Research Paper

SHOX2 is a Potent Independent Biomarker to Predict Survival of WHO Grade II–III Diffuse Gliomas

Yu-An Zhang a,i,1, Yunyun Zhou b,c,1, Xin Luo d, Kai Song e, Xiaotu Ma f, Adwait Sathe g, Luc Girard a, Guanghua Xiao b,d,h, Adi F Gazdar a,i,h,*

a The Hamon Center for Therapeutic Oncology Research, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA
b Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA
c Center of Biostatistics and Bioinformatics, University of Mississippi Medical Center, Jackson, MS 39216, USA
d Department of Bioinformatics, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA
e School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China
f Department of Computational Biology, St Jude Children’s Research Hospital, 262 Danny Thomas Place, Memphis, TN, USA
g The Center for Systems Biology, Department of Molecular and Cell Biology, The University of Texas at Dallas, Richardson, TX 75080, USA
h The Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA
i Department of Pathology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

Abstract

Background: Diffuse gliomas, grades II and III, hereafter called lower-grade gliomas (LGG), have variable, difficult to predict clinical courses, resulting in multiple studies to identify prognostic biomarkers. The purpose of this study was to assess expression or methylation of the homeobox family gene SHOX2 as independent markers for LGG survival.

Methods: We downloaded publicly available glioma datasets for gene expression and methylation. The Cancer Genome Atlas (TCGA) (LGG, n = 516) was used as a training set, and three other expression datasets (n = 308) and three other methylation datasets (n = 320), were used for validation. We performed Kaplan-Meier survival curves and univariate and multivariate Cox regression model analyses.

Findings: SHOX2 expression and gene body methylation varied among LGG patients and highly significantly predicted poor overall survival. While they were tightly correlated, SHOX2 expression appeared more potent as a prognostic marker and was used for most further studies. The SHOX2 prognostic roles were maintained after analyses by histology subtypes or tumor grade. We found that the combination of SHOX2 expression and IDH genotype status identified a subset of LGG patients with IDH wild-type (IDHwt) and low SHOX2 expression with considerably favorable survival outcomes. We further investigated the combination of SHOX2 with other known clinically relevant markers of LGG (TER1 expression, 1p/19q chromosome co-deletion, MGMT methylation, ATRX mutation and NES expression). When combined with SHOX2 expression, we identified subsets of LGG patients with significantly favorable survival outcomes, especially in the subgroup with worse prognosis for each individual marker. Finally, multivariate analysis demonstrated that SHOX2 was a potent independent survival marker.

Interpretation: We have identified that SHOX2 expression or methylation are potent independent prognostic indicators for predicting LGG patient survival, and have potential to identify an important subset of LGG patients with IDHwt status with significantly better overall survival. The combination of IDH or other relevant markers with SHOX2 identified LGG subsets with significantly different survival outcomes, and further understanding of these subsets may benefit therapeutic target identification and therapy selections for glioma patients.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Brain tumor gliomas include low grade (grade I) pilocytic astrocytomas, and the diffuse gliomas that include the grades II and III astrocytomas and oligodendrogliomas (referred to as lower-grade gliomas, LGG) and the highly malignant grade IV glioblastomas [GBM, grade IV, the World Health Organization (WHO) Classification of Tumors of the Central Nervous System (CNS)] (Louis et al., 2016; Louis et al., 2007).

http://dx.doi.org/10.1016/j.ebiom.2016.10.040
2352-3964/© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
LGG are diffusely infiltrative tumors and have highly variable, difficult to predict clinical courses, further compounded by inter-observer variability in histologic classification and grading (van den Bent, 2010; Louis et al., 2007). While some LGG have indolent outcomes, others rapidly progress to high grade GBM. GBM patients almost always die from their disease (Louis et al., 2007; Ostrom et al., 2015). The evolution of gliomas from grade II to grade III or IV are characterized by the stepwise acquisition of genetic alterations and a considerable worsening of prognosis, justifying studies to identify genetic alterations as potential biomarkers for prognosis and selection of targeted therapy and overall clinical management (Ellison, 2015). A relatively recent finding of major biological and clinical importance was the identification of mutations in the isocitrate dehydrogenase (IDH) enzyme genes IDH1 and IDH2. Somatic mutations, in particular of the IDH1 gene, are present in the majority of LGG, especially oligodendrogliomas, and have a positive effect on overall survival (Turkalp et al., 2014; Yan et al., 2009). They are rare in primary GBM and absent in pilocytic astrocytomas and are often associated with MGMT promoter hypermethylation, TP53 mutations as well as co-deletions of chromosome 1p or 19q (1p/19q codeletion). IDH mutations are an early, possibly driver, event for LGG (Watanabe et al., 2009), and clinical trials of IDH inhibitors are underway (Dimitrov et al., 2015). Many studies have demonstrated that survival outcome of LGG patients is significantly different based on the status of IDH gene mutation, 1p/19q codeletion, telomerase reverse transcriptase (TERT) promoter mutation, ATRX gene mutation, CpG island methylator phenotypes (CIMP), O-6-methylguanine-DNA methyltransferase (MGMT) promoter methylation, the neural stem cell gene nestin (NES) expression and mRNA expression signatures by multiple genes (Cancer Genome Atlas Research et al., 2015; Eckel-Passow et al., 2015; Ceccarelli et al., 2016; Chan et al., 2015; Noushmehr et al., 2010; Turcan et al., 2012; Hatanpaa et al., 2014; Siegal, 2015; Bao et al., 2014; Zhang et al., 2015). The classification by CIMP status after filtering IDH mutation status revealed biologically discrete subsets having different clinic survival outcomes in diffuse gliomas (Ceccarelli et al., 2016), supporting the principle that IDH mutation status plus other molecular biomarkers can enhance the prognostic value for certain molecularly distinct subsets of LGG patients. The importance of combining tumor molecular features with traditional diagnostic features such as histology and grading was recognized in the recently revised 2016 WHO classification systems of CNS tumors (Louis et al., 2016).

