Please cite this article in press as: Hazra et al., Identification of glioblastoma stem cell-associated lncRNAs using single-cell RNA sequencing datasets, Stem Cell Reports (2023), https://doi.org/10.1016/j.stemcr.2023.10.004

Stem Cell Reports
Report
ISSCR
OPEN ACCESS

Identification of glioblastoma stem cell-associated lncRNAs using single-cell RNA sequencing datasets

Rasmani Hazra,[1,2,*] Raditya Utama,[1] Payal Naik,[1] Alexander Dobin,[1] and David L. Spector[1,*]
[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA
[2]University of New Haven, Biology & Environmental Science Department, West Haven, CT 06516, USA
*Correspondence: rhazra@cshl.edu (R.H.), spector@cshl.edu (D.L.S.)
https://doi.org/10.1016/j.stemcr.2023.10.004

SUMMARY

Glioblastoma multiforme (GBM) is an aggressive, heterogeneous brain tumor in which glioblastoma stem cells (GSCs) are known culprits of therapy resistance. Long non-coding RNAs (lncRNAs) have been shown to play a critical role in both cancer and normal biology. A few studies have suggested that aberrant expression of lncRNAs is associated with GSCs. However, a comprehensive single-cell analysis of the GSC-associated lncRNA transcriptome has not been carried out. Here, we analyzed recently published single-cell RNA sequencing datasets of adult GBM tumors, GBM organoids, GSC-enriched GBM tumors, and developing human brain samples to identify lncRNAs highly expressed in GSCs. We further revealed that the GSC-specific lncRNAs *GIHCG* and *LINC01563* promote proliferation, migration, and stemness in the GSC population. Together, this study identified a panel of uncharacterized GSC-enriched lncRNAs and set the stage for future in-depth studies to examine their role in GBM pathology and their potential as biomarkers and/or therapeutic targets in GBM.

INTRODUCTION

Glioblastoma multiforme (GBM; a World Health Organization grade IV glioma) ranks as the deadliest primary malignant brain cancer, with nearly 25,700 new cases diagnosed each year in the United States (Ostrom et al., 2021). The currently available treatment regimen comprises maximal surgical resection and a combination of radiotherapy and chemotherapy, which extends patient survival to a median of only 14.6 months (Prager et al., 2020). Despite recent progress in understanding the tumor's biology and multimodal therapy options, GBM remains one of the most treatment-resistant malignancies, and even after successful treatment, the tumor inevitably recurs. Intratumor heterogeneity and therapeutic resistance in GBM are thought to be promoted by a cell population with stem-cell-like properties, including self-renewal and differentiation (Singh et al., 2003, 2004; Lathia et al., 2015). GSCs are characteristically similar to normal neural (NE) stem cells, as they possess self-renewal capacity and express the PROMININ-1 gene (PROM1), a marker for normal NE stem cells (Singh et al., 2004). The radiotherapy-resistant and chemotherapy-resistant GSC populations in GBM have been successfully isolated from GBM patient samples using cell surface membrane markers, including PROM1 (Singh et al., 2004; Piccirillo et al., 2006; Lathia et al., 2010), CD44 (Anido et al., 2010), FUT4 (Son et al., 2009), CD49f (Barbar et al., 2020), L1CAM (Bao et al., 2008), PDGFRA (Kim et al., 2012), EGFR (Emlet et al., 2014), and A2B5 (Tchoghandjian et al., 2010). Recent advances in technology, including single-cell RNA sequencing (scRNA-seq) and lineage-tracing techniques, have provided further evidence for the existence of GSC populations in GBM tumors (Patel et al., 2014). An increasing body of evidence has shown that GSCs play an indispensable role in tumor initiation, progression, treatment resistance, and recurrence (Singh et al., 2004; Bao et al., 2006; Chen et al., 2012), indicating that GSCs are a critical therapeutic target.

Until relatively recently, proteins have traditionally been thought to be the critical regulators of biological processes; however, there is mounting evidence that non-coding RNAs also play important regulatory roles in almost all biological processes (Batista and Chang 2013; Beermann et al., 2016; Slack and Chinnaiyan 2019). Advances in sequencing technologies have enabled the discovery of large numbers of non-coding RNA genes in the mammalian genome, which are transcribed but whose resulting transcripts do not code for proteins (Carninci et al., 2005; Katayama et al., 2005; Consortium et al., 2007). Long non-coding RNAs (lncRNAs) are defined as transcripts greater than 200 nt in length with no protein-coding potential because of their lack of an open reading frame (Wang and Chang 2011; Rinn and Chang 2012). They are transcribed by RNA polymerase II in the sense or antisense orientation and commonly originate from intergenic regions (Ulitsky and Bartel 2013). LncRNAs have been implicated in numerous regulatory molecular functions, including modulating transcriptional patterns, regulating protein activities, playing structural or organizational roles, altering RNA processing events, and serving as precursors to small RNAs (Guttman et al., 2010, 2011; Yao et al., 2019; Statello et al., 2021; Mattick et al., 2023). Many lncRNAs are expressed in a tissue-specific manner, and some show significant changes in expression during differentiation (Guttman et al., 2011; Bergmann et al., 2015). Moreover, altered lncRNA expression has been

Please cite this article in press as: Hazra et al., Identification of glioblastoma stem cell-associated lncRNAs using single-cell RNA sequencing datasets, Stem Cell Reports (2023), https://doi.org/10.1016/j.stemcr.2023.10.004

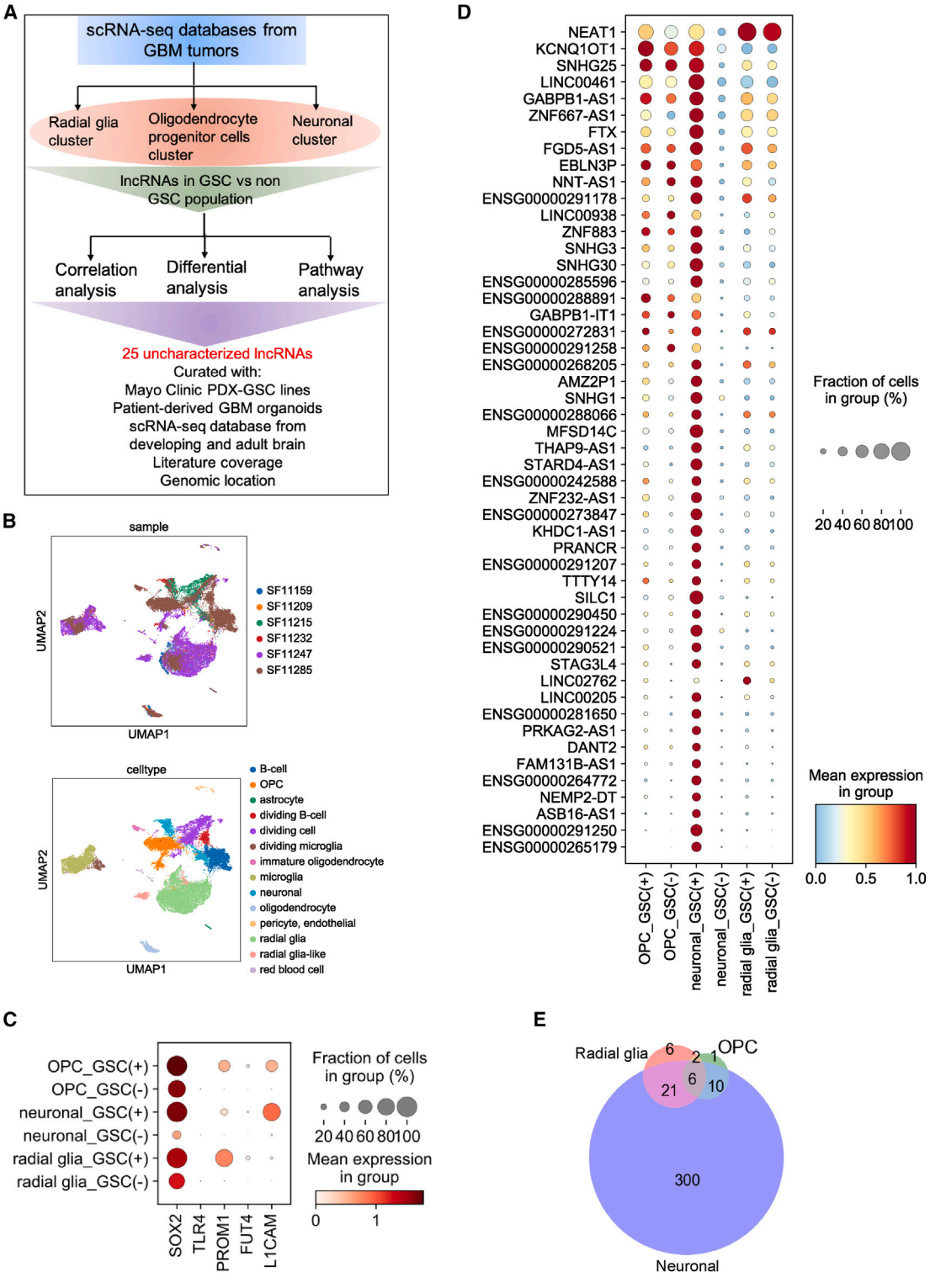


Figure 1. Identification of lncRNAs in multiple GSC populations in GBM tumors

(A) Workflow of GSC-enriched lncRNA analysis. The single-cell GBM dataset was reanalyzed to capture GSC populations in radial glia, oligodendrocyte precursor cells and neuronal clusters. Differential, pathway, and correlation analyses were performed on GSC vs. non-GSC

(legend continued on next page)

observed in several diseases, including cancer (Yang et al., 2014).

