCA) samples had low DNA methylation levels based on our probe (average beta: 0.05), classifying the samples into two groups was not feasible. Additionally, there were few vital events in prostate cancer in TCGA (10 Dead and 489 Alive), precluding any reliable estimation of survival advantage. However, the RNA expression level associated with the truncated isoform of *ZMIZ1* was higher in advanced stages of BRCA (III/IV) compared to earlier stages (I/II), as shown in Supporting Information Figure S8c. Furthermore, BRCA, BLCA and colon adenocarcinoma all demonstrated varying trends for improved survival at lower expression compared to higher expression levels (Supporting Information Figs. S8d–S8f), implying that higher RNA levels of the truncated *ZMIZ1* isoform negatively affects cancer patient survival.

Discussion

In our study, we sought to identify the clinical significance of intragenic methylation in cancer patients. We first evaluated the prognostic potential of intragenic DNA methylation at the genome-wide level for glioblastoma. Interestingly, we found that intragenic sites are enriched for highly variable methylated sites

(sites with high intertumoral heterogeneity in methylation across the studied patient population) compared to the 5' promoter gene region and can form a G-CIMP phenotype similar to what has been described before.¹¹ This appears to be the case in non-IDH-1 mutated samples as well. These intragenic sites are annotated with various types of epigenetic marks as well as RNA Polymerase II binding. Based on this evidence, we hypothesized that intragenic regions with strong regulatory signatures could be used as novel biomarkers for glioblastoma survival. By applying computational filters and integrating epigenetic profiles, we identified a DNA methylation probe in an alternative promoter region of *ZMIZ1* as a novel prognostic marker for glioblastoma. *ZMIZ1* methylation status was independently linked to improved survival in *IDH1* WT patients and was superior in comparison to *MGMT* methylation status (Supporting Information Table S3).

The prognostic significance of the alternative promoter region of *ZMIZ1* is further highlighted by its applicability to multiple tumor types. DNA methylation of the region correlated with distinct survival differences in astrocytoma, bladder cancer and renal cell carcinoma (Fig. 6c). These results suggested that the truncated *ZMIZ1* transcript may play more than a prognostic role and have

some biological relevance. Indeed, Rogers *et al.* recently noted that *ZMIZ1* is commonly mutated in squamous cell carcinoma (SCC) in an experimental model of the Sleeping Beauty (SB) transposon mutagenesis screen.²⁵ The mutation inserted by the SB transposon leads to a truncated *ZMIZ1* protein that exhibits greater stability and half-life in addition to spontaneous SCC in mice upon ectopic expression. It is possible that DNA methylation of the alternative promoter affects the amount of resulting protein. In fact, the *cg26654807* methylation probe that exhibited a prognostic significance in the tumors we studied here corresponds to the alternative promoter CGI of the truncated *ZMIZ1* transcript identified by Rogers *et al.*²⁵

In addition to our finding of *in vivo* role of *ZMIZ1* in tumor cell migration, there have been several reports about the importance of *ZMIZ1* in cancer. It has been shown to synergize with Notch homolog 1 (*NOTCH1*) to induce T cell acute lymphoblastic leukemia in mice, data pointing to a crucial role of *ZMIZ1* in cancer pathogenesis.²⁶ A low percentage of rearrangements that lead to a fusion protein *ZMIZ1*-Abelson murine leukemia viral oncogene homolog 1 have been identified as a distinct feature of B cell acute lymphoblastic leukemia.²⁷ While there is little known about the