The SHOX2 gene, located on chromosome 3q, is a member of the homeobox family of genes that encodes a transcriptional regulator and its expression and methylation have been studied in various tumors. The gene is involved in the development of skeletal, neural, and muscular tissues. Its expression is upregulated in glioblastomas compared to lower grade gliomas, suggesting a potential role in glioma progression. The promoter region of SHOX2 is highly methylated in glioblastomas, indicating a significant association with the disease.

### Fig. 1. Identification and correlation of differential SHOX2 methylation and expression in glial tumors. (A) SHOX2 gene structure and its CpG sites by the Illumina HumanMethylation450 BeadChip (Illumina) array (450 K) were based on the UCSC Genome Browser (http://genome.ucsc.edu) and the NCBI Gene site (http://www.ncbi.nlm.nih.gov/gene/6474). The 35 designated 450 K probes are indicated by vertical lines scattered throughout SHOX2 gene (Note: The 450K ProbeTarget ID cg00303223 was excluded due to its data unavailable in TCGA LGG and GBM 450K datasets). (B) Differential SHOX2 methylation among non-malignant brains (Non-M), pilocytic astrocytomas (PA), lower grade gliomas (LGG) and glioblastomas (GBM) using TCGA LGG and GBM 450 K datasets. Five high differential methylation probes were found. The pooled mean beta value of four probes (Probes 1, 2, 8 and 9, Fig. 1A) was used. (C) Differential SHOX2 mRNA expression using GSE16011 microarray data. (D) Correlation of SHOX2 methylation and expression using TCGA LGG 450 K and RNA SeqV2 datasets (see also Appendix p2–9 for dataset source and data analyses). r: Spearman r, 2-tailed. Bars: median values.
expression is highly restricted to craniofacial, brain, heart, and limb development (Blaschke et al., 1998; Clement-Jones et al., 2000). SHOX2 promoter DNA methylation has been identified as a diagnostic and prognostic biomarker for non-small cell lung cancer patients (Schmidt et al., 2010; Dietrich et al., 2012). Elevated SHOX2 expression is associated with tumor recurrence of hepatocellular carcinoma (Yang et al., 2013) and with poor survival in breast cancers (Hong et al., 2014). In our experience, some genes such as SCT and ITPKA, are frequently methylated in many invasive cancers, but their methylation in certain low-grade tumors is variable (Zhang et al., 2016; Wang et al., 2016). The availability of a well-studied large set of LGG having molecular data including exome sequencing and genome wide methylation (Cancer Genome Atlas Research et al., 2015), permitted us to examine the prognostic role of SHOX2 methylation and expression in gliomas and to correlate the data with other prognostic parameters. The primary aim of our study was to demonstrate the prognostic role of SHOX2 as a single indicator or in combination with IDH and other biomarkers for improving survival predictions for LGG patients.

2. Materials and Methods

2.1. Datasets

We examined publically available genome-wide methylation and expression data of nonmalignant brains and glioma tissues, the associated pathological molecular markers, and clinical variables from the following main sources: The Cancer Genome Atlas (TCGA) data portal, National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO), ArrayExpress, and the data as described previously (Ceccarelli et al., 2016). The detailed list of data sources and accession numbers and data process are provided in Appendix p1–4.

We used the TCGA LGG dataset as a training set to study the association between SHOX2 methylation (or expression) and overall survival in LGG patients. These findings were independently validated in three external expression microarray datasets (GSE61160, GSE16011, GSE30336 and REMBRANDT) and three external methylation datasets (GSE61160, GSE30338 and GSE58218). Details of the clinical trials and cancer centers from which the data were extracted are provided in Appendix p3–4. We examined enhanced prognostic values of SHOX2 in subsets of LGG classified by IDH or other known relevant markers using TCGA LGG datasets.