Individual lncRNA expression is often highly restricted to particular brain regions, and it has been suggested that lncRNAs offer more information about cell type identity during mammalian cortical development than protein-coding genes (Liu et al., 2016b). In addition, recent studies have shown that normal neurodevelopment in the human brain mirrors GBM development (Curry and Glasgow 2021). Therefore, the vast repertoire of lncRNAs in normal brain tissue indicates their potential importance in the context of GBM heterogeneity and treatment resistance. However, despite the large number of lncRNAs expressed in the brain, thus far just a handful have been shown to be involved in GSC biology. For example, the most abundant lncRNA, *MALAT1*, regulates proliferation and stemness in GSCs by activating the ERK/MAPK signaling pathway (Han et al., 2016). Knockdown (KD) of *XIST* in GSCs decreases proliferation, migration, and invasion via upregulation of miR-152 (Yao et al., 2015). Another abundant lncRNA, *NEAT1*, is highly expressed in GSCs. Depleting *NEAT1* inhibits GSC proliferation, migration, and invasion via upregulation of let-7e expression and downregulation of NRAS protein (Gong et al., 2016). Down-regulation of *SOX2OT* in GSCs reduced their proliferation, migration, and invasion, and induced apoptosis in GSCs (Su et al., 2017). In addition, the maternally expressed lncRNA *H19* is highly expressed in GSCs and GBM tissue and is involved in maintaining GSC stemness and proliferation. However, a detailed functional analysis of this mechanism has not been elucidated (Jiang et al., 2016). Given these few proof-of-concept examples, a deeper screening and identification/characterization of novel lncRNAs is required to identify appropriate GSC-specific targets in GBM.

Here, we identified a large number of presently uncharacterized lncRNAs associated with GSC populations in GBM. We performed a systematic bioinformatics analysis using three scRNA-seq databases of GBM patient tumors, GBM-patient-derived organoids (GBOs), and normal developing human brain (Nowakowski et al., 2017; Bhaduri et al., 2020; Jacob et al., 2020). We observed three main cell populations: radial glia (RG), oligodendrocyte precursor cells (OPCs), and neuron populations. Furthermore, we identified GSC and non-GSC cell types in every cluster using the cancer stem cell markers PROM1, CD44, FUT4, L1CAM, and SOX2 (Lathia et al., 2015). Using this analysis, we identified a large number of presently uncharacterized lncRNAs enriched with the GSC populations. Next, we validated the expression and subcellular localization of four selected lncRNAs using patient-derived GSC lines using qRT-PCR, single-molecule RNA fluorescence *in situ* hybridization (FISH), and subcellular fractionation. In addition, as proof of principle, we independently depleted the expression of two lncRNAs (*GIHCG* and *LINC01563*) in the GSC population and revealed that they each promote GSC proliferation, migration, and maintenance of stemness. In summary, our analyses identified a large number of lncRNAs associated with the GSCs transcriptome which should be pursued at the functional level and potentially as therapeutic targets in GBM.

RESULTS

Identification of lncRNAs in multiple GSC populations in GBM tumors

To identify novel lncRNA transcripts associated with GSCs in GBM tumors, we analyzed a publicly available scRNA-seq dataset of GBM tumors (Bhaduri et al., 2020). This dataset sequenced 32,877 cells from 11 GBM tumors obtained by surgical resection and processed immediately for RNA-seq (Bhaduri et al., 2020). Because of stark differences in cell number, we selected cell clusters to maintain proper balance in the dataset and to avoid sample bias in the downstream analysis (Figure 1A). We ran sequence alignment, clustering, and visualization pipelines to generate optimal lists of differentially expressed lncRNAs for various cell types (Table S1). Uniform manifold approximation and projection (UMAP)-based dimensionality reduction improved cluster visibility compared with t-distributed stochastic neighbor embedding (t-SNE), which helped decompose cell types and states. To identify the known cell types for each GBM cluster, we further applied cell-type markers (Figure S1A) provided by single nuclei transcriptomic analysis of developing human

populations from which significant differentially expressed lncRNAs were identified. Quantitative and qualitative filtering were applied using various algorithms to identify enriched lncRNAs.

(B) Sample composition and cell types were identified on UMAP using a combination of previously known markers including radial glia, oligodendrocyte precursor cells, neuronal cells, B cells, dividing cells, endothelial cells, pericytes, oligodendrocytes, astrocytes, and red blood cells.

(C) Expression and cellular fraction using GSC markers indicate a significant difference between GSC and non-GSC samples in the 3 main cell types.

(D) List of 50 significant differentially expressed GSC lncRNAs in the 3 cell types summarized on a dot plot providing comparisons in scaled expression and cellular fraction. Genes were clustered by their scaled expression across cells.

(E) Venn diagram representing the number of overlapping and unique differentially expressed lncRNAs from all three cell populations.

brain tissue (Nowakowski et al., 2017). From this analysis, we identified visually distinct main cell types enriched in given markers, including RG (identified by GFAP and HOPX), OPCs (PDGFRA), neuronal (DLX1, DCX, and NHLH1), microglia (C1QA), oligodendrocytes (MBP), astrocytes (GFAP), dividing cells (MKI67), pericytes (RGS5), endothelial cells (PECAM1), B cells (CD19), and red blood cells (GYPA) (Figures 1B and S1A). We identified a total of 1,426 expressed lncRNAs in GBM tumors. Of these 1,426 lncRNAs, 549 lncRNAs are associated with RG, 328 lncRNAs are in the neuronal cluster, and 549 lncRNAs are associated with OPC clusters (Figure S1B; Table S2). Next, for every cell type, we classified cells as either GSC positive (GSC) or GSC negative (non-GSC) using the stemness markers PROM1, FUT4, and L1CAM, in conjunction with SOX2 and not expressing TLR4 (Gimple et al., 2019) (Figure 1C). TLR4 is downregulated in GSCs (Alvarado et al., 2017) (Figure 1C). GSC-specific marker genes were extracted by implementing differential analysis on each cell type (see experimental procedures). Given the low expression level of most lncRNAs, we filtered the list solely on the basis of a significant false discovery rate (FDR), with values less than 0.1 (see experimental procedures). Furthermore, our analysis observed a small number of defined GSCs because of the stringent combination of GSC markers. Therefore, we combined and thus reduced the number of clusters to increase statistical power in differential analysis. This analysis identified three GSC populations: RG, OPCs, and neurons. We identified 35 lncRNAs significantly enriched in the RG GSC population out of 391 differentially expressed lncRNAs. Among these lncRNAs, we found that uncharacterized lncRNAs, including *LINC01088*, *LINC02762*, *GIHCG* (gradually increased during hepatocarcinogenesis), *ENSG00000288764*, and *ENSG00000274265*, were highly expressed in the RG population (Figures 1D, 1E, and S1C; Table S3). Similarly, 19 lncRNAs associated with GSCs were significantly enriched in the OPC population, and 337 lncRNAs were enriched in the neuronal population (Figures 1D, 1E, and S1C; Table S3). Among these lncRNAs, we found 3 uncharacterized lncRNAs (*LINC00667*, *LINC00662*, and *LINC02969*) that were highly expressed in the OPC population and more than 100 uncharacterized lncRNAs (Table S2) were highly expressed in the neuronal population. Our finding that the highest number of lncRNAs are enriched neuronal GSC population is consistent with previous findings that lncRNAs are generally expressed in a cell-type-specific manner (Cabili et al., 2011) and the majority of lncRNAs identified in mice to date are specifically expressed in neuronal cells (Mercer et al., 2008). Moreover, using the encephalization quotient (EQ), an indicator of brain size, a nonlinear correlation between the increased number of lncRNAs and brain size has been shown (Clark and Blackshaw 2017). In addition, it has shown that a large number of cognitive phenomena in the brain are associated with coordinated interactions of large numbers of neurons (Uhlhaas and Singer 2006). Therefore, considering all of these findings we hypothesize that the diversity and function of lncRNAs are critical to the complex cognitive ability of the brain and this may be due in part to the larger number of lncRNAs in neuronal cells.

We further compared the lists of significant lncRNAs across all three GSC populations to generalize these findings. We observed only 6 overlapping lncRNAs (*MEG8*, *CRNDE*, *MIR99AHG*, *HOTAIRM1*, *LINC00632*, and *CD27-AS1*) among all three populations: RG, OPC, and neuron (Figures 1B and S1B; Table S3). Interestingly, these lncRNAs are comparatively well studied and dysregulated in various cancers (Ellis et al., 2012; Terashima et al., 2018; Luo et al., 2020; Zhao et al., 2020; Tao et al., 2021; Zhou and Ma 2022). In addition, we identified 16 common lncRNAs between the OPC and neuron populations, whereas 27 common lncRNAs were identified between the RG and neuronal populations, and 2 common lncRNAs were identified between the RG and OPC populations (Figure 1E; Table S3). We further *in silico* validated this existing dataset with another publicly available dataset (Wang et al., 2020), in which 12,367 cells were used from 3 patients. Using the same analyses, we have identified similar cell type populations, including three main clusters: RG, OPCs, and neuronal (Figure S1D). Interestingly, we found qualitative agreements across three publicly available datasets despite the challenges of reproducibility in scRNA-seq. Additionally, in our analyzed data (Figures S1A and S1D), we observed that DCX (a neuronal marker) overlaps extensively with PDGFFRA and MBP oligodendrocyte markers), consistent with the analysis from Bhaduri et al. (2020). One possible explanation of these co-expression markers is the difference between "cell type identity" and "cell state" (Quake 2021). Therefore, we referred to the annotated cell populations here as a "type" but not a "cell state."