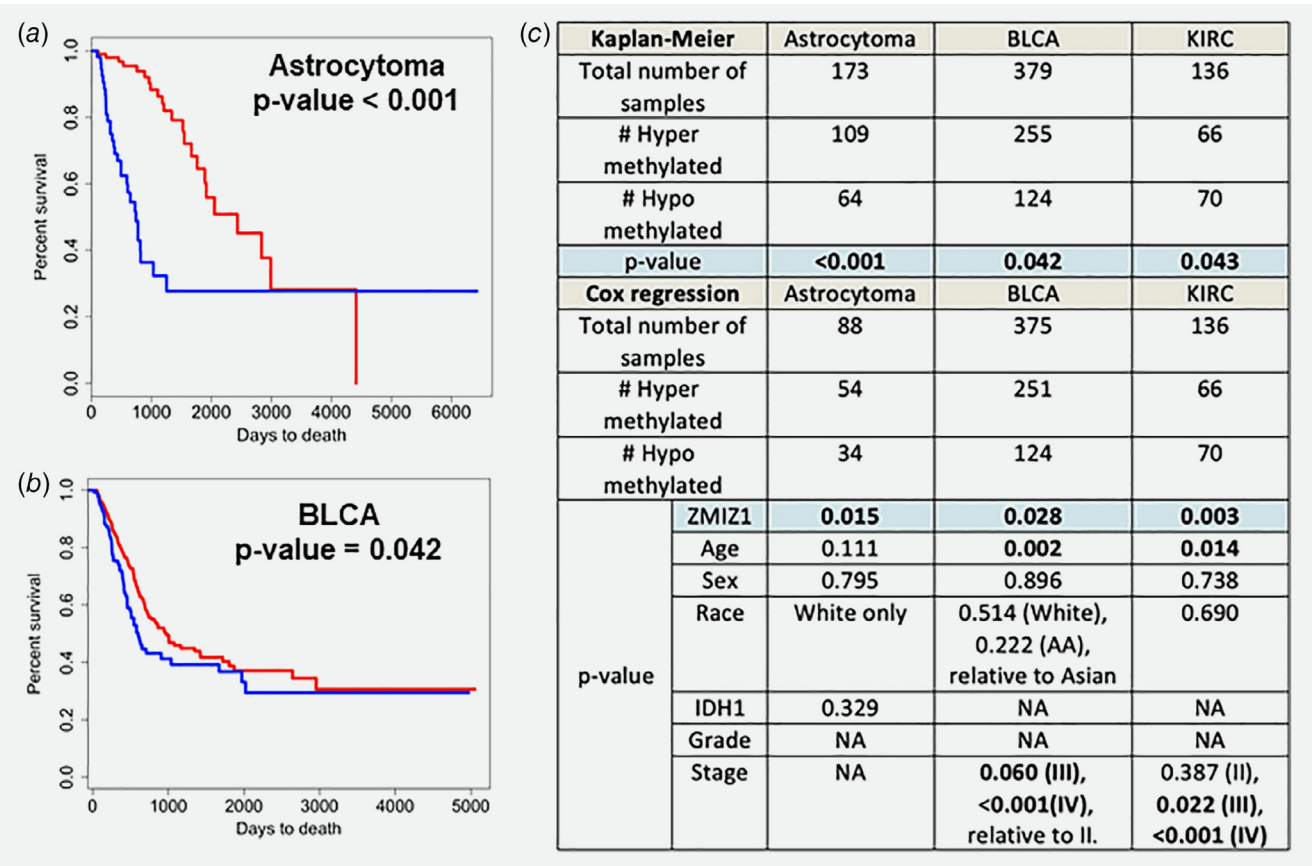


Figure 6. Survival analyses of multiple tumor types from TCGA. (a, b) The methylation at the alternative promoter of *ZMIZ1* (*cg26654807*) stratifies patients with (a) astrocytoma ($p < 0.001$), (b) bladder cancer (BLCA; $p = 0.042$). Red denotes patients with high methylation and blue represents patients with low methylation, (c) *ZMIZ1* methylation is a prognostic factor of patient survival in astrocytomas, bladder cancer (BLCA) and renal cell carcinoma (KIRC). Multivariate analysis confirms the independence of *ZMIZ1* methylation (*cg26654807*) as a prognostic marker of cancer survival. [Color figure can be viewed at [wileyonlinelibrary.com](https://onlinelibrary.wiley.com)]

function of ZMIZ1 protein, it has been found to be a regulator of SMAD3/⁴²⁸ and TP53,²⁹ both important cancer pathways. The most well-known function of ZMIZ1 is the sumoylation of androgen receptor.³⁰ Additionally, ZMIZ1 interacts with the SWItch/sucrose nonfermentable nucleosome remodeling complex to possibly influence chromatin remodeling.³¹ Our *in vivo* findings after ZMIZ1 knockdown suggest an additional role for the gene in tumor migration. Also, we note that the role of ZMIZ1 in normal development seems to involve vasculogenesis: homozygous ZMIZ1-null mice exhibit embryonal lethality by Day 10 due to abnormal vascular development.³²

Although our results revealed the clinical and functional significance of ZMIZ1 in cancer, future studies need to clarify certain aspects of the association between ZMIZ1 methylation and patient survival. First, our study primarily used TCGA data, which can lack certain patient characteristics and outcomes. Conducting genomic analyses on a prospectively enrolled cohort of patient tumor samples can help assess for outcomes beyond survival. Second, additional mechanistic studies are required to reveal the role of intragenic methylation on alternative transcription of

ZMIZ1 and how methylation interplays with genetic and other epigenetic factors to affect ZMIZ1 protein expression.

In conclusion, our study shows that ZMIZ1 intragenic methylation may be a novel prognostic marker in numerous solid malignancies. This is based on our finding that intragenic methylation is a possible regulatory mechanism in alternative transcription. Methylation levels of a truncated ZMIZ1 transcript were able to discriminate patient survival across multiple tumor types. Experimental data (*in vitro* and *in vivo*) from our study as well as from prior reports suggest that ZMIZ1 has a biological role in tumorigenesis as the truncated transcript creates a more stable protein, potentially leading to higher gene function or efficiency.²⁵ However, additional studies are needed to fully understand the prognostic and biological role of ZMIZ1 in brain tumors and other neoplasms.

Acknowledgements

We thank the Hunterian Neurosurgical Research Laboratory for helpful comments and discussions. Philanthropic grants from private families to support Hubble and Subependymoma research.