2.2. Statistical analysis

Two-tailed, Student t-test was performed to compare two groups of numerical values. To assess SHOX2 as a prognostic biomarker, patient samples analyzed were dichotomized into two groups designed as high- and low- groups, based on a SHOX2 cutoff value for their methylation or expression values. For the dataset containing numerical values, we used a data-driven approach to define an objective cutoff value by using model-based clustering method implemented by “mclust” package version 4.4 for R (Fraley and Raftery, 2002). By mclust clustering analysis, the SHOX2 values (methylation or expression) were fitted into a mixture of two normal distributions, one with high SHOX2 values and the other with low SHOX2 values. The same approach was used to define the cutoff value for younger and older patient groups by age. Overall survival time was calculated from the date of diagnosis until death or the last follow-up contact. Survival curves were estimated using the product-limit method of Kaplan-Meier (Kaplan and Meier, 1958) with the log-rank test. Univariate and multivariate Cox tests were performed to assess the relative contribution of the risk group when assessed alone or after adjusting for clinical variables (or other indicator) (Andersen and Gill, 1982, Therneau and Grambsch, 2000).

3. Results

3.1. SHOX2 Methylation or mRNA Expression in Glial Tumors

A cartoon and description of the SHOX2 gene are presented in Fig. 1A. From the TCGA datasets, varying degrees of methylation were noted at the 35 probes covering the entire gene (Fig. 1). We identified five probes, named as P1, P2, P8, P9 and P10 (Fig. 1A) which showed significantly high differential methylation in glial tumors (see details on Appendix p5–9). These probes were located between the last two exons.

**IDH (A): wt vs mut**

**SHOX2 Methylation (Methy: B) or Expression (Exp: C): High vs Low**

![Survival curves](image)

Fig. 2. Kaplan-Meier survival curve analyses of SHOX2 alone and the combination of SHOX2 with IDH mutation status in LGG patients using TCGA LGG datasets. The analyses of SHOX2-450K methylation (cutoff value of 0.357) and RNA-Seq expression data (cutoff value of 4.135) were presented in Figs. 2B1–3 and C1–3, respectively. IDHwt: IDH wild-type; IDHmut: IDH mutation (see also Appendix p2–9 for dataset source and data analyses).
and not at CpG island sites. As shown in Fig. 1B, SHOX2 methylation was absent or low in non-malignant brain tissues and pilocytic astrocytomas (PA), frequent and high in GBM, and with intermediate frequencies and values in LGG, respectively.

By microarray data, SHOX2 expression patterns were similar to methylation: low in non-malignant brain and PA cases, high in GBM and intermediate but variable in LGG (Fig. 1C). The significantly higher expression of SHOX2 in GBM compared to LGG was also

![Fig. 4. Kaplan-Meier overall survival curve analyses of SHOX2 methylation in additional external 450 K microarray methylation datasets besides the TCGA LGG dataset. The survival analyses were performed using SHOX2 pooled mean methylation probe values (see Fig. 1) of LGG samples from the combined datasets (GSE61160 and GSE30338). p: Log-rank test. See Appendix p2–4, 13–14 for detailed list of datasets and data source and further analyses data.](image)

**References:**

1. Zhang, Y.-A. et al. / EBioMedicine 13 (2016) 80–89
validated in additional microarray datasets (Appendix p10). The RNA-Seq data showed the same pattern of differential SHOX2 expression (Fig. 1D). In addition, by correlating the RNA-Seq data with their corresponding sample methylation data available, we found that the SHOX2 methylation and RNA-Seq expression values were highly positively correlated in both LGG and GBM tumors (Fig. 1D).

3.2. SHOX2 Methylation and mRNA Expression are Independent Potent Prognostic Markers for LGG

To explore whether SHOX2 methylation and expression are related to LGG patient survivals, we first analyzed TCGA LGG training datasets containing 516 cases. The samples were dichotomized into either high or low subgroups when classified by SHOX2 methylation or mRNA expression levels. We found that the methylation marker based on any of the five differentially methylated probes alone predicted overall survival of LGG patients, and use of the pooled mean beta values of four probes (Probes 1, 2, 8 and 9, Fig. 1A) showed significantly improved prognostic values as compared with the use of individual probes (Appendix p11–12) and other varied combinations of these probes (not shown). Thus the pooled mean beta values were used for further survival analyses of SHOX2 methylation. Our analyses demonstrated that SHOX2 methylation was associated with poor survival prognosis and was a potent independent prognosis marker for LGG (Wiestler et al., 2014; Noushmehr et al., 2010) as demonstrated in
three independent external 450 K methylation datasets (Figs. 4–5 and Appendix p13–14). SHOX2 methylation (of the gene body associated with increased gene expression) was negatively correlated with IDH status (methylation of promoter region of multiple genes associated with suppression of gene expression) (Fig. 5).

Similarly, SHOX2 high mRNA expression (SHOX2_high) was an even more potent independent marker in predicting worse overall survival in the same LGG dataset using TCGA LGG RNA-Seq dataset [n = 516, HR 5.08 (95% CI 3.53 to 7.33), p < 2.0E-16] (Fig. 2C1, Fig. 3A). These findings were independently validated in three additional external expression microarray datasets, as demonstrated by using SHOX2 as a single prognostic marker or in combination of SHOX2 with IDH status data available (Fig. 6 and Appendix p15–16).