Comparative analyses of lncRNA expression levels across public GBM, GBO, and TCGA datasets

We next corroborated our uncharacterized GSC-associated lncRNA set by performing a comparative analysis using scRNA-seq datasets from 8 GBO lines, The Cancer Genome Atlas (TCGA), a bulk RNA-seq database of patient-derived xenograft (PDX) GBM tumors, and a scRNA-seq dataset of the normal developing brain (Nowakowski et al., 2017; Jacob et al., 2020). The GBO lines displayed primary driver mutations in EGFR, NF-1, and PI3K (Jacob et al., 2020). By implementing similar pipelines, we were able to identify similar cell-type clusters in the organoid dataset using identical cell-type markers, such as RG (GFAP, HOPX, and VIM), OPC (PDGFRA), and neuronal clusters (DCX and DLX1) (Figures 2A, 2B, and S2A). We have also observed a highly distinct, dividing RG cell signature (GFAP and HOPX)

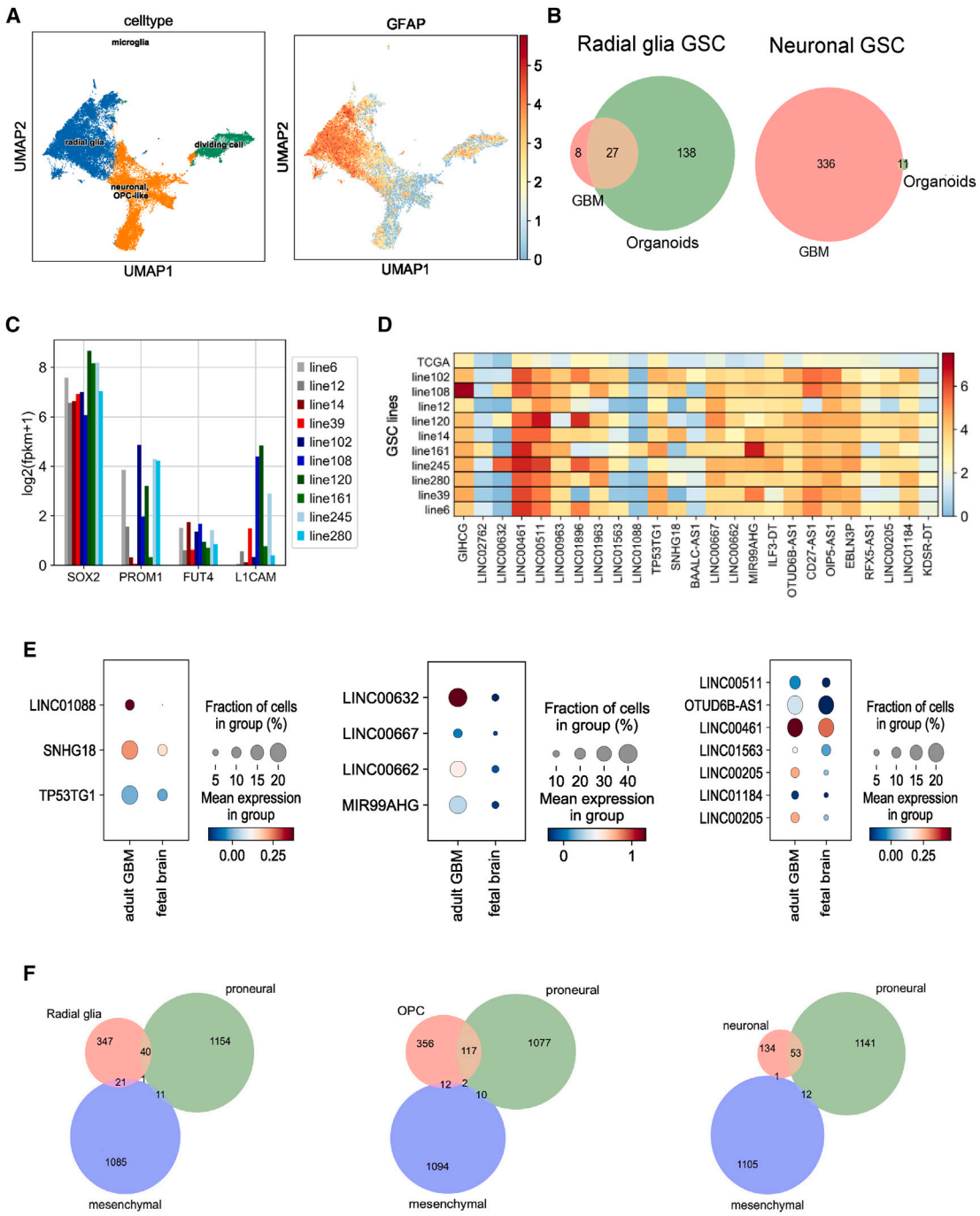


Figure 2. Comparative analyses of lncRNA expression levels across public GBM and GBO datasets

(A) Identification of a radial glial population in GBM-derived organoid dataset using known markers (GFAP).

(B) Venn diagrams representing the number of overlapping and unique differentially expressed lncRNAs in GBM tissue and GBM organoids.

(C) Expression levels of GSC markers across 10 Mayo Clinic cell lines from patient-derived xenograft (PDX) GBM samples.

(D) Expression levels of 25 uncharacterized significant differentially expressed GSC lncRNAs are shown on heatmap across 10 Mayo Clinic cell lines and bulk TCGA GBM dataset.

(E) Comparison between GBM and normal fetal brain are presented with the expression change and cellular fraction of differential GSC lncRNAs across 3 main cell types.

(F) Venn diagrams representing the overlaps between upregulated lncRNAs that are significant (FDR < 0.1) in mesenchymal and/or proneural subtypes and expressed in at least 5% of cells in radial glia, OPC, and neuronal GSC populations.

Please cite this article in press as: Hazra et al., Identification of glioblastoma stem cell-associated lncRNAs using single-cell RNA sequencing datasets, Stem Cell Reports (2023), https://doi.org/10.1016/j.stemcr.2023.10.004

identified by the MKI67 marker (Figure S2A), consistent with the analysis of Bhaduri et al. (2020). However, the expression of PDGFRA failed to capture a distinct OPC-like population (Figure S2A), which may be due to the selection in the cultured organoid medium. However, it requires a substantial functional assays for future study. We next compared the sets of GSC positive differentially expressed genes in RG, OPC, and neuron populations between GBM tumor and tumor-derived organoid datasets. Furthermore, by comparing GBM tumor and organoid datasets, we found 27 common lncRNAs in the RG population (including *GIHCG*, *LINC01088*, *LINC027622*, and *SNHG6*) and only one common lncRNA (*MIAT*) in the neuronal GSCs and no common lncRNAs in the OPC population (Figure 2B; Table S4). We observed a distinct RG population using the markers GFAP and HOPX in both tumor and tumor-derived organoid samples, but we did not observe distinct populations of OPC and neuronal GSCs in organoids using the markers PDGFRA (OPC) and DLX1, DCX, and NHLH1 (neuronal). This limitation may be due to variation in selection in organoid culture.

In general, the low expression of lncRNAs in most tissues, including GBM tissues, has been a great challenge in experimental validation. Thus, we selected the list of lncRNAs that have higher expression in tumor tissue and PDX GSC lines using TCGA and PDX bulk RNA-seq datasets (Vaubel et al., 2020). We also showed that the GSC markers SOX2, PROM1, FUT4, and L1CAM were expressed in GSC lines (Figure 2C). Together with TCGA data, we selected the 25 uncharacterized intergenic lncRNAs that are differentially expressed in GSCs and have high expression levels ($p < 0.05$ and $\log_2$[fold change] $\geq$ 2) (Figure 2D). Furthermore, to investigate the expression levels of GSC-associated lncRNAs in the normal human brain, we compared the expression level between two independent datasets generated from adult GBM tumors (Bhaduri et al., 2020; Wang et al., 2020), and a normal developing human brain dataset (Nowakowski et al., 2017). It is known that developing human brain expresses thousands of lncRNAs and abnormal lncRNA expression has been associated with a wide range of neurological disorders (Andersen and Lim 2018). Interestingly, we observed many of the lncRNAs identified from each of the three major clusters (RG, OPC, and neuronal) to be more highly expressed in the adult GBM tissue compared with the normal brain (Figures 2E and S2B), suggesting that these lncRNAs are upregulated in the GSC populations in tumor but not in the normal human brain.

Genome-wide transcriptome analysis revealed molecular classification of GBM into four major subtypes: mesenchymal (MES), classical (CL), proneural (PN), and NE on the basis of bulk tumor transcription profiles (Verhaak et al., 2010; Guardia et al., 2020). We used GBM samples from TCGA. Here we analyzed raw RNA-seq files to capture GBM subtype markers (see experimental procedures) and overlapped the markers with the list of lncRNAs expressed in at least 5% of each GSC population (RG, OPC, and neuronal). The choice of GBM subtype markers provided a distinct signature with the optimal number of clusters (Figure S2C). Thus, the clusters were annotated to define the subtypes and to predict the associated differential genes. We showed the overlaps between upregulated MES and/or PN lncRNAs and each GSC population lncRNAs (Figure 2F). We found 40 significant lncRNAs in PN and 21 in MES subtypes expressed in at least 5% of the RG GSC population. Similarly, there are 117 and 12 lncRNAs expressed in the OPC GSC population coinciding with PN and MES subtypes, respectively (Table S5). Finally, we identified 53 and 1 lncRNA expressed in neuronal GSC population overlapping with the PN and MES subtypes, respectively (Figure 2F; Table S5). Among the overlapping lncRNAs, *GIHCG* is expressed in both RG and OPC GSC populations while being simultaneously enriched in the PN subtype. In addition, these findings suggest that neuronal and OPC GSC transcriptional profiles are more similar to the PN subtype than the MES subtype.