References

- Esteller M, Cordon-Cardo C, Corn PG, et al. p14ARF silencing by promoter hypermethylation mediates abnormal intracellular localization of MDM2. *Cancer Res* 2001;61:2816–21.
- Baylin SB, Jones PA. A decade of exploring the cancer epigenome—biological and translational implications. *Nat Rev Cancer* 2011;11:726–34.
- Esteller M. Epigenetics in cancer. *N Engl J Med* 2008;358:1148–59.
- Jia Y, Yang Y, Zhan Q, et al. Inhibition of SOX17 by microRNA 141 and methylation activates the WNT signaling pathway in esophageal cancer. *J Mol Diagn* 2012;14:577–85.
- Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 2002;3:415–28.
- Jones PA, Baylin SB. The epigenomics of cancer. *Cell* 2007;128:683–92.
- Yan W, Wu K, Herman JG, et al. Epigenetic regulation of DACH1, a novel Wnt signaling component in colorectal cancer. *Epigenetics* 2013;8:1373–83.
- Hegi ME, Liu L, Herman JG, et al. Correlation of O6-methylguanine methyltransferase (MGMT) promoter methylation with clinical outcomes in glioblastoma and clinical strategies to modulate MGMT activity. *J Clin Oncol* 2008;26:4189–99.
- Esteller M, Garcia-Foncillas J, Andion E, et al. Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. *N Engl J Med* 2000;343:1350–4.
- Stupp R, Hegi ME, Mason WP, et al. Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol* 2009;10:459–66.
- Noushmehr H, Weisenberger DJ, Diefes K, et al. Identification of a CpG Island methylator phenotype that defines a distinct subgroup of glioma. *Cancer Cell* 2010;17:510–22.
- Brennan CW, Verhaak RG, McKenna A, et al. The somatic genomic landscape of glioblastoma. *Cell* 2013;155:462–77.
- Parsons DW, Jones S, Zhang X, et al. An integrated genomic analysis of human glioblastoma multiforme. *Science* 2008;321:1807–12.
- Turcan S, Rohle D, Goenka A, et al. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* 2012;483:479–83.
- Draht MX, Riedl RR, Niessen H, et al. Promoter CpG Island methylation markers in colorectal cancer: the road ahead. *Epigenomics* 2012;4:179–94.
- Maunakea AK, Nagarajan RP, Bilieny M, et al. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 2010;466:253–7.
- Illingworth R, Kerr A, Desousa D, et al. A novel CpG Island set identifies tissue-specific methylation at developmental gene loci. *PLoS Biol* 2008;6:e22.
- Rauch TA, Wu X, Zhong X, et al. A human B cell methylome at 100-base pair resolution. *Proc Natl Acad Sci U S A* 2009;106:671–8.
- Deaton AM, Webb S, Kerr AR, et al. Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Res* 2011;21:1074–86.
- Du P, Zhang X, Huang CC, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC bioinformatics* 2010;11:587.
- Kozielewski KL, Tzeng SY, Green JJ. A bioreducible linear poly(beta-amino ester) for siRNA delivery. *Chem Commun (Camb)* 2013;49:5319–21.
- Pruitt KD, Brown GR, Hiatt SM, et al. RefSeq: an update on mammalian reference sequences. *Nucleic Acids Res* 2014;42:D756–63.
- Karolchik D, Barber GP, Casper J, et al. The UCSC genome browser database: 2014 update. *Nucleic Acids Res* 2014;42:D764–70.
- Sturm D, Witt H, Hovestadt V, et al. Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *Cancer Cell* 2012;22:425–37.
- Rogers LM, Riordan JD, Swick BL, et al. Ectopic expression of Zmiz1 induces cutaneous squamous cell malignancies in a mouse model of cancer. *J Invest Dermatol* 2013;133:1863–9.
- Rakowski LA, Garagiola DD, Li CM, et al. Convergence of the ZMIZ1 and NOTCH1 pathways at C-MYC in acute T lymphoblastic leukemias. *Cancer Res* 2013;73:930–41.
- Soler G, Radford-Weiss I, Ben-Abdelali R, et al. Fusion of ZMIZ1 to ABL1 in a B-cell acute lymphoblastic leukaemia with a t(9;10)(q34;q22.3) translocation. *Leukemia* 2008;22:1278–80.
- Li X, Thyssen G, Beliakoff J, et al. The novel PIAS-like protein hZimp10 enhances Smad transcriptional activity. *J Biol Chem* 2006;281:23748–56.
- Lee J, Beliakoff J, Sun Z. The novel PIAS-like protein hZimp10 is a transcriptional co-activator of the p53 tumor suppressor. *Nucleic Acids Res* 2007;35:4523–34.
- Sharma M, Li X, Wang Y, et al. hZimp10 is an androgen receptor co-activator and forms a complex with SUMO-1 at replication foci. *EMBO J* 2003;22:6101–14.
- Li X, Zhu C, Tu WH, et al. ZMIZ1 preferably enhances the transcriptional activity of androgen receptor with short polyglutamine tract. *PLoS One* 2011;6:e25040.
- Beliakoff J, Lee J, Ueno H, et al. The PIAS-like protein Zimp10 is essential for embryonic viability and proper vascular development. *Mol Cell Biol* 2008;28:282–92.