While SHOX2 methylation and expression were highly correlated, SHOX2 expression appeared more potent than methylation for predicting LGG survival and complete expression datasets were more widely available. For these reasons above, we used SHOX2 expression marker for further studies.

We observed that SHOX2 expression marker had a prognostic value comparable to IDH status marker, a widely-accepted potent prognostic marker for LGG [IDH wild type (IDHwt) vs mutant type (IDHmut)] HR 6.52 (95% CI 4.49 to 9.47), p < 2.0E-16] (Figs. 2A, C1 and 3A). Interestingly, there was a significantly higher concordance rate (0.75) between IDHmut and SHOX2 low expression than that (0.65) between IDHwt and SHOX2 low expression, as compared with a significant but less degree concordance difference between IDHwt and SHOX2 high expression (0.13) vs SHOX2 low expression (0.06) using TCGA LGG expression dataset (Appendix p17). As compared with other known markers, such as TERT, MGMT, 1p/19q codel, TERT and NES, SHOX2 had the highest prognostic hazard ratio value as demonstrated in the univariate Cox proportional-hazards model analyses of each marker as a single indicator in TCGA LGG dataset (Fig. 3A).

We performed overall survival analyses of SHOX2 expression by histology (astrocytoma, oligodendroglioma, excluding oligoastrocytoma which is not recognized as a separate entity in the 2016 CNS tumor classification system) (Louis et al., 2016) or tumor grade (grades II and III). We found that SHOX2 expression was a potent survival indicator in both histology types and grades, especially in astrocytoma and grade III tumors (Fig. 7 and Appendix p18).

However, we did not find that SHOX2 expression was a prognostic marker for overall survival in high grade GBM patients in TCGA GBM dataset (Appendix p19), possibly because the vast majority of GBM samples had high SHOX2 expression values (Fig. 1D).

3.3. Combination of SHOX2 with IDH and Other Relevant Prognostic Markers

Next, we determined whether SHOX2 can improve prognostic values of IDH in LGG patients. While the poor prognostic SHOX2_high and IDHwt subgroups were frequently present together (Appendix p20), we found that approximately one third (27 out of 94) of the IDHwt cases which had SHOX2 low expression, were associated with an improved overall median survival of 9.6 years (p = 7.29E-05, log-rank test) (Figs. 2C2 and 3B). By contrast, combining SHOX2 expression or methylation had no or minimal effect on the favorable prognostic IDHmut subgroup.

The findings of the enhanced prognostic combination of SHOX2 and IDH markers led us to further study whether SHOX2 can aid the prognostic values of other LGG prognostic markers (NES expression, MGMT methylation, TERT promoter mutation or expression, 1p/19q codel and ATRX mutation). Using the TCGA dataset, we analyzed each marker by itself, and compared with the prognostic values of SHOX2 marker in the dichotomized subsets of LGG by these individual markers (Figs. 3B and 8). We found that 1p/19q codel or MGMT methylation as individual

![Fig. 7. Kaplan-Meier overall survival curve analyses of SHOX2 expression in LGG by histology and tumor grades (TCGA LGG dataset). The cutoff value of SHOX2 RNA-Seq expression data (4.135) was used to determine SHOX2 expression high vs low. The grades II or III tumor samples included all histology types included in the TCGA LGG RNA-Seq datasets analyzed. See Appendix p2–4. 18 for detailed list of datasets and data source and further analyses data.](https://example.com/7.png)
markers had a moderate prognostic effect, and NES, TERT and ATRX marker had no significant prognostic effect. In particular, we found that SHOX2 expression identified favorable prognostic subsets more significantly in the unfavorable prognosis subgroup determined by each individual marker, which was similar to our findings in the subset of IDHwt LGG (Figs. 3 and 8).

We observed that the prognostic value of other markers was improved in combination with IDH mutation status, as compared to utilization of each marker alone, especially TERT expression or mutation, MGMT methylation, ATRX mutation and NES expression (Appendix p21). Comparing the effects of SHOX2 and IDH makers while combining with other relevant markers, we found that SHOX2 expression

---

**Fig. 8.** Kaplan-Meier overall survival curve analyses of SHOX2 expression in the subsets of LGG when combined with other individual markers (TCGA LGG dataset). This figure presented the result of survival analyses of other markers (NES, MGMT, TERT, 1p/19q and ATRX) individually, and SHOX2 expression marker in the subsets of LGG sub-classified by these other individual markers. The cutoff value of SHOX2 RNA-Seq expression data (4.135) was used to determine SHOX2 expression high vs low. See Appendix p2–4 for detailed list of datasets and data source and Fig. 3B for NES expression cutoff value and additional data of each subset analyzed.
significantly enhanced the prognosis values of other relevant markers to a degree similar and comparable to IDH mutation status marker (Fig. 3B and Appendix p21).