Pathway analysis using GSC vs. non-GSC populations suggests involvement of the RG population in metabolic-related processes

To further investigate the underlying biological functions of the GSC-enriched lncRNAs in GBM, Gene Ontology (GO) term, Kyoto Encyclopedia of Genes and Genomes (KEGG), REACTOME, and HALLMARK pathway analyses were conducted using all genes ranked by statistical score (see experimental procedures) generated from GSC vs. non-GSC comparison separately in all three cell types (RG, OPC, and neuron). Interestingly, we did not observe any common pathways among all three GSC-associated cell types (Figures 3A–3C), suggesting the functional heterogeneity of GSC populations in GBM tumors. Biological pathways related to metabolic processes (oxidative phosphorylation, reactive oxygen species pathway, inner mitochondrial membrane protein complex, ATP synthesis coupled electron transport, and NADH dehydrogenase activity) were significantly enriched in the RG GSC population; on the other hand, OPC and neuronal populations were enriched in different pathways involved in cancer development and progression, such as angiogenesis, epithelial-to-MES transition, apoptosis, MTORC1 signaling, Hedgehog signaling, TGFB signaling, PI3K-AKT-MTOR signaling, G2M checkpoint, DNA repair, and E2F targets (Figures 3A–3C). Notably, the metabolic characteristics of GSC populations are highly heterogeneous in nature. The RG population is enriched in oxidative phosphorylation or mitochondrial metabolic pathway (Figure 3A). In contrast, the OPC population is enriched in the glycolytic pathway (Figure 3B), indicating the

Please cite this article in press as: Hazra et al., Identification of glioblastoma stem cell-associated lncRNAs using single-cell RNA sequencing datasets, Stem Cell Reports (2023), https://doi.org/10.1016/j.stemcr.2023.10.004

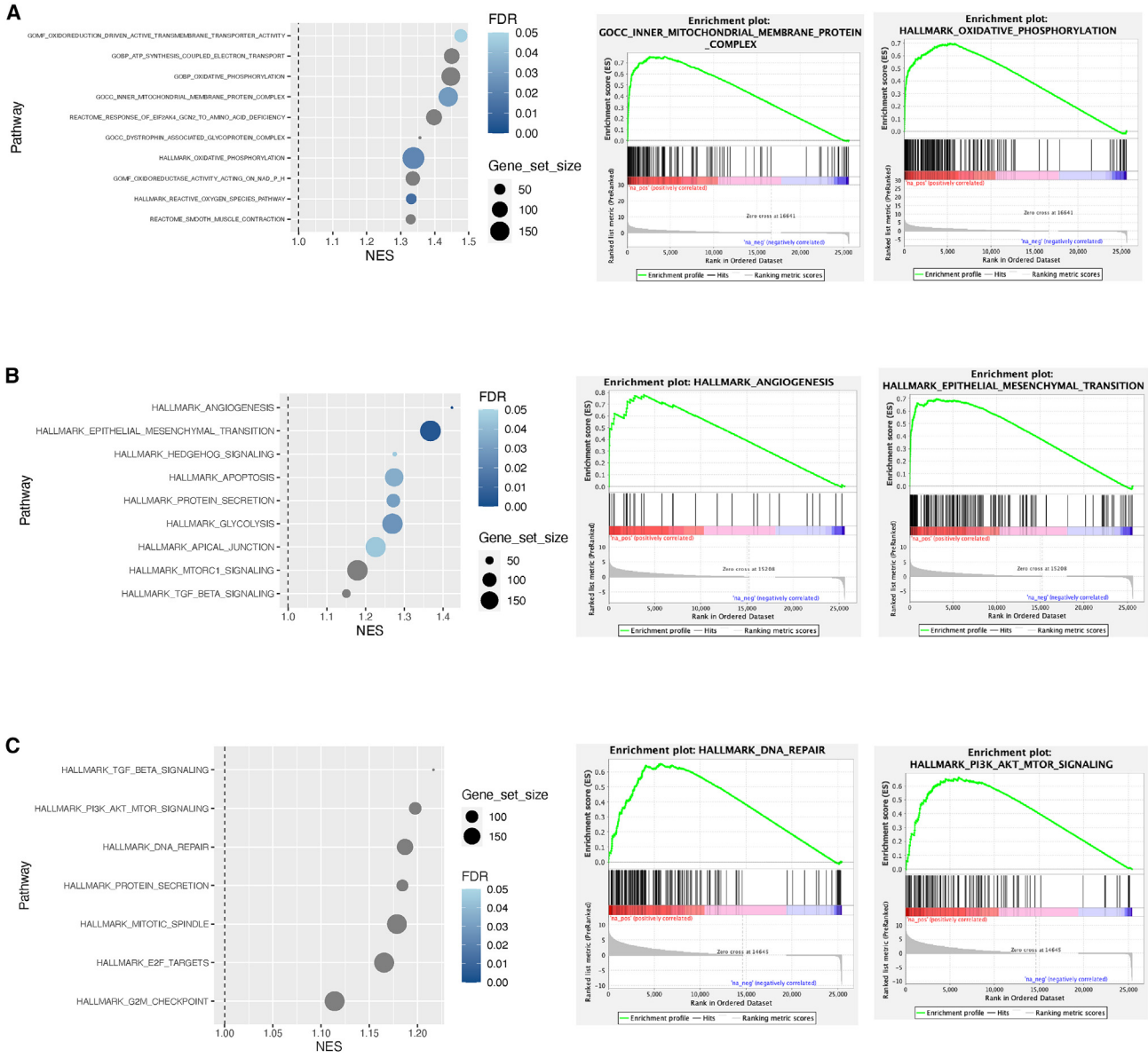


Figure 3. Pathway analysis using GSC vs. non-GSC populations suggests involvement of the radial glia population in metabolic-related processes

Radial glia GSC expression change is associated with pathways enriched in various metabolic functions including inner mitochondrial membrane protein complex and oxidative phosphorylation.

(B) Oligodendrocyte precursor GSC expression change is associated with pathways including angiogenesis and epithelial-to-mesenchymal transition.

(C) Neuronal GSC expression change is associated with pathways including TGF-beta signaling, DNA repair, and PI3K-AKT-MTOR signaling.

complex heterogeneous tumor microenvironment even within the same tumor.

Validation of GSC-associated lncRNAs

We first validated the PDX GSC lines using established GSC markers (NESTIN, SOX2, and PROM1) (Gimple et al., 2019). Nestin is a class six intermediate filament proteins expressed in NE stem cells (Lendahl et al., 1990) and is expressed in the cytoplasm of GSC lines (Figures 4A and S4A). SOX2 is a transcription factor that plays a critical role in maintaining the self-renewal capability of NE stem cells. Its activity is associated with maintaining the undifferentiated state of cancer stem cells in several tissues (Novak et al., 2020). We

Please cite this article in press as: Hazra et al., Identification of glioblastoma stem cell-associated lncRNAs using single-cell RNA sequencing datasets, Stem Cell Reports (2023), https://doi.org/10.1016/j.stemcr.2023.10.004

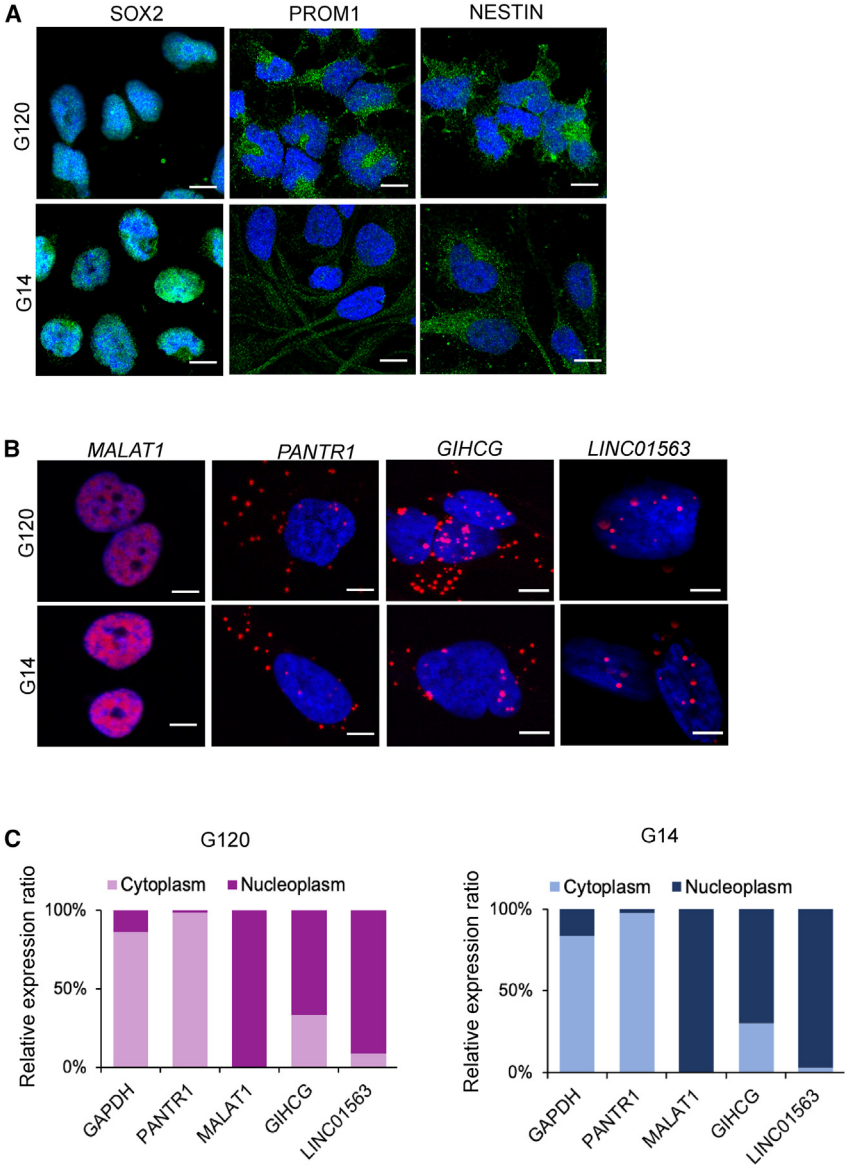


Figure 4. Validation of GSC-associated lncRNAs

(A) Immunolabeling of SOX2, PROM1, and NESTIN in GSC lines (GSC-120 and GSC-14). Scale bar, 12 μm.

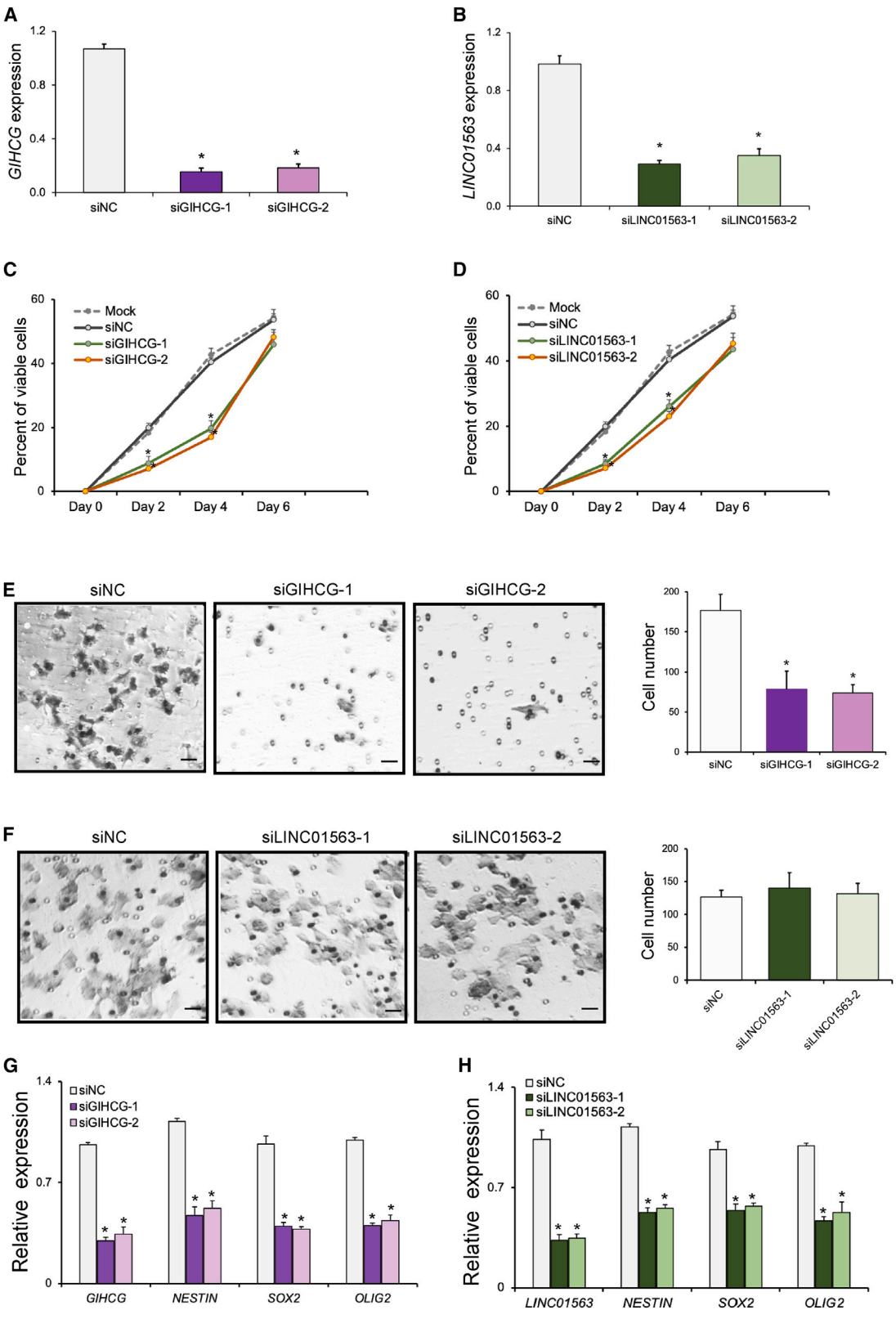
(B) Single-molecule RNA FISH images indicate localization of *MALAT1*, *PANTR1*, *GIHCG*, and *LINC01563* lncRNA transcripts (red dots) within the nuclei and/or cytoplasm in GSCs. Scale bars, 12 μm.

(C) Subcellular fractionation followed by qRT-PCR confirming the localization of *PANTR1*, *GIHCG*, and *LINC01563* lncRNA transcripts. *GAPDH* and *MALAT1* were used as cytoplasmic and nuclear markers for quality control, respectively. Data are presented as mean value ± SEM (n = 3 independent experiments).

found that SOX2 is highly enriched in the nucleus of all four GSC lines (Figures 4A and S4A). PROMININ1 (PROM1) is a transmembrane glycoprotein commonly used as a marker of normal and cancerous stem cells, particularly in central nervous system tumors, including GBM (Holmberg Olausson et al., 2014). On the other hand, PROM1 is expressed in both the cytoplasm and the cell surface of GSC lines (Figures 4A and S4A). To validate our GSC-specific lncRNA expression patterns, we performed single-molecule *in situ* hybridization and a cell fractionation assay followed by qRT-PCR for three lncRNAs—*MALAT1*, *GIHCG*, and *LINC01563*—and observed that *MALAT1* and *LINC01563* are highly enriched in the nuclear fraction, whereas *GIHCG* is enriched in both the cytoplasm and nuclear fractions (Figures 4B, 4C, S4B, and S4C).

GSC-associated lncRNAs promote GSC proliferation, migration, and stemness

To determine the functions of both *GIHCG* and *LINC01563* in GSC, we performed KD analysis using two independent, non-overlapping small interfering RNAs (siRNAs) targeting both *GIHCG* and *LINC01563* (designated siGIHCG-1 and siGIHCG-2; siLINC01563-1 and siLINC01563-2). The KD efficiency was measured using qRT-PCR. We found that both *GIHCG* and *LINC01563* were significantly downregulated



(legend on next page)

Please cite this article in press as: Hazra et al., Identification of glioblastoma stem cell-associated lncRNAs using single-cell RNA sequencing datasets, Stem Cell Reports (2023), https://doi.org/10.1016/j.stemcr.2023.10.004

in GSCs compared with a control siRNA-non-targeting control RNA (siNC) (Figures 5A and 5B). Next, we performed a CellTiter-Glo 2D cell viability assay using GIHCG-depleted GSCs and found that loss of *GIHCG* markedly reduced cell proliferation both at day 2 (56%) and at day 4 (61%) (Figures 5C and S5A). KD of *LINC01563* in GSCs significantly reduced proliferation at day 2 (42%) and at day 4 (53%) (Figures 5D and S5B). As cell migration is a critical process in GBM progression (Lefranc et al., 2005), we were interested in determining whether the loss of *GIHCG* and *LINC01563* in GSC might play a role in the migration event. We used a Transwell migration assay and found that loss of *GIHCG* in GSC resulted in a 44% reduction in migration ability compared with control GSC (Figure 5E). On the other hand, we did not observe any migratory defect upon deletion of *LINC01563* in GSC (Figure 5F). Furthermore, to determine the role of *GIHCG* and *LINC01563* in the maintenance of GSC stemness, we measured the expression of GSC markers, SOX2, NESTIN, and OLIG2 (Wang et al., 2018) using qRT-PCR and determined that KD of both lncRNAs significantly reduced the expression of stem markers compared with control (Figures 5G and 5H). Together, our results reveal that the lncRNAs *GIHCG* and *LINC01563* are required for GSC growth and maintenance.

DISCUSSION

The vast majority of the human genome is transcribed and among the transcriptome are thousands of lncRNAs, many of which play important roles in regulating different cellular processes including neurogenesis, neuronal differentiation, and brain development (Ma and Chang 2016; Mattick et al., 2023). In addition, there is increasing evidence that lncRNAs act as biomarkers, tumor suppressors, and oncogenes in various cancers (Cheetham et al., 2013; Huarte 2015; Arun et al., 2016; Diermeier et al., 2016; Chen et al., 2021). Recently, some *in silico* studies have identified lncRNAs in GBM tumors and cell lines using bulk RNA-seq from TCGA and microarray platforms (Liu et al. 2016a, 2020; Reon et al., 2016; Paul et al., 2018; He et al., 2020), although very few of them have been characterized in depth regarding their potential regulatory function in GBM. However, the bulk RNA-seq approach limits our understanding of GBM tumor heterogeneity, given its cellular complexity. Recent transcriptome analysis revealed that GBM tumors are highly heterogeneous (Patel et al., 2014). Moreover, lncRNAs tend to be expressed in a cell type- and/or tissue-specific manner, and remarkably, the highest number of conserved lncRNAs are expressed in the brain (Aprea and Calegari 2015). However, when they are examined using bulk tissue analysis, their average expression levels are lower compared with those of mRNAs. The possible explanation for this low expression of lncRNAs is either they are expressed at low levels in all cells uniformly or they are more highly expressed and functional in only a subset of cells, which is averaged out in bulk experimental studies (Mercer et al., 2008; Johnson et al., 2015; Molyneaux et al., 2015). As many lncRNAs are expressed in a cell-type-specific manner, identifying novel lncRNAs in different cell-type clusters by scRNA-seq provides a means to study their regulatory role among different cell types in GBM tumors and may offer a downstream lncRNA-based therapeutic approach.

Thus, we set out to provide a comprehensive analysis of the GSC transcriptome using scRNA-seq databases of GBM tumors and tumor-derived organoids. Although we confirmed several known cancer-associated lncRNAs, such as *NORAD*, *NEAT1*, *MALAT1*, *XIST*, *PVT1*, *MEG3*, *MEG8*, and *GAS5*, in our screen, most of the 1,426 differentially expressed lncRNA candidates have not been described previously in the context of glioblastoma. In addition, we have identified for the first time 374 uncharacterized lncRNAs associated with GSCs. In addition, we identified 61 significant lncRNAs (Table S5) that are associated with PN and MES GBM subtypes. Our study emphasizes the importance of unbiased screening approaches and represents a valuable resource for further study of the regulatory roles of GSC-associated lncRNAs in GBM initiation, maintenance, and progression.