3.4. Multivariate Cox Proportional-hazards Survival Analysis of SHOX2 and Other Markers

We performed multivariate Cox regression survival analysis to assess the relative contribution of the predicted risk groups classified by SHOX2 expression, after adjusting for clinical variables including age, gender, stage and histology subtypes as well as other prognostic markers. We demonstrated that SHOX2 expression and age were highly significant independent variables, while, gender, histology and tumor grade were of lesser significance (Table 1). Both SHOX2 methylation and expression remained a significantly predictive marker and SHOX2 expression had a stronger prognostic value comparable to IDH1 methylation (HR: 1.65 (95% CI 1.08 to 2.50, \( p = 1.94E-02 \)); expression: HR 3.38 (95% CI 2.25 to 5.09, \( p = 5.12E-09 \); IDH1 HR: 3.82 (95% CI 2.47 to 5.90), \( p = 5.12E-09 \)) (Table 1 and Fig. 9).

To determine the relative contribution of the risk groups by SHOX2 marker, after adjusting for the status of other known markers, we performed multivariate analyses under the same clinical variables setting as shown in Table 1 except including one of other prognostic markers (IDH1, 1p/19q codel, TERT, MGMT, NES and ATRX) (Fig. 8). Fig. 8 shows a comparison of the hazard ratio values of SHOX2 marker vs. other markers by multivariate analyses, and we demonstrated that under the combination of SHOX2 with the other prognostic markers, SHOX2 was also an independent prognostic factor. By contrast, IDH1, 1p/19q codel and MGMT but not TERT and NES showed moderate prognostic significance after adjusting by SHOX2 marker, based on their HR and \( p \) values (Fig. 9).

The overall survival of LGG patients was previously reported to be related with age and histology (Turcan et al., 2012; Cancer Genome Atlas Research et al., 2015). We found that SHOX2 high methylation and expression were associated with astrocytoma histology type (Appendix p22). SHOX2 methylation and expression levels were low and not age-dependent in normal and non-malignant brain tissues (Appendix p23–24), but appeared partially age-dependent in LGG tumors as reflected by a subset of LGG cases with high SHOX2 methylation and expression values available in all age groups. Thus we performed multivariate analyses in younger (<48 years) or older (≥48 years) subgroups, respectively. The cutoff value (48 years) was objectively determined by model based clustering analysis of age data in TCGA LGG dataset (see Methods Statistical analysis). We found that IDH1 and SHOX2 expression as individual markers both had higher HR values in the older patient subgroup than those in young patient subgroup, and SHOX2 had a slightly higher HR value with a higher significant \( p \) value than IDH1 in the younger patient subgroup (SHOX2: HR 3.00 (95% CI 1.54 to 5.87), \( p = 0.0013 \); IDH: HR 2.65 (95% CI 1.19 to 5.90), \( p = 0.0171 \)) (Appendix p25). While both SHOX2 and IDH were included in multivariate analyses, SHOX2 was more significant than IDH in the younger patient subgroup. The results of comparing SHOX2 with other markers in different age subgroups are presented in (Appendix p25–27).

4. Discussion

In May of 2016, the latest version of 2016 WHO Classification was published, and the major changes in this version, as it relates to the present report, was combining molecular markers with traditional histology classification to introduce a more clinically relevant classification (Louis et al., 2016). In particular, a) most grades II and III diffuse gliomas have mutations in IDH1 or IDH2 genes; b) most astrocytomas are 1p/19q intact and are often ATRX and 7p53 mutant; c) most or all oligodendrogliomas are 1p/19q co-deleted; and d) combined oligoastrocytomas are no longer recognized as an entity, but should be reclassified based on their molecular features. The grading system is maintained but its importance is de-emphasized. We present our analyses using traditional histology (omitting the category oligoastrocytoma) and grading, and the various molecular analyses including IDH, ATRX mutation status and 1p/19q co-deletions. The datasets we utilized were generated prior to the revised 2016 Classification, and from the sample data available, not all cases can be reclassified. We did not combine the traditional methods (histology and grading) with molecular methods for comparison with SHOX2 expression or methylation, as that would introduce too many variables and would result in subsets that could not be reclassified according to the 2016 Classification. However, because our approach utilized both the traditional approaches and the important molecular features utilized in the 2016 Classification (IDH and ATRX mutation status, and 1p/19q co-deletion status) we believe it is relevant to the new 2016 WHO classification.

The role of multiple prognostic markers has been investigated in LGGs. We investigated the role of SHOX2 expression and methylation in LGGs and their relationship to other prognostic and pathologic markers. We found approximately 20% of tumors had increased levels of expression and/or methylation. There was a high degree of positive correlation between the two parameters. At first this may appear paradoxical, as hypermethylation of the promoter region serves as a repressive epigenetic mark that down-regulates gene expression. However, the methylated regions of SHOX2 were located in the gene body and previous studies as well as a recent one by us have noted that gene body methylation, which is more prevalent in the genome than promoter hypermethylation, may be associated with increased gene expression (Jones, 2012; Wang et al., 2016). Multiple recent studies have highlighted the importance and promise of using molecular markers for LGG prognosis and potentially as aids for therapeutic selection (Cancer Genome Atlas Research et al., 2015; Eckel-Passow et al., 2015; Tertulian et al., 2014; Dimitrov et al., 2015; Ceccarelli et al., 2016). To date, IDH mutation status has been the most widely accepted and powerful prognostic factor, either alone or in combination with other factors. In this study, we identified that SHOX2 methylation or expression were potent independent prognostic indicator comparable to IDH for LGG patient survival. We used the TCGA dataset as our training set and six additional