A DAVID GO term pathway analysis using GSC vs. non-GSC populations from three different clusters (RG, OPC, and neuronal) showed that differentially expressed genes were significantly enriched in different cancer-related

Figure 5. GSC-associated lncRNAs promote GSC proliferation, migration, and stemness

(A and B) qRT-PCR analysis of siRNA knockdown of *GIHCG* (A) and *LINC01563* (B) in GSCs. A non-targeting control (NC)-siRNA was used as a control.

(C and D) CellTiter-Glo 2D luminescent cell viability assay. 10,000 GSCs were seeded into 96-well plates, and cell numbers were measured using the CTG assay at the indicated time points. The percentage of viable GSC populations (control vs. *GIHCG* [C] and *LINC01563* [D] knockdown-GSC populations) are shown ± SD (n = 3). *p < 0.05.

(E and F) Transwell migration assay: depletion of *GIHCG* and *LINC01563* significantly decreased cell migration. Scale bar: 150 μm. The average migratory GSCs (control vs. *GIHCG* and *LINC01563*-KD GSCs) is shown ± SD (n = 3). *p < 0.05.

(G and H) qRT-PCR analysis of stem markers, including NESTIN, SOX2, and OLIG2, using control (siNC) vs. *GIHCG*-KD (G) and *LINC01563*-KD GSCs. Data are presented as mean ± SEM from three independent experiments. **p < 0.05.

pathways. Interestingly, the RG population was enriched in the electron transport chain and oxidative phosphorylation pathways and mitochondrial metabolism (Figures 4A and S4A), consistent with the recent findings that oxidative phosphorylation or mitochondrial metabolism are the preferred energy source for cancer stem cells (Tuy et al., 2021). On the other hand, the glycolytic pathway is significantly enriched in the OPC GSC population, indicating that the metabolic state of GSCs differs in different cell populations. This finding demonstrates that metabolic characteristics in GSCs are highly heterogeneous within the tumor. Studies have shown that cancer stem cells can switch their metabolic state to favor either oxidative metabolism or glycolysis, and this switch relies on the tumor microenvironment (Zhu et al., 2020). Thus, targeting GSCs metabolism by potentially manipulating lncRNA expression may provide new and effective methods for treating GBM tumors.

This study allowed us to short list a number of presently uncharacterized GSC-associated lncRNAs that can serve as novel biomarkers for the diagnosis and prognosis of GBM. As proof of principle, we identified uncharacterized lncRNAs, GIHCG and LINC01563, that affect cell proliferation, migration, and maintenance of stemness in GSC. We suggest that these lncRNAs may be likely drivers of GBM tumor progression and future in-depth studies of molecular mechanism will reveal their therapeutic potential in GBM.

In conclusion, our *in silico* comprehensive analysis identified a group of GSC-associated lncRNAs in GBM. Future in-depth mechanistic studies of these uncharacterized lncRNAs will extend the field's understanding of GBM biology and offer new potential lncRNA biomarkers and therapeutic targets for GBM. Exploring the functional relevance of lncRNAs in GBM biology will provide better insights into this devastating disease and its potential treatment.

EXPERIMENTAL PROCEDURES

Resource availability

Corresponding author

Requests for materials and further information should be directed to Rasmani Hazra (rhazra@cshl.edu) and David L. Spector (spector@cshl.edu).

Materials availability

This study did not generate new unique reagents. GSC lines used in this study were acquired through a material transfer agreement with the Mayo Clinic (Rochester, MN). Requests for these GSC lines must be made directly to Dr. Jann N. Sarkaria at the Mayo Clinic.

Data and code availability

No custom codes were generated in this study.

Dataset preparation and sequencing for scRNA-seq

All datasets used in this study are publicly available. For scRNA-seq, the original main dataset (Bhaduri et al., 2020) consisted of 32,877 cells from 11 GBM tumors obtained by surgical resection and processed immediately for RNA-seq. We also obtained a dataset (Jacob et al., 2020) from glioblastoma-derived organoids (GSE141947). For the validation of those previous datasets, we used a scRNA-seq dataset consisting of 3 GBM patient samples (Wang et al., 2020) (GSE139448). We further used a single cell RNA sequencing dataset (Nowakowski et al., 2017) from developing normal human brain to compare lncRNA expression in normal vs. GBM samples. We used RNA-seq glioblastoma PDX GSC line datasets from the Mayo Clinic: (https://www.cbioportal.org/study/clinicalData?id=gbm_mayo_pdx_sarkaria_2019).

GSC lines

PDX GSC lines (GSC #06, GSC #14, GSC #120, and GSC #161) were obtained from the Mayo Clinic (Zhao et al., 2019) and cultured on laminin-coated plates (1–2 μg/cm^2) (#L2020; Sigma) or as neurosphere in StemPro NE stem cell medium. A serum-free medium (SFM) kit (#A1050901; Thermo Fisher Scientific) was used. Maintenance of the stem cell phenotype was verified by the expression of the stem cell markers SOX2, NESTIN, and PROM1.

Single-molecule RNA FISH

Single-molecule RNA FISH was performed according to the manufacturer's protocol for the RNAscope Fluorescent Multiplex Reagent Kit 320850 (#320850; Advanced Cell Diagnostics), as described previously (Hazra et al., 2022; Hazra and Spector 2022). Briefly, 5 × 10^4 GSCs were seeded onto laminin-coated coverslips (12 mm) in a 24-well plate for 24 h to reach 80% confluence, then fixed in freshly prepared 4% paraformaldehyde (PFA) (#19200; Electron Microscopy Sciences). Fixed cells were permeabilized with different concentrations (50%, 75%, and 100%) of ethanol for 2 min and 0.05% Triton X-100 for 10 min on ice before hybridization. The hybridization and signal amplification steps were performed according to the manufacturer's instructions, and nuclei were counterstained with DAPI. GSCs were imaged using a Zeiss LSM780 point scanning confocal microscope.

Nucleofection and siRNA-mediated KD in GSCs

Nucleofection using the Lonza 4D Nucleofector unit was used for the transfection of siRNAs in GSCs. Briefly, GSCs were harvested and counted, and cell viability was determined using trypan blue staining on a Countess automated cell counter (Life Technologies). For each reaction, 2 × 10^6 viable cells were resuspended in P3 Cell Line Solution (Lonza), mixed with 1 μM non-targeting control or 1 μM target-specific siRNAs, and transferred to nucleocuvettes for nucleofection using program code DS-120. Cells were subsequently transferred onto laminin-coated cell culture plates containing pre-warmed and supplemented growth medium. The next day, medium was replaced with fresh medium.

ON-TARGET-plus custom-designed siRNAs for GIHCG and LINC01563 non-targeting siRNA were purchased from Dharmacon Inc. siRNA sequences are listed below:

GIHCG-1: GGACUUAUCCGUUGAAUGAUU

GIHCG-2: GCAUGUGAAGACAAUUCUUUUUU
LINC01563-1: GGAAAGAAGAUGAGAAUAAUU
LINC01563-2: CUGAAAUAGAAGAGGAAUUUU

Statistical analysis

Statistical tests were performed and analyzed using Microsoft Excel and ggplot2. p values were calculated using the two-tailed paired Student's t test. Significance was defined as $p < 0.05$. All data are presented as mean ± SEM.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stemcr.2023.10.004.

ACKNOWLEDGMENTS

We thank members of the Cold Spring Harbor Laboratory (CSHL) Next-Gen Sequencing and Analysis Shared Resource for supporting the analysis. We thank Tse-Luen (Erika) Wee, director of the CSHL Microscopy Shared Resource for assistance with imaging. We also thank Andrea McMahon and Ann C.M. Tuma from Jann N. Sarkaria's laboratory (Mayo Clinic) for providing PDX GSC lines. This research was supported by National Cancer Institute (NCI) grant 2P01CA013106-Project 3 (to D.L.S.) and NIH grant R01HG009318 (to A.D.). The authors acknowledge the CSHL Cancer Center Sequencing Technologies and Analysis Shared Resource and the Microscopy Shared Resource for services (NCI grant 5P3OCA45508). The sequencing analysis was performed using equipment purchased through NIH grant S10OD028632-01.

AUTHOR CONTRIBUTIONS

R.H. conceptualized and designed the research, performed the experiments, analyzed the data, and wrote the manuscript. R.U. did the bioinformatics analysis and wrote part of the experimental procedures and results sections of the manuscript. P.N. performed the experiments. A.D. reviewed and edited the manuscript. D.L.S. was responsible for acquisition of funding, reviewing, and editing the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 2, 2023
Revised: October 4, 2023
Accepted: October 5, 2023
Published: November 2, 2023

REFERENCES

Alvarado, A.G., Thiagarajan, P.S., Mulkearns-Hubert, E.E., Silver, D.J., Hale, J.S., Alban, T.J., Turaga, S.M., Jarrar, A., Reizes, O., Longworth, M.S., et al. (2017). Glioblastoma Cancer Stem Cells Evade Innate Immune Suppression of Self-Renewal through Reduced TLR4 Expression. Cell Stem Cell *20*, 450–461.e4.

Andersen, R.E., and Lim, D.A. (2018). Forging our understanding of lncRNAs in the brain. Cell Tissue Res. *371*, 55–71.

Anido, J., Sáez-Borderías, A., Gonzàlez-Juncà, A., Rodón, L., Folch, G., Carmona, M.A., Prieto-Sánchez, R.M., Barba, I., Martínez-Sáez, E., Prudkin, L., et al. (2010). TGF-beta Receptor Inhibitors Target the CD44(high)/Id1(high) Glioma-Initiating Cell Population in Human Glioblastoma. Cancer Cell *18*, 655–668.

Aprea, J., and Calegari, F. (2015). Long non-coding RNAs in corticogenesis: deciphering the non-coding code of the brain. EMBO J. *34*, 2865–2884.

Arun, G., Diermeier, S., Akerman, M., Chang, K.C., Wilkinson, J.E., Hearn, S., Kim, Y., MacLeod, A.R., Krainer, A.R., Norton, L., et al. (2016). Differentiation of mammary tumors and reduction in metastasis upon Malat1 lncRNA loss. Genes Dev. *30*, 34–51.

Bao, S., Wu, Q., Li, Z., Sathornsumetee, S., Wang, H., McLendon, R.E., Hjelmeland, A.B., and Rich, J.N. (2008). Targeting cancer stem cells through L1CAM suppresses glioma growth. Cancer Res. *68*, 6043–6048.

Bao, S., Wu, Q., McLendon, R.E., Hao, Y., Shi, Q., Hjelmeland, A.B., Dewhirst, M.W., Bigner, D.D., and Rich, J.N. (2006). Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature *444*, 756–760.

Barbar, L., Jain, T., Zimmer, M., Kruglikov, I., Sadick, J.S., Wang, M., Kalpana, K., Rose, I.V.L., Burstein, S.R., Rusielewicz, T., et al. (2020). CD49f Is a Novel Marker of Functional and Reactive Human iPSC-Derived Astrocytes. Neuron *107*, 436–453.e12.

Batista, P.J., and Chang, H.Y. (2013). Long noncoding RNAs: cellular address codes in development and disease. Cell *152*, 1298–1307.

Beermann, J., Piccoli, M.T., Viereck, J., and Thum, T. (2016). Non-coding RNAs in Development and Disease: Background, Mechanisms, and Therapeutic Approaches. Physiol. Rev. *96*, 1297–1325.

Bergmann, J.H., Li, J., Eckersley-Maslin, M.A., Rigo, F., Freier, S.M., and Spector, D.L. (2015). Regulation of the ESC transcriptome by nuclear long noncoding RNAs. Genome Res. *25*, 1336–1346.

Bhaduri, A., Di Lullo, E., Jung, D., Müller, S., Crouch, E.E., Espinosa, C.S., Ozawa, T., Alvarado, B., Spatazza, J., Cadwell, C.R., et al. (2020). Outer Radial Glia-like Cancer Stem Cells Contribute to Heterogeneity of Glioblastoma. Cell Stem Cell *26*, 48–63.e6.

Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. *25*, 1915–1927.

Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., et al. (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559–1563.

Cheetham, S.W., Gruhl, F., Mattick, J.S., and Dinger, M.E. (2013). Long noncoding RNAs and the genetics of cancer. Br. J. Cancer *108*, 2419–2425.

Chen, J., Li, Y., Yu, T.S., McKay, R.M., Burns, D.K., Kernie, S.G., and Parada, L.F. (2012). A restricted cell population propagates glioblastoma growth after chemotherapy. Nature *488*, 522–526.

Chen, Y., Hao, Q., Wang, S., Cao, M., Huang, Y., Weng, X., Wang, J., Zhang, Z., He, X., Lu, H., and Zhou, X. (2021). Inactivation of the tumor suppressor p53 by long noncoding RNA RMRP. Proc. Natl. Acad. Sci. USA *118*, e2026813118.

Clark, B.S., and Blackshaw, S. (2017). Understanding the Role of lncRNAs in Nervous System Development. Adv. Exp. Med. Biol. *1008*, 253–282.

Consortium, E.P., Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

Curry, R.N., and Glasgow, S.M. (2021). The Role of Neurodevelopmental Pathways in Brain Tumors. Front. Cell Dev. Biol. *9*, 659055.

Diermeier, S.D., Chang, K.C., Freier, S.M., Song, J., El Demerdash, O., Krasnitz, A., Rigo, F., Bennett, C.F., and Spector, D.L. (2016). Mammary Tumor-Associated RNAs Impact Tumor Cell Proliferation, Invasion, and Migration. Cell Rep. *17*, 261–274.

Ellis, B.C., Molloy, P.L., and Graham, L.D. (2012). CRNDE: A Long Non-Coding RNA Involved in CanceR, Neurobiology, and DEvelopment. Front. Genet. *3*, 270.

Emlet, D.R., Gupta, P., Holgado-Madruga, M., Del Vecchio, C.A., Mitra, S.S., Han, S.Y., Li, G., Jensen, K.C., Vogel, H., Xu, L.W., et al. (2014). Targeting a glioblastoma cancer stem-cell population defined by EGF receptor variant III. Cancer Res. *74*, 1238–1249.

Gimple, R.C., Bhargava, S., Dixit, D., and Rich, J.N. (2019). Glioblastoma stem cells: lessons from the tumor hierarchy in a lethal cancer. Genes Dev. *33*, 591–609.

Gong, W., Zheng, J., Liu, X., Ma, J., Liu, Y., and Xue, Y. (2016). Knockdown of NEAT1 restrained the malignant progression of glioma stem cells by activating microRNA let-7e. Oncotarget *7*, 62208–62223.

Guardia, G.D.A., Correa, B.R., Araujo, P.R., Qiao, M., Burns, S., Penalva, L.O.F., and Galante, P.A.F. (2020). Proneural and mesenchymal glioma stem cells display major differences in splicing and lncRNA profiles. NPJ Genom. Med. *5*, 2.

Guttman, M., Donaghey, J., Carey, B.W., Garber, M., Grenier, J.K., Munson, G., Young, G., Lucas, A.B., Ach, R., Bruhn, L., et al. (2011). lincRNAs act in the circuitry controlling pluripotency and differentiation. Nature *477*, 295–300.

Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., et al. (2010). Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat. Biotechnol. *28*, 503–510.

Han, Y., Zhou, L., Wu, T., Huang, Y., Cheng, Z., Li, X., Sun, T., Zhou, Y., and Du, Z. (2016). Downregulation of lncRNA-MALAT1 Affects Proliferation and the Expression of Stemness Markers in Glioma Stem Cell Line SHG139S. Cell. Mol. Neurobiol. *36*, 1097–1107.

Hazra, R., Brine, L., Garcia, L., Benz, B., Chirathivat, N., Shen, M.M., Wilkinson, J.E., Lyons, S.K., and Spector, D.L. (2022). Platr4 is an early embryonic lncRNA that exerts its function downstream on cardiogenic mesodermal lineage commitment. Dev. Cell *57*, 2450–2468.e7.

Hazra, R., and Spector, D.L. (2022). Simultaneous visualization of RNA transcripts and proteins in whole-mount mouse preimplantation embryos using single-molecule fluorescence in situ hybridization and immunofluorescence microscopy. Front. Cell Dev. Biol. *10*, 986261.

He, Y., Wang, L., Tang, J., and Han, Z. (2020). Genome-Wide Identification and Analysis of the Methylation of lncRNAs and Prognostic Implications in the Glioma. Front. Oncol. *10*, 607047.

Holmberg Olausson, K., Maire, C.L., Haidar, S., Ling, J., Learner, E., Nistér, M., and Ligon, K.L. (2014). Prominin-1 (CD133) defines both stem and non-stem cell populations in CNS development and gliomas. PLoS One *9*, e106694.

Huarte, M. (2015). The emerging role of lncRNAs in cancer. Nat. Med. *21*, 1253–1261.

Jacob, F., Salinas, R.D., Zhang, D.Y., Nguyen, P.T.T., Schnoll, J.G., Wong, S.Z.H., Thokala, R., Sheikh, S., Saxena, D., Prokop, S., et al. (2020). A Patient-Derived Glioblastoma Organoid Model and Biobank Recapitulates Inter- and Intra-tumoral Heterogeneity. Cell *180*, 188–204.e22.

Jiang, X., Yan, Y., Hu, M., Chen, X., Wang, Y., Dai, Y., Wu, D., Wang, Y., Zhuang, Z., and Xia, H. (2016). Increased level of H19 long noncoding RNA promotes invasion, angiogenesis, and stemness of glioblastoma cells. J. Neurosurg. *2016*, 129–136.

Johnson, M.B., Wang, P.P., Atabay, K.D., Murphy, E.A., Doan, R.N., Hecht, J.L., and Walsh, C.A. (2015). Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. Nat. Neurosci. *18*, 637–646.

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C.C., Suzuki, M., Kawai, J., et al. (2005). Antisense transcription in the mammalian transcriptome. Science *309*, 1564–1566.

Kim, Y., Kim, E., Wu, Q., Guryanova, O., Hitomi, M., Lathia, J.D., Serwanski, D., Sloan, A.E., Weil, R.J., Lee, J., et al. (2012). Platelet-derived growth factor receptors differentially inform intertumoral and intratumoral heterogeneity. Genes Dev. *26*, 1247–1262.

Lathia, J.D., Gallagher, J., Heddleston, J.M., Wang, J., Eyler, C.E., Macswords, J., Wu, Q., Vasanji, A., McLendon, R.E., Hjelmeland, A.B., and Rich, J.N. (2010). Integrin alpha 6 regulates glioblastoma stem cells. Cell Stem Cell *6*, 421–432.

Lathia, J.D., Mack, S.C., Mulkearns-Hubert, E.E., Valentim, C.L.L., and Rich, J.N. (2015). Cancer stem cells in glioblastoma. Genes Dev. *29*, 1203–1217.

Lefranc, F., Brotchi, J., and Kiss, R. (2005). Possible future issues in the treatment of glioblastomas: special emphasis on cell migration and the resistance of migrating glioblastoma cells to apoptosis. J. Clin. Oncol. *23*, 2411–2422.

Lendahl, U., Zimmerman, L.B., and McKay, R.D. (1990). CNS stem cells express a new class of intermediate filament protein. Cell *60*, 585–595.

Liu, S., Mitra, R., Zhao, M.M., Fan, W., Eischen, C.M., Yin, F., and Zhao, Z. (2016a). The Potential Roles of Long Noncoding RNAs (lncRNA) in Glioblastoma Development. Mol. Cancer Ther. *15*, 2977–2986.