Table 1
The multivariate Cox proportional-hazards models analyses of SHOX2 methylation or expression for LGG patients (TCGA LGG dataset).

| Variable | SHOX2 methylation | | | SHOX2 expression | | |
|----------|-------------------|---|---|------------------|---|---|
|          | Hazard ratio (95% CI) | \( p \) Value | Hazard ratio (95% CI) | \( p \) Value |
| Patient#/Event#: 512/126 | | | | |
| SHOX2 high vs low | 1.65 (1.08 to 2.50) | 0.0194 | 3.38 (2.25 to 5.09) | 5.12E-09 |
| Age | 1.06 (1.04 to 1.07) | 0.0757 | 1.05 (1.04 to 1.07) | 7.88E-10 |
| Gender (male vs female) | 0.51 (0.33 to 0.78) | 0.0019 | 0.53 (0.34 to 0.81) | 0.0036 |
| Tumor Grade (3 vs 2) | 2.10 (1.37 to 3.23) | 0.0007 | 2.24 (1.47 to 3.41) | 0.0002 |

Hazard ratio for age variable: risk per 1 year. OA: oligoastrocytoma; A: astrocytoma; OD: oligodendroglioma. See Appendix p3–4 for sample source and data.
and signiﬁcant positive relationship with hepatocellular carcinomas (Yang et al., 2013; Hong et al., 2014). We observed an upward trend of expression from clinically benign pilocytic astrocytomas, intermediate to high malignant LGG datasets for validation of expression or methylation. While SHOX2 expression and gene body methylation were highly correlated, expression appeared to be the better prognostic marker and its data were more readily available. Thus we used expression for most additional studies. Importantly, we found that the combination of SHOX2 and IDH status identiﬁed a subset of IDHwt LGG with a considerably improved survival time. Of interest, another study utilizing molecular proﬁling indicated the presence of a subtype with improved prognosis (Ceccarelli et al., 2016). Conceivably, SHOX2 status can be determined by standard molecular assays such as PCR in a routine laboratory, which is clinically attractive.

We further investigated the combination of SHOX2 with other known clinically relevant markers of LGG including TERT, 1p/19q, MGMT, ATRX and NES, and found that SHOX2 was a potent indicator to identify subsets of LGG with signiﬁcant better survivals especially in the worse prognosis subgroup determined by individual markers. Our current study demonstrated that SHOX2 not only is an independent potent prognostic marker, but also has potential to reﬁne the molecular classiﬁcation of LGG in combination with other well established markers. The addition of SHOX2 expression identiﬁes small but highly signiﬁcant subgroups having good prognosis within the poor prognosis groups identiﬁed by IDH mutation status and several other commonly used prognostic markers. Thus, therapy options for these subgroups may be altered as a result of our observations, although prospective studies will be required to prove this. Finally, we demonstrated by multivariate survival analysis that SHOX2 was a potent survival prognosis marker comparable to IDH after adjusting for age and other clinical variables and signiﬁcantly different from all the other markers mentioned above.

As mentioned previously, SHOX2 may play a prognostic role in breast and hepatocellular carcinomas (Yang et al., 2013; Hong et al., 2014). SHOX2 was suggested to be a novel epithelial-to-mesenchymal transition inducer in breast cancer cells (Hong et al., 2014), and overexpression of SHOX2 was able to induce canine mesenchymal stem cell differentiation into native pacemaker cells (Feng et al., 2016). We observed an upward trend of SHOX2 aberrant hypermethylation and expression from clinically benign pilocytic astrocytomas, intermediate to high malignant LGG to high malignant GBM. This is in contrast to IDH mutations which are largely limited to LGG and only occasionally present in primary GBM (Yan et al., 2009). However, SHOX2 expression was not found to have prognostic signiﬁcance in the TCGA GBM dataset in this study, possibly because most GBM tumors have high SHOX2 expression. Our current ﬁndings suggest a potential oncogenic role of SHOX2 in glioma tumors, although the precise mechanism is largely unknown.

In conclusion, we have identiﬁed that SHOX2 expression or methylation are potent independent prognostic indicators for predicting LGG patient survival, and have potential to identify an important subset of LGG patients with IDHwt status with signiﬁcantly better overall survival. The combination of IDH or other relevant markers with SHOX2 identiﬁed LGG subsets with signiﬁcantly different survival outcomes, and further understanding of these subsets may beneﬁt therapeutic target identiﬁcation and therapy selections for glioma patients.

Markers Compared Variable Events/Patients HR (95% CI) p value

| SHOX2-alone | SHOX2 mRNA: high vs low | 51/101 vs 75/414 | 3.38 (2.25 to 5.09) | 5.12E-09 |
| IDH-alone | IDH: wt vs mut | 50/94 vs 74/418 | 3.82 (2.47 to 5.90) | 1.44E-09 |
| SHOX2 vs IDH | SHOX2 mRNA: high vs low | 51/101 vs 75/414 | 2.64 (1.61 to 4.34) | 0.0001 |
| SHOX2 vs IDH | SHOX2 mRNA: high vs low | 51/101 vs 75/414 | 2.87 (1.87 to 4.39) | 1.27E-06 |
| SHOX2 vs TERT | SHOX2 mRNA: high vs low | 51/101 vs 75/414 | 2.99 (1.63 to 5.46) | 0.0004 |
| SHOX2 vs MGMT | SHOX2 mRNA: high vs low | 51/101 vs 75/414 | 3.08 (2.01 to 4.71) | 2.09E-07 |
| SHOX2 vs TERT | SHOX2 mRNA: high vs low | 51/101 vs 75/414 | 3.34 (2.20 to 5.06) | 1.31E-08 |
| SHOX2 vs NES | SHOX2 mRNA: high vs low | 51/101 vs 75/414 | 3.40 (2.25 to 5.14) | 7.21E-09 |
| SHOX2 vs ATRX | SHOX2 mRNA: high vs low | 51/101 vs 75/414 | 3.91 (2.50 to 6.10) | 2.06E-09 |
| IDH vs SHOX2 | IDH: wt vs mut | 50/94 vs 74/418 | 2.53 (1.52 to 4.20) | 0.0003 |
| 1p/19q vs SHOX2 | 1p/19q: non-code vs code | 104/346 vs 22/169 | 2.30 (1.30 to 4.06) | 0.0042 |
| TERT(promoter) vs SHOX2 | TERT promoter: mut vs wt | 39/130 vs 35/162 | 1.09 (0.59 to 2.01) | 0.7887 |
| MGMT vs SHOX2 | MGMT: unmethylated vs methylated | 39/90 vs 87/425 | 1.82 (1.20 to 2.76) | 0.0048 |
| TERT(expression) vs SHOX2 | TERT mRNA: Not_Expressed vs Expressed | 67/288 vs 59/224 | 0.94 (0.61 to 1.44) | 0.7632 |
| NES vs SHOX2 | NES mRNA: high vs low | 68/255 vs 58/260 | 1.34 (0.93 to 1.93) | 0.1189 |
| ATRX vs SHOX2 | ATRX: wt vs mut | 79/325 vs 45/187 | 0.96 (0.61 to 1.51) | 0.8540 |

Author Contributions

YAZ and AFG contributed to the conception and study design, literature search, ﬁgures and tables’ presentation, data collection and assembly and analysis and interpretation, writing and critical reading of manuscript. YZ, XL and GX contributed to the data collection and assembly, and critical reading of manuscript.

Declaration of Interests

We declare no competing interests.

Funding Sources

National Institutes of Health (USA).
Partial funding for this study was provided by a grant from National Institutes of Health (1R01CA17221-01A1).

Appendix A Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ebiom.2016.10.040.

References

Andersen, P.K., Gill, R.D., 1982. Cox regression-model for counting-processes - a large sample processes. Ann. Stat. 10, 1100–1120.

Bao, Z.S., Li, M.Y., Wang, J.Y., Zhang, C.B., Wang, H.J., Yan, W., Liu, Y.W., Zhang, W., Chen, L., Jiang, T., 2014. Prognostic value of a nine-gene signature in glioma patients based on mRNA expression profiling. CNS Neurosci. Ther. 20, 112–118.

Blaschke, R.J., Monahan, A.P., Schiller, S., Scheichinger, B., Rao, E., Padilla-Nash, H., Ried, T., Rappold, G.A., 1998. SHOX, a SHOX-related homeobox gene, is implicated in craniofacial, brain, heart, and limb development. Proc. Natl. Acad. Sci. U. S. A. 95, 2406–2411.

Cancer Genome Atlas Research, N. Brat, D.J., Verhaak, R.G., Aldape, K.D., Yung, W.K., Salama, S.R., Cooper, L.A., Rheinbay, E., Miller, C.R., Vitucci, M., Morozova, O., Robertson, A.G., Noushmehr, H., Laid, P.W., Csernencsi, A., Wilk, P., Huse, J., He, J., Brat, D.J., Lehman, N.L., Barnholtz-Sloan, J.S., Zheng, S., Hess, K., ROA, G., Newton, Y., Radenbaugh, A., Pagnotta, S.M., Anjum, S., Wang, J., Manyam, G., Ostrom, Q.T., Gittleman, H., Fulop, J., Liu, M., Blanda, R., Perou, C.M., Schmidt, H.K., Ding, L., Wilson, K.K., Van Den Berg, D., Shen, H., Brettschloss, G., Noushmehr, H., Reis, G.F., Pico, A.R., O'Neill, B.P., Buckner, J.C., Giannini, C., Griffon, M., Lazar, A., Bai, X., Xiang, Y., Shi, H., Kim, J., Huang, S., 2014. SHOX2 is a direct miR-375 target and a novel epithelial-to-mesenchymal transition inducer in breast cancer cells. Neoplasia 16 (79–90), 1–e5.