Liu, S.J., Nowakowski, T.J., Pollen, A.A., Lui, J.H., Horlbeck, M.A., Attenello, F.J., He, D., Weissman, J.S., Kriegstein, A.R., Diaz, A.A., and Lim, D.A. (2016b). Single-cell analysis of long non-coding RNAs in the developing human neocortex. Genome Biol. *17*, 67.

Liu, Z., Wang, X., Yang, G., Zhong, C., Zhang, R., Ye, J., Zhong, Y., Hu, J., Ozal, B., and Zhao, S. (2020). Construction of

lncRNA-associated ceRNA networks to identify prognostic lncRNA biomarkers for glioblastoma. J. Cell. Biochem. *121*, 3502–3515.

Luo, T., Yan, L., and Liu, H. (2020). LINC00632 inhibits the malignant development of non-small cell lung cancer by downregulating miR-1203. J BUON *25*, 1517–1524.

Ma, Q., and Chang, H.Y. (2016). Single-cell profiling of lncRNAs in the developing human brain. Genome Biol. *17*, 68.

Mattick, J.S., Amaral, P.P., Carninci, P., Carpenter, S., Chang, H.Y., Chen, L.L., Chen, R., Dean, C., Dinger, M.E., Fitzgerald, K.A., et al. (2023). Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat. Rev. Mol. Cell Biol. *24*, 430–447.

Mercer, T.R., Dinger, M.E., Sunkin, S.M., Mehler, M.F., and Mattick, J.S. (2008). Specific expression of long noncoding RNAs in the mouse brain. Proc. Natl. Acad. Sci. USA *105*, 716–721.

Molyneaux, B.J., Goff, L.A., Brettler, A.C., Chen, H.H., Hrvatin, S., Rinn, J.L., and Arlotta, P. (2015). DeCoN: genome-wide analysis of in vivo transcriptional dynamics during pyramidal neuron fate selection in neocortex. Neuron *85*, 275–288.

Novak, D., Hüser, L., Elton, J.J., Umansky, V., Altevogt, P., and Utikal, J. (2020). SOX2 in development and cancer biology. Semin. Cancer Biol. *67* (Pt 1), 74–82.

Nowakowski, T.J., Bhaduri, A., Pollen, A.A., Alvarado, B., Mostajo-Radji, M.A., Di Lullo, E., Haeussler, M., Sandoval-Espinosa, C., Liu, S.J., Velmeshev, D., et al. (2017). Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. Science *358*, 1318–1323.

Ostrom, Q.T., Cioffi, G., Waite, K., Kruchko, C., and Barnholtz-Sloan, J.S. (2021). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2014-2018. Neuro Oncol. *23* (12 Suppl 2), iii1–iii105.

Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396–1401.

Paul, Y., Thomas, S., Patil, V., Kumar, N., Mondal, B., Hegde, A.S., Arivazhagan, A., Santosh, V., Mahalingam, K., and Somasundaram, K. (2018). Genetic landscape of long noncoding RNA (lncRNAs) in glioblastoma: identification of complex lncRNA regulatory networks and clinically relevant lncRNAs in glioblastoma. Oncotarget *9*, 29548–29564.

Piccirillo, S.G.M., Reynolds, B.A., Zanetti, N., Lamorte, G., Binda, E., Broggi, G., Brem, H., Olivi, A., Dimeco, F., and Vescovi, A.L. (2006). Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. Nature *444*, 761–765.

Prager, B.C., Bhargava, S., Mahadev, V., Hubert, C.G., and Rich, J.N. (2020). Glioblastoma Stem Cells: Driving Resilience through Chaos. Trends Cancer *6*, 223–235.

Quake, S.R. (2021). The cell as a bag of RNA. Trends Genet. *37*, 1064–1068.

Reon, B.J., Anaya, J., Zhang, Y., Mandell, J., Purow, B., Abounader, R., and Dutta, A. (2016). Expression of lncRNAs in Low-Grade Gliomas and Glioblastoma Multiforme: An In Silico Analysis. PLoS Med. *13*, e1002192.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. *81*, 145–166.

Singh, S.K., Clarke, I.D., Terasaki, M., Bonn, V.E., Hawkins, C., Squire, J., and Dirks, P.B. (2003). Identification of a cancer stem cell in human brain tumors. Cancer Res. *63*, 5821–5828.

Singh, S.K., Hawkins, C., Clarke, I.D., Squire, J.A., Bayani, J., Hide, T., Henkelman, R.M., Cusimano, M.D., and Dirks, P.B. (2004). Identification of human brain tumour initiating cells. Nature *432*, 396–401.

Slack, F.J., and Chinnaiyan, A.M. (2019). The Role of Non-coding RNAs in Oncology. Cell *179*, 1033–1055.

Son, M.J., Woolard, K., Nam, D.H., Lee, J., and Fine, H.A. (2009). SSEA-1 is an enrichment marker for tumor-initiating cells in human glioblastoma. Cell Stem Cell *4*, 440–452.

Statello, L., Guo, C.J., Chen, L.L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. Nat. Rev. Mol. Cell Biol. *22*, 96–118.

Su, R., Cao, S., Ma, J., Liu, Y., Liu, X., Zheng, J., Chen, J., Liu, L., Cai, H., Li, Z., et al. (2017). Knockdown of SOX2OT inhibits the malignant biological behaviors of glioblastoma stem cells via up-regulating the expression of miR-194-5p and miR-122. Mol. Cancer *16*, 171.

Tao, Y., Zhang, J., Chen, L., Liu, X., Yao, M., and Zhang, H. (2021). LncRNA CD27-AS1 promotes acute myeloid leukemia progression through the miR-224-5p/PBX3 signaling circuit. Cell Death Dis. *12*, 510.

Tchoghandjian, A., Baeza, N., Colin, C., Cayre, M., Metellus, P., Beclin, C., Ouafik, L., and Figarella-Branger, D. (2010). A2B5 cells from human glioblastoma have cancer stem cell properties. Brain Pathol. *20*, 211–221.

Terashima, M., Ishimura, A., Wanna-Udom, S., and Suzuki, T. (2018). MEG8 long noncoding RNA contributes to epigenetic progression of the epithelial-mesenchymal transition of lung and pancreatic cancer cells. J. Biol. Chem. *293*, 18016–18030.

Tuy, K., Rickenbacker, L., and Hjelmeland, A.B. (2021). Reactive oxygen species produced by altered tumor metabolism impacts cancer stem cell maintenance. Redox Biol. *44*, 101953.

Uhlhaas, P.J., and Singer, W. (2006). Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. Neuron *52*, 155–168.

Ulitsky, I., and Bartel, D.P. (2013). lincRNAs: genomics, evolution, and mechanisms. Cell *154*, 26–46.

Vaubel, R.A., Tian, S., Remonde, D., Schroeder, M.A., Mladek, A.C., Kitange, G.J., Caron, A., Kollmeyer, T.M., Grove, R., Peng, S., et al. (2020). Genomic and Phenotypic Characterization of a Broad Panel of Patient-Derived Xenografts Reflects the Diversity of Glioblastoma. Clin. Cancer Res. *26*, 1094–1104.

Verhaak, R.G.W., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P., et al. (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell *17*, 98–110.

Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Mol. Cell *43*, 904–914.

Wang, R., Sharma, R., Shen, X., Laughney, A.M., Funato, K., Clark, P.J., Shpokayte, M., Morgenstern, P., Navare, M., Xu, Y., et al. (2020). Adult Human Glioblastomas Harbor Radial Glia-like Cells. Stem Cell Rep. *14*, 338–350.

Wang, X., Prager, B.C., Wu, Q., Kim, L.J.Y., Gimple, R.C., Shi, Y., Yang, K., Morton, A.R., Zhou, W., Zhu, Z., et al. (2018). Reciprocal Signaling between Glioblastoma Stem Cells and Differentiated Tumor Cells Promotes Malignant Progression. Cell Stem Cell *22*, 514–528.e5.

Yang, G., Lu, X., and Yuan, L. (2014). LncRNA: a link between RNA and cancer. Biochim. Biophys. Acta *1839*, 1097–1109.

Yao, R.W., Wang, Y., and Chen, L.L. (2019). Cellular functions of long noncoding RNAs. Nat. Cell Biol. *21*, 542–551.

Yao, Y., Ma, J., Xue, Y., Wang, P., Li, Z., Liu, J., Chen, L., Xi, Z., Teng, H., Wang, Z., et al. (2015). Knockdown of long non-coding RNA XIST exerts tumor-suppressive functions in human glioblastoma stem cells by up-regulating miR-152. Cancer Lett. *359*, 75–86.

Zhao, S.G., Yu, M., Spratt, D.E., Chang, S.L., Feng, F.Y., Kim, M.M., Speers, C.W., Carlson, B.L., Mladek, A.C., Lawrence, T.S., et al. (2019). Xenograft-based, platform-independent gene signatures to predict response to alkylating chemotherapy, radiation, and combination therapy for glioblastoma. Neuro Oncol. *21*, 1141–1149.

Zhao, Y., Wang, W., Guan, C., Hu, Z., Liu, L., Li, W., and Jiang, X. (2020). Long noncoding RNA HOTAIRM1 in human cancers. Clin. Chim. Acta *511*, 255–259.

Zhou, L., and Ma, J. (2022). MIR99AHG/miR-204-5p/TXNIP/Nrf2/ARE Signaling Pathway Decreases Glioblastoma Temozolomide Sensitivity. Neurotox. Res. *40*, 1152–1162.

Zhu, X., Chen, H.H., Gao, C.Y., Zhang, X.X., Jiang, J.X., Zhang, Y., Fang, J., Zhao, F., and Chen, Z.G. (2020). Energy metabolism in cancer stem cells. World J. Stem Cell. *12*, 448–461.