Jones, P.A., 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat. Rev. Genet. 13, 484–492.

Kaplan, E.L., Meier, P., 1958. Nonparametric-estimation from Incomplete Observations. J. Am. Stat. Assoc. 53, 457–481.

Louis, D.N., Ohgaki, H., Wiestler, O.D., Cavenee, W.K., Burger, P.C., Jouvet, A., Scheithauer, B.W., Kleihues, P., 2007. The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol. 114, 97–109.

Noushmehr, H., Weisenberger, D.J., Diefes, K., Phillips, H.S., Pujara, K., Berman, B.P., Pan, F., Pellikoski, C.E., Sulman, E.P., Blat, K.P., Verhaak, R.G., Hoadley, K.A., Hayes, D.N., Perou, C.M., Schmidt, H.K., Ding, L., Wilson, K.K., Van Den Berg, D., Shen, H., Brettschloss, G., Noushmehr, H., Cope, L.M., Buckley, J., Herman, J.G., Baylin, S.B., Laird, P.W., Aldape, K. Cancer Genome Atlas Research, N. 2010. Identification of a CpG island methylator phenotype that defines a distinct subgroup of glioma. Cancer Cell 17, 510–522.

Ostrom, Q.T., Littlefield, H., Fujko, J., Liao, M., Blanda, R., Kramer, C., Wolinsky, Y., Kruchko, C., Barnholtz-Sloan, J.S., 2015. CBTRUS Statistical Report: primary brain and central nervous system tumors diagnosed in the United States in 2008–2012. Neuro-Oncol. (17 Suppl 4), 1–v62–1.

Schmidt, B., Liebenberg, V., Dietrich, D., Schlegel, T., Kniep, C., Seegebarth, A., Fleming, N., Seemüller, S., Dietler, J., Lewin, J., Tetran, R., Weiskamm, S., Wilde, U., Lilioglu, T., Raji, O., Walshaw, M., Fleischhacker, M., Witt, C., Field, J.K., 2010. SHOX2 DNA methylation is a biomarker for the diagnosis of lung cancer based on bronchial aspirates. BMC Cancer 10, 600.

Siegal, T., 2015. Clinical impact of molecular biomarkers in gliomas. J. Clin. Neurosci. 22, 437–444.

Therneau, T.M., Grambsch, P.M., 2000. Modeling Survival Data: Extending the Cox Model. Springer-Verlag, New York.

Turcan, S., Rohlie, D., Goenka, A.L., Wang, L., Fylma, E., Campos, C., Fabius, A.W., Lu, C., Ward, P.S., Thompson, C.B., Kaufman, A., Guryanova, O., Levine, R., Heguy, A., Viale, A., Morris, L.C., Huse, J.T. Mellinghoff, I.K., Chan, T.A., 2012. IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. Nature 483, 479–483.

Turkalj, Z., Karamandian, J., Das, S., 2014. IDH mutation in glioma: new insights and promises for the future. JAMA Neurol. 71, 1319–1325.

Van Den Bent, M.J., 2010. Interobserver variation of the histopathological diagnosis in clinical trials on gliomas: a clinician's perspective. Acta Neuropathol. 120, 297–304.

Wang, Y.W., Ma, X., Zhang, Y.A., Wang, M.J., Yatabe, Y., Lam, S., Girard, L., Chen, J.Y., Zhang, Y., Wang, Y.W., Yatabe, Y., Lam, S., Girard, L., Watanabe, T., Nobusawa, S., Kleihues, P., Ohgaki, H., 2009. IDH1 mutations are early genetic events in diffuse lower-grade gliomas. N. Engl. J. Med. 360, 765–773.

Xie, Y., Zhang, M.Q., Minna, J.D., Gazdar, A.F., 2016. Validation of SCT methylation as a hallmark biomarker for lung cancers. J. Thorac. Oncol. 11, 346–354.

Yang, T., Zhang, M., Wang, Y.W., Ma, X., Shen, Y.N., Yuan, X., Yang, G.S., Wu, M.C., Lu, J.H., Shen, F., 2013. Elevated SHOX2 expression is associated with tumor recurrence of hepatocellular carcinoma. Ann. Surg. Oncol. 20 (Suppl 3), 5644–5649.

Zhang, C.B., Zhu, P., Yang, P., Cai, J.Q., Wang, Z.L., Li, Q.B., Zhao, Z.S., Zhang, W., Jiang, T., 2015. Identification of high risk anaplastic gliomas by a diagnostic and prognostic signature derived from mRNA expression profiling. Oncotarget 6, 36643–36651.

Zhang, Y.A., Ma, X., Sathe, A., Fujimoto, J., Wistuba, I.I., Lam, S., Yatabe, Y., Wang, Y.W., Stastny, V., Gao, B., Larsen, J.E., Girard, L., Liu, X., Song, K., Behrens, C., Kalhor, N., Xie, Y., Zhang, M., Minna, J.D., Gazdar, A.F., 2016. Validation of SCT methylation as a hallmark biomarker for lung cancers. J. Thorac. Oncol. 11, 364–369.